Vol. 3, No.1-2, 2004

Journal of Global Positioning Systems

ISSN 1446-3156 (Print Version) ISSN 1446-3164 (CD-ROM Version)

International Association of Chinese Professionals in Global Positioning Systems (CPGPS)

Journal of Global Positioning Systems

Aims and Scope

The Journal of Global Positioning Systems is a peer-reviewed international journal for the publication of new information, knowledge, scientific developments and applications of the global navigation satellite systems as well as other positioning, location and navigation technologies. The Journal publishes original research papers, review articles and invited contributions. Short research and technical notes, book reviews, and commercial advertisements are also welcome. Specific questions about the suitability of prospective manuscripts may be directed to the Editor-in-Chief.

Publication and Copyright

The Journal of Global Positioning Systems is an official publication of the International Association of Chinese Professionals in Global Positioning Systems (CPGPS). It is published twice a year, in June and December. The Journal is available in both print version (ISSN 1446-3156) and CD-ROM version (ISSN 1446-3164), which can be accessed through the CPGPS website at http://www.cpgps.org/journal.php. Whilst CPGPS owns all the copyright of all text material published in the Journal, the authors are responsible for the views and statements expressed in their papers and articles. Neither the authors, the editors nor CPGPS can accept any legal responsibility for the contents published in the journal.

Subscriptions and Advertising

Membership with CPGPS includes subscription to the Journal during the period of membership. Subscriptions from non-members and advertising inquiries should be directed to:

CPGPS Headquarters
Department of Geomatics Engineering
The University of Calgary
Calgary, Alberta, Canada T2N 1N4
Fax: +1(403) 284-1980
E-mail: journal@cpgps.org

CPGPS Logo Design:
Peng Fang
University of California, San Diego, United States

Cover Design and Layout:
Satellite Navigation & Positioning Group
The University of New South Wales, Australia

Printing:
Satellite Navigation & Positioning Group
The University of New South Wales, Australia

Journal of Global Positioning Systems

Vol. 3, No.1-2, 2004

Table of Contents

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 1

Journal of Global Positioning Systems
2004
© CPGPS

Letter From the Guest Editor

Professor Chris Rizos, Guest Editor

School of Surveying and Spatial Information Systems, University of New South Wales, Sydney, NSW 2052, Australia.
e-mail: c.rizos@unsw.edu.au, Tel: +61 2 9385 4205; Fax: +61 2 9313 7493

I am pleased to be the guest editor for these special issues of the Journal of Global Positioning Systems to publish the selected papers from the 2004 International Symposium on GPS/GNSS (GNSS-2004), 6-8 December 2004, organised by the School of Surveying and Spatial Information Systems, University of New South Wales, Australia.

The GNSS-2004, attended by over 340 delegates from 29 countries, was the largest symposium in the Asia/Pacific region in 2004 dedicated to GNSS and wireless positioning. On the first day of the symposium the plenary session consisted of several presentations by invited speakers. The Civil GPS Service Interface Committee (CGSIC) convened a public meeting where representatives from the US Dept of Transport and US Coast Guard informed participants of recent developments in GPS policies and modernization. Presentations by several European speakers introduced the future GNSS "Galileo", scheduled for deployment by the end of the decade. The Symposium also featured technical workshops and around 200 oral and 'flashing poster' presentations.

Over 130 papers presented at the symposium were submitted to the special issues of the Journal of Global Positioning Systems (JGPS). However, given the limited space pages in these special issues many high quality papers could not be selected for publication. This has been a considerable administrative challenge and I wish to thank the reviewers for their assistance in reading and selecting the papers.

The 2004 and 2005 volumes of the JGPS feature the selected papers from the symposium on various GNSS topics, as well as research and development into a variety of other positioning technologies.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 2-11

Journal of Global Positioning Systems
2004

GNSS Indoor Location Technologies

Gérard Lachapelle

Position, Location And Navigation Group, Department of Geomatics Engineering, University of Calgary, Canada
Tel: 403 220 7104 E-mail: Lachapelle@geomatics.ucalgary.ca

Received: 15 November 2004 / Accepted: 3 February 2005

Abstract. This paper presents an overview of GNSS-based indoor location technologies. Current and emerging users and their potential requirements are first discussed. Signal attenuation and multipath caused under indoor environments are described. The basic method to acquire and track attenuated signals, namely longer integration of signal measurements, is summarized. The need for assisted GPS is addressed. Availability and accuracy performance currently achievable under various conditions (wooden structure building, single family residence, large sport facility) are illustrated through selected test results. The limitations of current technologies and potential enhancements are discussed. These include measurement noise, existing signal structure and future enhancements, frequency and time errors, user motion, sensor aiding such as ultra-tight integration, and solution reliability and continuity. The paper concludes with a discussion of receiver testing standards. The possibility of using a GNSS hardware simulator to create reproducible indoor environments in order to overcome the controllability issue encountered with real environments is analysed.

Key words: Indoor location, AGPS, HSGPS, aided GPS, Indoor GPS

1 Introduction

The need for indoor location was initially spurred by the mid-90s U.S. FCC decision to require mobile phone service providers to locate emergency (E911) callers with an accuracy requirement (October 1999 revision) of 100 m (67%) or 300 m (95%) for network-based solutions and 50 m (67%) or 150 m (95%) for handset-based solutions, the latter being the case with the use of GPS. The most promising technologies to achieve this on a continental level were cellular phone network-based TDOA methods and GPS if the latter could be made to operate under attenuated signal conditions such as urban canyons, forested areas and indoors. Attenuated signal environments are now often simply labelled “indoor” environments for the sake of simplicity. Early developments and testing of cellular phone network-based TDOA methods resulted in promising results with the use of GPS to precisely time synchronize signal transmissions (e.g. Klukas et al 1997, 1998). However, the additional cell equipment required proved to be a challenge. In addition, TDOA being an hyperbolic location method (similar to Loran-C for instance), observability was found to be low in a urban environment given the rapidly changing geometry of the available cells. Outside of urban areas, the cell geometry was simply not present to meet availability requirements. In parallel to the above developments, experiments to increase the integration time to potentially allow GPS signal reception under attenuated signal environments yielded promising results, especially with the use of assisted methods (Petersen et al 1997, Moeglein & Krasner 1998, Garin et al 1999). Given the existing space-based infrastructure and coverage advantages of GPS that makes possible an in-mobile phone solution, R&D efforts on improving GPS performance intensified rapidly.

It soon became obvious that continuous outdoor and indoor location availability could be used in a large number of applications such as personal digital assistant location, asset tracking, vehicular navigation, and no doubt many others to be discovered, in addition to emergency services. These applications now form part of a location-based services business that is expected to grow from USD 0.5 B in 2003 to USD 28 B in 2008. One of the most important technical questions arising is what are the performance levels users want. Standard location and navigation performance parameters are important, namely availability, accuracy, reliability and integrity, and

continuity (outdoor-to-indoor and indoor). The performance level for each of these parameters varies and, except for the case of E911, has not been rigorously defined yet, except as a wish list. A safe assumption is the higher, the better! A wish often expressed is "10 m in 10 s anywhere at any time".

2 Signal Attenuation and Integration Issues

Signal attenuation occurs in any medium other than free space. The atmosphere for instance attenuates the GPS signals by about 1 dB under normal circumstances. Signal attenuation caused by specific materials can be determined experimentally under controlled conditions (e.g. Klukas et al 2004). Gyprock and plywood for instance attenuate GPS signals by about 0.5 dB and 2.3 dB, respectively, while a cinder block causes attenuation ten times higher, at about 23 dB. Attenuation due to the forestry canopy is highly variable and depends on leaf size and thickness, moisture content, forest density and trunk size. Steel, reflective glass and other construction materials have yet other signal attenuation properties. The problem in characterizing signal attenuation and reflection indoors is the presence and arrangement of diverse materials. In situ measurements are preferable if a sufficiently sensitive receiver is available in the first place. If large attenuation and non-homogeneities occur, the signals measured by the receivers might be echo-only signals that may contain large errors, depending on indoor geometry, as illustrated in Figure 1. Thus, indoor location accuracy is not only a function of signal attenuation but also of building geometry, even when assuming fairly constant satellite geometry. This will be further illustrated through examples later. In addition to the level of attenuation and multipath, their spatial and temporal variations is of interest and importance as these affect position continuity.

Line-of-sight signals are already very week (around -160 dBW or -130 dBm) due to the large free space loss. This signal strength results in a SNR (signal-to-noise ratio) of -18 dB when a 2-MHz front-end filter is assumed. Signals are acquired and tracked by correlation and integration, resulting in a processing gain of 33 dB, if one assumes a pre-correlation bandwidth of 2 MHz and an integration time of 2 ms. This procedure increases the SNR to +15 dB which is the approximate lower limit allowable for acquisition (The lowest SNR required for tracking is approximately at least 3 dB lower than the above, the exact value depending upon the tracking loop configuration). When the signal is attenuated, integration can be done over a longer interval in order to increase the processing gain. Coherent integration can be done over intervals of up to 20 ms when bit synchronization is achieved due to the limited length of the GPS navigation message bits transmitted as part of the signal. If integration over a longer period is performed, a non-coherent technique that squares the in-phase (I) and quadra-phase (Q) measurements has to be used. Squaring also squares the noise, thus diminishing the net gain of this technique (e.g. Chansarkar & Garin 2000). The processing gain is given by

$$Gtot = 10 \log (Bpre \times T) + 10 \log (M) - Sqloss \quad (1)$$

where Bpre is the pre-detection bandwidth, T is the coherent integration time in ms, M is the number of non-coherent accumulations, and SQloss is the squaring loss due to non-coherent accumulation. If coherent integration is extended over the full 20 ms, the SNR gain, with respect to a 2-ms integration, is 10 dB. Coherent integration is often the only method used in tracking mode. For signal acquisition, non-coherent integration is often extended to a few to several hundred ms to reach the minimum SNR value of 15 dB.

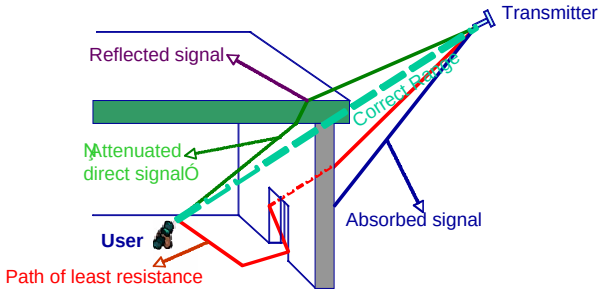


Figure 1: Possible GPS signal propagation paths into a building

A basic assumption, when long integration is considered, is that the Doppler error remains small or constant during the integration period. However the receiver frequency and time standard (FTS) drift and unpredictable user antenna motion can cause significant phase changes and can effectively limit the integration time. The FTS error can be reduced by using a higher grade unit. If only a better FTS is used and the user antenna is in motion, separating antenna motion and FTS drift will remain an issue. The use of an inertial measurement unit (IMU) to measure antenna motion will further improve performance, as will be discussed later. The antenna gain pattern is also an issue as the use of reflected signals will generally be the norm indoor, in which case these signals will arrive at the antenna from any direction. An antenna with a high gain at low elevation is therefore preferable, although this can result in a further loss (an ½ hemispherical antenna gain of 3 dB is usually assumed for signal strength calculations). The problem will be

compounded if the antenna is subject to inversion as in the case of a mobile phone.

2.1 Pseudorange accuracy versus signal attenuation

Under signal attenuation, the thermal noise significantly affects pseudorange accuracy. This is shown in Figure 2. The theoretical minimum pseudorange standard deviation, calculated using the Cramer-Rao lower bound for the BPSK C/A code is shown for two pre-correlation bandwidths, namely 2 and 16 MHz. A DLL (Delay Lock Loop) loop bandwidth of 0.3 Hz is assumed. Current HSGPS receivers use a 2 MHz pre-correlation bandwidth. C/No values of 24 and 8 dB-Hz correspond to attenuation levels of about 20 and 36 dB, respectively. The error growth as attenuation increases is indeed significant, namely from 3 m to 115 m for the case of a 2-MHz pre-correlation bandwidth. The effect on position accuracy would depend on the effect of multipath that would likely dominate the error budget. Figure 2 also shows the actual pseudorange standard deviation as obtained using a SiRF receiver under a signal attenuation range of 0 to 30 dB. The standard deviation is actually better than that predicted by the Cramer-Rao lower bound algorithm. This is likely due to the use of a proprietary filter inside the receiver.

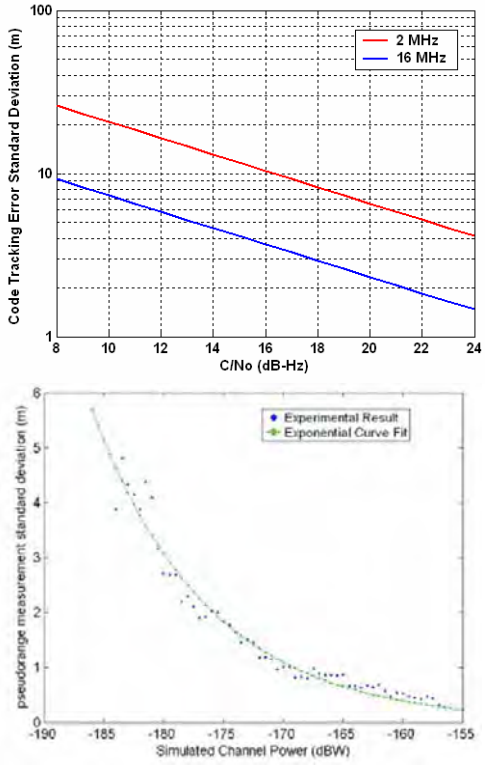


Figure 2: Pseudorange accuracy versus attenuation: Cramer-Rao lower bound (2 MHz pre-correlation bandwidth, 0.3 Hz DLL bandwidth), SiRF STARII measured accuracy

2.2 Signal acquisition and AGPS

Low pre-correlation SNR is one of the main issues for acquisition. Indeed, acquisition requires a post-correlation SNR value of around 14 dB, which is higher than for tracking, for safe (low false alarm or false detection) detection. Moreover, the acquisition process has to deal with the fact that it has no a priori knowledge of user motion, clock drift and bit synchronization (limiting coherent integration) in a standard standalone configuration. As a consequence, more advanced techniques have to be used. The main approaches to this problem are the use of large banks of correlators (van Diggelen and Abraham 2001) and assisted GPS (AGPS). The latter has become standard for mobile phone location. A GPS reference station provides various parameters to the user to reduce TTFF, such as ephemeris, almanac, initial position, reference time, Doppler information and DGPS corrections (Syrjärinne 2001, Karunanayake et al 2004, Weill et al 2004).

2.3 Using MEMS Inertial Measuring Units (IMUs)

An IMU consists of an assembly of gyros and accelerometers deployed in a configuration suitable to measure orientation change and acceleration in three dimensions. The minimal configuration for a six degree-of-freedom system consists of three units of each type. A more complex configuration may have redundant sensors on each axis (arrays of sensors) that can potentially yield superior performance. They come in a variety of performance classes, sizes, power consumptions and costs. Any class can be used to aid HSGPS. The better the class, the better the overall performance of the combined system will be, although no performance analysis of overall system improvement versus IMU class and signal attenuation is readily available. For low cost portable applications, miniature MEMS IMUs are relevant. Miniature MEMS gyros and accelerometers have currently dimensions of no less than 10 x 10 x 5 mm and consume $\leq$ 10 mA. In large quantities their prices are sufficiently low for consideration. IMUs can be used for tracking loop aiding in an ultra-tight integration mode, user motion detection during signal integration and bridging user positions during very short signal measurement gaps. The bandwidth and noise reduction can result in a gain of several dBs. They can at the same time be used in

parallel to increase position continuity and reliability as discussed later. A higher position rate can be made available if required.

There are however numerous challenges in effectively integrating MEMS IMUs with HSGPS. They are much affected by thermal effects, which is a major problem for numerous applications. If temperature effects are not compensated, a performance of the order of about 1 degree (or slightly better) per second is achievable. If temperature effects are compensated using even a simple linear model, a performance of better than 0.1 degree per second is expected. Another major issue with using an IMU is the determination of the initial attitude parameters. Also, the correlation of IMU velocity errors with oscillator frequency drift has to be dealt with. The algorithms and software will likely have to be application specific in order to optimise performance. The pay back is however potentially very significant.

3 Indoor Positioning Performance Issues and Examples

There are issues with all major positioning performance parameters when using GPS indoor. These can be summarized as follows:

Availability: The number of satellites available and their geometry limit this performance parameter. Low elevation satellites are not usually available indoor due to excessive signal attenuation. Low availability worsens geometry and reduces redundancy and therefore the effectiveness of RAIM algorithms. The use of miniature MEMS barometers to aid the height component is very cost effective and improves not only redundancy but also the horizontal dilution of precision (HDOP). A miniature barometer, if operating in differential mode, is sufficient to identify the relative height of a user with an accuracy of about 2 m. Thus, a change of floor can accurately be measured using only one MEMS barometer. The use of a clock constraint, if the latter is sufficiently accurate, also improves redundancy. If both a clock constraint and barometry are used, two satellites in a good geometry can deliver a horizontal position solution. A third satellite will enable a RAIM algorithm.

Accuracy: This performance parameter is affected by high noise, echo-only or high multipath signals and degraded geometry. This is why indoor GPS accuracy is much lower than outdoor as will be seen in the numerous examples described below. DGPS will help more than in the outdoors because position accuracy is total user range accuracy multiplied by the DOP. Since the DOP is generally poor, removing the user range error due to single point operation will help somewhat. The use of the height and clock constraints described above will also help. IMU aiding and filtering will improve relative accuracy. Other aiding methods such as pseudolites, UWB, cellular network TDOA methods, will also help accuracy, in addition to the other parameters.

Continuity: This is an especially serious problem indoor due to the rapid temporal and spatial decorrelation of multipath and the rapidly changing satellite geometry as signals come in and out. Epoch-by-epoch positions using an unconstrained least-squares approach show large jumps. The use of a Kalman filter with constraints adapted to the expected user dynamics will go a long way in dealing with this issue. The use of a self-contained low cost MEMS IMU will further improve positioning continuity.

Reliability: The use of RAIM is effective insofar as redundancy is available. Since redundancy is low indoor, applying RAIM may decrease availability and accuracy (Lachapelle et al 2004). Like in the case of continuity, adding self-contained sensors, external aiding and filtering with proper constraints will improve the situation. The use of combined GPS and Galileo will have a major impact on reliability (Kuusniemi et al 2004a, b, Lachapelle et al 2004).

Indoor Location Examples

Several examples are described herein to illustrate HSGPS performance under a variety of signal attenuation conditions, from relatively low to high signal attenuation building. Unless stated otherwise, the position solutions were derived using an unconstrained epoch-by-epoch least-squares algorithm to better assess the true epoch-by-epoch effects of high noise, multipath, satellite geometry and their temporal variations.

Kinematic Positioning in a wooden building

The results of this test, conducted with SiRF receivers, are shown in Figure 3. The building is a large barn made of wood with a roof of asphalt shingles. The attenuation was consistently below 10 dB. The data was processed in differential mode to eliminate atmospheric and orbital errors. A pedestrian walked along the outer walls numerous times with the equipment mounted in a backpack. A NovAtel BlackDiamond™ GPS/INS system was used to provide external reference positions (The pedestrian went outside between every run to allow GPS to update the tactical grade HG1700 INS unit. Each run was completed in a few minutes. The epoch-by-epoch least-squares horizontal positions, shown in Figure 3,

are accurate to about 5 m. An unconstrained least-squares solution is used to allow a more realistic performance analysis. In an operational environment, a Kalman filter would yield smoother relative positions. Availability with a HDOP of 4 or better is 95%. The accuracy degrades significantly when the HDOP exceeds 4. A similar test outdoor with similar HSGPS equipment would yield a corresponding accuracy of about 2 m. A standard receiver was used inside to confirm that no signals could be measured. These results therefore show the effectiveness of HSGPS in this type of indoor environment.

Location in a North American Residence

North American residences are typically constructed with a wooden frame, gyprock inside and wood, stucco or some other finishing outside. Depending on the region, walls vary in thickness between 15 and 25 cm and include insulating material. The roof material varies. In the example reported herein, the roof consists of concrete slates. A SiRF receiver was installed in the garage located under the main living room. The dimensions of the garage are approximately 8 x 12 x 3 m. The wooden garage door was kept closed during the test. The results are shown in Figure 4. The receiver was initialised outside and kept outside for about 30 minutes. Signal attenuation indoor varied between 10 and 30 dB, the latter number being at the tracking capability limit of the receiver (MacGougan et al 2002). Despite this high level of attenuation, the horizontal position accuracy was of the order of 10 m (RMS), with a maximum error of 30 m. This case illustrates well the expected location performance of an emergency caller inside such a building or for that matter of an asset tracking GPS device. The accuracy is high enough to identify the specific residence where the device is located.

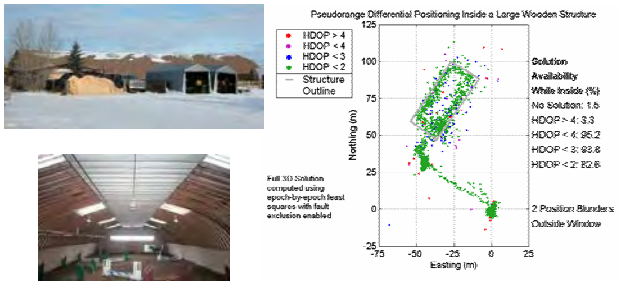


Figure 3: Kinematic test results inside a wooden building

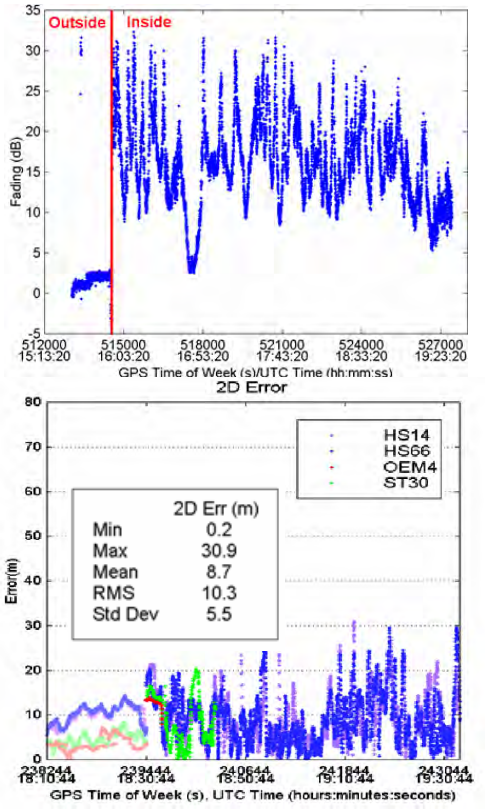


Figure 4: Static test results inside a North American residence

Positioning in the Calgary Olympic Oval

This test case is fully described by Dao et al (2004). The Calgary Olympic Oval recreational facility, shown in Figure 5a together with the two-antenna test equipment, was built for the 1988 Olympic Winter Games and includes a 400 m speed skating track bordered by a running track used for testing. The Oval is 25 m high at its centre, surrounded by concrete walls and a row of windows near the roof at a height of approximately 20 m. Surveyed points were established along the running track to evaluate the accuracy of HSGPS-derived positions. The test reported herein was conducted in kinematic mode. The passing times of the pre-surveyed reference points were tagged by the pedestrian conducting the test. Again HSGPS SiRF receivers were used. The C/No values for the satellites available during the test are shown in Figure 5a. Given that a normal C/No under LOS conditions is 44 dB-Hz, one observes that the attenuation ranged between 10 and 27 dB, about the same as in the case of the residence test case described earlier. The test was conducted with two systems, with an inter-antenna distance of 1 m to assess spatial multipath decorrelation and the impact of antenna diversity on position estimation under such conditions. The results of one test run are shown in Figure 5b, together with its

position performance statistics. The 2 DRMS accuracy with one antenna (black dots) is 45 m while the corresponding accuracy using the combined measurements (blue and green dots) from both antennas is about 25 m. Two important aspects of these results are that (1) antenna diversity can effectively average out some multipath effects in such an environment and (2) position accuracy and signal attenuation are only partly correlated, the impact of the indoor geometry and signal reflectivity also playing important roles.

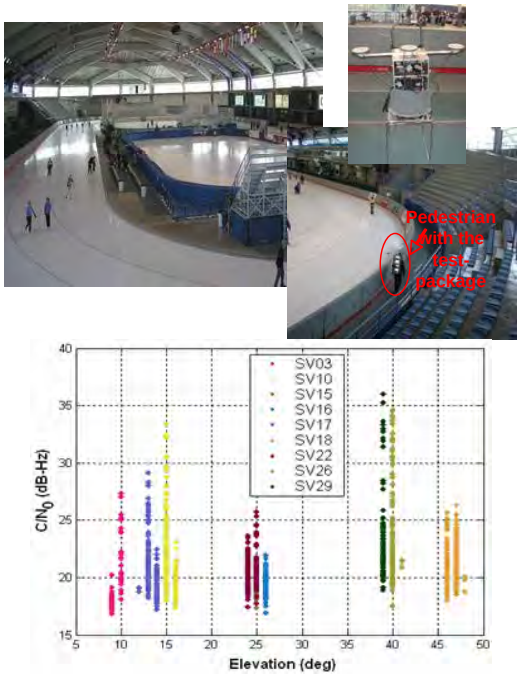


Figure 5a: Calgary Olympic Oval Kinematic test and signal attenuation

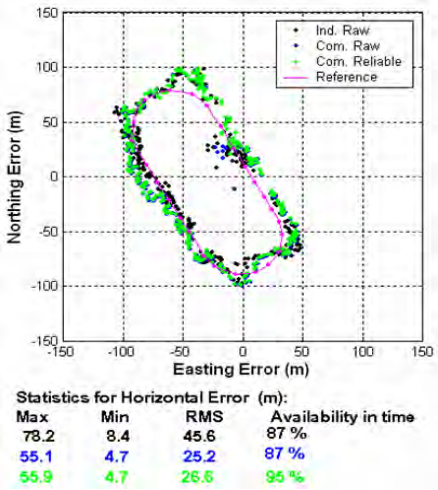


Figure 5b: Calgary Olympic Oval Kinematic test results

Location in a commercial building

This test was conducted inside the PLAN laboratory, located in the Calgary Centre for Innovative Technology (CCIT), as shown in Figure 6a. Signal attenuation, shown in Figure 6b, ranges from 15 to 30 dB. The epoch-by-epoch position scatter, also shown in Figure 6b, is relatively large and the 2D RMS accuracy is 75 m. The use of RAIM to identify unreliable solutions is effective but of limited help due to low redundancy, the user generally has the choice with a position with low or unavailable reliability measure or no position at all. Thanks to a rapid temporal decorrelation of multipath, the use of a batch type solution over short intervals (10 – 20 s) improves the accuracy to 58 m and, not shown here, the reliability.

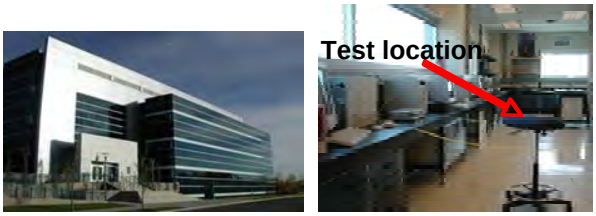


Figure 6a: Office building test set-up

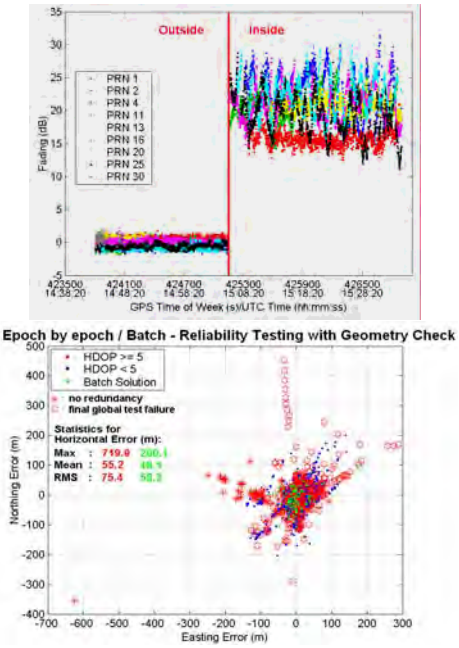
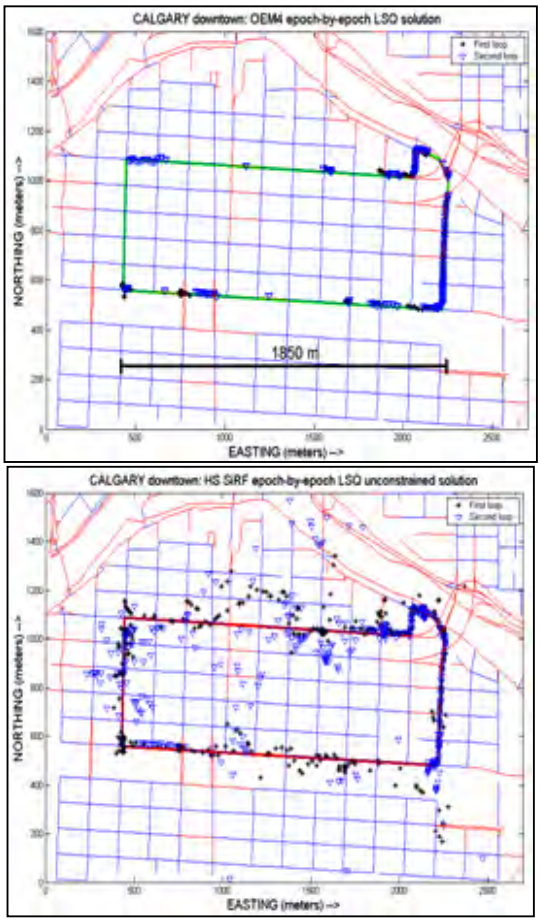


Figure 6b: Office building test results

Vehicular navigation in urban canyons

This example, reported by Mezentsev et al (2002) and illustrated in Figure 7, shows very well the advantages and limitations of HSGPS. A six km loop in

downtown Calgary was driven numerous times. That reported here is representative of the average performance. A standard receiver was used in addition to a SiRF receiver to show the performance differences. The UL (upper left) graph shows the trajectory in green and the fixes obtained with the standard receiver in blue. Availability is very low but accuracy very high. The UR graph shows the epoch-by-epoch least-squares fixes obtained with the HSGPS units. Availability is high but accuracy is low. No RAIM algorithm was applied and errors of hundreds of metres occur, due to echo-only signals reflected from buildings and, possibly, cross-correlation effects. The LL graph shows the HSGPS heavily filtered solution, which is generally quite good but benefits from the straight segments of the trajectory. The LR graph shows a solution (green) based on the integration of the HSGPS unit raw measurements with an automotive grade Murata rate gyro. In the latter case, availability is 100% and the accuracy is excellent, apart from a bias growing to 50 m in the LR section of the trajectory. Further augmentation by other vehicle's components such as the ABS would likely further improve accuracy and would certainly improve reliability, although the latter was not systematically analyzed in this example.



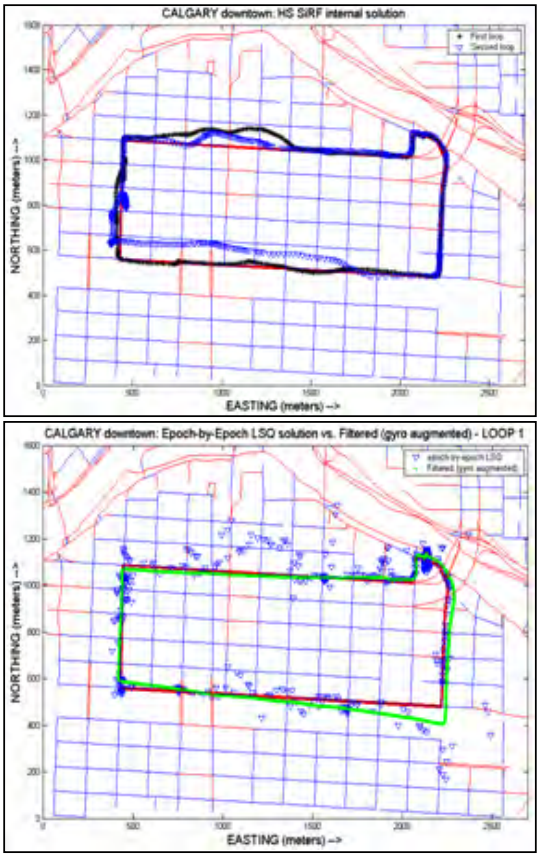


Figure 7: Vehicular navigation in urban canyons

Pedestrian navigation in urban canyons

This case is much more difficult than the vehicular case due to the more unpredictable nature of a pedestrian's trajectory. A totally unpredictable turning radius might be a good way to describe this case! In the example shown in Figure 8 and conducted by Mezentsev et al (2004), a full six degrees of freedom MEMS IMU attached to the user's waist is integrated with a HSGPS receiver in pedestrian dead reckoning (PDR) mode. The HSGPS receiver is a SiRF X-trac unit. The prototype test PDR unit consists of three gyros and three accelerometers built in a classical perpendicular triad scheme that represents a full six degrees of freedom low-cost IMU, as shown in Figure 8. The gyros used are Analog Devices ADXRS150 ±150°/s single chip rate gyros that are well suited for pedestrian navigation due to their small size (7 × 7 × 3 mm) and low power consumption (< 50 mW). The MEMS accelerometers used are VTI SCA 610 series. The total volume including a power regulation circuit is less than 100 cm^3. In a commercial product, the volume of such unit would likely be of the order of a few cm^3.

Since the IMU is attached to the body in a "semi-rigid" manner, its measured dynamics do indeed represent those of the user, which is not the case if the IMU was mounted in a PDA or mobile phone. The gyro triad is used to keep track of attitude. It is a complete tri-axis quaternion based attitude solution, so user's heading, pitch and roll are known. The accelerometers are used to detect steps. Also, optionally, they can be used to perform horizontal alignment at complete user stops. The user step length is initially assumed to be constant, say 70 cm. During use, it is calibrated when good GPS position solutions are available. In PDR mechanization, the position error is proportional to the distance traveled and not time as is the case in a classical INS mechanization. Given the position error growth characteristics of both approaches, the former (PDR) for a pedestrian significantly outperforms the latter, even with tactical grade systems over long periods of time (more than 1 minute of INS only navigation). The GPS-IMU PDR integration steps used are shown in Figure 8, together with the results of a 1.5 km loop in downtown Calgary. A RAIM algorithm was implemented to detect unreliable solutions. The unaided least-squares GPS position fixes are shown in green (reliable), red (Global test failure), and black (not enough redundancy to judge). The integrated HSGPS-IMU trajectory, obtained with a Kalman filter, is shown in blue. Its availability is nearly 100% while its maximum error reaches 50 m. An important question is what the performance gain would be if an ultra-tight integration was used.

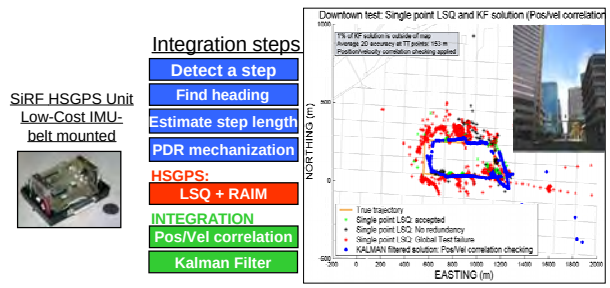


Figure 8: Pedestrian navigation in urban canyons

4 Testing Procedures and Standards

The test results shown in the previous section illustrate the difficulties associated with defining a sufficiently wide range of "standard" environments to fully test a commercial product purporting to meet pre-defined minimum operational performance standards (MOPS). Even if such environments were found, reproducibility would be an issue due to the spatial decorrelation of multipath, as was seen in the Calgary Olympic Oval for instance. In addition such environments would hardly be portable.

In order to establish common testing standards, the question arises as to whether it is possible to construct realistic environments on a simulator. If this were possible it would resolve the above issues. An initial attempt was made by Spirent with the development of a new generation of simulators. Using its SimGEN software, it is possible to simulate a wide range of multipath scenarios, with different degrees of obstructions up to echo-only signals, while varying signal strength, as shown in Figure 9. The next step is to "stochastically" reproduce a realistic environment. This means simulating an environment with the same stochastic signal fading, multipath and temporal variation properties as those observed in an actual environment. Early attempts have resulted in encouraging results (Lachapelle et al 2003) but there is still much work to accomplish to obtain satisfactory results.

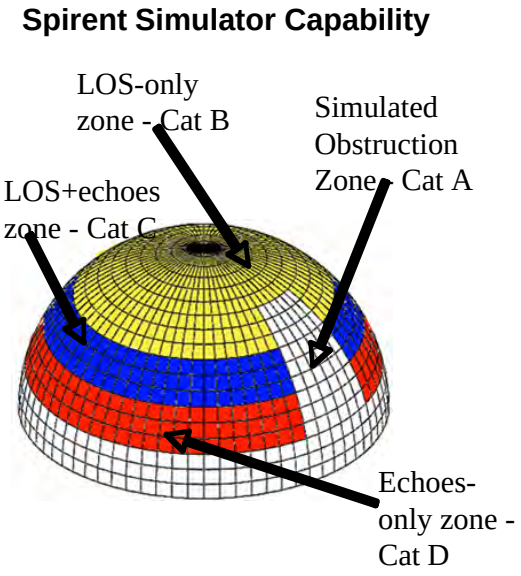


Figure 9: Indoor signal simulation scenario

5 Future Signals

GPS II and Galileo signals will have significant advantages over the current GPS C/A code modulated L1 signal. The availability of pilot channels will allow the use of a pure PLL (Phase Lock Loop) and will result in more robust carrier phase tracking, avoiding the squaring loss currently present due to the need of a Costas loop.

BOC (Binary Offset Carrier) modulation on the Galileo E5 and L1 (and GPS M-code) will improve mitigation of thermal noise, multipath and narrow band interference. Figure 10 shows the Cramer-Rao lower bounds for pseudorange accuracy for BOC(1,1)

modulation (that will be used for Galileo L1) versus BPSK(1) (currently used on the GPS L1 signal). The Cramer-Rao lower bound shown only deals with thermal noise effects. The standard deviation difference between a BPSK(1) modulation using a 2-MHz pre-correlation bandwidth and a BOC (1,1) modulation using a 16-MHz bandwidth is nearly one order of magnitude. How this will translate in actual position accuracy in the presence of large multipath and echo-only signals remains to be determined. It should also be noted that BOC signal tracking could lead to biased pseudorange measurement due to its multi-peak auto-correlation function. This might particularly be a threat when low SNR are considered.

Finally the use of secondary codes (on GPS L5 and Galileo L1, E5) will result in better narrow band interference mitigation and improved bit synchronization, although they might degrade the acquisition MTTF (Hegarty et al 2003; Macabiau et al 2003).

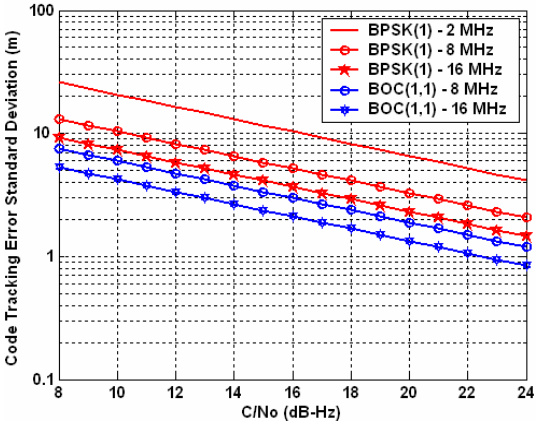


Figure 10: Cramer-Rao lower bounds – BPSK(1) versus BOC(1,1) signal modulation techniques

6 Conclusions

The past 10 years have seen the birth, development and deployment of the first generation of indoor GPS technology. The current limitations of this technology are severe, as compared to the level of performance achievable to outdoor users. Yet, the technology is largely responsible for the emerging location-based services market, which has enormous potential for growth and impact. This market is becoming increasingly demanding in terms of performance. Accuracy is highly addictive! The question now is what will the technological improvements be during the next 10 years. One can safely predict significant to major improvements in the following areas: better signal tracking, use of new GPS and Galileo signals, and effective use of self-contained sensors and external aiding. Will these improvements be sufficient to keep up with users expectations?

Acknowledgements

The assistance of the following researchers in the PLAN Group, Department of Geomatics Engineering, the University of Calgary, are acknowledged: Drs. C. Ma and M. Petovello, senior research associates, and D. Dao, O. Julien, D. Karunanayake, H. Kuusniemi, O. Mezentsev, and B. Zheng, MSc and PhD candidates.

References

Chansarkar, M. and L. J. Garin (2000): ***Acquisition of GPS Signals at Very LowSignal to Noise Ratio***. In Proceedings of the Institute of Navigation ION National Technical Meeting-2000 (January 26-28, 2000, Anaheim, California), pages 731–737.

Dao, D., H. Kuusniemi, and G. Lachapelle (2004): HSGPS ***Reliability Enhancements Using aTwin-Antenna System***. CD-ROM Proceedings of GNSS 2004 Conference, Rotterdam (16-19 May), 11 pages.

Garin, L. J., M. Chansarkar, S. Miocinovic, C. Norman, and D. Hilgenberg (1999): ***Wireless Assisted GPS-SiRF Architecture and Field Test Results***. In Proceedings of the Institute of Navigation ION GPS-99 (September 14-17, 1999, Nashville, Tennessee), pages 489–497.

Hegarty, C., M. Tran, and A.J. Van Dierendonck (2003): ***Acquisition Algorithms for GPS L5***, CD-ROM Proceedings of the European Navigation Conference GNSS, (Graz, Austria, April 22-), 14 pages.

Karunanayake, D., M.E. Cannon, G. Lachapelle, G. Cox (2004): ***Evaluation of AGPS in Weak Signal Environments Using a Hardware Simulator***. Proceedings of GNSS 2004 (Session A6, Long Beach, CA, 21-24 September), The Institute of Navigation, Fairfax, VA

Klukas, R., G. Lachapelle, and M. Fattouche (1998): ***Cellular Telephone Positioning Using GPS Time Synchronization***. GPS World Magazine, 9, 4, 49-54.

Klukas, R., G. Lachapelle, and M. Fattouche (1997): ***Field Tests of a Cellular Telephone Positioning System***. Presented at IEEE VTC '97 - 47th Annual International Vehicular Technology Conference, Phoenix, AZ, May 5-7.

Klukas, R., O. Julien, L. Dong, M.E. Cannon, and G. Lachapelle (2004): ***Effects of Building Materials on UHF Ranging Signals***. GPS Solutions, 8, 1, 1-8.

Kuusniemi, H., G. Lachapelle, and J. Takala (2004a): ***Reliability in Personal Positioning***. CD-ROM Proceedings of GNSS 2004 Conference, Rotterdam (16-19 May), 14 pages.

Kuusniemi, H., and G. Lachapelle (2004b): ***GNSS Signal Reliability Testing in Urban and Indoor Environments***. Proceedings of NTM 2004 (San Diego, January 24-28), 210-224.

Lachapelle, G., H. Kuusniemi, D. Dao, G. MacGougan, and M.E. Cannon (2004): ***HSGPS Signal Analysis and Performance Under Various Indoor Conditions***. Navigation, U.S. Institute of Navigation, 51, 1, 29-43.

Lachapelle, G., M.E. Cannon, R. Klukas, S. Singh, R. Watson, P. Boulton, A. Read and K. Jones (2003): ***Hardware Simulator Models and Methodologies for Controlled Performance Assessment of High Sensitivity AGPS Receivers***. CD-ROM Proceedings of GNSS 2003, The European navigation Conference (Graz, Austria, 22-25 April), Session A2 on Indoor Navigation, 21 pages.

Macabiau, C., L. Ries, F. Bastide, and J.-L. Issler (2003): ***GPS L5 Receiver Implementation Issues***, Proceedings of the US Institute of Navigation GPS/GNSS (Portland, OR, USA, Sept. 9-12), pp. 153-164.

MacGougan, G., G. Lachapelle, R. Klukas, K. Siu, L. Garin, J. Shewfelt, and G. Cox (2002) ***Performance Analysis of A Stand-Alone High Sensitivity Receiver***. GPS Solutions, Springer Verlag, 6, 3, 179-195.

Mezentsev, O., J. Collin, H. Kuusniemi, and G. Lachapelle (2004): ***Accuracy Assessment of a High Sensitivity GPS Based Pedestrian Navigation System Aided by Low-Cost Sensors***. Proceedings of the 11th Saint Petersburg International Conference on Integrated Navigation Systems, 24-26 May, 156-164.

Mezentsev , O., Y. Lu, G. Lachapelle, and R. Klukas (2002): ***Vehicular Navigation in Urban Canyons Using a High Sensitivity Receiver Augmented with a Low Cost Sensor***. Proceedings of GPS2002 (Session E1, Portland, OR, 24-27 September), The Institute of Navigation, Alexandria, VA, 363-369.

Moeglein, M. and N. Krasner (1998): ***An Introduction to SnapTrack Server-Aided GPS Technology***. Proceedings of the Institute of Navigation ION GPS-98 (September 15-18, 1998, Nashville, Tennessee), pages 333–342.

Peterson, B., D. Bruckner, and S. Heye (1997): ***Measuring GPS Signals Indoors***. Proceedings of the Institute of Navigation ION GPS-97 (September 16-19, 1997, Kansas City, Missouri), pages 615–624.

Syrjärinne, J. (2001): ***Studies of Modern Techniques for Personal Positioning***, Publication 319. Phd, Tampere University of Technology.

van Diggelen, F. and C. Abraham (2001): ***Indoor GPS, The No-Chip Challenge***. GPS World, 12(9), pages 50–58.

Weill, L, N. Kishimoto, S. Hirata, and K. Chin (2004): ***The next generation of a super sensitive GPS system***. Proceedings of GNSS 2004 (Session A5, Long Beach, CA, 21-24 September), The Institute of Navigation, Fairfax, VA.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 12-15

Journal of Global Positioning Systems
2004

GPS & Galileo: Prospects for Building the Next Generation of Global Navigation Satellite Systems

Glen Gibbons, Jr.

Group Editorial Director and Association Publisher, GPS World, Advanstar Communications, Eugene, Oregon
Email: ggibbons@advanstar.com , Tel: 541-984-5286, Fax: 541-984-5333

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. In the next 5 to 10 years, the world will experience the emergence of a true Global Navigation Satellite System (GNSS) — a compatible and, in many respects, interoperable system of systems. The U.S. Global Positioning System, Europe's Galileo, perhaps Russia's Glonass system, and regional augmentations including the Wide Area Augmentation System (WAAS), the European Geostationary Navigation Overlay Service (EGNOS), radiobeacon-based systems such as the U.S. Nationwide Differential GPS, and compatible commercial differential correction services will comprise this multifaceted GNSS. Common signal structures and frequency plans will enable combined user equipment that reduces the technical complexity and cost, while vastly expanding related applications. Additional satellites and signals, both more powerful and with improved designs, will increase the availability of robust signal reception outdoors and strengthen the potential of indoor positioning using only GNSS user equipment. But the path to the future is not without its risks: political, technical, economic, and cultural.

Key words: GPS, Galileo, interoperability, compatibility, GNSS.

One of the nice things about being a journalist is that you're not shackled to providing the same ratio of data to opinions as scientists and researchers. As we listen to the presentations over the next few days, I expect we'll hear about 10 to 100 statements of test results for every conclusion offered by a speaker. Journalists, on the other hand, can present large quantities of opinion only slightly seasoned by facts and still be considered to have done their jobs.

I'll try not to abuse that latitude, however, as I take up the subject of the prospects for building the next generation of global navigation satellite systems.

After nearly a decade of distrust and bickering, Europe and the United States are showing signs of real harmony in the matter of global navigation satellite systems. Last June, the two powers signed an agreement that lays the foundation for substantive cooperation on GPS and Galileo — not merely in system compatibility and interoperability, but also in matters of trade and security.

In certain respects, one can imagine no more unlike enterprises than the U.S. Global Positioning System and Europe's Galileo system. GPS is operated by the U.S.

military establishment as a public entity; Galileo will be managed by a private consortium as a public-private enterprise fully under civil control. GPS uses one time standard; Galileo, another. The geodetic coordinate frameworks are different. Not all the frequencies match up and signal designs will vary. GPS is operated as a national system; Galileo is multinational — encompassing not merely the 25 nations joined in the European Union, but also the People's Republic of China, Israel, India, and a half dozen or more other nations with whom the EU has been talking. GPS delivers signals in space for free; Galileo proposes to deliver certifiable, guaranteed fee-based services in addition to a free open-access signal.

And then, of course, there's the most obvious difference: GPS is a real, existing system with 29 satellites in orbit and tens of millions of users around the world. Galileo is a work in progress. Galileo is a developmental program with a couple of billion euros in its pocket, some leased channels on telecommunication satellites to support the European Geostationary Navigation Overlay Service, a bunch of components not yet assembled into the first Galileo spacecraft (out of 30 planned for a full constellation), and a patchwork of ground infrastructure.

Galileo's original completion date, originally planned for 2008, may slip until at least 2010.

Despite the substantial design differences, the two GNSS systems are basically variations on a common technological theme. Over the long run, the political, institutional, and commercial realm is where interoperability may meet its greatest challenge. Technical experts will continually fine-tune frequency plans and signal structures. Equipment manufacturers will come up with ever-better products based on those designs. Service providers and end users will apply them in unpredictable and imaginative ways. But everywhere, these efforts will be facilitated — or constrained — by the business models, the rules adopted on intellectual property rights, tax policies, security arrangements, carriage requirements and regulatory policy, control and management of the space and ground infrastructures, international participation in the GNSS programs, and so forth.

Ironically, GPS and Galileo have inhabited a looking-glass world in which the two sides were sometimes as divided by their similarities as they were united by their differences. George Bernard Shaw once described England and the United States as being two great nations divided by a common language. After a few days in Sydney, I've come to believe that we could probably substitute Australia for England and the observation would still be true.

But, in any case, the experience with dueling GNSSes has demonstrated a similar principle of contrary dynamics. Never were the two sides so far apart as when Europeans first wanted to put themselves into the same GPS control room as the Americans and, later, when they wanted to put certain Galileo signals on some of the same frequencies as certain GPS signals.

Back in the mid-1990s, a delegation of officials from Brussels came to Washington, D.C., to discuss the idea of European participation in the management and operation of GPS. The Europeans said they'd even be willing to help pay for the operation and modernization of the American system.

The first thing the Americans asked was Who are you and whom do you represent? The European Union? France? Germany? Italy? Brussels? Our NATO allies? The European Commission? The European Space Agency? Who are we talking to? And the next thing the Americans said was, we don't need your money and we don't want it if it means we have to give up an iota of control over a key national infrastructure. And besides, you haven't actually allocated any money for GNSS, no serious money, anyway; just some study funds.

Well, the Europeans went off and set about answering those questions and, in the meantime, came up with a GNSS of their own — Galileo. Along the way, they also created another practical example of how to go about building a political union. Nothing sorts out the rhetoric from the real stuff as having to build a tangible system and service. After the Airbus project and implementation of the euro currency, Galileo already stands as a notable example of successful common effort by the European Union. At least, so far.

Anyway, after a few years, the European Commission came back to the United States once again and said, okay, let's talk about GNSS now. And, because the EU appeared to have its diplomatic act together, the United States set up an interagency working group, led by the State Department, to meet with the Europeans. Nonetheless, for the next couple of years, the two sides seemed to be talking past each other: the Europeans wanted to talk about specific details of the technical designs of the systems. The United States insisted on discussing more general matters such as trade policy and regulatory issues first. This went on until two things happened: first, the EU made a firm commitment of funding to build Galileo. And, second, the Europeans went ahead and came up with a provisional design on their own. Now, what got the United States' attention was a part of the proposed Galileo design that overlaid the publicly regulated service (PRS), an encrypted security-oriented signal, on top of part of the new GPS military signal (M-code) planned for the L1 band.

Once again, a seeming common ground — use of the same radio frequency — became a point of contention. U.S. defense officials argued strongly that the PRS overlay would undermine GPS operators' ability to jam non-military signals in a theater of operations without interfering with the M-code. At that point, the two sides began talking about all of the issues at once. They set up technical working groups – which sometimes met under secret classified conditions -- to come up with mutually workable solutions. The United States even went so far as to propose that GPS would use a similar signal structure as Galileo — the binary offset carrier (BOC) — if the EU would agree to a narrower frequency plan that moved the PRS away from the M-code.

Now, the technical compatibility and interoperability of these two GNSSes for which the initial U.S.-EU agreement has laid the foundation will definitely bring GPS and Galileo closer together. On the other hand, the differences between the two complementary systems will tend to bring the GNSS world closer together. By complementary, I mean the two systems are similar enough to be compatible, but different enough to be useful. Separate GPS and Galileo signals, separate ground and space infrastructures, separate operating entities, and separate budgets. These things will build the global GNSS marketplace and user community faster than one system alone. That will occur as a result of the increased redundancy, signal availability, robustness, and

ultimately, user confidence that result from having compatible but independent systems.

Not only will they be complementary systems, but they will be two primary systems – that is, each on its own will be capable of providing a complete positioning and timing service. This simple and seemingly self-evident concept has not fully taken root yet. A few years ago, about the time the European system was designated Galileo, I was moderating a GNSS panel at a conference in Toulouse, France – a centers of the European space industry. An official from EADS, one of the leading European defense and aerospace companies, made the observation, "It will be desirable to have a back-up GNSS." And I said, "Yes, GPS as a back-up to Galileo, right?" And the EADS official looked at me quite blankly, because he had meant the opposite. At the time, part of the argument for building Galileo was that it would provide a back-up for GPS in case the U.S. system experienced a failure. That rationale and the fact that GPS came first and had been an operational system for many years has created the sense of its primacy – even among public and private advocates for Galileo. That unspoken attitude still persists in some quarters, and probably will persist until an operational Galileo system has achieved true parity – or even a superior position – with GPS. At which point, either system will serve as a back-up – as well as a complement -- for the other.

In addition to these benefits, Galileo will help keep the United States honest in its management of GPS. Not that I think the U.S. government has been noticeably dishonest or narrowly manipulative in this matter. Quite the contrary, the United States has been remarkably open-handed in ensuring access to GPS by users around the world. In fact, the rapid adoption and spread of GPS technology and applications could not possibly have taken place the way it did without that policy. And it is a precedent that I believe Galileo's leaders would benefit from considering further.

Over the years, the United States has been criticized for many things regarding its GPS policies and management. But one thing that it did get right – perhaps in large part accidentally and almost unwittingly – was to make the civil signal open and free to users around the world.

Nonetheless, unilateral control, like unilateral policy-making, of such a potent global utility is an invitation to complacence and unresponsiveness by the system operator. Monopolies also tend to pose threats to technological innovation and economic progress.

To this end, the mere discussion of a European system has already benefited GNSS users, and Galileo's implementation will extend those benefits. I believe the prospect of Galileo contributed to the urgency to craft the first comprehensive U.S. presidential policy on GPS in 1996, to eliminate selective availability in May 2000, and to modernize the Block IIR generation of satellites. I believe that approval of Galileo's implementation by the European heads of state and authorization of public funding will help keep GPS modernization, including the GPS III program, on track. Completion of Galileo within the near term will definitely accelerate growth in GNSS product and service markets, as well as drive new applications. It could also encourage the United States to change its launch policy from launch-on-need, that is, replacing satellites only as they fail. That means that many critical innovations in GPS signal and system design have to wait until after launch of all the unused satellites with technology that has been outstripped by technical advances. Successful completion of a Galileo constellation with new signals and higher power could encourage GPS' managers to launch on a planned schedule to more quickly install a new operational capability with the modernized GPS signals and satellites.

So, all this comes as good news for GNSS equipment manufacturers and users around the world. But many objectives must still be achieved and many obstacles, avoided, before compatible, interoperable GNSS becomes a reality. An example of the kinds of things that can derail this process can be seen in a recent article in a British newspaper. The article described an exchange between U.S. and European officials attending a conference on military space relations that led to one U.S. delegate suggesting that the United States would attack Galileo satellites if they continued transmitting signals that might be used by adversaries in a theater of conflict.

Now, I would not invest this anecdote with too much weight or power – even if it is completely true, even if these comments were actually made. I believe that they represent more an expression of anxiety than of intention. Indeed, I think that we should all share the anxiety of misuse of GNSS, whether GPS or Galileo. But, as I understand the NAVWAR scenarios, capabilities, and solutions developed by the U.S. Defense Department, the primary means to prevent hostile use of GNSS will be much more benign, limited in scope and targeted against the perpetrators and not GNSS system operators. And between any GNSS-related crisis and an assault on Galileo satellites or infrastructure stands the agreement signed in June, which established the official channel for relations between the European Union and the United States in GNSS matters.

So, what are some of the things that need to be done to continue the auspicious beginning on GPS and Galileo cooperation? Well, here a few suggestions:

** Establishing a permanent mechanism for regular political consultations on the GNSS agenda, which must inevitably evolve as the systems mature and modernize.

** Cooperation in system operations and open formal lines of communications, 24/7, between GPS and Galileo

controllers — whether that's a black box or a red telephone, or even an exchange of liaison officers in master control stations.

** A further agreement on security-related matters that sketches out the appropriate actions for possible threat scenarios. That could even include creation of a joint security board for assessing threats against either system, evaluating situations that might require jamming or degradation of civil signals, and recommending appropriate courses of action. Of course, actual events rarely take the exact form or follow the exact course anticipated by contingency plans. But forward-looking conventions would increase the state of readiness, the famikliarity of GNSS olperators and officials with their counterparts, and the capacity for responding to threat situations in whatever form they may arise.

** Clear statements on the reciprocal role that industrial partners from the United States and Europe can have in building and operating the other's GNSS. The Boeing Company has included Alenia Spazio and Alcatel Space on its GPS III team, EADS-Astrium is working with Boeing in GPS/Galileo matters, and the iNavsat consortium has done the same with Lockheed Martin, SiRF Technology, and NavTeq in its efforts to secure the Galileo concession.

But more needs to be done. The Galileo Joint Undertaking, the Galileo Supervisory Authority, or, if necessary, the European Council of Transport Ministers should provide a clear statement on status, ownership, and access to the Galileo equivalent of the GPS Interface Control Document (ICD). The GPS ICD provides complete technical specifications that enable manufacturers to build GPS equipment.

The two sides should also clarify the rules for U.S. companies' participation in building, maintaining, and operating Galileo. At the same time, the United States should clarify its guidelines on the export of GNSS-related technologies and the allowable scope for European industrial participation in the GPS III program.

** Agreements on carriage requirements for airplanes and commercial vessels that minimize the financial burdens on the transport companies and maximizes the use of combined GPS/Galileo equipment.

In closing, I'd like to end with a small warning. You often hear people say that the uses of GPS are limited only by the human imagination. I've used the expression myself. It has a wonderful gee-whiz quality to it -- Limited only by the human imagination. And with a second, interoperable GNSS, I guess that what? the human imagination will get twice as big, or GNSS innovations will take place twice as fast, or something like that?

But I had an experience a few days ago that has led me to think about that idea a little more closely: When I arrived in Sydney on Saturday morning, I went down to the Royal Botanical Gardens to try to walk off some of my jet lag. One of the many amazing facts that I learned there is that the world contains 80,000 species of edible plants. However, only 20 species make up 90 percent of the food actually eaten by the world's population – things like corn, wheat, and rice. So, before we start feeling too smug about the prospects of GNSS, and assuming that a second system is going to make things twice as good, I think we need to recall our track record with the human diet. Only implementing 20 out 80,000 options isn't so hot, and we need to do better with GNSS. And that will take not merely imagination, but hard work, good intentions, and sustained effort.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 16-24

Journal of Global Positioning Systems
2004
© CPGPS

The Deformation of Bromo Volcano (Indonesia) as Detected by GPS Surveys Method

Hasanuddin Z. Abidin, H. Andreas, M. Gamal

Department of Geodetic Engineering, Institute of Technology Bandung (ITB), Jl. Ganesha 10, Bandung 40132, INDONESIA
e-mail: hzabidin@gd.itb.ac.id Tel: 62-22-2534286; Fax: 62-22-2534286

M. Hendrasto, Ony K. Suganda, M.A. Purbawinata

Directorate of Vulcanology and Geological Hazard Mitigation (DVGHM), Jl. Diponegoro 57, Bandung 40132, INDONESIA

Irwan Meilano, F. Kimata

Research Center for Seismology and Volcanology and Disaster Mitigation (RCSVDM), Nagoya University, Nagoya, JAPAN

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Bromo is an active type-A volcano located inside Tengger caldera in East Java province of Indonesia. In her history, Bromo has erupted at least about 50 times since 1775. The last eruption occurred on June 2004. Monitoring of Bromo activities has been continuously done since early 1989 by using seismograph. EDM and GPS surveys have also been conducted since the last eruption in Dec. 2000. Up to now there have been four GPS surveys that have been conducted, namely on Dec. 2000, June 2002, August 2003, and June 2004, respectively. The obtained GPS and EDM results show that the deformation of Bromo volcano is typically in order of a few cm, with the inflation and deflation processes before and after the eruption. Estimated location of the pressure source is found to be beneath the active crater with depth of about 1 km below the caldera floor.

Key words: Bromo, volcano, deformation, GPS, EDM

1 Introduction

Indonesia has 129 active volcanoes and 271 eruption points as a consequence of interactions and collisions among several continental plates. The most actives volcanoes are shown in Figure 1, which one of them is Bromo volcano (Tengger caldera) in East Java. With the people of around 200 million and the fact that most populated island in Indonesia (i.e. Java) has the most number of active volcanoes then it is obvious that the Indonesian people are always live under threat of volcanic eruptions. According to (Katili & Siswowidjojo, 1994), around 10% of Indonesia people live in the area endangered by the volcanic eruptions, and about 3 million of them live in the danger zones. This fact alone suggests that in Indonesia the monitoring of volcano activity should be performed not only routinely but should also be done as good as possible.

Many methods have been utilized for observing and monitoring the volcano activity, and the efforts for establishing a good and reliable system for monitoring and predicting the volcano eruption is never ended (McGuire et al., 1995; Scarpa and Tilling, 1996). In relation with the deformation of volcano, it is already known that the explosive eruptions are usually preceded by the relatively large inflation of its body (Scarpa and Gasparini, 1996). In the case of the volcano, which has been quiet for sometimes, the deformation of its body is one of the reliable indicators of its reawakening activity. According to Van der Laat (1996), this point displacement could reach several tens of meter for silicic volcano with dome formation; while for volcano with magma chamber still very deep below or its magma movement is relatively slow, the observed deformation is relatively small, where its strain value sometimes is smaller than 0.1 ppm/year. Moreover, according to Van der Laat (1996) and Dvorak and Dzurisin (1997), the deformation of volcano body represented by the point displacement vectors and their velocity vectors could provide information on the characteristics and dynamics of magma chamber. In other word, the deformation information could be modelled to derive the location,

depth, shape, and size of the pressure source causing the deformation, which is usually called the magma (and hydrothermal) chamber; while the information on the deformation rate could be used to derive the pressure variations inside the magma chamber, which then can be used in predicting the magma supply rate to the volcano and its outgoing volume in case of eruption (Dvorak and Dzurisin, 1997).

Fig. 1 Location of major active volcanoes of Indonesia.

Volcano deformation monitoring could be performed by utilizing several methods, such as EDM (Electronic Distance Measurement), levelling, tiltmeter measurements, GPS surveys and InSAR (Interferometric Synthetic Aperture Radar) (Massonnet and Feigl, 1998). In this paper, the deformation characteristics of Bromo volcano as observed by GPS and EDM surveys will be presented and discussed. The explanation is mainly based on the results and experiences obtained from four GPS surveys that have been conducted, namely on December 2000, June 2002, August 2003, and June 2004, respectively; and four EDM surveys performed on December 2000, June 2002, August 2003 and June 2004.

2 Bromo Volcano

Bromo is an active type-A volcano located inside Tengger caldera in East Java province of Indonesian. It is the youngest of several other volcanic cones located inside a caldera, e.g. Widodaren, Kursi, Segorowedi, and Batok as shown in Figure 2. The towering conical peak of Semeru, Java's highest volcano, appears in the background. The caldera is 9 km x 10 km in size, surrounded by vertical cliffs of about 50 to 500 m in high, composed by volcanic rock layers of ancient Tengger Volcano (DVGHM, 2004). The caldera floor consists of widely distributed sand deposit in the northern part, known as A Sand Sea Caldera. In the eastern and southern part, the caldera floor covered by grass. Bromo crater has a size of 600m by 800m, with several explosion holes inside. Its summit has an altitude of about 2329 m above sea level and about 133 m above the caldera floor.

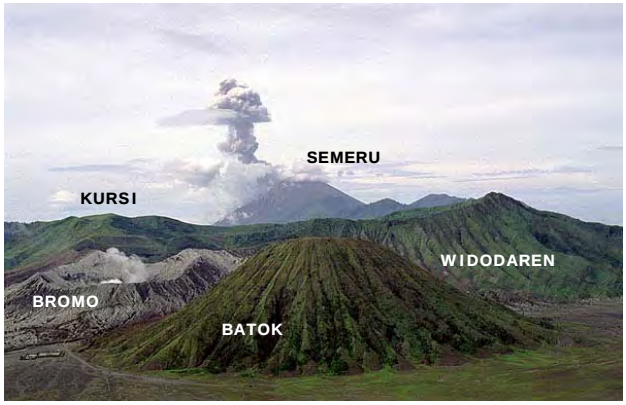


Fig. 2 Bromo and her surrounding,
adapted from (Decade Volcano, 2004).

In her history, Bromo has erupted at least about 50 times since 1775, as shown in Table 1. The eruptions were mostly Strombolian in type. On the basis of geologic data, historic activity of Bromo and morphology of Tengger Caldera, the potential hazards of Bromo eruption are ejected (glowing) rock fragments, heavy ashfalls and toxic gases. So far, lahar hazards have never occurred during Bromo eruptions.

Tab. 1 Eruption Years of Bromo Volcano.

1775(?), 1767(?), 1804, 1815, 1820, 1822-23, 1825, 1829, 1830, 1835, 1842, 1843, 1844, 1856, 1857, 1858, 1859, 1860, 1865, 1866, 1867, 1868, 1877, 1885, 1886, 1887, 1888, 1890, 1893, 1896, 1906, 1907, 1908, 1909, 1910, 1915, 1916, 1921, 1922, 1928, 1929, 1930, 1935, 1940, 1948, 1949, 1950, 1956, 1972, 1980, 1983, 1984, 1995, 2000, 2004

The two last eruptions occurred on December 2000 and June 2004, respectively. In the last eruption, Bromo was abrubtly erupted on 8 June 2004 at 15.26 local time for about 20 minutes. According to DVGHM (2004) the explosion was a phreatic type eruption revealed the ash column rose to 3000 m above the crater rim. Ash material and stones spread around the crater away in radius of about 300 m. On 9 June 2004 the volcanic activities were decreased and no more explosion was occurred. However there was still the white grey ash plume rose to about 10-25 m from the crater rim. On 10 and 11 June 2004 the volcanic activity of Bromo is still relatively low. On 13 and 14 June 2004 there are small explosions happened causing ash material to rise up to about 100 m. On 15 June 2004, volcanic activity of Bromo generally decreased, and its status hazard is reduced to lower level.

3 GPS and EDM Surveys in Bromo Volcano

Monitoring of Bromo activities has been continuously done since early 1989 by using seismograph. The GPS and EDM surveys have also been conducted since the last eruption in Dec. 2000. Up to now there have been four surveys conducted for GPS as shown in Table 2. GPS and EDM surveys were conducted in the same periods, except for the first survey and September 2002. The first EDM survey was conducted during the 2000 eruption period, i.e. December 2000.

Tab. 2 Timing of GPS and EDM Surveys.

GPS Survey	Survey Period	Observed
Survey - 1	11 Feb. 2001	POS, BTOK, BRMO, KRSI
Survey - 2	21-22 June 2002	POS, BTOK, BRMO, KRSI, BANT, WDDRN
Survey - 3	21-22 Agustus 2003	POS, BTOK, BRMO, KRSI, BANT, WDRN, KWAH

Configuration of GPS deformation monitoring network for Bromo volcano is shown in Figure 3. POS is the reference GPS point located around Bromo observatory at the rim of caldera. WDRN and BANT points are newly included in the second GPS survey. KWAH is the point on the crater's rim and established in the third GPS survey. The distances between POS and other GPS points are less than 10 km.

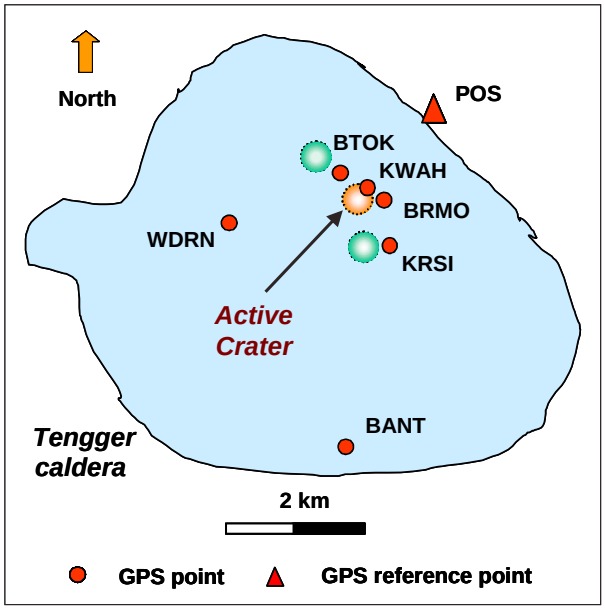


Fig. 3 GPS monitoring network in Bromo.

GPS surveys were conducted jointly by the Dept. of Geodetic Engineering ITB, DVGHM and RCSVDM Nagoya University. GPS surveys were conducted using GPS dual-frequency geodetic type receivers, with observation session ranging from 6 to 12 hours. GPS surveys a few days after the 2004 eruption, e.g. on 11, 12, 15 and 16 June 2004, were conducted only using two GPS receivers.

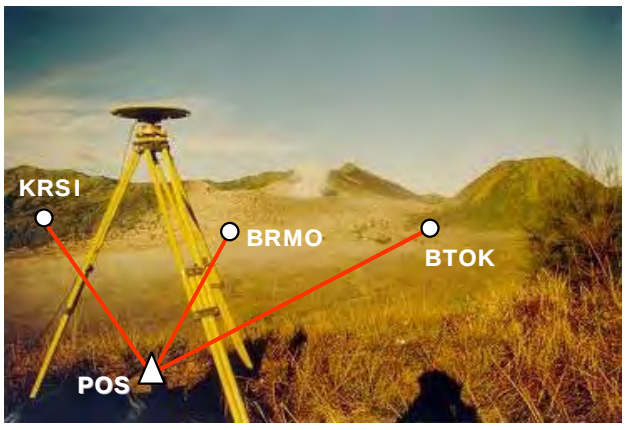


Fig. 4 EDM monitoring network in Bromo.

The EDM measurements were made between POS station and BATOK, BROMO, and KURSI stations, as shown in Figure 4. The EDM surveys were conducted using Leica DI 20 for the first survey and Leica DI 3000 for the second.

4 Data Processing and Results

GPS data processing were done using BERNESSE 4.2 scientific software (Beutler et al., 2001). Processing is radially performed from POS. The coordinate of POS itself is determined from BAKOSURTANAL, the IGS point located near Jakarta, capital of Indonesia. Precise ephemeris was used for the processing. The final results shows that several mm precision level were achieved for all components of coordinates, as shown in Figure 5.

These results also indicate that GPS data processing has properly done. Based on GPS coordinates estimated from four GPS surveys that have been conducted, several deformation characteristics of Bromo volcano can be inferred, as discussed in the following sub-chapters. The related results obtained from EDM surveys will also be given along the discussion.

4.1 Horizontal and Vertical Displacements

The displacement vectors (dE,dN,dh) are obtained by differencing the estimated coordinates of GPS stations obtained from two consecutive surveys. The displacement vectors obtained from four GPS surveys and their displacement magnitudes (δd) are shown in Table 3.

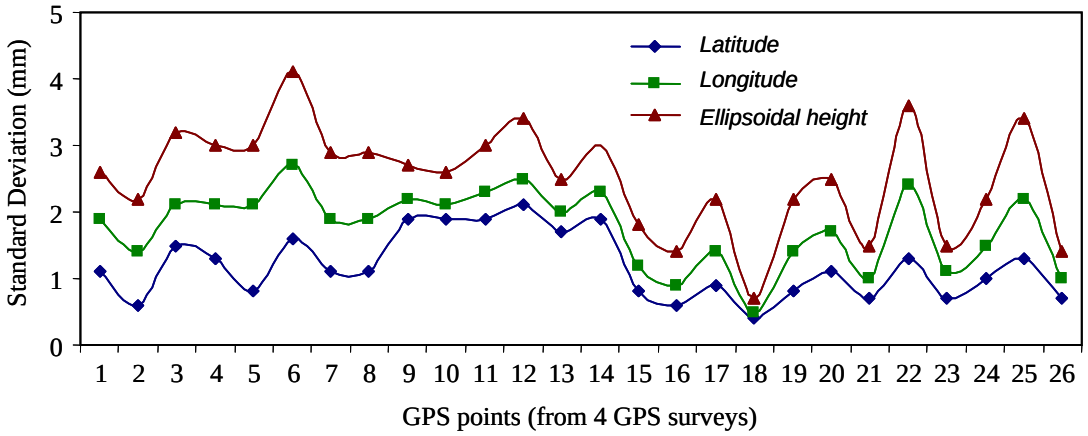


Fig. 5 Standard deviations of the estimated coordinates of GPS stations from 4 surveys.

Tab. 3 Coordinate differences and displacement of GPS stations and their standard deviations.

GPS Surveys : 1 to 2	**Coordinate differences (mm)**						**Displacements (mm)**	
	dE	**σ(dE)**	**dN**	**σ(dN)**	**dh**	**σ(dh)**	**δd**	**σ(δd)**
BTOK	-0.3	2.8	17.8	1.7	-0.4	4.0	17.8	1.7
BRMO	17.0	2.5	8.6	1.0	47.6	3.7	51.3	3.6
KRSI	4.0	3.4	3.5	2.2	29.5	5.2	30.0	5.2
GPS Surveys : 2 to 3	**Coordinate differences (mm)**						**Displacements (mm)**	
	dE	**σ(dE)**	**dN**	**σ(dN)**	**dh**	**σ(dh)**	**δd**	**σ(δd)**
BTOK	20.7	2.6	-8.0	1.6	16.8	3.7	27.8	3.4
BRMO	18.3	2.5	-14.0	1.2	23.1	3.5	32.6	3.1
KRSI	-7.3	3.0	-7.8	1.8	-152.6	4.6	153.0	4.6
BANT	6.8	2.9	-2.6	2.0	2.3	4.5	7.6	4.5
WDRN	-4.9	1.9	-19.9	1.3	23.8	2.9	31.4	2.4
GPS Surveys : 3 to 4	**Coordinate differences (mm)**						**Displacements (mm)**	
	dE	**σ(dE)**	**dN**	**σ(dN)**	**dh**	**σ(dh)**	**δd**	**σ(δd)**
BTOK	12.0	1.9	5.4	1.3	8.1	2.8	15.5	2.1
BRMO	-25.0	1.6	8.5	1.1	-2.3	2.3	26.5	1.4
KRSI	-2.8	1.9	35.4	1.2	12.9	3.0	37.8	1.5
BANT	-8.8	2.3	-15.7	1.7	63.4	3.6	65.9	3.5
WDRN	-15.0	1.4	-15.3	1.1	1.0	2.2	21.4	1.2

In order to statistically check the significance of the displacements derived by GPS surveys, the congruency test [Caspary, 1987] was performed on the following variable

$$\delta d_{ij} = (dE_{ij}^2 + dN_{ij}^2 + dh_{ij}^2)^{1/2} . \quad (1)$$

where δdij is the displacement of a station from epoch i to j. The null hypothesis of the test is that there is no displacement between the epochs. Therefore:

$$\text{null hypothesis} \quad Ho : \delta dij = 0 , \quad (2)$$

$$\text{alternative hypothesis} \quad Ha : \delta dij \neq 0 . \quad (3)$$

The test statistics for this test is:

$$T = \delta dij / \sigma(\delta dij) , \quad (4)$$

which has a Student's t-distribution if Ho is true. The region where the null hypothesis is rejected is [Wolf and Gilani, 1997] :

$$| T | > t_{df,\alpha/2} , \quad (5)$$

where df is the degrees of freedom and α is the significance level used for the test. In our case, for GPS baselines derived using 6 to 12 hours of GPS data with 30 seconds data interval, then df → ∞. Please note that a t-distribution with infinite degree of freedom is identical to a normal distribution. If a confidence level of 99% (i.e. α=1%) is used, then the critical value t ∞,0.005 ∞ is equal to 2.576 [Wolf and Gilani, 1997]. The displacements of GPS points and their standard deviations, i.e. δd and σ(δd), are shown in Table 3 above.

If the values are adopted for the congruency test, then the testing results show that significant displacement is found in all the stations in all three observed periods, except for station BANT in the second observed period (e.g. June 2002 – August 2003). Based on the testing results, it could be statistically concluded that with 99% confidence level there were deformation phenomena of Bromo volcano as observed by four GPS surveys that have been conducted. The magnitude of deformation is in general in the level of a few cm.

The example is shown in Figure 6, which is derived from the first and second GPS surveys conducted in February 2001 and June 2002. These displacements indicate the inflation of Bromo volcano during that observed period, after the eruption of December 2000.

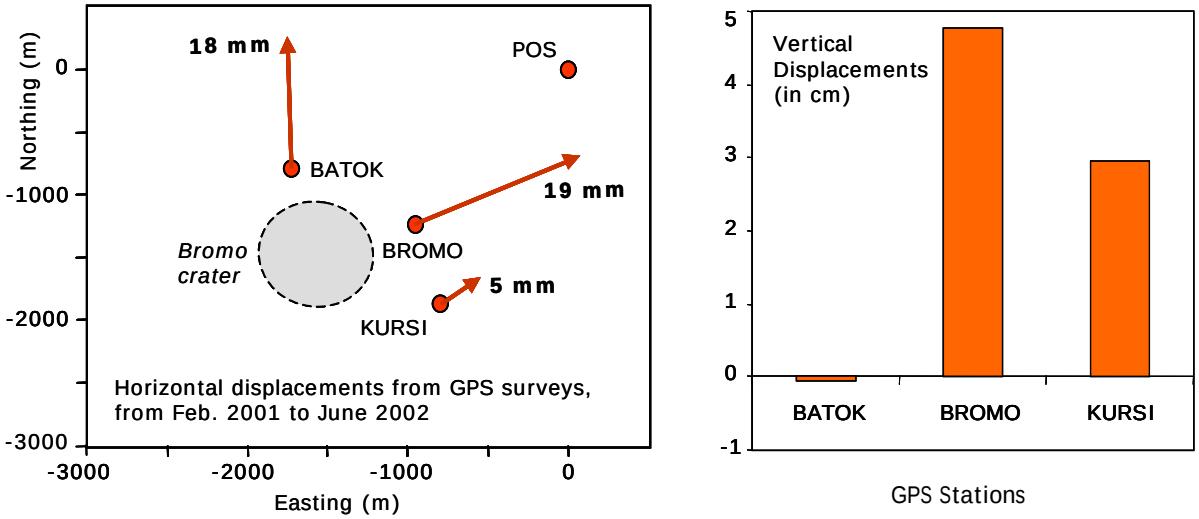


Fig. 6 Displacement vectors of from GPS surveys.

4.2 Horizontal Distance Changes

The horizontal distances changes computed based on the estimated GPS coordinates of the observed stations are also useful for studying the volcano deformation. Figure 7 gives an example of horizontal distance changes of (BRMO-BTOK) baseline from the first to the last GPS surveys. The baseline is closed to the active crater of Broom, and therefore should be more sensitive to deformation process of volcano.

Figure 7 shows that since February 2001, the inflation process of Bromo was occurring for about a few cm in three years before erupting again in June 8th, 2004. A few days after the eruption, a deflation process was occurring for about a few cm in just a week. The deflation process after the June 2004 eruption of Bromo is also indicated by the results of EDM measurement as shown in Figure 8.

The horizontal distance changes of all GPS baselines related to the June 8th, 2004 eruption are shown Table 4. The results indicate that the ground deformation related to June 8 eruption were relatively small, just in the order

of a few cm. It also shows that in the period between August 2003 and 24 June 2004 there was a sign of deflation of the active crater as also shown in Figure 7. The horizontal shortening of BTOK-BRMO and BTOK-KRSI baselines as much as 3.5 cm and 3.2 cm indicates it. More data however is required to reveal the more detail ground deformation pattern of Bromo volcano before and after June 8 eruption.

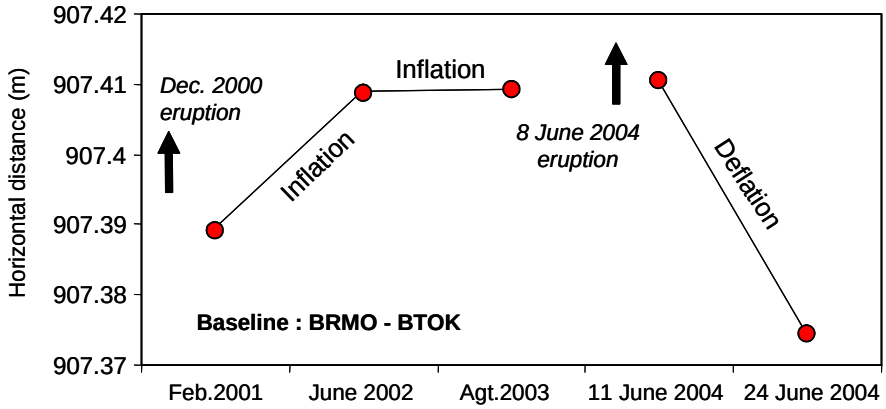


Fig. **7** Horizontal distance changes of (BRMO-BTOK) baseline.

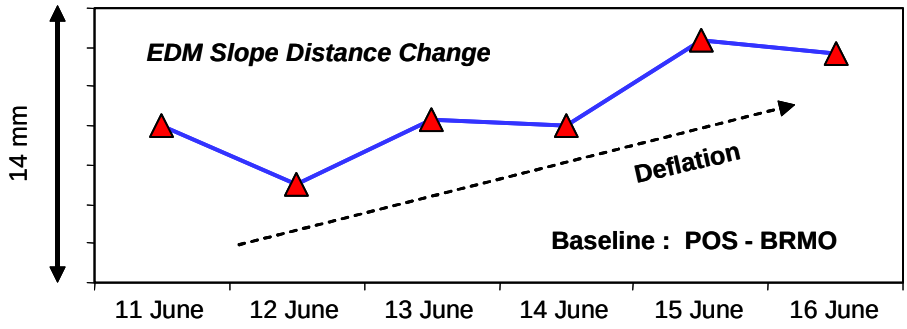


Fig. 8 Slope distance changes of (POS-BRMO) baseline as measured by EDM after the June 8th 2004 eruption.

Tab. 4 GPS derived horizontal distance changes (the eruption occurred on June 8th at 15:26 pm for about 20 minutes).

Baseline	Distance changes in cm w.r.t. August 2003				
	11 June	**12 June**	**15 June**	**16 June**	**24 June**
POS-BTOK	1.8	no data	no data	no data	-1.3
POS-BRMO	0.9	no data	0.0	no data	0.8
POS-KRSI	no data	-1.6	no data	-1.5	-3.2
POS-BANT	no data	no data	no data	no data	1.8
POS-WDRN	no data	0.1	no data	0.9	2.0
BTOK-BRMO	0.1	no data	no data	no data	-3.5
BTOK-KRSI	no data	no data	no data	no data	-3.2
KRSI-WDRN	no data	-1.2	no data	-0.3	0.6

4.3 Kinematic Positioning Results

In order to see the temporal variation of coordinate components two eeks after the eruption period, e.g. on June 23rd, GPS data were processed in differential kinematic mode using Bernesse 4.2 [Beutler et al., 2001]. The positioning results in North-South (NS), East-West (EW), and Up-Down (UD) components are given in Figures 9, 10 and 11, respectively.

The results given in these Figures show that two weeks after the eruption, i.e. on June 23rd, no deformation signals appear in the horizontal components of coordinates (NS and SW components). In the vertical component, however, there are variations shown with a few cm variations. These variations, however, may not be directly associated with the real vertical deformation of volcano, since no strong explosions were occurring during that day. More investigation is needed to

understand the temporal variations of UD components shown in Figure 11.

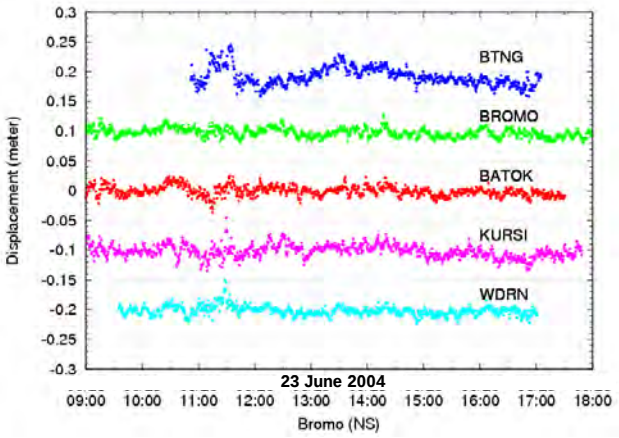


Fig. **9** Temporal NS Variation of GPS Stations (23 June 2004).

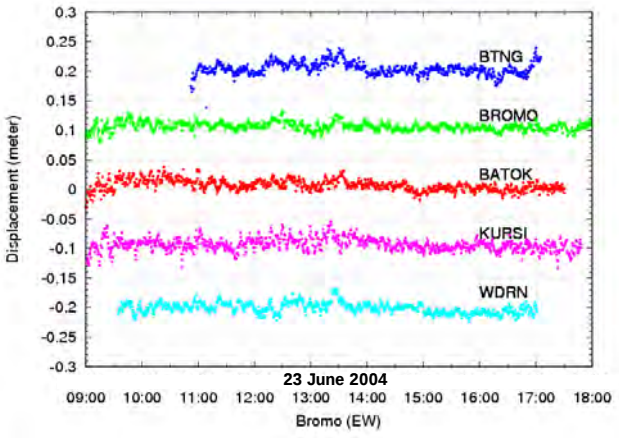


Fig. 10 Temporal EW Variation of GPS Stations (23 June 2004).

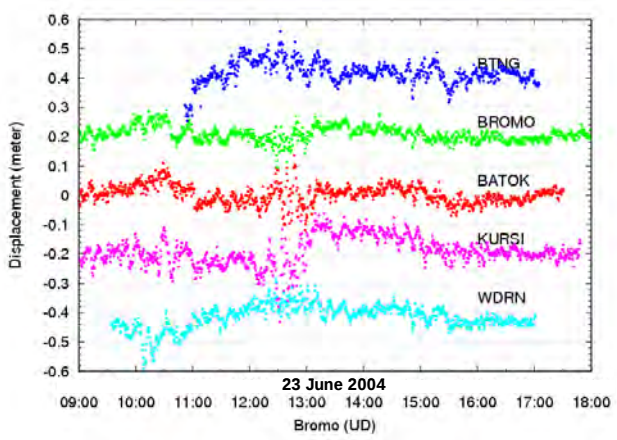


Fig. 11 Temporal UD Variation of GPS Stations (23 June 2004).

4.4 Pressure Source Characteristics

The other problem which is interesting to be investigated is related to the determination of pressure source characteristics, i.e. its depth, size, shape, and supply rate, from the GPS derived ground displacement vectors of the points located on the body of volcano and its surrounding area. Pressure source modelling of volcano deformation actually can be based on several models [Mogi, 1958; Okada, 1985; Trasatti, 2003].

It should be noted in this case that Mogi model [Mogi, 1958], as shown in Figure 12, is by far the most used and the simplest model to fit ground deformations in volcanic area. The model assumes that the Earth's crust consists of elastic half-space, the source of deformation is small and spherical as a point like-source with radial expansion, and it exerts hydrostatic pressure on the surrounding rocks. Whilst none of those assumptions strictly apply, many volcanoes show deformation patterns close to Mogi theoretical model [McGuire et al., 1995].

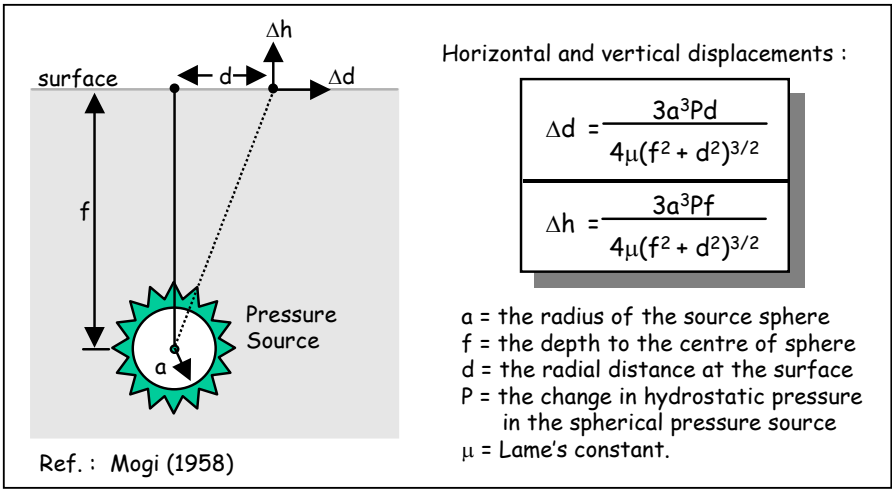


Fig. 12 Mogi Model.

Based on the results of first and second GPS surveys, Mogi model was then applied to estimate location of the pressure source. Only GPS horizontal displacement vectors were used in this estimation process and algorithm proposed by Nishi et al. (1999) was applieed. Estimated location of the pressure source is found to be beneath the active crater with depth of about 1 km below the caldera floor. Figure 13 shows estimated location of the pressure source. Its local coordinates (Easting, Northing) with POS as the initial point is about (-2 km, -2 km).

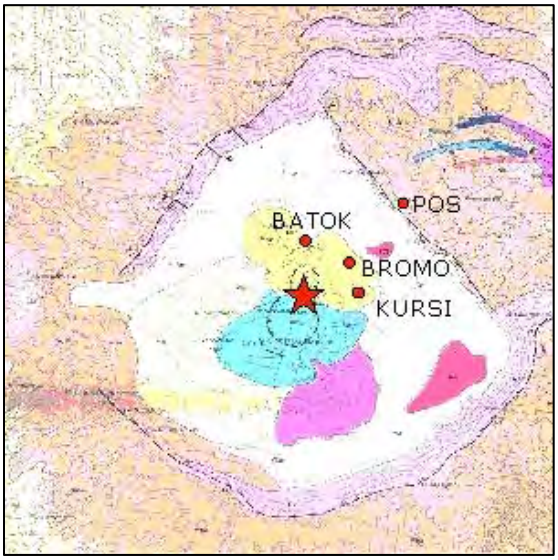


Fig. 13 Location of the pressure source of Bromo volcano (indicated by the star sign).

5 Closing Remarks

Based on the results of GPS surveys that have been conducted since February 2001 up to June 2004, it can be inferred that the deformation of Bromo volcano before and after eruption is relatively small, i.e. typically in order of a few cm; with the inflation and deflation processes before and after the eruption. By using Mogi model the estimated location of the pressure source is found to be beneath the active crater with depth of about 1 km below the caldera floor.

Based on the results obtained from deformation monitoring in Bromo and several other volcanoes in Indonesia, it can be concluded that GPS survey method is a reliable method for studying and monitoring volcano deformation. The method is capable of detecting the deformation signal that has a relatively small magnitude in the order of a few cm, or even several mm; although achieving this level of accuracy is not an easy task to do. In this case the use of dual frequency geodetic type receivers is compulsory along with good survey planning, stringent observation strategy, and stringent data processing strategy using the scientific software. Considering its relatively high accuracy, all-times weather-independent operational capability, wide spatial coverage, and its user friendliness, the use of repeated GPS surveys for volcano deformation monitoring is highly recommended.

Besides those technical aspects, in monitoring the volcano deformation using repeated GPS surveys there are some non-technical aspects that have to be considered and treated properly in order to achieve a relatively good monitoring performance. Based on our experiences gained from doing the surveys in Bromo and some other volcanoes, the operational issues such as team movement strategy, availability of sufficient power supply and local labours, preparation of logistic and accommodation for survey personnel, and communication mechanism among survey teams, are the issues which took the most efforts and times to handle and accomplish. The unfriendly and harsh environment of volcano also should be taken into consideration in selecting the survey team member.

Although GPS surveys could provide accurate and precise ground displacement vectors, however in order to have a better and more detail information on volcano deformation characteristics, GPS survey method should be integrated with other monitoring techniques such as EDM, levelling, tiltmeter measurements and InSAR (Interferometric Synthetic Aperture Radar) (Massonnet and Feigl, 1998). With many available data and information, a more reliable deformation and pressure source modelling can be better performed.

Acknowledgements

This research project is possible through the cooperation between the Department of Geodetic Engineering ITB, Directorate of Volcanology and Geological Hazard Mitigation (DVGHM), and Research Center for Seismology and Volcanology and Disaster Mitigation (RCSVDM), Nagoya University Japan. The administrative, transportation, accommodation and financials support given by these three institutions are highly appreciated. We also thank the ITB and RCSVDM students and the staffs of DVGHM and RCSVDM for participating in the GPS surveys operation in Bromo volcano.

References

Beutler, G., H. Bock, E. Brockmann, R. Dach, P. Fridez, W. Gurtner, U. Hugentobler, D. Ineichen, J. Johnson, M.

Meindl, L. Mervant, M. Rothacher, S. Schaer, T. Springer, R. Weber (2001). ***Bernese GPS software version 4.2***. U. Hugentobler, S. Schaer, P. Fridez (Eds.), Astronomical Institute, University of Berne, 515 pp.

Caspary, W.F. (1987). ***Concepts of Network and Deformation Analysis***. Monograph 11, School of Surveying, The University of New South Wales, Kensington, NSW, Australia, 183 pp.

Decade Volcano (2004). ***Homepage of Decade Volcano***. Address : http://www.decadevolcano.net/volcanoes/indonesia/tengger_caldera/bromo.htm. Accessed date : 1 October 2004.

DVGHM (2004). ***Homepage of Directorate of Volcanology and Geological Hazard Mitigation***. Address : http://www.vsi.esdm.go.id/. Accessed months : September – October.

Dvorak, J.J. and D. Dzurisin (1997). Volcano Geodesy : The Search for Magma Reservoirs and the Formation of Eruptive Vents. ***Review of Geophysics***, Vol. 35, No. 3, August, pp. 343 - 384.

Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins (1994). ***Global Positioning System, Theory and Practice***. Third, Revised Edition, Springer Verlag, Wien.

Katili. J.A. and S.S. Siswowidjojo (1994). ***Pemantauan Gunungapi di Filipina dan Indonesia***. Ikatan Ahli Geologi Indonesia (IAGI). ISBN: 979-8126-05-6. 321 h + xii.

Leick. A. (1995). ***GPS Satellite Surveying***. John Wiley & Sons. Second edition. New York. ISBN 0-471-30626-6. 560 pp.

Massonnet, D. and K.L. Feigl (1998). ***Radar Interferometry and its Application to Changes in the Earth's Surface***. Reviews of Geophysics, Vol. 36, No. 4, November, pp. 441-500.

McGuire, W.J. (1995). ***Monitoring active volcanoes – an introduction***, In Monitoring Active Volcanoes by B. McGuire, C.R.J. Kilburn, and J. Murray (Eds), pp. 1-31, UCL Press Limited, London, 421 pp.

McGuire B.. C.R.J. Kilburn. and J. Murray (Eds) (1995) ***Monitoring Active Volcanoes***. UCL Press Limited. London. 421 pp.

Mogi, K. (1958). Relation between the eruption of various volcanoes and the deformation of the ground surface around them. ***Bulletin of Earthquake Research Institute***, Vol. 36, pp. 99-134.

Nishi,K., H. Ono, H. Mori (1999). Global Positioning System measurements of ground deformation caused by magma intrusion and lava discharge: the 1990-1995 eruption at Unzendake volcano, Kyushu, Japan. ***Journal of Volcanology and Geothermal Research***, Vol. 89, Nos. 1-4, pp. 23-34.

Okada, Y. (1985). ***Surface deformation due to shear and tensile faults in a half-space***. Bulletin Seismological Society of America, Vol. 75, pp. 1135-1154.

Scarpa, R. and P. Gasparini (1996) ***A Review of Volcano Geophysics and Volcano-Monitoring Methods***. In Monitoring and Mitigation of Volcano Hazards by R. Scarpa and R.I. Tilling (Eds.), Springer Verlag, Berlin, pp. 3 - 22.

Scarpa. R. and R.I. Tilling (Eds.) (1996). ***Monitoring and Mitigation of Volcano Hazards***. Springer Verlag. Berlin. 841 pp.

Trasatti, E., C. Giunchi, M. Bonafede (2003). ***Effects of topography and rheological layering on ground deformation in volcanic regions***, Journal of Volcanology and Geothermal Research, Vol. 122, pp. 89-110.

Van der Laat, R. (1996) ***Ground-Deformation Methods and Results***. In Monitoring and Mitigation of Volcano Hazards by R. Scarpa and R.I. Tilling (Eds.), Springer Verlag, Berlin, pp. 147 - 168.

Wolf, Paul R. and C.D. Ghilani (1997). ***Adjustment Computations, Statistics and Least Squares in Surveying and GIS***. John Wiley & Sons, Inc., New York, 564 pp.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 25-31

Journal of Global Positioning Systems
2004
© CPGPS

An Assisted GPS Acquisition Method using L2 Civil Signal in Weak Signal Environment

Deuk Jae Cho

Department of Electronics, Chungnam National University, Korea
e-mail: panda@cslab.cnu.ac.kr; Tel: +82-42-825-3991; Fax: +82-42-823-4494

Chansik Park

School of Electrical and Computer Engineering, Chungbuk National University, Korea
e-mail: chansp@chungbuk.ac.kr; Tel: +82-43-261-3259; Fax: +82-43-268-2386

Sang Jeong Lee

Division of Electrical and Computer Engineering, Chungnam National University, Korea
e-mail: eesjl@cslab.cnu.ac.kr; Tel: +82-42-821-6582; Fax: +82-42-823-4494

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Recently, there has been increasing demands on the positioning capability in weak signal environment such as inside building and urban area. The present assisted GPS technology uses GPS L1 signals only. Meanwhile, according to the GPS modernization plan, Block IIR-M GPS satellite will be first launched in 2005, transmitting the civil code in L2 frequency as well as in L1 frequency with the updated signal structure. Since the L2 civil code has a worst-case cross correlation performance of 45 dB (over 251 times better than 21 dB cross correlation performance of the L1 C/A code), it will be much more effective in weak signal environment. This paper proposes an assisted GPS acquisition method using L2 civil signals. It will show that the acquisition success rate of the proposed assisted GPS acquisition method is better than that of the existing assisted GPS method using L1 signals in the same environment. The constellation of the next generation GPS satellites is scheduled to launch in 2005. Therefore, in order to design and test the assisted GPS acquiring the L2 civil signal, it is necessary to design a signal generator which can generate the L2 civil signal. The signal generator will be designed using the pseudo random noise (PRN) code generation method and navigation message protocol defined in GPS ICD PIRN 200C-007B. Finally, through the simulations using the designed signal transmitter, the success rate of the proposed assisted GPS acquisition method will be compared with that of the existing assisted GPS method to show the performance improvements.

Key words: Acquisition, L2 Civil Signal, Weak Signal

1 Introduction

A signal processing of GPS receiver is composed of signal acquisition, signal tracking and navigation in accordance with function. Particularly, the performance of the signal acquisition has influence on TTFF (Time to First Fix) and RF sensitivity of GPS receiver. The RF sensitivity of GPS receiver is defined as the minimum power for acquiring the GPS signal. The GPS L1 C/A signals and L2 civil signals (L2CS) in Block IIR-M satellites are guaranteed minimum -128.5dBm and -131.4dBm signal strength each into a 3dBi linearly polarized user receiving antenna at worst normal orientation when the satellite is above a 5-degree elevation angle (ICD PIRN-200C-007B, 2002). It is difficult for GPS receiver to acquire GPS signals in the case of being obstacles in the line of sight since the GPS signal strength is very low (Haddrell and Pratt, 2001). From this viewpoint, it can be said that the RF sensitivity of GPS receivers is the dominant factor that has influence on the performance of GPS receivers. Since the L2 civil code provides better protection (24dB) than C/A against code cross correlation and continuous wave interference, it will be much more effective in weak signal environment.

This paper proposes an assisted GPS acquisition method using L2 civil signals. In section 2, this paper summarizes the structure and the property of L2 civil signal

comparing with those of L1 C/A code. In order to design and test the assisted GPS acquiring the L2 civil signal not existing yet, it is necessary to design L2CS generator. So section 3 describes a software-based L2CS generator designed in this paper. Section 4 proposes an acquisition method for solving the problem of squaring loss in weak signal environment since the long coherent integration increases the number of frequency search cells and the non-coherent integration of weak GPS signals induces the squaring loss. In section 5, through the simulations using the designed signal generator, it will show that the acquisition success rate of the proposed assisted GPS acquisition method is better than that of the existing assisted GPS method using L1 signals in the same environment.

2 L2 Civil Signal Structure

The new signal structure adds M (Military) codes and enhances L2 civilian codes (Hartman et al, 2000). L2 civilian codes are composed of the L2 civil moderate (CM) and L2 civil long (CL) codes as part of the L2 civilian enhancements. The spectrums of current and proposed GPS signals are shown in Figure 1. And Table 1 shows the characteristics of L2CS and existing L1 C/A code.

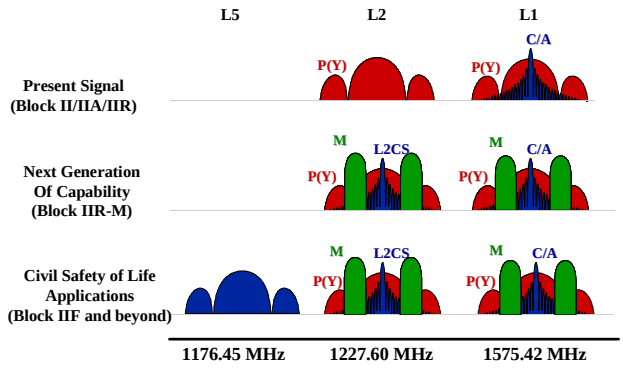


Fig. 1 Modernized GPS Signal Evolution

Tab. 1 Summary of Signal Characteristics

	L1 C/A	L2 CM	L2 CL	L2 CM/CL (TDM)
Code Type	Gold Code	Maximal Length Code	Maximal Length Code	Maximal Length Code
Chip Rate (Mchips/sec)	1.023	0.512	0.512	1.023
Code Length (Chips)	1,023	10,230	767,250	1,534,500
Repeat Rate (msec)	1	20	1500	1500
Carrier Frequency (MHz)	1575.42	1227.60	1227.60	1227.60
Bit Rate	50 bps	25 bps	No message	50 sps

Tab. 2 Cross Correlation Protection

Carrier Frequency (MHz)	Code Length (chips)	Code Clock (MHz)	Phases	Fully Available	Correlation Protection
1,575.42 (L1 C/A)	1,023	1.023	Bi-Phase	Now	> 21 dB
1,227.60 (L2CS)	10,230 (CM) 767,250 (CL)	1.023	Bi-Phase	~2013	> 45 dB

With the advent of the modernized GPS IIR-M satellites there will be an immediate benefit to all civilian GPS users including civil aviation. This is due to the characteristics of the L2C code on the L2 frequency. The L2C code signal is much more robust than the existing L1 C/A code and has much better cross correlation properties. The minimum L2C code cross correlation protection is 45 dB while 21 dB for the existing L1 C/A code as summarized in Table 2 (Fontana et al., 2001). This greater cross correlation protection is valuable in many environments where a weak GPS signal may be interfered with by another stronger GPS signal. It is beneficial to emergency indoor positioning or to personal navigation in wooded areas. Because of this great cross correlation protection, the L2C code signal also has a higher data recovery threshold and a better code tracking performance. The superior cross correlation properties also enable the GPS receivers to implement faster acquisition strategies because it can reduce the number of false alarms (Diggelen and Abraham, 2001).

The L2C code is composed of two multiplexed code. Each of two codes is a disjoint repeating segment of a maximal length code generated by a 27 bit shift register with 15 taps defined by a coder initial state which in turn determined by the satellite ID and code length. The diagram of L2C code generator is shown in Figure 2 (ICD PIRN-200C-007B, 2002). The CM code signal is a 10,230 chip sequence repeating every 20ms. The CL code signal is a 767,250 chip sequence repeating every 1.5 seconds.

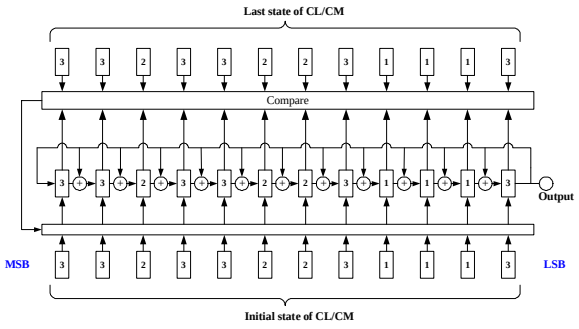


Fig. 2 L2C Code Generator

The L2C NAV (NAVigation) data can be either 50 bps data or 25 bps data, which is coded with a rate $1/2$ FEC (Forward Error Correction) convolutional coder. The coder state history is reset to zero at the beginning of each data message. The resulting 50 sps (symbol per second)

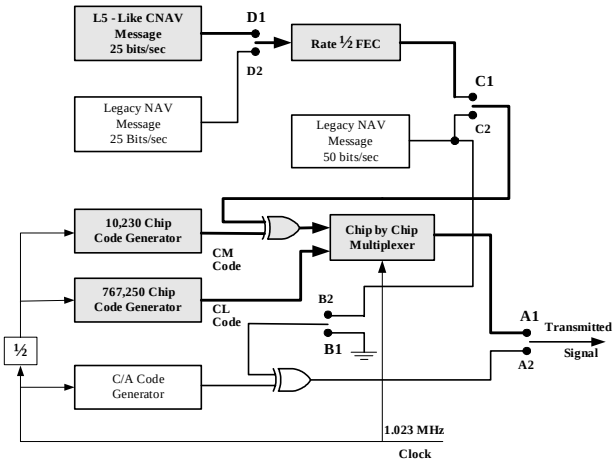


Fig. 3 L2C Signal Options in IIR-M Satellites

symbol stream is Modulo-2 added to the CM code. The resultant CM, CL bit-trains are combined using a time-division multiplex (TDM) method starting with the CM code. The combined bit-trains are used to modulate the L2 quadrature-phase carrier. The L2C NAV will have a flexible message structure controllable by the Control Segment. The structure of the navigation message for L2C, CNAV, is basically same as that of the L5 signal. It is more compact and more flexible than that of the current NAV message. Instead of a fixed message format, CNAV allows the Control Segment to specify the sequence and timing of each message component consisting of 300 bit subframe. Since the data rate of the L2C signal is 25bps, each subframe requires 12 seconds to be transmitted. The L2C signal options in IIR-M satellites are shown in Figure 3. The signal options are controlled by four switches whose preferred positions are A1, B1, C1, D1.

3 L2 Civil Signal Generator Design and Analysis

The constellation of the next generation GPS satellites is scheduled to launch in 2005. Therefore, in order to design and test the assisted GPS acquiring the L2 civil signal, it is necessary to design a signal generator which can generate the L2 civil signal. The signal generator is designed as shown in Figure 4.

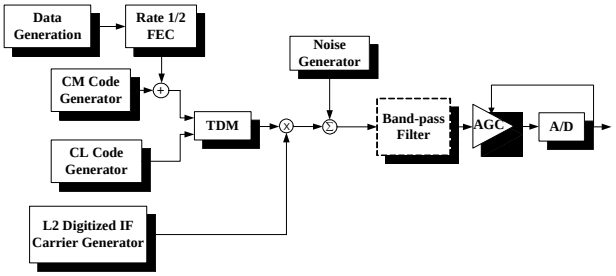


Fig. 4 Structure of GPS L2 Civil Signal Generator

Figure 4 shows the structure of the software-based L2C signal generator. In this signal generator, the noise generator has the zero-mean property as shown in Figure 5. And the output of the signal generator is shown in Figure 6.

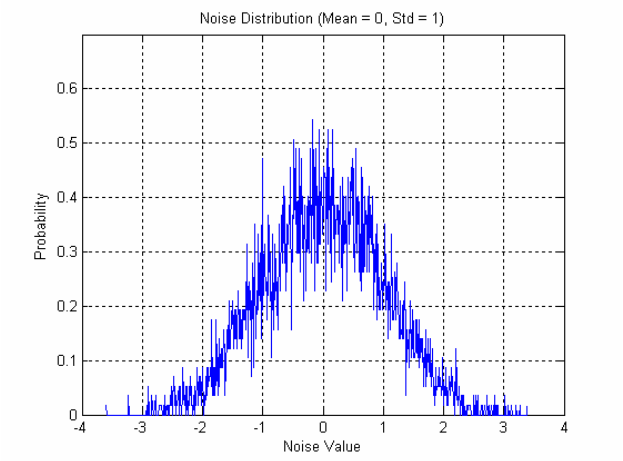


Fig. 5 The Output of Noise Generator

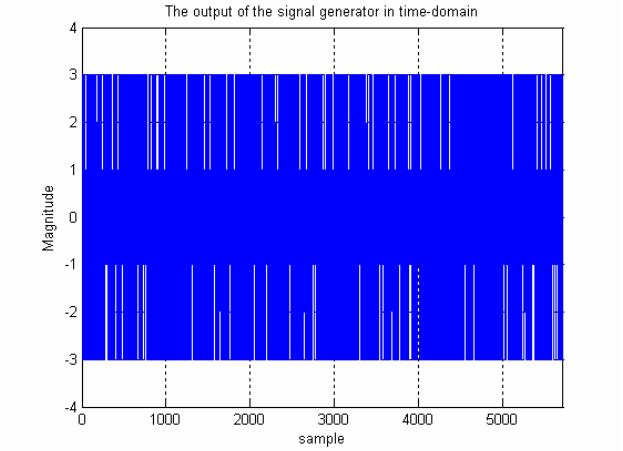


Fig. 6 The Output of the Signal Generator

4 A Proposed Assisted GPS Acquisition Method

4.1 The Squaring Loss

In general, in order to enhance RF sensitivity of GPS receiver, it is necessary to increase the correlation integration time over basic correlation time. Figure 7 shows a previously existing assisted GPS acquisition method using both the coherent integration and the non-coherent integration in weak signal environment. And Equation (1) and Equation (2) show the coherent integration and the non-coherent integration, respectively.

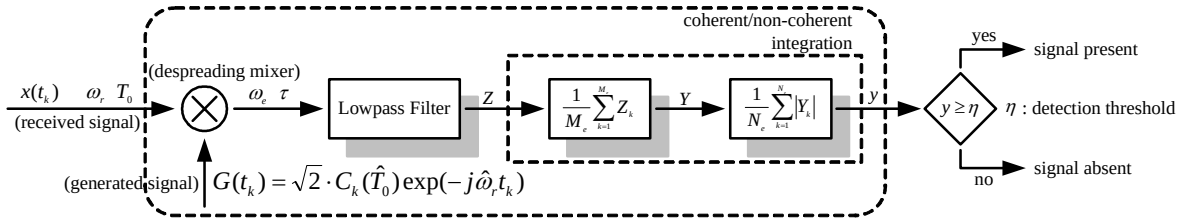


Fig. 7 The Previously Existing Signal Acquisition Method in Assisted GPS

After low pass filtering, the output of the coherent integration and the non-coherent integration are

$$Y = \frac{1}{M_e}\sum_{k=1}^{M_e} Z_k \tag{1}$$

$$y = \frac{1}{N_e}\sum_{k=1}^{N_e} \sqrt{\left(Y_k^i\right)^2 + \left(Y_k^q\right)^2} \tag{2}$$

where M_e and N_e are the number of the coherent integration and the non-coherent integration, respectively, Z_k is the output of the lowpass filter, Y and y are the output of the coherent integration and the non-coherent integration, respectively. Here Y is composed of the in-phase component, Y_k^i and the quadrature-component, Y_k^q.

The coherent integration is a technique integrating in-phase correlation result M_e times as given by Eq. (1). Therefore, it is assumed that there is no sign inversion of correlation by the navigation message bit transition or assistance of sign inversion information during the coherent integration. A relation of the number M_e of the coherent integration and an allowable carrier frequency error f_e is given by

$$L_{Acq}\big|_{\max} = 20\log_{10}\frac{\sin(\pi f_e T_i)}{\pi f_e T_i} \tag{3}$$

where $T_i = M_e \cdot T_p$ is the coherent integration time, T_p is the period of integration, and L_{Acq} is the acquisition loss. Eq. (3) shows that the longer integration time requires the less allowable carrier frequency error for the same signal acquisition loss.

Since there is no navigation bit stream during the initial acquisition time, the coherent integration technique is not proper to be adopted in the signal acquisition process of generic GPS receivers. For this reason, the generic GPS receiver performs the coherent integration with demodulating the navigation bit stream after the signal acquisition. Here the purpose of the coherent integration is to enhance the Signal-to-Noise (SNR) and to improve the quality of measurements.

The non-coherent integration is a technique integrating both the in-phase correlation result and the quadrature-phase correlation result as shown in Eq. (2). Therefore it is not necessary to know the navigation message bit transition. That is, the non-coherent integration is not influenced by sign inversion of the navigation message bit during the integration, and an allowable carrier frequency error is related not to the number of the non-coherent integration but to integration time of correlation values Y_k^i and Y_k^q. Therefore the non-coherent integration technique is adopted to enhance RF sensitivity of GPS receiver. But there is a disadvantage that the non-coherent integration induces the squaring loss for weak GPS signals. Particularly the squaring loss is the dominant factor among the acquisition losses of assisted GPS dealing with weak GPS signals.

The squaring loss is defined as the ratio of the SNR before the non-coherent integration for the SNR after the non-coherent integration.

$$L_{sq} = \alpha_c \Big/ \alpha_{nc} \tag{4}$$

$$\alpha_c = \sqrt{2 f_c T_p M_e}\,\frac{\sqrt{2P_s}}{\sigma_n}$$

$$\alpha_{nc} = 2\frac{\Gamma\left(1/2+1\right){}_1\mathrm{F}_1\left(-1/2;1;-\alpha_c^2/2\right)-\sqrt{\pi}}{\sqrt{4-\pi}}$$

where $\Gamma(\cdot)$ is the gamma function, ${}_1\mathrm{F}_1(\cdot)$ is the confluent hypergeometric function, α_c is the SNR before the non-coherent integration, and α_{nc} is the SNR after the non-coherent integration. Therefore, the squaring loss is given by

$$L_{sq} = \frac{\alpha_c\sqrt{4-\pi}}{2\left[\Gamma\left(\frac{1}{2}+1\right){}_1\mathrm{F}_1\left(-\frac{1}{2};1;-\frac{\alpha_c^2}{2}\right)-\sqrt{\pi}\right]} \tag{5}$$

And the squaring loss has properties as follows:

$$\frac{d}{d\alpha_c}L_{sq} < 0\ ,\quad L_{sq}\big|_{\alpha_c=\sqrt{10}} = 1 \tag{6}$$

From Eq. (6), it is explained that Eq. (5) is a monotonic decreasing function and the non-coherent integration induces the squaring loss when the Signal-to-Noise Ratio (SNR) is below $10^{1/2}$. As shown in Figure 8, if the SNR before the non-coherent integration is below $10^{1/2}$, the squaring loss exists. But if the SNR before the non-coherent integration is above $10^{1/2}$, the squaring loss does not exist.

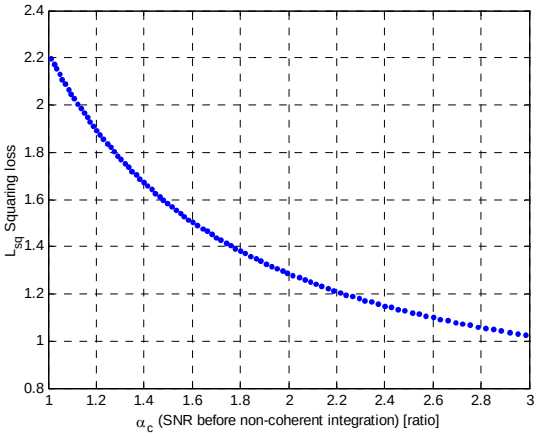


Fig. 8 The Squaring Loss vs. the SNR before Non-coherent Integration

4.2 The Proposed Assisted GPS Acquisition Method

The signal acquisition method proposed in this paper is shown in Figure 9. It integrates the inner products of adjacent two pair of in-phase components and quadrature-phase components sampled at different time instants as given by Eq. (7) (Lee et al., 2003).

$$Y_k = I_{k-1}I_k + Q_{k-1}Q_k \tag{7}$$

$$y = \begin{cases} \sum_{k=1}^{N_p} \sqrt{Y_k} & , Y_k \geq 0 \\ \sum_{k=1}^{N_p} -\sqrt{-Y_k} & , Y_k < 0 \end{cases}$$

where N_p is the number of the modified non-coherent integration.

When only the noise exists, the modified non-coherent integration method results in

$$y_n(t_k) = n_i(t_{k-1})n_i(t_k) + n_q(t_{k-1})n_q(t_k) \tag{8}$$

$$E(y_n(t_k)) = 0 \tag{9}$$

$$\mathrm{var}(y_n(t_k)) = \sigma^2 \tag{10}$$

The squaring loss does not occur because the inner product of adjacent samples has the zero-mean property as given by Eq. (9).

5 Performance Evaluation Test

To compare the signal acquisition performance of the proposed assisted GPS acquisition method using L2CS with that of the existing assisted GPS method using L1 signals in the same environment, this paper performed the signal acquisition test using signal generator designed in section 3.

First of all, this paper evaluates the GPS L1 signal acquisition performance using the previously existing acquisition method of Figure 7 and then evaluates the L1 and L2 civil signal acquisition performance using the proposed acquisition method.

It is assumed that the existing acquisition method of Figure 7 uses the navigation message bit information, and performs the 50 times non-coherent integration after

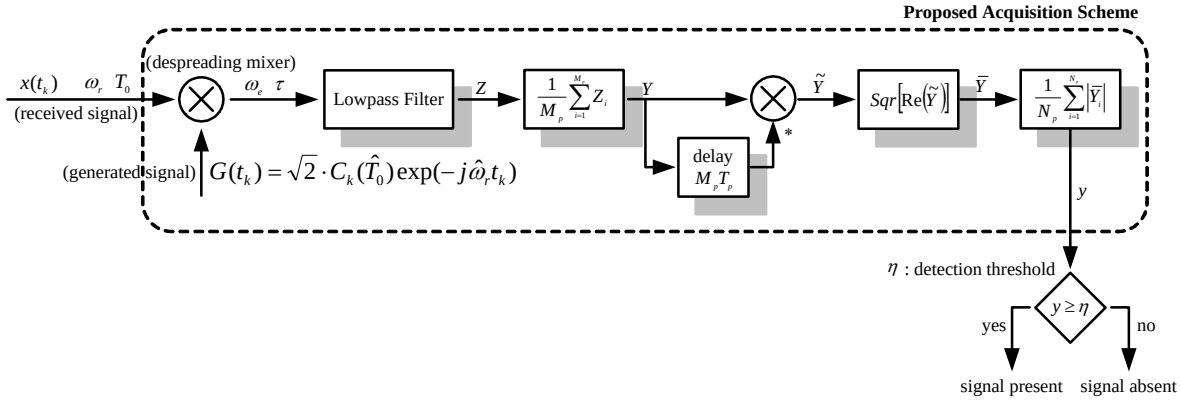


Fig. 9 The Proposed Signal Acquisition Method in Assisted GPS

coherent integration for 20 msec. The proposed method does not use the navigation message bit information, and performs the 200 times modified non-coherent integration after the coherent integration for 5 msec.

Figure 10 and Figure 11 show the noise distribution of the previously existing acquisition method and the proposed acquisition method respectively. As noted before, the proposed method has a nearly zero mean, but the previously existing method has a non-zero mean. It can be shown that the proposed method solves the squaring loss problem as expected.

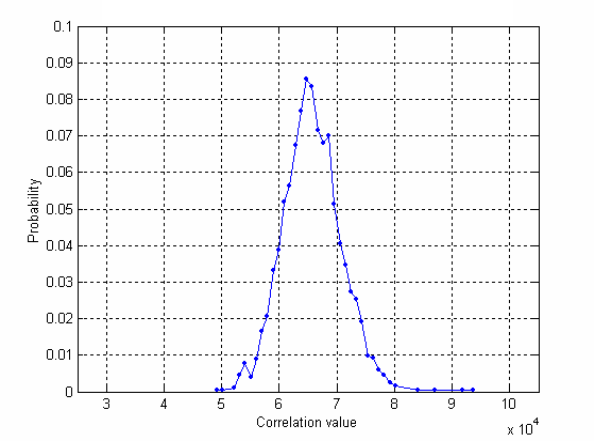


Fig. 10 Noise Distribution of the Existing Method

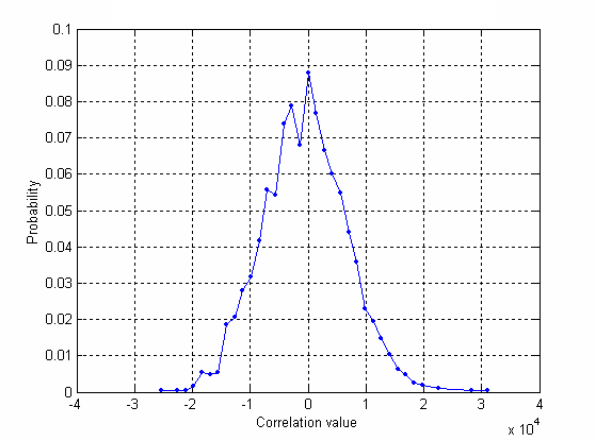


Fig. 11 Noise Distribution of the Proposed Method

Figure 12 shows the weak signal acquisition performance with respect to the input signal levels. Here, the input signal level is verified by comparing input level of the real GPS L1 signal with output level of the software based L1 signal generator through the post processing. And the software based L2CS generator designed in this paper adopts this experimental result.

The results of Figure 12 exclude the false alarm.

From Figure 12, it can be seen that the acquisition success rate of the proposed assisted GPS acquisition method is better than that of the existing assisted GPS method using L1 signals in the same environment. And it is evaluated that the assisted GPS using L2 civil signals is much more available than the assisted GPS using L1 civil signals in weak signal environment.

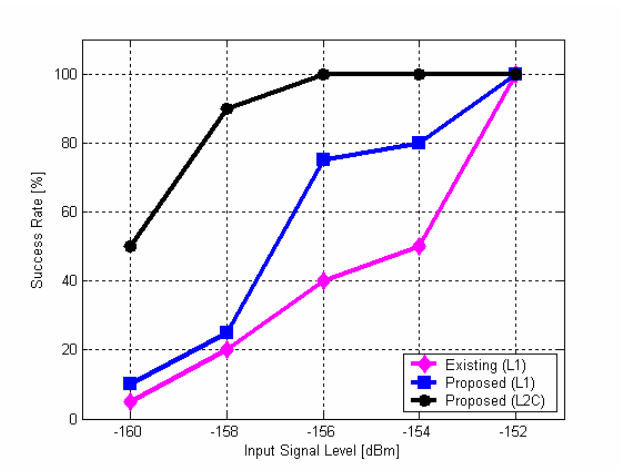


Fig. 12 The Signal Acquisition Success Rate

6. Conclusion

In this paper, the squaring loss of the previously existing assisted GPS is derived and the acquisition method for assisted GPS is proposed for solving the squaring loss problem in weak signal environment. It is shown that the proposed method solves the squaring loss problem by making the mean of the noise distribution to the integration output zero. Finally, the performance of the proposed acquisition method is verified by signal acquisition test using software based GPS signal generator designed in this paper. It is concluded that the proposed method for weak signal acquisition in assisted GPS shows much enhanced performance in not only L1 civil signal but also L2 civil signal.

References

C. Kee; H. Jun; D. Yun (2000): ***Development of Indoor Navigation System using Asynchronous Pseudolites***, Proceedings of the ION GPS-2000, 1038-1045.

F. van Diggelen; C. Abraham (2001): ***Indoor GPS Technology***, Presented at CTIA Wireless-Agenda, Dallas.

M.M. Chansarkar; L. Garin (2000): ***Acquisition of GPS signals at very low signal to noise ratio***, Proceedings of the ION 2000 National Technical Meeting, 731-737.

Richard D. Fontana; Wai Cheung; Tom Stansell (2001): ***The Modernized L2 Civil Signal***, GPS World Magazine, 28-34.

Richard D. Fontana; Wai Cheung; Paul M. Novak; Tom Stansell (2001): ***The New L2 Civil Signal***, Proceedings of the ION 2001 International Technical Meeting, 617-631.

Sang Jeong Lee; Deuk Jae Cho; Sang Hyun Park (2003): ***A Digital Correlator Design for Assisted GPS***, Proceedings of the International Symposium on Information Science and Electrical Engineering 2003 (ISEE 2003), ACROS Fukuoka, Japan, 275-278.

S.G. Glisic, et al. (1999): ***New PN code acquisition scheme for CDMA network with low signal-to-noise ratios***, IEEE Transactions on communications, Vol. 47, No. 2.

S. Soliman; S. Glazko; P. Agashe (1999): ***GPS receiver sensitivity enhancement in wireless applications***, Proceedings of 1999 IEEE MTT-S Symposium, 181-186.

T. Haddrell; A. R. Pratt (2001): ***Understanding the indoor GPS signal***, Proceedings of the ION 2001 International Technical Meeting, 1487-1499.

T. Hartman; L. Boyd; D. Koster; J. Rajan; J. Harvey (2000): ***Modernizing the GPS Block IIR Spacecraft***, Proceedings of the ION GPS-2000, 2115-2121.

W. Marquis (2001): ***M is for Modernization: Block IIR-M Satellites Improve on a classic***, GPS World Magazine, 38-44.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 32-39

Journal of Global Positioning Systems
2004

Using RFID for Accurate Positioning

Hae Don Chon, Sibum Jun, Heejae Jung, Sang Won An

Samsung Electronics Co., LTD
Email: {hd.chon, sibum.jun, cecil.chung, sangwon03.an}@samsung.com

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. In Korea, telematics is regarded as the technology to enhance and make everyday-driving experience more comfortable and safer. An essential part of the telematics is navigation and it is mainly based on GPS as the choice of positioning technology. The accuracy of GPS, however, is approximately ten to twenty meters. Combining with map-matching technologies, most navigation systems guide drivers with a best effort manner. In addition to telematics, RFID (Radio Frequency Identification) is an old but newly emerged technology. In this paper, we incorporate RFID technology into a navigation system to improve the accuracy. The skeleton of the idea is as follows: install RFID tags on roads in a certain way, store very accurate location information along with other necessary information in the tags, add an RFID reader module to the navigation system, and use this new location information along with GPS and a gyroscope to produce highly accurate location information. With this scheme, the accuracy of positioning can be dramatically improved, especially in tunnels and in downtown areas. Preliminary results show that this idea is feasible.

Key words: Positioning, RFID, Navigation, Telematics.

1. INTRODUCTION

Telematics can be defined as the technology that enables a vehicle to exchange data via wireless communication in the form of services or applications. Core components of a telematics system include positioning technology, Human-Machine Interfaces (HMI), and a navigation module. Over the last few years, telematics has drawn much attention in major industrial countries. Korea is one of the nations that endeavor to capitalize on telematics. In fact, the Ministry of Information and Communication of Korea (MIC) has recently released its master plan called the *IT839 Strategy* to help the IT industry develop new technologies (MIC, 2004). The IT839 Strategy is a plan for activating eight services (WiBro, DMB, Home Network, Telematics, RFID, W-CDMA, Terrestrial Digital TV, VoIP), implementing three infrastructures (Broadband Convergence Network, Ubiquitous Sensor Network, IPv6), and developing nine new growth engines (Next-Generation Mobile Communications, Digital TV, Home Network, IT SoC, Next-Generation PC, Embedded SW, Digital Contents, Telematics, Intelligent Service Robot). Telematics is one of the eight new services and also one of the nine new growth engines in the IT839 Strategy.

As mentioned above, positioning technology is a core component of telematics and most of telematics devices sold today rely on GPS as a choice of the positioning technology. With GPS, it is inherently difficult, if not impossible, to get a position in a tunnel or in downtown areas surrounded by skyscrapers. This issue in GPS would not be resolved for many years to come. To enhance the accuracy to a level at which lane-by-lane route guidance can be possible, using RFID technology for positioning is proposed.

RFID is not a new technology, rather its origin dates back to World War Two. Recently, RFID comes under the spotlight for its potential for changing the world by replacing barcodes and providing many services that are unknown today. In addition to the replacement of barcodes, RFID has gaining interests in mobile phone and consumer electronics industry (www.nfc-forum.org).

The RFID positioning can be divided into four steps: in the first step, install RFID tags on roads in a certain way, store very accurate location information along with other necessary information to the tags, add an RFID reader module to the navigation system, and use this new location information. Apart from the RFID system, we also propose to use a tag database.

Due to the memory constraint on the tag and the data size that needs be written in a tag, the use of a database for

tags is a necessary condition. In addition, the speed of the RFID communication also makes the use of the tag database indispensable.

Preliminary results show that the RFID communication speed does not solely depend on the bit data rate in the specification. Nonetheless, it will be shown that retrieving ID from a tag can be done fast even at a high velocity (by the high velocity we mean 150km/h). Another performance study on the tag database access time indicates that the access time is marginal.

In the RFID and telematics literature, there is not much research on RFID positioning pertinent to vehicle navigation. Kubitz *et al*. developed a technique for robot navigation using RFID (Kubitz *et al*., 1997). They, however, did not take the velocity of robots into account, whereas in vehicle navigation fast RFID communication would be crucial. Penttilä *et al*. presented a performance study on an RFID system, in which they achieved reliable identification accuracy at up to 40 km/h (Penttilä *et al.* 2004).

This paper is organized as follows: in Section 2 a brief overview of the RFID technology is given as background information. In Section 3, the idea of RFID positioning is described along with the tag database. The results of performance study are provided in Section 4. Conclusions and future work are described in Section 5.

2. RFID BACKGROUND

In this section a brief overview of RFID technology in general is given. An RFID system consists of tags, a reader with an antenna, and software such as a driver and middleware. The main function of the RFID system is to retrieve information (ID) from a tag (also known as a transponder). Tags are usually affixed to objects such as goods or animals so that it becomes possible to locate where the goods and animals are without line-of-sight. A tag can include additional information other than the ID, which opens up opportunities to new application areas.

An RFID reader together with an antenna reads (or interrogates) the tags. An antenna is sometimes treated as a separate part of an RFID system. It is, however, more appropriate to consider it as an integral feature in both readers and tags since it is essential for communication between them. There are two methods to communicate between readers and tags; *inductive coupling* and *electromagnetic waves*. In the former case, the antenna coil of the reader induces a magnetic field in the antenna coil of the tag. The tag then uses the induced field energy to communicate data back to the reader. Due to this reason inductive coupling only applies in a few tens of centimeter communication. In the latter case, the reader radiates the energy in the form of electromagnetic waves. Some portion of the energy is absorbed by the tag to turn on the tag's circuit. After the tag wake up, some of the energy is reflected back to the reader. The reflected energy can be modulated to transfer the data contained in the tag.

Three frequency ranges are generally used for RFID systems: low (100~500 kHz), intermediate (10~15 MHz), and high (850~950 MHz, 2.4~5.8 GHz). Detailed characteristics of these three frequency ranges along with examples of major applications can be found in *RFID Handbook* (Finkenzeller, 2002). The communication range in an RFID system is mainly determined by the output power of the reader to communicate with the tags. The field from an antenna extends into the space and its strength diminishes with respect to the distance to tags. The antenna design determines the shape of the field so that the range is also influenced by the beam pattern between the tag and antenna. Although it is possible to choose power levels for different applications, it is usually not allowed to have complete freedom of choice due to legislative constraints on power levels as in the case of the restrictions on carrier frequencies. More interested readers are referred to *RFID Handbook* (Finkenzeller, 2002) and an M.S. thesis from MIT (Scharfeld, 2001).

3. RFID POSITIONING

As mentioned in Section 1, we propose to use RFID technology for positioning. This technique, however, would not replace GPS rather it is a complementary technique. In this section, we give an overview of the RFID positioning and describe the feasibility of the idea.

3.1 Overview

The RFID positioning can be explained as follows: First, RFID tags need to be installed on a road in a manner which could maximize the coverage and the accuracy of positioning. This deployment scheme is discussed later in this subsection. Upon installation, necessary information such as coordinates of the location where the tag is installed needs to be written on each tag. The accuracy of this position information is very critical for this technique to be successful. The position information can be acquired by using DGPS or some other methods, which would take much longer time to compute the location. Contrary to GPS in navigation systems where real time positioning is necessary, the time for getting the accurate information would be tolerated since this computation would take place once.

Vehicles, then, need to be equipped with an RFID reader that can communicate with the tags on a road. No matter

how accurate the RFID positioning is, it only gives the position where the tags are. Therefore the vehicles need also to be equipped with a GPS receiver and inertial sensors such as a gyroscope for positioning when there are no tags around. While driving, the vehicles constantly monitor the presence of a tag. On detection, the reader retrieves the information from the tag including coordinates of the location, which are supposedly very accurate information. Figure 1 shows a scheme of deploying RFID tags on a road. In the figure, circles (blue dots) on each lane represent a tag and the tag itself is enclosed in a special purpose cat's eye. A cat's eye is a device built into a sturdy housing and placed on a road as a lane marker.

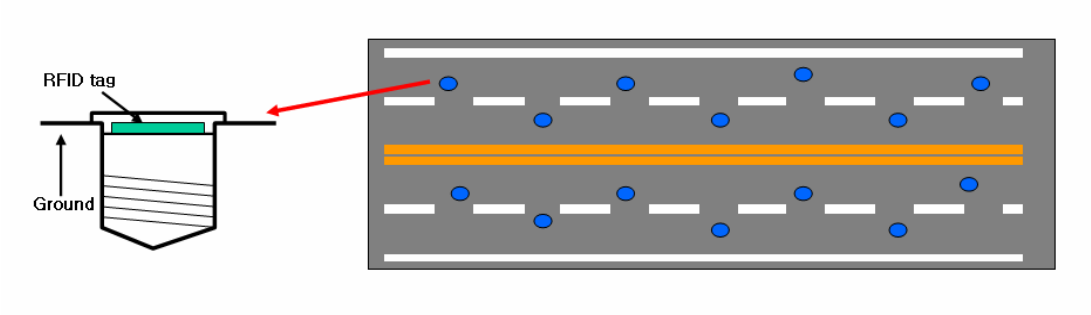


Figure 1. RFID Tags on a Road

The deployment should be done step by step: places such as tunnels from which getting GPS signal is not an option should be the first, intersections the next, urban areas, and then nationwide. Due to this nationwide scale, governmental actions are necessary. As part of the IT839 Strategy, the government of Korea funds a research project called *Development of Telematics Test Bed* being done by Electronics and Telecommunications Research Institute (Lee, 2004). Among other things, the research includes the development of RFID positioning.

3.2 Feasibility

In this subsection, the issues of feasibility of the RFID positioning are discussed. There has been work on using RFID for autonomous robots (Kubitz *et al.*, 1997). Contrary to the robot case, an RFID tag in our application will be installed on a road, where the operation environment for the tag is very harsh; high temperature in summer, low in winter, dusts, rain, snow, etc. Furthermore, vehicles equipped with an RFID reader that is compatible to the tags on the roads can move very fast; though speeding is illegal, some cars (Porsche) can go as fast as 300 km/h. To be more useful the tags should contain the information about the road property (number of lanes, which lane it is on, how far to the nearest intersection, etc) other than the location information. More data decreases the communication speed and requires more memory, which leads to high cost.

In summary, there are issues to be addressed before full-fledged deployment of RFID tags nationwide. The first issue is making RFID tags that can withstand a harsh environment. The second one is fast communication speed between readers and tags. The third one is the data size. Besides these three issues, there are other issues such as standards (frequency, data format, air interface, etc), which are not covered in this paper. Although it is very important, the environmental problem is beyond the scope of this paper. We only address the communication speed and the data size issues in this and the next section.

3.3 Tag Database

While there would be location information in a tag, it would be almost impossible to embed all the necessary information in a tag due to memory constraints and the dynamic nature of some information. Information such as absolute coordinates of the location will not be changed. On the contrary, relative coordinates and the property of the road on which the tag is could change some time (unlikely, though). Moreover, we can embed more useful information such as nearest museums, restraints, and gas stations. However, the contents of the information vary all the time. The data size as well as the dynamic nature of it prevents from writing all the information at the installation time. To address this issue, we devise a tag database which stores information corresponding to the tags available on the roads in a region (country for instance). The information stored in the tag database is whatever information on real tags and more such as point of interests.

Another reason for the necessity of the tag database comes from the speed of the RFID communication. Although we do everything we could to speed up the communication, it may not be fast enough to get all the information from a tag while driving at, for instance, 150km/h. As we show in Section 4, however, getting only identification (ID) is very feasible even at such a high velocity. Once the ID is retrieved, it can be efficiently

searched the tag database and extracted whatever information necessary.

The tag database is a collection of tags and a part of the digital map that a navigation system may carry. Generally, a digital map consists of cells each of which contains network information for route guidance. The network information is a graph with nodes and links.

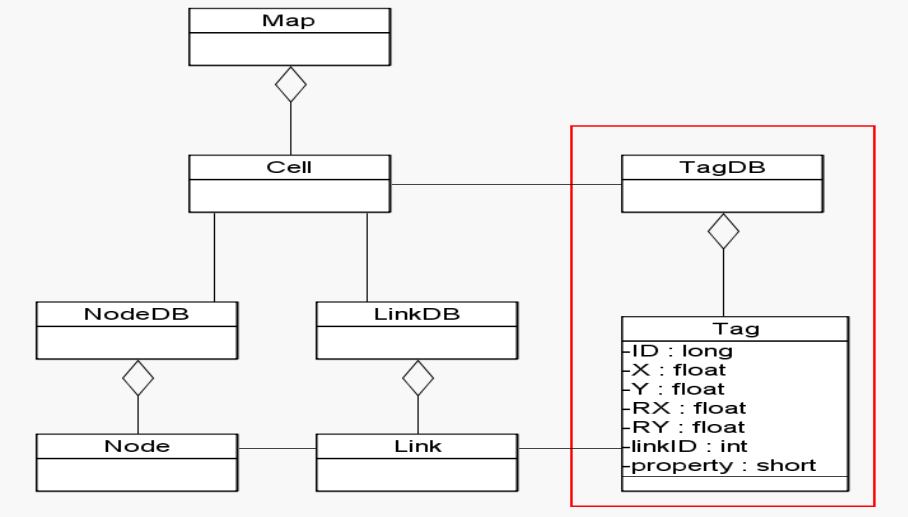


Figure 2. **Tag Database Class Diagram**

tags. For simplicity, we only show the attributes of the Tag object. As in the diagram, a Tag object includes ID, absolute coordinates X and Y, relative coordinates RX and RY, link ID where the tag is, and the property field. This last field is for the number of lanes of the link, type of the road (highway, local, etc), and so on. In Java language and most of other programming languages, type *long* is 8 bytes, type *float* and *int* are 4 bytes, and type *short* is 2 bytes. Summing up, the data size of a tag is 30 bytes. Therefore even with a million tags on the roads, thereby in the database, the size of the tag database is approximately 30MB. Since more and more embedded devices have large flash memories and even gigabytes of hard drives, the sheer size of the tag database would not be a big issue.

4. PERFORMANCE

We are in a very early stage of a research project and constructing right test equipments and environment. Nevertheless, preliminary results of experiments on communication speed and tag database access time are provided in this section. Authors of this paper did their best to produce accurate results.

Figure 2 shows a class diagram of the digital map. In the diagram, TagDB is an aggregation of Tag objects which represent tags in a real world.

Each cell has links to the collections of nodes, links, and

4.1 Communication Speed

Since there were no RFID tags and readers for the purpose of RFID positioning on a road, we have used a commercial one, KIS900RE from Kiscomm Co. (www.kiscomm.co.kr). Table 1 shows the specification of KIS900RE.

According to the specification, the bit data rate is 64 Kbps which would be fast enough to communicate several hundred bits at a high velocity. Figure 3 shows the distance needed to read 64 bits at various data rates with respect to vehicle velocities. In the figure, the data rate is assumed to be the sole factor of RFID communication. Only the distance for 64 bits is drawn since the data size of a tag ID is 8 bytes in our program. The purpose of the experiments explained in this section is to measure the time that actually takes to complete a reading transaction.

For the experiment, the radiation pattern of the antenna of KIS900RE is measured in a Satimo chamber (See Figure 4). Based on the radiation pattern, it can be assumed that the reliable reading angle from the center of the antenna is roughly 68° (assuming 3dB beam width).

The setting of the experiment is as follows: first a tag is dropped (freefall) at the height of 13 meters. Since the Kiscomm tag is very light and soft, it is attached a piece of paper that is hard enough. To make air resistance

marginal heavy coins are attached to the tag. The reader antenna is then placed at the height of one meter so that the distance between the center point of the antenna and the falling point is 12 meters. By the law of energy conservation, the velocity of the tag at the center of the antenna is 15.33 m/s or 55.21 km/h. See Figure 5 for the experiment setting.

$$9.8mh = \frac{1}{2}mv^2 \quad \text{Law of Energy Conservation}$$

Reader	Frequency	908.5 – 914 MHz (14 channels, 250 kHz spacing Frequency Hope)
	Radio Power Output	1W Conductive (Korean Band)
	Read Range	6 meters (Typically read range depends on Reader environment and the used tag)
	Operating Temperature Range	-10 deg to +50 C
	Humidity	5 to 95 % non-condensing
Tag	Tag ID	128 bits
	Bit Data Rate	64 Kbit/s
Antenna	Gain	6 dBi
	In Band VSWR	< 1, 2

Table 1. KIS900RE Hardware specification

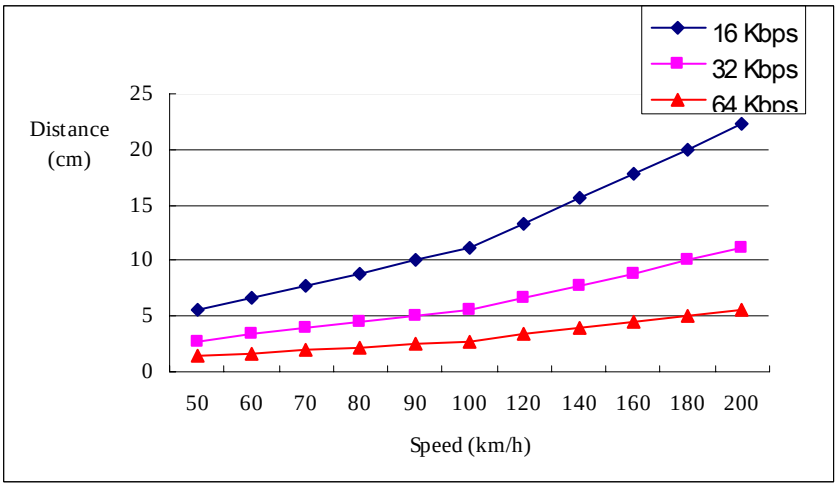
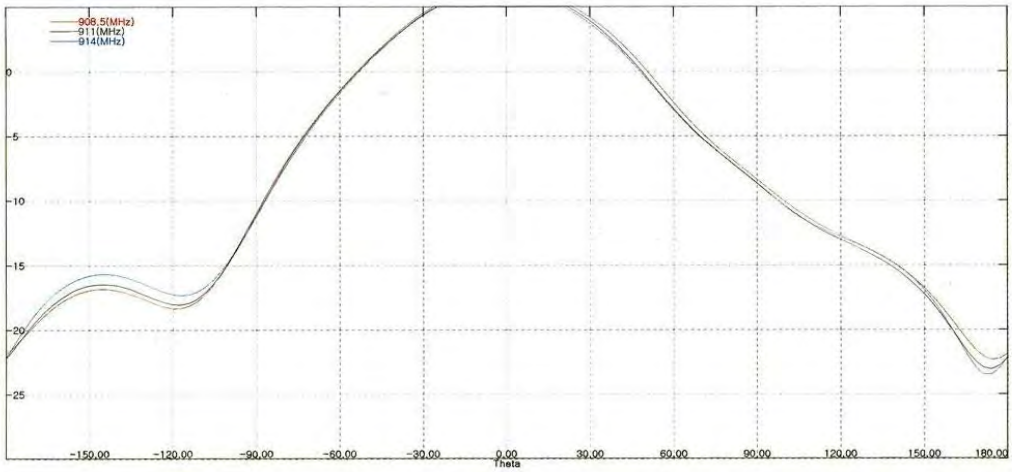


Figure 3. Distance for reading 64 bit

When passing the read range, the distance between the tag and the antenna is about 60cm. Considering the 68° of reliable reading angle, the distance from the entrance of the read range to the exit of the read range is 81cm ($2R = 2 \times d \times \tan 34°$). In the experiments, the reader (KIS900RE) was able to read a tag either two or three times in a freefall. Suppose that the tag passed the center line of the circle as in Figure 5 when it was read three times**.** Then the distance needed to complete one reading transaction can be computed as 27cm. At this time, the tag was moving at a velocity of 55km/h and the time actually taken for one reading transaction is 18 milliseconds. In the unit of bps, the reading speed is about 7.11Kbps since the data size of the tag is 128 bits. Note that this is very different from 64Kbps of the specification. Generally, RFID system has three phases of operation. The first is *charge-up phase* where the reader transmits energy and the tag absorbs and stores it until the internal supply voltage is high enough to turn on the circuitry. The second is *communication phase* which can be either inductive or back scatter coupling. The last is *discharge phase* where the whole circuit of the tag is reset and the supply capacitor is discharged. Therefore a possible explanation for the big difference is that 64 Kbps is measured only for communication phase and the charge-up and discharge phases take much more time than the communication phase.



Layer	Max value	Position	Min value	Position	BeamWidth	Average
908.5(MHz)	6.43	- 2.86 deg	- 22.36	174.28 deg	71.00 deg	- 0.22
911(MHz)	6.31	- 2.86 deg	- 23.08	174.26 deg	69.75 deg	- 0.41
914(MHz)	6.24	- 2.86 deg	- 23.50	171.45 deg	69.13 deg	- 0.50

Figure 4. Radiation pattern of the antenna of KIS900RE

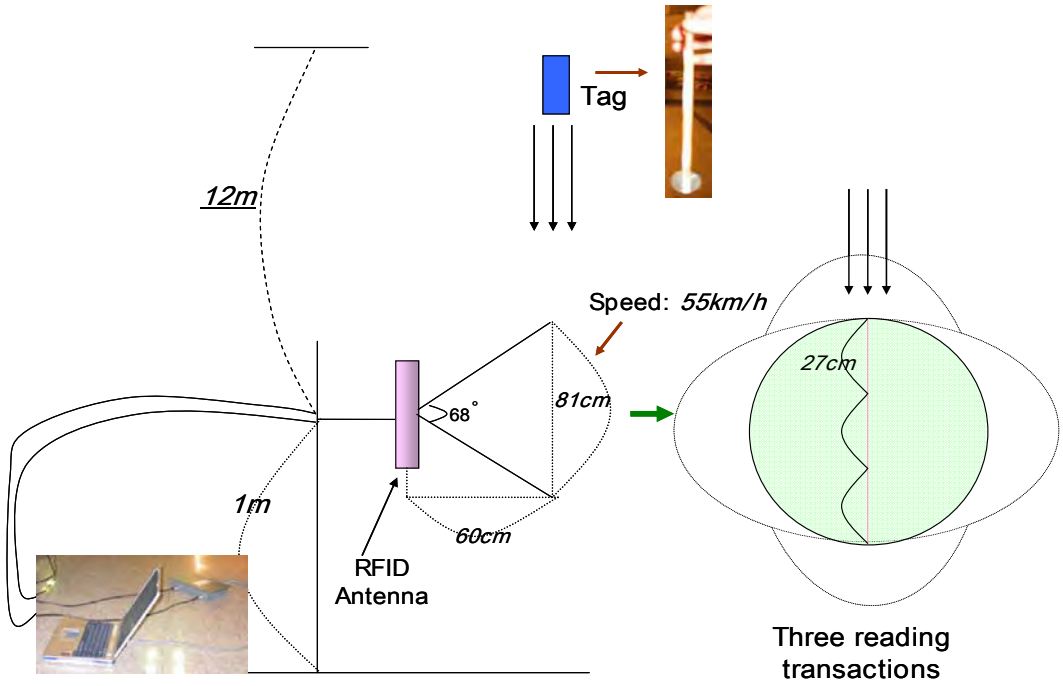


Figure 5. **Experimentation Setting**

The result of the experiments shows that the bit data rate is not the only contributor to the RFID communication speed. In fact, if the data rate was the sole factor in our experiment, then it would have been just 3cm for a reading transaction. As it turned out, however, it took nine times longer than that. This means that if we were to improve the communication speed, we need to focus on somewhere else such as capacitor charging time rather than the data rate. Nonetheless, it can be inferred from the experiment that it could need 81cm to read 128 bits at the velocity of 165km/h. Notice that the reader we tested and most of others have read ranges longer than 81cm. Therefore it is feasible to read an RFID tag at a high velocity up to 165km/h. However, this does not automatically mean that RFID readers can read tags on a road at a high speed. There are issues to address such as making tags withstand harsh condition. As we mentioned

earlier, this is a preliminary result and we intend to construct test equipments and environment for more accurate experiments.

4.2 Tag Database Access Time

In this section, we provide a result of the experiment on the tag database access time. The necessity of the tag database is explained in Section 3. What our main concern here is how fast the information corresponding to a tag in the tag database can be accessed. Therefore we concentrate on the tag database out of three databases in the map shown in Figure 2.

Given its ID, the time to extract position information of a tag was measured using a Java program. Given the number of tags (*dbsize*) in a cell, IDs were randomly generated and stored in the tag database. For the performance and memory constraint reasons that are normal on embedded devices such as telematics terminals,

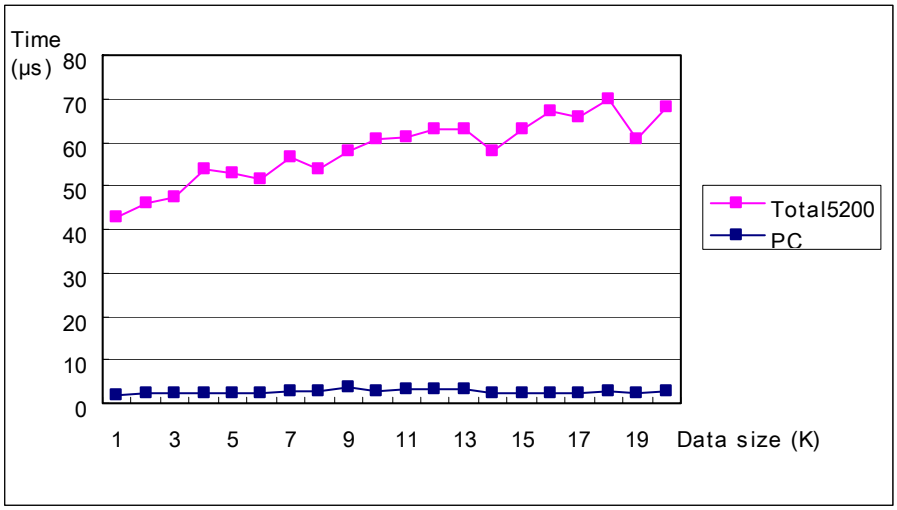


Figure 6. Average Access Time for a Tag

a *sorted array* was chosen as the data structure of the tag database. The key of the sorting was the ID of tags. After construction of the tag database, some number (*idsize* = *10000*) of IDs were also randomly generated and stored in a separate array for the fast execution. The time to extract *idsize* tags was then measured and the average was taken. When searching the database, the binary search algorithm was used. Given *dbsize*, this execution was repeated twenty times and *dbsize* varies from 1000 to 20000.

The program was run on a Total5200 development platform. Total5200 is a test platform for telematics from Freescale Semiconductor, Inc (www.freescale.com). It consists of 400MHz MPC5200, 64 MB SDRAM, 64MB Flash memory, J9 (IBM's Java Virtual Machine), and more (the detailed specification can be found on their website). For the comparison purpose, the program was also run on a PC with Intel Pentium4 2.8 GHz and 512MB. The results are shown in Figure 6.

Notice that the running time increases as the data size increases, which was expected. Notice also that the running time on a PC is less than 3.5 microseconds and the running time on Total5200 is less than 70 microseconds regardless of the data size. The 70 microseconds is a negligible time given that 18 milliseconds are for a reading transaction in the experiment. Therefore we can conclude that once a tag ID is acquired, extracting necessary information from a database would not take any meaningful time.

5. CONCLUSIONS and FUTURE WORK

In this paper, the idea of using RFID technology for positioning was proposed. Preliminary results on RFID communication speed and tag database access time were shown. For the communication speed, the experiment showed that 18 milliseconds were needed to complete a reading transaction. With the result, it was computed that the communication speed is 7.11Kbps as opposed to 64Kbps in the specification. It can be concluded that it is feasible to retrieve bits for ID of a tag at high velocities provided that operational tags in harsh conditions are installed. We also argued that a database for information corresponding to tags on roads is necessary. The performance study on the tag database showed that the access time is insignificant compared to a reading transaction. Therefore once an ID is retrieved by the RFID reader, getting the necessary information from the database does not pose any problem.

Although it is our conclusion that the idea of RFID positioning is feasible, there are many issues to resolve before any full-fledged deployment. Some of the issues are (1) making a tag to withstand harsh conditions, (2) testing communication speed at up to 150 km/h, (3) developing a scheme to combine an RFID reader, a GPS receiver, and a gyroscope to produce a consistent and accurate position information, and (4) developing a tag database management algorithm and program for future use. As future work, these issues will be investigated.

ACKNOWLEDGEMENTS:

We give our thanks to Yunkyoung Ko for his cooperation in measuring the radiation pattern of the RFID antenna.

REFERENCES

Finkenzeller K, ***RFID Handbook***: Fundamentals and Applcations in Contactless Smart cards and Identification (2002)

Kiscomm co., LTD, KIS900RE View, www.kiscomm.co.kr

Kubitz O, Berger M, Perlick M and Dumoulin R (1997) ***Application of Radio Frequency Identification Devices to Support Navigation of Autonomous Mobile Robots***, IEEE 47th Vehicular Technology Conference, 126-130

Lee J (2004) ***Development of Telematics Test Bed***, TTA Telematics Standards & Technology Seminar (www.tta.or.kr, in Korean)

Ministry of Information and Communication (2004) ***IT 839 Strategy Master Plan***, www.mic.go.kr

Penttilä K, Sydänheimo L, and Kivikoski M (2004) ***Performance development of a high-speed automatic object identification using passive RFID technology***, Proceedings of the 2004 IEEE International Conference on Robotics & Automation, 4864-4868

Scharfeld T (2001) ***An Analysis of the Fundamental Constraints on Low Cost Passive Radio Frequency Identification System Design***, MIT M.S. Thesis

Siegemund F and Flörkemeier C (2003) ***Interaction in Pervasive Computing Settings using Bluetooth-Enabled Active Tags and Passive RFID Technology together with Mobile Phones***, Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 378-387

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 40-44

Journal of Global Positioning Systems
2004
© CPGPS

JAMFEST - A Cost Effective Solution to GPS Vulnerability Testing

Lt Col Eric Lagier

46th Test Group, 871 DeZonia Ave, Holloman AFB, NM 88330
email: Eric.Lagier@46tg.af.mil; Tel: (505) 572-1367; Fax: (505) 572-1575

Capt Desiree Craig

46th Test Group, 871 DeZonia Ave, Holloman AFB, NM 88330
email: Desiree.Craig@46tg.af.mil; Tel: (505) 572-1243; Fax: (505) 572-1575

Paul Benshoof

746th Test Squadron, 1644 Vandergrift Rd, Holloman AFB, NM 88330
email: Paul.Benshoof@46tg.af.mil; Tel: (505) 679-1769; Fax: (505) 679-1759

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. From May 24-28, 2004, the 746th Test Squadron, located at Holloman Air Force Base (AFB), New Mexico (NM), planned and executed an innovative Global Positioning System (GPS) jamming program at White Sands Missile Range, NM. This program, known as JAMFEST, was aimed at providing low to no cost, realistic, GPS jamming scenarios for testing GPS-based navigation systems, as well as, training personnel in unique GPS denied environments. Through sponsorship from the GPS Joint Program Office, White Sands Missile Range, and the 46th Test Group, the 746th Test Squadron was able to provide this opportunity at a significantly reduced cost to each participant. During JAMFEST, the 746th Test Squadron hosted twelve simultaneous, yet very diverse customers, including multi-service Department of Defense (DoD) organizations, several defense contractors, and civil organizations. Their objectives ranged from training personnel on the effects of GPS jamming to characterizing the performance of prototype advanced anti-jam technologies against operationally realistic threats. To accomplish these goals, participants drove, flew, or walked through 59 jamming scenarios specifically tailored to stress the systems under evaluation. These tests would have cost a total of $660,000 or more if conducted separately. However, JAMFEST achieved the same objectives for approximately $85,000 in available funds coupled with discounted or donated services totaling $175,000. This paper details overall test and participant objectives, strategies, conduct, and addresses future JAMFEST activities.

Key words: JAMFEST, vulnerability, test, GPS

1 INTRODUCTION

The 746th Test Squadron (746 TS) has conducted complex GPS jamming experiments since the early 1990s and played a key role “behind the scenes” creating multiple high-profile jamming environments for programs such as the Joint GPS Combat Effectiveness (JGPSCE) exercises and Quick Reaction Tests. These programs, conducted to support real-world operations, enhanced the 746th TS’s ability to recreate realistic jamming environments and resulted in the 746 TS earning the reputation as the recognized experts for open-air GPS vulnerability testing.

The 746 TS conducted JAMFEST as an opportunity to broaden both the operational and test communities’ awareness of GPS vulnerabilities by offering a cost-effective, operationally realistic venue to facilitate testing and training objectives. This opportunity was truly important to the operational and test communities because GPS signals use very low power and are vulnerable to both intentional and unintentional interference. These effects can adversely impact the position and timing accuracy of various receivers and navigation devices employed by military and civilian users.

Because of this adverse impact, it was critical to the

success of the program that the jamming environment be both operationally realistic and beneficial to the military and civilian users. Specifically, it involved coordinating frequency clearances, securing range space, developing jamming scenarios, deploying personnel and equipment, operating the threat assets, data reduction and analysis, reporting, and securing funding for the program. By capitalizing upon their unparalleled experience base, the 746 TS easily met and overcame these tasks.

The 746 TS designed innovative vulnerability scenarios that streamlined test conduct into a one-week event that maximized set-up efficiency and significantly reduced costs to the participants. In doing so, the May 04 JAMFEST capitalized on the ability to share common jamming scenarios, which enabled users to participate in unique environments at a fraction of the normal cost and allowed organizations that normally would not consider participating in such an event the opportunity to take part. In fact, by streamlining the program in this highly efficient manner, the team reduced a $660,000 program into a $175,000 program. After excluding donated services and benefits from the White Sands Missile Range, and the GPS Joint Program Office, the total program cost was $85,000.

This was the first in a series of recurring events; the next event is scheduled in Nov 04.

2 OBJECTIVES & RESOURCES

The overall objective of JAMFEST was to provide and characterize the GPS jamming environment in multiple configurations to enable the participants to test, train, or gain experience in a GPS jammed scenario. Each participant used JAMFEST to execute their own objectives, which included the following:

Evaluate the effects of jamming on a representative set of GPS receivers to determine the effective range from the jammers and the power level that disrupts GPS tracking;

Evaluate potential benefits of anti-jam technology available to civil operators;

(1.) Collect performance data against specific targets/environments that will confirm proper operation of the overall locator system and sub-system;

(2.) Subject anti-jam systems under test to high GPS jamming/Signal (J/S) environments and compare results;

(3.) Collect jamming environment truth data to improve and verify laboratory modeling and simulation tools, vulnerability prediction analysis, and mission planning software;

(4.) Validate tactics, techniques and procedures (TTPs) using hand held receivers (HHRs).

To effectively execute these objectives, the 746 TS employed multiple test assets to configure an operationally representative GPS jamming environment. The ground jamming configuration was set up on White Sands Missile Range (WSMR).

One of the primary test resources used to create the jamming environment was the Portable Field Jamming System (PFJS). The PFJS (see Figure 1) is a modified Ford 350 van with a full suite of GPS Electronic Warfare (EW) equipment, which included TMC Advanced Threat Emulators (TATEs) and TAVIA-32 Emulators (TAVIAs) as well as a variety of high power adjustable amplifiers. The onboard EW equipment was programmed to provide a wide range of jamming scenarios and signal modulations. The system records time-tagged amplifier power output for test analysis and time correlation to the test item.

Figure 1 Portable Field Jamming System (PFJS)

Another key resource employed was the Tactical Field Jamming System (TFJS). The TFJS (see Figure 2) is designed to supply the same capabilities as the PFJS, but in a vehicle capable of accessing terrain that is more rugged. Each TFJS is a modified High Mobility Multi-Purpose Wheeled Vehicle (HMMWV) that comes equipped with a full suite of GPS EW equipment, which includes TATEs and TAVIAs, as well as a variety of high power adjustable amplifiers. Due to the TFJS's ability to be positioned in areas inaccessible to most vehicles, these jammers were set up in remote territory and controlled via radio modem.

Figure 2 Tactical Field Jamming System

Portable Box Jammers (PBJ) (see Figure 3) in conjunction with the PFJSs and TFJSs, were set up along designated range roads and remote locations to help create the jamming environment. Each PBJ is a stand-alone jamming system designed to supply the same capabilities as the PFJS and TFJS, but in a smaller, more versatile package. Each system was equipped with portable generators, a portable antenna mast and tower, and a full suite of GPS EW equipment that included TATEs and TAVIAs, as well as a variety of high power adjustable amplifiers.

Figure 3 Portable Box Jammer

The 586th Flight Test Squadron (586 FLTS), a sister squadron to the 746th TS, characterized the jamming field using the C-12J (see Figure 4). The C-12J is a Beechcraft 1900 twin turbo-prop aircraft that has been modified to for GPS/inertial guidance and navigation components and systems tests. Its capacities include a 16,600-pound maximum gross weight and a maximum of four test stations or equipment pallets. The aircraft was configured with controlled reception pattern antenna (CRPA) ports and fixed reception pattern antenna (FRPA) ports on the top and bottom of the fuselage.

During JAMFEST, the C-12J carried 746 TS equipment designed to collect airborne reference measurements of the GPS jamming environment. It flew data collection sorties that spanned the airspace and altitudes used by the systems under test.

Figure 4 C-12 J Aircraft

In any test environment where navigation aids are evaluated, it is paramount that the truth reference data is preserved and collected. This is particularly difficult to achieve in a live GPS jamming environment, because many reference systems use GPS to obtain an accurate truth source. To overcome this obstacle, the 746 TS developed the CIGTF Reference System (CRS); this was the reference system used for JAMFEST. The CRS is a rack-mounted (see Figure 5), loosely/tightly-integrated system, consisting of navigation sensors/subsystems, Data Acquisition System (DAS), and post-mission processing mechanization (see Figure 6). Figure 5 CIGTF Reference System

Figure 5 CIGTF Reference System

The DAS, a DOS-based computer, performs the primary functions of data collection and real-time control for the following subsystems: (1) Embedded Global Positioning System (GPS)/Inertial Navigation System (INS) (EGI) navigation system, (2) GPS receiver/receivers, (3) Standard Navigation Unit (SNU) INS, and (4) Cubic CR-100 Range/Range Rate Interrogator/Transponders System (RRS). Other subsystems supported in the CRS

architecture are the GPS Environment Monitoring System (GEMS), data link, altitude encoder, and Satellite Reference Station (SRS) receiver supporting differential GPS (DGPS) algorithms. The post-mission processing mechanization utilizes combinations of the subsystem measurements in an extended Kalman filter/smoother algorithm to produce an optimal reference trajectory.

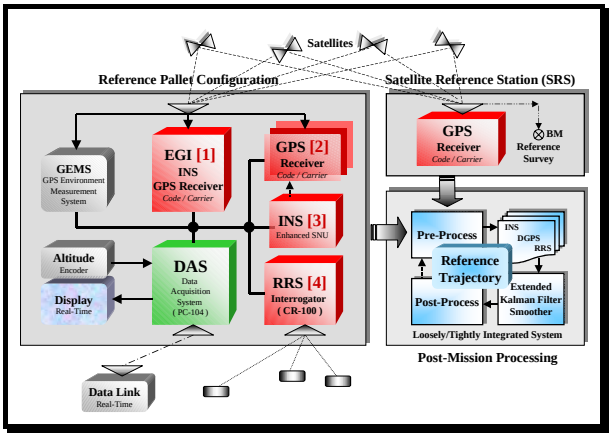


Figure 6 CRS Processing Mechanization

Nominal performance accuracies of the reference trajectory characterized for JAMFEST are detailed in Figure 7.

Subsystem Configuration	RMS Position (m)		
	Horz	Vert	3D
[1] [2] GPS *Code*	2.00	2.25	3.25
[1] [2] DGPS *Code*[1]	1.75	1.75	2.50
Carrier [2]	0.30	0.20	0.35
[4] RRS	1.40	1.00	1.70
SRS Range Constraints: [1] 300-500 nm [2] 50-100 nm			

Subsystem Configuration	RMS Velocity (m/s)			
	East	North	Up	3D
[1] INS/EGI	0.010	0.010	0.010	0.017
[3] INS/ESNU	0.005	0.005	0.005	0.010
Attitude Accuracy: 20 arcsec (Roll,Pitch,Heading)				

Figure 7 Reference System Accuracies

3 EVENT CONDUCT

JAMFEST testing began on 24 May 04 at 2000 MST and spanned 5 days. A total of 12 military organizations, DoD contractors, and civil agencies participated, all with very different goals and objectives.

During the test week, 746 TS engineers conducted GPS jamming operations from 2000 to 0400 hours on each test day, and characterized the jamming field with ground and aviation monitoring equipment. Additionally, the 746 TS deconflicted all customer flight and ground operations and provided on-site technical experts to help resolve customer difficulties and ensure each objective was met. In some cases, this required significant instrumentation and analysis support.

On each test day, two types of scenarios were offered: (1) Operationally realistic and (2) Experimental scenarios. Operationally realistic scenarios included threat laydowns consisting of one, four, and seven jammers broadcasting on L1 and L2 frequencies and using a variety of waveforms and power levels. Experimental scenarios, on the other hand, were useful for research and development efforts requiring high jamming levels capable of stressing robust anti-jam electronics. These scenarios were achieved by using seven close-proximity jammers focused in the same direction.

Most JAMFEST participants utilized their own test beds and recorded their own receiver data and reference information. These participants either mounted their equipment in rental vehicles, government vehicles and aircraft or walked through the jamming environments. In other cases, the 746th TS provided support to participants who could not supply their own test beds, data collection systems, or reference data. In this situation, customer assets were mounted into the 746 TS land navigation vehicles. Customer assets were connected to FRPAs, CRPAs or prototype antennas, depending on the customer's desires and asset availability.

The jamming scenarios were carefully developed to maximize efficiency and meet everyone's goals. A total of 59 jamming scenarios with different threat laydowns were executed during the test week. Jammer placement was carefully planned to maximize the number and variety of scenarios offered while minimizing relocation and set-up time. Figure 8 depicts three sample jammer placement scenarios.

Utilizing a configuration similar to the one depicted, permitted the execution of one jammer, three jammer and seven jammer scenarios without relocating any of the jammers. This offered the most scenario flexibility while limiting the number personnel required to operate the jammers at these locations. For example, in a one jammer scenario, only the jammer at TX 8 may be used or in a three jammer scenario, the jammers at TX 6, TX 7 and TX 1 may be used. Lastly, in a seven jammer scenario the jammers at TX 1, TX 2, TX 3, TX 4, TX 5, TX 6, and TX 8 may be used. Typically, when these jammers were turned on, vehicles would drive down the corresponding range road, park at a predesignated location or fly though the jamming field. While most participants drove or flew during testing, other participants tested hand-held receivers and walked near the jamming field.

During testing, the jammer configuration alternated between operationally realistic and experimental scenarios. All scenarios utilized a variety of waveforms

at low, medium, high, and ramped power levels. The specific waveforms broadcast included Bi-Phase Shift Key, Broadband, Partial Band, Continuous Wave and Swept Continuous Wave, and Pulsed Continuous Wave on both L1 and L2 frequencies.

Figure 8 Sample Jammer Laydown

Another jammer laydown used, involved placing multiple jammers along a predesignated range road all pointed in the same direction. Participants drove into and out of the field to test their equipment in a concentrated GPS jamming environment. Regardless of the means or scenarios used, the participants successfully met their objectives.

Following each test day, 746 TS personnel checked ground jammer logs and collected reference data for accuracy and proper format and provided this information to each participant in the form of a data package. The purpose of the data package was to accurately document the event, cite any necessary deviation(s) from the test plan, detail exact scenarios as they are transmitted, and provide reference data that describes the signals received. Information in the data package was sufficient for each participant to evaluate their own data and generate defensible conclusions.

4 SUMMARY

JAMFEST serves as an affordable avenue to identify system limitations in a GPS jamming environment so that system designers and users can begin to identify and mitigate vulnerabilities in their specific applications. This is particularly valuable information to civil users who otherwise would not have access to such vulnerability scenarios. After participating in JAMFEST customers are better armed with realistic vulnerability data, to better understand their system limitations, work to mitigate these effects, and apply backup systems or procedures as appropriate.

In addition to civil GPS users, JAMFEST also benefited operational military units who are likely to experience GPS jamming during operational conflicts but may not have actually experienced the effects of jamming during training maneuvers. Training in such electronic warfare environments raises vulnerability awareness and affords the opportunity to devise, implement, and practice countermeasures.

All participants reached their objectives and praised the planning, organization and execution of JAMFEST. The participants agreed that the event was worthwhile and indicated that they would be interested in attending future JAMFESTs.

As long as there is interest in GPS vulnerabilities, the 746 TS plans to conduct JAMFEST events. Although the first JAMFEST was in essence “free” to the participants, there may be a nominal fee associated with future test events. The fee is contingent upon sponsorship from other organizations and the complexity of testing or analysis desired by the participants.

Future JAMFESTs will focus on expanding the participant base to include not only the operational and test communities, but add US allies, civil users such as the Dept of Transportation, power and telecommunications industries. The next JAMFEST is planned for November 2004 with another following in mid-May 2005.

ACKNOWLEDGEMENTS:

The authors would like to acknowledge the support of the GPS Joint Program Office, White Sands Missile Range, and the 46th Test Group.

REFERENCES

NAVSTAR GPS System Protection Guide, GPS JPO, LAAFB, CA; 13 Jun 97.

46 TGI 99-1, Safety Review Process; 01 Jun 01.

46 TGI 99-101, Test Management; 30 Aug 01.

AFI 10-1101, Operations Security; 31 May 01 (including AAC Sup 1; 08 Apr 99 vice 31 Mar 95).

AFI 35-205, Air Force Security and Policy Review Program; 02 Nov 95 (Rev. 17 Jan 97).

AFI 91-202, The US Air Force Mishap Prevention Program; 01 Aug 98.

MIL-STD-882D, Standard Practice for System Safety; 10 Feb 00

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 45-48

Journal of Global Positioning Systems
2004

Aperture Jitter Effects in Software Radio GNSS Receivers

Andrew Dempster

School of Surveying and Spatial Information Systems, University of New South Wales, Sydney 2052
e-mail: a.dempster@unsw.edu.au Tel: +61 2 93854182; Fax: +61 2 93137493

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Increasingly, software radio techniques are being used in the implementation of communications receivers in general, and GNSS receivers in particular. In such a receiver, the received signal is sampled as close to the receive antenna as possible, and all subsequent processing uses digital signal processing (DSP) techniques. The sampling clock will suffer from phase noise instabilities, leading to a phenomenon known as aperture jitter. This paper examines the effects of aperture jitter for a number of "typical" software radio GNSS receivers. A jitter specification is derived which restricts the noisy effects due to jitter to 10dB below thermal noise. It transpires that regardless of the new signals that are selected to accompany it, it is the L1 signal that drives this jitter specification.

Key words: Software radio, GNSS receiver, aperture jitter.

1 Introduction

1.1 Bandpass Sampling

With the increase in the variety of signals required to be processed by radio receivers, versatile receiver designs are gaining in popularity. Software radio (SWR) receivers (which perform the "radio" functions in a processor) and software-defined radio (SDR) receivers (where these procedures are also performed digitally but using hardware controlled and configured in software) are two approaches that deliver this versatility. In both these approaches, the ideal is to convert the received signal to a sampled digital signal as closely as possible to the receive antenna. With the advent of new GPS signals, and a range of new Galileo signals, GNSS is an application where SWR and SDR are likely to have an impact.

Sampling the received signal using low-pass sampling (LPS), the more usual interpretation of Nyquist's sampling theorem, requires a sampling rate of twice the maximum frequency of interest. For GNSS, this is of the order of 3.2Gs/s. Analogue-to-digital converters (ADCs) do exist that will operate at these frequencies (Lee, 2003, Poulton, 2003), but they are expensive and consume a great deal of power. A cheaper, more efficient method of conversion in a SWR receiver is "bandpass sampling" (BPS) (Vaugan, 1991), where, consistent with Nyquist's sampling theorem, it is possible to sample an RF signal at twice the bandwidth of that signal. That signal is the "aliased" into the baseband used for that sampling rate. For bandpass sampling to be successful, out of band signals must be attenuated to reduce aliasing of unwanted signals and noise, requiring a high-Q bandpass filter. This is a more demanding design requirement than the low pass anti-aliasing filter required for LPS at the same sampling frequency. When several bands are required for downconversion, the minimum sampling rate is twice the sum of the bandwidths. However, there are constraints that mean that the sampling rate is usually higher than that minimum.

Software receivers have been designed for the GPS L1 signal (Akos, 1996, Ledvina, 2003, Lin, 2001) and for the L5 signal (Ries, 2002). Bandpass sampling dual band receivers have also been designed, for L1 GPS and Glonass (Akos, 1999), and for GPS L1 and L3 (Thor, 2003).

1.2 Aperture Jitter

Aperture jitter is defined as the sample-to-sample variation in time between the effective points at which the samples are actually taken (Sheingold, 1986). The effects of aperture jitter can be modelled as noise (we will use the terminology of (Sun, 2004). If the signal into the analogue-to-digital converter (ADC) is $y(t)$, then the jitter noise can be modelled as:

$$\varepsilon_\tau(n) = y(t_n + \tau_n) - y(t_n) \quad (1)$$

where t_n is the ideal sampling instant and τ_n is the offset to that instant due to jitter. The power of this jitter noise signal is:

$$N_\tau = E\left[\varepsilon_\tau^2(n)\right] - E\left[\varepsilon_\tau(n)\right]^2 \quad (2)$$

Previous work has identified two formulae for this jitter noise power. Both assume a sinusoidal signal and zero-mean gaussian jitter around fixed "perfect" sampling points. This latter assumption has been found experimentally to be reasonable (Shinagawa, 1990), although the spectrum is not white if a phase-locked loop is used to generate the clock (Da Dalt, 2002), or if the jitter is accumulative(Awad, 1998). The first formula assumes that $2\pi f_c \tau_n << 1$, where f_c is the carrier frequency of the sampled signal, i.e. that the jitter offset is much less than a period of the carrier (Shinagawa, 1990):

$$N_\tau = 2\pi^2 f_c^2 \sigma_\tau^2 A^2 \quad (3)$$

where σ_τ^2 is the variance of τ_n and A is the amplitude of the sinusoid. The second formula makes no assumption about the size of τ_n (Awad, 1998):

$$N_\tau = A^2\left(1 - e^{-2\pi^2 f_c^2 \sigma_\tau^2}\right) \quad (4)$$

Other simple non-sinusoidal signals such as square and triangular waves have also been examined with regard to their jitter noise power (Kobayashi, 1999). Comparisons have been made between the noise generated by jitter and thermal (Shinagawa, 1990) and quantisation (Kobayashi, 1999) noise.

2 Example GNSS Receivers

2.1 Frequency Bands

Sampling requirements for a number of GNSS example systems are examined, restricting interest to GPS and Galileo signals that are available to commercial users.

Commercial satellite navigation receivers have until recently largely been restricted to using the GPS signal at L1 (carrier 1575.42MHz, chipping rate 1.023Mcps) (ICD, 2003). Soon GPS will provide a similar signal on L2 (carrier 1227.6 MHz) (ICD, 2003, Fontana, 2001) and a new signal on L5 (carrier 1176.45 MHz, chipping rate 10.23Mcps) (ICD, 2002, RTCA, 2000, van Dierendonck 2000).

Galileo will have several signals available to commercial users. Open services (OS) are available on E5 (1164-1215MHz, nominal carrier 1189MHz) and E2-L1-E1, known for convenience as L1 (1559-1592Mz). Commercial services (CS), for which a fee is required but are not restricted to security services, are available on E6 (1215-1300MHz, carrier 1278.75MHz) and L1. At the time of writing, the above frequency allocations were the latest to have been formally released (Hein, 2002, Hein, 2003), although they will change because of agreements between the Galileo and GPS teams (GPSW, 2004, 2004A). For instance, the OS L1 signal bandwidth has been reduced by a factor of 2 so that it has a 1.013MHz chipping rate, and a binary offset code of 1.013MHz, a so-called BOC(1,1) code.

Combinations of these signals were selected that are likely to be common within a "GNSS" receiver. Example 1 is the most expensive receiver processing GPS L1, L2, L5 and Galileo OS and CS, using L1, E5 and E6. Example 2 uses only the free-to-air GPS L1 and L2 and Galileo OS, on L1 and E5. Example 3 uses L1 and E5 (ignoring L2) and example 4 uses L1 and L2 (i.e. the GPS bands that are "currently" available, and incorporating the Galileo L1). Example 5 is cheapest arrangement, simply using "L1 only" and incorporating the new Galileo signal.

Table 1. Frequency bands for example receiver designs

Band	E5 (L5)	L2	E6	L1
Fmin	1164	1217	1260	1573
fmax	1214	1238	1300	1577
Ex. 1	X	X	X	X
Ex. 2	X	X		X
Ex. 3	X			X
Ex. 4		X		X
Ex. 5				X

The frequencies required for the four examples are shown in Table 1. It can be seen that because the frequency bands are contiguous, in all cases, only two bands are required. Example 1, for instance, must receive in the range 1164-1300MHz and 1573-1577MHz.

2.2 Sampling Rates

For simplicity, as the type of results sought need only to be indicative, bandpass sampling is assumed, but only the minimum sampling rate required by Nyquist's theorem is considered (i.e. the full set of constraints identified in (Vaugan, 1991) are not all accounted for). These minimum sampling rates are shown in Table 2.

2.3 Data Bandwidths

As we have seen, aperture jitter tends to be modelled as noise. This noise can be considered to be additive with the thermal noise present in the receiver. For the new signals, which have data in one channel and a dataless

pilot signal in quadrature, the effects of noise are greater in the data channel, so we will consider the bandwidth of the data channel as being the relevant bandwidth in which thermal noise appears. The data bandwidths of the GNSS signals are shown in Table 3.

Table 2. For the 5 examples, the lower bandwidth, the upper bandwidth, the total bandwidth, and the minimum sampling rate, considering only Nyquist's theorem. All figures in MHz

	BW 1	BW 2	Total BW	f_{smin}
Ex. 1	136	4	140	280
Ex. 2	74	4	78	156
Ex. 3	50	4	54	108
Ex. 4	21	4	25	50
Ex. 5	-	4	4	8

Table 3 GNSS data bandwidths

GPS band [ICD (2003), van Dierendonck (2000)]	Data bandwidth (Hz)
L1	100
L2	100
L5	2000
Galileo band [Hein (2002), Hein (2003)]	
L1	500
E5	500
E6	2000

2.4 Signal Strengths

We will be comparing SNR due to noise to SNR due to jitter. In order to do this, we will need to know the signal levels for each of the signals we are examining. These are shown in Table 4.

Table 4 GNSS signal strengths

GPS signal [ICD (2003), van Dierendonck (2000)]	Signal strength (dBW)
L1	-160
L2	-160
L5	-154
Galileo signal [Hein (2002), Hein (2003)]	
L1	-155
E5	-152
E6	-155

3 Derivation of Acceptable Aperture Jitter

It is assumed that the jitter is relatively small with respect to the carrier frequency, and hence equation (3) rather than (4) is used to evaluate the contribution to the noise power made by aperture jitter. The SNR due to jitter arising from (3) is:

$$SNR_j = \frac{S}{N_\tau} = \frac{\frac{A^2}{2}}{2\pi^2 f_c^2 \sigma_\tau^2 A^2} = \frac{1}{\left(2\pi f_c \sigma_\tau\right)^2} \tag{5}$$

In order to keep this contribution to an insignificant size, the jitter noise power is restricted to be 10dB down from the thermal noise power, as in equation (6).

$$SNR_j \geq 10\frac{S}{N_{th}} \tag{6}$$

where S is the signal power from Table 4, and $N_{th} = kTB$ is the thermal noise power, with the bandwidths of interest B being taken from Table 3. Thus, using equation (6), it is possible to evaluate the required jitter standard deviation allowable for each GNSS band of interest, by using

$$\sigma_\tau \leq \frac{1}{2\pi f_c}\sqrt{\frac{N_{th}}{10S}} \tag{7}$$

Evaluating this expression for the values selected in the previous sections gives the jitter requirements in Table 5.

Table 5 Jitter requirements (standard deviation) for each of the signal bands

GPS signal	Jitter requirement (ps)
L1	2.05
L2	2.63
L5	6.15
Galileo signal	
L1	2.58
E5	2.42
E6	6.35

These jitter requirements do not vary very much with the loosest (6.35ps) being only 3.1 times the tightest (2.05ps). The tightest of these jitter requirements is for the GPS L1 band, which is probably not surprising as it is the oldest of the signals. However, it has been included in all of our examples, so the 2ps jitter requirement is the one that drives the specification for all of the examples. This requirement can be turned into a phase noise standard deviation requirement of the sampling clock:

$$\sigma_{\theta\tau} = 2\pi f_s \sigma_\tau \quad (8)$$

where f_s is the bandpass sampling frequency. Interestingly, as the sampling frequency increases, the phase noise requirement gets looser. Therefore, in phase noise terms, the most difficult receiver to design is what in other ways is the easiest receiver to design, the "L1-only" example, giving a phase noise requirement of 50μrads.

4 Conclusion

The analysis evaluated the jitter requirement such that noise due to sampling jitter at the carrier frequencies of GNSS signals was 10dB less than the thermal noise. For all GNSS bands, this requirement was of the order of picoseconds. Several combinations of bands were examined but they all used the L1 band, which has the tightest requirement for jitter, a standard deviation less than 2ps. However, that requirement is not much less than for L2 and E5. The jitter analysis assumed that $2\pi f_c \tau_n << 1$, and for the tightest specification $2\pi f_c \sigma_\tau = 0.02$, which satisfies that assumption.

References

Akos, D M, and J B Y Tsui (1996), ***Design and Implementation of a Digital Digitization GPS Receiver Front End***, IEEE Trans Microwave Theory and Techniques, vol 44, no 12, Dec 1996

Akos, Dennis M, et al (1999), ***Direct Bandpass Sampling of Multiple Distinct RF Signals***, IEEE Trans Communications, vol 47, n0 7, July 1999, pp983-988

Awad, Selim Saad (1998), ***Analysis of Accumulated Timing-Jitter in the Time Domain***, IEEE Trans Instrumentation and Measurement, vol 47, no 1, Feb 1998, pp69-73

Da Dalt, Nicola, et al (2002), ***On the Jitter Requirements of the Sampling Clock for Analogue-to-Digital Converters***, IEEE Trans Circuits and Systems I, vol 49, no 9, Sep 2002, pp1354-1360

Fontana, R D, et al (2001), ***The Modernized L2 Civil Signal***, GPS World, Sep 1, 2001

GPSW (2004), ***GPS and Galileo Reach Signal Agreement***, Global View, GPS World, Mar 2004, pp12-13

GPSW (2004A), **EC: GPS-Galileo Deal Has Mutual Benefits**, Global View, GPS World, Apr 2004, p 13

Hein, G, et al (2002), ***Status of Galileo Frequency and Signal Design***, Proc ION-GPS 2002

Hein, G, et al (2003), ***Galileo Frequency and Signal Design***, GPS World, June 2003, pp30-37

ICD-GPS-705 (2002), ***Interface Control Document: Navstar GPS Space Segment/ Navigation L5 User Interfaces***, US DOD, 29 Mar 2002

ICD-GPS-200C (2003), ***Interface Control Document: Navstar GPS Space Segment/ Navigation User Interfaces***, US DOD, IRN-200C-005R1, 14 Jan 2003

Kobayashi, Haruo, et al (1999), ***Aperture Jitter Effects in Wideband ADC Systems***, Proc ICECS '99, vol 3 ,5-8 Sept. 1999 pp1705-1708

Ledvina, B.M., M.L. Psiaki (2003), S.P. Powell, and P.M. Kintner, ***A 12-Channel Real-Time GPS L1 Software Receiver***, Proc. ION NTM, Jan 22-24, 2003. Anaheim, CA

Lee, J, et al (2003), ***10 Gsample/s optoelectronic A/D converter***, Electronics Letters, vol 39, no 23, 13 Nov 2003

Lin, D M, and J B Y Tsui (2001), ***A Software GPS Receiver for Weak Signals***, Microwave Symposium Digest, 2001 IEEE MTT-S International , vol 3, 20-25 May 2001, pp2139 – 2142

Poulton, K, et al (2003), ***A 20GS/s 8b ADC with a 1MB Memory in 0.18μm CMOS***, Proc Int Solid-State Circ Conf, 2003, IEEE

Ries, L, et al (2002), ***A Software Receiver for GPS-IIF L5 Signal***, Proc ION-GPS 2002, Sep 2002, pp1540-1553

RTCA SC-159 (2000), ***DO-261 Navstar GPS L5 Signal Specification***, 14 Dec 2000,

Sheingold, Daniel H, (ed) (1986), "Analog-Digital Conversion Handbook, 3rd ed., Prentice Hall, 1986

Shinagawa, Mitsuru, et al (1990), ***Jitter Analysis of High-Speed Sampling Systems,*** IEEE J of Solid State Circuits, vol 25, no 1, Feb 1990, pp220-224

Sun, Yi-Ran, and Svante Signell (2004), ***Effects of Noise and Jitter on algorithms for Bandpass Sampling in Radio Receivers***, Proc. ISCAS 2004, IEEE, Vancouver, Canada, 2004, pp I-761 – I-764

Thor, Jonas, and Dennis M Akos (2002), ***A Direct RF Sampling Multifrequency GPS Receiver***, Proc Position Location and Navigation Symposium, 2002 IEEE, 15-18 April 2002 pp44-51

Van Dierendonck, A J, (2000), ***The New L5 Civil GPS Signal***, GPS World, Sept 2000, pp64-71

Vaugan, Rodney G et al (1991), ***The Theory of Bandpass Sampling***, IEEE Trans Signal Processing, vol 39, no 9, Sept 1991, pp1973-1984

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 49-56

Journal of Global Positioning Systems
2004

Benefits of a Reconfigurable Software GNSS Receiver in Multipath Environment

Fabio Dovis, Marco Pini, Massimiliano Spelat

Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
Tel: +39 011 2276 415 Fax: +39 011 564 4099 Email: name.surname@polito.it

Paolo Mulassano

Istituto Superiore Mario Boella (ISMB), Via P. Boggio 61, 10138 Torino, Italy
Tel: +39 011 2276 414 Fax: +39 011 2276 299 Email: paolo.mulassano@ismb.it

Received: / Accepted: 15 November 2004 / Accepted: 3 February 2005

Abstract. The increased interest for applications based on both navigation and communication systems, represents an important driver for the design and implementation of innovative receivers architectures. The realization of civil GPS applications, the advent of the European navigation system Galileo, and the integration of localization services in communication network make the reconfigurability an indispensable requirement for the development of innovative Global Navigation Satellite Systems (GNSS) platforms. In addition, it must be pointed out that several problem as indoor positioning and multipath recovery, are pushing the research activity in order to provide users of flexible devices able to adapt their functionalities according to the environment. Considering this complex scenario, the Software Defined Radio (SDR) approach constitutes an interesting perspective to develop modular architectures. In this paper, the implementation of a reconfigurable user terminal integrating both navigation and communication capabilities will be discussed. The work will be presented focusing the attention on software-designed functionalities and the Navigation Unit will be analyzed and tested. An example of adaptability of the receiver to the operating environment will be presented. The reconfigurable module for multipath mitigation in the tracking phase will be described with particular attention to the implementation aspect, and some simulation results will be presented.

Keywords. GNSS Receiver, Software Defined Radio, Reconfigurability, Multipath.

1 Introduction

Satellite navigation, positioning and timing has already found widespread applications in a large variety of fields. It is well known that the GNSS programme has given an important contribution to the research activity, and in particular on topics based on the development of user terminals. The European Galileo project constitutes, for the next decade, an essential contribution to the satellite navigation market. In this context, it has to be remarked that in addition to the realization of basic system segments, a relevant effort is being devoted to the design of *Local Elements* (LE) for providing assistance data to users in the coverage area. As a direct consequence, the study and design of advanced terminals able to deal with both the *Signal-In-Space* (SIS) and other classes of signals transmitted by stations and satellite augmentations (EGNOS), become indispensable. In particular, it is expected that at the LE level, *User Terminal* (UT) will have to deal with the innovative Location Based Services that are being developed for the cellular network, with particular attention to emergency services (e.g. the European emergency number E-112).

At the same time, the GPS Joint Program Office has launched the "Modernization Phase", which consists in the study of new frequencies and spreading codes for additional military and civil signals. Given these innovations, a civil user will receive additional SIS for improving accuracy, integrity and continuity of services; on the other hand, military users will increase the security with a shorter time to first fix and a faster signal acquisition procedure.

Considering this complex scenario, flexibility represents an essential feature for the UTs. This peculiarity can be

obtained using a Software Defined Radio (SDR) approach, which allows for the development of reconfigurable devices realized on a modular architecture. It is clear that the reconfigurability aspect can play a key role in specific applications where the features of the environment surrounding the UT, change with time. A typical example is a user moving in an indoor scenario. It is important to remark that the indoor performances of GNSS are strongly limited by several factors. Acquisition of weak signals is critical, visibility of a sufficient number of satellites is not guaranteed, and the SIS received at the antenna propagates via multiple paths. In this case the reconfigurability allows to adapt the receiver functionalities to the actual operating conditions.

Following these considerations in according to the most recent research works on GNSS receiver (Akos, 2003), the paper discusses a receiver architecture based on SDR techniques where functionalities are software-implemented on modular platform composed by *Field Programmable Gate Array* (FPGA) and high speed *Digital Signal Processor* (DSP).

In the paper the concept of reconfigurability assumes two different meanings:

1. *Stand-by reconfigurability*. In this case the user is allowed to upload new software routines which correspond to new receiver functionalities (e.g., an innovative PVT computation strategy);
2. *On-fly reconfigurability*. In this case the Algorithm Control Unit (see Fig. 1) has a set of pre-loaded modules (corresponding to different receiver functionalities or strategies) and it is able to select in real time the best set of routines in order to match the requirements (i.e. trading-off time and power consumption with processing of the SIS, use of acquisition strategies tailored to the environment, etc).

In the first part of this paper, on the basis of previous works (Mulassano et al., 2001) (Dovis et al., 2002), the flexibility of a reconfigurable platform based on SDR technique (Buracchini, 2000) is introduced with particular attention to the implementation of the Navigation Unit. From this standpoint, in order to explain advantages given by the on-fly reconfigurability, the paper is focused on the analysis of a routine based on a multiple Delay Lock Loop (DLL) structure for reducing the effect of multiple paths. Considering the implementation phase, the architecture block diagram is discussed, and simulation results are presented with particular attention to the optimization procedure.

2 Analysis of GNSS Modular Receiver Architecture for Advanced Algorithms Implementation

The conceptual scheme of a reconfigurable terminal for positioning applications, based on SDR techniques is shown in Figure 1. As described in (Mitola, 1995), one of the main advantages of SDR is the design of a flexible system able to adapt in real time its functionalities in order to match the requirements. At the same time it allows to optimize the trade-off among speed, power dissipation and programmability.

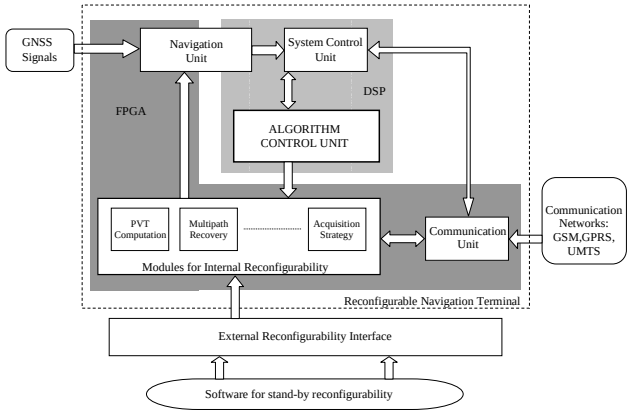


Fig. 1 GNSS receiver functional architecture.

The architecture proposed in Fig. 1 represents the functional diagram of a hybrid *Navigation/Communication* (NAV/COM) user terminal based on FPGA and DSP, as considered for the work presented in this paper. It is clear that the user terminal functions are distributed among the different functional blocks and modules:

1. The on-fly reconfigurability is managed by the Algorithm Control Unit, an advanced routine implemented on DSP that represents the core of the innovative aspect of the architecture proposed. It allows for the real-time configurability, and consists in a set of commands set for the control of internal modules available for the System Control Unit.
2. The Navigation and Communication Units constitute the core blocks of the system; in fact, they are able to acquire and tracking signals incoming from satellites and *Base Tranceiver Stations* (BTS) on the basis of standard routines or advanced algorithms provided by the Internal Modules. Both the NAV and COM blocks are implemented on the FPGA, and they are controlled by the System Control Unit based on DSP.

3. All modules for the on-fly reconfigurability are software uploaded (on the FPGA) and, during the stand-by phase, can be changed, upgraded and integrated employing the external reconfigurability interface. They represent complex algorithms used by the receiver System Control Unit for contrasting the performance worsening due to the critical environment, saving on the power consumption when the complex processing is not required by the environmental conditions.

Taking into account the previous considerations, it is clear that the main feature proposed by this architecture is the capability of processing satellite SIS and aiding signals allowing mutual interaction between the System Control Unit and the Algorithm Control Unit. Focusing the attention on the innovative aspects based on advanced Internal Modules, it is significant to analyze the multipath recovery algorithm because represents a complex situation that highlights the adaptability of the receiver functionalities. In particular, several simulation results are presented, and the architecture proposed is discussed taking into account the implementation procedure. In order to understand what are the potentialities of the SDR implementation, in the next session the architecture of the Navigation Unit is analyzed and some performance results are highlighted.

3 The Navigation Unit

The Navigation Unit shown in Fig. 1 represents a device able to acquire and process GNSS signals.

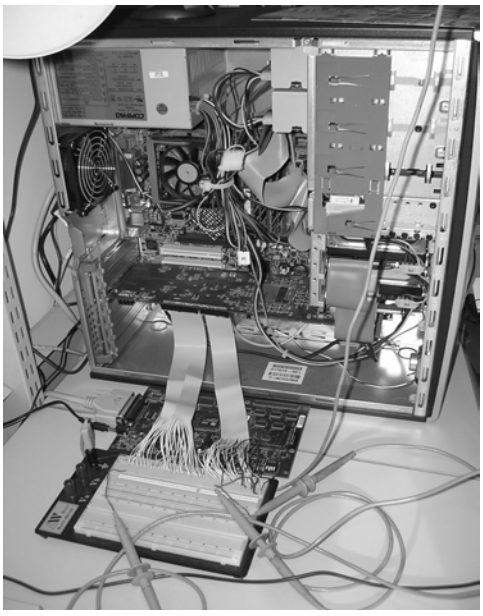


Fig. 2 SDR platform for the Navigation Unit implementation.

Given the flexibility of the overall GNSS receiver architecture presented in the previous section, this unit has been designed and implemented on reconfigurable hardware according to the SDR principles. On the basis of the data processing rate, the structure of the Navigation Unit has been mapped on specific devices, which allows the management of signals at different frequencies. In particular, the resources partitioning has been realized considering a Xilinx FPGA Virtex XCV2000EBG560 and a Texas Instruments DSP C6711 (floating point).

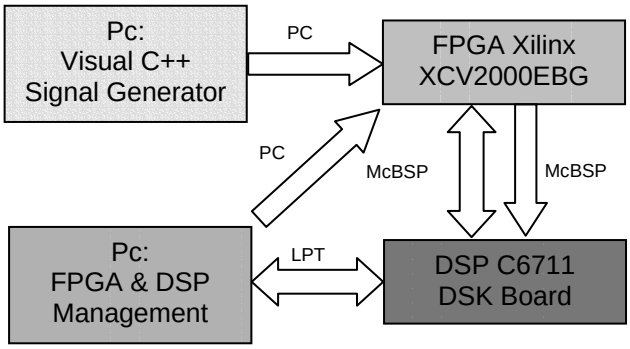


Fig. 3 Communication link between programmable devices.

Fig. 2 and Fig. 3 show the platform employed for the implementation, and the block diagram highlights the communication links between the FPGA, the DSP and a common PC.

3.1 Architecture Description

A detailed block diagram for the Navigation Unit Architecture in the case of single channel is reported in Fig. 4. *The Radio Frequency* (RF) front-end, which is responsible for reception, filtering and A/D conversion of incoming GNSS signals, is modelled by a software simulator in C language; in particular, in-phase (I) and quadrature (Q) samples of satellite signals are computed and written into a binary file on the basis of the signals properties specified in the *Signal-In-Space Interface Control Documents* (SIS ICD). This file is then read and data are stored into the FPGA on-board *Random Access Memory* (RAM) in order to assure data stream continuity at the input of the base-band stage. I and Q samples are processed by the base-band channel implemented on the FPGA, which is in charge of correlating these samples with a local replica of the satellite signal (generated on the FPGA itself). Correlator outputs are fed to the DSP in order to apply advanced algorithms for computing corrections for both the acquisition and the tracking phases.

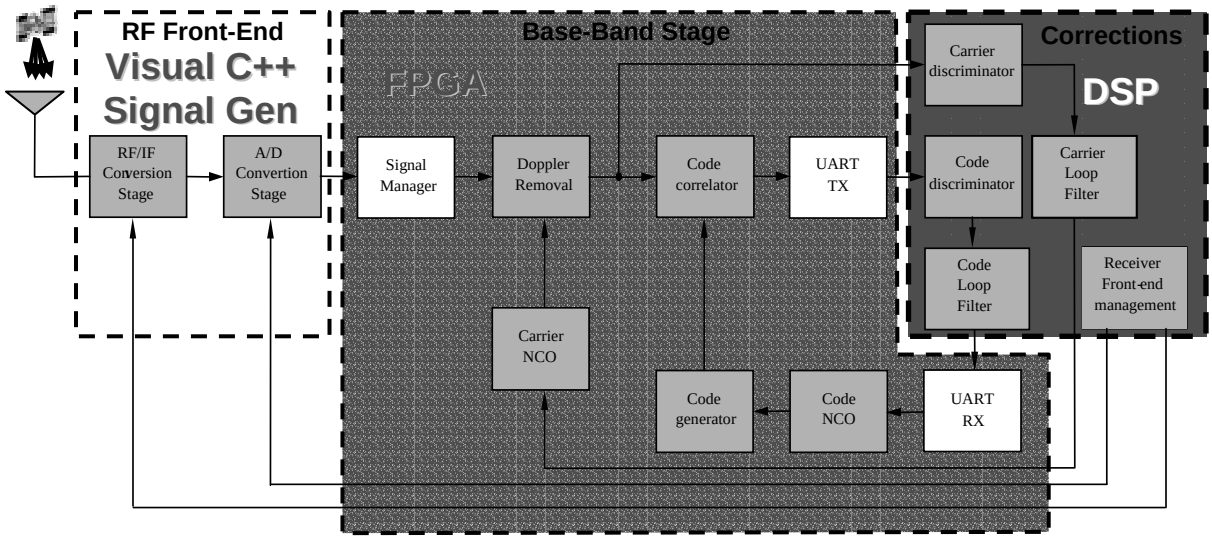


Fig. 4 Navigation Unit architecture.

Finally, these values are sent back to the *Numerically Controlled Oscillator* (NCO) which drives the local code generator. Due to the fact that the internal clock of the Navigation Unit is provided by the FPGA on-board oscillator, every data transfer between FPGA and DSP allows to operate in real-time.

As shown in Fig. 4, the base-band channel and the corrections stage constitute the core of the Navigation Unit because they realize the double loop for the tracking of both code and carrier: the Delay Lock Loop (DLL) and the Phase Lock Loop (PLL) respectively. During the acquisition phase, a specific DSP algorithm is responsible for the search of satellites in view and in particular, for the raw alignment between the incoming and the local code with a certain error. After the acquisition, the code and doppler shift computed are used to initialize the tracking loops, and both the discriminator and the loop filter algorithms (DSP implemented) are responsible for the fine alignment of the codes.

As shown in Fig. 4, it is important to highlight that the design of the base-band structure implemented on the FPGA also consists in the realization of the "Signal Manager" block. It manages the I and Q samples stored into the RAM and sends them to the input of the channel.

Finally, in order to assure a correct communication between the base-band stage (FPGA) and the correction stage (DSP), two interfaces for the serial data transmission (called UART TX and UART RX) have been realized.

Considering the overall architecture implemented on the Xilinx programmable hardware, which is constituted by 2,541,952 gates, it is important to provide an estimation of the resources used in order to understand the complexity of the Navigation Unit. Tab. 1 summarizes the FPGA implementation report in the case of one base-band channel.

Section (one channel)	***Num. Of Slices***	***Num. Of Gates***	***Complexity***
Signal Manager Block	110	14,520	0.5 %
UART TX	42	5,544	0.22 %
UART RX	50	6,600	0.26 %
Base-Band Channel	1,320	174,240	6.85 %
Overall Structure	1,522	200,904	~ 7.83 %

Tab. 1 Navigation Unit: resources analysis.

3.2 Performance Results for the Navigation Unit

In order to test the Navigation Unit performance the ability of acquiring a code and of maintaining the lock have been evaluated, The test of the SDR platform has been realized using the advanced *Real Time Data eXchange* (RTDX) controls which allow to create a buffer for the communication link between the DSP and a generic Host Application, which in this case is constituted by a Matlab routine.

In presence of noise the sample stream has been obtained considering a value of C/N_0 equal to 60 dBHz at the input of the RF filter. In particular, the double side band of the IF filter (see Fig. 4) has been set to 8 Mhz and a 3rd order Butterworth filter has been employed.

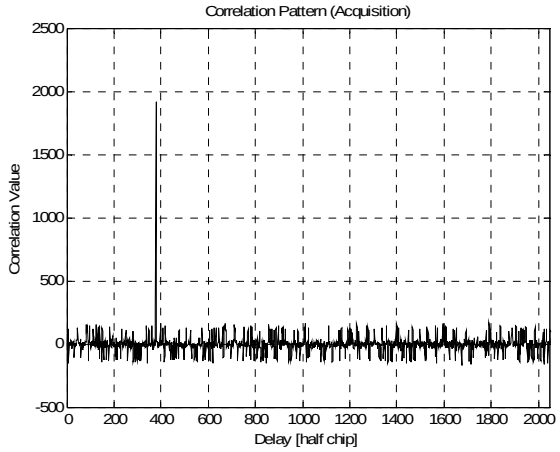


Fig. 5 Acquisition phase: correlation pattern.

The sampling frequency used for the satellite signal generation within the simulator, has been set equal to 2.182 MHz and therefore, the system operates with 2.13 samples per chip.

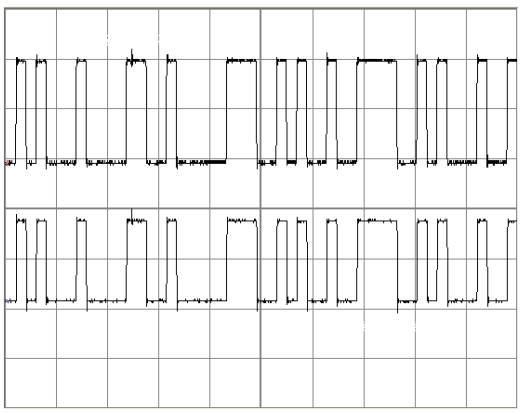


Fig. 6 Tracking phase: codes alignment.

With a FPGA on-board clock of 65.472 MHz, samples at the input of the base-band channel are correlated with the local code over an integration time of *one code period*; it means that the maximum value for the output of the correlator can be 2,182. The actual value at the output of the correlator is acquired from the DSP, which compares it with the proper threshold in order to capture the peak in the Correlation Pattern (depicted in Fig. 4 in absence of noise). The DSP also shifts the local code replica (steps of ½ chip) in order to span all the possible delays.

At the end of the acquisition phase, the delay between the two codes is used to track the signal. During this phase the correlation value increases because some DSP algorithms drive the local code generator in order to obtain the fine alignment (Fig. 6). In particular, an "Early-Late Envelope Normalized" discriminator (Equation 1) with a 2nd order FIR loop filter are implemented.

$$Discr\_output = \frac{\sqrt{I_E^2 + Q_E^2} - \sqrt{I_L^2 + Q_L^2}}{\sqrt{I_E^2 + Q_E^2} + \sqrt{I_L^2 + Q_L^2}} \qquad (1)$$

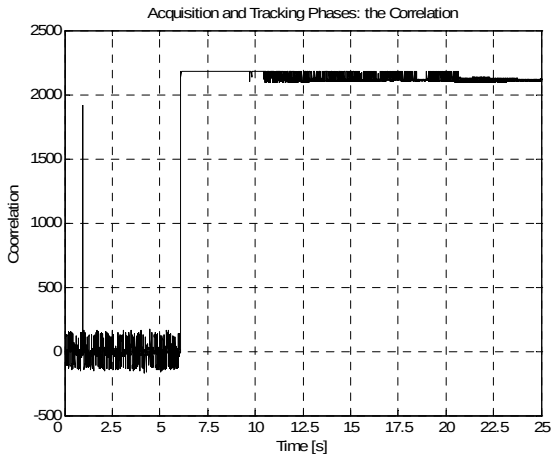


Fig. 7 Correlation during the acquisition and the tracking phases for signal without noise. In absence of noise the acquisition threshold is set equal to 1800.

Fig. 7 and Fig. 8 depict the Navigation Unit performance results in terms of correlation values. Considering the tracking phase, it is possible to see that in presence of noise (Fig. 8) the correlation reaches a mean value of 600, while in the theoretical situation (Fig. 7) the mean of the correlation is near the maximum value of 2,182 according to the results obtained by simulation. It is then possible to conclude that this architecture allows for the realization of advanced modules for the implementation of complex algorithms based on satellite navigation signals.

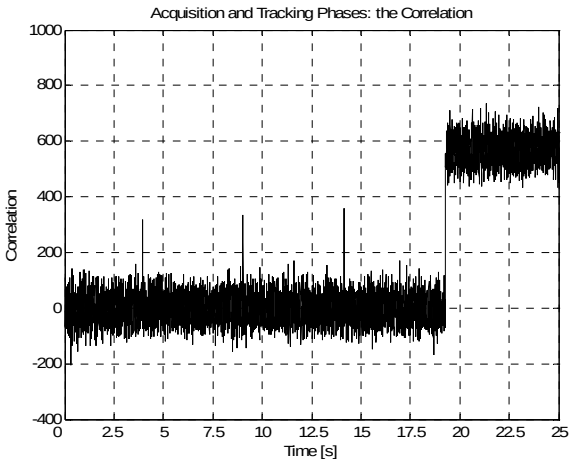


Fig. 8 Correlation during the acquisition and the tracking phases for signal with noise. In this case the threshold used during the acquisition is decreased in order to detect the correlation peak.

4 Multipath Tracking Algorithm for Reducing the Code Alignment Error

The reconfigurable structure described in the previous section has the goal to allow the implementation of advanced Internal Modules, as for example the multipath recovery algorithm.

It is well know that multipath errors become particularly significant in urban or indoor navigation, where the presence of many scatterers around the receiver may be relevant. When a SIS replica, delayed within one chip with respect to *Line Of Sight* (LOS), arrives at the receiver the S-curve of the DLL is distorted. The punctual local code replica of the DLL can not be exactly aligned to the LOS and a systematic error in the pseudorange estimation is experienced. In the past several techniques to reduce the multipath error have been proposed. Among them it is possible to cite the narrow correlator (Braasch, 1996), which allows to get a better alignment reducing the spacing between the early and late local replicas. An innovative system for multipath recovery based on a multiple DLL architecture has been introduced in (Dovis *et al.*, 2004), and its implementation in the SDR receiver is considered in this paper. It is important to highlight that the new architecture is quite different from common rejection techniques. In fact, it is based on the philosophy of the RAKE receiver for communications and employs specific DLLs to track the replicas in order to cancel the distortions of the S-curve of the tracking loop. The architecture implemented in the reconfigurable hardware is depicted in Fig. 9.

A key role is played by the multipath Monitoring Unit which is in charge of activating the multistage architecture according to the scenario. In particular it has to:

- estimate the number of replicas impinging at the antenna employing an algorithm, as for example (Laxton *et al.*, 1977).
- compare them with a predefined threshold to roughly determine their relevance to the distortion of the S-curve.
- activate the multistage procedure for the tracking loop.

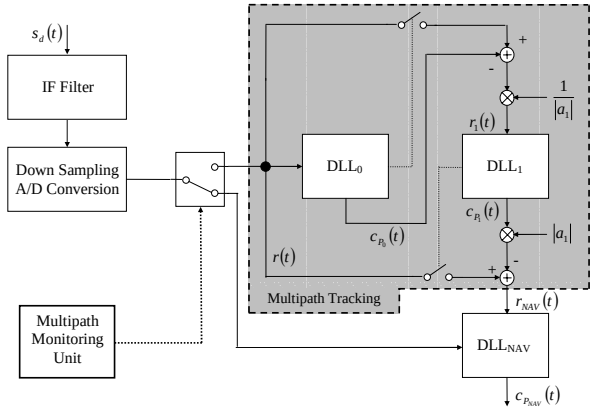


Fig. 9 Multiple DLL architecture for one channel within the navigation receiver.

For sake of presentation of the multistage procedure, let consider a generic user in an indoor scenario, where it is clear that the receiver performance are degraded by multipath due to the presence of obstacles around the antenna. In this scenario, the reconfigurable terminal is able to activate the Module for the Multipath recovery considering the *n*-th paths estimated as arriving at the antenna. In other words, the complexity of the algorithm is increased with respect to the number of multiple paths.

Fig. 9 shows the innovative architecture in the case of one multipath arriving at the receiver, but it can be easily extendable to the general situation in which the input signal arrives at the antenna via *n*-th different paths.

The local code replica of the DLL_0 is usually delayed with respect to the LOS code due to the distortion of the S-curve. Such a code instance is subtracted from the filtered incoming code *r(t)*, giving a signal $r_1(t)$ as input for the DLL_1, which is therefore able to track the multipath. After a transient phase, the first *DLL* tracks the incoming LOS and the second multipath replica. The local code of DLL_1, that follows the evolution of multipath, is then subtracted from the incoming code *r(t)* and the resulted signal is sent as input for the

DLL_{NAV}. The local code of this last *DLL* can be used for pseudorange estimation.

The innovative DLL architecture described aims at reducing the bias error in the code alignment. Therefore, the mean of the tracking error between the LOS code and the DLL_{NAV} local code has been evaluated considering different values of C/N_0 (which represents the ratio between the LOS power and the noise power spectral density at the input of the RF filter) and different multipath delays.

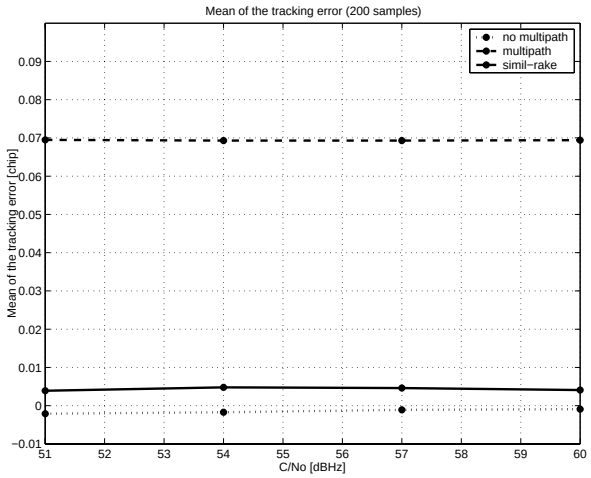


Fig. 10 - Comparison of the mean of the tracking jitter: common DLL with no multipath (dotted line), common DLL with multipath ($\tau_1 = 0.7$ chip) (dashed line), innovative multiple DLL (Fig. 9) with multipath ($\tau_1 = 0.7$ chip) (continuous line).

Fig. 10 shows the result obtained by simulation of the module in case the multipath and LOS signals have the same phase rotation, the multipath is 0.7 chip delayed with respect to LOS and its amplitude is half the LOS amplitude. It can be noted that using a single *DLL* narrow correlator (0.125 chip spacing), no reasonable tracking is possible; in fact, the mean of the tracking jitter is about 0.07 chip apart from the correct zero crossing (dashed line in Fig. 10) for all the C/N_0 values considered. Fig. 10 also reports the performance of a single *DLL* when no multipath is present (dotted line), and the performance of the innovative architecture (solid line). It can be observed from the picture that employing the new scheme, the systematic error can be significantly reduced for all C/N_0 values.

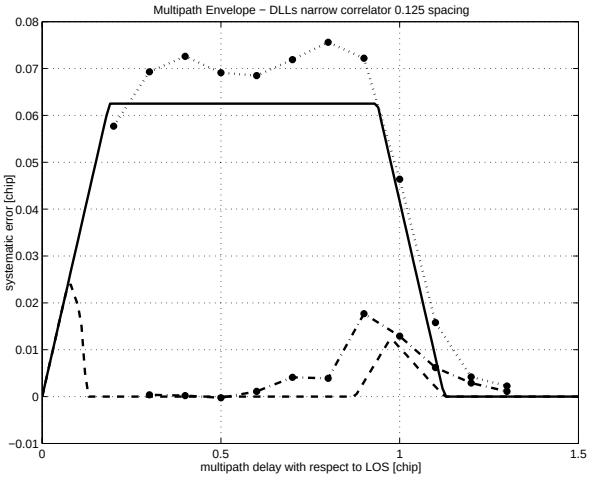


Fig. 11 - Theoretical and simulated multipath envelopes of a single narrow correlator (solid and dotted lines), theoretical and simulated multipath envelopes of the simil-RAKE architecture (dashed and dashed-dot lines).

Fig. 11 shows the theoretical multipath envelops of a single DLL narrow correlator and of the whole simil-RAKE architecture compared to the systematic error obtained by simulation. Both simulation results and theoretical envelopes show that using the innovative architecture, it is possible to drastically reduce the systematic error for all the multipath delays considered. The distance between the simulated values and the theoretical curves can be explained considering that simulations have been performed in presence of noise and taking into account the *Intermediate Frequency* (IF) filter effect on the incoming signal.

This structure can be simply extended and adapted to the external environment estimating the number of multipaths (thanks to the Multipath Monitoring Unit). Therefore, it is clear that it can play a key role in the realization of an advanced module for the on-fly reconfigurability procedure. The drawback of the proposed architecture resides in the increased computational complexity, and consequently to the power required to the FPGA/DSP architecture. It is important to point out that following the core idea behind the general statement of SDR, the implementation must be performed with particular attention to the *mapping* procedure onto the programmable platform, in order to assure the maximum degree of modularity.

5 Conclusions

This paper has discussed the implementation of innovative functionalities for the realization of GNSS receiver according to the SDR philosophy. The study has been conducted highlighting novel aspects of the

architecture which allows several degrees of reconfigurability. In particular, the design faced the choice of the best set of internal modules for satisfying the requirements in terms of precision, power consumption and time to fix. The implementation of the Navigation Unit has been analyzed in details because it represents the core for navigation modules. As a significant case-study of the reconfigurability opportunity, the adaptive multipath rejection algorithm has been discussed, presenting the simulation results of the implemented architecture.

Acknowledgements

Authors would like to thank ISMB (Istituto Superiore Mario Boella) research center for its support. In particular, all simulations and tests have been performed in the Navigation Lab within the Services and Applications Lab of ISMB. In addition, Massimiliano Spelat expresses many tanks indeed to ISMB for the financial support to his Ph.D. program.

References

Akos D M (2003) ***The role of Global Navigation Satellite Systems (GNSS) software radios in embedded systems***, GPS solution, vol. 7, no. 1

Braasch M. W. (1996) ***Performance Comparison of Multipath Mitigating Receiver Architecture***, IEEE PLANS 1996

Buracchini E (2000) ***The Software Radio Concept,*** Commun. Mag., vol.38, no.9, pp. 138-143

Dovis F, Gramazio A, Mulassano P (2002) ***SDR Technology Applied to Galileo Receivers,*** ION GPS 2002

Dovis F, Mulassano P, Pini M (2004) ***Multiple DLL Architecture for Multipath Recovery in Navigation Receivers***, IEEE VTC 2004

Laxton M C, De Vilbiss S L (1977) ***GPS Multipath Mitigation During Code Tracking,*** in Proc. of the American Control Conference

Mitola J (1995) ***The Software Radio Architecture,*** IEEE Commun. Mag., vol.33, no.5, pp. 26-38

Mulassano P, Dovis F, Avagnina D, Gramazio A (2001) ***REGAL: A Reconfigurable Receiver for GPS and Galileo,*** GNSS 2001

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 57-62

Journal of Global Positioning Systems
2004
© CPGPS

Differential LORAN for 2005

Benjamin B. Peterson, Kenneth Dykstra

Peterson Integrated Geopositioning, LLC, 30 Pond Edge Drive, Waterford, CT 06385, USA
e-mail: benjaminpeterson@ieee.org Tel: + 0118604428669; Fax: +011

Kevin M. Carroll, Anthony H. Hawes

U.S. Coast Guard Loran Support Unit, 30 12001 Pacific Avenue, Wildwood, NJ 08260, USA
e-mail: ahawes@lsu.uscg.mil Tel: + 0116095237321; Fax: +01l6095237320

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Note: *The views expressed herein are those of the authors and are not to be construed as official or reflecting the views of the Commandant, U.S. Coast Guard, U.S. Federal Aviation Administration or the U.S. Departments of Homeland Security or Transportation.*

Abstract. A multimodal group of engineers, scientists, and industry representatives, including the U.S. Coast Guard (USCG) and Federal Aviation Administration (FAA) completed a major effort to define and analyze the performance of a new Enhanced Loran system as a backup for the navigation and timing services provided by the NAVSTAR Global Positioning System (GPS) provided services. Each mode of transportation has defined requirements that the new Enhanced Loran must meet to be acceptable in the radionavigation mix of systems. The group developed a set of requirements for Loran maritime navigation in terms of availability, accuracy, integrity and continuity for the Harbor Entrance and Approach (HEA) requirements defined in the Federal Radionavigation Plan (FRP). This paper discusses the goals of the Loran Support Unit for Fiscal Year 2005 (FY05), and the program to support these goals. The factors related to achieving the objective of moving Differential Loran from the proof-of-concept stage to an operational status will be discussed. Also covered are the results of an initial survey of the Inner Harbor at Boston, MA, USA.

Key words: Loran, radionavigation, GPS, timing.

1 Introduction

The Loran Integrity Performance Panel (LORIPP) and Loran Accuracy Performance Panel (LORAPP) determined that an improved version of the LORAN-C system, called Enhanced LORAN, could meet the operational requirements of the HEA for maritime positioning use and the FAA-derived Required Navigation Performance of 0.3 NM (RNP 0.3). The U.S. Department of Transportation (DOT) Volpe Center completed a benefit-cost analysis covering this move, with favorable results. Both reports were completed and delivered to the Office of the U.S. Secretary of Transportation in March of 2004. At the time of this writing, the Loran community awaits a public decision regarding the future of the LORAN system.

Although a definitive direction for LORAN has not been decided, the USCG Loran Support Unit (LSU) has continued research and development into the Enhanced Loran architecture. Having completed the aforementioned reports, a transition is underway from the proof-of-concept stage to a quasi-operational status, which will promote receiver development and other LORAN research.

2 Differential LORAN

The basic concept of Differential LORAN is to provide two sets of phase corrections to improve the navigation accuracy from the current 0.25 NM level to approximately 20 meters. One set of corrections is called Additional Secondary Factors (ASFs) which are defined as the phase differences between an all seawater propagation path and the actual propagation path and are functions of the ground conductivity and terrain along the

path. These ASFs will be obtained by detailed surveys of the coverage area. In addition, there are temporal changes in the observed phase caused by changes in index of refraction along the propagation path and variations in transmitter bias. These variations will be measured at a fixed local monitor site, and communicated to users via modulation of the LORAN signal. For a detailed description of this data channel the reader is referred to Peterson et al (2004).

3 Goals

There are two main goals for FY05. The first goal is to establish Differential LORAN on a 24/7 real-time basis for selected areas of the Northeastern U.S. Previous tests were done either in post-processing or during limited time periods in which Differential LORAN data was broadcast over the air waves from the experimental transmitter at the LORAN Support Unit. While these relatively short broadcasts were useful to demonstrate that the technology was feasible, continuous broadcasting of real-time data is needed in order to refine the implementation. This has the added potential benefit of promoting receiver development.

The second goal is to develop the procedures and working knowledge necessary to establish Differential LORAN in an area. Knowledge gained from the marine and aviation surveys can be integrated in support of this goal. In addition to scientific concerns, some practical considerations may drive the final shape of the new Loran system.

4 Program

Differential LORAN is a technology that is applied to both timing and navigation applications. Consequently, two types of monitor sites have been identified: 1) Tier I sites which possess a GPS independent, highly accurate source of absolute time (within 10 ns of UTC(USNO)) facilitated by one or more atomic time standards disciplined using Two-Way Satellite Time Transfer (TWSTT), and Tier II sites which have a less accurate and possibly GPS dependent source of absolute time. Tier I and Tier II sites are nominally called “timing” and “navigation” monitor sites, respectively. The Tier I sites will support both timing and navigation users. If GPS service is lost the Tier II sites will revert to pseudorange vice absolute corrections whereby one correction will be set to zero, all others calculated as relative corrections, and the corrections will be useful to navigation users but not to timing users. The message format includes bits to notify users of the type of base station the corrections come from and whether on not the GPS time reference is available.

The Northeastern U.S is the area of the country with the highest seasonal variation in phase propagation. Planned monitor sites include: U.S. Department of Transportation Volpe National Transportation Systems Center (USDOT Volpe), to support some marine surveys in the Boston, MA area, The USCG Loran Support Unit (LSU), Wildwood, NJ, The USCG LORAN monitor site at Sandy Hook, NJ, due to its proximity to the metropolitan New York City area, and the United States Naval Observatory (USNO), where official time for the U.S. is maintained.

Boston Harbor will be the initial location for a marine survey. A navigation monitor site has been established at the Volpe Center to support surveys in the area. Once the Boston survey research is complete and as time permits, it is desired to apply the newly refined procedures to another metropolitan area such as New York.

5 Issues

Communications Network: Due to the topography of the areas surveyed, monitor sites may be placed in remote areas and at locations with varied methods of access to the Internet. This requires the establishment of an ad hoc network in which data sources can be added, removed, or moved easily. This capability requires a specialized computer network structure. A next-generation IT network for the Enhanced LORAN system is being developed at the USCG Loran Support Unit, however it is not due to become operational until FY 2007. An interim solution that will allow for real-time data broadcast is being developed at the LSU.

Monitor Site Density: The seasonal variation in phase propagation is region-dependent. Differential LORAN technology reduces the error due to this variance. However, for a given area and a given location within the area, the accuracy achieved using the correction from a monitor site degrades with distance from the site.

6. Survey Considerations

There are several factors to consider when executing a marine survey. Some of the most important ones are discussed here.

(1) Geographic Survey Boundaries: The single most basic question to answer in conducting a marine survey is: what are the boundaries of the area to be surveyed? As an example: consider the Chesapeake Bay, VA area, which is large and has many tributaries and other waterways connected to it. A decision needs to be made concerning the areas of a waterway that require Differential LORAN.

(2) Seasonal Variations: The phase of the signal from a given LORAN station and a given observation point varies temporally. When conducting a marine survey, it is necessary these temporal changes be measured at the local monitor site and that these variations in phase be taken into account in processing the survey data. Once a survey has been completed, a table of geographic points and associated nominal ASF values are calculated. Once calculated, this table or "grid" is loaded into a user receiver module. A navigation monitor site sends out the temporal corrections for the area covered by the grid. In the receiver, the temporal corrections are used to increment or decrement the base offset for the grid values as a whole. This method is effective as long as the phase variation is relatively uniform throughout the geographic region that the grid covers. It is assumed that the temporal variations in phase are constant over the coverage area of a particular monitor site. To verify that this is valid for a particular coverage area it is necessary to survey the area at multiple times during the year.

(3) Grid density: This factor is influenced by the spatial gradient of the ASF for a given area. A spatial gradient develops when there is a significant difference in the land path between a given LORAN station (LORSTA) and two points. Assuming that it is desirable to have a uniform level of accuracy for the area that a grid covers, the existence of a gradient is problematic since it means that the grid points must be closer together for the high-gradient regions of the area. Another solution is to divide the area into sub-grids of different point spacing, or simply restrict grids to cover areas where the ASF gradient is below a certain threshold. Finally the grid must be in a format amenable to receiver manufacturers.

(4) Source of Ground Truth/Geographic datum: There are two possible sources of ground truth for the ASF surveys: the USCG maritime Differential GPS system and the Wide Area Augmentation System (WAAS) operated by the FAA. DGPS is based on the North American Datum of 1983 (NAD 83) and WAAS is based on the World Geodetic System 1984 (WGS 84). Both systems have comparable accuracy. In the surveys done this far, we have logged both DGPS and WAAS data simultaneously and have compared the two sets of fixes and compared the differences to that predicted by the differences between NAD 83 and WGS 84.

7 Loran Data Channel (LDC) Test for 2005

The real-time dissemination of Differential Loran data (i.e.: moving data from multiple monitor sites to a central database and broadcasting the same data from a LORAN station) will represent a major move forward for Differential Loran, allowing more effective test of the technology and process, and will support additional research in the field. The success of this endeavor depends on proper integration of specialty software and COTS hardware.

LORSTA Seneca, NY is the planned first broadcast node in this network. Initially, observations from monitor sites at the US DOT Volpe Center at Boston, MA and USNO at Washington, D.C. will be broadcast from this station.

Communications between the monitor sites, a central server and LORSTA Seneca, NY will be crucial to the success of this endeavor. Currently, the operational network for the LORAN system is being used for the present Loran data collection efforts. There are three obstacles to using this scheme for real-time corrections. First, the architecture of the current operational network coupled with the protocols employed is not amenable to the type of data requirements for research. Second, the security policy for the operational LORAN network does not permit adding users on an ad hoc basis and with varying security assurance levels, and does not allow access from the Internet. Third, the remote possibility that a catastrophic network glitch could be caused by this research makes using the operational network an un-attractive option. For these reasons, it was decided that a network other than the operational network would be used. Due to the prohibitive cost of acquiring another research network for this specific purpose, it was decided to use the Internet for communications during this test and research phase.

LSU has undertaken the effort to determine the requirements for the next-generation Loran network, which will support Differential Loran messaging; however the planned operational phase is for FY2007. An interim, Internet-based solution is being developed at LSU to facilitate research and monitoring of the differential messages. This communications scheme will allow dissemination of real-time differential corrections.

8 Architecture of Differential LORAN Data Network

In general, Differential LORAN is being implemented for this experiment in the following way: Monitor sites (navigation or timing) are placed at strategic locations near certain waterways. The sites produce Loran observations at a specific reporting interval which are immediately sent to a central computer at LSU via the Internet. Upon arrival at LSU, the observations are logged and immediately relayed to the applicable LORSTA (initially LORSTA Seneca) for broadcast. So there are three types of nodes in the aforementioned network: monitor, central, and broadcast nodes. Only one central node (the server) exists. The location of the monitor nodes is influenced mainly by available space/real-estate, proximity to desired coverage area (for

navigation sites), and proximity to existing sources of high-quality oscillators (e.g.: cesium clocks).

9 Required Equipment

The equipment being used for this experiment is mostly commercial off-the-shelf (COTS). The nodes are connected via the Internet. The central node requires the least amount of equipment, consisting of a fast computer running connected to the Internet and running specialized software to relay the differential messages. The broadcast node requires a computer to receive the messages from the server and encode them for transmission to the standard equipment at the Loran Station. The computer at this node is also connected to a Loran receiver, and a source of absolute time. Finally, an uninterruptible power supply (UPS) will be used to prevent unnecessary loss of power. The monitor node requires a computer connected to a source of UTC and a Loran receiver. A very stable oscillator is required for a timing monitor site. A UPS is also used at this type of node.

10 Boston Harbor Survey

An initial survey of the Inner Harbor at Boston, MA, USA was conducted on July 17, 2004. Although previous marine surveys have been conducted, this survey helped bring some lingering issues to the fore. ASFs are calculated and organized by cells in a two dimensional grid of latitude and longitude. Cell size is a variable to be determined, and it may vary from port to port or even within a port. Specialized software has been developed to perform some calculations on the raw survey data. The software calculates and plots for each cell:

a. Number of samples
b. Mean
c. Standard deviation
d. Maximum difference to any adjacent cell

Figures 1 through 8 illustrate the analysis for Boston harbor. Figure 1 shows the path of the survey on a nautical chart.

Figures 2 and 3 show the number of data points per cell for cell sizes of 0.005 and 0.002 degrees respectively. Figure 4 shows the mean ASF for the 9960Y signal. The ASFs are relative or pseudo-ASFs meaning that they are all relative to the 9960M signal which has its ASF set to zero. The values are therefore the difference between the 9960Y (Carolina Beach) ASF and the 9960M (Seneca) ASF and are negative due the larger portion of land in the path from Seneca to Boston. Figures 5 and 6 show the maximum absolute value of the difference in ASF to any of the eight adjacent cells for cell sizes of 0.002 and 0.005 degrees respectively. Figures 7 and 8 show the standard deviation of ASF for cell sizes of 0.002 and 0.005 degrees respectively. The intent is to determine whether enough data was collected, the data collected is valid, and that the cell density is sufficient such that variations within a cell or between adjacent cells are adequately bounded.

Figure 1. Path of Boston Harbor Survey

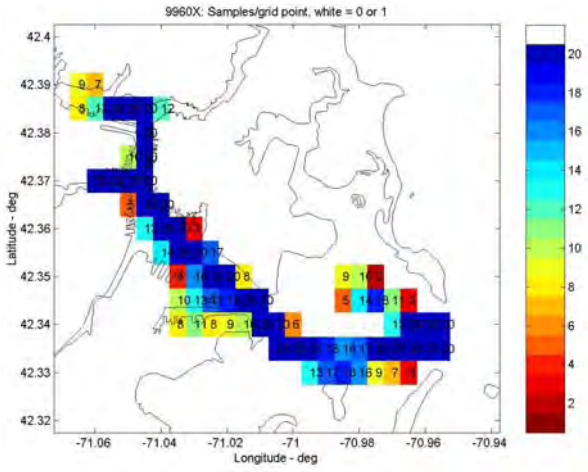


Figure 2. Number of Data Points per Grid Cell (Cell Size 0.005 degrees)

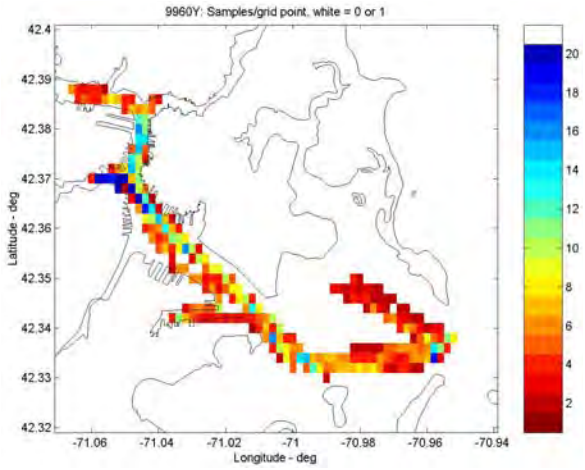


Figure 3. Number of Data Points per Grid Cell (Cell Size 0.002 degrees)

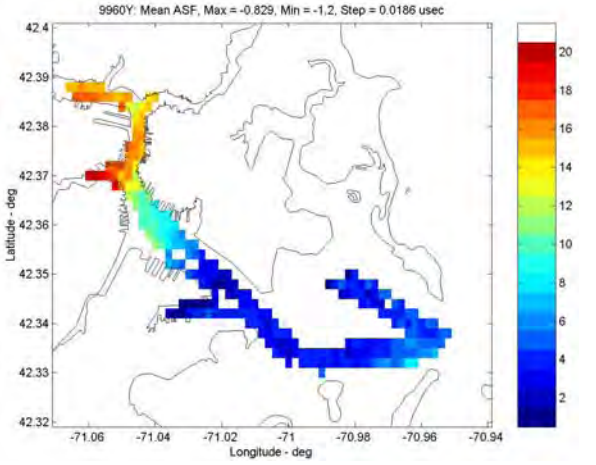


Figure 4. Average ASF for 9960Y (Carolina Beach) Signal

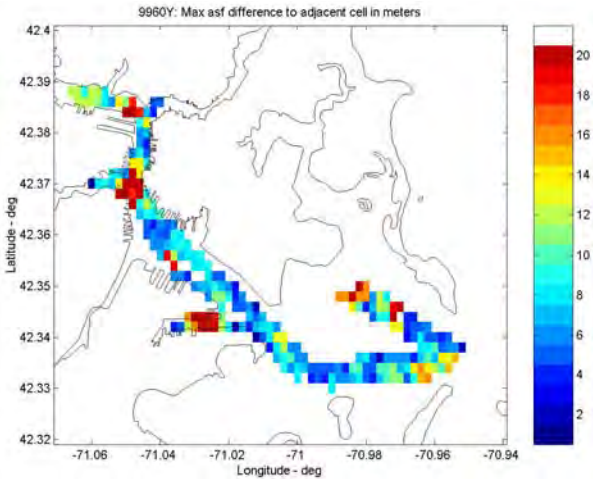


Figure 5. Difference in ASF Between Adjacent Grid Cells (cell size 0.002 degrees)

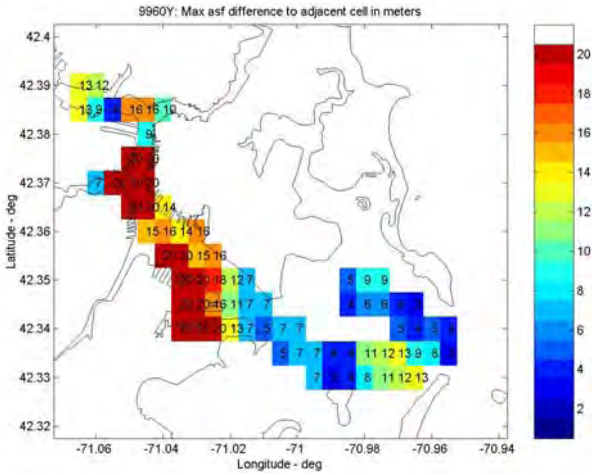


Figure 6. Difference in ASF Between Adjacent Grid Cells (cell size 0.005 degrees)

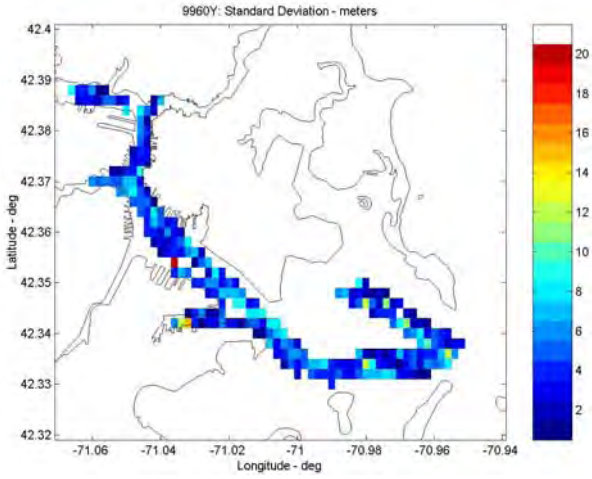


Figure 7. Standard Deviation of ASF By Cell (cell size 0.002 degrees)

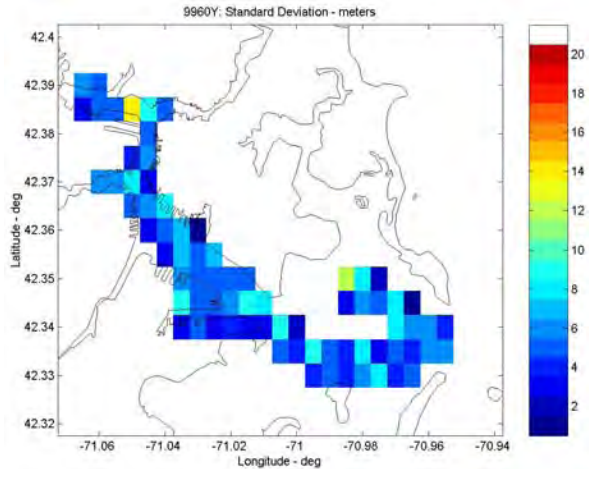


Figure 8. Standard Deviation of ASF By Cell (cell size 0.005 degrees)

11 DGPS vs. WAAS

Figure 9 shows the comparison of DGPS and WAAS positions for the survey. The mean East difference is 0.05 m with a standard deviation 0.24 m and the mean North difference is -1.05 m and with a standard deviation 0.26 m. The values predicted by HTDP.exe from NGS Geodetic Tool Kit (www.ngs.noaa.gov) are 0.18 m East and -1.01 m North.

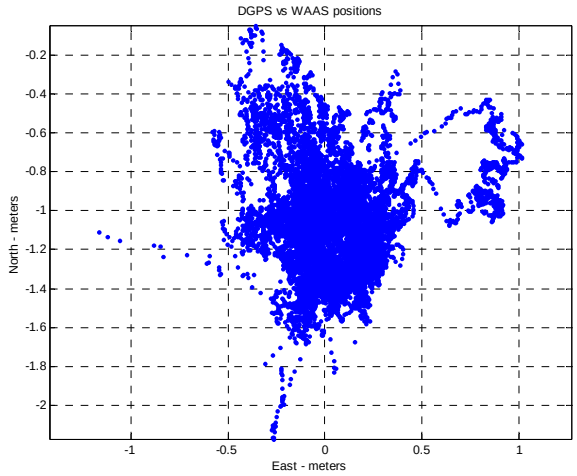


Figure 9. Difference between GPS and WAAS positions for Ground Truth

12 Conclusions and Recommendations

We have presented an outline of the effort to take differential LORAN from the proof of concept stage to an operational system. The main issues discussed include the communications network necessary to broadcast real time differential data and the methodology of conducting and analysing ASF surveys.

Acknowledgements

This work was supported by the US Federal Aviation Administration LORAN evaluation program. The program manager is Mitchell Narins.

References

Peterson B, Dykstra K, Carroll K, Hawes A, and Swaszek P (2004) ***Differential Loran-C***, Proceedings of GNSS 2004, Copenhegan, Denmark, May 2004.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 63-69

Journal of Global Positioning Systems
2004
© CPGPS

An Open GNSS Receiver Platform Architecture

Frank Engel[†], Peter Mumford[‡], Kevin Parkinson[‡], Chris Rizos[‡], Gernot Heiser[†]

[†] Embedded Real-Time and Operating Systems Program (ERTOS), National ICT Australia (NICTA)
e-mail: {frank.engel, gernot.heiser}@nicta.com.au Tel: + 61 2 9358 7339; Fax: +61 2 9385 7942

[‡] Satellite Navigation and Positioning Group (SNAP), The University of New South Wales (UNSW)
e-mail: {p.mumford, c.rizos}@unsw.edu.au, kevin@dynamics.co.nz Tel: + 61 2 9358 4182; Fax: +61 2 9313 7493

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. In this article we present the concept of a FPGA-based GPS receiver architecture with the aim of providing a framework for investigating new receiver architectures for current and upcoming GNSS standards. This development system facilitates researchers to prove new receiver concepts using real signals, which nowadays can only be simulated using tools such as Matlab. One will be able to work with the satellites as soon as they are operational, rather than having to wait for the availability of commercial products. The system allows individual development of signal processing solutions for base-band processing. A soft-core processor implements higher layer services that provide data to the user.

Key words: GPS Receiver, GNSS, FPGA, Soft-Core Processor

1 Introduction

The *Global Positioning System (GPS)* is a satellite-based, all-weather, round-the-clock navigation system developed by the U.S. Department of Defence. Since the early 1980s GPS has been available to the civilian community and is the most effective navigation and positioning system ever developed. Once a tool for experts and an exotic gadget for nerds, it is rapidly becoming a commodity. Location-aware devices are an essential part of many ubiquitous computing scenarios, and the recent mandate by the US FCC (FCC, 2004) that all emergency calls from mobile phones must be locatable has increased the pressure on improving the technology. Satellite-navigation receivers will in the future be embedded in a large number of devices of widely different uses.

1.1 Background

Although GPS is the only fully operational 'first generation' *Global Navigation Satellite System (GNSS)*, since the mid-1980s signals from the Russian Federation's *Glonass* system have also been available (Glonass, 2004). However, since 1995 Glonass has been starved of funding and the constellation has progressively degraded. Russia's President has announced that the full Glonass system will again be operational by 2008. In January 2004 India signed a Memorandum of Understanding with Russia to collaborate on the Glonass refurbishment.

Europe is currently developing its own 'second generation' GNSS, known as *Galileo* (Galileo, 2004), for deployment between 2006 and 2008. In 2003 China signed onto the Galileo program.

There are a number of additional satellite systems in operation today, or to be soon available that complement GPS, Glonass and Galileo. These systems are generically referred to as *Space Based Augmentation Systems (SBAS)* (SBAS, 2004) because they transmit extra signals that certain receivers can decode and use (together with the global GNSS signals) to achieve improved positioning performance. To augment GPS for aviation users, the U.S. and Europe have deployed the *Wide Area Augmentation System (WAAS)* and the *European Geostationary Navigation Overlay System (EGNOS)* respectively (WAAS/EGNOS, 2004). Japan, India and China have indicated that they too will develop regional SBAS by the end of this decade. The most advanced plans are Japan's *Quasi-Zenith Satellite System (QZSS)* (Petrovski, 2003), a constellation of 3 to 7 satellites that will broadcast GPS (and ultimately Galileo) signals from orbits optimised for East Asian.

In the near to medium-term future the market for satellite navigation technology is expected to continue to experience major growth (Canalys, 2004). The greater availabil-

ity of signals from these upcoming systems and the reduction in costs of GNSS receivers will support new application fields like indoor positioning for consumer devices. Besides the positioning data, these systems will support extra, low data rate channels. These channels can be used to receive personalised information similar to a pager system, which prepares the ground for new services offering this information.

1.2 Project Objectives

The aim of this project is to develop a platform for investigating new algorithms and systems based on current and upcoming GNSS receiver standards. This platform could serve as a generic base for GNSS research that supports research into signal processing issues, protocols and early implementation of emerging standards:

- A reference and testbed environment to investigate interoperability and standard compliance.
- An IP block readily accessible to industry partners for development of GNSS and GNSS-enhanced devices.
- Considering the wide range of next generation GNSS signals a solution is required that support flexible hardware and software design. As an example, a system that can incorporate extra processing power or tracking channels, as the need arises.

The article is structured as follows. Section 2 introduces the requirements of such a platform concept. Section 3 explains project stages in more detail. Section 4 discusses future investigations into improved GNSS receivers and applications/products. Intermediate project results and a conclusion finalises this article.

2 Platform Reqirements

2.1 Components

In general a GNSS receiver consists of three major components:

- *RF front-end to receive GNSS signals and converting into baseband*: Since satellite signals are very weak, the development of front-end components still requires custom circuit design.
- *Baseband processing to generate raw / preprocessed data from received signals*: It comprises strong digital signal processing either executed by custom hardware or a digital signal processor (DSP).
- *Control processing to manage receiver activity and post-process data*: Besides converting raw data into e.g. position velocity and time, tasks like maintaining receiver status (e.g. tracking satellites) is usually executed by a micro controller (µC) running special firmware.

According to this structure, a generic receiver platform should support interfaces to various RF front-ends, signal processing hardware, and (µC supported) software layers for control processing and applications.

2.2 Configurability

State of the art field programmable gate array devices (FPGAs) such as Altera's Stratix series provide high computational capacities of up several million logic element equivalents (Stratix, 2004). Besides pure configurable logic they also provide on-chip memory and DSP capabilities. Thus they are very well suited for complete system-on-chip solutions (SoC) without the need of silicon fabrication. Furthermore, FPGA vendors offer CPU cores integrated into the FPGA either as hard-core or soft-core. While hard-core processors are optimally integrated into the FPGA fabric, soft-core processors allow a fine grain tailoring towards the design requirements by introducing caches and specialised instructions etc.

Following this strategy, baseband processing could be implemented by FPGA logic, supported by hardware macros, while control processing could be implemented by a soft-core processor such as Altera's NiosII. Thus an FPGA device serves as the central processing element on this reconfigurable platform. It connects to RF front-ends as well as different communications interfaces (e.g. TCP/IP, RS232, USB).

3 Approach

The general approach is to break the development process into three stages. These stages relate to the migration of the hardware platform from a Mitel Superstar receiver towards the FPGA solution on a custom designed circuit board. The stages are identified as:

- Stage 1: SuperStar-NiosII;
- Stage 2: GP2021_Emulator-NiosII and;
- Stage 3: Custom GNSS Platform.

The Mitel GPS Architect (sometimes referred to as Orion) has become some sort of standard GPS software suite for the Zarlink chipsets. It was available as a development kit with a licence allowing royalty free distribution of compiled binaries (GPSArch, 1997). The SNAP Group purchased a kit in 1999. The GPS Architect has been selected for testing purposes for the reasons mentioned above. Altera's products 'Quartus' and the 'NiosII IDE' are the primary tools used for programming the FPGA and code generation (Altera, 2004).

3.1 Mitel GPS Architect

The original Mitel GPS Architect software was designed to run on an ARM60 processor, with a Mitel (now Zarlink) GP2021 (GP2021, 2001) baseband processor and GP2010 or GP2015 RF front end (GP2015, 2002). Version 6.17 of the GPS Architect code was ported to the Sigtec MG5001 GPS receiver in 2003. This receiver has a Zarlink GP4020 chip with an ARM7 core and GP2015 RF front end (GP4020, 2002). The GP4020 has the same correlator design as the GP2021. This version of the GPS Architect software was selected for porting to the NiosII embedded processor.

There are significant differences between the NiosII and ARM7 architectures and instruction sets, as well as differences in the development tools and embedded libraries. Addressing these differences, getting familiar with the new tools and creating interfaces to hardware such as the baseband signal processing (or correlator) block, constitutes the bulk of the porting effort. Although, only a few changes to the application software are required, system initialisation, the Real Time Operating System (RTOS), chipset interfacing and serial port communication require significant changes to the code base since those parts are highly architecture related.

3.2 Altera Tools

Altera provides Quartus, a tool for the generation of FPGA designs and the NiosII IDE for the generation of code to run on the NiosII soft-core processor. Quartus contains the 'SOPC Builder' used for generating NiosII core FPGA design files (Altera, 2004). The NiosII development tools provide the Hardware Abstraction Layer (HAL) Application Programming Interface (API) that incorporates the Newlib ANSI C embedded library. In addition, the HAL provides some UNIX-style functions. This provides a higher level of abstraction than the ARM tools, facilitating the removal of the GPS Architect's booting, initialization and interrupt assembler code, as well as most of the serial port communication code. Altera also provide a development board with a Stratix FPGA device and associated memory and I/O facilities called the 'NiosII Development Board - Stratix Edition' (EvalBoard, 2004). This is a very convenient platform to get projects up and running quickly. It is used for Stage 1, and Stage 2.

3.3 Stage 1: SuperStar-NiosII

The first stage involves porting the GPS Architect software to a NiosII processor connected to an existing GP2021/GP2015 hardware platform. To do this, a "hack" is performed into a Superstar receiver to effectively replace the ARM60 with a NiosII processor. This is achieved by removing the ARM60 and memory chips and connecting the GP2021 address, data and some control lines to a header on the NiosII Development Board.

RT-Exec Code

The GPS Architect software has a minimal RTOS closely integrated with the rest of the code. The task-switching core of this RTOS is coded in assembler, and this is completely rewritten for the NiosII processor.

SuperStar-Interface

An interface to read and write to the GP2021 baseband signal processor and other registers is required, as these are not memory-mapped registers as in the GP4020 case. This 'glue logic' is realised in the FPGA as a small VHDL module connected between dedicated parallel I/O ports on the NiosII and the data, address and control lines of the GP2021. Read and write functions provided by the HAL API are used to send and receive serial data via the NiosII UART's.

3.4 Stage 2: GP2021_Emulator-NiosII

The second stage is to replace the external GP2021 baseband signal processor by an FPGA baseband processor in the Stratix device. The FPGA baseband processor is designed as a GP2021 emulator such that the code developed in Stage 1 can be reused with changes required to the SuperStar-Interface only. For this stage, only the RF front-end chip GP2015 of the SuperStar receiver is used. The sampled signals (sign and magnitude) from the GP2015 are connected to FPGA pins and these signals are routed internally to the FPGA baseband processor.

GP2021 Emulator

In order to emulate the GP2021 silicon inside the FPGA, a separate baseband processor is being developed to provide all of the functions as expected by the ported software. This block will replace the GP2021 baseband processing functions using internal FPGA connections to the NiosII core via an on-chip Avalon Bus. These functions,

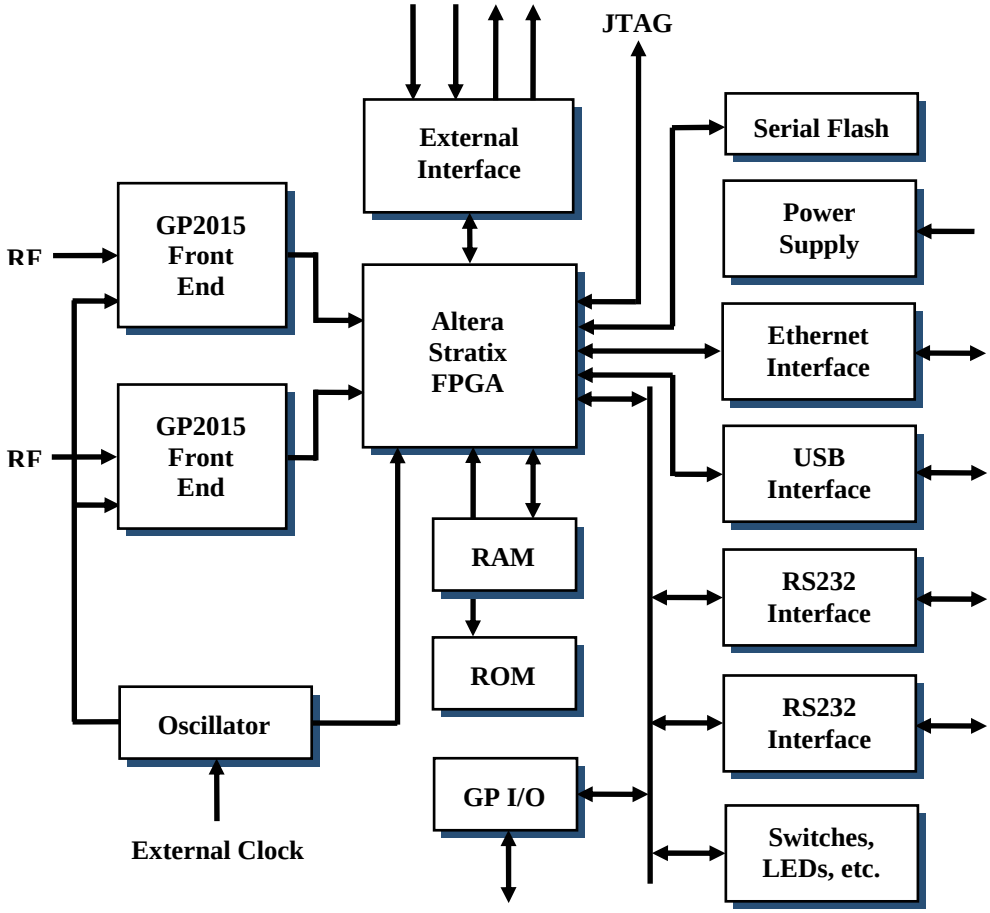


Fig. 1 Block Diagram of the Platform

such as the multi-phase clock generator, time base generator, numerically controlled oscillators, correlator modules etc. are developed in HDL to initially have identical functionality to the GP2021. The system clock and RF signal connections from the GP2015 GPS RF front-end chip to the baseband processor are still required from the SuperStar receiver board.

Obviously not all of the GP2021 functions need to be emulated, given that the chip was originally designed as a multi-purpose chip to support other microprocessor interface schemes, and also Glonass signals. Improvements in the core baseband processing functions of the GP2021 will come later as the design diverges from the ported software in Stage 3 of the development. While Glonass signal processing capability is possible within the FPGA, support for this may be considered in the future because at the moment the different RF front-end structure is not easily accessible.

3.5 Stage 3: Custom GNSS Platform

In Stage 3, SuperStar receiver and NiosII Development Board are replaced with a custom designed board shown in Figure 1. This board has all the regular features such as I/O, memory, status LED's and switches, an Altera Stratix device, headers for expansion and multiple RF front-ends. In fact it has many similarities to the NiosII Development Board. This will help facilitate the move from Stage 2.

The board has two RS232 serial ports along with Ethernet 10/100 and a USB port. The UART components for the serial ports are instantiated into the FPGA as peripherals connected to the on-chip Avalon Bus, while both the Ethernet and USB functions are provided by external chips connected to the FPGA to become peripherals accessible in the NiosII address space.

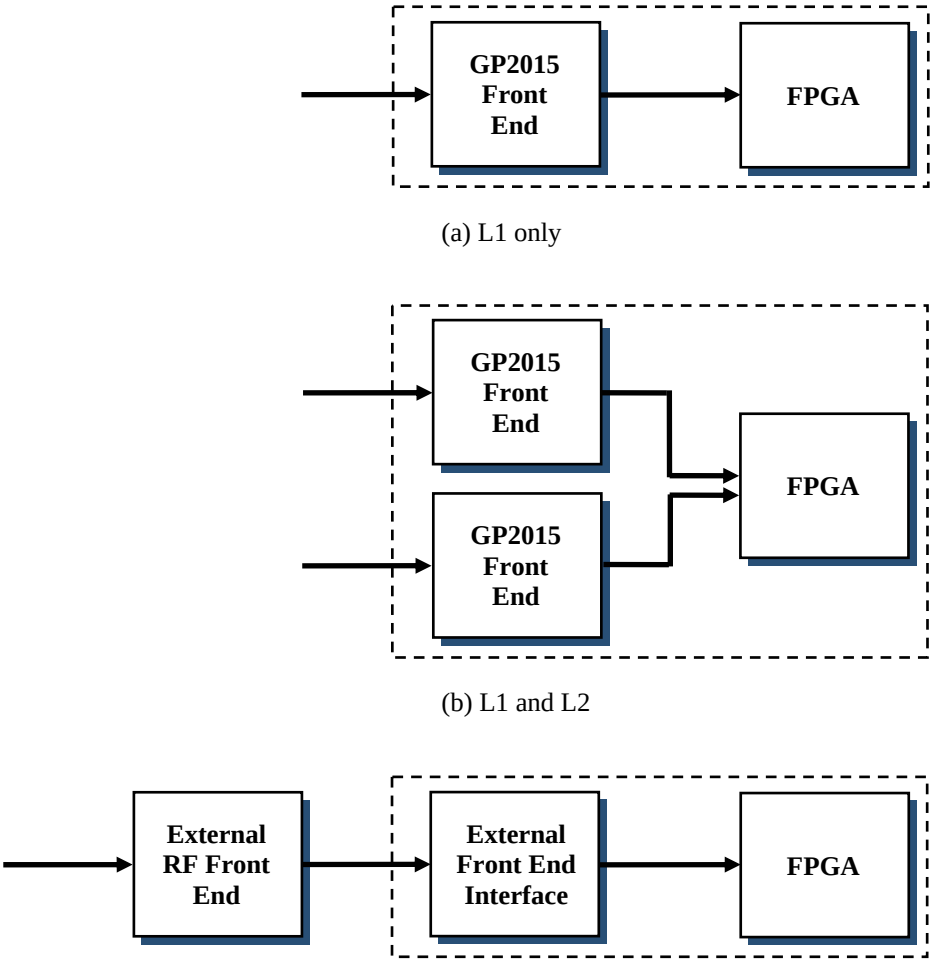


Fig. 2 Signal Path Options

Providing flexibility in RF front-end choice facilitates access to the new signals as well as new designs as they emerge. The custom board incorporates two GP2015 RF front-ends, and a special header to accommodate an external RF front-end or any other extension board. This arrangement potentially allows simultaneous reception of multiple signals at different frequencies. Figure 2 illustrates some possible signal paths.

The clock signals for the on-board and external RF front-end chips are driven from either a common temperature compensated oscillator or an external signal source via a small coaxial connector. This allows the use of an optional extra stable frequency source where required. In both cases the clock signal is also connected to the FPGA phase locked loop circuitry to allow multi-phase clock generation within the device.

4 Further Development

This project has the fundamental objectives of proving the concept and developing a platform for future investigations. Due to the configurable nature of FPGAs, there are many possibilities for investigation including:

- *Baseband signal processing block design:* This design is not restricted to a GP2021 emulator, as many design enhancements are possible. Ideas include; improved tracking in weak signal and multi-path environments.
- *DSP search engine*: The Stratix FPGA device has DSP blocks, which could be used to create a search engine for signal acquisition and tracking, particularly for weak signals.

- *Develop the GPS Architect software for better performance or specific functionality:* The SNAP Group has developed improvements and additions such as better tracking algorithms, which could be implemented into the standard GPS Architect software (Lee, 2003 and Mumford, 2003 and Satirapod, 2003). A licence is required to develop applications using the GPS Architect software suite. However, binaries developed by a licensed user can be distributed without royalties. There may be a way of distributing core functionality as a binary API, allowing the development of particular critical sections such as the tracking loops.

- *Replace the GPS Architect with another software suite:* for control, tracking and navigation solution functions on the NiosII processor. The OpenGPS project (OpenGPS, 2004) is a possible candidate here.

- *Raw data collection and packaging for PC based soft receiver processing:* The sign and magnitude data from the RF front-end can be collected and packed up for transmission via Ethernet to a host PC or stored onto USB media.

- *Signal interference detection:* The custom board is ideal for developing signal interference and jamming detectors.

5 Intermediate Results

Stage 1 of this project is almost finished. All architecture dependant software parts (See Sec.3.3) are ported. Tests show the system is capable of tracking of at least 9 satellites and able to run more than 16 hours without failing. This means a reference design is available which will be used to evaluate accuracy of the baseband processor emulation, the main task of Stage 2.

As for Stage 2, the development and testing of the initial designs for the baseband signal processor to emulate the functionality of the GP2021 have begun. So far a single channel correlator is under development. First simulation results and size estimates show that the design will fit into a Stratix device and meet timing requirements. The DSP blocks available within the Stratix device could be optionally used and investigations are under way to evaluate how their use may contribute to improve performance.

The custom GNSS platform board is in the early stages of development.

6 Conclusion

In this paper an approach of designing an FPGA based GNSS receiver platform is described. The project has been broken down into three stages so that particular problems can be isolated and addressed at each stage. In Stage 1 the focus is on porting the GPS Architect software to the NiosII processor architecture. Once the GPS Architect is running smoothly on the NiosII Development Board with a GP2021 baseband processor and a GP2015 RF front-end attached, Stage 2 can be tackled. This stage involves replacing the GP2021 with an FPGA baseband signal processor emulating the main features of the GP2021. Finally, the developed design can be ported to the platform, a FPGA based prototyping board.

The outcome of this project will be a development platform for investigations into GNSS receiver designs for current and future signals. It is the intention of this project to foster research into improvements of existing GPS receiver solutions, new GNSS signals and to help realising product ideas by providing a sophisticated hardware/software infrastructure.

References

Altera (2004): ***Altera Inc. – Homepage.*** Altera Inc. http://www.altera.com .

Canalys (2004): ***EMEA mobile GPS navigation market races ahead.*** GPS market analysis, http://www.canalys.com/pr/2004/r2004093.htm .

EGNOS (2004): ***European Geostationary Navigation Overlay System (EGNOS).*** European Space Agency (ESA), http://esamultimedia.esa.int/docs/egnos/estb/egnos_pro.htm .

EvalBoard (2004): ***NiosII Development Kit - Stratix Edition.*** Altera Inc., http://www.altera.com/products/devkits/altera/kit-nios_1S10.html .

FCC (2004): ***Wireless Enhanced 911 Rules (E911).*** US Federal Communications Commission (FCC), http:// www.fcc.gov/911/enhanced/ .

Galileo (2004): ***Galileo.*** European Space Agency (ESA), http://www.esa.int/esaNA/galileo.html .

Glonass (2004): ***GLONASS Satellite Navigation System.*** Russian Federation Ministry of Defence, http://www.glonass-center.ru .

GP2021 (2001): ***GP2021 GPS 12-Channel Correlator.*** Zarlink Semiconductor Inc., Data Sheet (DS4077, Issue 3.2), April 2001.

GP2015 (2002): ***GP2015 Miniature GPS Receiver RF Front End.*** Zarlink Semiconductor Inc., Data Sheet (DS4374, Issue 3.1), Feb. 2002.

GP4020 (2002): ***GP4020 GPS Receiver Baseband Processor.*** Zarlink Semiconductor Inc., Data Sheet (DS5134, Issue 4.4), May 2002.

GPSArch (1997): ***GPS Architect - 12 Channel GPS Development System.*** Zarlink Semiconductor Inc., Data Sheet (DS4605, Issue 2.5), March 1997.

Lee H.K.; Rizos C.; Jee, I.G. (2003): ***Design and analysis of DGPS filters with consistent error covariance information***. In Proceedings of 6th Intl. Symposium. on Satellite Navigation Technology Including Mobile Positioning and Location Services, Melbourne, Australia, 22-25 July 2003, paper 47.

Mumford P. (2003) ***Timing characteristics of the 1PPS output pulse of three GPS receivers***. In Proceedings of 6th Intl. Symposium. on Satellite Navigation Technology Including Mobile Positioning and Location Services, Melbourne, Australia, 22-25 July 2003, paper 45.

OpenGPS (2004): ***OpenGPS***, http://www.gfz-potsdam.de/pb1/staff/gbeyerle/opengps/ .

Petrovski I. (2003): ***QZSS - Japan's New Integrated Communication and Positioning Service for Mobile Users.*** GPSWorld, 14(6):24-29.

SBAS (2004): ***Satellite Base Augmentation System (SBAS),*** Eurocontrol SBAS Project, http://www.eurocontrol.fr/ projects/sbas/ .

Satirapod C.; Khoonphool R.; Rizos C. (2003): ***Multipath mitigation of permanent GPS stations using wavelets.*** Proceedings of 2003 Intl. Symposium. on GPS/GNSS, Tokyo, Japan, 15-18 Nov. 2003, 133-139.

WAAS (2004) ***Wide Area Augmentation System (WAAS),*** US Federal Aviation Administration (FAA), http://gps.faa.gov/Programs/WAAS/waas.htm .

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 70-78

Journal of Global Positioning Systems
2004
© CPGPS

Analysis of Biases Influencing Successful Rover Positioning with GNSS-Network RTK

Hans-Jürgen Euler

Leica Geosystems AG, Heinrich-Wild-Strasse, CH-9435 Heerbrugg (Switzerland)
e-mail: Hans-Juergen.Euler@leica-geosystems.com; Tel: +41(71)727 3388: Fax: +41(71)726 5388

Stephan Seeger

Leica Geosystems AG, Heinrich-Wild-Strasse, CH-9435 Heerbrugg (Switzerland)
e-mail: Stephan.Seeger@leica-geosystems.com; Tel: +41(71)727 3863; Fax: +41(71)726 5863

Frank Takac

Leica Geosystems AG, Heinrich-Wild-Strasse, CH-9435 Heerbrugg (Switzerland)
e-mail: Frank.Takac@leica-geosystems.com; Tel:+41(71)727 4476 ; Fax: +41(71)726 6476

Received: 15 November 2004 / Accepted: 16 February 2005

Abstract. Using the Master-Auxiliary concept, described in Euler et al. (2001), Euler and Zebhauser (2003) investigated the feasibility and benefits of standardized network corrections for rover applications. The analysis, focused primarily in the measurement domain, demonstrated that double difference phase errors could be significantly reduced using standardized network corrections. Extended research investigated the potential of standardized network RTK messages for rover applications in the position domain (Euler et al, 2004-I). The results of baseline processing demonstrated effective, reliable and homogeneous ambiguity resolution performance for long baselines (>50km) and short observation periods (>45 sec). In general horizontal and vertical position accuracy also improved with the use of network corrections. This paper concentrates on the impact of wrongly determined integers within the reference station network on RTK performance. A theoretical study using an idealized network of reference stations is complemented by an empirical analysis of adding incorrect L1 and L2 ambiguities to the observations of a real network. In addition, the benefits of using network RTK corrections for a small sized network in Asia during a period of high ionospheric activity is also demonstrated.

Key words: Master-Auxiliary concept, dispersive errors, non-dispersive errors, approximation, influence of wrong ambiguities.

1 Introduction

Standardization of RTCM SC104 network RTK messages is still in progress. In the absence of a standard, this paper uses the Master-Auxiliary concept (MAC) as described in Euler et al. (2001) to analyse the effect of various biases on network RTK positioning performance. MAC closely resembles the format adopted by the RTCM working group as the basis for network RTK messages.

MAC uses so-called dispersive and non-dispersive phase correction differences to compress network RTK information without the need for standardized correction models. To understand how MAC compresses this information consider the single difference L1 phase equation $\Delta\Phi^{j}_{km,1}$ for stations *k (the reference)* and *m (the auxiliary)* and a satellite *j*.

$$\Delta\Phi^{j}_{km,1}(t) = \Delta s^{j}_{km}(t) + \Delta\delta r^{j}_{km}(t) + c\cdot\Delta dt_{km,1}(t) + \Delta T^{j}_{km}(t) - \frac{\Delta I^{j}_{km}(t)}{f_1^{2}} + \frac{c}{f_1}\cdot\Delta N^{j}_{km,1} + \Delta\varepsilon_1 \quad (1)$$

where

Δs^{j}_{km} geometric range term including antenna phase centre variations which have been applied by the network processing software.

$\Delta\delta r^{j}_{km}$ broadcast orbit error.

Δdt_{km} receiver clock error.

ΔT_{km}^{j} tropospheric refraction error.

ΔI_{km}^{j} frequency dependent ionospheric delay.

ΔN_{km}^{j} frequency dependent integer ambiguity.

$\Delta\varepsilon$ frequency dependent random measurement error.

t epoch.

c speed of light.

f_1 frequency of L1.

Replacing the index of the frequency dependent terms with '2' yields an analogous equation for the L2 single difference phase. Reducing Formula 1 by the slope distance, receiver clock error and the ambiguity term yields the ambiguity-levelled correction difference $\delta\Delta\Phi_{km,1}^{j}$

$$\delta\Delta\Phi_{km,1}^{j}(t) = \Delta s_{km}^{j}(t) - \Delta\Phi_{km,1}^{j}(t) + c\cdot\Delta dt_{km,1}(t) + \frac{c}{f_1}\cdot\Delta N_{km,1}^{j} \quad (2)$$

The correction difference described in Formula 2 is separated into a dispersive component, consisting mainly of ionospheric refraction, and a non-dispersive component consisting primarily of tropospheric refraction and orbit errors in order to reduce the amount of data transmitted to the rover. The equations for the dispersive and non-dispersive components are given in Formula 3 and Formula 4 respectively in units of meters.

$$\delta\Delta\Phi_{km}^{j,\ disp} = \frac{f_2^{\,2}}{f_2^{\,2} - f_1^{\,2}}\delta\Delta\Phi_{km,1}^{j} - \frac{f_2^{\,2}}{f_2^{\,2} - f_1^{\,2}}\delta\Delta\Phi_{km,2}^{j} \quad (3)$$

$$\delta\Delta\Phi_{km}^{j,\ non-disp} = \frac{f_1^{\,2}}{f_1^{\,2} - f_2^{\,2}}\delta\Delta\Phi_{km,1}^{j} - \frac{f_2^{\,2}}{f_1^{\,2} - f_2^{\,2}}\delta\Delta\Phi_{km,2}^{j} \quad (4)$$

This alternate representation of the correction differences has some specific benefits. Unlike the correction differences described in Formula 2, changes in the dispersive and non-dispersive components vary at different rates. Non-dispersive errors change slowly over time, while dispersive errors vary more rapidly, especially in times of high ionospheric activity. Therefore, the throughput of the data-link can be maximised by optimising the individual transmission rates of the dispersive and non-dispersive observables. In addition to the correction differences, the raw carrier phase observations for the master reference station, described via RTCM v3.0 standard messages 1003 or 1004 (RTCM 2004), must also be streamed to the rover.

Using the phase data of the master station and the correction differences, the rover can re-assemble and apply the raw phase information of the auxiliary stations in conventional baseline processing schemes. Alternatively, optimal correction differences can be approximated for any position in the network and used to improve the positioning performance of the rover. As with other network RTK methods that model dispersive and non-dispersive errors (e.g. VRS), MAC requires the correction differences to be related to a common integer ambiguity level (see Euler et al., 2001). An incorrectly determined L1 and/or L2 single difference ambiguity between the master and an auxiliary station will eventually manifest itself in the position solution. The effect of wrong ambiguities on network RTK positioning is the focus of the next section.

2 The Influence of Incorrect Ambiguitites on Correction Differences and the Position Solution

2.1 Impact of Wrong Ambiguities on Dispersive and Non-Dispersive Corrections

Tab. 1 and Tab. 2 show how an incorrect L1 and/or L2 single difference ambiguity affects the dispersive and non-dispersive correction differences described in Formulas 3 and 4 respectively. For simplicity, the magnitude of the ambiguity error is restricted to ±1 cycle.

Tab. 1. Impact of a wrong L1 (ΔN_1) and/or L2 (ΔN_2) single difference ambiguity on the dispersive correction difference (in units of L1 cycles).

ΔN_1 \ ΔN_2	0	+1	−1
0	0	≈ 1.98	≈ −1.98
+1	≈ −1.54	≈ 0.44	≈ −3.53
−1	≈ 1.54	≈ 3.53	≈ −0.44

The impact of wrong ambiguities on the correction difference observables depends on the combination of the L1 and L2 errors. For example, in the dispersive case (Tab. 1) a maximum error of ±3.53 L1 cycles occurs when the incorrect L1 and L2 ambiguities are of equal magnitude but opposite sign. Similarly in the non-dispersive case (Tab. 2), a maximum error of ±4.53 L1 cycles also occurs when the L1 and L2 ambiguity errors are of equal magnitude but opposite sign. The magnitude of the ambiguity error is also amplified in the correction differences when only a single L1 or L2 bias is present. In these cases, the amplification factor is in the order of approximately 2. On the contrary, a reduction in the magnitude of the correction difference errors results when the single-difference wide lane ambiguity is correct; that

is, if the L1 and L2 ambiguity errors are of equal magnitude and sign.

Tab. 2. Impact of a wrong L1 (ΔN_1) and/or L2 (ΔN_2) single difference ambiguity on the non-dispersive correction difference (in units of L1 cycles).

ΔN_1 \ ΔN_2	0	+1	−1
0	0	≈ −1.98	≈ 1.98
+1	≈ 2.55	≈ 0.56	≈ 4.53
−1	≈ −2.55	≈ −4.53	≈ −0.56

Normally, optimal correction differences are approximated for the rover's position. The effect of wrong ambiguities on approximated correction differences will depend on the algorithm used to model the dispersive and non-dispersive corrections in the area bounded by the reference stations. The next section investigates the propagation of correction difference biases for a two-dimensional (2-D) linear approximation.

2.2 Approximation of Correction Differences

Numerous algorithms can be employed for approximating optimal network corrections at the rover. For example, Euler et al. (2003) and Euler et al. (2004-I) compare the effectiveness of a distance weighted approximation technique with a 2-D linear plane represented by

$$b_L(x, y) = a_0 + a_1 x + a_2 y \tag{5}$$

where

b_L linear surface.
a_i coefficients defining the plane.
x, y coordinates of the approximated point.

The case of a quadratic approximation is detailed in Euler et al. (2004-II). Only the linear approximation, represented by Formula 5, will be used to approximate correction differences in this paper. To measure how wrong ambiguities at a reference station propagate to the rover in the linear case consider the hypothetical network of 6 reference stations as shown in Fig. 1.

Reference stations P_1, P_3 and P_5 lie at the vertices of an equilateral triangle Δ and stations P_2, P_4, and P_6 lie at the midpoints of Δ. Due to the symmetry of the network there are only two scenarios that have to be considered in the analysis: an error introduced at one of the reference stations located at the vertices of Δ (e.g. P_1) and an error introduced at one of the reference stations located at the midpoints of Δ (e.g. P_2).

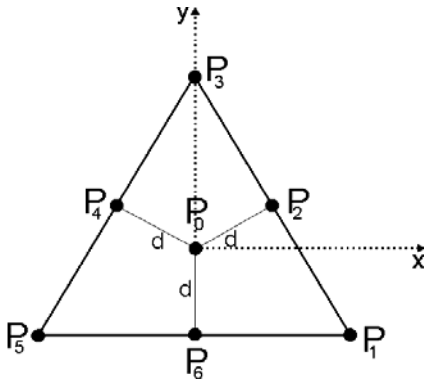


Fig. 1. Hypothetical network of 6 reference stations and one rover station located at the centroid of the figure.

Let the station coordinates be $P_i = (x_i, y_i)$ where $i = 1,\ldots,6$ for the reference stations and $i = 0$ for the rover station. For simplicity, let $P_0 = (0,0)$. If d is the distance from P_0 to P_2, P_4 and P_6, respectively, then the plane coordinates of the reference stations in Fig. 1 are

$$\begin{aligned} P_1 &= \left(\sqrt{3}, -1\right)d, \quad P_2 = \left(\sqrt{3}/2, 1/2\right)d, \quad P_3 = (0,2)d \\ P_4 &= \left(-\sqrt{3}/2, 1/2\right)d, \quad P_5 = \left(-\sqrt{3}, -1\right)d, \quad P_6 = (0,-1)d. \end{aligned} \tag{6}$$

Let the value given at reference station i (e.g. an L1 or L2 phase correction) be $b_i \in \Re$. We want to approximate the values b_i at (x_i, y_i) by a function $b(x, y)$ so that

$$b(x_i, y_i) = b_i + \varepsilon_i \tag{7}$$

where ε_i is the approximation error. The linear approximation given in Formula 5 can be rewritten as

$$b_L(x, y) = (1, x, y) \cdot (a_0, a_1, a_2)^t \tag{8}$$

Expanding the polynomial Formula 7 for all i results for $b(x, y) = b_L(x, y)$ in

$$M_L \cdot (a_0, a_1, a_2)^t = b + \varepsilon_L \tag{9}$$

where

$$M_L = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ 1 & x_4 & y_4 \\ 1 & x_5 & y_5 \\ 1 & x_6 & y_6 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{pmatrix}, \varepsilon_L = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix} \tag{10}$$

The polynomial coefficients $a_L \equiv (a_0, a_1, a_2)^t$ that minimize $\|\varepsilon_L\|^2$ are given by

$$a_L = \left(M_L^t M_L\right)^{-1} M_L^t b \,. \tag{11}$$

Substituting (x_i, y_i) from Formula 6 into M_L yields the following expression for $\left(M_L^t M_L\right)^{-1} M_L^t$:

$$\left(M_L^t M_L\right)^{-1} M_L^t = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{2\sqrt{3}}{15d} & \frac{\sqrt{3}}{15d} & 0 & -\frac{\sqrt{3}}{15d} & -\frac{2\sqrt{3}}{15d} & 0 \\ -\frac{2}{15d} & \frac{1}{15d} & \frac{4}{15d} & \frac{1}{15d} & -\frac{2}{15d} & -\frac{2}{15d} \end{pmatrix} \tag{12}$$

Assume that the values b_i differ from the correct values $\bar{b}_i$ by Δb_i, i.e. $b_i = \bar{b}_i + \Delta b_i$. Substituting $\bar{b} + \Delta b = (\bar{b}_1 + \Delta b_1, \ldots, \bar{b}_6 + \Delta b_6)^t$ into Formula 11 leads to a natural splitting of the polynomial coefficients a_L into terms $\bar{a}_L$ belonging to $\bar{b}$ and terms Δa belonging to Δb.

$$\begin{aligned} a_L &= \left(M_L^t M_L\right)^{-1} M_L^t \left(\bar{b} + \Delta b\right) \\ &\overset{!}{=} \bar{a}_L + \Delta a_L \end{aligned} \tag{13}$$

where

$$\begin{aligned} &\bar{a}_L = \left(M_L^t M_L\right)^{-1} M_L^t \bar{b} \text{ and} \\ &\Delta a_L = \left(M_L^t M_L\right)^{-1} M_L^t \Delta b \,. \end{aligned} \tag{14}$$

Thus, following from Formula 8 the change in the approximated value at $P = (x, y)$ due to the reference station biases Δb is given by

$$\Delta b_L(x, y) = (1, x, y) \cdot \Delta a_L \tag{15}$$

The analysis is restricted to the following two cases:

Case 1: error δ introduced at P_1,
i.e. $\Delta b = (\delta, 0, 0, 0, 0, 0)^t$

Case 2: error δ introduced P_2,
i.e. $\Delta b = (0, \delta, 0, 0, 0, 0)^t$.

Substituting Δb for case 1 and case 2 into Formula 14 yields together with Formula 15 the following general expressions for the change in the approximated value at $P = (x, y)$:

Case 1: $$\Delta b_L(x, y) = (1, x, y) \cdot \left(\frac{1}{6}, \frac{2\sqrt{3}}{15d}, -\frac{2}{15d}\right)^t \delta \tag{16}$$

Case 2: $$\Delta b_L(x, y) = (1, x, y) \cdot \left(\frac{1}{6}, \frac{\sqrt{3}}{15d}, \frac{1}{15d}\right)^t \delta \tag{17}$$

For the rover station $P_0 = (0,0)$, located at the centre of the network, Formulas 16 and 17 reduce to the following simplified expressions

Case 1: $$\Delta b_L(0,0) = \frac{1}{6}\delta$$

Case 2: $$\Delta b_L(0,0) = \frac{1}{6}\delta$$

In fact further analysis, which has been omitted from the text for the sake of brevity, using additional reference stations placed on a circle around the rover shows that the linear approximation attenuates the introduced error by a factor of about $1/n$ where n is the number of stations used in the network.

More generally, Fig. 2 illustrates the magnitude of the approximated error for any station $P = (x, y)$ with $d = \delta = 1$.

Note that the approximated error is largest in the vicinity of the biased station. More accurately, the magnitude of the error depends on the distance of the approximated station from the biased reference station along the line of maximum gradient. The closer the approximated station is, the larger the error.

The preceding theoretical study demonstrates the propagation of reference station biases for the linear approximation; however, it does not measure the impact of these biases on the network RTK performance. To complement the theoretical study, the following section empirically investigates the actual impact of wrong ambiguities on the final position solution.

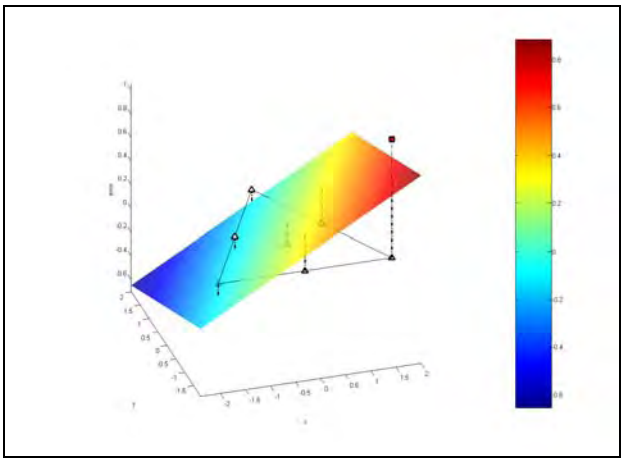


Fig. 2. Case 1: the approximated error at a station $P = (x, y)$ resulting from a bias at one of the reference stations P_1, P_3 or P_5.

2.3 Impact Of Incorrect Ambiguities On The Position Solution.

For the empirical analysis, four hours of 1 Hz dual-frequency data was collected on the network of 4 stations depicted in Fig. 3, which form part of the SAPOS permanent reference station network in Bavaria, Germany.

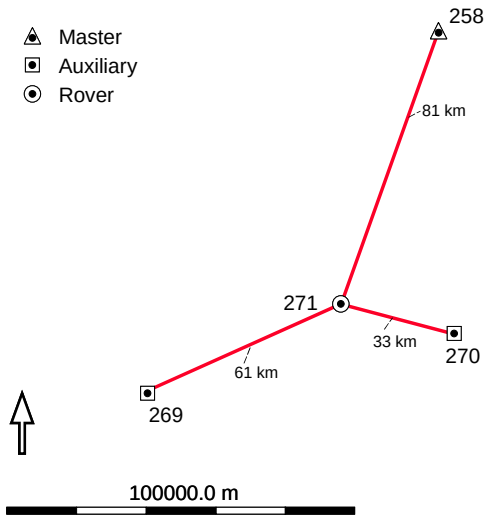


Fig. 3. Distribution of reference stations in the test network. Station 258 was the designated master and station 271 was used to simulate the rover.

The double-difference phase ambiguities between the reference stations 258, 269 and 270 were determined and dispersive and non-dispersive corrections differences, described by Formulas 3 and 4 respectively, were computed using station 258 as the master. Optimal network corrections were approximated for the rover station (271) and subsequently applied to the data. The 2-D linear approximation given by Formula 5 was used for the interpolation process. The corrected data was then processed using discrete observation times of 60 seconds. All the solutions were fixed correctly; a fixed solution being deemed correct if the difference of the height component to the true height of the rover station was less than ±5 cm.

In the first experiment, a bias of +1 cycle was added to the L2 ambiguity of satellite 6 at station 270 before computing the dispersive and non-dispersive correction differences. The L2 frequency was chosen to simulate the real-world situation where the fixing and keeping of L2 ambiguities is generally more problematic compared to L1. Since L2 has no civil code and the P-code is encrypted, proprietary tracking techniques are used to recover the range and phase information. This process yields L2 observables with a higher relative noise and causes the phase measurement to be more susceptible to cycle-slips than L1 at low elevations. Using the biased data, optimal network corrections were computed and the observations reprocessed as previously described. Fig. 4 shows the fixed solutions for the observation period in relation to the elevation of the biased satellite and time. Correctly fixed solutions are shown as solid green squares while solutions with incorrectly fixed ambiguities are represented as hollow red squares.

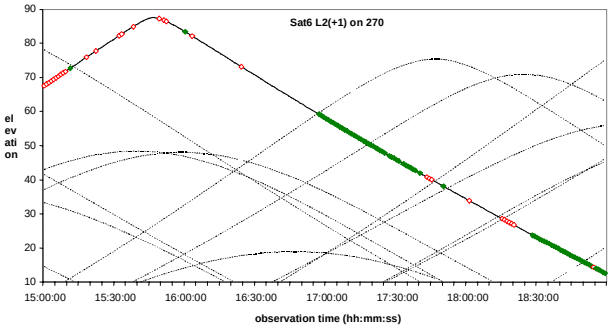


Fig. 4. The number of fixed solutions using corrected data. A bias of +1 cycle was added to the L2 ambiguity of satellite 6 at station 270 before computing the correction differences. Correctly fixed solutions are shown as solid green squares and incorrectly fixed solutions are represented as hollow red squares.

Ambiguity resolution was generally problematic especially when the biased satellite was above an elevation of 60 degrees. The problem to fix ambiguities was expected since an incorrect L2 or L1 ambiguity is amplified in the dispersive and non-dispersive correction differences by a factor of almost 2 (see Section 2.1). According to Tab. 1 and Tab. 2, one could expect fewer problems fixing if both the L1 and L2 ambiguities were biased so that the widelane ambiguity is still valid, since the influence of this bias is reduced by a factor of approximately 2 in the respective correction differences. To test this hypothesis, biases of +1 cycle were added to the L1 and L2 ambiguities of satellite 6 at station 270 prior to generating the correction differences. Optimal correction differences were then applied at the rover and the data reprocessed. The results are depicted in Fig. 5.

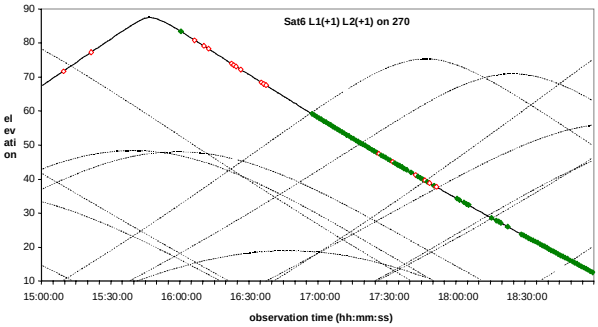


Fig. 5. The number of fixed solutions using corrected observations. Biases of +1 cycles were added to the L1 and L2 ambiguities of satellite 6 at station 270 prior to computing the correction differences.

Ambiguity resolution is still problematic when the biased satellite is above an elevation of 60 degrees. However, below this elevation the ambiguity resolution performance has improved. There are more correctly fixed solutions and importantly the number of wrongly fixed solutions has decreased by approximately 1/2. One could expect further improvement in the fixing performance if the biased station was further from the rover due to the distance dependency inherent in the linear interpolation algorithm (see Section 2.2). Station 269 is approximately twice the distance from the rover as station 270. Biases of +1 cycle were added to the L1 and L2 ambiguities of satellite 6 at this station instead of 270 and the data processed as before.

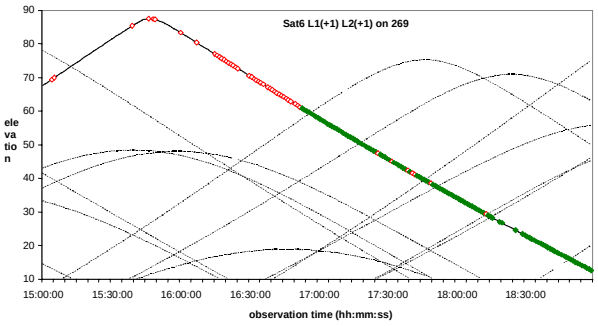


Fig. 6. The number of fixed solutions using corrected observations. Biases of +1 cycles were added to the L1 and L2 ambiguities of satellite 6 at station 269 prior to computing the correction differences.

Again, successful ambiguity resolution is still problematic above 60 degrees. However, below this elevation there are more correctly fixed solutions compared to the results of the previous test (Fig. 5), especially in the elevation band between 25 and 40 degrees. These results also highlight a trait common to the previous two tests; the impact of wrong ambiguities on the position solution is less when the bias is present on a low elevation satellite. To test this premise, a bias of +1 cycle was added to the L2 ambiguity of satellite 10 at station 269. Satellite 10 reached its highest elevation of approximately 20 degrees midway through the observation period. Optimal correction differences were computed for the rover as usual and the data processed for the interval when satellite 10 was above the elevation cut-off. The fixed solutions are shown in Fig. 7.

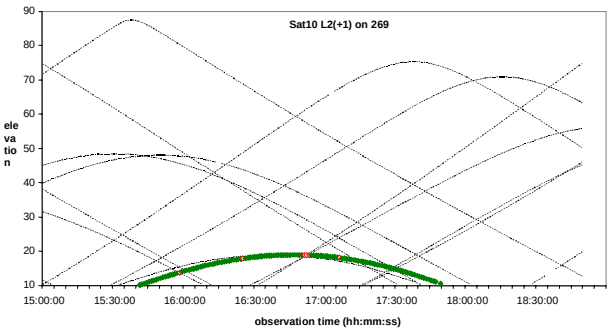


Fig. 7. A bias of +1 cycle was added to the L2 ambiguity of satellite 10 at station 269.

In comparison to the results of the first experiment, described in Fig. 4, the biased satellite has a minimal impact on ambiguity resolution performance; although, it should be noted that a few solutions were fixed incorrectly. For completeness, biases of +1 cycle were added to the L1 and L2 ambiguities of satellite 10 at station 269 and the data reprocessed. The results are given in Fig. 8.

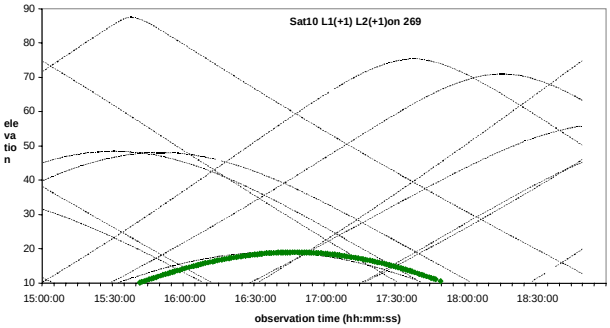


Fig. 8. The number of fixed solutions using corrected observations. Biases of +1 cycle were added to the L1 and L2 corrections of this satellite at station 269.

As expected, when the wide lane ambiguity is correct the biased satellite has virtually no impact on ambiguity resolution performance. In order to analyse the effect of biases on low elevation satellites in more detail, the dispersive and non-dispersive errors for the master-rover baseline (258-271) were grouped into elevation bins of 1 degree according to the elevation of the lowest satellite used to build the double difference. For each elevation bin, the average and mean true errors were calculated where the mean true error is given by:

$$\bar{\varepsilon} = \sqrt{\frac{[\varepsilon\varepsilon]}{n}} \tag{18}$$

and ε is the true error and n is the number of observations. The graph of the average and mean true dispersive errors for the unbiased data is shown in Fig. 9.

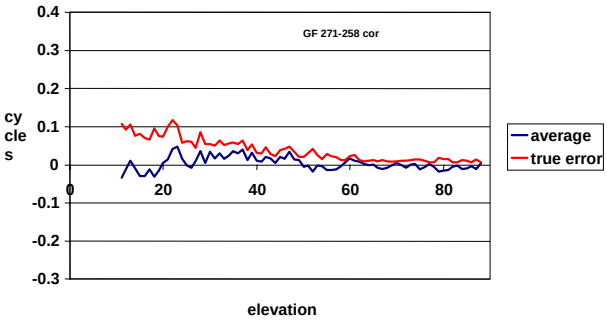


Fig. 9. Corrected average and mean true dispersive errors for the baseline 271-258.

The dispersive errors are generally less than 0.1 cycles and apparently random as indicated by the average error line. In addition, the magnitude of the errors decreases linearly with increasing elevation as shown by the mean true error line. For comparison, the average and mean true dispersive errors are also shown when the bias of +1 cycle was added to the L2 ambiguity of satellite 10 at station 269.

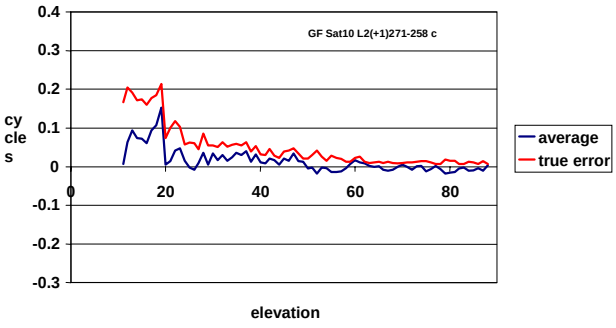


Fig. 10. Corrected average and true dispersive errors for the baseline 271-258. A bias of +1 cycle was added to the L2 ambiguities of satellite 10 at station 269.

As expected, only the dispersive errors below an elevation of 20 degrees are affected. In this elevation band, a bias of approximately +0.1 cycles has been added to the dispersive errors. Combined with an elevation dependent observation weighting strategy, which is common to many baseline processing algorithms, the effect of this bias on the position solution is further reduced. Conversely, the observations of high elevation satellites are given a higher weight by the processor and have a greater influence on the position solution. The problem can be compounded for if the biased high elevation satellite happens to be chosen as the reference. Consider the following average and mean true dispersive errors for the same baseline with an L2 ambiguity bias of +1 cycle added to the observations of the reference satellite 21.

The resulting dispersive errors are biased by approximately +0.3 cycles across virtually all elevations. This bias is larger by a factor of 3 compared to the previous example for the low elevation satellite. By itself, this bias will have a negative impact on ambiguity resolution and when coupled with an elevation dependent weighting strategy effective high-precision positioning can be severely hampered.

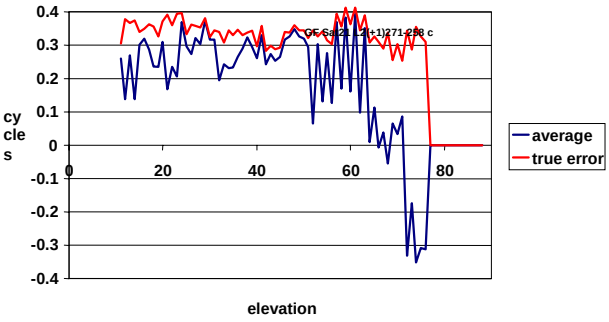


Fig. 11. Corrected average and mean true dispersive errors for the baseline 271-258. A bias of +1 cycle was added to the L2 ambiguities of the reference satellite 21 at station 269.

Several conclusions can be drawn from the empirical analysis presented in this section. First, the influence of wrongly fixed ambiguities on the position solution is greater for high elevation satellites. The problem can be compounded if the reference satellite ambiguity has been fixed incorrectly. Second, the influence of a wrongly fixed reference station ambiguity on positioning performance at the rover is distance dependent. The closer the rover is to the reference station, the larger the effect. However, this conclusion is heavily dependent on the approximation algorithm used to derive optimal network corrections for the rover station. Thirdly, the influence of a wrongly fixed wide lane ambiguity is less than a single L1 or L2 ambiguity bias. The strength of these conclusions should be considered together with the fact that the results can be heavily influenced by satellite-station geometry and the choice of the approximation and processing algorithms employed. In addition, several other biases can have a negative impact on positioning performance, high ionospheric activity for example, which is the subject of the following section.

2 Analysis of the Effect of high Ionosphere on Network RTK

For a network of 6 stations located in Hong Kong, 4 hours of 1 Hz data were collected on the 14th November 2003 between 04:00 am and 08:00 am during a period of known high ionospheric activity (IPS Radio and Space Services, 2003). According to the data archive of the Center for Orbit Determination in Europe (CODE) at the Astronomical Institute of the University of Berne, the TEC (Total Electron Content) value at the respective location and time was about 450. For comparison, the TEC value at the same time for the mid of Germany was about 44. The stations of the network are depicted in Fig. 12.

Cycle slips were removed from the raw data prior to the estimation of the double-differenced phase ambiguities between the reference stations. The resulting ambiguity-levelled data was used to form dispersive and non-dispersive correction differences. Station HKSL represents the master reference station. The remaining stations serve as auxiliaries, except for station HKKT, which is the designated rover. The length of the master-rover baseline is 16.4km. The closest reference station to the rover is HKLT approximately 7.8km away. This baseline, being the shortest, would be used in a usual baseline algorithm where no network corrections were applied and therefore serves as a reference for analysing the benefit of using approximated corrections.

The percentages of fixed ambiguities resulting from different processing strategies on the shortest baseline and also for the master-rover baseline are listed in Tab. 3. These percentages were used as the measure of processing performance.

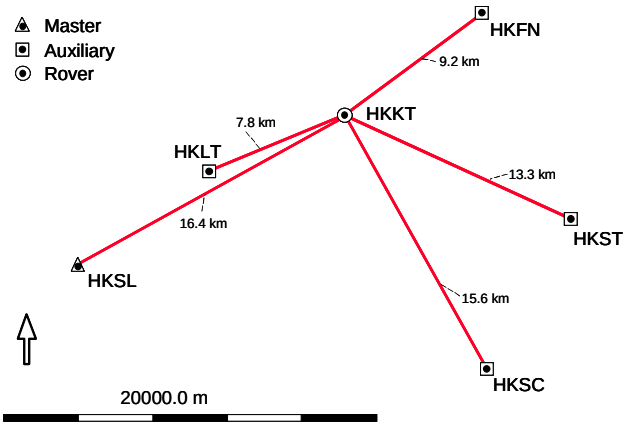


Fig. 12. Distribution of auxiliary reference stations in relation to the rover station HKKT. Station HKSL was the designated master reference station.

Tab. 3. Percentage of fixed ambiguities for the shortest baseline in the Hong Kong network and for the baseline between rover and master for different processing strategies.

Baseline	*Processing Mode*	*Percentage of Fixed Ambiguities*
HKKT - HKLT	No corrections, no stochastic modelling	17.5
HKKT - HKSL	No corrections, ionospheric activity low	32.8
HKKT - HKSL	No corrections, ionospheric activity medium	76.9
HKKT - HKSL	Applied corrections (D=1, ND=15), no stochastic modelling	97.2
HKKT - HKSL	Applied corrections (D=1, no ND), no stochastic modelling	94.4
HKKT - HKSL	Applied corrections (D=1, ND=1), no stochastic modelling	100
HKKT - HKSL	Applied corrections, (D=1, no ND), ionospheric activity low	100

D=x, ND=y determine the update rates in seconds for dispersive (D) and non-dispersive (ND) corrections. In all cases an observation time of 45 seconds and an elevation mask of 10 degrees was chosen.

RTK systems are usually tuned for most general observing conditions. Users should not be required to change special processing parameters, especially when they have no indication when and what to change. Therefore, baselines shorter than 10 km are usually processed in real-time without stochastically modelling the ionosphere (in order not to confuse multipath or obstructions with ionospheric noise). For this reason no stochastic modelling was used on the shortest baseline HKKT – HKLT. Baselines between 10 and 20 km are already in the range where ionospheric biases are more likely to be present and affect positioning results. For such baselines, real-time algorithms would stochastically model the ionosphere using parameters associated with a low ionospheric activity setting. In post-processing, an operator might check which ionospheric setting produces the optimal solution, as demonstrated in Tab. 3 where the low and medium ionospheric activity settings were tested for the short baseline. However, this approach is not feasible in real-time since the system deals with short occupation times and has no indication of long-time behaviour. To compare the effectiveness of network corrections with a usual approach, the master-rover baseline was processed with the standard ionospheric settings that would be used if no corrections were applied.

As described in the introduction, the update rates of the dispersive and non-dispersive corrections are chosen differently to maximise the data throughput. For this experiment, an update rate of 1 Hz was always used for the dispersive errors. For the non-dispersive errors an update rate of 15 seconds was compared with an update rate of 1 second. In addition, the effect of applying dispersive but no non-dispersive corrections (D=1,no ND) was also tested. The percentage of fixed solutions increases from 17.5% on the shortest baseline HKKT – HKLT to 97.2% on the rover-master baseline HKKT – HKSL when an update rate of 15 seconds for the non-dispersive corrections (which is a typical update rate for a real-world application) was used. Increasing the update rate for the non-dispersive errors to 1 Hz increases the number of fix solutions to 100%. It should be emphasized that these results are achieved without stochastic modelling. When the ionospheric activity setting for ionospheric stochastic modelling is set to low, 100%

fixed solutions are also obtained without any non-dispersive corrections. These results compare favourably to the 32.8% of fixed solutions obtained on the short baseline when no network corrections were applied. This analysis illustrates the benefits of using network corrections in the presence of relatively high ionospheric disturbances even for a small sized network.

3 Conclusions

The effect of various combinations of wrong L1 and L2 integers on the correction difference observables, which could be introduced by either wrong initial fixing or undetected cycle slips, was tabulated and the propagation of these biases in a 2-D linear approximation was analysed for an idealized network of reference stations. The actual impact on ambiguity resolution was investigated by comparing unbiased results with computations where artificial ambiguities were added to the observations. The processing results showed that a bias on only one frequency affects the final performance of a real-time system more than when the widelane is still valid; that is when identical biases for both frequencies are present. Therefore, correction differences with a correct widelane ambiguity, which are provided for in the current RTCM SC104 network RTK message proposal, could help to fix the ambiguities of low elevation satellites. These observations may eventually have to be down-weighted.

Reference station networking is usually considered as an approach for achieving better RTK performance over long baseline lengths. It is often argued that establishing a reference station network with inter-station distances of only a couple of 10kms is excessive. However the example of the Hong Kong network, with an average reference station separation of less than 15km, shows that ionospheric disturbances can be severe in equatorial regions and hamper effective positioning on baselines usually considered as unproblematic. The additional information provided by surrounding reference stations increased the performance of RTK positioning from very unfavourable, with only 32.8% fixed solutions, to high performance with a success rate of 100%.

Acknowledgements. We thank Mr. Andreas Brünner of SAPOS Bayern for kindly supplying us the Bavarian data set and Mr. Simon Kwok of the Geodetic Surveying Section, Hong Kong Lands Department for the Hong Kong data set.

References

CODE data archive, ftp://ftp.unibe.ch/aiub/CODE/

Euler, H.-J., Keenan, C.R., Zebhauser, B. E., Wübbena, G. (2001) ***Study of a Simplified Approach in Utilizing Information from Permanent Reference Station Arrays***, ION GPS 2001, September 11-14, 2001, Salt Lake City, UT

Euler, H.-J., Zebhauser, B. E., Townsend, B., Wübbena, G. (2002) ***Comparison of Different Proposals for Reference Station Network Information Distribution Formats***, ION GPS 2002, September 24-27, 2002, Portland, OR

Euler, H.-J., Zebhauser, B.E. (2003) ***The Use of Standardized Network RTK Messages in Rover Applications for Surveying***, ION NTM 2003, January 22-24, 2003, Anaheim, CA

Euler, H.-J., Seeger, S., Zelzer, O., Takac, F., Zebhauser, B. E. (2004-I) ***Improvement of Positioning Performance Using Standardized Network RTK Messages, ION NTM 2004***, January 26-28, 2004, San Diego, CA

Euler, H.-J., Seeger, S., Takac, F. (2004-II) ***Influence of Diverse Biases on Network RTK***, ION GNSS 2004, Long Beach, CA

RTCM (2004) ***RTCM Recommended Standards for Differential GNSS (Global Navigation Satellite Systems) Service, Version 3.0***, RTCM paper 30-2004-SC104-STD.

Zebhauser, B. E., Euler, H.-J., Keenan, C.R., Wübbena, G. (2002) ***A Novel Approach for the Use of Information from Reference Station Networks Conforming to RTCM V2.3 and Future V3.0***, ION NTM 2002, January 28-30, 2002, San Diego, CA.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 79-84

Journal of Global Positioning Systems
2004
© CPGPS

First results from Virtual Reference Station (VRS) and Precise Point Positioning (PPP) GPS research at the Western Australian Centre for Geodesy

N. Castleden

Western Australian Centre for Geodesy, Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia
e-mail: castledn@vesta.curtin.edu.au; Tel: +61 8 9266 7559; Fax: +61 8 9266 2703

G.R. Hu

Western Australian Centre for Geodesy, Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia
e-mail: hug@vesta.curtin.edu.au; Tel: +61 8 9266 7559; Fax: +61 8 9266 2703

D.A. Abbey

AAMHatch Pty Ltd, 23 Hamilton Street, Subiaco, WA 6008, Australia
e-mail: d.abbey@aamhatch.com.au; Tel: +61 8 9381 4133; Fax: +61 8 9381 6161

D. Weihing

Geodetic Institute, University of Karlsruhe, Englerstr.7, D-76128 Karlsruhe, Germany
e-mail: dianaw@gik.uni-karlsruhe.de; Tel: +45 721 608 2305; Fax: +45 721 608 6552

O. Øvstedal

Department of Mathematical Sciences and Technology, Agricultural University of Norway, P.O. Box 5003, N-1432 Ås, Norway
e-mail: ola.ovstedal@imt.nlh.no; Tel: +47 6494 8876; Fax: +47 6494 8810

C.J. Earls

AAMHatch Pty Ltd, 23 Hamilton Street, Subiaco, WA 6008, Australia
e-mail: c.earls@aamhatch.com.au; Tel: +61 8 9381 4133; Fax: +61 8 9381 6161

W.E. Featherstone

Western Australian Centre for Geodesy, Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia
e-mail: W.Featherstone@curtin.edu.au; Tel: +61 8 9266 2734; Fax: +61 8 9266 2703

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Over the past 18 months, a team in the Western Australian Centre for Geodesy at Curtin University of Technology, Perth, has been researching the optimum configurations to achieve long-range and precise GPS-based aircraft positioning for subsequent airborne mapping projects. Three parallel strategies have been adopted to solve this problem: virtual reference stations (VRS), precise point positioning (PPP), and multiple reference stations (MRS). This paper briefly summarises the concepts behind the PPP and VRS techniques, describes the development and testing of in-house software, and presents the latest experimental results of our research. Current comparisons of the PPP and VRS techniques with an independently well-controlled aircraft trajectory and ground-based stations in Norway show that each deliver precisions of around 3 cm. However, the implementation of more sophisticated error modelling strategies in the MRS approach is expected to better deliver our project's objectives.

Key words: GPS, kinematic positioning, VRS, PPP

1 Introduction

For many years, robust and reliable long-range kinematic positioning has been the desire of airborne mapping projects that utilise relative carrier-phase GPS as the primary positioning technique (e.g., Cannon et al., 1992; Colombo and Evans, 1998). However, this has been difficult to achieve in a production environment because of the de-correlation of, notably atmospheric, GPS-related errors as the rover-to-basestation receiver distance increases. This is somewhat problematic in Australia, where airborne mapping projects are usually undertaken in remote areas in which geodetic control is sparse and ground access is difficult, thus significantly increasing project costs.

There are currently three viable approaches to tackle this problem:

1. The multi-basestation or multiple reference station (MRS) approach (e.g., Han and Rizos, 1996; Wübbena et al., 1996; Raquet, 1998; Wanninger, 1999, 2002; Vollath et al., 2000; Fotopoulos and Cannon, 2001);
2. The virtual reference station (VRS) concept (e.g., Vollath et al., 2000; Higgins, 2002; Wanninger, 2002; Hu et al., 2003); and
3. The precise point positioning (PPP) technique (e.g., Kouba and Heroux, 2001; Gao and Shen, 2001; Witchayangkoon, 2000).

This paper describes initial results of our attempts to implement the VRS and PPP approaches; our MRS approach will be described in a later paper. Here, we briefly describe the VRS and PPP concepts, followed by results of some of our initial experiments using our in-house-developed software. These show that both techniques can deliver ~3 cm precision, but we expect that improved error modelling from the MRS approach may be better still.

2 The VRS concept

The VRS concept is a derivative of the MRS approach, but differs in that the surrounding reference GPS receivers are instead used to determine 'synthetic' dual-frequency code and carrier-phase GPS data at a virtual basestation that is located close to the user's receiver. The user's and the virtual basestation GPS data are then processed in single baseline mode to determine the coordinates of the user's receiver. This can be achieved in near-real-time (akin to real-time kinematic, RTK) or post-processed modes, but the near-real-time mode requires high-bandwidth telecommunications among the reference stations, master control station (where the virtual GPS data are computed) and subsequently to the user.

Of pragmatic benefit, the VRS approach does not require an actual physical reference station close to the user, but does require a MRS network of GPS stations surrounding the area of interest. The VRS concept allows the user to access data from a virtual GPS reference station at any location interpolated within the network coverage area. Also, the VRS approach is more flexible in terms of permitting users to use their current receivers and software without involving any special software or communications equipment (if used in post-processed mode) to simultaneously manage data from all the reference stations.

With VRS, users within the MRS network can operate at distances greater than conventional RTK or fast/quick-static GPS modes (typically 10 km and 20 km, respectively) without degrading accuracy. Under ideal conditions, the VRS approach can deliver single-point coordinate accuracies of a few centimetres for a MRS network of reference stations separated by 50-70 km.

With regard to our application-specific problem of aircraft positioning, it should be noted that the VRS concept was initially developed for fast/quick-static or RTK users working in fairly small (less than 10 km) areas. In such conditions, one approximate user position is sufficient for the process of VRS data generation. For long-range pure kinematic GPS positioning, however, the rover may be moving over longer distances. Therefore, it is necessary to frequently update the position of VRS so that the VRS-to-rover separation is not too large, ideally less than 10 km.

As such, we have modified the VRS method proposed by Wanninger (2002) for long-range post-processed kinematic GPS applications. In this modified method, the VRS is referred to a fixed position according to the rover's initial approximate position, and the corrections are applied to the rover's trajectory. When the rover's current approximate position becomes greater than 10 km from the initial VRS position, a new VRS is created. This process continues each time the distance reaches 10 km, essentially making the VRS 'follow' the aircraft's trajectory.

3 The Curtin-AAMHatch VRS software

Post-processing VRS software for medium- to long-range (30-100 km) kinematic positioning has been developed in the Western Australian Centre for Geodesy at Curtin University of Technology, Perth. The software generates modified (as described above) VRS data for the pure kinematic user in RINEX format so that the VRS-to-user distance is always less than 10 km. Because of the elevation difference between the ground-based and airborne GPS receivers, the tropospheric delay must be treated carefully.

Based upon our earlier tests, the UNB3 models (Collins and Langley, 1996) with Niell mapping functions (Niell, 1996) are used in the in-house software to compute the *a priori* troposphere delay. The post-processing mode of operation allows for more careful treatment of the GPS data and avoids the high-bandwidth telecommunication problems commonly encountered in RTK applications. For instance, we use the International GPS Service's (IGS) Final Product orbits instead of the broadcast GPS ephemerides.

4 VRS experiments and results

A dual-frequency airborne GPS dataset from 08:25 to 09:25 (GPS time) on 8 May 2002 was arbitrarily selected from a test flight in Norway, with a data collection rate of 1 Hz [for further details about the whole flight period, see Kjorsvik et al. (2004)]. This test used three Norwegian reference stations (SAND, SORH and OE23; Fig 1) to form a MRS network, also with the sampling rate of 1 Hz. The reference stations were equipped with dual-frequency Trimble MS750 receivers and TRM41249.0 antennas. The distance between the reference stations varies from 60-70 km. Figure 1 shows the test network and the trajectory of the flight during the test period.

The VRS data were generated using our modified VRS methods and software, then the aircraft trajectory determined in a single-baseline mode with respect to each 'following' VRS station. The reference trajectory of the aircraft is known to an accuracy of several centimetres from comparisons with different software and independently determined positions from an aerial triangulation (Kjorsvik et al., 2004). The modified-VRS-estimated positions were then compared with the reference trajectory.

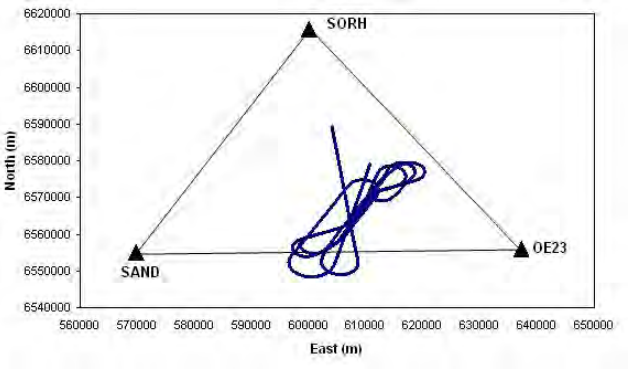


Fig 1. The one-hour aircraft flight trajectory (blue) and the three-reference-station Norwegian network used for the modified VRS tests (UTM projection)

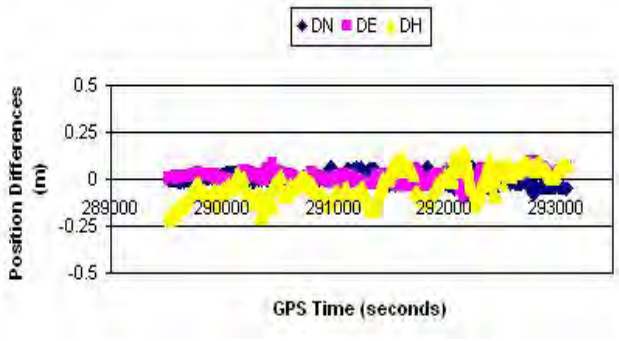


Fig 2. Epoch-by-epoch differences (metres) between the modified VRS results and the reference trajectory for the one-hour airborne kinematic GPS test in Norway

	STD	Mean	Max	Min
North	0.031	-0.005	0.068	-0.076
East	0.031	0.010	0.097	-0.107
Height	0.078	-0.035	0.157	-0.219

Table 1. Descriptive statistics of the differences (metres) between the modified VRS results and the reference trajectory for the one-hour airborne kinematic GPS test in Norway

Figure 2 shows the position differences in the easting, northing and height components, and Table 1 summarizes the results of the comparisons conducted during the one-hour test period. From Table 1, the standard deviations of north, east and height components are 31 mm, 31 mm and 78 mm, respectively, when using the modified VRS method. Importantly, this is commensurate with the expected accuracy of the reference trajectory.

As expected, the precision of the height component is a factor of 2 to 3 times worse than for the horizontal components (Table 1). Since the accuracy of GPS positioning is inherently worse in the height component, it is necessary to pay careful attention to the *a priori* tropospheric delay, which is the subject of our current work (e.g., Hu et al., 2005). The oscillations of the horizontal coordinate precision in Fig 2 are probably caused mainly by multipath from the aircraft. The main sources for the vertical oscillations are most probably due to the combination of multipath and residual tropospheric errors.

5 The PPP concept

Classical GPS point positioning, also known as standalone positioning, involves only one code-tracking GPS receiver. That is, one GPS receiver simultaneously tracks the code pseudoranges from four or more GPS satellites to determine its position (e.g., Hofmann-Wellenhof et al., 2001). After the discontinuation of

Selective Availability, the expected horizontal positioning accuracy of this classical point positioning approach is about 22m (2DRMS) or better (e.g., Shaw et al., 2000; Witchayangkoon, 2000).

However, it is also possible to determine single point positions from both code and carrier-phase data. Using precise GPS orbit and clock products (i.e., time corrections to the satellite clocks) available from the IGS, it has been shown that carrier-phase-based single point positioning can be improved to decimetre accuracy levels (e.g., Witchayangkoon, 2000; Gao and Shen, 2001; Kouba and Heroux, 2001). This is through the use of ionosphere-free, undifferenced code pseudorange and carrier-phase measurements, together with the precise IGS orbit and clock products. This approach is commonly known as Precise Point Positioning (PPP).

The PPP approach is very similar to single-point code pseudorange positioning achieved from a hand-held GPS receiver, except that it adds the more precise carrier-phase GPS data, as well as improved orbits and satellite clock corrections. Under ideal conditions, the PPP concept can deliver single point accuracies of 5-10 cm *irrespective of baseline length*.

PPP is relatively new and is now attracting much attention internationally (e.g., Bisnarth et al., 2002). Importantly, it requires only one dual-frequency carrier-phase GPS receiver and thus avoids the expense and logistics of deploying a network of GPS receivers surrounding the area of interest, as is needed for the MRS and VRS techniques. However, it uses a (global) MRS network by 'stealth' through the data provided by the IGS.

6 The Curtin-AAMHatch PPP software

We are developing in-house software completely from scratch to implement the PPP technique, in post-processing mode, according to two mathematical models.

- *PPP Strategy 1* takes the model of Kouba and Heroux (2001), which uses ionosphere-free linear combinations of L1 and L2 carrier-phase data, but cannot deliver integer-ambiguity-fixed solutions (i.e., it only delivers float solutions).
- *PPP Strategy 2* uses the model of Gao and Shen (2001), which uses a ionosphere-free linear combination of L1 and L2 carrier-phase data that reduces code pseudorange noise and tries to return the integer properties of the ambiguities.

Currently, only PPP Strategy 1 is coded in our software; Strategy 2 is under development.

Since the PPP technique does not enjoy the cancellation of errors found in relative/differential GPS techniques, several corrections have been included in our software. In addition to the IGS Final products described earlier, these are: relativity corrections for the satellites' orbital eccentricity and the Earth's rotation (the Sagnac correction), satellite antenna offsets from their centre of gravity (notably the so-called Y-bias), tropospheric delays, phase windup (essentially the carrier-phase at the satellite) and solid-Earth tides. A sequential filter and smoothing procedure has also been implemented in our PPP software, where the ambiguities are essentially carried forward from epoch to epoch.

7 PPP experiments and results

While we seek airborne kinematic positioning, we have undertaken our PPP Strategy 1 software testing with ground-based static GPS data, but processed in kinematic mode. Here, the 'error' values are the epoch-by-epoch differences between the PPP-estimated position and the well-controlled ground-based station coordinates. While we have compiled a large number of well-coordinated static GPS datasets, the few results presented below for the test data in Norway are typical of those achieved with our other datasets. The use of an epoch-by-epoch PPP solution partially simulates a (ground-based) kinematic example.

Five GPS stations (HONE, ARNE, KONG, SORH, SAND; Fig 3) from a MRS network in Norway were used. Although the distances between these five stations range between 63 km and 132 km, this does not matter because the PPP method is baseline-length-independent. Each station was equipped with Trimble MS750 dual-frequency receivers and TRM41249.0 antennas. The data set used to test the PPP software was measured between 07:45 and 10:15 (GPS time) on 8 May 2002 with a sampling rate of 1 Hz.

The 3D coordinates of these five reference stations are known to ~1 cm in the horizontal and ~3 cm in the vertical on the EUREF89 datum. These were determined from processing three days of static GPS data with the GIPSY scientific software. The north, east and up component 'error' values, epoch-by-epoch, were computed by subtracting the PPP-estimated position from the reference coordinates. One graphical example is presented in Fig 4.

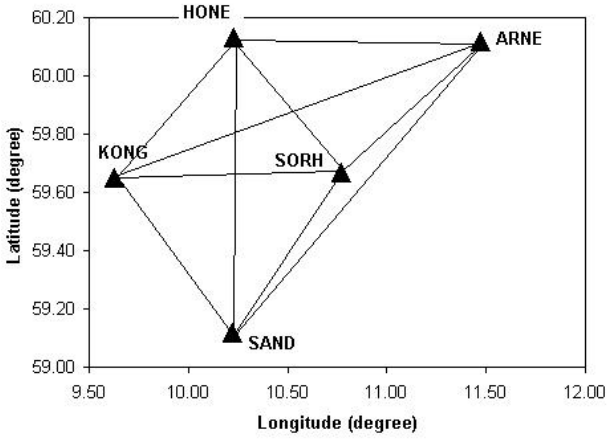


Fig 3. The five-reference-station Norwegian network used for the PPP tests (0.1° ~ 11 km)

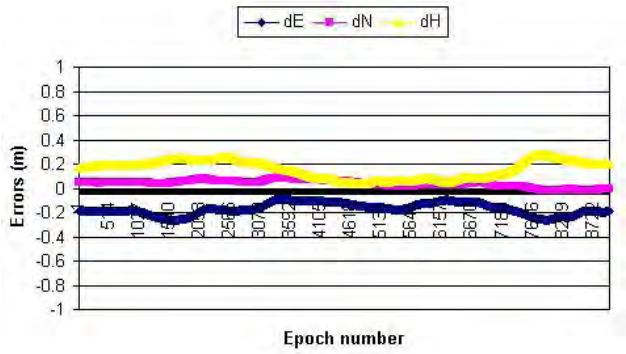


Fig 4. Epoch-by-epoch differences (metres) between the PPP results and the known coordinates of the static station KONG for the three-hour (simulated) kinematic GPS test in Norway

Station	STD			Mean		
	North	East	Height	North	East	Height
KONG	0.028	0.048	0.074	0.036	-0.170	0.157
HONE	0.030	0.032	0.063	0.042	-0.198	0.034
ARNE	0.033	0.030	0.053	0.038	-0.128	-0.016
SORH	0.023	0.033	0.035	0.062	0.006	0.118
SAND	0.005	0.005	0.002	-0.008	-0.034	0.116

Table 2. Descriptive statistics of the differences (metres) between the PPP results and the known coordinates of the static stations for the three-hour (simulated) kinematic GPS tests in Norway

Table 2 summarises the PPP 'error' values for all five stations. It can be seen from Table 2 that the PPP Strategy 1 solutions achieve around 30 mm precisions for the north, east and height components in terms of standard deviations, but this depends on the GPS data quality from the stations. Curiously, the height errors are less than expected in some cases.

8 Concluding remarks

We have presented some of our preliminary results for long-range kinematic GPS positioning using two post-processing software packages developed in-house at the Western Australian Centre for Geodesy. For the modified VRS software, test results using one-hour of 1 Hz real airborne kinematic GPS data in Norway show ~3 cm precision for the horizontal components and ~8 cm precision for the height component. For the PPP software, epoch-by-epoch static (i.e., simulated ground-based kinematic) test results using three hours of 1 Hz data in Norway indicate ~3 cm precision in all three coordinate components. While these results are extremely encouraging, and commensurate with studies conducted by other authors, we are now focussing on the post-processed MRS approach (e.g., Hu et al., 2005), which will allow the use of network constraints and improved (particularly atmospheric) error modelling.

Acknowledgements:

The Australian Research Council and AAMHatch Pty Ltd jointly fund this research

References

Bisnarth SN, Beran T, Lngley RB (2002) ***Precise platform positioning with a single GPS receiver***, GPS World 13(4): 42-49.

Cannon ME, Schwarz KP, Wei M (1992) ***A consistency test of airborne GPS using multiple monitor stations***, Bulletin Géodésique 66(1): 2-11

Collins P, Langley R (1996) ***Limiting factors in tropospheric propagation delay error modelling for GPS airborne navigation***. Proceedings of The Institute of Navigation 52nd Annual Meeting, 19–21 June, Cambridge, Massachusetts, USA, pp 519-528

Colombo OL, Evans AG (1998) ***Testing decimeter-level, kinematic, differential GPS over great distances at sea and on land***, Proceedings of ION GPS 1998, 15-18 Sept, Nashville, Tennessee, USA, pp 1257-1264

Fotopoulos G, Cannon ME (2001) ***An overview of multi-reference station methods for cm-level positioning***, GPS Solutions 4(3): 1-10

Gao Y, Shen X (2001) ***Improving ambiguity convergence in carrier phase-based precise point positioning***. Proceedings of ION GPS 2001, 11-14 September, Salt Lake City, Utah, USA, pp 1532-1539

Han S, Rizos C (1996) ***GPS network design and error mitigation for real-time continuous array monitoring system***, Proceedings of ION GPS-1996, Sept 17-20, Kansas City, Missouri, USA, pp 1827-1836

Higgins MB (2002) ***Australia's changing surveying infrastructure from marks in the ground to virtual reference stations***. Proceedings of FIG XXII International Congress, April 19-26, Washington, USA, CD-ROM paper TS5.6

Hofmann-Wellenhof B, Lichtenegger H, Collins J (2001) ***Global positioning system: theory and practice***. 5th edn. Springer, Vienna New York

Hu GR, Khoo HS, Goh PC, Law CL (2003) ***Development and assessment of virtual reference stations for RTK positioning***. Journal of Geodesy 77(5-6): 292-302

Hu G, Abbey DA, Castleden N, Featherstone WE, Earls CJ, Ovstedal O, Weihing D (2005) ***An approach for instantaneous ambiguity resolution for medium- to long-range multiple reference station networks***. GPS Solutions 9(1):1-11

Kjorsvik N, Ovstedal O, Svendsen JGG (2004) ***Long range differential GNSS for positioning of airborne sensors***. in: Sansò, F. (ed) A Window on the Future of Geodesy, Springer, Berlin Heidelberg New York (in press)

Kouba J, Héroux P (2001) ***Precise Point Positioning using IGS orbit and clock products***. GPS Solutions 5(2): 12-28

Niell AE (1996) ***Global mapping functions for the atmosphere delay at radio wavelengths***. Journal of Geophysical Research 101(B2): 3227-3246

Raquet J (1998) ***Development of a method for kinematic GPS carrier phase ambiguity resolution using multiple reference receivers***, UCGE report 20116, Dept of Geomatics, Univ of Calgary (http://www.geomatics.ucalgary.ca/gradtheses.html).

Shaw M, Sandhoo K, Turner D (2000) ***Modernization of the Global Positioning System***. GPS World 11(9): 36-44

Vollath U, Buecherl A, Landau H, Pagels C, Wagner B (2000) ***Multi-base RTK positioning using virtual reference stations***. Proceedings of ION GPS-2000, 19-22 Sept, Salt Lake City, Utah, USA, pp 123-131

Wanninger L (1999) ***The performance of virtual reference stations in active geodetic GPS-networks under solar maximum conditions***. Proceedings of ION GPS-1999, Nashville, Tennessee, USA, pp 1419-1427

Wanninger L (2002) ***Virtual reference stations for centimeter-level kinematic positioning***, Proceedings of ION GPS-2002, 24-27 Sept, Portland, Oregon, USA, pp 1400-1407

Witchayangkoon B (2000). ***Elements of GPS Precise Point Positioning.*** Ph.D. dissertation, Dept of Spatial Information Science and Engineering, Univ of Maine, USA, 265 pp

Wübbena G, Bagge A, Seeber G, Böder V, Hankemeier P (1996) ***Reducing distance dependent errors for real-time precise DGPS applications by establishing reference station networks***. Proceedings of ION GPS-1996, Sept 17-20, Kansas City, Missouri, USA, pp 1845-1852

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 85-94

Journal of Global Positioning Systems
2004
© CPGPS

Augmentation of Low–Cost GPS Receivers via Web Services and Wireless Mobile Devices

Roger Fraser, Adam Mowlam, Philip Collier

Department of Geomatics, The University of Melbourne, Australia
e-mail: rfraser@sli.unimelb.edu.au Tel: 61+ 3 8344 4509; Fax: 61+ 3 9347 2916

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Low–cost GPS receivers can produce positions almost instantly, however they have a limited use and application due to the impact of random and systematic errors associated with real time autonomous positioning. To achieve higher levels of accuracy and precision, some other form of correction or augmentation information must be applied. There are various real time augmentation alternatives, such as WAAS/LAAS, integrated sensors and systems, receiver based optimal estimation algorithms, and potentially, combined GNSS. To improve the accuracy of low–cost GPS receivers, a feasible option is Differential GPS (DGPS). A popular means for transferring real time DGPS corrections is via the RTCM SC–104 protocol over radio transmission. In recent times, the Internet has been shown to be an efficient and reliable form of data communication. In this paper, the Web services architecture is examined as a viable protocol and communication alternative for disseminating DGPS augmentation information over the Internet. Preliminary results from a simple prototype indicate that Web services offers a practical, efficient and secure method for exchanging CORS network data, and augmenting GPS enabled mobile devices capable of wirelessly reaching the internet. Web services are further shown to provide advantages for disseminating other GPS related data, such as IGS satellite orbit data, carrier–phase data for location–centric augmentation, and a host of other LBS information.

Key words: Low–cost GPS, Internet DGPS, Web Services, Mobile Devices

1 Introduction

There are numerous types of GPS receivers in the current international marketplace, ranging from inexpensive, low accuracy handheld devices to expensive, high precision geodetic equipment. By and large, low–cost GPS receivers (whether sold as a plug–in hardware device or as a complete navigation and positioning receiver) have almost assumed mass market status in the consumer electronics industry. Recent advances in micro and wireless technology, reductions in consumer costs, and the apparent growth of the Location Based Services (LBS) industry have somewhat fuelled the need for mobile (information communications and technology) consumers to become "location aware".

Notwithstanding current levels of maturity in GPS hardware and algorithms, low-cost GPS receivers still suffer from large positioning errors mainly attributable to atmospheric effects, broadcast ephemeris errors, multipath and receiver noise. As a result, they have a limited use in real time applications despite being able to produce positioning and navigation information almost instantaneously. When higher levels of precision and accuracy are required, some form of correction or augmentation must be applied so as to reduce the influence of random and systematic errors on the autonomous positioning solution.

There are a host of augmentation options, in general, that are available for minimising the influence of random and systematic errors in GPS positioning. These include Local Area Augmentation Systems (DGPS, VRS, Network RTK, Pseudolites, US LAAS); Wide Area Augmentation Systems (EGNOS, MSAS, US WAAS, GAGAN); integrated positioning and navigation sensors (gyro, precise clocks, INS, digital compass); various correction algorithms (integrity monitoring, noise filtering, atmospheric modelling, code–carrier phase smoothing, multipath detection); and in the foreseeable future, integrated satellite systems (GPS, Galileo, GLONASS, QZSS). To improve the accuracy of low–cost GPS receivers, the most practicable form of augmentation is via DGPS corrections obtained either from a local broadcasting service or from a nearby

reference station. Historically, local area DGPS corrections have been provided through the RTCM protocol over the conventional means of radio transmission.

In recent times, research and development has seen the implementation and analysis of real time Internet based DGPS systems. For instance, the German Federal Agency for Cartography and Geodesy (BKG) together with other partners have developed a dissemination standard called Networked Transport of RTCM via Internet Protocol (NTRIP) for the real time streaming of DGPS or RTK corrections to mobile receivers (Lenz, 2004). Similarly, the Jet Propulsion Laboratory (JPL) has developed a Global Differential GPS (IGDG) system, which facilitates the exchange of corrections to the orbits and clocks of the GPS constellation over the Internet (Muellerschoen et al, 2002). These two developments, together with other independent tests (Gao et al, 2002, Kechine et al, 2003, for example) show that the Internet is able to provide satisfactory levels of accuracy and latency in the dissemination of GPS augmentation information.

In all cases mentioned, the dissemination of corrections is provided over the Internet via constant streaming of data. The purpose of this paper is to examine the XML Web services architecture as an alternative mechanism for exchanging DGPS augmentation information between Continuously Operating Reference Station (CORS) networks and mobile devices over the Internet. The key difference between disseminating data via Internet streaming data and Web services is that the latter requires two–way "request and response" communication. To illustrate the capabilities and performance of an augmentation system based on the Web services architecture, a simple prototype has been developed. The differential positioning technique implemented within the prototype uses the DGPS code range model described by Hofmann-Wellenhof et al (2001). The Web service prototype is consistent with international Internet technology standards established by the Open Geospatial Consortium (OGC) and the World Wide Web Consortium (W3C) and as such, may be accessed and implemented using any standards–compliant software and/or platform.

Principal results from prototype testing confirm that the Web services architecture is a viable option for exchanging GPS augmentation information over the Internet. Tests conducted over peak and off–peak periods using a mobile phone and GPRS show a typical latency of 1–3 seconds in medium– to high–density urban environments. The disadvantages of using the Web services architecture for an Internet based DGPS system relate primarily to the bloat in the message caused from XML tags.

Before the prototype and preliminary test results are discussed in detail, some basic principles concerning Web services, low–cost GPS receivers and DGPS augmentation are outlined.

2 Web Services — Interoperability, Lifecycle and Encryption Concepts

Web services provide a standardised way of interoperating between different software applications, running on various platforms and/or frameworks distributed on a network or over the internet. In particular, a Web service is a piece of application logic residing at a specific network location (or network address) that is accessible to programs via standard internet protocols. When called, a Web service performs one or more functional tasks and then sends a response back to the calling agent. During the process, a Web service may call other Web services and/or run other software applications. The response may be in the form of an entire data set such as a map, a database entry, or simply the result of a computation.

Web services use standardised protocols (i.e. OGC, W3C) so that raw data can be exchanged between operating systems and software components, whether compatible or not, in a platform independent way. The key to the success of Web service interoperability therefore, is in the implementation of open standards for data encoding, communication and transport protocols. The interoperability stack shown in Figure 1 illustrates the standards based enabling technologies upon which Web services are implemented and deployed.

Universal Discovery Description and Integration (UDDI) enables the publishing of services in a directory (similar in concept to the yellow pages), making it possible for other developers to locate and consume a Web service. A Web service's interface is described using Web Service Description Language (WSDL). A WSDL document describes everything a developer requires to enable software to interact with the service, such as input messages, the format of its responses to those messages, protocols that the service supports and where to send the messages.

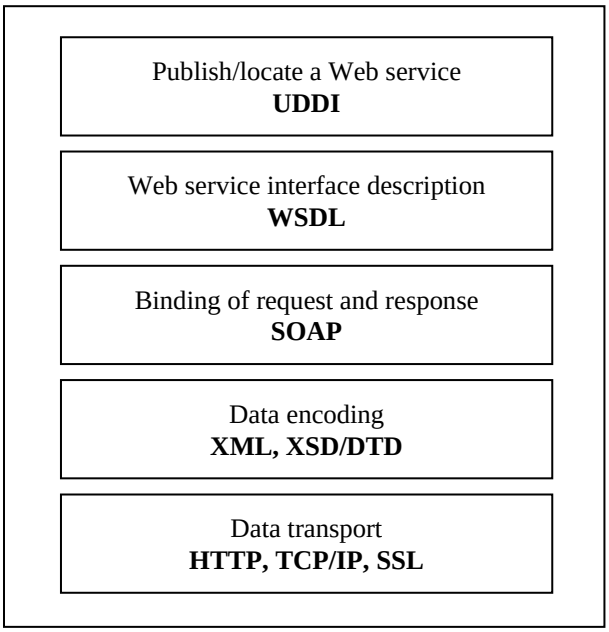


Fig. 1 Web services interoperability stack

Simple Object Access Protocol (SOAP) is a standardised, lightweight protocol for connecting service endpoints distributed over network environment. The SOAP envelope is the cornerstone to interoperability as it enables a Web service to be used by any system capable of sending and receiving plain text over a distributed computing framework. With the exception of pure binary data, the structured data and SOAP envelope are encoded in XML.

Formats for data encoding are described in a schema language such as XML Schema (XSD) or Document Type Definition (DTD). HyperText Transfer Protocol (HTTP) and Transmission Control Protocol/Internet Protocol (TCP/IP) permit the connectivity between software components and Web services by enabling them to send and receive messages. The lifecycle of a Web service request and response is shown in Figure 2.

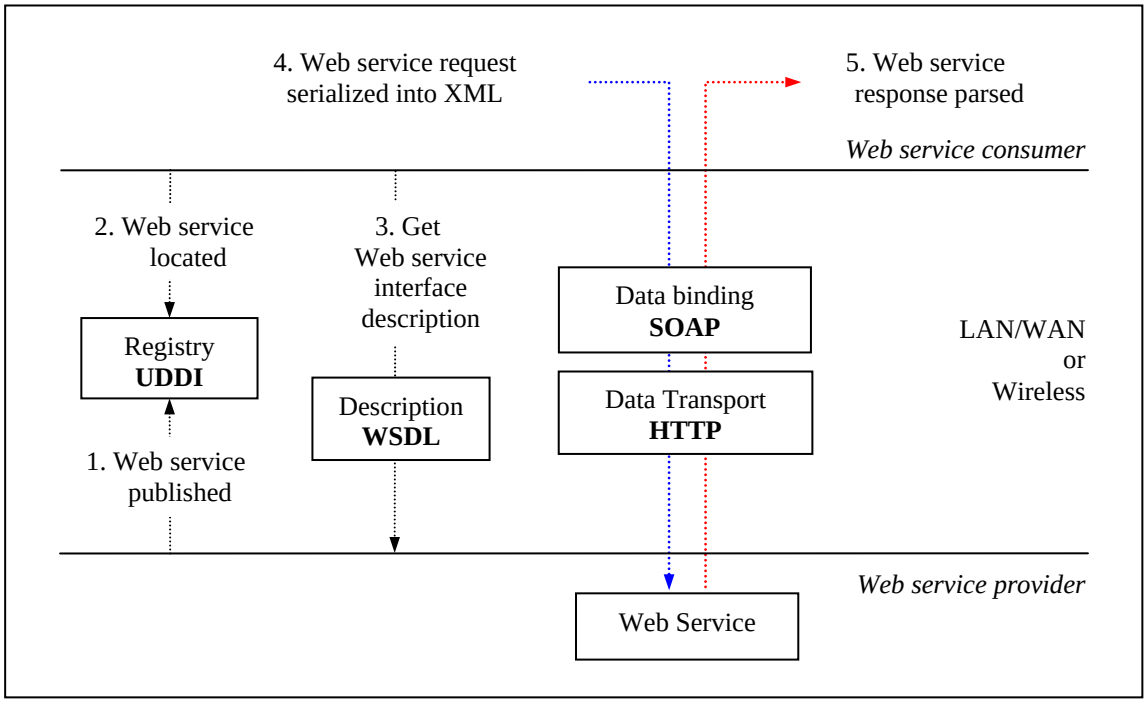


Fig. 2 Web services lifecycle.

The Web service lifecycle is based on the concept of bind–once, consume–many. Thus, once a Web service has been located (2) and software written to consume it according to its interface description (3), the software sends a request (4) to a Web service and waits for a response (5) synchronously or asynchronously. Steps (4) and (5) are repeated as many times as required. For every request, raw data is serialized (or encoded) into XML format and bound in a SOAP envelope, which in turn is conveyed over the Internet (or network) using HTTP and TCP/IP.

A topic of contention that is raised frequently with Web services is that of security — that is, the unauthorized viewing and/or manipulation of the transported message. Standards for the encryption of the data transport layer, such as SSL/TLS for point–to–point security already exist. Since these standards only support end–to–end security, encryption of the data itself must be undertaken

if information is to be kept secure. To support this need, a draft recommendation exists for XML encryption [W3C, 2002], and relies on client–server encryption algorithms and symmetric/asymmetric keys schemes. Under the XML encryption standard, arbitrary data, XML elements, and/or XML element content can be secured from unauthorised viewing. Additional checks can be made at each end to verify whether an encrypted stream has been manipulated or not.

In the interest of monitoring and charging for Web service usage in a commercial environment, transaction monitoring procedures may be easily implemented. In addition, authentication procedures may be applied so that a Web service can be protected from unauthorised usage when publicly exposed over the Internet and internal networks. XML parsers can be implemented at both the client and server ends of a Web service connection to validate the well formed–ness of the SOAP message. This reduces the risk of a corrupted message from being sent or received undetected.

3 Low–Cost GPS Receivers

For the purposes of this paper, low-cost GPS receivers are regarded as those typically used for general purpose positioning and navigation, low accuracy data capture, reconnaissance, bushwalking, marine and in-car navigation, and the like. The accuracy and precision obtainable from low-cost GPS receivers, following the discontinuation of Selective Availability (SA), is well known (Satirapod et al, 2001, for example). With an unobstructed sky, low-cost GPS receivers have an expected horizontal and vertical positioning precision in the range of 10-20 metres and 20-30 metres respectively. After applying local DGPS corrections, many spatially correlated (or systematic) errors can be removed and hence, it is not unreasonable to expect some low-cost GPS receivers to achieve a horizontal positioning accuracy better than 5 metres.

A question that is often raised is: "*How can low–cost GPS receivers be augmented with DGPS corrections via Web services rather than by conventional radio transmission*?" A practicable means for augmenting low–cost GPS receivers using the Web services technique is via custom software which is able to read the output GPS measurement information. Low–cost GPS receivers typically output GPS information in one or more proprietary protocols, depending on the receiver make and chipset. Currently, a majority of low–cost GPS receivers provide positioning information by way of the National Marine Electronics Association (NMEA) standard. Under the NMEA–0183 standard, elementary positioning and navigation information, such as position, velocity, direction and observed satellites, is provided through ASCII sentences. In the context of Location Based Services (LBS) applications, this protocol is sufficient for low–accuracy (or unaugmented) positioning requirements. However, where custom augmentation algorithms are to be implemented to improve the accuracy and precision of the autonomous positioning solution, a more extensive protocol (i.e. one that can provide the original code and/or carrier–phase observations) is required.

The low–cost GPS receiver used in this investigation contains a SiRF chipset and outputs GPS information in the SiRF Binary protocol (SiRF, 2004). The SiRF Binary protocol provides an extensive list of input and output messages concerning the original measurements, the positioning and navigation solution, and commands for configuring the GPS chipset. For instance, the SiRF binary protocol supports the output of pseudorange and (integrated) carrier–phase measurements, subframes 1 to 5 of ICD–GPS–200C, receiver clock bias and drift, and various other observation and navigation messages (SiRF, 2004). The SiRF protocol also provides the ability for receivers to be supplied with DGPS corrections from either the US WAAS or broadcasting beacons in RTCM SC–104 format.

Although a (low–cost) SiRF–based receiver has been used for this investigation, it can be shown that any GPS–enabled device capable of reaching the Internet (wirelessly or not) can be used to take advantage of a Web services augmentation system. It follows that the "right" protocol for custom software to implement DGPS corrections via a Web services augmentation system ultimately depends on the mathematical model embodied within the method of augmentation. Since the method of augmentation used in this paper is based on correcting pseudoranges, any receiver which outputs time–stamped pseudorange measurements could be used.

4 Correlated Errors Accounted for by DGPS Corrections

As discussed previously, to improve the accuracy of real time autonomous GPS positioning, DGPS corrections obtained from a local reference station (or CORS network) may be applied. In principle, the technique of DGPS minimises the influence of errors which are spatially correlated between two or more simultaneously operating receivers. Differential pseudorange corrections are computed by (stationary) receivers located over known reference marks and represent the difference between a computed range and (measured) code pseudorange. Accordingly, pseudorange corrections quantify the impact of the errors affecting the code range measurement at that specific location and time. After removing the influence of random and site specific errors contained within the reference receiver's measurements (i.e. receiver noise, receiver clock error and multipath),

the pseudorange corrections are then transmitted and applied to nearby (roving) receivers to yield a better estimate of their position.

In general, the spatially correlated GPS errors for which (differential) pseudorange corrections can compensate include satellite clock errors, broadcast ephemeris errors, and signal propagation errors (i.e. ionospheric and tropospheric refraction). Over short distances in real time (or after a short latency of around a few seconds), the effect of these errors can be greatly reduced by DGPS techniques. However, as the distance between the roving and reference station receivers increases, there is a natural degradation in the level of correlation and thus the accuracy of the pseudorange correction diminishes. Furthermore, since the correlated errors are subject to temporal variation (i.e. by changes in the atmosphere, satellite orbit deviation and clock offset), the level of correlation may be further reduced as the time separating the observations increases. Experience has shown that for low-cost GPS applications, the decorrelation of errors according to time may be neglected up to a period of about 10-20 seconds without significantly reducing the accuracy.

The DGPS corrections disseminated in the prototype discussed in this paper are based on the RTCM SC–104 (version 2.3) Type 1 message and include a pseudorange and range rate correction for each measurement at time (t). At the receiver, the corrected pseudorange $R^j{}_B(t)_{corr}$ is obtained from (Hofmann–Wellenhof et al, 2001):

$$R^j{}_B(t)_{corr} = R^j{}_B(t) + PRC^j(t) \qquad (1)$$

where $PRC^j(t)$ is the predicted differential pseudorange correction and is given by:

$$PRC^j(t) = PRC^j(t_0) + RRC^j(t_0)(t - t_0) \qquad (2)$$

where $PRC^j(t_0)$ and $RRC^j(t_0)$ are the pseudorange and range rate corrections observed at the reference station at time (t_0) and t is the time of measurement at the roving receiver.

5 Prototype for a Web Services DGPS System

To demonstrate the feasibility of a Web services DGPS system, a simple prototype has been developed. The prototype is comprised of three components: a GPS and Web enabled client (PDA), a server-enabled PC, and a single CORS receiver. Two Web services are hosted on the server — (1) a CORS Web Service for logging the reference station receiver–determined pseudorange corrections and (2) a DGPS Web Service for disseminating those corrections to roving receivers. Accordingly, software has been developed for the PDA to facilitate the request, response and application of DGPS corrections. The architecture of the prototype is given in Figure 3.

The specifications of the prototype components developed for this paper are given in Table 1.

Tab. 1 Prototype component specifications

Component	Specifications
PDA (client)	iPAQ 3850, Bluetooth expansion pack, custom software
Low–cost GPS receiver	Pretec CompactGPS (SiRF Binary & NMEA 0183)
Mobile Phone	Ericcson Bluetooth phone (GPRS-only internet account)
PC (Web server)	2.0 GHz Pentium 4, MS W2K/IIS 5.0/ASP.NET, custom software
CORS receiver	Leica GPS System 500

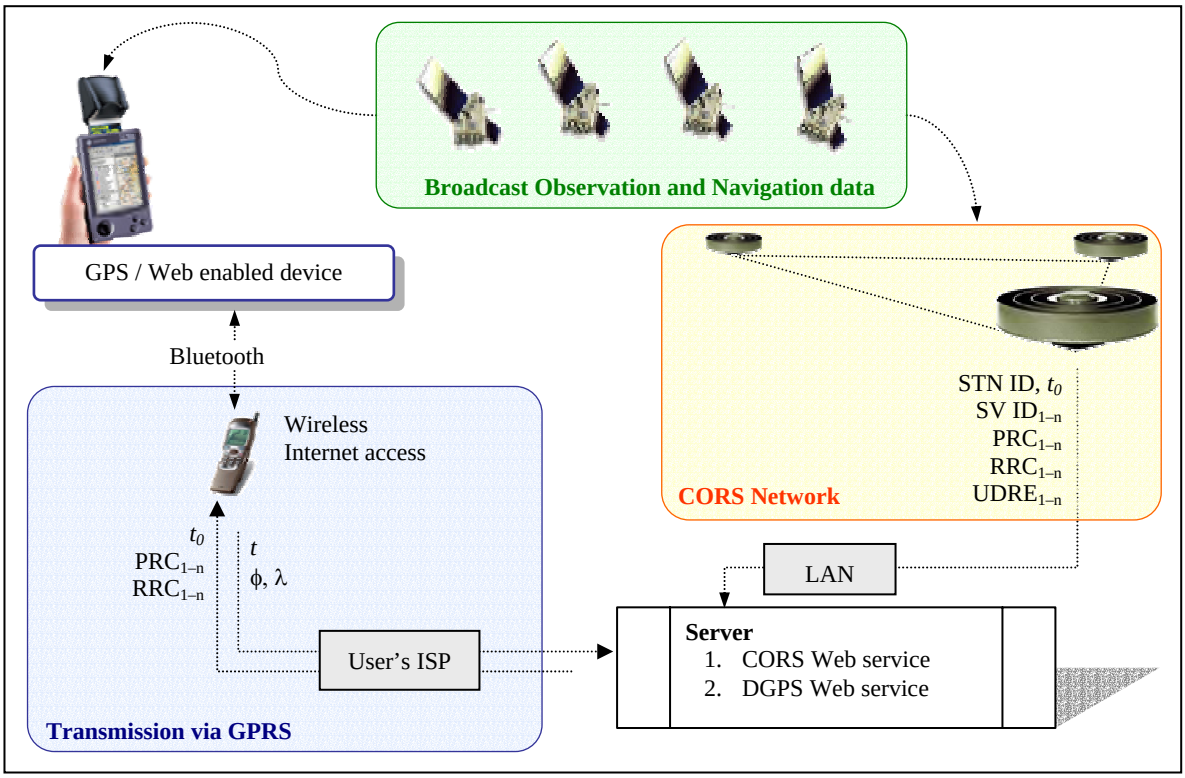


Fig. 3 Architecture of the Web services DGPS augmentation prototype

5.1 CORS Network and Logging software

For the purposes of demonstration, only a single CORS receiver was used for the prototype. The reference station receiver was configured to output a Type 1 RTCM message (Differential DGPS corrections) via a RS232 serial port every second. Software was developed to read the output RTCM messages and in turn send the corrections to the CORS Web service. Figure 4 shows the flow of tasks performed by the reference station software.

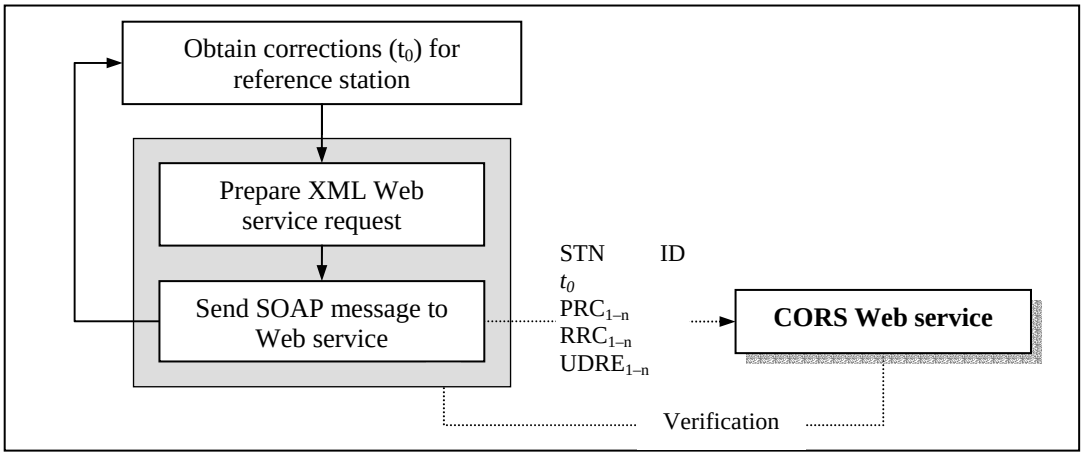


Fig. 4 Flow of tasks performed by the reference station receiver software

5.2 Client software

The client (or mobile device) software has been developed to perform three basic tasks:

1. Obtain the smoothed pseudoranges and position from the SiRF protocol
2. Handle the request and response of differential GPS corrections
3. Apply the pseudorange corrections and display the augmented GPS position

To obtain a set of DGPS corrections, GPS time t and an approximate position (ϕ, λ) are sent to the Web service. Each request is made asynchronously to allow the client to continue operating without waiting for a response. When a response is received, a new position is computed from the corrected pseudoranges via equations (1) and (2). Prior to making a call to the DGPS Web service, the client verifies whether an internet connection has been established. Figure 5 shows the general flow of the prototype client software.

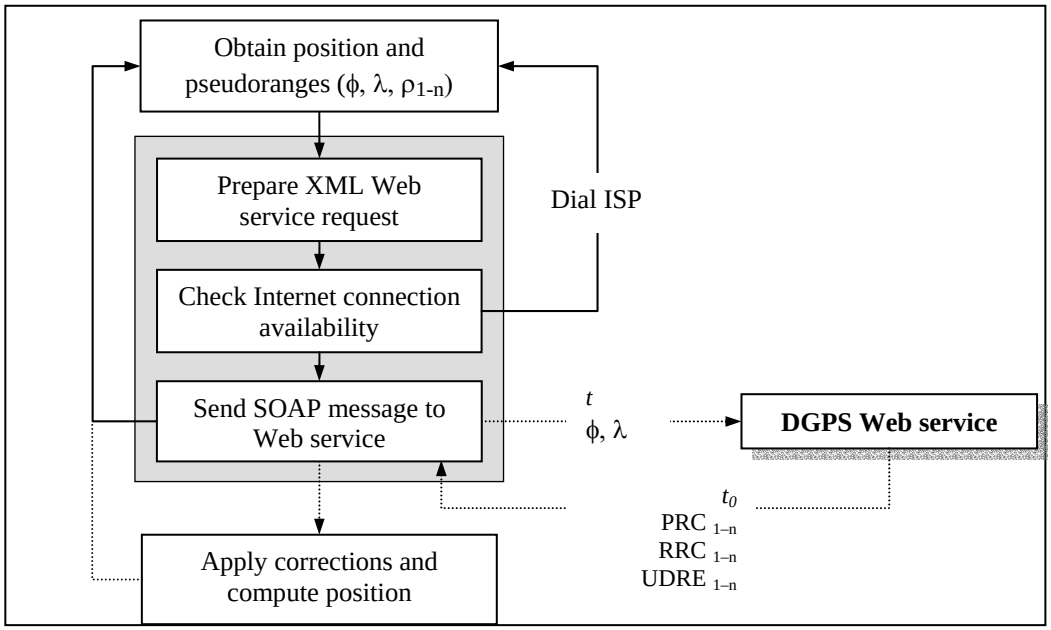


Fig. 5 Flow of tasks performed by the client software

5.3 Differential GPS and CORS Web Services

To provide real time DGPS augmentation, the two Web services (hosted on the server) operate independently to facilitate the logging and dissemination of real time corrections as described in §5.1 and §5.2. Firstly, when sent by the custom reference station software, the computed pseudorange corrections are logged by the CORS Web service according to the GPS time and satellite ID. Secondly, when and as requested by the roving client, the DGPS Web service retrieves the pseudorange corrections based on the GPS time and sends that data to the roving client.

Although the client's approximate position is sent to the DGPS Web service, it is not used in this prototype. The purpose of sending the client's position is simply to illustrate that, given an appropriate mathematical model for multi–station interpolation, the DGPS Web service could be extended to provide a set of location–centric pseudorange corrections. This is discussed further in §7.

6. Preliminary Test Results

6.1 Latency in data Transmission

Critical to real and near real time applications is the delay (or latency) in the time taken for a DGPS correction to be received by the rover. In contrast to the streaming technique, the latency considered in this paper is the time taken for (1) the request to be sent to the DGPS Web service, (2) a correction to be determined and (3) the corresponding corrections to be sent back to the client. Figure 6 shows the latencies in transmission of DGPS corrections during peak and off–peak periods using a mobile phone and a GPRS connection.

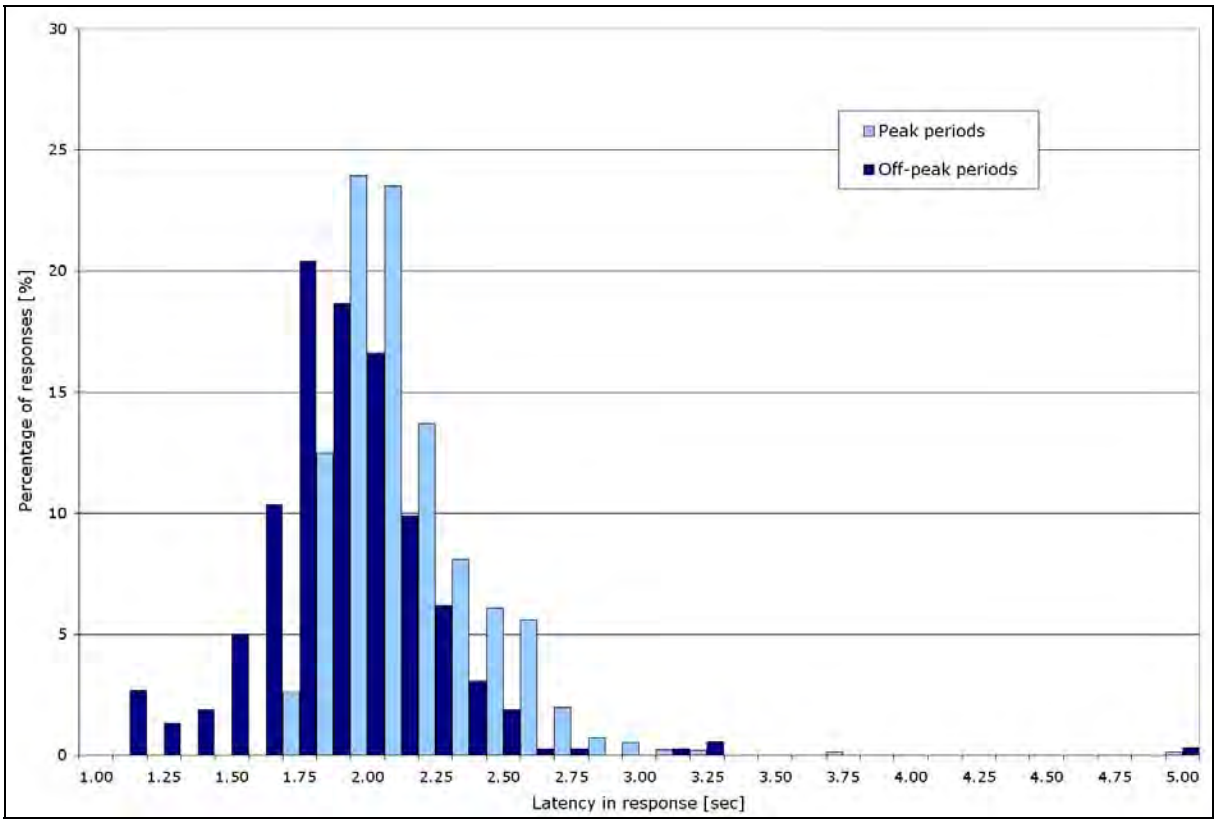


Fig. 6 Latency in request/response of DGPS pseudorange corrections for 6 satellites using a mobile phone (over GPRS) during peak and off–peak periods

The average latencies from various tests show a mean of 1.86 seconds during off–peak periods and 2.12 seconds during peak periods. These results are based on a series of tests conducted over two weeks and are representative of the Melbourne CBD environment only. In all cases, none of the messages were found to be corrupted or degraded in any way. The latencies associated with the sending of reference station information to the CORS Web service (over a 1.5 Mb/s connection) is almost instantaneous and such, is not presented herein.

Previous research and development in Internet–based DGPS systems using data streaming has indicated similar results (Weber et al, 2004, for example). However, the (two–way) Web services architecture generally shows greater latencies over those achieved through (one–way) Internet streaming. Notwithstanding the influence of bandwidth, network congestion, mobile phone infrastructure performance, server performance and the like, the primary cause of the additional latency comes as a result of XML tag bloat and sending data in plain text. If each message were compressed into binary format, the latencies could be reduced. The latency may be further reduced if the entire message was sent as a character delimited string. However, sending multiple data elements in a single string would be contrary to the intention of the OGC and W3C interoperability standards.

6.2 DGPS Corrections

Table 2 shows the positioning accuracy results achieved by the DGPS Web service prototype described in this paper. The results show the difference between the average (or expected) values and standard deviations obtained from the autonomous GPS and DGPS-corrected positions over 30 minutes (1 second observations) with a 15° elevation mask.

As shown by Table 2, the effect of the spatially correlated GPS errors is greatly reduced by the DGPS technique. However, as indicated by the standard deviations in both GPS and DGPS estimates, the DGPS-corrected positions still suffer from large random errors.

Tab. 2 Comparison of results from Autonomous GPS and DGPS–corrected positions.

		Easting	Northing	E height	ΔE	ΔN	Δh
GPS	μ	320469.61	5814432.62	79.52	2.34	–0.43	16.17
	σ	1.73	2.56	3.74			
DGPS	μ	320466.70	5814432.98	66.13	-0.57	-0.07	2.77
	σ	1.67	2.58	3.64			

7 Web Services for Other GPS Applications

This paper has focussed on the single baseline (differential) pseudorange correction technique for augmenting autonomous GPS positions. However, the principle of the Web services architecture may be used for exchanging a variety of GNSS related data, and applied to other GNSS positioning applications. Typical applications may include multi–station relative positioning, the dissemination of satellite orbit and clock corrections for Precise Point Positioning (PPP), and the monitoring of CORS network and receiver performance.

The one–way streaming of data over the Internet has been shown by other research to be a satisfactory method of disseminating DGPS corrections. As described in §6, results from this prototype indicate that Internet streaming offers marginally better latencies than the Web services architecture. However, where the method of augmentation requires the specific location of the roving receiver relative to one or more reference stations, two–way communications is needed. For instance, the interpolation of the differential correction for a rover based on the nearest surrounding CORS receivers requires the rover's approximate location. As indicated by the prototype, the Web services architecture facilitates applications of this nature by enabling the user to supply measurements of any type, such as approximate position, observed pseudoranges and/or carrier–phase measurements. In the interest of multi–station RTK and Virtual Reference Station (VRS) applications, Web services can provide an efficient, secure and reliable source of platform–independent (or receiver independent) communication.

In certain environments where the effects of spatial decorrelation are at relatively high levels, the differential correction approach described here may provide little or no improvement to positioning accuracy. To improve positioning quality in this instance, a practicable alternative is the integration of precise satellite orbit and clock parameters, rather than the less–precise broadcast ephemeris commonly used in the positioning solution. It is well known that, given the availability of precise orbit and clock parameters from IGS, the influence of broadcast ephemeris errors on an absolute positioning solution can be significantly reduced. Though not included in this paper, in conjunction with the developments undertaken for this prototype, a Web service for downloading and disseminating SP3 ultra rapid orbit information has also been developed. Whilst it has not been implemented as yet, this work indicates that an "ultra–rapid clock and orbit" Web service would be a viable option for Web–enabled receivers located in areas where local CORS network corrections are unavailable.

Finally, Web services offer significant advantages for platform independent monitoring of CORS network receivers. Notwithstanding the fact that many systems have long been in place for exchanging data between reference station receivers, Web services offer the ability to integrate receiver types of any make, provided that there is a means for sending data over TCP/IP. It follows that Web services can be used to interact with (or configure) remote GPS receivers from a central processing system securely from any location, wirelessly or not. Experience has shown that Web services communications over a (wired) 1.5 Mb/s Internet connection is almost instantaneous, showing that multi–reference station application performance would not be significantly degraded.

8 Discussion and Conclusion

This paper shows that XML Web services are a simple, efficient and secure means for exchanging GPS augmentation information over the Internet. Web services offer some considerable advantages. Web services offer a platform independent way of augmenting GPS and Web enabled mobile devices capable of reaching the Internet (wirelessly or not) with correction information from existing CORS networks. It follows that software to consume a Web service can be developed by almost any programming language. Almost any hardware device capable of sending data over TCP/IP may be used. Since the Web services architecture allows for two–way communication, location–centric applications (i.e. multi–reference station approach) can benefit from this method of Internet augmentation. Web services can further provide opportunities for implementing encryption

algorithms, secure transaction monitoring, and user authentication. Being a transactional service, the user pays only for what data is requested and received. Hence, where DGPS corrections are shown to be stable over longer periods (i.e. 10–20 seconds), the user may opt to restrict the number of calls made to the Web service. Typical latencies for Web services communications can, in general, be expected to decline as mobile communications infrastructure is upgraded.

However, the exchange of GPS augmentation information via the Web services architecture also has its disadvantages. Strictly speaking, Web services architecture introduces larger latencies than streaming methods due to XML–message bloat. Since the Web services architecture disseminates information via two–way "request and consume" communications, additional latencies are introduced. For sub–second applications, Web services offer less than optimal performance over wireless connections. Whilst not necessarily restricted to the Web services approach, the main disadvantages come as a result of:

- Poor wireless/mobile coverage across rural areas which are dominant throughout Australia.
- Network congestion has the potential for introducing large latencies, even for simple data transmission.
- There can be no guarantee that a data packet will arrive by a certain time, if at all.

Acknowledgements

The authors would like to gratefully acknowledge the scholarship funding provided by Natural Resources, Mines and Energy (NRM&E, QLD) and the department of Geomatics, University of Melbourne. All hardware and software development tools used in this investigation have also been provided by NRM&E and the Department of Geomatics.

References

Gao Y and Liu Z (2001): ***Differential GPS Positioning over Internet***. Journal of Geospatial Engineering, Vol. 3, No. 1, 1–7.

Hofmann–Wellenhof B, Lichtenegger H and Collins J (2001): ***GPS Theory and Practice (5th Ed., revised)***. Springer–Verlag Wien, New York.

Kechine MO, Tiberius CCJM and van der Marel H (2003): ***Network Differential GPS: Kinematic Positioning with NASA's Internet-based Global Differential GPS***. Journal of Global Positioning Systems, Vol. 2, No. 2, 139–143

Lenz E (2004): ***Networked Transport of RTCM via Internet Protocol (NTRIP)***. Proceedings of the FIG Working Week 2004, Athens, Greece, 22–27 May.

Monteiro LS, Moore T and Hill CJ (2004): ***What is the Accuracy of DGPS?*** Presented at the European Navigation Conference GNSS 2004, Rotterdam, The Netherlands, May.

Muellerschoen R J, Bertiger WI and Lough M (2000): ***Results of an Internet-Based Dual-Frequency Global Differential GPS System***. Proceedings of the IAIN World Congress, San Diego, CA, June.

RTCM SC–104, (2001): ***RTCM Recommended Standards for Differential GNSS Service.*** Version 2.3, 20 August, Radio Technical Commission for Maritime Services, Alexandria, Virginia.

Satirapod C, Rizos C and Wang J (2001): ***GPS Single Point Positioning With SA Off: How Accurate Can We Get?*** Survey Review, 36(282), 255-262.

SiRF (2004): ***SiRF Binary Protocol Reference Manual***. SiRF Technology, San Jose, CA, Revision 1.1, February 2004.

Weber G, Dettmering D and Gebhard H (2004): ***Networked Transport of RTCM via Internet Protocol (NTRIP).*** unpublished, Available online: http://igs.ifag.de/pdf/NtripPaper.pdf.

W3C (2002): ***XML Encryption Syntax and Processing – W3C Recommendation***. Eastlake D, Reagle J (Eds), 10 December, Available online: http://www.w3.org/TR/xmlenc-core/.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 95-100

Journal of Global Positioning Systems
2004
© CPGPS

Performance Analysis of Precise Point Positioning Using Rea-Time Orbit and Clock Products

Yang Gao

Department of Geomatics Engineering, University of Calgary, Calgary, AB, Canada
e-mail: gao@geomatics.ucalgary.ca Tel: +1-403-2206174 ; Fax: +1-403-2841980

Kongzhe Chen

Department of Geomatics Engineering, University of Calgary, Calgary, AB, Canada
e-mail: kzchen@ucalgary.ca Tel: +1-403-2204916 ; Fax: +1-403-2841980

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. The real-time availability of precise GPS satellite orbit and clock products has enabled the development of a novel positioning methodology known as precise point positioning (PPP). Based on the processing of un-differenced pseudorange and carrier phase observations from a single GPS receiver, positioning solutions with centimeter to decimeter accuracy can be attained globally. Such accuracy can currently be achieved only through differential processing of observations acquired simultaneously from at least two receiver stations. The potential impact of PPP on the positioning community is expected to be significant. It brings not only great flexibility to field operations but also reduces labor and equipment cost and simplifies operational logistics by eliminating the need for base stations. This paper will address issues related to precise point positioning and perform data analysis to assess the performance of different application solutions from PPP using real-time precise orbit and clock corrections. They include the discussion of an algorithm for un-differenced data processing, error source and mitigation, and critical elements related to real-time GPS orbit and clock products. Numerical results will be presented to show the positioning accuracy attained with datasets acquired from different environments using real-time precise orbit/clock products currently available. Features of a software package that has been developed at the University of Calgary for precise point positioning will also be described.

Key words: GPS, Precise Point Positioning, Un-differenced, Precise Orbit and Clock

1 Introduction

Current carrier phase based GPS kinematic positioning systems are primarily based on double differencing data processing approach which is able to provide centimetre to decimetre accurate positional accuracy in real-time. It has found wide applications from geodetic survey, mapping, resources exploration, deformation monitoring construction to aircraft landing. The differential process however requires simultaneous observation of common GPS satellites at both a base station (a reference site with precisely know coordinates) and rover user stations. This not only complicates the data acquisition process but also reduces the applicability of the method to many other potential applications. Since the reduction of common errors is very much dependent on the inter-station baseline lengths, the base and rover station separation must be short typically in the range of about 20 kilometres. Further, the need for a base station would increase the cost in equipment and labour and inconsistency using different base stations.

The availability of precise GPS satellite orbit and clock products has enabled the development of a novel positioning methodology known as precise point positioning (PPP). Based on the processing of un-differenced pseudorange and carrier phase observations from a single GPS receiver, this approach effectively eliminates the inter-station limitations introduced by differential GPS processing as no base station is necessary. As a result, it offers an alternative to differential GPS that is logistically simpler and almost as accurate (Zumberge et al., 1997; Kouba & H¨¦roux,

2001). Although PPP does not require any base station, it requires accurate knowledge of the GPS satellite coordinates and the state of their clocks.

The performance of PPP for positioning determination has been demonstrated in various papers, e.g. Zumberge et al., 1997; Kouba & H¨¦roux, 2001; Gao and Shen, 2002; Gao et al., 2003, using post-mission precise orbit and clock from IGS. The potential impact of PPP on the positioning community is expected to be significant. It brings not only great flexibility to field operations but also reduces labor and equipment cost and simplifies operational logistics by eliminating the need for base stations. Following the availability of real-time precise GPS satellite orbit and clock products from some organizations, the interest to apply PPP to real-time kinematic positioning is currently strong as a next generation real-time kinematic (RTK) methodology.

This paper will address issues related to precise point positioning and conduct data analysis to assess performance of different application solutions from PPP using real-time precise orbit and clock corrections. They include the discussion of an algorithm for un-differenced data processing, error source and mitigation, and critical elements related to real-time GPS orbit and clock products. Numerical results will be presented to show the positioning accuracy attained with datasets acquired from different environments using real-time precise orbit/clock products currently available. Features of a software package that has been developed at the University of Calgary for precise point positioning will also be described.

2 Precise Point Positioning (PPP) Method

In the following, the method of PPP is described along with mathematical equations. With a dual-frequency GPS receiver, the following ionosphere-free combinations can be applied to facilitate PPP positioning using un-differenced observations.

$$P_{IF} = \frac{f_1^2 \cdot P_1 - f_2^2 \cdot P_2}{f_1^2 - f_2^2} = \rho + cdt + d_{trop} + dm_{IF} + \varepsilon(P_{IF}) \quad (1)$$

$$\begin{aligned} \Phi_{IF} &= \frac{f_1^2 \cdot \Phi_1 - f_2^2 \cdot \Phi_2}{f_1^2 - f_2^2} \\ &= \rho + cdt + d_{trop} + \frac{cf_1 N_1 - cf_2 N_2}{f_1^2 - f_2^2} + \delta m_{IF} + \varepsilon(\Phi_{IF}) \end{aligned} \quad (2)$$

where P_i is the measured pseudorange on L_i (m); Φ_i is the measured carrier phase on L_i (m); ρ is the true geometric range (m); c is the speed of light (m/s); dt is the receiver clock error (s); d_{trop} is the tropospheric delay (m); f_i is the frequency of L_i (m); N_i is the integer phase ambiguity on L_i (cycle); dm_i is the multipath effect in the measured pseudorange on L_i (m); δm_i is the multipath effect in the measured carrier phase on L_i (m) and $\varepsilon(.)$ is the measurement noise (m).

Satellite orbit and clock errors are not present in equation (1) and (2) since they can be removed by the use of precise orbit and clock products. The remaining receiver clock and tropospheric delays in equations (1) and (2) are to be estimated in PPP. A choke-ring antenna should be used in the presence of significant multipath.

The estimation of tropospheric gradients is beneficial for both GPS positioning and tropospheric delay estimation (Bar-sever *et al.*, 1998). The following equation can be used to model the tropospheric effect (McCarthy and Petit, 2003):

$$\begin{aligned} d_{trop} &= m_h(e) D_{hz} + m_w(e) D_{wz} \\ &+ m_g(e)[G_N \cos(a) + G_E \sin(a)] \end{aligned} \quad (3)$$

where D_{hz}, D_{wz} are the zenith hydrostatic and wet delay; G_N, G_E are the horizontal delay gradient in north and east direct; $m_h(e)$ is the hydrostatic mapping function; $m_w(e)$ is the wet mapping function and $m_g(e)$ is the gradient mapping function; a, e are the azimuth and elevation angles.

In this research the following gradient mapping function has been used (Chen and Herring, 1997):

$$m_g(e) = \frac{1}{\sin(e)\tan(e) + 0.0032} \quad (4)$$

and the Saastamoinen model has been applied to model the zenith hydrostatic delay (McCarthy and Petit, 2003):

$$D_{hz} = \frac{0.0022768\ P_0}{1 - 0.00266\cos 2\phi - 0.00028\ H} \quad (5)$$

where P_0 is the pressure in millibars; ϕ is the latitude and H is the height above the geoid (km).

The unknown vector in the PPP processing include three position coordinate parameters, a receiver clock offset parameter, a wet zenith tropospheric delay parameter, two tropospheric gradient parameters and float ambiguity terms in ionosphere-free combinations (equal to the number of satellites used in estimation).

3 Real-Time Precise GPS Orbits and Clocks

List in Table 1 is the source of precise orbit and clock products from IGS and other organizations. We notice that only JPL and NRCan are currently providing real-time precise orbit and clock data, known as IGDG and GPS•C respectively. The precise orbit and clock data from JPL is generated based on data from a network consisting of about 70 globally distributed reference stations and their accuracy are about 20 cm for orbits and 0.5 ns for clocks. Its latency is about 4 seconds and the date interval is 1 second (Muellerschoen, 2003). The precise orbit and clock data from NRCan is generated based on data from a network consisting of about 20 globally distributed reference stations with accuracy for orbits about 10 cm and clocks about 1 ns respectively. Still under development, the data latency for NRCan's precise orbit and clock is at the level of several hours and the update interval is 2 seconds (Héroux, 2004).

JPL real-time precise orbit and clock data is now available for commercial applications and will be used in this paper to assess the performance of different application solutions from PPP. JPL IGDG real-time precise orbit and clock corrections were acquired over Internet from a JPL server at a rate of 1 Hz.

Tab. 1 Precise orbit and clock accuracy
(unit: cm for orbit and ns for clock)

Sources	Accuracy		Latency	Update	Interval	
	Orbit	Clock			Orbit	Clock
IGS Final	<5	<0.1	13 days	Weekly	15 min	5 min
IGS Rapid	<5	0.1	17 hours	Daily	15 min	5 min
IGS UltraEST	<5	0.2	3 hours	12 hours	15 min	15 min
IGS UltraPRD	10	5	None	12 hours	15 min	15 min
IGDG (Global)	20	0.5	~4 sec	1 sec	29 sec	1 sec
GPS•C (Global)	20	1	~8 hours	2 sec	20 sec	2 sec

4 Numerical Results and Analysis

In the following, data processing and analysis are conducted to assess performance of different application solutions from PPP using JPL real-time precise orbit and clock corrections. Results in different positioning modes and other application solutions including receiver clock offset and water vapor estimates are presented.

P3 Software

A software package called P^3 has been developed at the University of Calgary for precise point positioning that runs on Microsoft Windows operating system family. The software is able to output solutions of different application parameters including position, zenith tropospheric delay and receiver clock offset estimates. Processing can be done in post mission or in real-time, and the program can be run in either static or kinematic mode. Backward processing is supported to reduce errors associated with solution convergence. A sample screenshot of the software during processing is shown in Fig. 1.

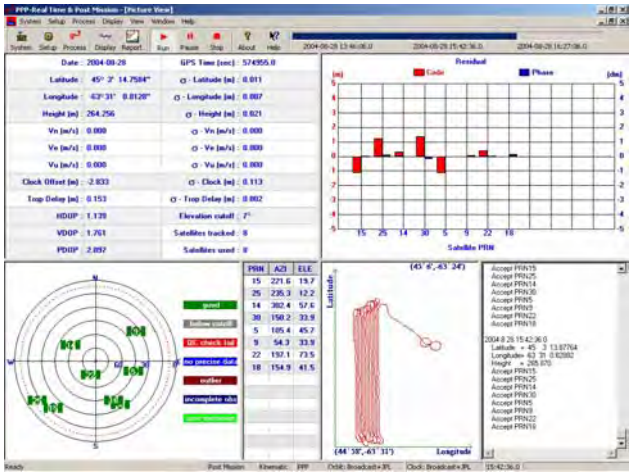


Fig. 1 P3 interface

Static Control Survey

In this test, one day of GPS data acquired on August 4, 2004 at IGS station ALGO was processed. The data from a IGS station was selected because the coordinates of all IGS stations are precisely determined everyday with respect to ITRF2000 which is also the reference frame that has been used by JPL in the generation of real-time orbit and clock corrections. The GPS data at ALGO and the station coordinates were downloaded from the IGS website while the JPL real-time corrections were retrieved from JPL server. The position results are shown in Figure 2 and the accuracy statistics is given in Table 2. It is seen that the coordinate estimates could converge to centimetre level within 20 minutes. After the convergence, all position coordinate components are accurate at sub-centimetre level. The results in Table 1 indicate that PPP is capable of providing real-time centimetre level accuracy for static control survey.

Tab. 2 Static positioning accuracy

	RMS (m)	BIAS (m)	STD (m)
Latitude	0.009	0.008	0.003
Longitude	0.010	0.003	0.009
Height	0.007	0.000	0.007

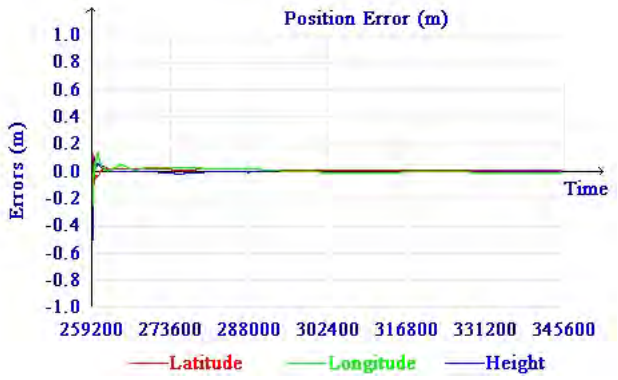


Fig. 2 Static positioning errors

Vehicle Kinematic Positioning

In this test, a kinematic positioning with a vehicle was conducted on September 30th, 2003. The vehicle was driven along the highway at a speed of 80 km/h near Springbank, Alberta. In order to establish a reference trajectory for the vehicle, a reference receiver was set up at one control point of the Springbank baseline network so double difference data processing could be performed to establish a reference for accuracy assessment. Both the control point and vehicle used a Javad Legacy dual-frequency receiver with the same type of antenna. A CDPD radio was used to receive JPL IGDG real-time precise orbit and clock corrections via the Internet. The sample rate of the two GPS receivers was set to 1 Hz. The PPP solutions are obtained using P3 software while the double difference solutions are obtained using a commercial software package from Waypoint Consulting Inc. With a relatively short baseline length (about 7km), the ambiguity-fixed results were available and can be served as the ground-truth to assess the positioning accuracy of PPP solutions.

The positioning differences between PPP and double difference solutions are shown in Figure 3 and the accuracy statistics is given in Table 2. They indicate that centimetre accurate positioning results have been obtained in real-time using precise point positioning method.

Tab. 3 Vehicle kinematic positioning accuracy

	RMS (m)	BIAS (m)	STD (m)
Latitude	0.009	0.008	0.003
Longitude	0.010	0.003	0.009
Height	0.007	0.000	0.007

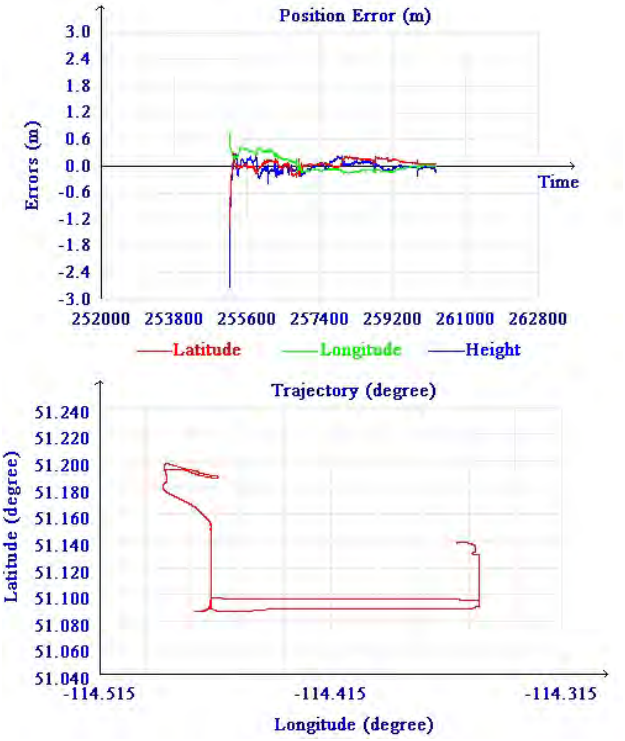


Fig. 3 Vehicle trajectory and positioning errors

Airborne Kinematic Positioning

The airborne dataset was collected on August 28, 2004 at 40 kilometers north of Halifax, Nova Scotia. A Novatel GPS receiver (Black Diamond) and antenna (model 512) were set up on a helicopter. Another Novatel DL-4 receiver and antenna with ground plane were served as base station. The sample rate of the two GPS receivers was 1 Hz. The helicopter was typically flying at an altitude of 250 meters above ground level at 50 knots. The distance between the rover and base is less than 10 kilometers. The double-differenced with ambiguity-fixed trajectory is served as ground-truth.

As shown in Figure 4 and Table 4, centimetre accurate positioning results have been achieved in using real-time precise orbit and clock products and point positioning method.

Tab. 4 Aircraft positioning accuracy

	RMS (m)	BIAS (m)	STD (m)
Latitude	0.009	0.008	0.003
Longitude	0.010	0.003	0.009
Height	0.007	0.000	0.007

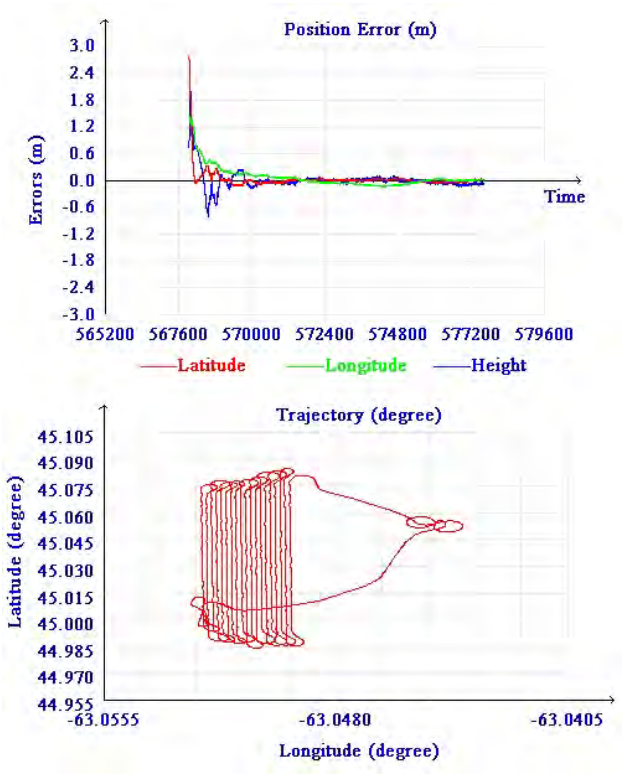


Fig. 4 Aircraft trajectory and positioning errors

Receiver Clock Estimation

In addition position determination, PPP can also output receiver clock offset solution which has the potential to support precise timing applications. Since JPL IGDG corrections are generated using a high-precision clock at IGS station AMC2 (equipped with a hydrogen maser external frequency) as the reference clock (Muellerschoen, 2003), we can assess the accuracy of receiver clock offset estimation from our PPP by processing the GPS data from AMC2. The resultant receiver clock estimates from PPP solutions for AMC2 station should theoretically equal zero using JPL IGDG precise orbit and clock corrections and the variations in the solutions directly reflect the quality of the clock solutions from PPP.

In this test, the receiver clock offset was estimated as white noise using GPS data from ACM2 station acquired on June 12, 2004. Shown in Figure 5 are receiver clock offset estimates at ACM2 station. Table 5 provides the statistics of the estimation accuracy. The results indicate that PPP is capable of providing real-time sub-nanosecond accurate receiver clock estimates as a promising tool for time transfer. In order to use the estimates for clock comparisons, all instrumental biases should be calibrated (Petit *et al.*, 2001). Special cables that are less temperature sensitive may be required (Larson *et al.*, 2000).

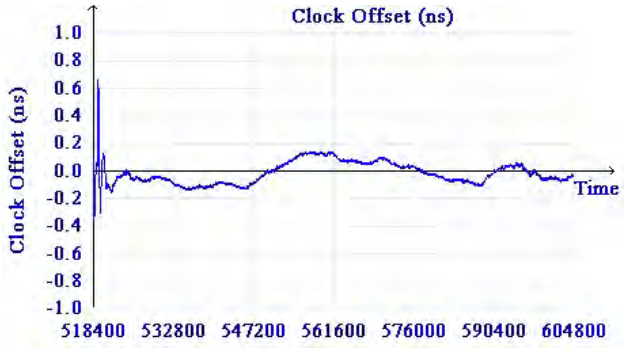


Fig. 5 Receiver clock offset estimates

Tab. 5 Receiver clock offset estimation accuracy

Products	RMS (ns)	BIAS (ns)	STD (ns)
JPL IGDG	0.077	0.018	0.075

Water Vapor Estimation

In this test, a Javad JPSLEGANT antenna was set up on a pillar on the roof of the Engineering Building at the University of Calgary with precisely known coordinates. JPSLEGANT is an antenna with a flat ground plane so it can partly mitigate the multipath effects. A GPS data acquisition at a sampling interval of 10 seconds was conducted on September 5th 2004. For performance analysis, a Radiometrics™ 1100 water vapour radiometer (WVR) (Radiometrics Corp.) and a Paroscientific™ MET3A meteorological sensor located on the same roof have been applied to provide "true" precipitable water vapor (PWV) and pressure measurements. The radiometer was set up to make direct measurements of line-of-sight slant water vapor to all GPS satellites. Since the WVR tracks each satellite for approximately 40 seconds and consequently it takes about 6 minutes to track all satellites in view in a given cycle, the PWV measurements for each 6-minute cycle of observations were averaged and then compared with the average value of PPP-derived PWV estimates over the same time period. The average PWV measurements from the radiometer, the average PPP-derived PWV, and the differences between them are shown in Figure 6. The accuracy statistics were shown in Table 6.

The results indicate that the PWV difference between the WVR truth measurements and GPS estimates is less than 1 millimetre, with very small bias at the level of about 0.3 millimeter. The results demonstrate the potential to determine PWV to an accuracy of 1 mm in real-time using precise orbit and clock products and PPP methodology. This can satisfy the required accuracy for GPS meteorological applications (Gutman and Benjamin 2001). The results are also comparable to the traditional double-difference method, where accuracies of 1~2 mm are achieved with very long baselines (Tregoning *et al.*, 1998). An advantage of the PPP approach is that no local

reference stations are required (as no differential techniques are employed), and this method can be readily adopted at isolated sites – e.g. in a sparse GPS network (Gao *et al.*, 2004).

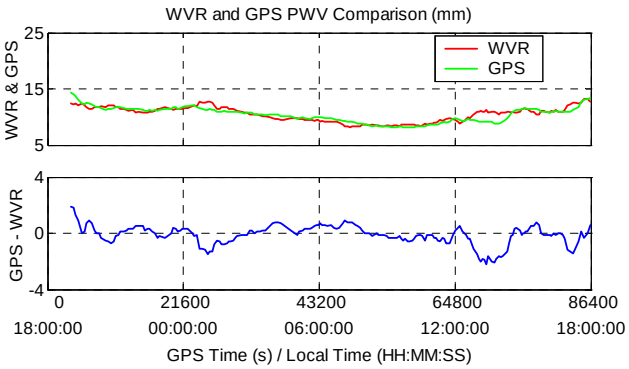


Fig. 6 WVR and GPS PWV comparison (Sept 5/04)

Tab. 6 PPP derived PWV accuracy

Products	RMS (mm)	BIAS (mm)	STD (ms)
JPL IGDG	0.77	0.28	0.72

5 Conclusions

The performance of different application solutions, including position determination under different dynamics environments, water vapour and receiver clock parameters estimation, using precise point positioning methodology has been assessed using real-time precise orbit and clock corrections. For position determination, centimetre accuracy is obtainable which is comparable to conventional double difference differential positioning. For receiver clock estimates, an accuracy of sub-nanosecond has been demonstrated by comparing to a very accurate clock at an IGS station. The precipitable water vapor (PWV) estimates from PPP agree with PWV measurements from a water vapour radiometer at 1 millimeter level.

The performance analysis presented in this paper demonstrate the potential of precise point positioning for real-time precise positioning, time transfer and water vapour estimation. Since no base stations are required for precise point positioning method, it is expected that the new method will bring greater operational flexibility while significant reduced costs to those applications in the future.

Acknowledgements

Financial support via a research grant from GEOIDE is acknowledged. Susan Skone and Natalya Nicholson are acknowledged for providing the radiometer and meteorological data, Jet Propulsion Laboratory is acknowledged for providing the real-time precise orbit and clock corrections, and Paul Mrstik and Sarka Friedl from Mosaic Mapping Systems Inc. are thanked for providing the aircraft dataset used in the data analysis.

References

Bar-Sever YE, Kroger PM and Borjesson JA (1998): ***Estimating Horizontal Gradients Of Tropospheric Path Delay With A Single GPS Receiver***. Journal of Geophysical Research, Vol. 103, No. B3, pp. 5019-5035.

Chen G and Herring TA (1997): ***Effects Of Atmospheric Azimuthal Asymmerty On The Analysis Of Space Geodetic Data***, Journal of Geophysical Research, 102, No. B9, 20, 489–20,502, 1997.

Gao Y, Skone S, Chen K, Nicholson NA and Muellerschoen R (2004): ***Real-Time Sensing Atmospheric Water Vapor Using Precise GPS Orbit and Clock Products***, Proceedings of ION GNSS 2004, Long Beach, California, September 21-24, 2004.

Gao Y, Chen K and Shen X (2003): ***Real-Time Kinematic Positioning Based on Un-Differenced Carrier Phase Data Processing***, Proceedings of ION National Technical Meeting, Anaheim, California, January 22-24, 2003.

Gao Y and Shen X (2002): ***A New Method for Carrier Phase Based Precise Point Positioning***, Navigation, Journal of the Institute of Navigation, Vol. 49, No. 2.

Héroux P (2004) personal communication.

Kouba J and Héroux P (2001): ***GPS Precise Point Positioning Using IGS Orbit Products***, GPS Solutions, Vol.5, No.2, pp. 12-28.

Larson K, Levine J, Nelson L and Parker T (2000): ***Assessment Of GPS Carrier-Phase Stability For Time-Transfer Applications***, IEEE Trans. Ultrason., Ferroelect., Freq. Control. Vol. 47, pp. 484–494.

Muellerschoen RJ (2003): personal communication.

Petit G, Jiang Z, White J, Beard R, and Powers E (2000): ***Absolute Calibration Of An Ashtech Z12-T GPS Receiver***. GPS Solutions, Vol 4, No. 4, pp. 41–46.

Radiometrics Corp. (1999): ***WVR-1100 Total Integrated Water Vapor and Liquid Water Radiometer Manual***.

Zumberge JF, Heflin MB, Jefferson DC, Watkins MM and Webb FH (1997): ***Precise Point Positioning For The Efficient And Robust Analysis Of GPS Data From Large Networks***. Journal of Geophysical Research, Vol. 102, 5005-5017.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 101-105

Journal of Global Positioning Systems
2004
© CPGPS

Alternative Positioning Method using GSM Signals

Goh Pong Chai

Nanyang Technological University, Singapore
e-mail: cpcgoh@ntu.edu.sg Tel: + 65-67905247; Fax: + 65-67910676

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Location-Based-Services (LBS) have not made any significant foray into commercial applications following the mandatory 911 requirements for emergency location initiated by US. Commercial adoption of LBS has been hampered by cost issues and technical difficulties faced by the various modes of positioning. For LBS to be really usable, positioning, irrespective of the technology, has to be a seamless and transparent process and made available under most environments and to within acceptable accuracy. Several services for positioning technologies had been researched and developed around GSM radio network, each with varying success. This paper describes a commercial application that has been developed on the Symbian platform. It allows user-friendly access to maps on a phone and incorporates a proprietary location algorithm that resides within the phone. The algorithm is a significant improvement from basic cell-id location methods.

Key words: LBS, GSM, Phone-Location, Cell ID, Symbian

1 Introduction

Precise positioning technologies have seen developments reached a plateau after Global Positioning System (GPS) was developed and commercially deployed since the mid 70s. The main thrust for alternative positioning techniques comes from the Federal Communications Commission's (FCC) wireless Enhanced 911 (E911) rules mandating the effectiveness and reliability of wireless 911 services by providing 911 dispatchers with additional information on wireless 911 calls. The wireless E911 program is divided into two parts - Phase I and Phase II. Phase I requires carriers, upon appropriate request by a local Public Safety Answering Point (PSAP), to report the telephone number of a wireless 911 caller and the location of the antenna that received the call. Phase II requires wireless carriers to provide far more precise location information, within 50 to 100 meters in most cases (http://www.fcc.gov).

Various positioning technologies were subsequently developed to ensure compliance, and/or to circumvent technological limitations of each technology. Most GSM operators settled for less accurate network positioning solutions such as Cell-ID, EOTD and TDOA in order to meet FCC E911 requirements. Unfortunately, most of these attempts did not perform to standards that are commercially viable. Trails conducted by a Singapore telephone operator revealed a wide range of achievable positional accuracy under varied conditions. The vagaries of positions obtained render such technologies unacceptable for many location based services (LBS). According to a research study of 25 carriers in the European Union carriers responded that they believe that accuracy would increase the attractiveness of some applications to consumers and revenue generation (http://www.trueposition.com).

An improved positioning technology (codenamed gprX) based on GSM technology, was developed by the Nanyang Technological University (NTU) and agis, a privately held company based in Singapore. A complete model and processing algorithms were developed for the whole of Singapore which allows a position to be determined entirely on the processing capability of the mobile handset. Efforts to developed passive systems that do not need to register to the base station network such as those using hyperbolic location positioning techniques are typically done within controlled environments (Senturk, 2002) (Mizusawa, 1996). Koets (2002) reported mean location accuracy of 89 meters in the trials done in San Antonio.

gprX is a passive system that does not require support services from the telephone operator's backend network. Unlike most other GSM network solutions, the technology takes away the burden of relying on operator's network services. Neither does it require

additional hardware beyond the basic mobile handset. The position of a handset can be determined using gprX to a much better accuracy than what is achievable today under similar environments. Accuracy can be further improved to an order of 30m under good network conditions and using a terrain constrained solution. In a worst case scenario, when operating in a single base mode (only one base station is available), the returned position is typically 60% better than otherwise achieved by traditional cell-ID methods. With gprX, widespread location based services (LBS) is now a real offering. Current implementation requires Symbian OS or Pocket Mobile OS (some hardware dependencies of the PocketPC Mobile).

2 Review of Current Positioning Technologies

2.1 GPS-Phone

This solution uses a combined GSM+GPS handset, the former provides the communications link while the latter provides the positioning capability. This hybrid solution fails in several areas:

Battery life of the device is too short.

It works only in open spaces with clear view of the sky.

Time to first fix (TTF) is not fast enough

2.2 A-GPS

This is another attempt to improve the coverage problem and slow TTF encountered by the hybrid GSM+GPS system by employing very sensitive GPS receivers that may pick up GPS signals indoors (close to the windows and open access). The weak signals are partially overcome, or assisted, by augmenting the system with relevant data that are transmitted to the hardware. Assisted data includes information such as GPS ephemeris and the approximate location. However, the method still needs GPS signals, albeit weak. When GPS signals are totally lost, A-GPS fails.

2.3 Terrestrial Based Network Solutions

The basic idea behind these technologies is that of trilateration, using signals from ground based radio stations such as GSM networks, TV signals and special ground based transmitters to measure ranges to known locations. In theory, if the time from a base station to a receiver is known, then all that is needed is to know are the station locations and the ranges to at least 3 of them in order to compute the receiver's location. Technologies in this category include:

Phone Network Solutions: Enhanced Observed Time Difference (EOTD), Time Difference of Arrival (TDOA), Angle of Arrival (AOA) and Cell of Origin (COO) are variants of such implementations. To compute a distance based on time difference, the stations must be synchronized. For EOTD, this requires supplementing the network with additional hardware/services so that the timings may be synchronized. For TDOA, each base station is synchronized with a GPS receiver or atomic clock.

Pseudolites: These are essentially ground-based transmitters that produce unsynchronized GPS-like signals. The technology uses a reference station for generating corrections. Generally, these are expensive and difficult to maintain. There are now attempts to develop low-cost, synchronised pseudolites. If successful, these could be a major breakthrough in positioning technology.

None to these technologies today is able to provide a low-cost, plug & play location solution that can be easily deployed. Perhaps the closest today may be that of COO. The advantage of COO is that it is pervasive, works in all areas with signal coverage (including in-buildings). Two main problems pertaining to this humble COO solution are, firstly, it can be extremely inaccurate where the base station spacing is sparse or where the signal propagated by a base covers a large area and secondly, it requires knowledge of the location of the base stations. Deployment becomes problematic because the telephone operator is in full control of the services and no access can be granted without their permission. Accuracy using COO is uncertain and can in the order of several kilometres, rendering the service virtually useless.

3 gprX Technology

gprX positioning technology is a proprietary technology that uses only information and signals that are available on the phone to enhance the quality of COO. The world today is wirelessly connected by two major technologies, GSM and CDMA. The billionth GSM user was connected in Q1 2004 – just a dozen years after the commercial launches of the first GSM networks. Today, GSM accounts for 73% of the world's digital mobile market and 72% of the world's wireless market (http://www.gsmworld.com).

GSM base stations are strategically located so as to provide sufficient signal overlaps by neighbouring stations to ensure smooth handovers to the next cell when a mobile handset moves from one point to another and also to provide for optimal voice quality. When a mobile unit is switched on, it is automatically registered to a

particular network and maintains a list of cells based on signal parameters and handover protocols. Unfortunately, signal propagation and signal pattern at any location is dynamic, and can be quite erratic, depending on ambient conditions. By studying the characteristic behaviour of signals at a location (affected by the ambience and terrain), a model was developed and appropriate algorithms were developed to map the observable parameters to probable locations. When a handset is in motion, characteristic patterns are generated and this can be assigned to the geographical location computed by the algorithm.

3.1 gprX Implementation

This process is currently implemented as a client software resident in the handset. A thin client solution may also be offered so that most of the processing is taken off the mobile handset and moved to the server. Most high end phones available today are equipped with sufficient processing power to run the software locally. The client software continually monitors the network parameters and other information provided by the GSM network and computes the most probable location in real-time. Once the model has been calibrated, it does not require knowledge of where the base stations are located and is therefore totally independent of the network operator's network services. The dynamic nature of the GSM signals is continually monitored and re-calibrated. Re-calibrated model parameters are returned to the handset which will improve its location capability with time.

3.2 Constrained Solutions

In situations where physical constraints are known, then it is possible to locate with almost 100% certainty where the handset is located. Buildings which are served by only one single base station (and repeaters) fall into this category. For improved performance, additional geometric and/or terrain constraints may also be imposed to improve the positioning accuracy. With known constraints, the computed location can be dramatically improved. Underground roads and train tunnels are geometric constraints that may be used to improve positioning quality.

3.3 Poor GSM Network

In the worst case scenario when only one base station is within monitoring range, the algorithm cannot return an improved position, then the COO weighted position will be returned. This is not the location of the base station but the location of a computed and averaged high sensitivity point. In all cases, it is a much better solution than the existing COO provided by the operator's network services which will only indicate the base-station location. Current tests showed that an improved COO gprX solution is at least 60% better than traditional COO location (Figure 1).

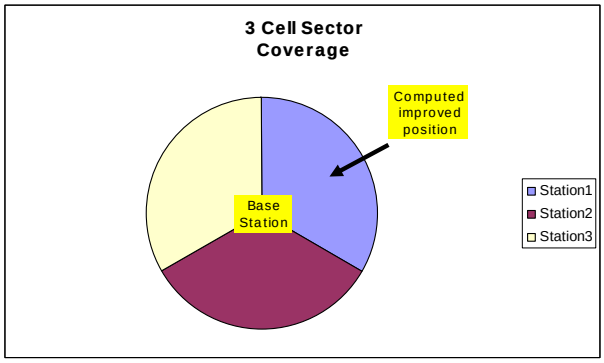


Fig1 Improved COO Location

3.4 Transient Signal Errors

In areas of low population density, it is usual to have long-range GSM stations and repeaters which are located far from its actual location. In such situations, the COO technique returns locations with errors of a few kilometres. In Singapore for example, a handset that at Orchard road may be tracked by a base located at Bukit Timah, a good 6 kilometres away. A filtering algorithm embedded in gprX removes such transient errors.

3.5 gprX Pre-Release Processes

gprX services are released after performing a sequence of pre-release analysis and formulation of algorithm and a model suitable for the particular terrain. Depending on requirements and geographical coverage, the pre-release exercise typically takes about a month to complete (Figure 2a). This is sufficient to for a model to be developed and to start a gprX session. A subsequent post-release automated process will continue to re-model the parameters continuously so that positioning accuracy will improve with time.

3.6 gprX Post-Release (run-time) Processes

After gprX services are released and implemented, a self-calibration process is implemented to improve the initial modelling parameters and also to provide a feedback loop for changes in the network. For example, when a new base station is added or the terrain is changed drastically because of new high-rise buildings, this self-enhancement process will capture these changes and feedback into the

model. Typically, the computed results will improve with time as this post-release phase kicks in (Figure 2b).

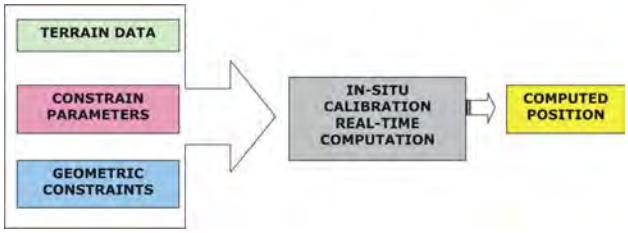


Fig 2a Pre-Release Process

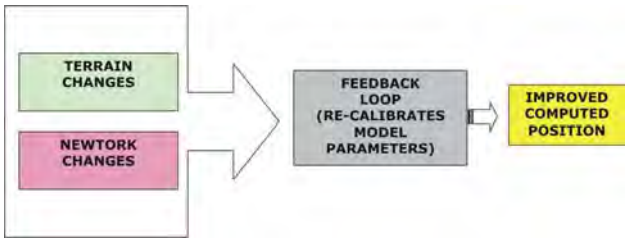


Fig 2b Post Release Process

3.7 Test Results

Using a calibrated model for Singapore, a drive test was done using a GPS to track an accurate position while the gprX provides location from only the phone. Figure 3 shows a comparison of a gprX location track (red circles) superimposed on a GPS track (blue points). In all tests, similar excellent results were obtained.

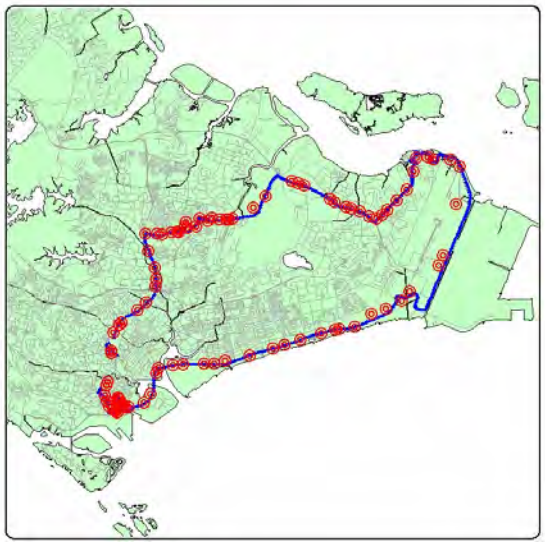


Fig 3 A Test Result

4. Conclusion

Hitherto, no single tracking technology has been able to fulfil the quest to determine accurate positions without the use of complicated and expensive hardware and under operating uner most environments. The FCC's mandate has not really been fulfilled even after massive development efforts to find the holy grail of positioning. While GPS is able to determine very good location fixes, it suffers from its inability to work under cover, relatively high costs and being hardware dependent. Other technologies have proven to be limited in terms of its overall performance in price and accuracy. There are many advantages of gprX:

Telco/Network Independent. There is no longer a need to engage the telephone operators and depend on expensive backend services.

No Additional Hardware.

Pervasive Coverage. Works anywhere, across countries where there are GSM signals, including in-buildings.

Good Accuracy in most situations. Sub-100m accuracy is achievable in most environments. High accuracy location data increases the value of mobile location services, enhancing the user experience and maximizing revenue.

Scalable and Cost Effective. Start-up costs are necessary but pales in comparison with other current technology available.

Seamless, Client Server Integration. Interfaces for system access and data distribution provide mechanism for third party software to be built upon the basic gprX services.

Scalable. Not limited by network infrastructure.

Acknowledgements

gprX was developed jointly by the author and Agis Private Limited (http://www.asiagis.com.sg), a Singapore domiciled company. Agis has extensive geographic database of Singapore and Malaysia and was instrumental in the development of the model and software development of services on mobile devices.

References

Mizusawa GA (1996): ***Performance of Hyperbolic Position Location Techniques for Code Division Multiple Access.*** Virginia Polytechnic Institute and State University. PhD Thesis.

Senturk H (2002): ***Performance Evaluation of Hyperbolic Position Location Technique in Cellular Wireless Networks***. Air Force Institute of Technology Wright-Patterson AFB.

Koets MA (2002): ***Development of Inverse Hyperbolic Positioning Using the GSM Cellular Telephone Network***. Southwest Research Institute, 16-9222.

Reference web-sites:

http://www.trueposition.com: (2004) U.S. Poised to Capitalize on Location Services, White Paper

http://www.fcc.gov

http://www.gsmworld.com

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 106-114

Journal of Global Positioning Systems
2004
© CPGPS

FPGA Implementation of a Single Channel GPS Interference Mitigation Algorithm

Gabriel Bucco, Matthew Trinkle, Doug Gray and Wai-Ching Cheuk

CRC for Sensor Signal and Information Processing, Dept of Electrical and Electronic Engineering, University of Adelaide, Adelaide SA,
e-mail: dgray@eleceng.adelaide.edu.au Tel: +618 8303 6425; Fax: +618 8303 4360

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. The FPGA (Field-Programmable Gate Array) implementation of an adaptive filter for narrow band interference excision in Global Positioning Systems is described. The algorithm implemented is a delayed LMS (Least Mean Squares) adaptive algorithm improved by incorporating a leakage factor, rounding and constant resetting of the filter weights. This was necessary as the original adaptive algorithm had stability problems : the filter weights did not remain fixed, and tended to drift until they overflowed, causing the filter response to degrade. Each model was first tested in Simulink, implemented in VHDL (Verilog Hardware Description Language) and then downloaded to an FPGA board for final testing. Experimental measurements of anti-jam margins were obtained.

Key words: Adaptive filtering, FPGA, LMS, GPS, Interference Mitigation

1 Introduction

Single channel adaptive filtering techniques have been shown to be an effective technique for mitigating multiple narrowband interferences to GPS systems (Robert, 1999, Landry et al., 1997). Since they can be seamlessly inserted between the existing GPS antenna and receiver, (Trinkle et al., 2001, Dimos et al., 1995), they offer a cost effective solution that involves minimum system disruption. However to become a fully practical solution the size and power demands of their hardware implementation should be minimised. FPGAs (Field-Programmable Gate Arrays) offer the potential for achieving the goals of small size, weight and power consumption and in this paper the implementation of an adaptive filter using an FPGA device is described.

In Section 2 an experimental system, termed mini-GISMO, is described and an overview of the system architecture is presented. The use of interpolation and decimation filters within the FPGA is also described.

The main adaptive algorithm implemented is the delayed LMS (Least Mean Squares) adaptive algorithm (Haykin, 2002). As discussed in Section 3 this algorithm is well suited to FPGA implementations. However, particularly in the presence of strong interferences, the original adaptive algorithm had stability problems (Sethares et al., 1986), as on convergence, the filter weights did not remain fixed, and tended to drift until they overflowed, causing the filter response to degrade. In Section 4 it is shown that incorporating a leakage term (Nascimento et al.,1999) and rounding instead of truncating resulted in the weights remaining near the optimal values. However, this solution introduced memory effects, which produced a second null when the interference frequency was changed. Resetting the weights every second removed this problem and appeared to have the least stability effects, as a short pulse in the output every second didn't cause any undesirable results in this algorithm. Also, the bit allocations were optimised to reduce the quantisation error. By reducing the quantisation noise power a smaller leakage factor is required to stabilise the adaptive algorithm resulting in a slower drift of the weights towards zero.

Finally, in Section 4.3, experimental results testing the ability of mini-GISMO to reject interferences whilst receiving GPS signals are presented.

2 System Description

The mini-GISMO system consisted of two boards : an RF and a digital board as shown in Figure 1 below.

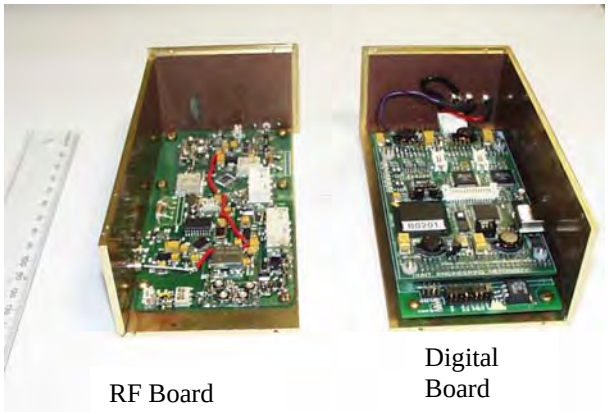


Fig. 1 mini-GISMO

An overview of the complete system is shown in Figure 2. The RF board converts the incoming GPS signal to an intermediate frequency of 22 MHz before feeding it to an A/D converter on the FPGA board. The output of the A/D converter is then passed into the FPGA, where a decimation filter, adaptive filter and interpolation filter are implemented. The final output from the interpolation filter is then passed back out through a D/A converter and converted back up to the GPS carrier frequency of 1.5754 GHz by the RF board.

2.1 Filter Design and Implementation

To avoid expensive and bulky analogue filters with very sharp transition bands the anti-aliasing filtering was effected using digital oversampling, i.e., the analogue signal was sampled at a much higher rate (32MHz) than twice the GPS signal bandwidth (2MHz). This over-sampled digital signal was then digitally filtered and sub-sampled to an 8MHz sampling rate in the decimation filter. By using a combination of analogue and digital filtering, sharp transition bands could be achieved .

2.2 Analogue Anti-Aliasing Filter

The ADC samples at 32 MHz, and the GPS IF frequencies range from 21-23 MHz. Thus the frequency band of interest aliases to 9-11 MHz. The analogue anti-aliasing band-pass filter is also required to have good attenuation between 9 and 11 MHz and between 41 and 43 MHz as noise from these bands also aliases into the final GPS signal bandwidth.

2.3 Decimation Filter

The decimation filter reduces the input sample rate from the ADC from 32 to 8 MHz by sub-sampling. To avoid additional aliasing at this stage a further anti-aliasing filter needs to be implemented. Since the down-converted GPS spectrum is from 9-11 MHz, the anti-aliasing filter needs to have good rejection below 7 and above 13 MHz to avoid aliasing into the decimated GPS bandwidth.

A linear-phase FIR filter design with least-squares error minimization was used to obtain the coefficients, which were then normalised to 14 bit signed integers. The frequency response of the resulting decimation filter is shown in Figure 3 and a diagram showing the entire frequency down-conversion scheme is shown in Figure 4.

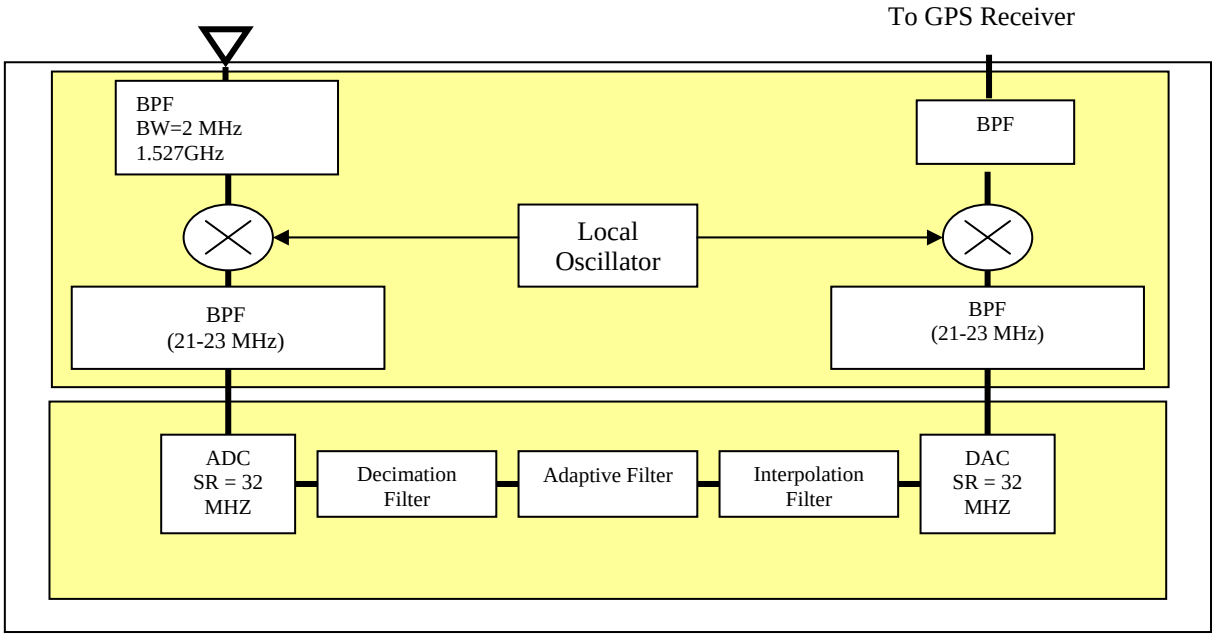


Fig. 2 mini-GISMO overview

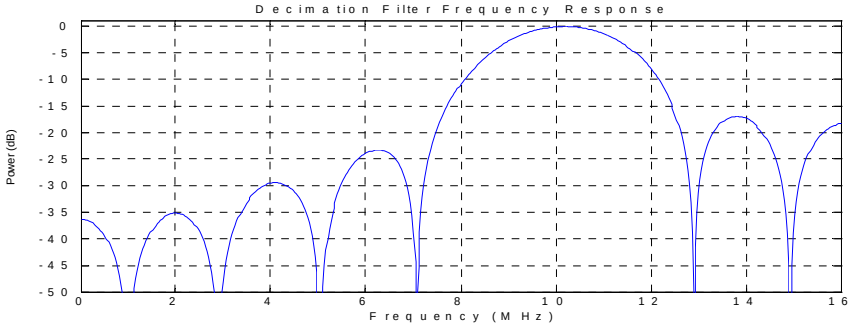


Fig. 3 Frequency response of decimation filter

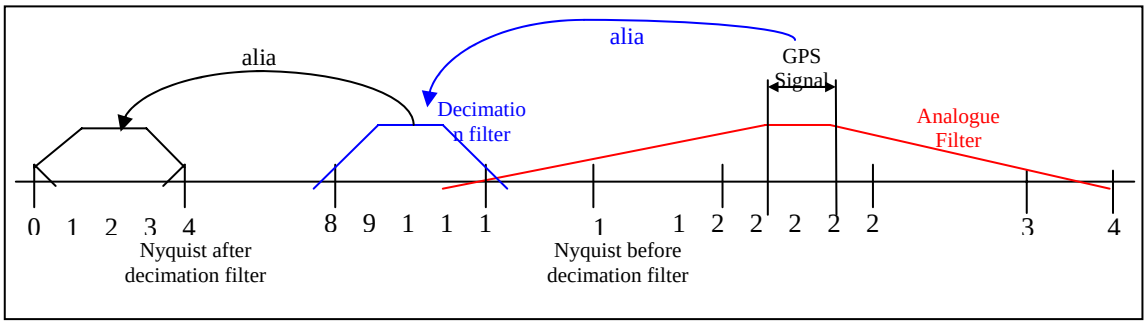


Fig. 4 Frequency conversion scheme

2.4 Interpolation Filter

Interpolation is effected by adding 3 zeros between contiguous samples of the output of the adaptive filter thus taking the input sample rate of 8 MHz to an output sample rate of 32 MHz. However this interpolation introduces unwanted image frequencies and these must be filtered out. This is achieved by the use of an interpolation filter, which effectively inverts the effect of the decimation filter, i.e., it translates the adaptively filtered GPS signal from base-band to an intermediate frequency of 10 MHz. The interpolation filter transfer function and hence filter coefficients are identical to those of the decimation filter.

2.5 FPGA Filter Implementation

A block diagram showing the interconnections between the main system blocks in the FPGA are shown in Figure 5. The data is clocked in at a 32 MHz rate, and the decimation filter creates a Ready signal at a quarter the input clock rate, or 8 MHz. The output data is clocked by this signal, i.e., at 8 MHz, and both these signals are used to drive the adaptive filter module. The data into the interpolation filter (DIN2) is obtained from the adaptive filter output, and changes at 8 MHz. The filter itself is clocked at 32 MHz, which is also the output data rate.

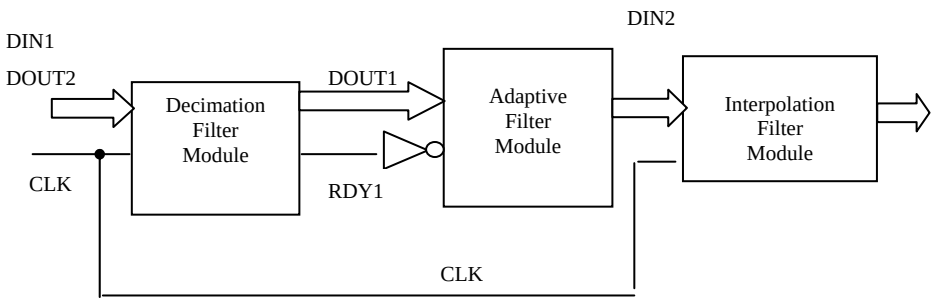


Fig. 5 Main system blocks in the adaptive filter FPGA implementation

2.6 FPGA Decimation/Interpolation Filter Testing

The frequency responses of the decimation/interpolation filter were measured by passing white noise into the system and measuring the output spectrum. The results are shown in Figure 6 where the power spectrum of the input 'white noise' is shown in black and the output spectrum in blue. The input white noise spectrum is measured after the analogue anti-aliasing filter. The results show that the filters give at least 15-20 dB of attenuation from the pass-band to the cut-off regions.

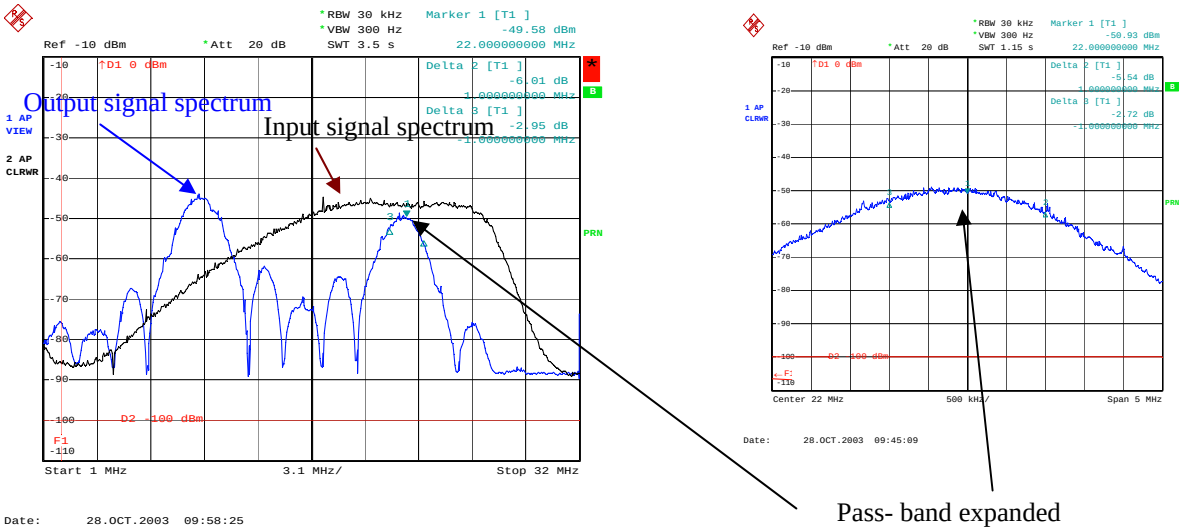


Figure 6 Output power of interpolation filter when input is white noise

A comparison of Figure 3 (theoretical) and Figure 6 (measured) indicate a good agreement, allowing for the appropriate shifts of the frequency axis, as the attenuation side lobe levels appear consistent. The insert in Figure 6 shows an expanded view of the frequency response in the pass-band of the decimation/interpolation filter. Whilst it shows that the attenuation in the 21-23 MHz pass-band is quite flat, the attenuation at the edges reaches almost 3dB on one side, and 5.5dB on the other side, which is a bit uneven. The response could be improved by adding more taps, but due to space limitations on the FPGA, the limit was set to 13 taps.

3 Adaptive Filter

The adaptive filter implemented in the FPGA is the symmetric delayed LMS adaptive prediction-filtering algorithm, shown in Figure 7 below. The implementation of the LMS algorithm for updating the filter weights is a modified version of the leaky LMS leakage algorithm that also included periodic re-initialisation of the weights.

3.1 Adaptive Filter FPGA Implementation

The Adaptive Filter was implemented in schematic VHDL using four main blocks as illustrated in Figure 8. The following diagram shows these blocks, the Coefficient Update Block, the Input Signal Delay Block, the FIR Algorithm Block and the Tree Adder Block.

Each of the modules carries out the following operations

Coefficient Update Block : Updates weights using a variant of the LMS algorithm and scalings.

Input Signal Delay Block : Delays input signal according to a 33 tap symmetric FIR filter.

FIR Algorithm Block : Carries out weight vector multiplications and scalings

Tree Adder Block : Carries out summations

When designing the scaling both truncation and rounding were investigated.

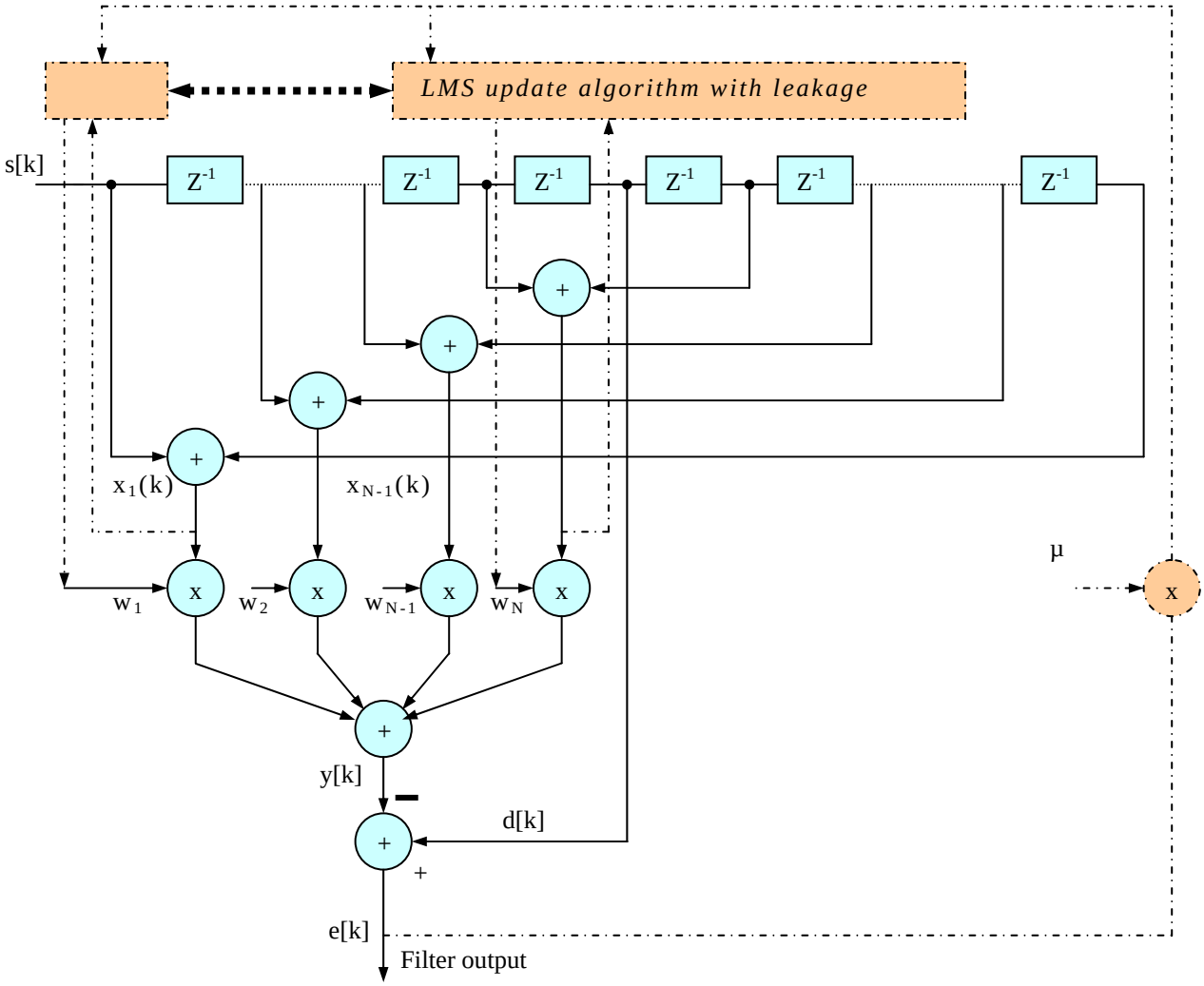


Fig. 7 Symmetric Adaptive Prediction Filter

4. Adaptive Filter Development and Testing

The adaptive filter was first tested using a Simulink model to evaluate the effect of various quantisation schemes, (Sherwood et al., 1987), convergence times and weight vector drift. Then the VHDL code was downloaded onto the FPGA and the experimental system was evaluated. These results are presented in the following sections.

4.1 Simulink Tests

The adaptive filter uses 33 taps and has a sampling rate of 8 MHz, thus the convergence time should be between 10.8μs and 21.7μs for typical interferences. Whilst Simulink analysis confirmed convergence times of around 20μs it also confirmed that weight vector drift, due to finite precision arithmetic, in the standard delayed LMS algorithm was a serious problem. The drift caused most of the weights to become very large, thus distorting the frequency response of the filter. This is illustrated in Figure 9 below where the transfer function of the standard delayed LMS adaptive filter is plotted over a 10 second period.

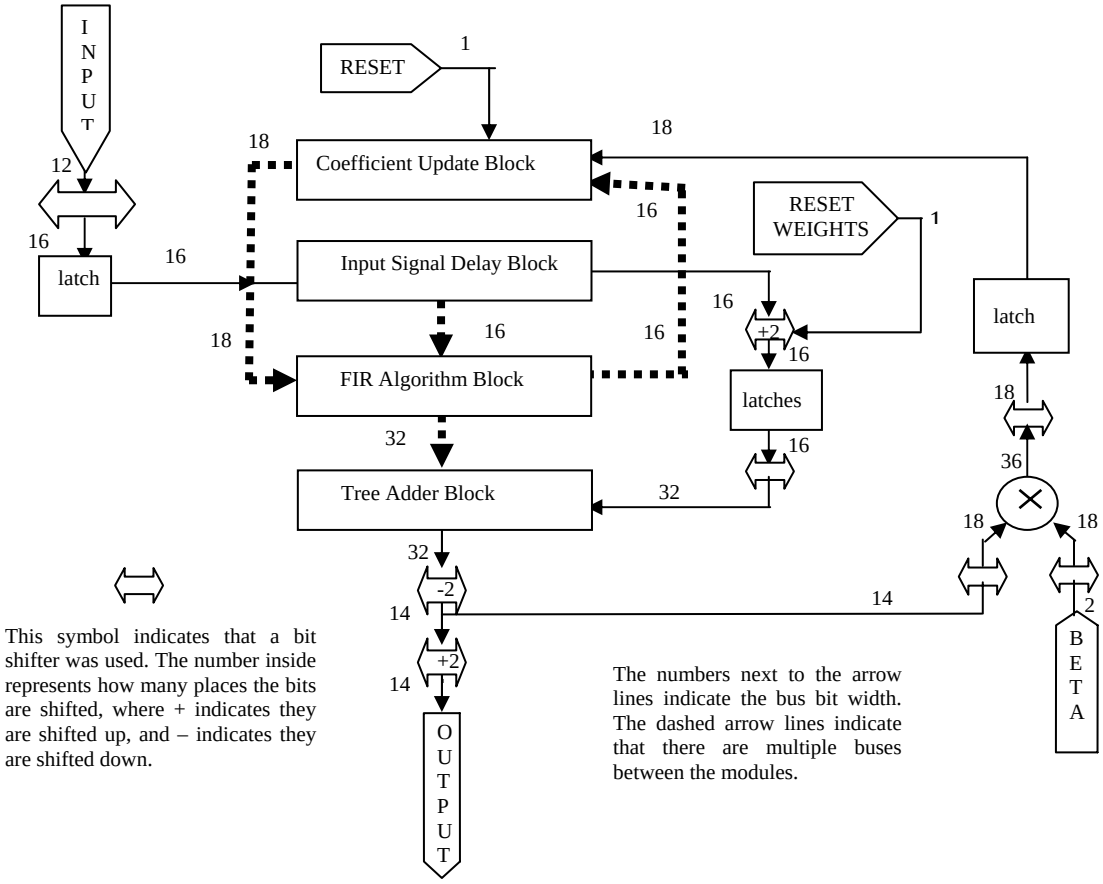


Figure 8 Adaptive Filter Module

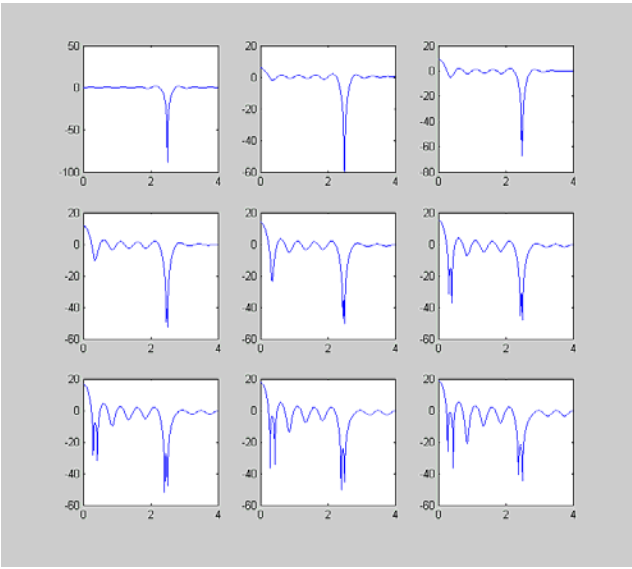


Figure 9. Effect of weight vector drift on transfer function of standard delayed LMS

The top left corner shows the optimum result, after convergence, with a relatively flat response and a deep, narrow notch. Over time, the ripples start to grow, the notch gets wider and starts to divide, and some of the ripples start forming secondary notches. This shows the degradation of the response, as the weights drift during a period of about 10 seconds. The size of the ripples in the standard LMS algorithm is also interference to noise ratio, (INR), dependent. As the INR increases the adaptive algorithm places less emphasis on maintaining a flat frequency response. This is illustrated in Figure 10 where the INR for the red curve is 21dB and that for the blue curve is 30 dB.

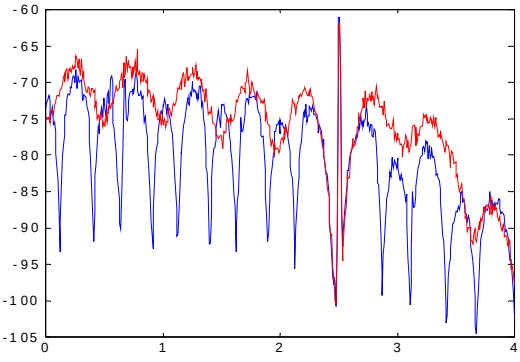


Figure 10. Effect of increased INR on ripples

One way of minimizing weight vector drift investigated was to use a modified version of the leaky LMS algorithm. The effect of the leakage term is to eventually draw the weights towards zero, preventing them from overflowing. Simulink analysis confirmed this behaviour. However whilst this approach is good in the short term, over a long period of several seconds many components of the weight vectors get drawn to zero, reducing the effective number of filter taps. Thus although the sharp null-steering capability is retained, a lumpier frequency response curve results. This is illustrated below in Figure 11. The behaviour of the leaky algorithm was not found to be strongly dependent on the INR.

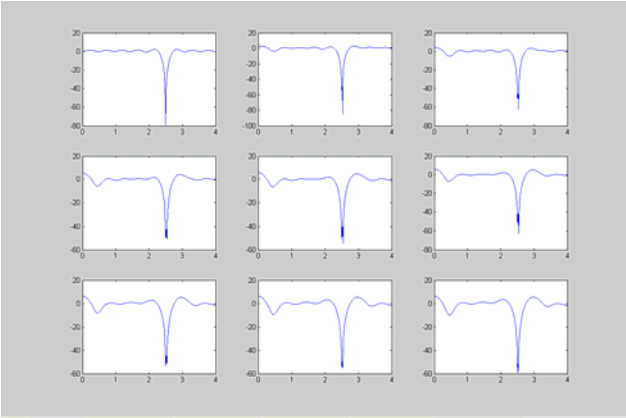


Figure 11. Effect of weight vector drift on transfer function of delayed LMS with leakage.

Rounding rather than truncation was also investigated and found to be particularly effective in reducing weight vector drift. Truncation always pulls the numbers in the same direction, causing the weights to also start drifting in the same direction, whereas the random nature of rounding keeps the weights oscillating around a constant value. Thus in theory, rounding should mitigate weight vector drift as it removes the biased nature of truncation.

4.2 FPGA Tests

The leaky adaptive filter code incorporating leakage was downloaded onto the FPGA board, and an interference signal with an INR of 21 dB was used. The output of the mini-GISMO system was input to a spectrum analyser and spectra were recorded approximately every second. During the collection period of about 10 mins the frequency of the interference was varied to check how the adaptive filter handled a changing interference frequency. In the results presented below the logged spectra are plotted as a colour image with the horizontal axis representing frequency bins, the vertical axis time and the intensity is colour coded with the colour bar indicating strength in dBs.

Figure 12 presents results obtained implementing the delayed LMS algorithm with leakage. The results show that although the leakage term has been effective in reducing weight vector drift thus maintaining cancellation of the interference, the spectrum still has a number of deep ripples (red) and rolls off on the right faster (blue) than on the left. However, the notch appears to be quite narrow, quickly tracks and effectively knocks out the interference signal.

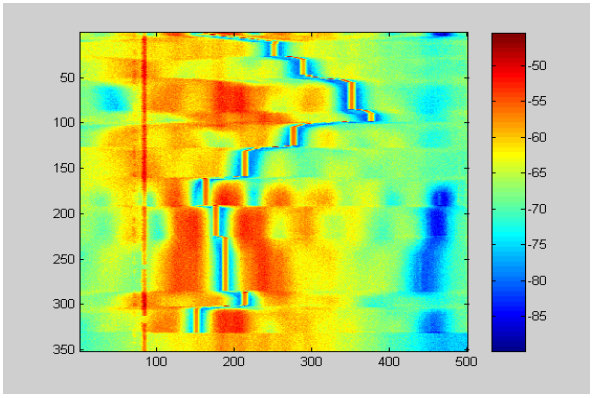


Figure 12. FPGA results - delayed LMS with leakage

As discussed in Section 4.1 rounding with no leakage was particularly effective in reducing weight vector drift whilst maintaining a flat frequency response. This was tested using the mini-GISMO system with an interference whose INR was 30dB. The results, Figure 13, indicate the output spectrum is flatter than that of Figure 11 but the filter now exhibits a relatively long memory effect, maintaining a notch at the previous interference frequency for some time and only slowly converging to the new interference frequency. As the INR increased this memory effect was found to become more prominent. That is, as the time taken for the notches to combine increases, the response became less flat and the notch less sharp.

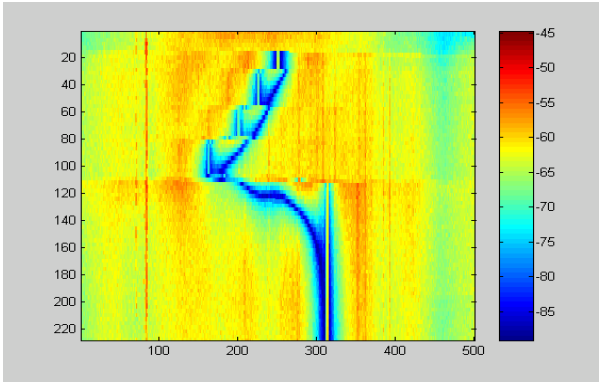


Figure 13. FPGA results - delayed LMS with rounding

4.3 Mini-GISMO Implementation and Testing

The investigations detailed in Section 4.2 indicated that whilst either adding a leakage term and modifying the quantisation scheme reduced the effect of weight vector drift, there were disadvantages to both approaches. The scheme chosen for the final implementation was the following combination of the above approaches.

(a) A leakage term was used.

(b) The weights were reset to zero every second.

(c) Due to FPGA capacity constraints combined rounding and leakage was not implemented. However a detailed analysis of the quantisation effects of the LMS algorithm was carried out and an improved bit allocation scheme, which lowered the quantisation noise due to truncation, was adopted. This reduced the effect of frequency response deterioration over time.

A typical example of the power spectrum for two different INRs is shown in Figure 14.

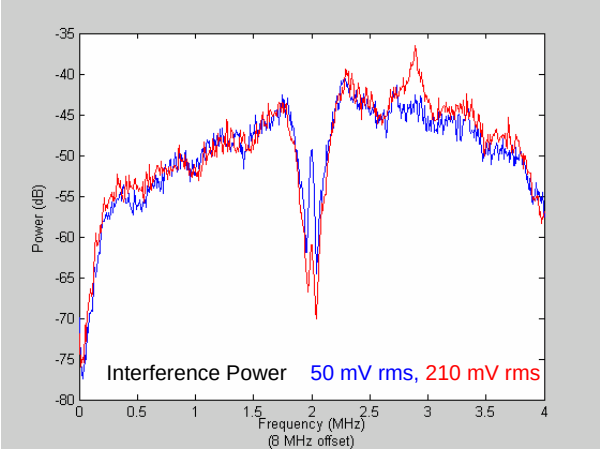


Figure 14. FPGA results – optimised algorithm

The frequency responses when the interference signal is at the same frequency but different power levels are shown. Both interference power levels give a similar result, apart from a deeper notch for the stronger interference. The responses are relatively flat, without any extra ripples or notches forming in the pass-band region

This adaptive algorithm was downloaded onto the PROM and testing was carried out with mini-GISMO inserted between an antenna and a small GPS receiver. An insertion loss of approximately 3dB was measured by measuring the GPS signal strengths prior to and after insertion of mini-GISMO. A continuous-wave interference was added to the antenna output prior to the mini-GISMO and the output spectrum was logged.

The results for a changing interference frequency of power –60dBm are illustrated in Figure 15. The spectra are relatively flat, the notch is quite narrow and is quite deep. However there is still some residual distortion and ripple around the notch frequency as seen by the cyan colouring. This distortion was worse when the interference frequency was close to the centre of the pass-band however this did not seem to cause a significant loss in GPS signal strength as the receiver was still able to maintain lock on the satellite signals.

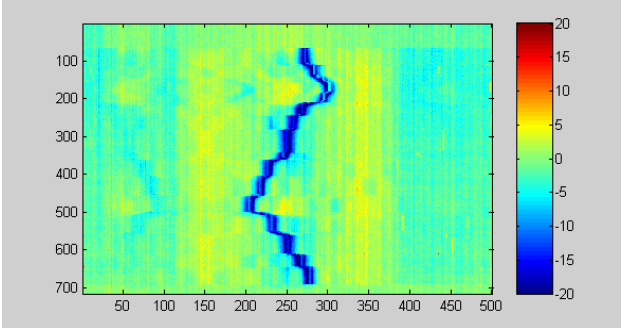


Figure 15. mini-GISMO results –interference power of –60dBm

In a separate test the strengths of four GPS signals in the presence of –60dBm and –70dBm interferences were recorded as the frequency of the interference was changed. These results are plotted in Figure 16 below.

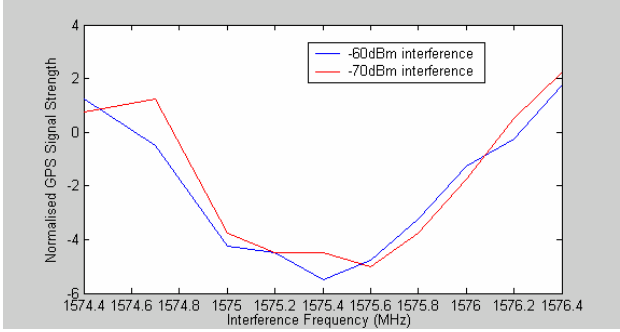


Figure 16. Measured GPS signal strengths as interference frequency was changed

The GPS signal strengths were averaged and normalised by the signal strength in the absence of the interference. These results indicated that the system whilst effectively cancelling the interference was causing a loss in the GPS signal strength varying from 0dB at the band edges to –5 dB at the centre of the band. This variation across the band is probably due to the adaptive filter attempting to flatten the coloured response across the pass-band caused by the decimation filter and resulted in an apparent gain when the interference frequency was at the high end of the pass-band.

Finally the anti-jam gain of mini-GISMO was measured by comparing the interference power levels at which an

unprotected and protected receiver lost lock on the GPS signals. Without mini-GISMO the GPS receiver lost lock when the interference power level was greater than about –95 to -90dBm whilst with mini-GISMO inserted the threshold was about –57 dBm indicating that mini-GISMO provided about 35dB of anti-jam gain.

5 Conclusions

The single channel delayed LMS adaptive algorithm is an effective technique for removing narrow-band interfering signals from GPS receivers. It can be effectively implemented using FPGA technology that can be seamlessly inserted between a GPS antenna and receiver. Measurements indicate that the experimental mini-GISMO system can provide around 35dB of anti-jam gain against such interferences.

Careful design of the delayed LMS algorithm was required to prevent overflows due to weight vector drift. Several improvements were tested and the results showed that incorporating a leakage factor, replacing truncation by rounding and periodically resetting the weights all contributed to mitigating this problem. In the final FPGA implementation resetting the weights every second, incorporating a leakage factor and careful allocation of the bit allocations when truncating produced good results.

References

Robert S. (1999): ***The Impact of Interference on Civil GPS***. ION 55th Annual Meeting, June 1999 pp. 821-828

Landry R.; Renard A. (1997) ***Analysis of Potential Interference Sources and Assessment of Present Solutions For GPS/GNSS Receivers***, 4th Saint-Petersburg International Conference on Integrated Navigation Systems, May 1997.

Trinkle M.; Gray D. (2001) ***GPS Interference Mitigation; Overview and experimental Results***. SatNav 2001, July, 2001.

Dimos G.; Upadhyay T.; Jenkins T. (1995) ***Low-Cost Solution to Narrowband GPS Interference Problem*** Proc IEEE 1995 Naecon Meeting, vol I pp145-153 May 1995.

Haykin S. (2002) ***Adaptive Filter Theory 4th Edition***, Prentice Hall.

Sethares W.; Lawrence D.; Johnson C.;, Bitmead R.; (1986) ***Parameter Drift in LMS Adaptive Filters*** IEEE Transactions on Acoustic, Speech and Signal Processing Vol. ASSP-34, No 4, August 1986.

Nascimento V.; Sayed A. (1999) ***Unbiased and Stable Leakage-Based Adaptive Filter*** IEEE Transactions on Signal Processing, Vol. 47, No. 12 December 1999.

Sherwood D.; Bershad N. (1987) ***Quantization effect in the Complex LMS Adaptive Algorithm: Linearization Using Dither-Theory*** IEEE Transaction on Circuits and System, Vol. CAS-34, No. 7, July 1987.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 115-131

Journal of Global Positioning Systems
2004
© CPGPS

An analysis of the effects of different network-based ionosphere estimation models on rover positioning accuracy

Dorota A. Grejner-Brzezinska[1], Pawel Wielgosz[1,2], Israel Kashani[1,3], Dru A. Smith[4], Paul S. J. Spencer[4], Douglas S. Robertson[4] and Gerald L. Mader[4]

[1] The Ohio State University, SPIN LAB, 470 Hitchcock Hall, 2070 Neil Ave., Columbus, OH 43210-1275
e-mail: dbrzezinska@osu.edu; Tel: (614) 292-8787; Fax: (614) 292-2957

[2] University of Warmia and Mazury – UWM

[3] Israel Institute of Technology – Technion

[4] National Geodetic Survey – NGS

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. The primary objective of this paper is to test several methods of modeling the ionospheric corrections derived from a reference GPS network, and to study the impact of the models' accuracy on the user positioning results. The five ionospheric models that are discussed here are: (1) network RTK (NR) carrier phase-based model — MPGPS-NR, (2) absolute, smoothed pseudorange-based model — MPGPS-P4, (3) IGS Global Ionosphere Model — GIM, (4) absolute model based on undifferenced dual-frequency ambiguous carrier phase data — ICON, and (5) carrier phase-based data assimilation method — MAGIC. Methods 1–4 assume that the ionosphere is an infinitesimal single layer, while method (5) considers the ionosphere as a 3D medium.The test data set was collected at the Ohio Continuously Operating Reference Stations (CORS) network on August 31, 2003. A 24-hour data set, representing moderate ionospheric conditions (maximum Kp = 2o), was processed. The ionospheric reference "truth" in double-difference (DD) form was generated from the dual-frequency carrier phase data for two selected baselines, ~60 and ~100 km long, where one station was considered as a user receiver at an unknown location (simulated rover). The five ionospheric models were used to generate the DD ionospheric corrections for the rover, and were compared to the reference "truth." The quality statistics were generated and discussed. Examples of instantaneous ambiguity resolution and RTK positioning are presented, together with the accuracy requirements for the ionospheric corrections, to assure integer ambiguity fixing.

Key words: Network-based RTK, ionospheric models, ambiguity resolution.

1 Introduction

One of the major limiting factors in GPS-based precise positioning is the ionosphere-induced propagation delay that, if not properly accounted for, may result in significant positioning errors. This is particularly true for single frequency data, reducing the length of the effective baseline to 10–15 km. While dual-frequency carrier phase measurements can form an ionosphere-free linear combination that removes first-order ionospheric errors, the integer ambiguities can be fixed only for short baselines, since the ionospheric error decorrelates with distance. The ionospheric signal delay is a function of the total electron content (TEC). TEC is defined as the total number of electrons contained in a column with a cross-sectional area of 1 m^2 along the signal path. TEC displays primarily day-to-night variations, but also depends on the geomagnetic latitude, time of year, and the sunspot cycle. It is measured in electron/m^2, where one total electron content unit (TECU) is defined as 10^{16} electron/m^2. One meter of the ionospheric delay (or advance) on the first GPS frequency corresponds to 6.16 TECU, or one TECU causes 0.162 m delay.

The ability to remove the ionospheric delay from GPS data can increase the performance of the integer

ambiguity resolution (AR) process, and improve the computational efficiency of the search process. However, due to the high-level variability of TEC, empirical ionospheric models do not provide sufficient accuracy to support high-precision positioning applications. On the other hand, if external information on the integrated TEC between pairs of satellites and receivers is provided based on real observation data, the base-rover separation can be significantly extended, resulting in a virtually "distance-independent" precise positioning. The external ionospheric information provides an important constraint for accurate and rapid carrier phase AR. An alternative approach is to estimate the double-difference (DD) ionospheric delay parameters in the positioning adjustment model. However, the underlying mathematical model becomes weaker, and as a consequence the required observation sessions become longer (Odijk, 2000); thus, this method does not apply to rapid static or kinematic algorithms where the occupation time is of the order of seconds to minutes.

Several methods have been proposed to estimate and model the ionospheric corrections from the ground-based GPS data from the Continuously Operating Reference Stations (CORS) network, and the mathematical representation of the ionospheric electron density field have been studied (see, for example Odijk, 2000 and 2001; Schaer, 1999; Wielgosz *et al*., 2003; Kashani *et al*., 2004a; Smith, 2004; Spencer *et al*., 2004) to support network-based real-time kinematic (RTK) (see, for example, Vollath *et al*., 2000; Rizos, 2002; Wanninger, 2002; Grejner-Brzezinska *et al*., 2004a and 2004b). The most challenging among these methods is the instantaneous AR approach (Kim and Langley, 2000), where for each individual observation epoch a new integer ambiguity solution is obtained using only the current epoch data (Bock *et al*., 2003). Consequently, this method is resistant to negative effects of cycle slips and gaps, and can provide centimeter-level positioning accuracy immediately, without any delay needed for initialization (or re-initialization). However, to successfully resolve the ambiguities instantaneously over long distances, external atmospheric corrections of high accuracy are required (Odijk, 2001; Kashani *et al*., 2004b).

This paper presents the accuracy analysis of several methods of ionospheric correction modeling in comparison to reference "truth," using the Multi Purpose GPS Processing Software (MPGPS™) developed at the Satellite Positioning and Inertial Navigation (SPIN) group at The Ohio State University. The ionospheric corrections are represented by the DD ionospheric delays estimated directly from dual-frequency carrier phase data constrained by the integer ambiguities. Examples of the positioning accuracy, supported by instantaneous AR, achieved as a function of the ionospheric correction quality are also presented and discussed. The five ionospheric models tested are listed below.

MPGPS-NR — Network RTK (NR) carrier phase-based model, decomposed from DD ionospheric delays (Kashani *et al*., 2004a),

MPGPS-P4 — Absolute, smoothed pseudorange-based method (Wielgosz *et al*., 2003),

IGS GIM — International GPS Service (IGS) global ionospheric map (GIM). IGS GIM is a combination of several different ionosphere models provided by the IGS Ionosphere Associate Analysis Centers (Schaer, 1999),

ICON — Absolute model based on undifferenced dual-frequency ambiguous carrier phase data (Smith, 2004),

MAGIC — Pseudorange-leveled carrier phase-based data assimilation method (Spencer *et al*., 2004).

Methods 1–4 assume that the ionosphere is an infinitesimal single layer, while method 5 considers the ionosphere as a 3D medium. In terms of spatial coverage, methods 1 and 2 are considered local; methods 4 and 5 are regional, while method 3 offers global coverage. MAGIC and ICON (4 and 5) are the two NGS ionosphere models derived for the continental United States. These two models use CORS GPS data, and provide the ionospheric corrections with a three-day delay. Both models are prototypes, a part of ongoing research projects, and are currently available to the general public for testing and evaluation purposes (http://www.noaanews.noaa.gov/stories2004/s2333.htm).

2 Approach and methodology

This section provides a brief description of the algorithms and methods applied to the aforementioned ionosphere estimation models. In addition, the AR and rover positioning algorithms are also briefly discussed. More details can be found in the references provided in each section.

2.1 The ionospheric models

2.1.1 The reference "truth" model — MPGPS-L4

The fundamental mathematical model for the network-based adjustment, using dual-frequency carrier phase and pseudorange data from the reference stations, is described by Eq. (1). The system of equations (1) written for the entire network is solved for each epoch of observations using a generalization-based sequential least-squares adjustment with stochastic constraints (Kashani *et al*., 2004a).

$$\begin{aligned}
&\Phi_{1,ij}^{kl} - \rho_{ij}^{kl} - (\alpha_i^k T_i - \alpha_i^l T_i - \alpha_j^k T_j + \alpha_j^l T_j) + I_{ij}^{kl} - \lambda_1 N_{1,ij}^{kl} = 0 \\
&\Phi_{2,ij}^{kl} - \rho_{ij}^{kl} - (\alpha_i^k T_i - \alpha_i^l T_i - \alpha_j^k T_j + \alpha_j^l T_j) + (\upsilon_1^2/\upsilon_2^2) I_{ij}^{kl} - \lambda_2 N_{2,ij}^{kl} = 0 \\
&P_{1,ij}^{kl} - \rho_{ij}^{kl} - (\alpha_i^k T_i - \alpha_i^l T_i - \alpha_j^k T_j + \alpha_j^l T_j) - I_{ij}^{kl} = 0 \\
&P_{2,ij}^{kl} - \rho_{ij}^{kl} - (\alpha_i^k T_i - \alpha_i^l T_i - \alpha_j^k T_j + \alpha_j^l T_j) - (\upsilon_1^2/\upsilon_2^2) I_{ij}^{kl} = 0
\end{aligned} \quad (1)$$

where:

i, j receiver indices,

k, l satellite indices,

$\Phi_{n,ij}^{kl}$ DD phase observation on frequency n (n=1,2),

$P_{n,ij}^{kl}$ DD code observation on frequency n,

ρ_{ij}^{kl} DD geometric distance,

$T_{i,j}$ tropospheric total zenith delay (TZD),

α_i^k troposphere mapping function,

I_{ij}^{kl} DD ionospheric delay,

υ_1, υ_2 L1 and L2 frequencies,

λ_1, λ_2 GPS frequency wavelengths on L1 and L2,

$N_{1,ij}^{kl}, N_{2,ij}^{kl}$ DD carrier phase ambiguities on L1 and L2.

The unknown parameters are undifferenced total zenith delay (TZD), provided for individual stations ($T_{i,j}$), DD ionospheric delays (I_{ij}^{kl}), and DD ambiguities ($N_{1,ij}^{kl}$, $N_{2,ij}^{kl}$). The coordinates of the permanent stations (CORS) are considered known (obtained from a 24-hour solution using the BERNESE software (Hugentobler *et al*., 2001)), which makes the AR for the reference network much easier to perform, even for long-range distances (i.e., ~200 km between the CORS stations). All the parameters are constrained to some a priori information, which may consist of empirical values (e.g., 30–50 cm for the DD ionospheric delays, depending on the baseline length, local time and satellite elevation angle), as well as variance-covariance matrix for the known CORS coordinates. The Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) is used to fix the ambiguities to their integer values (Teunissen, 1994). The validation procedure used is the AR success rate (Teunissen *et al*., 2002), which represents the probability of estimating the correct integers.

After the DD ambiguities associated with the reference receivers have been fixed to their correct integer values, the "true" DD ionospheric delay can be correctly estimated using the geometry-free linear combination described by Eq. (2):

$$L_{ij,4}^{kl} + \xi_4 I_{ij}^{kl} - (\lambda_1 N_{1,ij}^{kl} - \lambda_2 N_{2,ij}^{kl}) = 0 \quad (2)$$

where:

$$L_{ij,4}^{kl} = L_{ij,1}^{kl} - L_{ij,2}^{kl} \text{ and } \quad \xi_4 = 1 - \frac{\upsilon_1^2}{\upsilon_2^2} \approx -0.647$$

It can be seen from Eq. (2) that if the dual-frequency ambiguity parameters $N_{1,ij}^{kl}$ and $N_{2,ij}^{kl}$ are fixed to their integers, the only remaining unknown parameter, I_{ij}^{kl}, representing the DD ionospheric delay on L1, can be estimated with a few-millimeter accuracy, corresponding primarily to the noise on the DD carrier phase observables.

2.1.2 The carrier phase DD model — MPGPS-NR

The network-based approach presented is Section 2.1.1 was also used to derive the network-based ionospheric corrections (MPGPS-NR). In this model, the zero difference (ZD) ionospheric delays (i.e., one-way, biased) were obtained by decomposing the network-derived DD delays, and interpolated for the rover location (KNTN). This was done for all satellites/epochs used in the analyses presented in Section 3. The network solution did not include the selected rover observations. It should be mentioned that for an individual baseline and *n* DD delays the rigorous decomposition is not possible without the a priori knowledge of at least two ZD values; there are only *n-2* linearly independent DD observation equations, thus the system is undetermined. To regularize the normal matrix, independent constraints on at least two ZD delays are needed, or loose constraints can be introduced to the diagonal of the normal matrix. Both methods result in a biased estimate. Although the reconstruction of the DD delays from the interpolated ZD delays may still include some amount of residual biases, the resulting DD delays are generally of sufficient accuracy to enable fast AR with only a few epochs of data (see Kashani *et al*., 2004a).

2.1.3 The smoothed pseudorange model — MPGPS-P4

In this approach (Wielgosz *et al.*, 2003) dual-frequency GPS carrier phase data are used to smooth the pseudorange observations collected at the reference station network (Springer, 1999). After the smoothing procedure, the pseudoranges are effectively replaced by the carrier phase observations with approximated (real-value) ambiguities. The differential code biases (DCBs) for the satellites are provided by IGS (ftp://gage.upc.es/pub/gps_data/GPS_IONO), and for the receivers, are derived from the calibration performed with the BERNESE software (Hugentobler *et al.*, 2001). The instantaneous absolute ionospheric delay I_i^k is computed from Eq. (3):

$$I_i^k = (\tilde{P}_{i,4}^k - c(\Delta b^k + \Delta b_i)) / \xi_4 \quad (3)$$

where:

i, j, k, l indices are the same as in Section 2.1.1,

$\tilde{P}_{i,n}^k$ carrier-smoothed code observation on frequency n (n=1,2),

$\tilde{P}_{i,4}^k$ geometry-free linear combination of smoothed code observations $\tilde{P}_{i,4}^k = \tilde{P}_{i,1}^k - \tilde{P}_{i,2}^k$,

c speed of light,

Δb^k, Δb_i DCB for satellite k, and receiver i, respectively,

ξ_4 coefficient converting ionospheric delay on P_4 to P_1 delay (Eq. (2)).

The ionospheric delay is the following function of TEC (Schaer, 1999):

$$I_i^k = \pm \frac{C_x}{2} TEC \upsilon_1^{-2} = \xi_{TEC} TEC \quad (4)$$

where the proportionality factor $\frac{C_x}{2}$ = 40.3 $\times 10^{16}$ ms^{-2}/TECU, and the ionospheric delay caused by 1 TECU on $L1$ is ξ_{TEC} = 0.162 m/TECU.

2.1.4 The IGS GIM

The IGS GIMs are derived as a weighted combination of several different models developed independently by the IGS Ionosphere Associate Analysis Centers. The combined maps have a spatial resolution of 2.5° and 5.0° in latitude and longitude, respectively, and a 2-hour temporal resolution (Feltens and Jakowski, 2002). IGS GIMs assume a single layer model with the layer height, H, at 450 km. To convert the vertical TEC (VTEC) from GIMs into the line-of-sight slant TEC, a modified single-layer model (MSLM) mapping function is adopted (Eq. 5).

$$F(Z) = \frac{1}{\cos(z')} \quad (5)$$

with:

$$\sin(z') = \frac{R}{R+H} \sin(\alpha z)$$

$$TEC = VTEC / \cos(z')$$

where: $F(z)$ is the mapping function, R is the Earth's radius, z is the vertical angle to the satellite, z' is the angle between the topocentric direction to the satellite and the normal to the ionospheric layer through the pierce point (piercing angle), and α is a scaling factor. GIMs provide absolute TEC in the IONEX format (Schaer *et al.*, 1998). In this study, the ionospheric delays were interpolated for the rover and base receiver locations using kriging; they can subsequently be used to form DD corrections in the rover-positioning step.

2.1.5 The carrier phase absolute model — ICON

This method of computing absolute (unambiguous) levels of TEC (and subsequently L1 and L2 delays) from a ground-based network of GPS receivers requires dual-frequency carrier phase data. The ionosphere is assumed to lie in a single layer of constant ellipsoidal height of 300 km, and the geographic locations of the GPS ground stations must allow simultaneous observation of satellites by a number of stations. Slant TEC derived from GPS measurements is converted to VTEC, as shown in Eq. (5). Considering dual-frequency data, one can derive Eq. (6) to estimate the TEC difference between consecutive epochs m and n, where κ corresponds to $C_x / 2$ in Eq. (4). Using Eq. (5), the TEC difference can be further converted to the difference in VTEC, as shown in Eq. (7), which represents the ambiguous VTEC as a function of the unknown absolute TEC at epoch m. In these formulas, φ_1 and φ_2 are the carrier phase data in cycles.

$$\Delta TEC^{m,n} = TEC^m - TEC^n = \left(\frac{1}{\kappa}\right)\left(\frac{1}{\upsilon_1^2} - \frac{1}{\upsilon_2^2}\right)^{-1} \left(\Delta\varphi_1^{m,n} - \Delta\varphi_2^{m,n}\right) \quad (6)$$

$$\Delta VTEC^{m,n} = VTEC^n - VTEC^m = TEC^n \cos(z'^n) - TEC^m \cos(z'^m) = \Delta TEC^{m,n} \cos(z'^n) + TEC^m (\cos(z'^n) - \cos(z'^m)) \quad (7)$$

In order to estimate the bias (as only TEC time-difference can be obtained from (6) or (7)), a concept of track crossover is introduced, which is the fundamental notion of this method. The term crossover is related here to any two satellite tracks formed by the ionosphere pierce points, that fall within some acceptably small tolerance (in both space and time) of one another. Considering three tracks forming a triangle, as shown in Figure 1, each of the three tracks has one unknown bias (b_1, b_2 and b_3), corresponding to the absolute TEC value at some initial epoch. If these biases were known, the absolute values of TEC and VTEC at every epoch can be computed using the biases and ΔVTEC defined in Eq. (7). Each crossover A, B and C in Figure 1 forms a unique constraint for the system (Eq. 8), which allows for absolute estimation of VTEC. Adjoining "triangles," shown in Figure 1, provide more constraints introducing redundancy to the system. For more details on this method, see (Smith, 2004). In this study, the ionospheric delay values were computed for entire continental U.S., using around 340 stations; normally, for stations not included in the computations, TEC (and delays) can be interpolated for the station/satellite pair.

$$\begin{aligned} VTEC_1^A &= VTEC_3^A = VTEC^A \\ VTEC_1^B &= VTEC_2^B = VTEC^B \\ VTEC_2^C &= VTEC_3^C = VTEC^C \end{aligned} \tag{8}$$

2.1.6 The carrier phase tomographic model — MAGIC

A method used here to reconstruct the 3D Earth ionospheric electron density field using a land-based network of GPS receivers is referred to as the data assimilation method. It uses a Kalman filter and can combine data from various sources to obtain inversions in three dimensions (Spencer *et al*., 2004). The primary problem of the data assimilation method, which requires some form of mathematical regularization, transpires from the fact that the method is essentially a very poorly constrained linear least squares problem. Continuity relationships, defining the smoothness or entropy of the solution, are often used regularization techniques in tomographic methods (2D and 3D), while in cases of data assimilation methods, an a priori model estimate of the solution along with an estimate of its covariance are used. MAGIC uses an optional mapping function to alter the representation of the Kalman filter state vector in terms of a set of discrete radial empirical orthonormal functions (EOFs) to enable a more concise representation of the state vector (the ionospheric electron density field) in three dimensions (Spencer *et al*., 2004). The EOFs were formed by applying singular value decomposition to a set of model profiles generated by IRI95 (International Reference Ionosphere; see, for example, Bilitza, 1997). The dominant term, EOF_1, represents a mean ionospheric profile. Examples of empirical orthonormal functions are illustrated in Figure 2.

The carrier phase geometry-free linear combination (L4) leveled to the pseudorange is used as the GPS observable. The solutions are quantized with 15-minute time-steps. In this study, about 150 CORS and IGS stations were used for the TEC (converted to ionospheric delay) estimation for the continental US. For more details on this method, see (Spencer *et al*., 2004).

2.2 The kinematic positioning algorithm

The concept of long-range instantaneous RTK GPS presented here is based on the atmospheric corrections derived from the GPS observations collected by the reference network that supports the rover positioning. The data reduction algorithm operates in the DD mode, which requires receiving observations from at least one reference station, together with the atmospheric corrections, i.e., tropospheric, T_{ij}^{kl}, and ionospheric, I_{ij}^{kl}, as shown in Eq. 9.

$$\begin{aligned} \Phi_{1,ij}^{kl} - \rho_{ij}^{kl} - T_{ij}^{kl} + I_{ij}^{kl} - \lambda_1 N_{1,ij}^{kl} &= 0 \\ \Phi_{2,ij}^{kl} - \rho_{ij}^{kl} - T_{ij}^{kl} + (f_1^2 / f_2^2) I_{ij}^{kl} - \lambda_2 N_{2,ij}^{kl} &= 0 \\ P_{1,ij}^{kl} - \rho_{ij}^{kl} - T_{ij}^{kl} - I_{ij}^{kl} &= 0 \\ P_{2,ij}^{kl} - \rho_{ij}^{kl} - T_{ij}^{kl} - (f_1^2 / f_2^2) I_{ij}^{kl} &= 0 \end{aligned} \tag{9}$$

The fundamental observation equations for pseudorange and carrier phase observations, parameterized according to the generalization-based approach, are presented in Eq. 9, where the notation follows the previously explained paradigm. Stochastic constraints are applied to the atmospheric corrections and the LAMBDA method is used for the AR. All the processing is carried out at the rover receiver in the instantaneous mode. The success of the instantaneous GPS positioning over long baselines depends on the ability to resolve the integer phase ambiguities. The performance of the method strongly depends on the quality of the atmospheric corrections provided from the network. If high quality corrections are available, the method becomes virtually distance-independent. The analyses presented in Section 3 are aimed at estimating the required quality of these corrections to assure seamless, instantaneous positioning without any initialization, as required in the on-the-fly (OTF) technique.

All the processing algorithms currently implemented in the MPGPS™ software include the following modules: long-range instantaneous and OTF RTK GPS, precise point positioning (PPP), multi-station DGPS, ionosphere

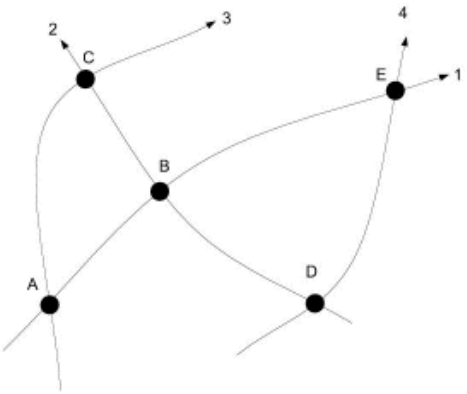


Fig. 1 A "triangle" ABC formed by 3 tracks and 3 crossovers; Adjoining "triangles" showing four tracks (unknowns) and five crossovers (observations), provide redundancy (Smith, 2004).

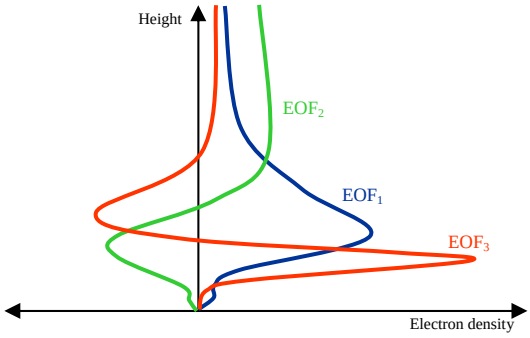


Fig. 2 Examples of empirical orthonormal functions (Spencer *et al.*, 2004).

Fig. 3a The CORS subnetwork used in the experiments.

Fig. 3b The baselines analyzed in the experiments.

modeling and mapping, and troposphere modeling. The software operates in static, kinematic and instantaneous modes, and can provide solutions in the network as well as in the baseline mode. More details on the instantaneous positioning algorithms can be found in (Kashani *et al*., 2004a and 2004b; Wielgosz *et al*., 2004; Grejner-Brzezinska *et al*., 2004a and 2004b).

3 Test data and experimental results

3.1 Test data set and data processing strategy

A 24-hour GPS data set collected by the Ohio CORS stations on August 31, 2003, with a 30 s sampling rate, was used in the experiments described in the following sections. The 24-hour data set allowed for a comparison of the time windows with different ionospheric TEC levels and varying GPS constellations. Figure 3 illustrates the selected reference subnetwork (3a) and the baselines processed and analyzed here (3b). Station KNTN was selected as an unknown rover, whose "true" coordinates were obtained and analyzed from a 24-hour static solution of the Ohio CORS network using the BERNESE software. The MPGPS™ software was used to first process the five-station network (KNTN, COLB, SIDN, DEFI, TIFF) to derive the tropospheric and ionospheric corrections that formed the reference "truth" (MPGPS-L4) in the test comparing the ionospheric models. Next, the four-station network, where the simulated "rover" was removed from the solution, was used to derive the atmospheric reference corrections (TZD and ZD ionospheric corrections - MPGPS-NR), which were then interpolated to the rover location, and applied in the baseline solutions presented in Section 3.3.

3.2 The ionospheric model analysis

The five ionospheric models (MPGPS-P4, MPGPS-NR, IGS GIM, ICON and MAGIC), as explained earlier, were used to derive DD ionospheric corrections for two baselines formed by the rover and two different reference stations (baseline KNTN-SIDN is ~60 km long, and KNTN-DEFI is ~100 km long). Figures 4 and 6 illustrate the estimated corrections for both baselines, while Figures 5 and 7 display the residuals with respect to the reference "truth" (MPGPS-L4). The time scale is shown in UTC time, with the zero epoch corresponding to UTC midnight (five hours ahead of Eastern Standard Time, i.e., the local time).

Tab. 1 DD ionospheric delay residual statistics (±5 cm and ±10 cm cut-off) for 24 h.

	Residuals in % within predefined limits, 24 h			
	KNTN-SIDN (~60 km)		**KNTN-DEFI (~100 km)**	
	±10 cm	**±5 cm**	**±10 cm**	**±5 cm**
MPGPS-P4	66.8	31.5	75.3	48.7
MPGPS-NR	99.3	94.2	99.3	94.2
IGS GIM	94.9	71.4	81.7	54.3
ICON	58.4	31.9	58.2	32.5
MAGIC	98.0	83.3	90.1	67.1

Tab. 2 DD ionospheric delay residual statistics ±5 cm and ±10 cm cut-off) for 2-hour session 04:00–06:00 UTC.

	Residuals in % within predefined limits, 04:00–06:00 UTC			
	KNTN-SIDN (~60 km)		**KNTN-DEFI (~100 km)**	
	±10 cm	**±5 cm**	**±10 cm**	**±5 cm**
MPGPS-P4	79.8	71.5	79.8	48.7
MPGPS-NR	97.0	84.1	97.0	84.1
IGS GIM	90.7	63.7	70.8	37.3
ICON	70.0	40.8	42.0	22.2
MAGIC	94.5	78.6	86.1	62.4

Tab. 3 DD ionospheric delay residual statistics (±5 and ±10 cm cut-off) for 2-hour session 18:00–20:00 UTC.

	Residuals in % within predefined limits, 18:00–20:00 UTC			
	KNTN-SIDN (~60 km)		**KNTN-DEFI (~100 km)**	
	±10 cm	**±5 cm**	**±10 cm**	**±5 cm**
MPGPS-P4	59.9	39.9	66.0	36.3
MPGPS-NR	100.0	100.0	100.0	100.0
IGS GIM	98.1	69.8	88.6	55.9
ICON	75.2	26.1	93.5	63.4
MAGIC	99.9	91.5	98.2	83.5

As can be observed in Figures 4 and 6, the ionosphere variability changes during the course of the day, and seems to be slightly more variable during the local night, as compared to the local day. It is quite opposite to the standard ionospheric behavior; however, the ionospheric gradients were indeed higher during the night, indicating greater ionospheric activity. The maximum Kp index for that day was 2o.

It can be observed from Figures 5 and 7 that the MPGPS-P4 model displays a very smooth signature of differences from the reference "truth"; this is primarily due to the fact that both solutions are based essentially on actual carrier phase data (i.e., phase-smoothed pseudoranges). The OSU network RTK model (MPGPS-NR) shows a good fit to the reference "truth," while the IGS GIM model displays bigger departures from the reference solution. This can be explained by the fact that GIM has lower spatial and temporal resolution, and thus it is subject to some smoothing effects. Still, the fit is approximately at the level of about 1.0 L1 cycle.

The NGS ICON model displays a rather flat spectrum of differences with respect to the reference "truth"; however, some biases are visible in the figures. The ionospheric signature of both the "truth" and the ICON models are very similar, as both were derived from carrier phase data. The ICON solution may be subject to incorrectly resolved biases, primarily due to the very strict procedure of cycle slip detection and fixing, and due to using a simple cosine mapping function at crossovers. This matter is currently under detailed investigation, and once it is resolved, this model is expected to provide high-quality ionospheric corrections for precise positioning.

The comparison of MAGIC to the reference "truth" indicates a fit similar to the signature shown for the IGS GIM. Again, this model is subject to smoothing due to

the time quantization of the final output, as already explained. Still, the fit is good, reaching a maximum of around 0.5–1.0 L1 cycle. It should be pointed out that the apparent discontinuities, which can be observed in the ionospheric model plots are in fact not real, and are due to the fact that the reference satellite was changed every two hours, so each 2-hour session displays an essentially different DD ionosphere.

In analyzing the differences between each model and the reference "truth" the following should be considered: the MPGPS-NR differences with respect to the "true" DD ionosphere are not base-rover distance-dependent, as the errors in the DD ionospheric correction come only from the undifferenced (and thus, biased) ionospheric correction interpolated for the rover location. The accuracy of the ionospheric estimate at the reference stations is considered uniformly biased per satellite; thus the bias is removed by forming the DD in the rover positioning procedure. In case of the MPGPS-P4 model the errors are also distance-independent since they arise from the undifferenced ambiguity estimation for a particular satellite-station pair during the carrier phase smoothing procedure. The biased ambiguities, and thus the ionospheric delay estimates, are station-dependent and completely random. The behavior of the ICON model is similar to that of MPGPS-P4. Namely, the undifferenced ambiguities are satellite-station-specific, and the amount of bias that affects them is random. To the contrary, both MAGIC and GIM models do not display station- or station-satellite dependence, but both are rather distance-dependent, and thus, the error in the ionospheric correction is bigger for longer baselines.

The summary statistics for the ionospheric model comparison is shown in Table 1, which covers the entire 24-hour session. Since the behavior of the ionosphere and the models changes over the course of the day (see Figures 4–7), two representative 2-hour windows (04:00–06:00 UTC, corresponding to roughly the local midnight; and 18:00–20:00 UTC, representing the local daytime window) were selected, and the statistics are summarized in Tables 2 and 3. The tables show the percentage of the model differences from the reference "truth" that falls within ±5 cm and ±10 cm predefined limits.

3.3 Instantaneous RTK Positioning Analysis with MPGPS-NR Model

In order to assess the quality of the ionospheric correction in more absolute terms, the final rover positioning test was performed. The two baselines, selected earlier, were used to derive the coordinates of KNTN (the rover) using the MPGPS-NR ionospheric corrections for the two representative 2-hour windows, as explained in the previous section. These windows correspond to the worst (04:00–06:00 UTC) and the best (18:00–20:00 UTC) quality of the MPGPS-NR ionospheric corrections (see Tables 4 and 5). The instantaneous positioning module from the MPGPS^{TM} software was used (i.e., single-epoch solution without OTF initialization), and the accuracy of the positioning results is illustrated in Figures 8–11. In order to emphasize the importance of the correct measure of the quality of the applied ionospheric corrections, several experiments were performed with the DD ionospheric corrections treated as fixed (0 sigma, which means the ionosphere is removed from the functional model) and as stochastic constraints, with the sigma varying from 1 cm to 5 cm. Examples of these analyses are shown next.

Figure 8 displays the *n*, *e* and *u* residuals of the rover coordinates with respect to the known station coordinates for the ~60 km baseline, where the initial standard deviation of the ionospheric corrections of 0 cm was used in the rover positioning algorithm. Clearly, the assumption that the external ionosphere is errorless was not correct for the "worst" window (04:00–06:00 UTC), while the "best" window (18:00–20:00 UTC) provides an excellent solution (with 100% of instantaneous AR success rate) under this assumption. The best solution for the "worst" window was obtained when the ionospheric correction was constrained to 5 cm, while this constraint was too loose for the "best" window, as shown in Figure 9. However, the instantaneous AR success rate for the "worst" window was only around 75 %. In case of the 100 km baseline, the sigma of 0 cm provided unsatisfactory results for both windows analyzed (see Figure 10), while a sigma of 1 cm for the ionospheric constraint was the best choice for the "best" window (with 100% of instantaneous AR success rate), and a sigma of 5 cm delivered the best positioning results (with around 70% of instantaneous AR success rate) for the "worst" window, as shown in Figure 11.

The positioning results discussed in this section were all derived using the MPGPS-NR ionospheric corrections that, based on the statistics provided in Tables 4 and 5, display the best fit to the reference "truth." The time windows analyzed in Tables 4 and 5 correspond to the 2-hour sessions illustrated in Figures 8–11. Clearly, the MPGPS-NR solution provides DD ionospheric corrections that fit the reference "truth" the best, and are usually sufficient to provide instantaneous centimeter-level positioning of the user. The other models discussed here offer lower rate of success in supporting instantaneous AR, and therefore the suitability of these models to support fast OTF AR is currently under investigation and will be reported in the next publication.

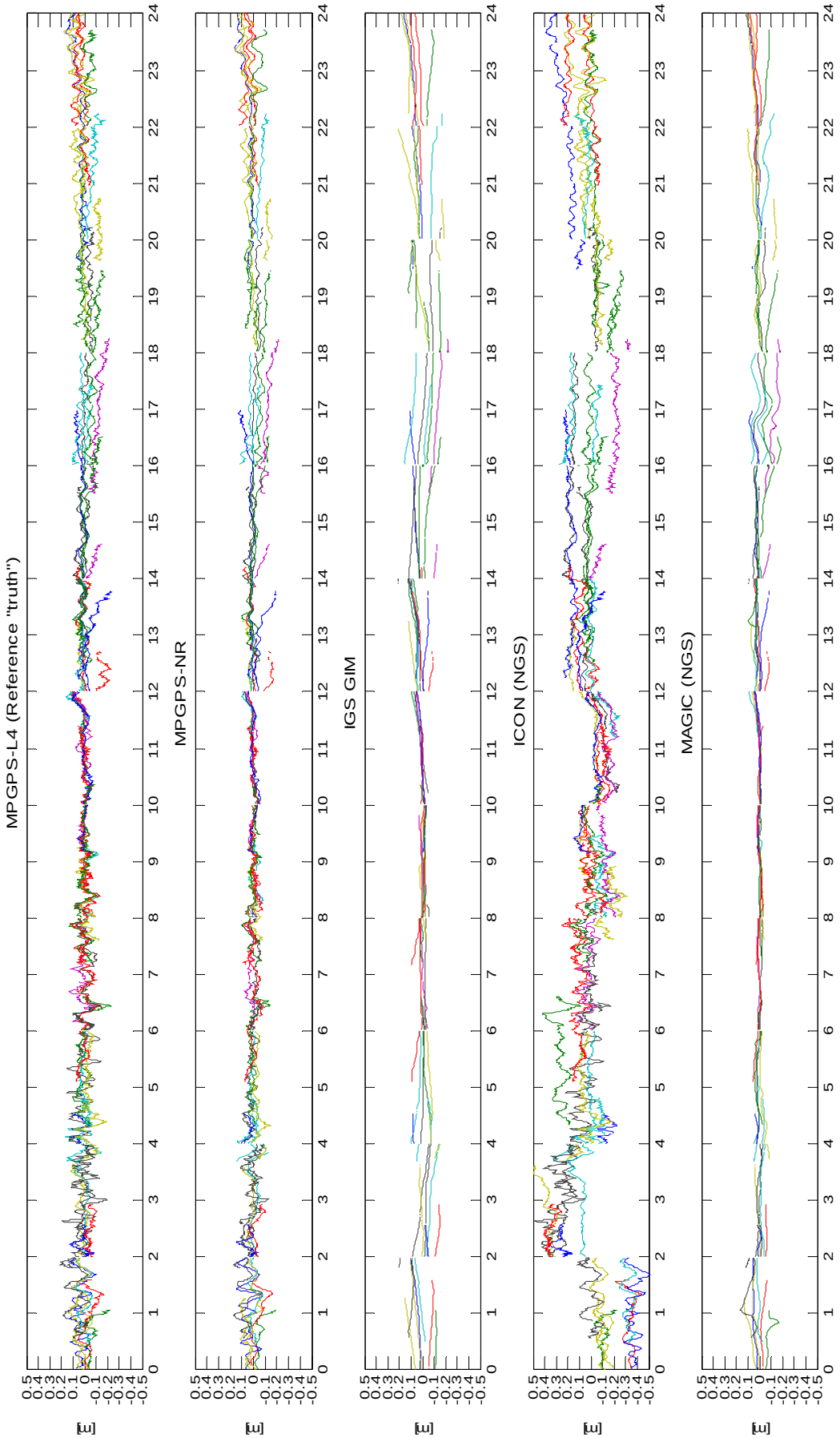


Fig. 4 Estimated DD ionospheric corrections for the analyzed models — 24 h, KNTN-SIDN (~60 km).

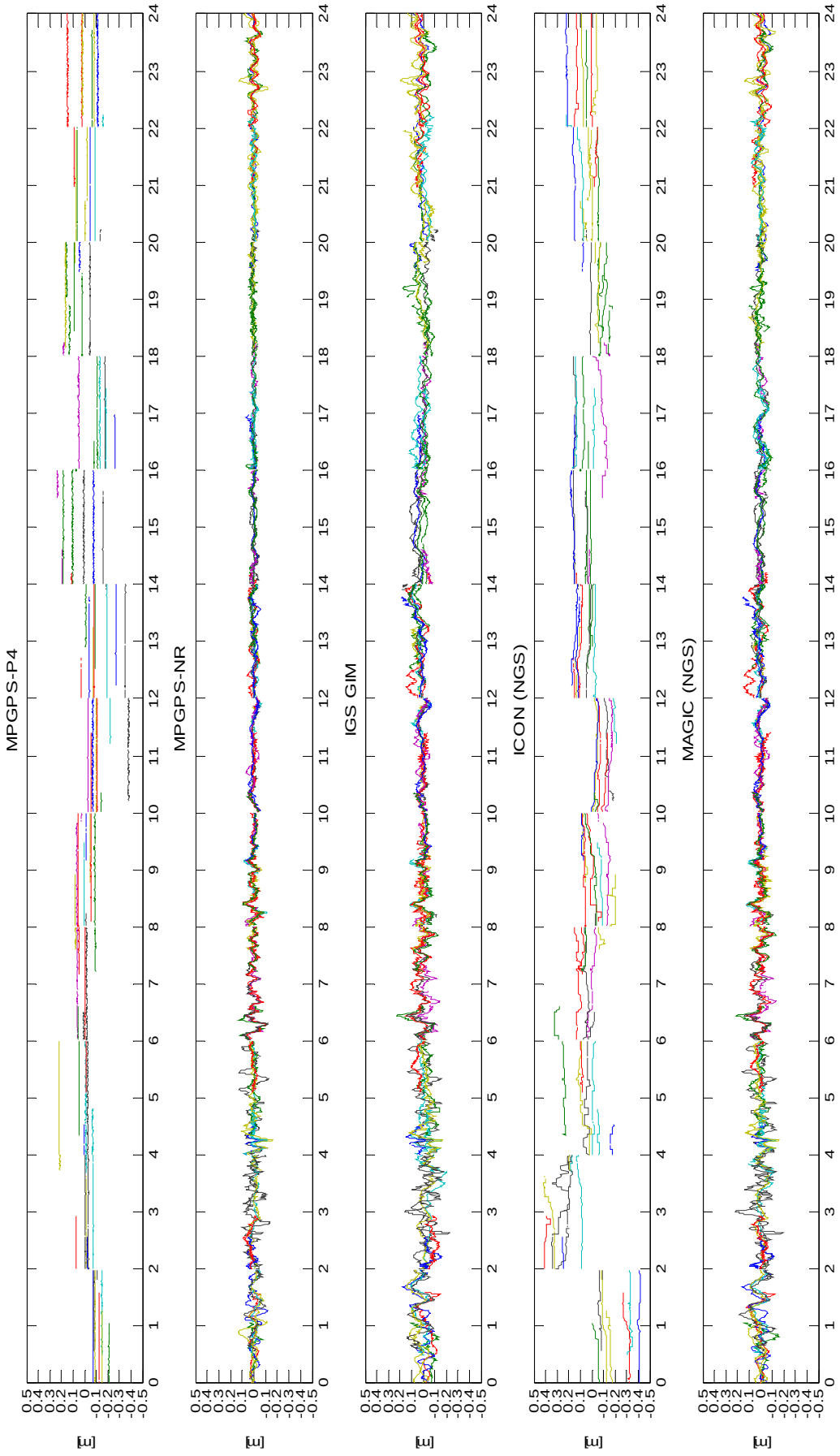


Fig. 5 DD ionospheric residuals with respect to the reference "truth" — 24 h, KNTN-SIDN (~60 km).

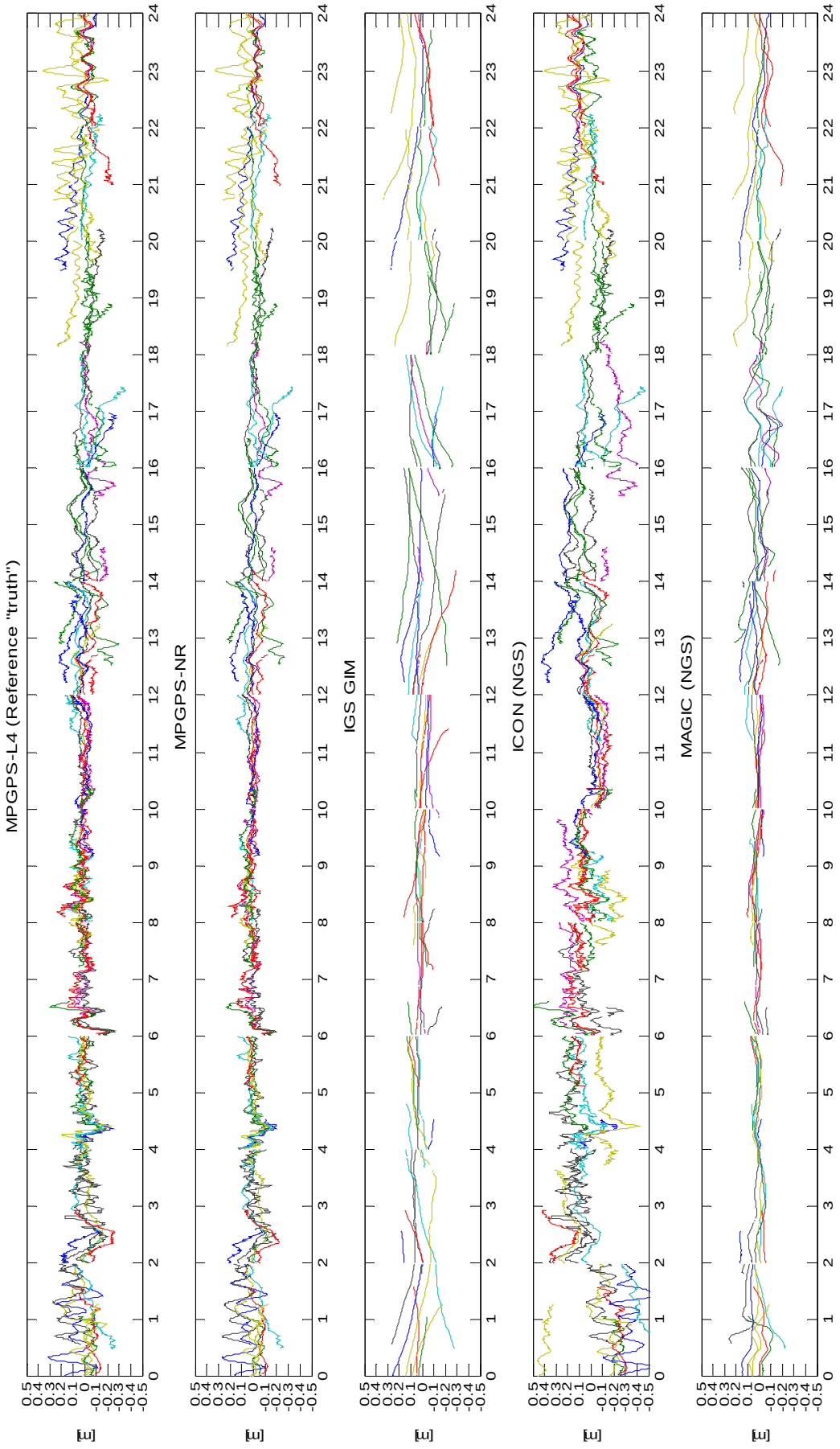


Fig. 6 Estimated DD ionospheric corrections for the analyzed models — 24 h, KNTN-DEFI (~100 km).

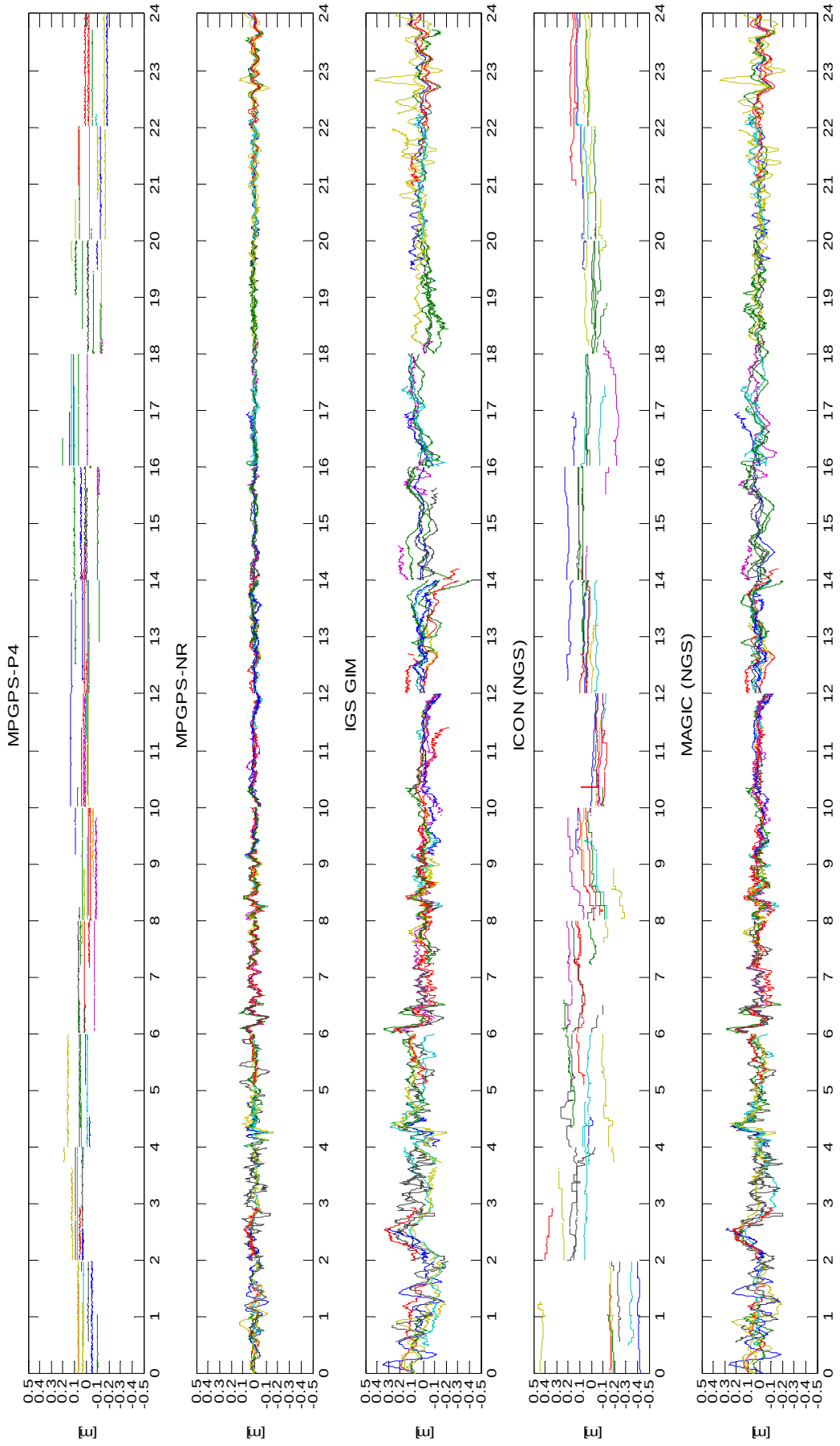


Fig. 7 DD ionospheric residuals with respect to the reference "truth" — 24 h, KNTN-DEFI (~100 km)

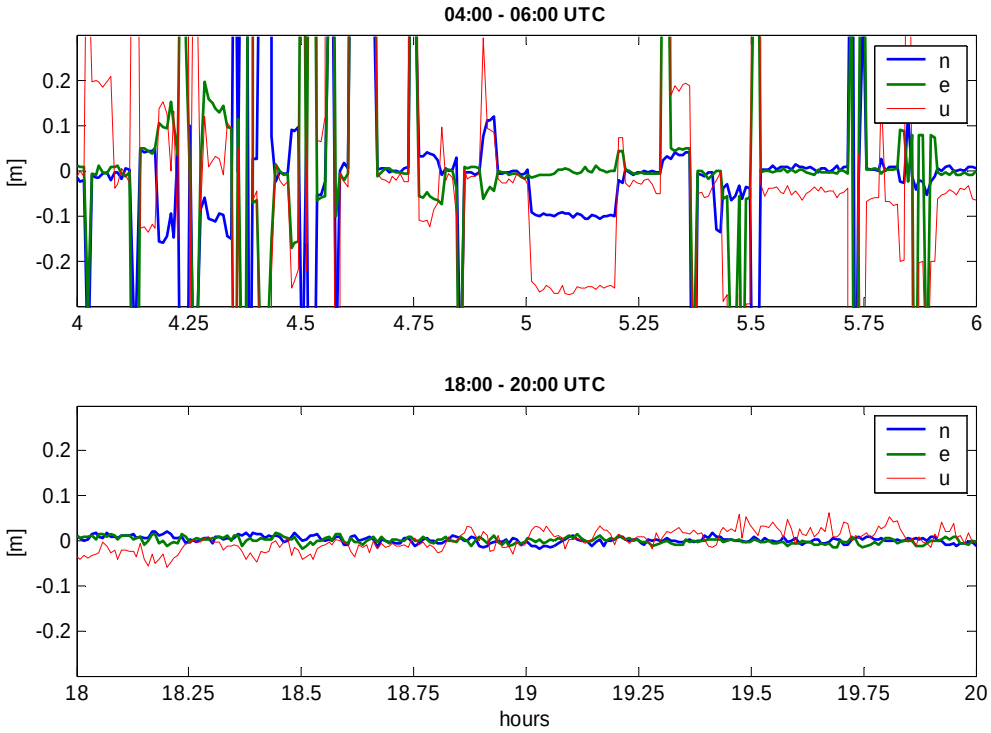


Fig. 8 Instantaneous RTK position residuals with respect to the known coordinates; 2-hour windows, KNTN-SIDN (~60 km); 0 cm constraint (1 sigma) was applied to the ionospheric corrections.

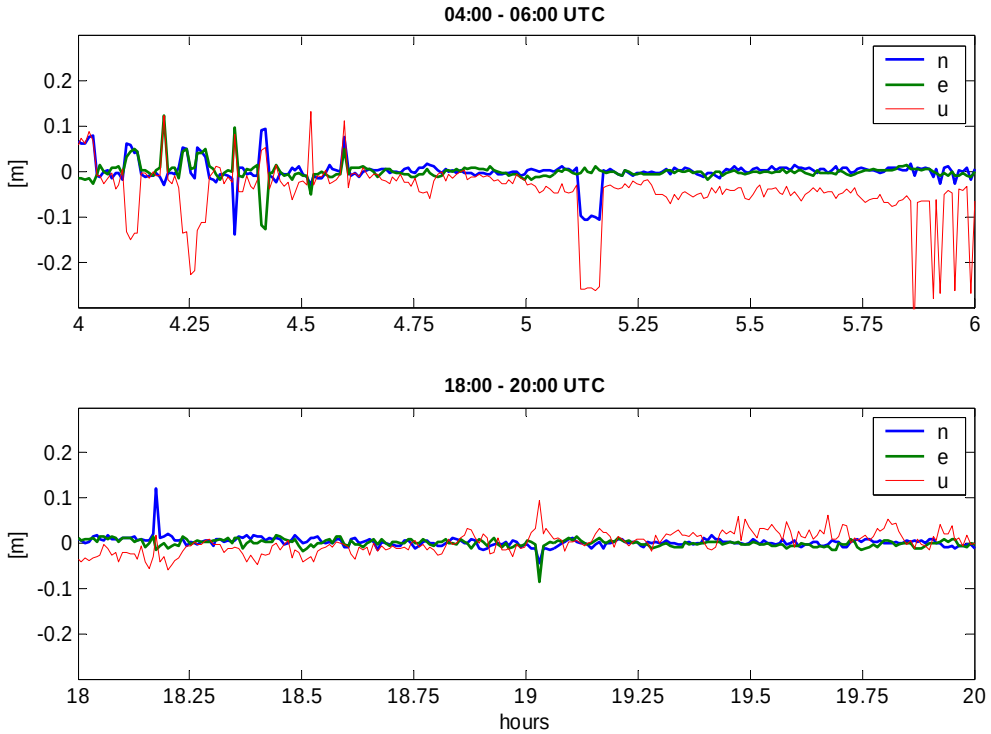


Fig. 9 Instantaneous RTK position residuals with respect to the known coordinates; 2-hour windows, KNTN-SIDN (~60 km); 5 cm constraint (1 sigma) was applied to the ionospheric corrections.

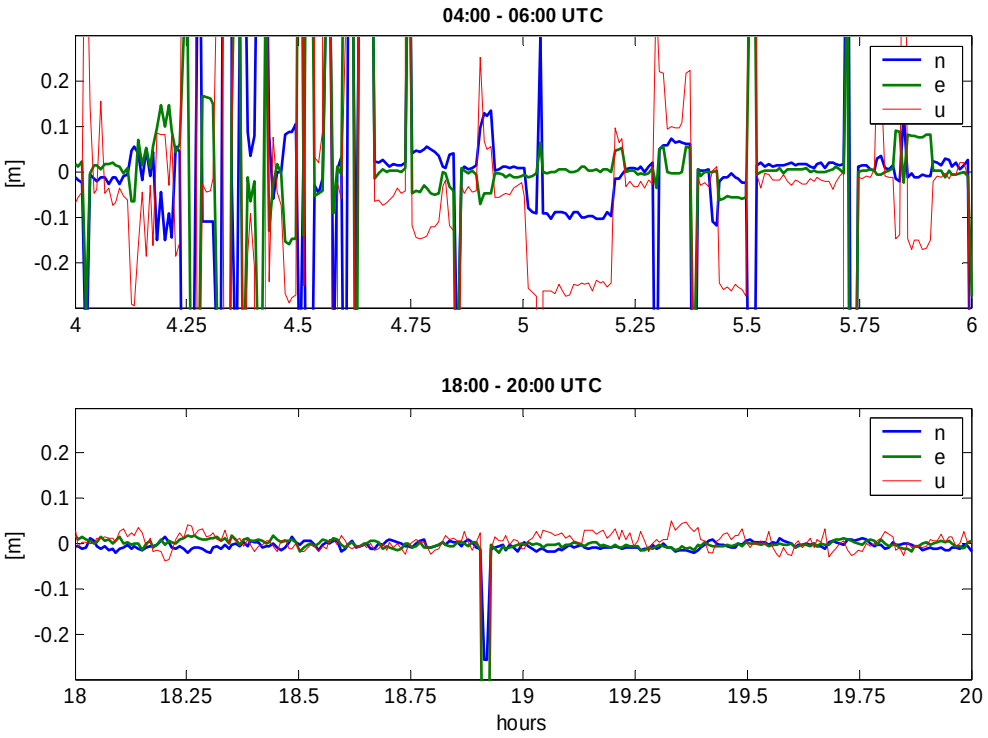


Fig. 10 Instantaneous RTK position residuals with respect to the known coordinates; 2-hour windows, KNTN-DEFI (~100 km); 0 cm constraint (1 sigma) was applied to the ionospheric corrections.

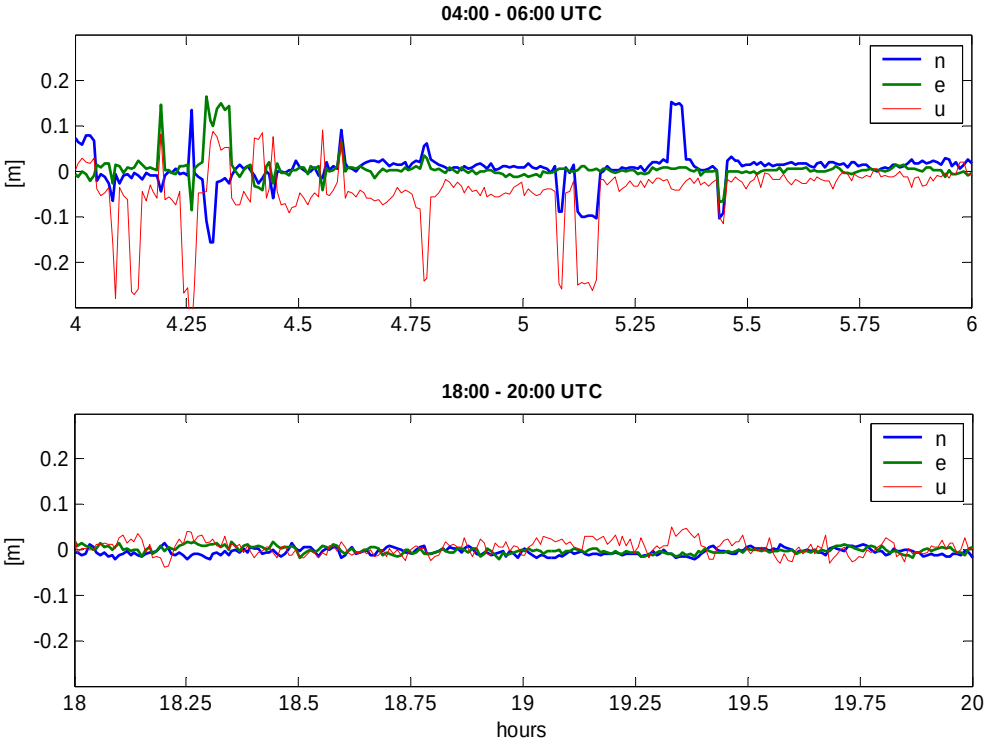


Fig. 11 Instantaneous RTK position residuals with respect to the known coordinates; 2-hour windows, KNTN-DEFI (~100 km); 1 cm constraint (1 sigma) for the ionospheric correction was applied to the “best” window (bottom) and 5 cm for the “worst” window (top).

Tab. 4 Mean and standard deviation (std) of DD ionospheric residuals with respect to the reference "truth" for two 2-hour windows and ~60 km baseline (KNTN-SIDN).

KNTN-SIDN (~60 km)											
04:00–06:00 UTC						18:00–20:00 UTC					
mean [m]						mean [m]					
PRNs	MPGPS P4	MPGPS NR	IGS GIM	ICON	MAGIC	MPGPS P4	MPGPS NR	IGS GIM	ICON	MAGIC	PRNs
28 - 4	-0.01	-0.00	0.02	-0.01	-0.01	0.04	0.00	0.06	0.07	-0.01	25 - 1
28 - 7	-0.03	0.01	-0.02	0.04	-0.01	0.16	0.01	0.07	-0.11	-0.01	25 - 2
28 - 8	0.01	0.01	0.06	-0.17	-0.00	0.18	-0.01	-0.03	-0.13	0.03	25 - 5
28 - 9	0.05	-0.00	-0.05	0.24	0.00	0.08	0.01	-0.00	-0.03	-0.01	25 - 6
28 - 11	-0.07	-0.00	-0.05	-0.06	-0.02	-0.04	0.01	-0.03	0.01	0.03	25 - 14
28 - 20	0.22	0.00	-0.03	0.08	0.01	0.09	0.01	0.04	-0.06	-0.00	25 - 16
28 - 24	-0.01	0.01	0.05	0.09	0.02	0.16	0.01	0.03	-0.06	0.01	25 - 20
						0.13	-0.00	-0.02	-0.15	0.02	25 - 23
						0.02	0.01	-0.04	-0.10	0.01	25 - 30
std [m]						std [m]					
28 - 4	0.00	0.02	0.03	0.01	0.03	0.00	0.01	0.02	0.00	0.02	25 - 1
28 - 7	0.00	0.06	0.07	0.02	0.07	0.00	0.02	0.03	0.02	0.03	25 - 2
28 - 8	0.00	0.04	0.05	0.01	0.04	0.00	0.01	0.01	0.01	0.01	25 - 5
28 - 9	0.00	0.03	0.04	0.01	0.04	0.00	0.02	0.02	0.00	0.02	25 - 6
28 - 11	0.00	0.04	0.04	0.01	0.03	0.00	0.01	0.02	0.01	0.03	25 - 14
28 - 20	0.00	0.04	0.05	0.04	0.04	0.00	0.01	0.02	0.01	0.02	25 - 16
28 - 24	0.00	0.02	0.03	0.01	0.03	0.00	0.01	0.03	0.01	0.03	25 - 20
28 - 4	0.00	0.02	0.03	0.01	0.03	0.00	0.02	0.03	0.01	0.02	25 - 23
						0.00	0.01	0.03	0.02	0.02	25 - 30
						0.00	0.01	0.02	0.00	0.02	25 - 1

Tab. 5 Mean and standard deviation (std) of DD ionospheric residuals with respect to the reference "truth" for two 2-hour windows and ~100 km baseline (KNTN-DEFI).

KNTN-DEFI (~100 km)											
04:00–06:00 UTC						18:00–20:00 UTC					
mean [m]						mean [m]					
PRNs	MPGPS P4	MPGPS NR	IGS GIM	ICON	MAGIC	MPGPS P4	MPGPS NR	IGS GIM	ICON	MAGIC	PRNs
28 - 4	-0.01	0.00	0.00	0.03	0.01	-0.10	0.01	0.08	0.06	-0.01	25 - 1
28 - 7	0.06	0.01	0.05	0.19	0.02	0.09	0.01	-0.02	-0.07	0.02	25 - 2
28 - 8	-0.03	0.01	-0.01	0.01	0.07	-0.14	-0.01	-0.02	-0.12	0.01	25 - 5
28 - 9	0.05	0.00	0.08	0.17	0.03	0.12	0.01	-0.01	-0.09	-0.01	25 - 6
28 - 11	0.04	0.00	0.13	0.07	0.06	-0.02	0.01	-0.04	-0.03	0.00	25 - 14

28 - 20	0.16	0.00	0.07	-0.12	0.01	0.03	0.01	-0.06	-0.01	-0.00	25 - 16
28 - 24	0.00	0.01	0.05	0.10	0.02	-0.13	0.01	0.04	0.04	0.02	25 - 20
						-0.13	-0.00	-0.13	-0.08	0.01	25 - 23
						-0.06	0.01	-0.04	-0.02	0.04	25 - 30
std [m]						**std [m]**					
28 - 4	0.00	0.02	0.05	0.01	0.06	0.00	0.01	0.02	0.01	0.03	25 - 1
28 - 7	0.00	0.06	0.07	0.03	0.07	0.00	0.02	0.04	0.01	0.04	25 - 2
28 - 8	0.00	0.04	0.09	0.01	0.10	0.00	0.01	0.03	0.01	0.03	25 - 5
28 - 9	0.00	0.03	0.05	0.02	0.05	0.00	0.02	0.03	0.02	0.04	25 - 6
28 - 11	0.00	0.04	0.07	0.01	0.08	0.00	0.01	0.02	0.01	0.02	25 - 14
28 - 20	0.00	0.04	0.08	0.03	0.06	0.00	0.01	0.04	0.01	0.03	25 - 16
28 - 24	0.00	0.02	0.05	0.02	0.04	0.00	0.01	0.02	0.01	0.03	25 - 20
28 - 4	0.00	0.02	0.05	0.01	0.06	0.00	0.02	0.04	0.03	0.05	25 - 23
						0.00	0.01	0.03	0.01	0.04	25 - 30
						0.00	0.01	0.02	0.01	0.03	25 - 1

4 Summary and conclusions

The analysis of the quality of the network-based ionospheric correction derived from five independent models was presented, and the applicability of these models to support instantaneous AR and kinematic positioning were tested and discussed. The CORS and/or IGS reference network GPS data were used to derive the model corrections. One of the CORS stations in Ohio was selected as an unknown rover location, and its coordinates were estimated in the simulated kinematic mode using the OSU-developed MPGPS™ software. The successful instantaneous AR was achieved with the MPGPS-NR model, whose quality (1 sigma) is estimated as 1–6 cm, as shown in Tables 4 and 5. Clearly, this accuracy, corresponding to around one quarter of the L1 cycle, can generally assure an accurate instantaneous AR for longer baselines; however, the level of success is a function of the level of ionospheric variability. The remaining models are currently tested for their applicability to OTF AR in terms of the speed of AR, i.e., the number of epochs needed to find integers and the resulting quality of the position coordinates. These findings will be reported in a subsequent publication.

Acknowledgement:

This project is supported by NOAA, National Geodetic Survey, N/NGS (project NA04NOS4000067).

References

Bilitza D. (1997): ***International Reference Ionosphere — Status 1995/96***. Advances in Space Research, Vol. 20, No. 9, 1751–1754.

Feltens J.; N. Jakowski (2002): ***The International GPS Service (IGS) Ionosphere Working Activity***. SCAR Report No. 21.

Bock Y.; de Jonge P.; Honcik, D.; Fayman, J. (2003): ***Wireless Instantaneous Network RTK: Positioning and Navigation***, Proceedings of ION GPS/GNSS, September 9–12, Portland, OR, pp. 1397–1405.

Grejner-Brzezinska D.A.; Kashani I.; Wielgosz P. (2004a): ***Analysis of the Network Geometry and Station Separation for Network-Based RTK.*** Proceedings of ION NTM, January 26–28, San Diego, CA, 469–474.

Grejner-Brzezinska D. A.; Kashani I.; Wielgosz P. (2004b): ***On Accuracy and Reliability of Instantaneous Network RTK as a Function of Network Geometry, Station Separation, and Data Processing Strategy***. Submitted to GPS Solutions, September 2004.

Hugentobler U.; Schaer S.; Fridez P. (2001): ***BERNESE GPS Software Version 4.2***. Astronomical Institute, University of Berne, Switzerland.

Kim, D.; Langley, R.B. (2000): ***GPS Ambiguity Resolution and Validation: Methodologies, Trends and Issues***. Proceedings of 7th GNSS Workshop and International Symposium on GPS/GNSS, Seoul, Korea, November 30 – December 2, 213–221.

Kashani I.; Grejner-Brzezinska D.A.; Wielgosz P. (2004a): ***Towards instantaneous RTK GPS over 100 km distances***. Proceedings of ION 60th Annual Meeting, June 7–9, 2004, Dayton, Ohio, 679–686.

Kashani I.; Wielgosz P.; Grejner-Brzezinska D.A. (2004b): ***The Effect of Double-Difference Ionospheric Corrections Latency on Instantaneous Ambiguity Resolution in Long-Range RTK***. ION GNSS 2004, September 21–24, Long Beach, California, 2881–2890.

Odijk D. (2000): ***Weighting Ionospheric Corrections to Improve Fast GPS Positioning Over Medium Distances***. Proceedings of ION GPS, September 19–22, Salt Lake City, UT, 1113–1123.

Odijk D. (2001): ***Instantaneous GPS Positioning Under Geomagnetic Storm Conditions***. GPS Solutions, Vol. 5, No. 2, 29–42.

Rizos C. (2002): ***Network RTK research and implementation: A geodetic perspective***. Journal of Global Positioning Systems, Vol. 1, No. 2, 144–150.

Schaer S. (1999): ***Mapping and Predicting the Earth's Ionosphere Using the Global Positioning System***. Ph.D. Thesis, Astronomical Institute, University of Berne.

Schaer S.; Gurtner W.; Feltens J. (1998): ***IONEX The IONosphere Map EXchange Format Version 1***. Proceedings of the IGS Analysis Center Workshop, ESA/ESOC, February 9–11, Darmstadt, Germany, 233–247.

Smith D.A. (2004): ***Computing unambiguous TEC and ionospheric delays using only carrier phase data from NOAA´s CORS network***. Proceedings of IEEE PLANS 2004, April 26-29, Monterey, California, 527-537.

Spencer P.S.J.; Robertson D.S.; Mader, G.L. (2004): ***Ionospheric data assimilation methods for geodetic applications***. Proceedings of IEEE PLANS 2004, Monterey, California, April 26-29, 2004, 510-517.

Springer T.A. (1999): ***Modeling and Validating Orbits and Clocks Using the Global Positioning System***. Ph.D. dissertation, Astronomical Institute, University of Berne, Switzerland.

Teunissen P.J.G. (1994): ***A new method for fast carrier phase ambiguity estimation***. Proceedings of IEEE PLANS'94, Las Vegas, NV, April 11–15, 562–573.

Teunissen P.J.G.; Joosten P.; Tiberius C. (2002): ***A Comparison of TCAR, CIR and LAMBDA GNSS Ambiguity Resolution***. Proceedings of the ION GPS 2002, 24–27 September 2002, Portland, OR, 2799–2808.

Vollath U.; Buecherl A.; Landau H.; Pagels C.; Wagner B. (2000): ***Multi-Base RTK Positioning Using Virtual Reference Stations***. Proceedings of ION GPS, 19–22 September 2000, Salt Lake City, UT, 123–131.

Wanninger L. (2002): ***Virtual Reference Stations for Centimeter-Level Kinematic Positioning***. Proceedings of ION GPS 2002, September 24–27, Portland, Oregon, 1400–1407.

Wielgosz P.; Grejner-Brzezinska D.A.; Kashani I.; (2003): ***Regional Ionosphere Mapping with Kriging and Multiquadric Method***s. Journal of Global Positioning Systems, Vol. 2, Issue 1, 48–55.

Wielgosz P.; Grejner-Brzezinska D.; Kashani I. (2004): ***Network Approach to Precise GPS Navigation***. Navigation, Vol. 51, No. 3, 213–220.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 132-142

Journal of Global Positioning Systems
2004
© CPGPS

Process for improving GPS acquisition assistance data and server-side location determination for cellular networks

Neil Harper

Nortel, Nortel Building, Northfields Ave, University of Wollongong, NSW 2500
e-mail: nharper@nortel.com; Tel: +61 2 42242800; Fax: +61 2 42242801

Peter Nicholson

Nortel, Nortel Building, Northfields Ave, University of Wollongong, NSW 2500
e-mail nichol@nortel.com; Tel: +61 2 42242800; Fax: +61 2 42242801

Peter Mumford

School of Surveying & Spatial Information Systems, University of New South Wales, Sydney, NSW 2052
e-mail: p.mumford@unsw.edu.au; Tel: +61 2 9385 4189; Fax: +61 2 9313 7493

Eric Poon

Nortel, Nortel Building, Northfields Ave, University of Wollongong, NSW 2500
e-mail: pooneric@nortel.com; Tel: +61 2 42242800; Fax: +61 2 42242801

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. This paper introduces a process for analysing and improving the quality of the acquisition assistance data produced for Assisted GPS (A-GPS) positioning in cellular networks. An experimental test bed is introduced and a series of experiments and results are provided. The experiments validate the GPS acquisition assistance data in open-sky conditions. Accuracy results for initial testing of our server location determination engine in a range of different environments are also given with results of a long-term run. Acquisition assistance data provides the GPS handset with information that allows it to detect the GPS signals more quickly and allows detection of much weaker signals. It does this by providing information to the handset about where to look for the signals. The A-GPS server is a mobile location server determining the location of devices within a cellular network. In order to measure the quality of the acquisition assistance data produced by the A-GPS server a piece of hardware has been developed called the "A-GPS Trainer". This takes assistance data and uses it to lock on to the satellites. It then provides code phase measurements back to the server for it to do a location calculation. The trainer also reports on the amount of time it takes to lock on to the signals of the individual satellites. It can run multiple calculations over a period of time and report the results.

Key Words: A-GPS, Assistance Data, GPS acquisition assistance data.

1. Introduction

Assisted-GPS is considered by many the best solution to meet the accuracy requirements of the U.S. Federal Communications Commission's E-911 mandate. The E-911 mandate requires telecommunications carriers to give emergency services providers accurate location information in a timely manner. Additionally, the accuracy of GPS facilitates Location Based Services that have good economic potential.

The two different paradigms in cellular networks are MS-Based Assisted-GPS and MS-Assisted Assisted-GPS. In MS-Based A-GPS, the network provides the handset with detailed assistance data to help the handset to lock on to the satellites and perform the position calculation.

In the case of MS-Assisted A-GPS, the network provides a smaller amount of assistance data to help the handset to lock on to the satellites. The handset then provides the satellite measurements to the network for the network to perform the position calculation.

There are advantages and disadvantages of both MS-Assisted A-GPS and MS-Based A-GPS. Some of the advantages of MS-Assisted A-GPS are that less

information is transmitted over the air interface, and there is potential on the server to use other information such as network timing information from base stations or a Digital Elevation Model. Another benefit is that the position calculation is centralised so that any software problem can be fixed in one place on the server instead of millions of customer handsets.

GPS acquisition assistance data is used by an MS-Assisted A-GPS handset to lock on to satellite signals. This paper describes a test bed and a process for tuning the server-side algorithm which produces the acquisition assistance data.

In an emergency situation, the time it takes to lock on to the satellites is very important and in this paper the time to lock on to the satellites in open sky conditions from a cold start is optimised. The results show that acquisition assistance data speeds up the acquisition by about 200 times with our hardware.

A measure called the 'acquisition assistance usefulness factor' is introduced. This factor is the ratio of the relative speed of the GPS fixes scaled by the proportion of the satellites in view that the device locks on to. It defines how good the acquisition assistance data is for a particular GPS fix.

The other major part of an MS-Assisted A-GPS solution is the position calculation function. In this paper some results of testing at the University of New South Wales (UNSW) and the results of a long-term run (18 months) are given.

2. Background

Nortel Networks manufactures various 3GPP-compliant cellular network elements. These elements include Serving Mobile Location Centers (SMLC), Radio Network Controllers (RNC) and Standalone A-GPS SMLCs (SAS). These elements generate GPS acquisition assistance data for cellular phones with embedded A-GPS capability. The GPS assistance data speeds up acquisition of GPS signals and increases GPS yield in weak signal environments.

It was necessary to develop a process and a test environment that allowed the effectiveness of the supplied GPS assistance data to be measured. The ultimate measure of the value of the assistance data is to make empirical observations on the impact to sensitivity and time-to-fix with an A-GPS receiver. The testing process and environment developed facilitates many such observations being made, ensuring a valid set of statistical measurements.

The test environment provides a means of measuring the effects of optimising the assistance data. By measuring the results of changes to the supplied assistance data, a validation can be performed to ensure that the data provides optimal results in practice. In addition, the test environment provides a practical regression test suite to ensure future versions of the assistance data calculation software produce correct assistance data.

Assistance data helps to significantly reduce the search space required by the GPS enabled handset (Zhao 2002, Djuknic and Richton 2001). Fast, accurate fixes can be made using a combination of good assistance data and high sensitivity receivers (van Diggelen and Abraham 2001, and Heinrich *et al* 2004).

Much of the literature concentrates on improvements to the handset. This paper is focused on improvements that can be made on the server.

3. GPS acquisition assistance data

The cellular network can provide assistance data to the A-GPS cellular handset through the air when it is in contact with a base station. Assistance data types that can be sent to the handset includes the ephemeris, reference time, ionosphere model, UTC model, real time integrity and acquisition assistance.

The format of the GPS acquisition assistance data is defined by the relevant standards body; the specifics of the encoding and field ranges depend on the access technology. The protocol for the GSM system for example is defined in 3GPP TS 04.21 and is referred to as RRLP (Radio Resource LCS protocol). Other protocols include Positioning Calculation Application Part (PCAP) for UMTS networks (between the SAS and the RNC) and Position Determination Service Standard for Dual Mode Spread Spectrum Systems (PDDM) for CDMA networks.

A high-level diagram of an SMLC network is shown in Fig. 1. In this scenario, the SMLC receives a request for assistance data for a particular handset. The request effectively includes the initial location estimate and the uncertainty of the estimate, which is based on the size of the serving cell.

This initial location estimate and the uncertainty are used to calculate the acquisition assistance data using information from the Wide Area Reference Network (WARN). This acquisition assistance data is used to populate the message that is sent back to the network and on to the handset.

The fields that the server provides are summarised in Tab. 1. They are named similarly and have similar ranges across the different network access technologies

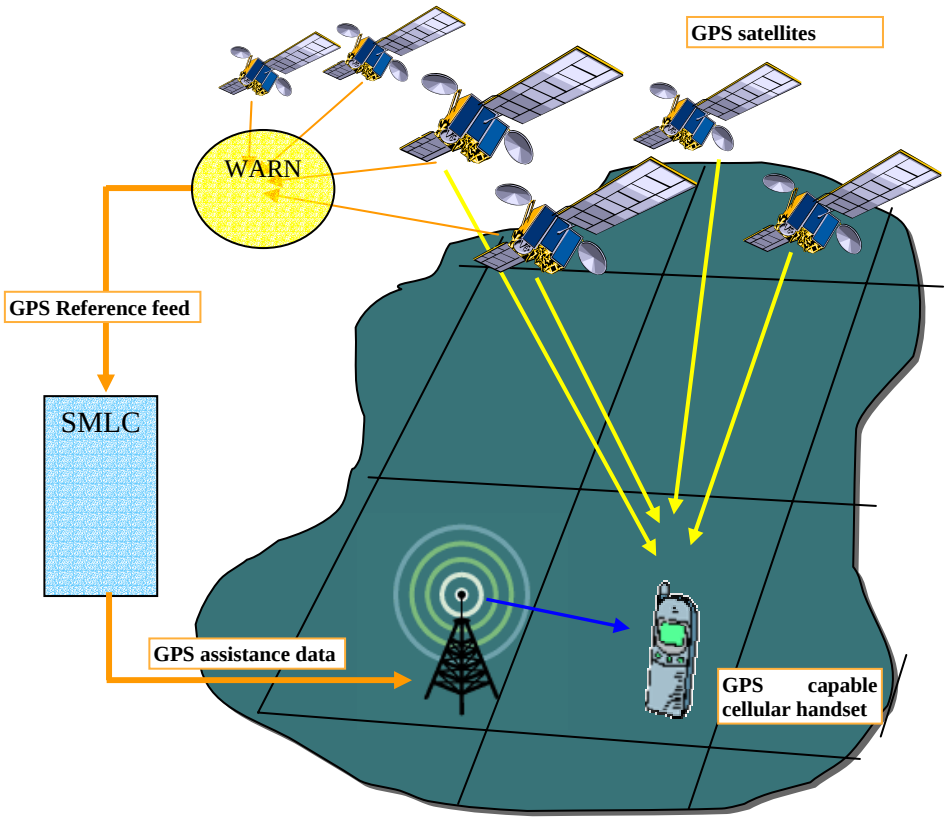


Fig. 1. A-GPS in a cellular network

Tab. 1. The fields for GPS Acquisition assistance data

Field	Approximate Range	Notes
GPS Time	0 to 604800.0	The GPS time of week in seconds at which the assistance data is relevant
Satellite Information		The following fields appear for each satellite in the message:
Satellite ID		The satellite ID
Doppler	-5120 to 5117.5 Hz	The predicted doppler at the specified time
Doppler (1st order)	-1 to 0.5 Hz	The rate of change of doppler
Doppler Uncertainty	12.5 to 200 Hz	The doppler uncertainty value
Code Phase	0 to 1022 chips	The centre of the code phase search window (in chips)
Integer Code Phase	0 to 19	The number of 1023 chip segments
GPS Bit number	0 to 3	The GPS bit number (20 1023-chip segments)
Code Phase Search Window	1 to 1023	The width of the search window for code phase
Azimuth	0 to 360	The horizontal distance in degrees of the satellite from North from the estimated location of the handset on the ground
Elevation	0 to 90	The angular distance of the satellite in degrees above the horizon from the estimated location of the handset on the ground

The server estimates the information including the search windows in the acquisition assistance data (Doppler uncertainty and code phase search window) using the initial location estimate and its uncertainty only. The

server makes no assumptions about the handset hardware since there may be many different types of handset in a network. When the handset receives the acquisition assistance data, it adjusts the windows based on its hardware.

For example, the measured Doppler is affected by the local handset TXCO (Temperature Controlled Crystal Oscillator). The size of the Doppler search window needs to be interpreted in relation to the handset's estimate of its local oscillator error. If a handset that has a local oscillator error of up to 100 Hz uses the predicted search window of 12.5 Hz, it is unlikely to find the satellite.

Similarly, the code phase values are highly dependent on the clock error of the handset at the millisecond level because the PRN replica is driven by the local handset clock. The handset needs to know its time accurately in order to make use of the code phase predictions as is. In a synchronised network, such as CDMA, the handsets are synchronised with the base stations, which are synchronised to GPS time, so the code phase can be used.

In a non-synchronised network, such as GSM, the handset needs to treat supplied code phase measurements as relative code phase offsets. To do this, it locks on to the first satellite using its Doppler information and a potentially large search in the time (code phase) domain. Once it locks on to this first satellite, it calculates the difference between the code phase measurement for this satellite and that supplied in the assistance data. This offset is then applied to all of the other code phase estimates in order to determine where to search for those satellites. In a non-synchronised network, it will take a longer time to lock on to the first satellite but once it gets that one, the narrow code phase search using the assistance data can be applied to the other satellites.

4. GPS Acquisition Results from logged GPS measurement data

This section reports on the results of calculating the GPS acquisition assistance data and comparing some of the fields in the acquisition assistance data against measurements that have been logged from a GPS receiver over a long period of time. The results show that the acquisition assistance data calculated is close to the measured values of the fields.

GPS measurements have been logged from a reference receiver for more than 18 months and approximately 130 million measurements of individual satellites have been made. The GPS reference receiver tracks in open-sky conditions and operates continuously.

Software reads through each of these measurements and, for each observation, compares the calculated GPS acquisition assistance data with the measurements. The calculated Doppler is compared with the measured Doppler and the predicted range (used to estimate the code phase for acquisition assistance data) with the measured pseudorange. The predicted range is compared with the measured pseudorange because the reference receiver does not supply code phase. Before they can be compared, the predicted range is converted into a predicted pseudorange by applying the clock offset of the receiver calculated by the position calculation function. This means that both ranges are offset by the same clock error and can be compared.

The results in Tab. 2 show that there is a very small difference in the calculated Doppler and range compared to the measured values. This gives us confidence that the basic algorithm used to calculate the Doppler and range (from which code phase is derived) is correctly coded and that the assistance data can be tested using the A-GPS Trainer described in the next section.

The first column Tab. 2 shows the difference between the predicted Doppler and the measured Doppler for over 130 million individual satellite observations. The average error is 0.46 Hz with a standard deviation of 0.35 Hz. There were 15 outliers that had more than 10 Hz of error and they have been split out into a separate statistic shown in the column titled 'Doppler error over 10 Hz'. There is presently no real-time integrity information such as provided by the Wide Area Augmentation System (WAAS) to determine whether there was a problem with the calculation, measurement or satellite for these outliers.

The predicted range error is shown as 15.21 metres with a standard deviation of 12.47 metres. There were 3 records with an error over 150 metres and they are shown in the last column of the table.

Tab. 2. The difference between calculated and measured values

	Doppler error	Doppler error over 10 Hz	Predicted range error	Predicted range error over 150 metres
Number of Records	130667033	15	130667045	3
Average	0.46	393.43	15.21	4363.20
Standard Deviation	0.35	1096.96	12.47	7286.02
Minimum	1.4E-09	10.08	3.97	155.97
Maximum	9.96	4341.67	149.82	12776.37

5. A-GPS Trainer

The A-GPS Trainer is a custom device designed and developed by the U NSW Satellite Navigation and Positioning group within the School of Surveying & Spatial Information Systems in consultation with Nortel Networks. It is a device for testing GPS acquisition assistance data.

The device consists of a Sigtec MG5001 GPS receiver running a modified version of the Mitel GPS Architect software.

The Mitel software has been extensively modified to facilitate control and reporting of signal acquisition. A custom interface has been developed that allows the device to run in autonomous GPS mode or in A-GPS mode. This alows a comparison of the times to make different types of fixes. A model of the device is shown in Fig. 2.

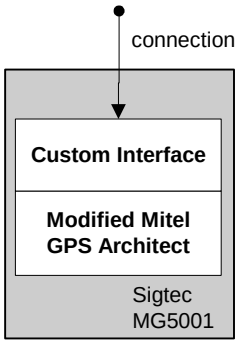


Fig. 2. A-GPS Trainer model

During a typical measurement session, the A-GPS Trainer is sent acquisition assistance data and then attempts to acquire and track those satellites and report back. The Doppler search begins at the given Doppler value and searches each side of this value until the maximum width is reached. The size of this Doppler search window depends on the estimated receiver clock error as well as the Doppler window size given in the assistance data. If a signal is not detected, the search starts again.

The code search uses the standard Mitel GPS Architect method, and starts at one end of the window and slides across to the other. An improvement could be made on this method by starting the code search at the given code phase value and then searching more widely on each side until the search window is reached. This improvement is possible because the true location of the handset is more likely to be at the centre of the search window than near the boundary so the lock should be made faster. This is a device improvement however and since the focus of this paper is measuring server-side improvements it is not described further.

The information reported at the end of a measurement session includes (for each satellite allocated) pseudorange, the time taken to acquire locks on the signal, Doppler, code phase and SNR.

The A-GPS Trainer treats the code phase values as absolute code phase values. This means that the code phase search uses the code phase values in the assistance data as absolute values in order to centre the search. The problem is that the generation of the code phase replica in the handset is driven by the handset clock which is subject to error. If the clock is out by just half of a millisecond then the code phase will be out by half of its range of 511 chips.

The handset needs to treat the supplied code phase values as relative values. It will do a complete search in the time domain and when it locks on to the first satellite, the code phase for the centre of the search windows for the other satellites can be set relative to that satellite using the relative differences in the supplied acquisition assistance data.

The A-GPS trainer presently does not have this facility which means that the code phase is of less value and the lock is slightly slower than it could be.

6. A-GPS Train Driver tool

The A-GPS Train Driver is a software tool that calculates the acquisition assistance data using the ephemeris for the satellites in view extracted from an independent GPS receiver that has already locked on to the satellites. The A-GPS Train Driver then provides the GPS acquisition assistance data to the A-GPS Trainer.

The A-GPS Train Driver parses the information returned by the A-GPS Trainer and reports on the Carrier and Code (CC) and Carrier, Code, Frame and Bit (CCFB) lock times of individual satellites in view. The Train Driver can also invoke the position calculation function using the GPS code phase measurements made by the A-GPS Trainer.

Fig. 3 shows the connection between the various components of the test system. On the right hand side is an image of the system in the lab.

In order for the A-GPS Trainer to make use of the code phase supplied in the assistance data, a synchronised network is crudely simulated. In a synchronised network, the handset has accurate time. In order to simulate this using the A-GPS Trainer, two pieces of information are used.

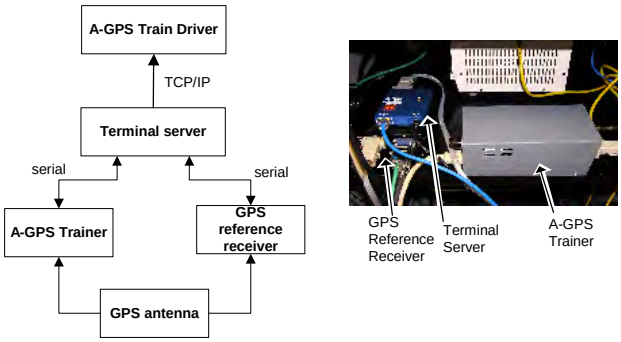


Fig. 3. A-GPS Train Driver tool and its connections

Firstly, a statistical measure is made of the time that it takes to generate the acquisition assistance data up until the A-GPS Trainer acknowledges the receipt of the assistance data. The time for which the acquisition assistance data is calculated is advanced using this information.

Secondly, information about the clock in the handset is used to artificially adjust the code phase. To do that, a autonomous position calculation is performed, which allows the A-GPS Trainer to update its clock. This is followed by a series of GPS fixes where assistance data is supplied. After each fix using the assistance data, the position calculation function is invoked to calculate the clock offset of the A-GPS Trainer. This offset is applied to the code phase measurements sent to the A-GPS Trainer for the next assistance data based request.

Fig. 4 shows graphically the simulated synchronised network process. Note that the A-GPS Trainer is run on a machine that has a Network Time Protocol (NTP) server attached in order for it to have accurate time.

7. Quality Measurement of Assistance Data

The model for analysing the acquisition assistance data and collecting data is shown in Fig. 5. Time increases along the axis from left to right.

Initially, the A-GPS Trainer is issued with a “clear” command so that the next GPS fix will be a cold start. The A-GPS Trainer then performs an autonomous GPS fix. This gives a cold-start baseline for the amount of time that the A-GPS Trainer takes to lock on to the satellites without assistance data.

Another “clear” command is then issued and acquisition assistance data is calculated before supplying it for a fix that starts at time *T1*. The A-GPS Trainer locks on to as many satellites as possible and the process completes at time *T2*. The A-GPS Trainer then reports on its results.

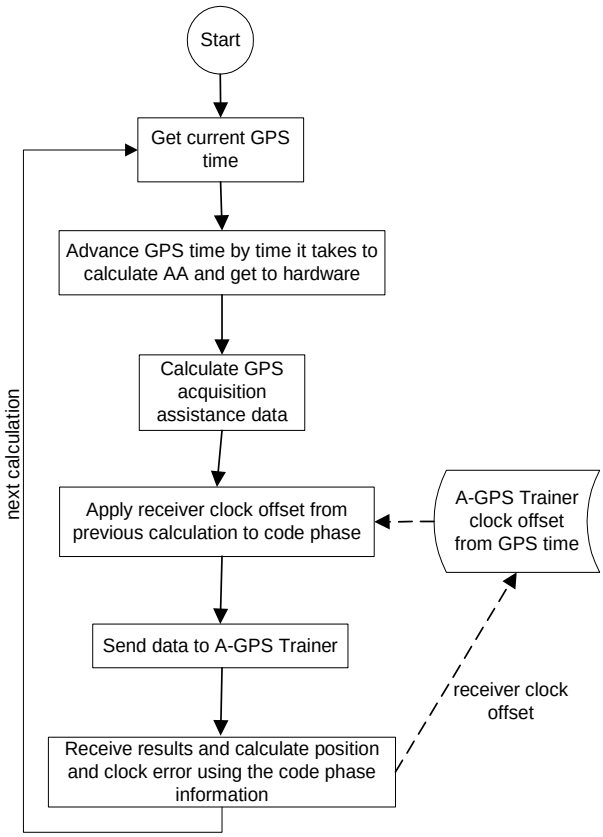


Fig. 4. Synchronised network simulation using the A-GPS Train Driver

The acquisition assistance data is calculated again but this time for time *T2*. This is so that the measured satellite information can be compared with the assistance data for the time when the measurement was actually made. During testing, there is limited value in comparing the measured values against the assistance data at *T1* because there are cases when the A-GPS Trainer is not able to lock on to one or more satellites and the difference between *T1* and *T2* can be up to 30 seconds. This large elapsed time is due to the fact that the A-GPS Trainer is configured to lock on to all of the satellites supplied and for some of the test cases satellites that are not in view are supplied.

When analysing the results, *T1* is only looked at for sanity and a comparison between the calculations at *T2* and the measured values are made.

From this data, there are two measures of the quality of the acquisition assistance data. One is the relative amount of time that it takes to lock on to the first four satellites between a fix with assistance data and a fix without assistance data. The other is the difference between the assistance data calculated at time *T2* and the measurements reported by the A-GPS Trainer.

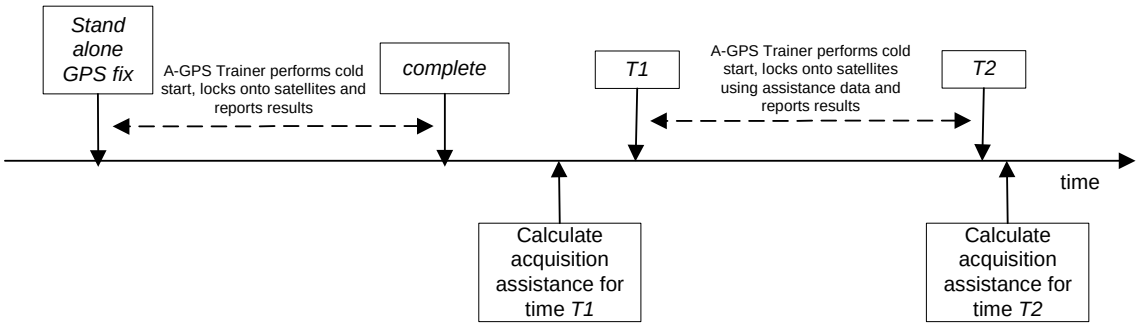


Fig. 5. Data Capture Cycle

From the relative times between the autonomous fix and the assisted fix, the speed-up factor is calculated using Equation 1. This is a measure of the number of times faster a fix with assistance data is.

Note that the time for the first four satellites to get CCBF lock is used for an autonomous fix. An autonomous fix needs to get frame lock in order to get the ephemeris data. This contrasts to a fix with assistance data which only needs CC lock since it only needs to measure the code phase and Doppler information and return this to the network to do the position calculation.

$$SpeedUpFactor = \frac{CCBFstndAloneTime}{CCassistedTime} \qquad (1)$$

A new measure called the Acquisition Assistance Usefulness Factor is defined. This measure weights the speed up factor by the number of satellites that it manages to lock on to when provided with assistance data. This then gives a measure of how good the assistance data is across all the satellites in view of the antenna. The Acquisition Assistance Usefulness Factor is calculated using Equation 2.

$$AcqAssistUseFactor = SpeedUpFactor * \frac{numberOfSatsLockedOn}{numberOfSatsInView} \qquad (2)$$

The Acquisition Assistance Usefulness Factor is the same as the speed up factor when the device locks on to all the satellites in view. The scale factor part of Equation 2 reduces the speed up factor relative to the number of satellites that it locks on to.

Due to the constantly changing satellite constellation, the time that a GPS fix takes can vary significantly from one period to another, especially in the case of a cold-start autonomous fix. So all measurements are made multiple times and the averages are reported.

8. Improving assistance data

The process used is a feedback loop where results are analysed, and then algorithms and code are modified based on this analysis.

The parameters are varied and the results are used for an analysis phase before making improvements to the algorithms.

The general model is shown in Fig. 6. The full set of scenarios are run, results are analysed over all the scenarios and areas of potential improvement identified. Areas of potential improvement include:

Satellites that don't achieve lock

Satellites that take a long time to lock on

Satellites where the assistance data at *T2* does not match what the A-GPS Trainer measures

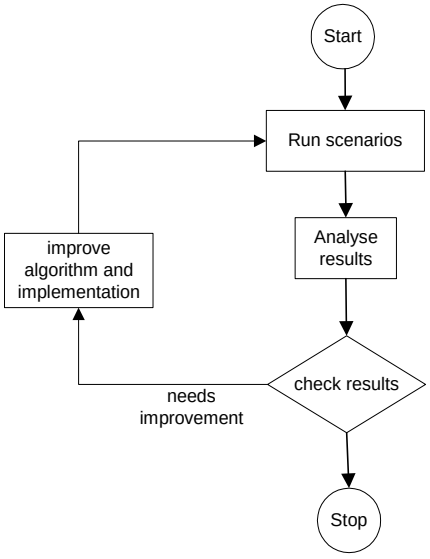


Figure 6. Process

The experiments described in the rest of this paper deals with how the initial location affects the GPS fix. The experiments are focused in the following areas with different sizes of the circle surrounding the initial location estimate:

Testing with the centre of the initial location estimate co-

located with the antenna

Testing individual pieces of assistance data by zeroing the Doppler or the code phase values

Testing with the antenna within the initial location estimate but in different horizontal locations and altitude relative to the antenna

Testing with the centre of the initial location estimate a fixed distance from the antenna where the antenna is outside of the initial location estimate

9. Results

The A-GPS Trainer tests shown in this section were performed using a simulated synchronised network and the results are an average of 100 measurements. The results shown in this section have already been through the process discussed in Section 8.

During the experiments the improvement process described in Section 8 was found to be useful for identifying and tracking down problems with the custom developed firmware of the A-GPS trainer. Once the problems with the A-GPS Trainer were identified using the process and subsequently fixed, then acquisition of the satellites in open-sky conditions "just worked". That is, the A-GPS Trainer was able to lock on to the satellites very quickly.

Further work needs to be done in weak signal environments in order to find areas of improvement to the process and the GPS acquisition assistance data.

9.1 Tests with the GPS antenna at centre of initial location estimate

During this test, the GPS antenna is at the centre of the initial location estimate ellipse. The radius of the circle for the initial location estimate is 5000 metres. The results in Tab. 3 show that it is 185.6 times faster to lock on to the satellite using the GPS acquisition assistance data compared to an autonomous GPS fix. The data shows that over the 100 tests, it never misses any of the satellites that are in view of the antenna.

The difference between the Doppler calculated at *T2* and the measured Doppler is out by 198 Hz with a small standard deviation of 7 Hz. This is due to the local oscillator error of the TXCO and is quite consistent over time as shown by the small standard deviation. The code phase difference on the other hand varies significantly because it is affected more by the clock error of the device, which is quite variable.

Tab. 3. The GPS lock information for GPS antenna at the centre of the location estimate

Avg autonomous time seconds	Avg time using acquis seconds	Speed-up Factor	Avg number of Sats in view	Avg number of Sats missed	Acquisition Assistance Usefulness Factor	Avg Doppler diff (stdev)	Avg code phase diff (stdev)
185.8	1.0	185.6	10.0	0.0	185.6	198 (7)	171 (245)

9.2 Tests for relative contribution of Doppler and code phase

During this test, information is provided to the A-GPS Trainer that contains zero for either code phase or Doppler fields of the acquisition assistance data. It can be seen that code phase has relatively low importance to the acquisition of the signal for the acquisition algorithm used in the A-GPS Trainer. The A-GPS Trainer uses a code sweeping technique which is driven by the Doppler search. The A-GPS Trainer searches each Doppler and then within each Doppler searches for the code. A more sophisticated GPS receiver may employ a matched-filter based implementation which searches all possible code phases simultaneously (Syrjarinne 2000).

The acquisition assistance usefulness factor is slightly lower because a larger search in the code phase domain is required and very occasionally the A-GPS Trainer is unable to lock on to one of the satellites.

Doppler on the other hand is very important for the A-GPS Trainer because if it is set to zero then no fix is possible because it does not lock on to the satellites. This is because the searching algorithm in the A-GPS Trainer is driven by the Doppler.

The results for zero Doppler show that the average number of satellites in view is 9.3 and it misses 8.1 of them. Further analysis of the data shows that the satellites that the A-GPS Trainer could acquire were the ones with

Doppler around 0 Hz. A potential improvement to the test tool is to improve the algorithm to supply incorrect Doppler, for example, setting the Doppler in the assistance data to 5000.0 when the doppler is around 0 results in the A-GPS Trainer being unable to lock on to any satellite.

Tab. 4. The GPS lock information for zero code phase and Doppler

	Avg autonomous time secs	Avg time using acquis secs	Speed-up factor	Avg no of sats in view	Avg no of Sats missed	Acquisition Assistance Usefulness Factor	Avg Doppler diff (stdev)	Avg Code phase diff (stdev)
Zero code phase	144.3	1.2	122.7	10.1	0.1	121.5	216 (3)	343 (204)
Zero Doppler	206.3	No Fix	N/A	9.3	8.1	N/A	N/A	N/A

9.3 Tests for offset in the initial location

In this test, the A-GPS Trainer is provided with acquisition assistance data calculated at different distance offsets from the true location of the antenna. The results show that there is very little difference in the A-GPS Trainer's ability to lock on when the acquisition assistance data is calculated for a location away from the receiver until the distance error is more than 100 kilometres. It is at this point that the acquisition assistance usefulness factor drops dramatically because the device can't lock on to all the satellites. Even though the A-GPS Trainer only needs to lock on to a minimum of 4 satellites, this is still a problem because if any of the useful satellites are obscured by obstacles then the handset won't be able to lock on to them and hence enable a location fix.

Tab. 5. The GPS lock information for GPS antenna different distances from the location estimate

Initial location offset distance (direction)	Avg autonomous time secs	Avg time using acquis secs	Speed-up factor	Avg no of sats in view	Avg no of sats missed	Acquisition Assistance Usefulness Factor	Avg Dopp diff (stdev)	Avg code phase diff (stdev)
5 Km (altitude)	215.7	1.2	173.2	10.1	0.5	165.2	220 (17)	295 (265)
10 Km (horizontal)	183.2	1.1	169.9	10.6	0.3	165.3	209 (18)	193 (239)
50 Km (horizontal)	113.4	1.4	83.3	10.4	0.5	79.2	198 (168)	354 (239)
100 Km (horiz)	151.5	1.1	135.0	10.0	0.1	133.4	219 (45)	272 (237)
500 Km (horiz)	241.4	No Fix	N/A	9.3	8.7	N/A	251 (136)	327 (264)
1000 Km (horiz)	169.5	4.7	36.3	10.8	5.6	17.5	278 (155)	348 (234)

10. Position Calculation Accuracy Results

The other major component of an MS-Assisted A-GPS solution is the position calculation function. The position calculation function is a combined least squares implementation consisting of a mathematical model and stochastic model as described in Parkinson *et al* (1996) and Harvey (1998).

10.1 Testing at UNSW

The position calculation function was tested against ground truth and a post-processing software package called "P4" at points around the UNSW Campus. The

School of Surveying and SIS has established a network of ground points with known coordinates (to centimetre accuracy) over a wide range of environments.

The position calculation function determined coordinates over a subset of these points with reasonable GPS signal conditions complied with the E-911 regulation statistical method (as described in OET BULLETIN No. 71 (2000)). The comparison between the position calculation function and "P4" established the position calculation function as a good performer on real world data as it outperformed P4 on 76% of epochs in which both had a solution. The position calculation function had an average position error of 12 metres with a standard deviation of 34 metres over a total of 2500 GPS measurements on 42 ground points.

10.2 Long-term run

The position calculation has been running since September 2003. A test tool gets the ephemeris and ionosphere data at startup and every 10 minutes while it is running from a NovAtel GPS receiver on the roof at Nortel Networks Wollongong. Once the test tool has the ephemeris and ionosphere data, it requests pseudoranges and invokes the position calculation function to calculate the location once per second.

The current statistics of this process are as follows. The position calculation uses the pseudorange from all the satellites in view, which is generally 8 to 12 satellites:

- Total number of position calculations: 30717770
- Average error in metres from ground truth: 4.08
- Standard deviation of the error in metres from ground truth: 2.38
- RMS error: 4.72
- Minimum number of metres of error from the ground truth: 0.001
- Maximum number of metres of error from the ground truth: 62.8

11. Conclusion

In this paper, a simple process has been described for improving the quality of GPS acquisition assistance data supplied to a handset. A test bed to facilitate the testing has been introduced and a series of experiments to show how acquisition assistance data can enable a fast location fix in open-sky conditions have been described. Some results of our position calculation are also given.

With GPS acquisition assistance data, the hardware can lock on to satellites up to 200 times faster than without GPS acquisition assistance data. The usefulness of acquisition assistance data degrades when the location of the handset is not known to within 100 kilometres.

The process for improving the quality of the GPS acquisition assistance data proved useful in identifying problems with the A-GPS Trainer firmware. Once these problems were fixed, the GPS acquisition assistance provided to the A-GPS enabled it to lock on to all of the satellites in view very quickly. This demonstrates that the GPS acquisition assistance that is supplied to the A-GPS Trainer is valid. It also validates our interpretation of the 3GPP specification against an independent implementation. This will reduce the chance of interoperability problems with A-GPS enabled cellular handsets.

A new test bed has been recently acquired which includes a Motorola FS Oncore development kit. The Motorola hardware is much more typical of the hardware in a cellular phone. The hardware is being interfaced to our test bed and will be the basis for further tuning of our GPS acquisition assistance data generation algorithms. It is also more portable, so weak signal conditions will be analysed. We will also have the opportunity to compare response times between MS-Based A-GPS and MS-Assisted A-GPS.

Acknowledgements

Chris Rizos (UNSW) for assisting with numerous technical issues, especially when the position calculation function was being developed. Martin Dawson (Nortel) for his suggestions for the focus of this paper and management support. Stewart Needham (Nortel) for valuable suggestions about the draft and management support. David Evans (Nortel) for management support. John Walsh (Nortel) for editing the final version of the paper.

References

3GPP TS 04.31 3rd Generation Edge Project; Technical Specification Group GSM/EDGE Radio Access Network; ***Location Services (LCS); Mobile Station (MS) – Serving Mobile Location Centre (LCS) Radio Resources LCS Protocol (RRLP)***

3GPP TS 25.453 3rd Generation Edge Project; Technical Specification Group Radio Access Network; ***UTRAN Iupc interface Positioning Calculation Application Part (PCAP) signalling***

3GPP2 C.S0022-0-1 3rd Generation Partnership Project 2; ***Position Determination Service Standard for Dual Mode Spread Spectrum Systems (PDDM)***

Djuknic G, Richton R (2001), ***Geolocation and Assisted GPS***, IEEE Comminications, February 2001

Harvey BR (1998), ***Practical Least Squares for Statistics and Surveyors,*** The University of New South Wales School of Geomatic Engineering Monograph 13

Heinrich G, Eissfeller B, Pany T, Weigel R, Ehm H, Schmid A , Neubauer A , Rohmer G , Ávila-Rodríguez J, ***HIGAPS A Highly Integrated Galileo/GPS Chipset for Consumer Applications***, GPS World September 2001

Syrjarinne J, ***Possibiliies for GPS Time Recovery with GSM Network Assistance***, ION GPS 2000, 19-22 September 2000, Salt Lake City, UT

Parkinson and Spilker, Eds., Axelrand and Enge, Assoc. Eds (1996) ***Global Positioning System: Theory and Applications Volume I***,American Institute of Aeronautics and Astronautics

van Diggelen F and Abraham C (2001) ***Indoor GPS Technology***, CTIA Wireless-Agenda, Dallas, May 2001

Zhao Y (2002), ***Standardization of Mobile Phone Positioning for 3G Systems***, IEEE Communications Magazine, July 2002

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 143-153

Journal of Global Positioning Systems
2004

Recent Improvements to the StarFire Global DGPS Navigation Software

Ronald R. Hatch

NavCom Technology, Inc., 20780 Madrona Avenue, Torrance, CA 90744
e-mail: rhatch@navcomtech.com, Tel: 310.381.2603, Fax: 310.381.2001

Richard T. Sharpe

NavCom Technology, Inc., 20780 Madrona Avenue, Torrance, CA 90744
e-mail: tsharpe@navcomtech.com, Tel: 310.381.2601, Fax: 310.381.2001

Received: 6 Dec. 2004 / Accepted: 2 Feb. 2005

Abstract. A general review of NavCom Technology's StarFire Global DGPS system is followed by a description of a number of improvements which have been either recently introduced or are in the process of being introduced. These improvements include: (1) an improved mode switching between various differential aiding signals and between dual-frequency and single-frequency operation when the L2 signal is lost; (2) a high-rate, high-accuracy, and efficient time-difference of carrier-phase position propagation process, which is used to generate the position coordinates between the one-second epochs; (3) an improved RAIM measurement error detection process; (4) a simplified process of computing the earth tides caused by both the sun and the moon; and (5) a built-in RTK capability (referred to as RTK Extend) which can make use of the synergism between the Global and RTK correction streams to continue RTK accuracy for up to 15 minutes when the RTK corrections are lost due to obstructed line-of-site or other problems with the local RTK corrections. Each of these will be addressed at least briefly. The more significant improvements will be addressed at greater length.

Key words: Global DGPS, RAIM, RTK, StarFire

1 Introduction

The StarFire Global Differential GPS (GDGPS) system provides sub-decimeter horizontal navigation accuracy anywhere in the world other than the two polar regions. The StarFire GDGPS system was developed at NavCom Technology and Ag Management Solutions, both John Deere companies, primarily for agricultural and marine applications. However, the unprecedented accuracy and global coverage have resulted in a multitude of applications, both commercial and military.

The StarFire system depends upon very accurate corrections to the satellite clocks and to the broadcast satellite orbits. These corrections make use of orbit prediction technology developed at the Jet Propulsion Laboratory (JPL) for the National Aeronautics and Space Administration (NASA). NavCom has licensed the Real Time GIPSY (RTG) software developed at JPL for use in the generation of real-time clock and orbit corrections for all the GPS satellites. In addition, NavCom has contracted to receive the data from NASA's dual-frequency reference receivers located around the world. These tracking sites are augmented by more than 23 NavCom installed and maintained sites, bringing the total number of global tracking sites to more than 60.

The StarFire user navigation receivers are high-accuracy, dual-frequency receivers. Both the receiver hardware and the internal navigation software were developed by NavCom to meet stringent accuracy goals. By using dual-frequency receivers, the effects of ionospheric refraction can be removed directly. This eliminates the largest source of error which competing Wide-Area Differential GPS (WADGPS) systems suffer.

As indicated above, the root-mean-square (rms) horizontal accuracy in each axis is better than 10 centimeters. The rms vertical accuracy is normally better than 20 centimeters, and one of the improvements described below improves this accuracy significantly. At NavCom continuing improvements are being made to enhance the robustness and accuracy. After describing the overall system, several of these improvements will be

considered in detail.

2 StarFire Overview

In addition to the GPS system itself, the StarFire GDGPS system is composed of seven major components. These are:

- Reference Network – reference receivers that continuously provide the raw GPS observables to the Hubs for processing. These observables include dual frequency code and carrier measurements, ephemerides, and other information.
- Processing Hubs – facilities at which the GPS observables are processed into DGPS corrections. There are two geographically separated, independent Hubs that operate fully in parallel, with each continually receiving all the measurement data and each computing corrections that are sent to the uplink facilities for the satellites. The Hubs are also the control centers for StarFire, from which the system operators monitor and manage StarFire. An automated Alerts System is used to advise operators of problem conditions as soon as they arise so that corrective actions may be taken.
- Communication Links – provide reliable transport mechanisms for the GPS observables and for the computed corrections. A wide variety of links are used to ensure that data are continuously available at the Hubs and that corrections can always be provided to the Land Earth Stations (LES) uplink sites.
- Land Earth Stations – satellite uplink facilities that send the correction data received from the Hubs to the geostationary satellites. The StarFire equipment at the LESs also makes decisions on which of several streams of correction data are best and should be broadcast.
- Geostationary Satellites – used to distribute the corrections to users via L-band broadcasts. The corrections from the LESs are broadcast on L-band frequencies to the users. Three Inmarsat geostationary satellites are used to provide correction coverage over most of the Earth (the areas North of +76 degrees latitude and South of –76 degrees latitude are not covered by geostationary satellites).
- Monitors – user receivers distributed throughout the world that use the broadcast corrections and provide their navigation information to the Hubs in real time. The Monitors are used by StarFire to continuously observe the operation of the system and provide immediate automated feedback about any problems that may arise. The Monitors are intended to behave just like user receivers in the field.
- User Equipment – uses the broadcast corrections along with local GPS observables to produce very precise navigation. The user equipment makes dual frequency GPS observations that remove ionospheric effects, and these refraction-corrected observations are combined with the broadcast corrections in a Kalman filter.

These system elements are illustrated in Figure 1, which shows in schematic form the overall system architecture. The monitor sites are collocated with the NavCom reference sites. The system concept is similar to the regional augmentation systems being deployed by various governments, e.g. the FAA's Wide Area Augmentation System (WAAS), the European EGNOS and Japan's MSAS. The government generated WADGPS corrections use predicted models of the ionospheric effect and broadcast data for the user to estimate the ionospheric effect. Because the StarFire system uses dual-frequency receivers, the ionospheric effects are measured and removed. This gives more accurate results and also requires far fewer reference stations than are required by the above government run WADGPS corrections.

2.1 StarFire ground reference networks

There are 23 NavCom reference sites located on four continents. The NavCom sites are configured somewhat different than the NASA global sites which are receiver only sites. Each of the NavCom reference sites has an identical set of equipment including:

a) two redundant NCT 2000D GPS reference receivers which send a full set of dual frequency observables for all satellites in view to both of the redundant processing hubs,

b) a fully packaged production StarFire user equipment unit which serves as an independent monitor receiver,

c) communications equipment (routers, ISDN modems),

d) a remotely controlled power switch and UPS module.

In the US network the communication lines used to link the reference sites with the processing hubs are frame relay private virtual circuits. Each frame relay circuit is backed up with an ISDN dial-up line, which can be activated in the event of a frame relay connection failure. The same implementation is used for the communication lines to and from the hubs and the uplink facilities, with the exception that the backup is automatic for these more critical links. The other regional networks are essentially

identical with the US network with the exception that Internet and satellite links are used rather than frame relay and ISDN.

A StarFire user set is located at each reference site and is used as a monitor unit to report the health of the system. These units receive the broadcast correction stream from the communications satellite, perform differential GPS navigation and report their positioning results back to the processing hubs, using the same communication lines as the reference receivers. In addition to the DGPS positioning results, the monitor data includes the received signal strength of the L-band communications satellite, packet error statistics, age of differential corrections, signal strengths for the received GPS satellites, PDOP and other operating parameters. This data, from all of the reference sites, is continuously monitored by an Alert Service processor which automatically generates E-mail and pager messages to on-call network service engineers in the event of a DGPS service failure.

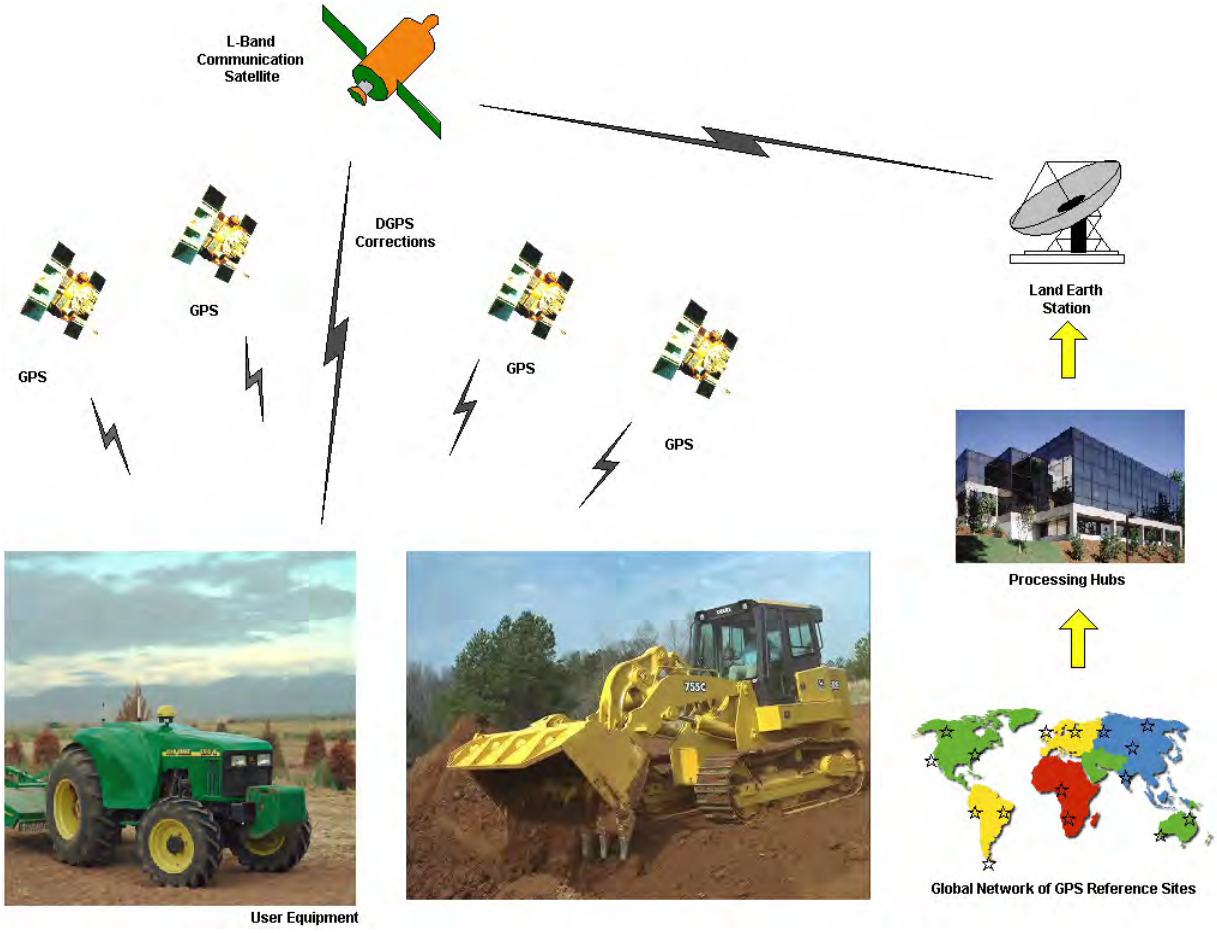


Figure 1: Overview of the StarFire WADGPS and GDGPS system

The global reference sites operated by JPL for NASA, identified by the red flags in Figure 2, along with the NavCom reference sites, identified by the green flags, are used in computing the global corrections. The Internet and satellite links are used to bring the dual-frequency measurements from these global references back to the processing hubs at Torrance and Moline. Since the hubs receive data from more than 60 reference sites, the corrections are robust; and the loss of data from one or more sites would not affect the accuracy significantly. Duplicate processing at both Torrance and Moline is used so that an automatic switch of the data uploaded to the communication satellites occurs if a failure is detected in one of the hubs.

2.2 Hub processing software

As indicated above, the StarFire system provides a set of corrections for each satellite that is good globally. The corrections are in the form of a three-dimensional correction to the satellite position and a correction to the satellite clock. The algorithms used at the processing hubs to compute the correction streams use the following inputs:

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 143-153

Journal of Global Positioning Systems
2004
© CPGPS

Figure 3: The StarFire Network of Reference Stations

a) dual frequency measurements, i.e. C/A code pseudoranges, L1 carrier phase, P2 code pseudoranges and L2 carrier phase, for all of the GPS satellites tracked at the reference sites, which are sent to the hubs at 1Hz in real time from all sites;

b) the broadcast ephemeris records from the reference receivers, which are delivered in real time;

c) a configuration file defining the precise location (±2cm) of each of the reference receiver antennas, as determined from network solutions based on the IGS worldwide control stations.

The dual-frequency measurements are used to form smoothed, refraction corrected pseudoranges which are free of ionosphere delay and, due to extended smoothing with the carrier phase, virtually free of multipath. The StarFire processing software used within the processing hubs at Torrance and Moline is that developed by JPL and licensed to NavCom. Orbit corrections for each satellite are generated and transmitted each minute and clock corrections for each satellite are generated and transmitted every two seconds. The correction data is sent via the land-lines to the uplink facility for broadcast from the geostationary communication satellites.

The StarFire corrections are efficient in that only one set of corrections per satellite is required. Competing systems require a multiple set of corrections, which requires computing a weighted correction at the user set. This is costly in terms of communication bandwidth and user computational requirements. In addition, improvements to the StarFire correction generation can be made without impacting the thousands of user sets deployed worldwide.

2.3 StarFire user equipment

Several different implementations of the StarFire user equipment are available. The most widely sold equipment is the compact rugged unit with all elements internal to the package. Three elements are contained in the package. First, it contains a tri-band antenna, which receives the L1 and L2 GPS frequencies together with the L-band Inmarsat frequencies. The gain pattern of the antenna yields a fairly constant gain even at low elevation angles. This allows the corrections broadcast over the Inmarsat channel to be received at high latitudes (low elevation angle) without requiring extra signal strength

from the geostationary Inmarsat satellite. Second, it contains an Inmarsat L-band receiver module to acquire, decode and send the corrections received to the GPS receiver. It is frequency agile and, under software control, can track any frequency in the Inmarsat receive band. Finally, the package contains a NCT 2000D GPS receiver, designed and built by NavCom.

An alternate "black box" package contains only the Inmarsat L-band receiver module and the NCT 2000D GPS receiver. The antenna is packaged separately for more flexibility, such as is required for aircraft installations.

Finally, a new StarFire integrated package is nearing completion. It consists of a tri-band antenna together with a new combined Inmarsat L-band module and GPS receiver module. The new package will also contain an optional inclinometer used in agricultural applications to map the GPS antenna position to the ground, i.e. it adjusts for the varying tilt of the agricultural implement as it passes over rough ground.

2.3.1 The NCT 200D dual-frequency GPS engine

The NCT 2000D is a compact, high-performance, dual frequency GPS engine aimed at OEM applications. The receiver is mounted inside the lower housing and interfaces to the digital board of the L-band receiver via an RS232 serial port. The StarFire corrections are input from the L-band receiver and 1, 5, or 10Hz PVT data is output via the external interfaces (RS-232 and CAN Bus). The NCT 2000D has ten full dual-frequency channels and two WAAS channels. It produces GPS measurements of the highest quality suitable for use in the most demanding applications, including millimeter level static surveys.

The NCT 2000D includes a patented multipath reduction technique built into the digital signal processing ASICs of the receiver. This greatly reduces the magnitude of multipath distortions on both the CA code and P2 code pseudorange measurements. When combined with extended dual-frequency code-carrier smoothing, multipath errors in the code pseudorange measurements are virtually eliminated. It also includes a patented technique used to achieve near optimal recovery of the P code from the anti-spoofing Y-code, resulting in more robust tracking of the P2/L2 signals. The compact size (4" x 3"x 1") of the NCT 2000D allows it to be readily integrated into the StarFire package. Finally, a high resolution 1pps output signal, synchronized to GPS time, is provided by the NCT 2000D. This same signal is used by the L-band communications receiver to calibrate its local oscillator, which aids in the acquisition of the StarFire correction signal. This technique has been patented by NavCom.

The measurement processing of the NCT 2000D software in the StarFire user equipment is somewhat unique. Dual-frequency code and carrier-phase measurements are used to form smoothed, refraction corrected, code pseudoranges. But, unlike competing systems, a greater reliance is put on the carrier-phase measurements and a floating ambiguity estimate is made of the whole-cycle ambiguities. All of the measurement data from all satellites is used to make this ambiguity estimate as accurate as possible. In addition, a constrained estimate of the tropospheric refraction is made, using the data from all the satellites. This constrained solution removes some of the unmodeled tropospheric refraction effects. The resulting PVT estimates are output at either 1, 5, or 10 Hz under software control.

2.4 StarFire positioning accuracy

Figure 3 shows typical position accuracy obtained by using the global StarFire corrections for 24 hours at an Australian site. The one-sigma accuracy per horizontal axis rarely exceeds 10 cm. As expected, the accuracy is not dependent upon geographical location. The accuracy obtained by using the StarFire global corrections within a StarFire receiver is unmatched by any other global system. Furthermore, very few regional networks can match these results.

2.5 Applications of the StarFire System

StarFire user equipment is now used in a large number of agricultural applications. These include yield mapping, field documentation, operator-assisted steering and automatic steering. Automatic steering is probably the most rapidly growing new application. However, it is also used in a wide variety of other applications. These include: (1) land survey and geographic information systems, (2) construction equipment guidance and control, (3) marine survey and resource exploration, (4) hydrographic mapping and dredging systems, (5) tracking of valuable cargo, (6) high precision airborne survey, and (7) specialized military applications on land, at sea, and in the air.

3 Recent improvements to the StarFire user software

A number of improvements to the user positioning software inside the StarFire unit have been made and are in the process of being verified and released. These include: (1) improved mode switching logic to minimize the loss of accuracy when it is necessary to change the operating mode of the receiver; (2) use of the change in

carrier-phase measurements to implement a high-rate, high-accuracy, and efficient "position change" process which is used to update the position coordinates between major epochs; (3) the implementation of a "receiver autonomous integrity monitoring" (RAIM) process which makes use of the high-accuracy change in the carrier-phase to detect measurement errors; (4) a simplified process for computing earth tides caused by the Sun and the Moon; and (5) the addition of a "real-time kinematic" RTK mode which can make use of the synergism between the global RTG correction stream and the RTK correction stream. Each of these five improvements are described in more detail below.

3.1 Improved Mode Switching Logic

There are a large number of conditions under which the navigation software must provide a position solution. These include the availability of several different correction streams made available from the various government sponsored augmentation systems. In addition, in some environments, signal blockage may cause the L2 measurements to be lost, even while the L1 measurements are still retained. If too few measurements are available or the vertical dilution of precision (VDOP) is too large, it may be desirable to operate in an "altitude hold" mode. A complicated hierarchy of modes results from the interplay of all these factors. In the past the transition between various modes retained no history (knowledge of the position variance-covariance matrix) of the prior mode. Failure to retain knowledge of the position accuracy during the mode transition allowed large position jumps during the transition, particularly when the transition was between dual-frequency and single-frequency modes while using the RTG correction stream.

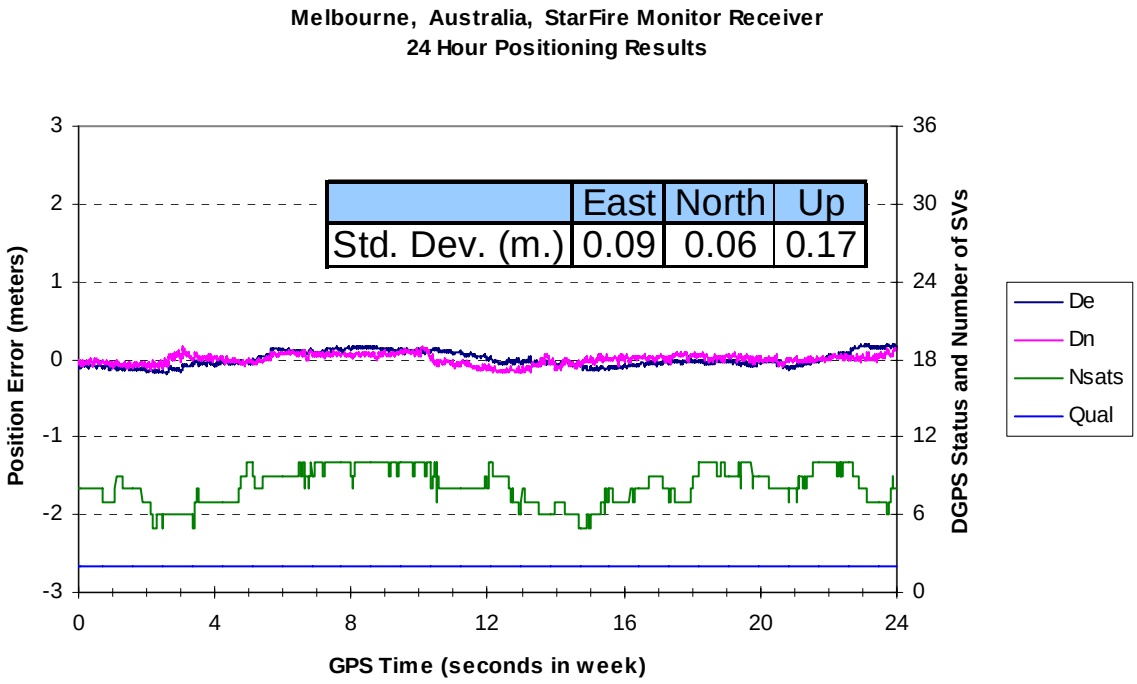


Figure 3 Sample RTG StarFire results in Australia

There is a significant "pull-in" time before the full ten-centimeter accuracy of the StarFire RTG global DGPS can be attained. This pull-in time comes from the need to solve for a floating ambiguity to the refraction-corrected carrier-phase measurements. Until that floating ambiguity can be determined with the requisite accuracy, the position solution will depend in large measure upon the less accurate refraction-corrected code measurements. It takes some significant geometry change (movement of the GPS satellites) before these ambiguities can be successfully resolved to an accuracy of a few centimeters. With no accuracy history, if the L2 signals were lost for even a few seconds, the mode would switch to a single-frequency mode and a significant loss of accuracy would result. Thus, when dual-frequency operation was resumed, the initial poor accuracy would require a new pull-in time before the full accuracy was again obtained.

In the new logic the change in the carrier-phase measurements, even if only available on the L1 signals, allows the propagation of the position accuracy to degrade only slowly. Thus, if the outage of the L2

measurements is only brief, when the dual-frequency measurements are resumed, the position accuracy will be only and no new pull-in time will be required before full accuracy is restored.

3.2 Propagation of Position Using the Change in the Carrier-Phase Measurements

The change in the carrier-phase measurements can be used as a measure of the change in position (and clock). This allows a highly-efficient, high-accuracy update of the position to be computed at a high rate [Hatch, et al., 2004]. In the new StarFire navigation software, this technique is used to compute the position at high rates, e.g. 5 to 10 Hz. At the slower 1 Hz rate the normal least-squares computation is performed. As part of that low-rate computation, the gain matrix used during the high-rate is computed once and applied successively at the high-rate. An explanation of the fundamental process is provided below. The above reference gives further details on how to handle the loss of measurements from one or more satellites during the interval between the low-rate epochs. The addition of a measurement from a new satellite can simply be delayed until the next low-rate epoch.

The computation of the position during the low-rate epoch (for a simplified least-squares solution) is derived below. The fundamental measurement equation for a specific satellite is given by:

$$z = \mathbf{h}\mathbf{x} + \eta \qquad (1)$$

Where: **x** is the state correction vector (change in position and clock) value to be computed; **h** is the measurement sensitivity vector, which characterizes the effect of any errors in the state vector upon the measurement; is the measurement noise; and *z* is the measurement innovations, i.e. the difference between the measurement and the expected value of the measurement given the current estimate of the state vector (position and time). The value of *z* also includes the StarFire corrections to the measurement.

Equation (1), when expanded into matrix form to represent the set of equations from all tracked satellites, becomes:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n} \qquad (2)$$

The least-squares solution to Equation (2), which minimizes the effect of the noise vector, **n**, is given by:

$$\mathbf{x} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{z} \qquad (3)$$

where the superscript, *T*, represents the transpose, and the superscript, *-1*, represents the inverse.

The matrix operations can be performed to give simpler forms of Equation (3):

$$\mathbf{x} = \mathbf{A}\mathbf{H}^T\mathbf{z} \qquad (4)$$

or

$$\mathbf{x} = \mathbf{B}\mathbf{z} \qquad (5)$$

where: $\mathbf{A} = (\mathbf{H}^T\mathbf{H})^{-1}$ and $\mathbf{B} = \mathbf{A}\mathbf{H}^T$.

The matrix **B** has four rows, corresponding to the three position coordinates and the clock. It has as many columns as there are satellite measurements available. It is stored for subsequent use in the high-rate propagation computation.

The high-rate computation uses the change in the carrier-phase measurements to compute the change in position (and clock) over the high-rate epoch intervals. Having stored the **B** matrix used in Equation (5) above, the change in carrier phase over the high-rate epoch can be used to propagate the position forward in time with high accuracy and with minimal computations. The high accuracy is a result of the low noise in the carrier-phase measurements. The first computation required is to compute the innovations (pre-fix residuals), **z** for use in Equation (5). The change in the measured carrier phase (delta phase) for each satellite is a major component of the innovations. Generally, only one correction to the delta-phase measurements is required to make the innovations accurate enough to maintain centimeter accuracy over major epoch intervals of one second. Specifically, one must subtract from the delta phase the change in the radial distance to the satellite over the high-rate epoch interval. The specific equation to compute the innovations for each satellite is:

$$z_i = (\phi_i - \rho_i) - (\phi_{i-1} - \rho_{i-1}) \qquad (6)$$

where represents the phase measurement scaled by the wavelength and represents the range to the satellite. The subscript *i* indicates the current epoch and *i-1* the prior epoch.

The difference between the current phase measurement and the range to the satellite can be stored for use in the subsequent epoch. The satellite position computation should be optimized for this high-rate computation. There are a number of methods described in the literature to optimize this computation.

Most of the normal corrections to the innovations that are required at the low-rate major epochs are not required for the innovations at the high rate. This is because they generally change by less than the centimeter level over the one-second interval between the major epochs. This generally includes: 1) the satellite clock error; (2) the ionospheric and tropospheric refraction; and 3) the corrections from StarFire (or any other correction

source). Of these factors, the largest may well be the satellite clock errors. They can contribute an error which can approach one centimeter over a one-second major epoch interval. But it is quite easy to incorporate the satellite clock correction if desired. It is simply the clock frequency offset times the high-rate epoch interval.

Once the innovations are computed using Equation (6), the values can be used in Equation (5), together with the stored value of the **B** matrix, to give the high-rate correction to the position and clock. This propagates the position with the high accuracy of the carrier-phase measurements, and the computational load is minimized. Once the high-rate propagation of position has been initialized, the low-rate computation is implemented as a simple correction to that propagated position.

3.3 Receiver Autonomous Integrity Monitoring (RAIM)

As indicated in the prior section, the position determined in the StarFire solution results from the propagation of the position using the change in the carrier-phase measurements. The low rate computation is simply a correction to that propagated position. Because the measurements of the carrier-phase are so precise, they open the possibility of a very effective means of detecting and even correcting measurement errors. The technique summarized below has been described in detail in Hatch et al. [2003], along with some sample computations.

To detect an erroneous measurement one needs to analyze the post-fix residuals. To the extent the solution is accurate, the residuals become an estimate (the negative) of the noise in Equation (2). The post-fix residuals can be computed directly from the innovations or pre-fix residuals. Given the correction to the state vector, x, the post-fix residuals, r, are obtained by rearranging Equation (2):

$$\mathbf{r} = \mathbf{z} - \mathbf{H}\mathbf{x} \tag{7}$$

But using equation (5) to replace the state update, x, in equation (7) gives:

$$\mathbf{r} = \mathbf{z} - \mathbf{H}\mathbf{B}\mathbf{z} \tag{8}$$

This in turn can be simplified to:

$$\mathbf{r} = \mathbf{S}\mathbf{z} \tag{9}$$

Where: S = (I – HB).

The matrix, S, maps the pre-fix residuals into the post-fix residuals. S is square and the number of rows and columns are equal to the number of satellite measurements.

With the high accuracy of the carrier-phase measurements, i.e. the low noise in Equation (2) and the small residuals in Equation (9), a measurement error can generally be detected by computing the value of the RMS residuals and comparing them to a threshold value. The RMS residuals are given by:

$$\delta = \sqrt{\frac{\mathbf{r}^T\mathbf{r}}{n}} \tag{10}$$

Where n is the total number of measurements.

Assuming a single fault, which is by far the most probable situation, the specific measurement at fault can be determined by forming the correlation between the residuals and each column (or row) of the S-matrix.

$$\rho^j = \frac{\sum_i r_i s_i^j}{|r||s^j|} \tag{11}$$

In this equation j is the correlation coefficient associated with the satellite j measurement; ri is the ith element of the residual vector and s_i^j is the ith element of the jth column of the S-matrix; $|r|$ and $|s^j|$ are the lengths of the two column vectors, i.e. the square root of the sum of the squares of the elements.

Note that the two lengths can be obtained from simple alternate equations:

$$|r| = \delta\sqrt{n} \tag{12}$$

$$|s^j| = \sqrt{s_j^j} \tag{13}$$

Given that a fault has been detected by the size of the rms residuals, the faulty satellite will be identified by the correlation coefficient from Equation (11) with absolute value closest to one. This test will fail to identify the faulty measurement if there are only five measurements in the total set. When only five measurements are available, all of the correlation coefficients from equation (13) will be very close to one or minus one. Only the lengths of the columns (or rows) of the S-matrix will vary when measurements to only five satellites are available.

When there are more than five measurements available, the faulty satellite can be identified by the largest correlation value; and the size of the error (i.e. the number of cycle slips) can be computed from the ratio of the residual vector length (Equation 12) to the length of the associated column of the S-matrix (Equation 13).

$$e^j = \frac{\delta\sqrt{n}}{\sqrt{s_j^j}} \tag{14}$$

The sign of the error in Equation (14) is obtained by using the associated sign of the correlation value. This error in the measurement from the jth satellite can be tested to see if it is very close to an exact integer. If it is, a correction to the measurement can be made. The revised residuals can be computed from the following equation:

$$r' = r + s^j e^j \tag{15}$$

where sj is the jth column of the S-Matrix.

The revised rms residuals computed from this new residual vector should now pass the threshold test and indicate that the fault has been removed.

A discussion of unusual situations, where the error detection and removal described above might fail or require special logic, is found in the more detailed paper [Hatch et al. 2003].

3.4 Removing the effects of solar and lunar earth tides

While the StarFire orbit and clock computations licensed from JPL included adjustments for the earth tides caused by the sun and the moon, no corresponding adjustment was included in the initial StarFire navigation software developed by NavCom. For almost all farming applications the small cyclic variation in height over a 12 hour period was of little consequence or interest. However, in other applications precise height location is desired. That small but significant height variation caused by the earth tides was confirmed by computing daily values of the RMS height variation over a complete 24 hour period. It was found that this daily RMS height variation was a strong function of the phase of the moon.

Figure 4 shows that when the sun and the moon are not aligned the RMS height variation is generally between 10 and 15 centimeters. When the sun and moon are aligned, either on the same side of the earth (new moon) or on opposite sides of the earth (full moon), the RMS variation is about twice as big. Because the plane of the moon's orbit is not aligned with the ecliptic (earth's orbital plane), the tides can be larger at one of the two alignments than at the other.

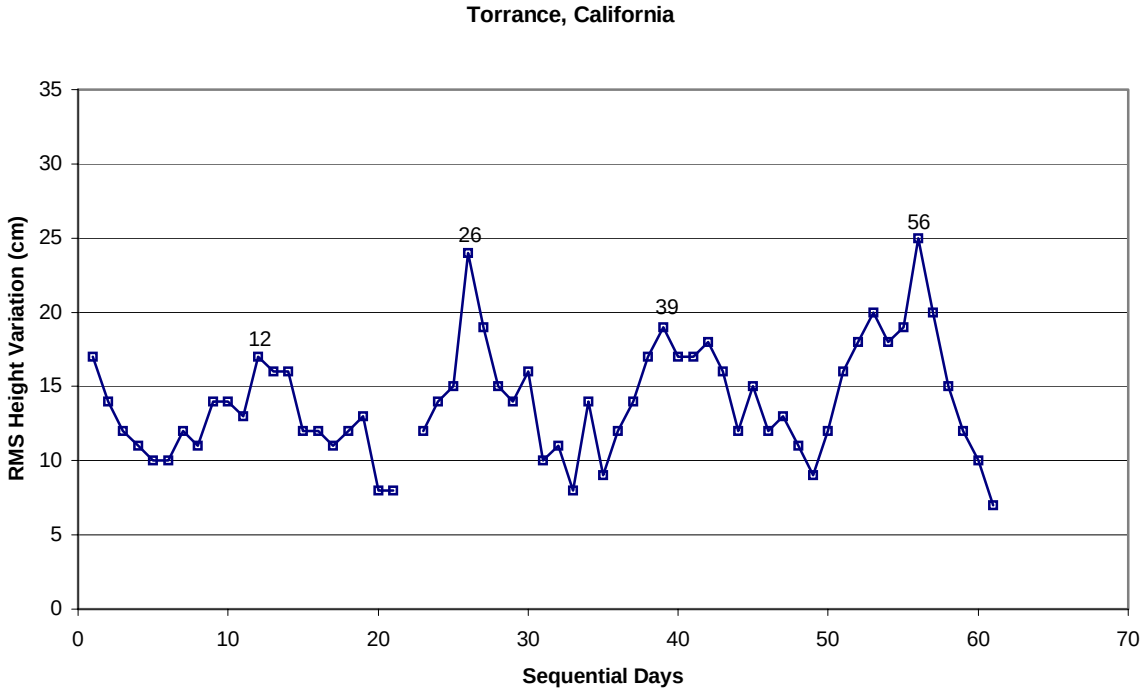


Figure 4 RMS height variation over 60 days at Torrance, California

Having determined that the solid earth tides were significant, a computationally efficient method of removing them within the navigation software was undertaken. The algorithms developed by Sinko [1995] were selected, and four relatively minor modifications were implemented to improve their computational efficiency. The concept of the modifications was relatively simple. First, the equations were updated to use the initial epoch of 2000 rather than the 1900 epoch. Second, to cut down on the computational load,

the equations were separated into those requiring frequent update (every few minutes) from those requiring only daily updates. Third, at the general cost of less than a centimeter in accuracy, it was found that a spherical earth model could be used rather than a flattened ellipsoid. This significantly simplified some of the higher rate computations. Finally, the equations were modified to avoid where possible the computation of the specific angular values. Instead, the sine and cosine of angular values were retained. This allowed the subsequent computation of the sine and cosine of sums (or differences) of angles to use the product rule for sines and cosines rather than computing the sine and cosine for the sum (or difference) directly. Using this process a large number of trigonometric functions and their inverses were avoided.

Testing the simplified Sinko model within the software with external off-line computations of the earth tides verified that the implementation was generally within one centimeter of the more precise off-line computations. The latest release of the StarFire navigation software now incorporates this earth-tide computation.

3.5 Extending the RTK performance

As one of many options, the StarFire navigation software has for some time included a high-accuracy RTK capability. This past year a significant improvement was made to the RTK performance. Specifically, the short-term accuracy performance of the StarFire GDGPS can be used to augment the RTK performance when the RTK corrections are lost for any reason. This capability is of particular benefit in rolling hills or other conditions in which the line of sight between RTK reference receiver and the RTK navigation receiver is often obstructed.

The StarFire performance, as illustrated in Figure 3, looks pretty noisy. However, if one were to amplify the horizontal scale by 144 to look at any fifteen-minute segment, the performance would appear to be virtually noise free but with a small very slowly changing bias. The new “RTK Extend” feature embedded within the navigation software takes advantage of this “low noise, but biased” characteristic of the StarFire GDGPS performance. Specifically, the offset between the RTK performance and the StarFire performance is continually computed when both the RTK and StarFire navigation are enabled. Applying this offset to the StarFire navigation when the RTK correction link is lost allows the navigation to proceed with virtually no loss in accuracy.

The process is a little more complicated than indicated above only because of the normal pull-in time required of the StarFire navigation. To avoid the pull-in time on the navigation receiver, the RTK solution is used to “quick start” the StarFire navigation. The quick start capability depends upon initializing the position to an accurate absolute position with an associated small valued variance-covariance matrix. Unfortunately, the RTK navigation solution is not an absolute position unless the RTK base position is an accurate absolute position. This problem can be resolved by using a StarFire receiver at the base station. Requiring that the base station operate 15 to 30 minutes before the RTK extend feature is enabled is generally not an onerous requirement. This allows the base station sufficient pull-in time to determine the absolute offset of the initialized RTK base station. This offset is transmitted along with the RTK correction stream to allow the StarFire navigation to be initialized with the absolute position, i.e. with the RTK solution offset by the measured absolute error at the base is used to initialise the StarFire navigation within the user set.

4 Conclusions

The StarFire Global DGPS was reviewed and its performance illustrated. Recent improvements to the navigation software were described. These included five new features. The first was to avoid losing the historical information of position and its associated accuracy when switching between modes in the receiver. This simple improvement provides a much more accurate solution under constrained visibility situations where switching between single-frequency and dual-frequency operation is quite common.

The second improvement was the implementation of an effective high-rate propagation of position using the change in the carrier-phase measurements. The accuracy of the carrier-phase measurements, combined with the ability to use the same gain matrix generated at the low-rate epochs, allows the technique to be both very efficient and very accurate.

The accuracy of the carrier-phase measurements used in the prior improvement allows those same measurements to be used to implement a very simple and effective RAIM and measurement correction scheme. With five or more satellites, small errors or carrier cycle slips can be detected. With six or more satellites, carrier cycle slips of a single cycle can normally be isolated to the specific satellite and corrected.

The implementation of an earth tides algorithm has improved the height accuracy significantly and improved the horizontal position marginally. The Sinko [1995] algorithm was simplified and implemented within the navigation software.

Comparison with off-line computations of the earth tides were generally at the centimeter level.

Finally, the synergistic characteristics of the StarFire Global DGPS solution and the RTK solution were exploited to give a "quick start" to the StarFire solution and an accurate "RTK extend" to the RTK solution. This capability allows the StarFire RTK solution to continue without interruption for up to 15 minutes when the RTK base station corrections are lost. This capability is particularly useful for where the line-of-site communications link is obstructed by rolling hills or other obstacles.

Acknowledgements:

The developments described herein were largely funded by and performed for the Ag Management Solutions group of John Deere. The initial earth-tides development was stimulated by feedback from C&C Technologies, Inc.

References

Hatch RR, Sharpe RT, and Yang Y (2003): ***A simple RAIM and fault isolation scheme,*** Proceedings of the 16th Int. Tech. Mtg. of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 Sept., 801-808.

Hatch RR, Sharpe RT, and Yang Y (2004): ***An innovative algorithm for carrier-phase navigation,*** Proceedings of the 17th Int. Tech. Mtg. of the Satellite Division of the U.S. Inst. of Navigation, Long Beach, California, 21-24 Sept.

Sinko J (1995): ***A compact earth tides algorithm for WADGPS,*** Proceedings of 8th Int. Tech. Mtg. of the Satellite Division of the U.S. Inst. of Navigation, Palm Springs, California, 12-15 Sept., 35-44.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 154-166

Journal of Global Positioning Systems
2004
© CPGPS

Limitations of Pseudolite Systems Using Off-The-Shelf GPS Receivers

Mustafa Ozgur Kanli

School of Surveying and Spatial Information Systems, The University of New South Wales, Sydney, Australia
e-mail: m.kanli@student.unsw.edu.au Tel: + 61-2-6229-1780; Fax: + 61-2-6229-1778

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Pseudolites (PLs) are ground-based transmitters that transmit GPS-like signals. They have been used to test GPS system elements and to enhance GPS in certain applications by providing better accuracy, integrity and availability through the use of PL signals in addition to the GPS signals. PLs are also a promising technology for providing positioning in indoor, high multipath environments where GPS signals are generally unavailable or severely attenuated and of questionable quality. In experiments to date, researchers have almost exclusively used PLs that transmit C/A code on L1/L2 in order to use existing off-the-shelf GPS receivers. This is because no hardware modifications to the GPS receiver are necessary and only minor changes to the receiver firmware are needed to track a PL's signal. However, there are some fundamental issues that limit the effectiveness of a PL system using C/A code on L1/L2. These include the legality of transmitting on L1/L2, cross-correlation between PL and GPS signals, saturation of GPS receiver front-ends, and the limited multipath mitigation offered by C/A codes. When combined with other problems inherent to all PL systems such as near-far, multipath, and synchronization, the issues in using L1/L2 C/A code PL systems further complicates the design and deployment of such systems and places limits on its operational effectiveness. This paper presents the issues which limit PL systems that use GPS hardware and explores the impact of these issues on some common PL applications.

Key words: pseudolite, cross-correlation, jamming.

1 Introduction

The concept of using ground-based GPS signal transmitters originated from the time when GPS was first being developed. These transmitters, which came to be known as Pseudolites (PLs), enabled developers to test GPS concepts before the first satellites were even launched. Since then, PLs have been found to improve geometry for accurate positioning, especially in the vertical component. They provide additional ranging signals to augment existing GPS signals, increasing availability. In some environments, such as indoors, GPS signals are heavily attenuated and are of questionable quality. In these cases, PLs have been used to provide alternative ranging signals. A comprehensive summary of PL technology, including PL designs and applications, is provided by Wang (2002). There, the reader will also find many helpful references.

Almost all purpose-built PLs transmit on the L1 frequency. Some also transmit on L2, or even both. Those that are simply comprised of a signal generator may transmit on any range of frequencies. Most L1/L2 PLs typically use spreading codes from the same family as the C/A codes used in GPS satellites. This practice enables the PL to closely resemble a GPS satellite, and this is important to allow existing off-the-shelf (OTS) GPS receivers to track a PL's signal. Only slight changes to a GPS receiver's firmware is required to achieve this. First is a modification to the receiver's search database, allowing it to look for local PLs as well as GPS satellites. If the PL modulates data onto its signal, software to correctly interpret this data also has to be written. Some other thresholds and assumptions may also need to be adjusted. Importantly, no hardware changes to a receiver's correlators or front-end circuitry are needed. This practicality simplifies experimentation and leaves more time to be spent on research.

Unfortunately, this convenience comes with a price. The L1 and L2 frequency bands are protected and reserved for radionavigation; legal issues hinder unlicensed transmissions in these bands. PL signals are often considerably stronger than GPS signals. This can result in jamming, which will deny GPS to other nearby receivers that don't participate in PL navigation. Typical OTS GPS receivers may also be overwhelmed by the strength of PL

signals; they expect weak GPS signals. Saturation of the front-end can result, which will decrease accuracy, reliability and introduce other complications. For indoor PL navigation, the C/A code chipping rate used in GPS does not sufficiently limit range errors due to multipath. These factors contribute to other complications that can affect the operational effectiveness of PLs.

This paper will begin with a discussion of issues inherent to all PL systems; these are near-far, multipath and synchronization. Then, issues specific to L1 C/A code PLs will be presented. A brief look at the impacts of these issues on some common applications will follow. The paper will then conclude with a section discussing the recent developments in PL technology. Other competing positioning systems will also be briefly discussed, as well as some possible future developments.

2 Issues in pseudolite systems

2.1 General Issues

This section will describe the general issues confronted by PL systems.

2.1.1 Near-far

GPS satellites are located in near-circular orbits around the Earth with a radius of about 26,560 km. This puts the typical distance between a user receiver and a visible satellite overheard at about 20,000 km. Changes in this distance due to typical user and satellite motion are insignificant compared to the overall separation distance. Also, the radiation patterns of GPS satellite antennas are shaped to spread the RF signal almost uniformly over the surface of the Earth. These factors ensure that a user receiver can expect to see GPS signal strengths around -160 dBW from all visible satellites. In contrast, PLs are located much closer to user receivers so that user movements can cause significant variations in the distance between PLs and user receivers. Because of this, user receivers will see large variations in PL signal strengths. For example, for an isotropic PL antenna the signal strength 30 m away will be 24 dB stronger than the signal strength 500 m away. This is known as the near-far effect.

The repercussions of the near-far effect on GPS signals are primarily jamming and limited operating distance. For instance in a standalone PL network each PL will, within an area of close proximity, jam the signals from other distant PLs. Receivers in such a network will observe varying PL signal strengths at different locations. Off-the-shelf GPS receivers expect signals from all sources to be of approximately equal strength. This will only occur at points equi-distant to all PLs. PL signals can be more than 30 dB stronger than GPS signals; which is greater than the worst case code separation (21.6 dB) between C/A codes. In an integrated GPS/PL system, PLs can thus potentially jam GPS signals at close proximity. The output power of the PL can be calibrated to minimize jamming, but this limits the operating range of the PL.

The most popular solution used to mitigate the near-far problem, as presented by Klein and Parkinson (1984), is to pulse the PL signals at fixed cycle rates (typically 10%). This provides a 10 dB improvement in the signal-to-interference level. Transmitting the signals at a frequency offset from GPS L1 (while remaining within the same frequency band as GPS) and using longer sequence spreading codes have also been suggested as potential solutions. However, not all off-the-shelf GPS receivers can accommodate the use of different spreading codes and/or the frequency offsets. A different approach, proposed by Madhani et al (2001), uses successive interference cancellation, where the receiver accurately regenerates the interfering PL signal before subtracting it from the total received signal to yield a partially cleaned version of the received signal. This approach relies on the ability of the receiver to accurately track both the interfering signal and the desired weaker signal. In order to do this, the receiver requires an ADC with sufficient resolution and an RF front-end with sufficient dynamic range to accommodate both sets of signals. Most off-the-shelf GPS receivers, however, don't have this capability. Shaping the radiation pattern of PL transmit antennas is another approach used in the mitigation of the near-far problem (Söderholm et al, 2001). This approach aims to provide a near-uniform spread of PL signal strength covering only a specific desired area, achieved by appropriate orientation of the transmit antennas of the PLs. While effective within the desired area, near-far conditions are still apparent on the fringes of, and outside, this operational area. Custom antenna designs are required to provide sufficient coverage of each environment. Thus, this approach will not be suitable for dynamic environments such as in shipping yards, where moving freight containers cause considerable changes to the landscape. Furthermore, the deployment of PLs is restricted by the very nature of this near-far mitigation technique.

2.1.2 Multipath

A pseudo-range measurement involves determining the propagation time of a ranging signal along a direct line-of-sight path from its source antenna to a receiver's antenna. Due to reflective objects in the environment, the signals at a receiver's antenna can be composed of both direct ranging signals and any number of reflected (multipath) ranging signals. These multipath signals are

delayed relative to the direct signal and have differing amplitude, phase, and polarization, characterized by the reflecting surface and the number of reflections. Pseudo-range errors of tens of metres often result from the presence of multipath signals, which in most cases is the largest error source. Multipath signals give rise to larger pseudo-range measurement errors than carrier phase measurement errors, which are of the order centimetres. Multipath signals may also destructively interfere with direct signals resulting in multipath fades. A good introduction to multipath effects is given by Braasch (1996).

Unlike near-far, the problem of multipath is an issue for GPS users as well as for designers of PL systems. The severity of multipath varies with the environment in which the positioning signals are applied. As an example, when PLs are used for augmenting GPS in aircraft approach and landing, a typical source of multipath is long delay ground bounce. In this case, multipath signals are typically delayed by more than a ranging code chip period. Theoretical texts may show that the multipath signal will have no effect. However, C/A code correlation side-lobes provide a way for relatively strong long delay multipath signals to induce errors of several metres. It should be noted that a GPS receiver may choose to ignore satellites at low elevations to exclude the possibility of ground bounce signals introducing pseudo-range errors. For PL systems however, all signal sources are almost always at low elevations significantly increasing the amount of multipath.

A different case is a PL system for indoor positioning, where there is a wide range of reflectors in close proximity of the receiving antenna. Examples include reinforced concrete structures, home/office furnishings and antenna mountings (man or machine). These reflectors generate a prevalence of short delay multipath, where ranging signals are delayed by less than a ranging code chip period. Short delay multipath typically has a greater impact on pseudo-ranges than long delay multipath, and it is also harder to mitigate its effects. For example, a single multipath signal with ¼ the amplitude of the direct signal can induce a 40 m range error (Misra and Enge, 2001).

Multipath mitigation can be performed on several levels. First, multipath signals can be attenuated selectively at the antenna. A choke-ring antenna is one device that can achieve this (Kunysz, 2003). It consists of concentric rings of grounded metal around the antenna element. These attenuate signals that arrive from low elevations relative to the axis of the metal rings. Choke-ring antennas are effective in attenuating ground bounce multipath signals. Other directional antenna designs have been used to attenuate signals from non-line-of-sight paths (Stolk and Brown, 2003). Of course, such techniques are only useful if the multipath signals arrive from the directions they are anticipated from.

Modelling multipath errors in software in order to minimize its effects is another type of mitigation. One approach includes the simulation of different types of reflectors at a variety of distances from receiver antennas. Receiver tracking algorithms can then be modelled to determine the response of the tracking loop to the reflected signals (Weiss et al, 2003). Modelling multipath or multipath estimation may be effective for controlled, static environments but it is unsuitable for dynamic, indoor and high multipath environments. In such cases, receivers moving through a reflective environment will encounter multipath effects that deviate from the model.

Multipath mitigation is also performed at the level of the receiver's correlators. Ordinary receivers perform correlation with a pair of early and late tracking arms separated by a chip period. These tracking arms straddle the correlation peak by maintaining the level difference between the pair at zero. In the presence of a multipath signal, however, keeping the level difference at zero may not imply that the tracking arms straddle the true correlation peak. Decreasing the separation between the tracking arms is one way to achieve a better result in the presence of medium delay multipath signals. The Narrow Correlator™ operates in this fashion (Van Dierendonck et al, 1992). Other correlation techniques such as the Double Delta correlator, the Early/Late Slope technique and the Early1/Early2 tracker have also been devised. These can dramatically decrease the maximum pseudo-range error due to multipath. Irsigler and Eissfeller (2003) present a good comparison of the multipath mitigation of different correlation techniques mentioned here. It is important to note that techniques like the Narrow Correlator™ require a receiver with a front-end having a wide bandwidth relative to the chipping rate of the ranging code used. For C/A code, typical narrow correlator receivers have pre-correlator bandwidths of 16 MHz. The consequence of having such a wide bandwidth is that signals from adjacent radio bands are more likely to cause interference. Also, not all OTS GPS receivers employ such advanced multipath mitigation techniques.

2.1.3 Pseudolite synchronization

Fundamental to the operation of GPS is satellite clock predictability. Without precisely synchronized satellite clocks, precise time transfer and accurate stand-alone navigation would be impossible. The cesium and rubidium atomic clocks onboard GPS satellites have stabilities of the order of 1 part in 1013. To keep the clocks synchronized, the GPS Operational Control Segment (OCS) uploads satellite clock corrections to the satellites at least once a day. These corrections, which are

part of the navigation message, are used by receivers to correct for the satellite clock drifts.

While atomic clocks have been used in certain PL applications where a ranging signal of high quality is desired (Soon et al, 2003), most PLs typically use inexpensive temperature compensated crystal oscillators (TCXOs) to provide their reference frequency. Typical TCXOs have a stability of around 1 part in 106, more than six orders of magnitude worse than atomic standards. Because of such insufficient stability in their reference frequencies, PLs that operate asynchronously cannot provide accurate stand-alone navigation. In such cases double-differencing with a reference receiver is commonly used to eliminate PL and receiver clock biases. For a real-time solution this requires a wireless data link between the user and reference receivers. This adds operational constraints and affects performance, depending on data link range, integrity and latency. Furthermore the reference receiver must have visibility of all PLs for accurate measurements. This may mean having several reference receivers, for instance, in an indoor PL network that spans several rooms. This of course will add cost and complexity to the system. Also, timing information is eliminated in the double-differencing procedure. So, asynchronous PL operation will be unsuitable for those applications in which precise time transfer is important.

Asynchronous operation of PLs can also complicate the pulsing technique often used for the mitigation of the near-far effect. Consider the case of two PLs that are both pulsed with a duty cycle of 10% over the C/A code duration. If the PL pulsing scheme is asynchronous, the transmission cycles of both PLs may overlap for durations that vary according to the drift of both PL clocks. This will impact on how well a receiver tracks both PL signals. Since a receiver can only correlate for 10% of either PL's transmission, overlapping of the two signals can cause a significant increase in interference, especially in a saturated receiver. In the worst case both PLs will be mutually jamming each other.

An example of asynchronous PLs that are used to augment GPS for an aircraft precision approach and landing is given by Soon et al (2003). Kee et al (2000) present a successful account of indoor positioning using only asynchronous PLs. Cobb (1997) discusses a synchronous pseudolite, known as a synchrolite. A synchrolite acts as an electronic mirror to reflect GPS satellite signals from a known point on the ground. A synchrolite consists of a co-located GPS receiver and a PL transmitter. The receiver is able to determine the precise code phase and carrier frequency of the satellite signal that it is tracking. This information is then used to generate the PL transmissions that are synchronous to the GPS signals. Synchrolite signal measurements are typically differenced from satellite signal measurements to eliminate spatially correlated errors.

Söderholm and Jokitalo (2002) present a synchronous PL network for indoor positioning. The PLs in this network are synchronized by a Master Control Station (MCS) that tracks all PL signals using a reference receiver. Clock corrections are communicated to all the PLs via hard wired links. This centralized approach to synchronization that mirrors the GPS OCS has several disadvantages. First, the reference receiver must have visibility of all PLs in the network. In the OCS this is achieved by distributing monitoring stations around the world. Similarly, the MCS needs to collect measurements from as many reference receivers as are required to track all PLs in the network. Second, the synchronization control loop of the MCS must operate much more frequently than its counterpart, the OCS. This is because PL clocks are much less stable than the atomic standards used in GPS satellites. Obviously, adding more PLs to the network would also mean more computations for the MCS, and this reduces the scalability of the system. Although hard wired data links were presented, these could be substituted by wireless links to eliminate the process of laying communication cable during installation. However, running such a high frequency synchronization control loop over communication links to remote PLs will undoubtedly add noise due to communication latency.

A decentralized approach to synchronization has been demonstrated by Kee et al (2003) and Barnes et al (2003). In a decentralized system, each PL has a co-located receiver that tracks its own signal as well as a reference signal. The reference signal may be the transmission of a PL selected to be "master", or that of a GPS satellite. PL clock corrections can then be determined by taking single-difference measurements of pseudo-range and Integrated Carrier Phase (ICP) between the reference signal and the local PL signal. Because the receiver is co-located, control loop latencies are minimized and the PL network is highly scalable; any new PLs can be added to the network without being limited by data links or being concerned about the visibility of PLs to a reference receiver. Having a co-located receiver, however, also adds to the cost and complexity of PLs.

In a decentralized system, various synchronization topologies can be formed to suit the application. For instance PLs are not required to track the "master" PL to be synchronous to the network. A PL can be selected to synchronize to another PL that is already synchronous to the master PL. This propagation of time (i.e. PL synchronization) has been demonstrated by Barnes et al (2003). Clearly, propagation of PL synchronization allows for easy expansion of PL networks. Consider an indoor PL network for a warehouse that requires expansion to cover outdoor adjacent delivery sites. Additional PLs can be placed outside that only need to

track one indoor PL to achieve synchronization with the existing network. Such a scheme is comparable to ad-hoc wireless communication networks in which routing information propagates to reflect changes in topology.

It must be said that propagation of synchronized PLs requires further study. The process of synchronizing a 'slave' PL to a 'master' signal will undoubtedly produce noise in the slave PL's signal, the amount of which will vary according to factors such as the bandwidth of the synchronization loop. If another 'slave' PL is to then synchronize to the signal of the first 'slave' PL, the amount of noise in the transmission of the new PL will increase. This can, of course, affect the accuracy of the code and carrier measurements made from this signal.

2.1.4 Location errors and modelling

Near-far, multipath and synchronization are considered to be the major issues that have to be addressed in PL systems. However, there are other issues that arise, some depending on the particular application of the PL system. One issue concerns the accuracy of the specified location at which a PL transmit antenna is mounted. A study of the impact of PL location errors on positioning was presented by Lee and Wang (2002). The observed effects vary with the geometry between the PLs and the user receivers. For a static receiver, the PL location error will impose a bias in the PL measurements. For a non-static receiver, single-differenced measurement errors can be up to twice as large as the PL location error, depending on the geometry. Thus mounting the PL antenna accurately on a stable platform is necessary for high precision applications.

Another issue that needs to be addressed for wide area PL systems is tropospheric modelling. Signal propagation delays through the troposphere vary with air pressure, temperature and relative humidity. These delays can be corrected for in the receiver by modelling the delay. The tropospheric models used in GPS assume signals will originate from 20,000 km in space, and the modelled delays are highly dependant on satellite elevation. Because PL networks aren't as large, the same tropospheric models as used in GPS won't be as accurate. Tropospheric delays can amount to 32 cm per km, so for wide-area PL networks spanning several kilometres these delays must be factored to obtain accurate measurements. An overview of error modelling in PL applications is presented by Dai et al (2001).

2.2 Issues specific to L1 C/A code pseudolites

2.2.1 Legal

The rational, equitable and efficient use of the radiofrequency spectrum is coordinated by the International Telecommunications Union (ITU). Member states belonging to the ITU co-operate on agreements concerning issues such as frequency allocation and the registration of radio frequency assignments in order to avoid harmful interference between radio stations. These agreements are then implemented at a national level in the form of laws and regulations that are applicable only to the governed territory. As such, there may be slight differences in the way frequency allocation is regulated among the ITU's member states.

The frequency band to which L1 belongs, 1559 – 1610 MHz, is reserved by the ITU for aeronautical radionavigation and radionavigation-satellite (space-earth and space-space) services. The aeronautical radionavigation service refers to electronic aids such as marker beacons and some aeronautical mobile communications that form an integral part of aeronautical radionavigation systems. The radionavigation-satellite service refers to satellite based radio-determination systems such as GPS and GLONASS.

The Federal Communication Commission (FCC) is the government body in the United States of America that regulates the use of the spectrum according to ITU agreements. The FCC limits the use of the band 1559-1626.5 MHz to airborne electronic aids for air navigation and any associated land stations (47 CFR 87.475). Under this regulation an L1 PL used in augmenting GPS in an application such as aircraft approach and landing may be considered as an aeronautical radionavigation aid. The FCC states further that transmissions by aeronautical radionavigation land stations must be limited to aeronautical navigation (47 CFR 87.471). So, while aeronautical applications of L1 PLs may comply with existing FCC regulations, it seems that other applications of L1 PLs are not permitted. However, the case may be that L1 PLs deployed to provide air navigation may also be useful for non-aeronautical purposes as well.

The FCC assigns the band 1215-1300 MHz primarily for the military services and only permits limited secondary use by other government agencies in support of experimentation and research programs (47 CFR 2.106). This regulation rules out any prospect of using L2 PLs, and hence the benefits from using dual-frequency L1/L2 PLs, for any non-military application or for any extended period of time. Dual-frequency PLs would allow on-the-fly ambiguity resolution and also provide redundancy to protect against multipath fades.

For researchers, however, FCC regulations permit experimental radio service licences. Such a licence may allow researchers to use any government or non-government frequency, as designated by the FCC, provided that the need for the frequency requested is fully justified by the applicant (47 CFR 5). Licences are valid for either 2 or 5 years with renewals awarded only upon an adequate showing of need. It must be said that the FCC will only permit an Experimental Radio Service licence on the condition that harmful interference will not be caused to any station operating in the frequency bands allocated.

The Australian Communications Authority (ACA) implements regulations based on ITU agreements in Australia. As per the ITU allocation, the frequency band 1559 – 1610 MHz in Australia is also reserved for aeronautical and radionavigation-satellite services. However, with regards to all frequency bands, section 10(1) of the Australian Radiofrequency Spectrum plan states that: 'A frequency band may be used for an unspecified service if the unspecified service uses the frequency band to support a specified service'. Hence it seems that L1/L2 PLs may be legally used to augment GPS under existing laws. However, the use of L1/L2 PLs in any such augmentation must not cause harmful interference to the GPS service (section 10(5)). There is no clarification on whether L1/L2 PLs can be used for other purposes; however, the ACA also provides experimental radio service licences for research purposes with conditions similar to those imposed by the FCC.

A recent piece of legislation announced by the ACA concerns the possession of radionavigation-satellite service (RNSS) jamming devices (ACA, 2004). While the operation of RNSS jamming devices was already prohibited, this new legislation allows for the prosecution of a person who supplies a RNSS jamming device or possesses a RNSS jamming device for the purpose of supply. The ACA defines a RNSS jamming device as "a device that is designed to have an adverse effect on the reception by RNSS receivers of RNSS radiocommunications". While L1/L2 PLs aren't explicitly designed for this purpose, they can easily be used to prevent the reception of GPS signals. For instance the popular IN200D PL by IntegriNautics, which has a peak output power of +6 dBm, can jam GPS signals at distances of up to 10 km. Even an L1-only PL can inhibit the reception of L2 signals by jamming satellite C/A codes on L1. This can happen because dual-frequency receivers use the shorter C/A codes to acquire the week-long P codes on L2. With such possibilities in mind, an inquiry was put forward to the ACA as to whether L1/L2 PLs would be classified as RNSS jamming devices. If so, this would certainly be of considerable concern to researchers using/developing L1/L2 PL technology in Australia. However, no reply was received by the time of writing this paper.

From a rather brief look at radiofrequency spectrum regulation in the USA and Australia, it is clear that current legislation does not permit the widespread use of L1 or L2 PLs. Commercially developed PLs often have sections in their documentation that highlight this legal issue. For instance the quick-start application note for the IN200D by IntegriNautics states that users must check the local laws and acquire any necessary licences and that the licences be posted in public view during use. The author is also aware of PL experiments performed for Air Services Australia (ASA) during which a Notice to Airmen (NOTAM) was issued to warn of possible local-area interference to GPS signals. Such precautionary measures are considered important by governing authorities, even if interference to GPS signals may be minimal. The legal issues involved in using L1/L2 PLs are rarely discussed in PL research papers; the most comprehensive overview found by the author is presented by Cobb (1997).

2.2.2 Jamming GPS

GPS satellite signals must compete with each other and with any other signals present in the L1/L2 frequency bands. When competing amongst other satellite signals, the auto-correlation properties of C/A codes allow a GPS receiver to differentiate between the signals of the satellites that are in view. C/A code cross-correlation averages to about -30 dB when considering all Doppler and time offsets; so there is about 30 dB of separation between the satellite signals (the worst case cross-correlation happens to be -21.6 dB).

It was stated earlier that GPS signals have a typical strength of -160 dBW at the surface of the Earth. In contrast, the power level of thermal noise at the antenna of a GPS receiver is -205 dBW/Hz, which is -142 dBW for the C/A code bandwidth of 2.046 MHz. The noise figure of the amplifier immediately following the antenna must also be considered. For typical OTS GPS receivers this will increase the thermal noise to about -138 dB. This means that GPS satellite signals are about 22 dB lower than the noise floor. Since C/A codes are 1023 chips long, a GPS receiver can achieve a processing gain of about 30 dB if it integrates over a single C/A code period (1ms). Typical OTS GPS receivers can integrate for up to 20ms which corresponds to a processing gain of 43 dB. This processing gain allows the GPS receiver to raise the satellite signals up above the noise floor. The result is a post-correlation signal to noise ratio (SNR) of about 21 dB.

In the ideal case, where there are no other signals present in the L1 band, GPS receivers won't have any problem acquiring or tracking unobstructed satellite signals. Typical signal tracking loops of OTS GPS receivers can acquire and track signals with a post-correlation strength

of 6 dB above the noise floor. This leaves a margin of about 15 dB. As satellite signal strengths vary due to factors such as satellite elevation and signal obstruction, a receiver will still be able to track the signals provided the theoretical 15 dB margin is not exceeded.

Forssell and Olsen (2003) performed a study on the susceptibility of commercial GPS receivers to various jamming signals. Modulated and unmodulated continuous carrier waves as well as band-limited white noise centred on the L1 frequency were subjected to three different types of receivers. It was found that for the OTS receiver used, GPS signals were unable to be tracked in the presence of modulated interfering signals that were 36 dB stronger than the GPS signals. White noise required being about 53 dB stronger than GPS signals to have the same effect. The study, however, did not look at spread spectrum modulated interference signals, such as from a PL. Although PL signals can be considered noise-like, PL transmissions are modulated with C/A codes and as a result cross-correlation effects with GPS signals will become significant. Because of this PL signals are likely to interfere at power levels much less than that of white noise as reported by Forssell and Olsen.

Protection from interference to GPS signals can be implemented at different levels. One level is at the receiver's antenna. A temporal (also spectral or notch) filter at the antenna can provide some protection against continuous wave CW interference, though such a filter may be overwhelmed by interference from multiple sources/jammers. With an antenna that has multiple elements, interfering signals can be attenuated by adjusting a weight applied to the outputs of each element. Some techniques that use this approach are cancellation, spatial temporal adaptive processing (STAP) and spatial frequency adaptive processing (SFAP). Cancellation has limited use, and although STAP and SFAP both can provide 25-40 dB of jamming suppression, they do require significant computation as well as high RF dynamic range. Furthermore, the issue of phase perturbations also has to be resolved.

Interference can also be suppressed internally within the receiver. Narrow-width correlators, for example, provide better immunity to interference than ordinary correlators. This was shown in the experiments performed by Forssell and Olsen (2003). A more simple yet effective method is to increase the integration time. A typical OTS receiver is limited to integrating for up to 20ms due to the presence of the 50Hz navigation data bits, which cause 180 degree phase shifts of the carrier. This limit may be overcome by data wipe-off, or by using non-coherent integration. Data wipe-off presumes prior knowledge of data bit values from an external source, for which a data link is required in real-time receivers. Non-coherent integration, meanwhile, can lead to squaring loss and thus an increase in noise. Nevertheless, although increased integration times are achievable, it is done so at the expense of the rate of measurement and navigation solution computation. A better alternative in interference cancellation involves coupling inertial navigation sensors (INS) with the receiver tracking internals at varying degrees. The most benefit is gained from tightly-coupled schemes, though with this approach comes added cost and complexity. Further details of interference cancellation techniques mentioned here are presented by Rounds (2004a and 2004b).

Although a variety of interference suppression techniques are available, they are often limited to military or survey-grade GPS receivers, which are generally quite expensive. Typical OTS GPS receivers do not provide sufficient immunity against interference to be of any practical use with PL systems. To overcome this limitation PL output can be set to a low enough power level that avoids interference, though this can be a major operational restriction depending on the type of application. Another technique for minimizing interference is to pulse the PL signals. Pulsing at low duty cycles, which is often used to mitigate near-far effects, allows PL signals to be stronger than GPS signals without denying GPS to non-participating receivers.

To illustrate, consider a PL pulsing at a 10% duty cycle over the C/A code period. This means that it is on for 100μs and off for 900μs. In this pulsing scheme, over a C/A code period the GPS signals' duration is 10 times longer than that of the PL signal. This corresponds to a 10 dB difference in favour of the GPS signals. So, if a PL signal whose power is 10 dB greater than GPS signals is pulsed, to a GPS receiver the pulsed signal will appear to have the same power level as the GPS signals. In this case no interference will be caused, since the PL signal power is averaged over the 1ms integration period. For stronger PL signals, though, there will be interference effects. A receiver tracking GPS signals will correlate over only the GPS signals for 90% of the time. However, for the 10% during which the PL is transmitting, the receiver will also correlate over the PL's in-band transmissions. If the PL signal is significantly stronger than the GPS signal, a reduced correlation gain in the GPS signal will be observed during the PL pulse. The effects of cross-correlation between the PL and GPS signals will also become quite significant. These combined factors result in a lower SNR for the GPS signals as observed by a receiver. The longer the pulse duration, the higher the reduction of SNR due to increased interference.

Fig. 1 (Cobb, 1997) illustrates the relationship between PL pulse duration and GPS signal SNR (S/N in the figure), as observed by a typical GPS receiver. The PL signal power is assumed to be at saturation level; which is equal to the level of thermal noise 23 dB greater than the GPS signals. The minimum SNR threshold at which a

receiver is able to track signals is indicated on the figure to be about 6 dB. According to Cobb, this threshold is exceeded for pulse duty cycles greater than 20% (curve labelled 'without blanking'). The minimum pulse duty cycle required to track a PL signal is also determined from the figure to be 10%. This means that at most two such PLs outputting high power may operate within the same area without denying GPS to non-participating receivers.

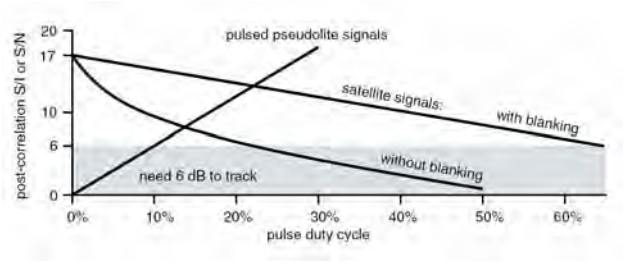


Fig. 1 PL Pulse Duty Cycle Trade-off (Cobb, 1997).

2.2.3 Cross-correlation

The C/A codes used in GPS are taken from a family of codes known as Gold codes. C/A codes were especially chosen for their good multiple access properties over their 1023 chip period. Their relatively short code length, which is clocked at a 1.023 Mbps rate, also permits rapid acquisition. Cross-correlation of C/A codes is dependant on the delay offset between any given pair of codes. The maximum cross-correlation of C/A codes is -24 dB, and this worst-case is likely to occur 25% of the time. The rest of the time cross-correlation is -60 dB; over time, the average cross-correlation can be taken to be about -30 dB. The different C/A code used to modulate each satellite signal isn't the only major distinction between any two signals. The orbit of GPS satellites impose a Doppler frequency offset (up to +/- 6 kHz) on signals that also affects cross-correlation. The worst case cross-correlation over the expected Doppler offsets is -21.6 dB. With GPS, the worst case cross-correlation is typically not considered too important. This is because any coincidence of Doppler offset and code offset that manifests in a worst case is only transitory due to the constant motion of the satellites. PLs on the other hand are static devices. Therefore, a worst case combination of frequency and code offset is unlikely to vary as much or as often. Any variation will be entirely due to user movements, which generate Doppler frequencies considerably less than the motion of satellites. And, since a C/A code chip at 1.023 Mbps is 300 m long, the resulting changes in code phase due to typical user movement are also considerably less.

In the example pulsing scheme of the previous section, the PL signal pulsing was performed relative to the C/A code period. Receivers will always integrate for at least one C/A code period; typically over one period during acquisition before increasing to as much as 20ms once bit synchronization is achieved. So, the presented pulsing scheme ensures that a PL signal pulse will always be present during every integration period. If the PL pulse duty cycles are not kept synchronous with the C/A code period, then the case may arise where PL pulses are absent from some integration periods but dominant in other ones. This may cause some receivers to have trouble with acquisition and tracking of PL signals (Cobb, 1997).

For accurate and reliable tracking of C/A code modulated signals it is important to correlate over the entire code sequence. This is the only way to fully exploit the orthogonal property of the C/A codes. Pulsing PL signals thus presents a problem, since only a portion of the C/A code is transmitted during the active-phase of a pulse cycle. During the off-phase a receiver tracking the PL will correlate against noise, GPS signals or the transmission of another PL, depending on the environment. In such a case, cross-correlation noise will dominate. This effect places a hard limit on the minimum pulse duty cycle. Some duty cycles as low as 7% have been used while still reliably tracking C/A code PLs (Cobb, 1997). More typical, though, is the RTCM recommended duty cycle of 1/11 (~10%). This allows for correlation with 93 chips of each PL signal's code, and allows a maximum of 11 PLs to be used at any given time if one so desires (eg in an indoor PL only application).

With simple synchronous pulsing at a 1/11 duty cycle, a PL will transmit the same 93 code chips every C/A code period. This, however, does not utilize the whole code sequence and may cause aliasing in the signal. A better implementation would use a different code interval every code period. For the 1/11 pulse duty cycle, the entire code sequence can eventually be transmitted in 11 code periods, which correspond to 11ms. Since typical GPS receivers integrate for up to 20ms this method will allow any receiver to correlate over the entire code sequence of a pulsed PL signal.

If the C/A code is chipped at faster rates more code chips can be transmitted during each pulse. For instance, chipping at 10.23 Mbps will allow the complete transmission of the C/A code sequence at the same pulse rate presented earlier (100µs on, 900µs off). This offers a huge advantage in correlation. A faster chipping rate also increases the maximum number of PLs that can be accommodated in a synchronous pulsing system. Although some GPS receiver hardware may permit modifications to allow for faster chipping rates, the same cannot be said for all typical receivers. Another solution to allow higher pulse duty cycles, proposed by Cobb (1997), involves disabling the correlators during periods of high cross-correlation interference. For instance a correlator channel assigned to GPS signals can be

disabled during PL active pulse cycles. The effect of this is to prevent correlation with the stronger PL signals, thus eliminating cross-correlation influences. Inversely, correlators assigned to PL signals can be disabled during pulse-off cycles to eliminate cross-correlation noise in the same manner. The expected benefit of this technique, referred to as correlator blanking, is depicted in Fig. 1. The satellite signal SNR curve (labelled 'with blanking' in the figure) shows that a higher pulse duty cycle of up to 60% can be achieved. Cobb also briefly discusses an implementation of correlator blanking, stating that a saturation detector at the front-end could serve as an indication of PL pulses. Correlators can then be enabled or disabled accordingly. Unfortunately, this requires significant modification to a typical receiver, but may be easily incorporated into newer receiver designs. A problem with this implementation is that it does not address how two pulses from different PLs can be distinguished. All correlators assigned to a PL will be enabled during saturation. If there are two PLs the receiver will correlate over pulses from both devices, increasing cross-correlation noise. Also, as the distance to a PL increases the PL signal will cease to saturate the receiver. In this case, the saturation detector will not respond to the PL signal, leaving the correlator disabled during the PL pulse. Rather than helping, correlator blanking can prevent the acquisition and tracking of a PL signal; not correlating at all is worse than correlating against interference for 90% of the time. A further extension to this tragic case is when a pulse from a near PL enables the correlator assigned to the far PL, this time leaving the receiver to correlate against the wrong signal.

2.2.4 C/A code chip rate in high multipath

Multipath is a major concern for PL systems that operate indoors and in other cluttered environments. These environments generate multipath signals that may outnumber the direct signal by many orders of magnitude, and each may have different delays, strengths and polarizations. According to a simple analysis, assuming the multipath signal is never stronger than the direct signal, the peak multipath error can equal one half of a chip length (Braasch, 1996). This evaluates to nearly 150 m at the C/A code chipping rate. The actual error can be a lot worse when considering the myriad of different reflections in, say, a warehouse or urban canyon. It is therefore unlikely that a chipping rate of 1.023 Mbps can provide an acceptable level of multipath rejection in high multipath environments. Faster chipping rates however, will have reduced peak multipath errors. For example, by the same analysis, the 10.23 Mbps rate used for P codes has a peak multipath error of only 15 m.

2.2.5 Receiver front-end saturation

A typical GPS receiver front-end consists of an antenna, followed immediately by a low noise amplifier (LNA), then a filter, one or more stages of down-conversion, and finally an analog-to-digital converter (ADC). An automatic gain control (AGC) may also be implemented to keep the input signal within the dynamic range of the front-end. (Dynamic range refers to the range of signal powers that a component (or system of components) can process without generating distortion.)

Recall that the GPS signals at the Earth's surface are extremely weak. When unobstructed, these signals are approximately of equal strength, except for minor variations due to changes in satellite elevation. GPS receiver manufacturers take advantage of these characteristics of GPS signals to simplify the design of their receivers. One feature that is common of typical GPS receivers is the low dynamic range of the front-end. Since GPS signals are buried 22 dB below the noise floor at a relatively constant level, it makes little economic sense to design an RF front-end that compensates for large signal power deviations when only slight changes in noise level due to changes in temperature is expected.

A receiver will see only thermal noise when no transmissions other than the GPS signals are present in the L bands. In the presence of strong interference, like a PL signal 30 dB above noise, the receiver's AGC will reduce gain to compensate for increased signal power. The AGC tries to scale the input signal, which is expected to be thermal noise, over the full dynamic range of the ADC. The result, though, is that the thermal noise and GPS signals are attenuated along with the interfering signal. If the attenuation exceeds about 15 dB then the receiver will be unable to track GPS signals. Furthermore, if the reduction of gain is insufficient, the interfering PL signal will also be clipped at the extremities of the ADC's range. The receiver then is said to be in saturation.

In the presence of such interference, the RF components used in a typical GPS receiver can be operating outside of their dynamic range. In effect, they are operating beyond the limits of their specification, in which their function may become non-linear. One important component, the mixer used in down-conversion, should always be operating in its linear range. Otherwise, if the input power exceeds its threshold, compression of the received signal can result that would generate harmonics which cause distortion. Sustained excessive interference may even cause damage to some components.

Typical GPS receivers in saturation will perform quite poorly simply because they are not designed to operate in this condition. In these receivers, interference and cross-correlation effects are magnified. Detection thresholds must be increased to avoid locking onto false peaks.

Furthermore, a receiver can only track one saturating PL signal at a time, and only if it significantly stronger than the other PL signals. Two saturating PL signals of the same strength will interfere with each other so that neither can be accurately tracked. This has implications for pulsed PL networks. As mentioned in section 2.1.3, the active-phase of a pulse cycle from multiple PLs may overlap in time due to PL clock drift. This will decrease PL signal SNR, due to partial jamming, and may also cause tracking errors or prevent acquisition. The pulsing scheme of multiple PLs should be synchronized so that duty cycles do not coincide. In wide-area deployment of PLs attention must also be given to propagation delays that may cause some overlapping of adjacent synchronous PL pulses. A more thorough discussion of the effects of GPS receiver saturation is presented by Cobb (1997).

3 Impact of L1 C/A code issues on pseudolite applications

This section will briefly outline how the presented issues may impact on PL usage for several different applications.

3.1 Aircraft approach and landing

Legal: L1 C/A code PLs used for aiding aircraft navigation appear to be legal under FCC and ACA regulations.

Jamming/Saturation: Radionavigation systems for aircraft are subject to stringent performance requirements. Therefore, any factors that limit the accuracy and reliability of these systems are of great concern. PLs must provide adequate coverage to enable aircraft to acquire and track their signals well before the landing approach. This involves distances of up to 15 km. In order to provide this coverage, PLs must transmit at high power, meaning receivers at closer distances will be saturated. The saturation of receivers by L1 C/A code PL signals limit the number of PLs that can be deployed around an airfield to about 2. Users in this system will be denied more PLs that could offer better geometry and provide extra ranging signals for reliability. The jamming effect of high power PL signals combined with saturation effects also degrades the performance of participating and non-participating receivers.

3.2 Integrated GPS/pseudolite positioning

Legal: According to current FCC regulations, operating an L1 PL is illegal without an experimental radio station (or other) licence. It is unclear as to whether any changes in regulations will be made. For research work which typically involves limited intervals of PL transmission this legal obstacle may be considered minor. However, from a commercial perspective this is a major hurdle that stands in the way of implementing a round-the-clock L1 PL service which is integrated with GPS. Furthermore, under current ACA regulations designers and suppliers of L1 PLs, which may be considered as RNSS jamming devices, are liable to prosecution and heavy fines.

Jamming/Saturation: From a consumer's perspective, those who have bought and paid for GPS receivers that are either unable to track PL signals or are unable to handle even minimal interference from PL signals, are unlikely to support the deployment of L1 C/A code PLs in their area. PLs may instead be configured to transmit at low power, although this requires their deployment in increased numbers to provide the required coverage.

3.3 Indoor pseudolite-only systems

Legal: Operating L1 PLs indoors without a licence, even at low power levels, is still considered illegal.

Jamming/Saturation: It is difficult to maintain constant signal levels in cluttered indoor environments, even if the near-far effect was eliminated. PL signals will be received at power levels varying over a very wide range, especially if signals are required to propagate through walls and ceilings. A weak signal from an adjacent room will most likely be jammed by a closer PL. Also, current OTS receivers with their low dynamic range will exhibit difficulties in coping with the high range of PL signal strengths. Receivers may frequently move in and out of saturation. GPS signals indoors are so heavily attenuated that they are unusable by most receivers. From this perspective, the jamming of indoor GPS signals by L1 PLs could be forgiven. However, new Assisted-GPS (AGPS) techniques have been developed that enable special GPS receiver designs to acquire and track weak indoor signals. In this case, jamming GPS signals is still an issue for indoor systems.

Multipath: By far the greatest limitation of L1 C/A code PLs in indoor applications is the low multipath mitigation offered by its signal structure. The chipping rate of the C/A code is just too low, leaving the receivers vulnerable to high multipath range errors.

4 Future directions

Research into the use of PLs for many different applications has been conducted for over two decades. Throughout this time almost all of the experimentation has involved PLs using the civilian signal structure of GPS. The primary reason for this was that it enabled the

use of existing receiver hardware with only minor firmware modifications. Innovative methods of incorporating PLs into navigation applications were developed. Promise in this new technology led several companies providing positioning solutions to develop and sell L1 C/A code PLs. However, to date the use of PLs has been limited to the realm of research or to highly specialized applications. The author believes that this is due to the various issues which limit the effectiveness of L1 C/A code PLs as discussed in the previous sections.

With the development of Galileo and the modernization of GPS, some of the issues in this paper will be addressed. One is the faster chipping rate on civilian signals that will reduce maximum multipath errors; and the other is the two civilian frequencies that provide a means for ambiguity resolution. When both these systems are operational, PLs augmenting Galileo and modernized GPS on the new civilian signals will be able to provide some performance gains over the current PLs augmenting GPS. Unfortunately, no benefit will be gained with regards to jamming, interference or receiver saturation. Furthermore, if a provision for the use of PLs to augment the new signals is not defined, the same legal obstacles will remain in place, limiting their widespread use.

A current trend in PL research and development is the movement away from preserving backward compatibility with existing GPS receivers. Most of the ideas are not new, but the driving force behind them is a recent realization that a signal structure used for global positioning cannot always be used for local positioning, especially indoors. One of the ideas being implemented is multi-frequency transmissions in the Industrial Scientific and Medical (ISM) band (Zimmerman et al, 2000). Operating in the ISM band requires no licence, as long as the maximum power output is limited to less than 1 Watt (FCC-47CFR15.247). The major advantage gained is the legal freedom to transmit on multiple bands, including 915 MHz, 2.45 GHz and 5.8 GHz. Receivers can exploit the frequency diversity to resolve carrier cycle ambiguities. On the other hand, there may be interference issues with other devices that also use the ISM band, such as cordless phones and 802.11b/g.

Locata Corporation, an Australian company based in Canberra, is currently developing PL technology for operation in the ISM band. Locata Corporation is building on the success of its synchronous L1 prototype PLs that have proven accurate performance in an indoor, high multipath environment (Barnes et al, 2004). A Locata PL, called a LocataLite, is an intelligent transceiver that transmits on dual frequencies in the ISM band. A LocataLite is able to synchronize to the transmissions of other LocataLites, forming a synchronous positioning network called a LocataNet. A receiver device, called a Locata, can determine its position within the LocataNet using both code and carrier phase measurements. The main advantage offered by a LocataNet is its synchronous signals, which allow standalone navigation and precise time transfer.

Novariant (formerly IntegriNautics) has also developed an off-frequency PL system called a Terralite™ (Novariant, 2004). A Terralite™ transmits a proprietary signal called XPS. At the time of writing, little public information was available on the frequency or signal structure of the XPS signal, though the requirement for a licence was announced. Novariant also provides a tri-frequency receiver that is able to track both L1 and L2 GPS signals in addition to the XPS signal. Novariant is targeting specialized heavy industries, such as mining, where GPS is one of the primary positioning systems used. In such industries considerable investment has been made into GPS receivers and even into L1 C/A code PLs in order to address their positioning needs. However, in light of the issues discussed in the previous sections, this can only provide limited performance. The Terralite™ XPS signal offers an additional signal designed to overcome the restrictions of L1 PLs that are currently used. Furthermore, the ability of the XPS receiver to also receive both L1 and L2 signals means that a client's existing GPS infrastructure can still be utilized.

Multi-frequency PLs can complicate the design of the front-end of receivers that are intended to process them. For example the Novariant XPS receiver, which must acquire L1, L2 and a third frequency. The number of components required will increase; especially if separate RF paths are to be used for each frequency. This can add significant cost to receivers. Currently, research is being performed on reducing the need for a complicated front-end by using direct sampling instead (Psiaki et al, 2003). Another development in receiver technology is the software correlator receiver. These receivers perform all signal processing tasks after digital conversion, including correlation, within a multi-purpose processor. Software receivers can offer flexibility in adapting to the various signal modulation plans of future PL systems. However, significant processing power will be required for processing signals with high chipping rates. Pany and Eissfeller (2004) address this performance issue by using sub-Nyquist sample rates.

In the past, jamming of GPS signals has been primarily a concern for the military. However, beyond military concerns are civilian vulnerabilities related to critical infrastructure. Cell phone networks, commercial fishing, transportation, emergency response, air navigation and air traffic control; all have developed a reliance on GPS for timing and navigation. The list of dependencies will grow as accurate position information becomes more tightly integrated into business practices. This is expected to spur the development of more robust receivers to address some of the issues in jamming and interference.

For indoor positioning, the high multipath environment necessitates the use of more robust signal structures. An example is Locata Corporation's synchronously pulsed signals with fast chipping rates. Also, their use of multiple frequencies provides redundancy to protect against multipath fades. Another is presented by Progri et al (2004), who propose a very unique indoor positioning system using ultra-wideband (UWB) signals. This system uses orthogonal frequency division multiplexing (OFDM) with frequency division multiple access (FDMA) that minimizes cross-channel interference. The UWB signals also offer good protection from multipath fades.

Other techniques for indoor positioning have been proposed. One is Siemen's Local Positioning Radar (LPR) system (Siemens, 2004). This system comprises of transponders that are deployed around the area of interest. A base station transmits pulses which are received and retransmitted by each transponder. The base station will receive each retransmitted signal at different times according to its distance from each transponder, hence providing range information. Another is a UWB time-of-arrival (TOA) technique (Fontana and Gunderson, 2002). In this system, passive receivers in a chained network detect the pulsed transmissions from active UWB tags. TOA data from the receivers are then processed to determine the position of each active UWB tag.

Assisted GPS (AGPS) techniques provided by Global Locate, SiRF and Q-Comm use massively parallel correlator architectures to detect the heavily attenuated GPS signals indoors (Global Locate, 2004). To do this, a data link provides known navigation messages to the AGPS receiver, enabling it to integrate for longer periods. However, considerable infrastructure is needed to provide navigation messages to all users.

Rosum Corporation proposes the use of digital television signals to augment GPS (Rosum, 2004). TV signals use a variety of VHF and UHF bands, hence giving protection against multipath fades. Their signals are also stronger than GPS signals by about 40 dB. However, this technique also requires significant infrastructure to monitor channel stability and timing information. Also, currently receivers using this technology are unable to calculate their own position; this is performed at a location server and sent back to a receiver because current TV signals are not synchronous to a common clock.

In some of the techniques above, the position of a location device is determined by the system, not by the device itself. This may be desirable for applications such as centralized asset management, but would be unsuitable for other applications such as autonomous robotics or guidance and control. Furthermore, the reliance on data links for navigation messages or timing information may be a hindrance. For these and other reasons, the author believes that it is unlikely for any single positioning system to be the answer to every application's needs. What the author does believe is important, though, is that the users of positioning technology should have full control over their signals. Only then can a positioning system be optimized to best suit one's needs.

5 Conclusions

L1 C/A code PLs simplify experimentation with PL technology. They allow the convenient use of existing OTS GPS receivers. However, this practicality has disadvantages. One involves the legal issues of transmitting on L1, a band protected globally by legislation. L1 PLs can also jam GPS signals, denying GPS to non-participating receivers. The C/A code chipping rate used by these PLs also does not provide sufficient protection against multipath. Typical OTS GPS receivers are easily saturated by strong PL signals; this limits the accuracy and reliability of the PL system. These factors add further complications to general PL issues such as near-far, multipath and synchronization. A current trend in PL development serves to address these complications by moving off the L1 frequency and using more robust signal structures.

7 Acknowledgements

The author would like to thank Dr. Jimmy Lamance, Dave Small, Dr. Joel Barnes, and Dr Dorota Grejner-Brzezinska for their valuable input and assistance.

References

ACA (2004): ***Media Release No. 73***, 1 September, http://www.aca.gov.au/aca_home/media_releases/media_enquiries/2004/04-73.htm

Barnes J.; Rizos C.; Wang J.; Small D.; Voigt G.; Gambale N. (2003): ***High precision indoor and outdoor positioning using LocataNet,*** 2003 Int. Symp. on GPS/GNSS, Tokyo, Japan, 15-18 November, 9-18

Barnes J.; Rizos C.; Kanli M.; Small D.; Voigt G.; Gambale N.; Lamance J.; Nunan T.; Reid C. (2004): ***Indoor industrial machine guidance using Locata: A pilot study at BlueScope Steel,*** Proceedings of the 60th Annual Meeting of the U.S. Inst. Of Navigation, Dayton, Ohio, 7-9 June, 533-540.

Braasch M.S. (1996): ***Multipath Effects,*** Global Positioning System: Theory and Applications I, AIAA, 547-568.

Cobb H.S. (1997): ***GPS pseudolites: Theory, design, and applications,*** PhD Dissertation, Stanford University.

Dai L.; Wang J.; Tsujii T.; Rizos C. (2001): ***Pseudolite Applications in Positioning and Navigation: Modelling and Geometric Analysis,*** Int. Symp. on Kinematic Systems

in Geodesy, Geomatics & Navigation, Banff, Canada, 5-8 June, 482-489.

Fontana R.J.; Gunderson S.J.; (2002): ***Ultra-WideBand Precision Asset Location System,*** IEEE Conference on Ultra Wideband Systems and Technologies, May 2002, Baltimore, MD.

Forssell B.; Olsen T.B. (2003): ***Jamming GPS: Susceptibility of Some Civil GPS Receivers,*** GPS World, Vol. 14, No. 1, January, 54-58.

Global Locate (2004): ***Global Locate: Home,*** http://www.globallocate.com

Irsigler M.; Eissfeller B. (2003): ***Comparison of Multipath Mitigation Techniques with Consideration of Future Signal Structures,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 September, 2584-2592.

Klein D.; Parkinson B.W. (1986): ***The use of pseudo-satellites for improving GPS performance,*** Global Positioning System (red book), Vol III, Institute of Navigation, 135-146.

Kee C.; Jun H.; Yun D. (2000): ***Development of Indoor Navigation System using Asynchronous Pseudolites,*** Proceedings of US Institute of Navigation GPS-2000, Salt Lake City, Utah, 19-22 September, 1038-1045.

Kunysz W. (2003): ***A Three Dimensional Choke Ring Ground Plane Antenna,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 September, 1883-1888.

Wang J.; Lee H.K. (2002): ***Impact of Pseudolite Location Errors on Positioning,*** Geomatics Research Australasia, 77, 81-94.

Madhani P.H.; Axelrad P. (2001): ***Mitigation of the Near-Far Problem by Successive Interference Cancellation,*** Proceedings of the 14th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Salt Lake City, Utah, 11-14 September, 148-154.

Misra P.; Enge P. (2001): ***Global Positioning System: Signals, Measurements, and Performance,*** Ganga-Jamuna Press, Lincoln, Massachusetts, 2001.

Novariant (2004): ***Novariant: Mining – Products,*** http://www.novariant.com/mining/products/index.cfm

Pany T.; Eissfeller B. (2004): ***Code and Phase Tracking of Generic PRN Signals with Sub-Nyquist Sample Rates,*** Journal of the Institute of Navigation, Vol. 51, No. 2, Summer 2004, 143-159.

Progri I.F.; Michalson W.R.; Cyganski D. (2004): ***An OFDM/FDMA Indoor Geolocation System,*** Journal of the Institute of Navigation, Vol. 51, No. 2, Summer 2004, 133-142.

Psiaki M.L.; Powell S.P.; Jung H.; Kintner Jr. P.M. (2003): ***Design and Practical Implementation of Multi-Frequency RF Front Ends Using Direct RF Sampling,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 September, 90-102.

Rosum (2004): ***Rosum Corporation. Reliable Location. Inside and Out,*** http://www.rosum.com

Rounds S (2004a): ***Jamming protection of GPS receivers; Part I: Receiver Enhancements,*** GPS World, Vol. 15, No. 1, January.

Rounds S (2004b): ***Jamming protection of GPS receivers; Part II: Antenna Enhancements,*** GPS World, Vol. 15, No. 2, February, 38 - 45.

Siemens (2004): ***Methods and Apparatus for Determining a Location of a Device,*** United States Patent Application Publication No. US 2004/0157625 A1.

Söderholm S.; Juhola T.; Saarnimo T.; Karttunen V. (2001): ***Indoor Navigation Using a GPS Receiver,*** Proceedings of the 14th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Salt Lake City, Utah, 11-14 September, 1479-1486.

Söderholm S.; Jokitalo T. (2002): ***Synchronized Pseudolites - The Key to Indoor Navigation,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 September, 226-230.

Soon B.H.K.; Poh E.K.; Barnes J.; Zhang J.; Lee H.K.; Lee H.K.; Rizos C. (2003): ***Flight test results of precision approach and landing augmented by airport pseudolites***, Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Institute of Navigation, Portland, Oregon, 9-12 September, 2318-2325.

Stolk K.; Brown A. (2003): ***Phase Center Calibration and Multipath Test Results of a Digital Beam-Steered Antenna Array,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Institute of Navigation, Portland, Oregon, 9-12 September, 1889-1897.

Van Dierendonck A.; Fenton P.; Ford T. (1992): ***Theory and Performance of Narrow Correlator Spacing in a GPS Receiver,*** Navigation, Volume 39, Number 3, 265-283.

Wang J (2002) ***Pseudolite Applications in Positioning and Navigation: Progress and problems,*** Journal of Global Positioning Systems, 1(1): 48-56.

Weiss J.P.; Anderson S.; Axelrad P. (2003): ***Multipath Modeling and Test Results for JPALS Ground Station Receivers,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 September, 1801-1811.

Yun D.; Kee C. (2002): ***Centimeter Accuracy Stand-alone Indoor Navigation System By Synchronized Pseudolite Constellation,*** Proceedings of the 16th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Portland, Oregon, 9-12 September, 213-225.

Zimmerman K.R.; Cohen C.E.; Lawrence D.G.; Montgomery P.Y.; Cobb H.S.; Melto W.C. (2000): ***Multi-Frequency Pseudolites for Instantaneous Carrier Ambiguity Resolution,*** Proceedings of US Institute of Navigation GPS-2000, Salt Lake City, Utah, 19-22 September, 1024-1030.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 167-172

Journal of Global Positioning Systems
2004
© CPGPS

Real-time Experiment of Feature Tracking/Mapping using a low-cost Vision and GPS/INS System on an UAV platform

Jonghyuk Kim, Matthew Ridley, Eric Nettleton[1], and Salah Sukkarieh

ARC Centre of Excellence for Autonomous Systems, The University of Sydney, Australia
e-mail: {jhkim,m.ridley,salah}@acfr.usyd.edu.au Tel: +61 (0)2 9351 8515; Fax: +61 (0)2 9351 7474,

[1]Advanced Information Processing Department, Advanced Technology Centre, BAE Systems, UK
e-mail: eric.nettleton@baesystems.com

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. This paper presents the real-time results of an air-to-ground feature tracking algorithm using a passive vision camera and a low-cost GPS/INS navigation system on a UAV (Uninhabited Air Vehicle) platform. The vision payload is able to observe a number of ground features, and the GPS/INS navigation system is used in conjunction with a waypoints-based guidance and flight control module. Due to limited processing resources, the vision node employs a simple but fast method of point based feature extraction algorithm. The feature tracking performance is greatly affected by the accuracy of the on-board navigation system. Conversely though, it can be used as a performance indicator of the navigation filter by comparing it with the truth feature location and some simple geometry. This paper will present the results of targeting performance against known location of features, and hence verifying the accuracy of the real time GPS/INS system

Key words: low cost GPS/INS, vision sensor, features tracking, real-time navigation, UAV

1 Introduction

The multi-target (or multi-feature) tracking problem is to track all features of interest within some accuracy, and therefore build a picture of all objects in that area. It is an integral part of surveillance systems employing one or more sensors to interpret the environment. Typical on-board sensor systems, such as radar, infrared, vision, and vision laser provide measurements from features of interest. There has been extensive research in the area of multi-feature tracking. This work has concentrated on topics such as computational efficiency, data association, model accuracy, multiple-model techniques, multiple hypothesis techniques and spatial representations as in the work of Blackman, 1999, and Bar-Shalom and Blair, 2000.

In real-world applications, however, the tracking performance is not only affected by the tracking algorithms but also by the accuracy of the on-board navigation system. This is due to the feature registration process. Even a small attitude error in the navigation system can cause a significant deviation in the feature location amplified by the range information. Hence most commercial remote mapping/tracking systems are equipped with a high-grade navigation system to minimise the effects of the vehicle attitude error. If an Uninhabited Air Vehicle (UAV) is used as the platform for the feature-tracking, it poses greater restrictions due to the availability of a compact, low-power navigation system. Commercially available navigation systems for UAVs are very limited or too expensive for most academic research purposes.

This paper will present a cost-effective airborne feature-tracking system by incorporating a low-cost inertial and Global Positioning System (GPS) sensor for UAV navigation, and by employing a low-cost passive vision camera for feature observation. Figure 1 illustrates the system architecture of the feature tracking/mapping system for the Brumby, a UAV platform developed in the University of Sydney. For the purpose of modularity and scalability, the sensor node is designed in a decentralised fashion, which allows it to be connected or disconnected easily. The navigation solution is computed from the flight control computer, which performs GPS/INS fusion, and guidance and control for autonomous flight. The

sensor payloads connected to the vehicle bus use this navigation solution for the feature tracking, radar gimbal control, and time synchronisation.

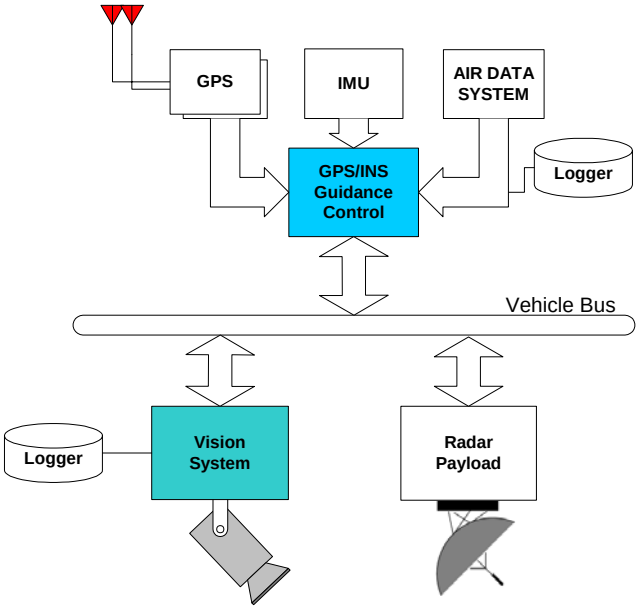


Fig. 1 Modular structure of the feature tracking and navigation system

Section 2 will present the vision system including the sensor and tracking algorithm and Section 3 will provide the GPS/INS integration. In Section 4, details of the flight vehicle and on-board system will be presented. Section 5 will present the real-time flight results based on the Brumby UAV, and then Section 6 will provide conclusions and future work.

2 Vision System

2.1 Passive Vision Camera

An on-board vision sensor provides feature observations to the feature-tracking computer. The vision system makes use of a low cost, lightweight, monochrome CCS-SONY-HR camera from Sony as shown in Figure 2. This imaging sensor has a resolution of 600 horizontal lines using a 12V power source. It has a composite video output, which gives images at up to 50Hz, or 25Hz when the images are interlaced. This occurs as the odd and even lines are used to form separate images 20ms apart. The vision sensor is mounted in the second payload bay of the Brumby Mk-III, immediately behind the forward bulkhead. The sensor is mounted pointing down as shown in Figure 2.

Typical airborne images from this sensor are shown in Figure 3. Artificial landmarks were placed on the ground before flight. These are plastic sheets for easier identification from the vision system. Due to limited

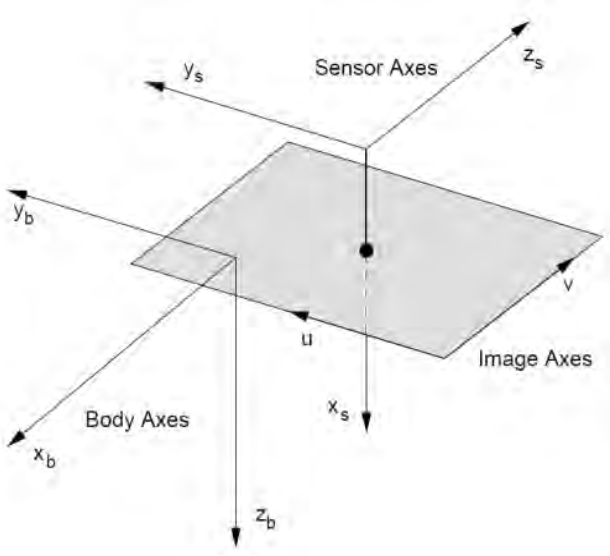


Fig. 2 Vision camera used (top) and body and camera frames whose x-axis is aligned to point downward (bottom)

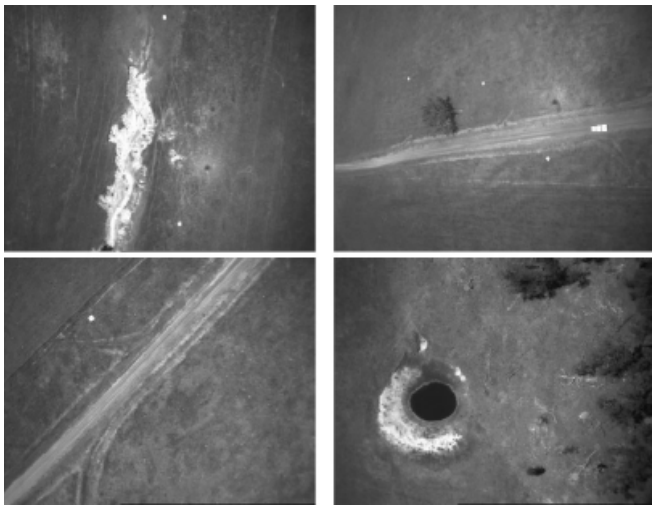


Fig. 3 Aerial images during flight test which show several white artificial features as well as some natural features such as road, dam, and trees.

processing resources, a simple but fast method of point based feature extraction is employed. All pixels above a threshold are converted into line segments. A range gate

performs data association on these segments and the centre of mass of the pixels is obtained. The mass, aspect ratio and density of the cluster of pixels is then utilised for feature identification. The bearing and elevation to the feature can then be generated. Although the vision sensor does not provide range directly, an estimated value is generated based on the known size of the features.

2.2 Feature Tracking Algorithm

The locations of the sensor and platform are provided from the GPS/INS system through the vehicle bus. This location information is used to convert all relative observations to a global Cartesian frame in which tracking takes place. This conversion is performed in the sensor pre-processing stage, which makes the filter observation model a simple linear model. In global coordinates, the x and y position and velocity are modelled as an integrated Ornstein-Uhlenbeck process as in paper by Stone et al, 1999. This process models the velocity as Brownian motion, which can be bounded by appropriate choice of the model parameter γ. The z position is modelled as a simple Brownian process. This can be expressed as:

$$\mathbf{x}(k+1) = \mathbf{F}(k)\mathbf{x}(k) + \mathbf{G}(k)\mathbf{w}(k) \quad (1)$$

$$\mathbf{z}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k)\,, \quad (2)$$

where the state vector is

$$\mathbf{x}(k) = \begin{bmatrix} x(k) & \dot{x}(k) & y(k) & \dot{y}(k) & z(k) \end{bmatrix}^T . \quad (3)$$

The state transition matrix for this system is given by (Ridley, 2002)

$$\mathbf{F}(k) = \begin{bmatrix} 1 & \Delta t & 0 & 0 & 0 \\ 0 & F_v & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta t & 0 \\ 0 & 0 & 0 & F_v & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{G}(k) = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{\Delta t}(1-F_v) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \sqrt{\Delta t}(1-F_v) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$\mathbf{Q}(k) = \begin{bmatrix} \boldsymbol{\sigma}_{vx}^2 & 0 & 0 \\ 0 & \boldsymbol{\sigma}_{vy}^2 & 0 \\ 0 & 0 & \boldsymbol{\sigma}_z^2 \end{bmatrix}$$

with F_v being defined as $e^{-\Delta t\gamma}$ using Brownian motion parameter.

To simplify the filter observation model, the sensor observations in range, bearing and elevation are converted into Cartesian coordinates $\begin{bmatrix} x & y & z \end{bmatrix}^T$ in a global reference frame during the sensor pre-processing stage. The observations are in the form $\mathbf{z}(k) = \begin{bmatrix} x & y & z \end{bmatrix}^T$, hence the observation matrix and noise strength matrix are

$$\mathbf{H}(k) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{R}(k) = \begin{bmatrix} \boldsymbol{\sigma}_x^2 & \boldsymbol{\sigma}_{xy} & \boldsymbol{\sigma}_{xz} \\ \boldsymbol{\sigma}_{yx} & \boldsymbol{\sigma}_y^2 & \boldsymbol{\sigma}_{yz} \\ \boldsymbol{\sigma}_{zx} & \boldsymbol{\sigma}_{zy} & \boldsymbol{\sigma}_z^2 \end{bmatrix}. \quad (5)$$

The noise strength matrix are computed by using the Jacobians of the polar to Cartesian transformation function, hence it contains cross-correlation terms.

Using this feature model and vision observation model, the tracking filter estimates the position and velocity of the features on the ground. Data association between the observation and feature are performed by using the innovation gate method within the tracking filter.

3 Navigation System

3.1 Inertial Navigation

The inertial navigation algorithm is required to predict the high-dynamic vehicle motions using the Inertial Measurement Unit (IMU). In this implementation a quaternion-based strapdown INS algorithm formulated in earth-fixed tangent frame is used (Kim, 2004):

$$\begin{bmatrix} \mathbf{p}^n(k) \\ \mathbf{v}^n(k) \\ \mathbf{q}^n(k) \end{bmatrix} = \begin{bmatrix} \mathbf{p}^n(k-1) + \mathbf{v}^n(k-1)\Delta t \\ \mathbf{v}^n(k-1) + [(\mathbf{q}^n(k-1) \otimes \mathbf{f}^b(k)) \\ \qquad \otimes (\mathbf{q}^n)^*(k-1) + \mathbf{g}^n]\Delta t \\ \mathbf{q}^n(k-1) \otimes \Delta\mathbf{q}^n(k-1) \end{bmatrix}, \quad (6)$$

where $\mathbf{p}^n(k), \mathbf{v}^n(k), \mathbf{q}^n(k)$ represent position, velocity, and quaternion respectively at discrete time k, Δt is the time for the position and velocity update interval, $(\mathbf{q}^n)^*(k)$ is a quaternion conjugate for the vector transformation, $\otimes$ represents a quaternion multiplication, and $\Delta\mathbf{q}^n(k)$ is a delta quaternion computed from gyroscope readings during the attitude update interval.

3.2 GPS/INS Integration

In the complementary GPS/INS architecture, the fusion filter estimates the errors in INS by observing vehicle

states. Hence the state is defined as the error in the position, velocity and attitude expressed in the navigation frame:

$$\delta\mathbf{x}(k) = [\delta\mathbf{p}^n(k) \quad \delta\mathbf{v}^n(k) \quad \delta\boldsymbol{\psi}^n(k)]^T . \tag{7}$$

The system dynamic and observation equations in discrete time can be written by

$$\delta\mathbf{x}(k+1) = \mathbf{F}(k)\delta\mathbf{x}(k) + \mathbf{G}(k)\mathbf{w}(k) \tag{8}$$

$$\delta z(k) = \mathbf{H}(k)\delta\mathbf{x}(k) + \mathbf{v}(k) , \tag{9}$$

where the system transition matrix, the system input-noise matrix, and noise strength are given by (Kim, 2004)

$$\mathbf{F}(k)=\begin{bmatrix} \mathbf{I} & \Delta t\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \Delta t\mathbf{f}^n \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \mathbf{G}(k)=\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \sqrt{\Delta t}\mathbf{C}_b^n & \mathbf{0} \\ \mathbf{0} & -\sqrt{\Delta t}\mathbf{C}_b^n \end{bmatrix}$$

$$\mathbf{Q}(k)=\begin{bmatrix} \boldsymbol{\sigma}_{\delta f}^2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_{\delta\omega}^2 \end{bmatrix} , \tag{10}$$

with $\mathbf{C}_b^n$ being the direction cosine matrix formed from the quaternion, $\boldsymbol{\sigma}_{\delta f}$ and $\boldsymbol{\sigma}_{\delta\omega}$ representing noise strengths of acceleration and rotation rate respectively. The GPS observation is position and velocity, hence the observation matrix becomes

$$\mathbf{H}(k) = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix}, \mathbf{R}(k) = \begin{bmatrix} \boldsymbol{\sigma}_p^2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_v^2 \end{bmatrix}. \tag{11}$$

The filter input, $\delta\mathbf{z}(k)$ is formed by subtracting the position and velocity of the GNSS from the INS indicated position and velocity, and is then fed to the fusion filter to estimate the errors in the INS.

The high-rate inertial navigation loop provides a continuous navigation solution using the inertial measurements and the filter estimates the inertial errors whenever GPS information is available.

4 Physical System

4.1 UAV Platform

The ANSER (Autonomous Navigation and Sensing Experiment Research) project was conducted with multiple Brumby Mk III UAVs. Mk III offers a maximum speed of 180 km/hr, and roll rate of up to 300 °/sec. The airframe is modular in construction to enable the replacement or upgrade of each component. The Brumby series UAVs are compared in Figure 4 and their specifications are summarized in Table I. The Brumby UAV is a delta fixed-wing, pusher aircraft. The delta wing design requires no tail and is compact for a given wing area and has a minimal component count. The pusher design leaves a clear space in the nose for the radar and other sensors. The engine is located at the back and sensors in the front section of the aircraft stay free from exhaust contamination. The larger and transparent nose cone of the Mk III has a rear fairing to blend the nose cone to the fuselage.

Tab. 1 The Brumby performance characteristics.

Model	Engine	Max Speed	Payload	Endurance
Mk-I	74cc	180km/h	5kg	20min
Mk-II	80cc	180km/h	14kg	20min
Mk-III	150cc	180km/h	20kg	45min

Fig. 5 Brumby UAV series: Mk I (left), Mk II (centre), Mk III (right),

4.2 On-board Computing System

The hardware of the feature-tracking and navigation system are installed on the fuselage of the Brumby Mk III. The embedded PC104 platform is used as a flight control system performing navigation, guidance and control by fusing data from the IMU, GPS receivers and two tilt sensors. The IMU, from Inertial Science Inc., is very light and has a small form-factor, which makes it suitable for UAV applications. Two CMC Allstar GPS receivers are stacked on the flight control computer with the antennae installed on each of the wings. The vision camera is installed next to the IMU to minimise the lever-arm offset. The camera is connected to a secondary PC104 vision computer which performs the feature extraction and tracking tasks. Each computing node communicates by the Ethernet bus.

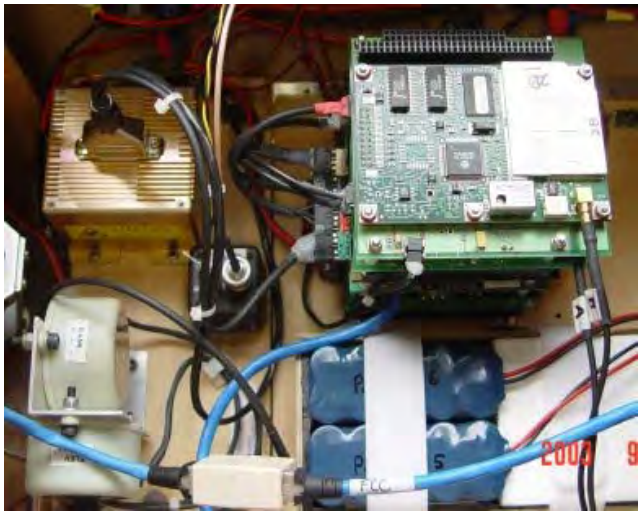


Fig. 6. Flight control system with IMU (top left), tilt sensor (bottom left), GPS (top right) and vision camera (centre).

5 Results

Intensive flight tests were performed to demonstrate the ANSER program at the test site. The results shown in this paper are from the real-time flight test on June 2002.

Figures 6 and 7 illustrate real-time feature observations, which are converted from range, bearing and elevation to Cartesian coordinates, then transformed to the navigation frame. The real-time GPS/INS navigation solution is used for the coordinate transformation. During level flight paths, it can be observed that the x and y position of the observed features are fairly close to the true feature positions. This is firstly due to the high accuracy of the bearing and elevation solution in the camera which was 0.16° and 0.12° respectively, and due to the consistent accuracy in the navigation solution. During banking however, large horizontal errors are introduced as can be seen clearly in Figure 8. This is due to the large range errors reflected in the horizontal plane. The range information extracted based on the size of the feature gives extremely poor quality information ranging from 20m to 100m. In addition, during the experiments the GPS satellite coverage was quite poor, with only 6 satellites in view. When the aircraft banked it often lost lock of some of these satellites, which degraded the height estimate, and subsequently the estimated height of the features. This highlights the importance of an accurate estimate of the platform state in feature-tracking as any error here will result in an error in the feature location.

During high banking lots of spurious observations, such as water reflection, are detected. These can be seen in the left corner of flight path and these cause the estimated ranges to be extremely noisy and unreliable at banking. Hence observations taken when the roll angle was greater than 30° are discarded in the tracking system. The results after this filtering still indicate that there are several areas, particularly on the lower left side of the plot, where clusters of observations are seen away from features. Rather than being spurious observations, these are actually natural features such as patches of sand (Nettleton, 2003), which appear in the images very similar to the artificial white features. They are detected consistently during every flight.

Figures 8 and 9 show the tracked feature positions within the vision node during the flight test. As the targets are known to be stationary, the tracking process model is tuned to decay velocity to zero in the filter prediction. Therefore, the errors in velocity states are essentially zero over the duration of the flight. The horizontal plot shows feature positions during the first three rounds. The estimated feature positions are close to the true positions, but some covariance ellipses fail to include the true position. The main reason is due to the poor range performance coupled with the INS error especially in roll angle. With successive observations of the feature, the effect of the range error can be reduced. The final tracking result after nine rounds shows that most of estimated feature positions are within the 2σ uncertainty boundary of the true position. These results show that the low-cost GPS/INS navigation system developed can be effectively used for the airborne feature-tracking purpose.

6 Conclusions

This paper presented real-time results of the airborne feature-tracking system on the UAV platform. The system incorporated a cost effective vision system and GPS/INS navigation system. The vision system provides bearing and elevation observation as well as range information based on known feature size information.

The feature position is computed using the GPS/INS solution and then it is used as the observation of the tracking filter. In spite of the large range error in the vision and the low cost sensors used, the tracked feature positions showed quite promising performance with several metres in accuracy. This also validates the accuracy of the on-board navigation system.

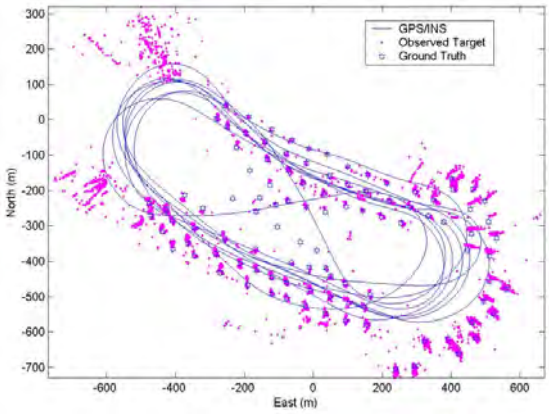


Fig. 7 Vision observations plotted in the navigation frame. Horizontal feature position shows good performance due to the accurate bearing and elevation observation.

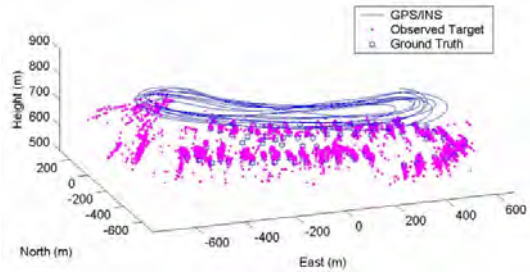


Fig. 8 3D-view of the vision observations in the navigation frame. The vertical position shows large errors due to the poor range accuracy in vision system.

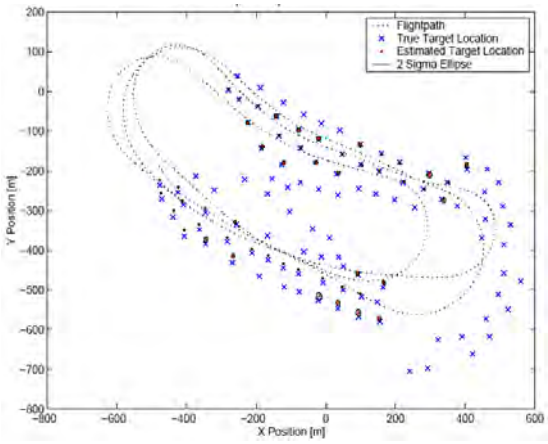


Fig. 9 Feature tracking result of final feature positions with 2σ uncertainty ellipsoids. Observations when the aircraft is banking at greater than 30deg are ignored in the tracking filter (plots from Nettleton, 2003).

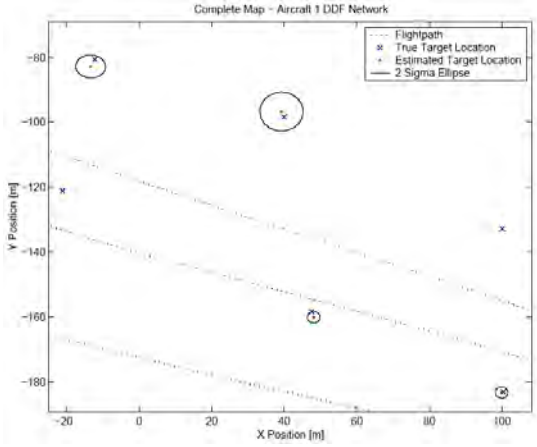


Fig. 9 Enhanced view of some features with 2σ uncertainty ellipsoids.

Acknowledgements

This work is supported in part by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government. The Autonomous Navigation and Sensing Experimental Research (ANSER) Project is funded by BAE SYSTEMS UK and BAE SYSTEMS Australia.

References

Blackman S.: (1986) ***Multiple-Target Tracking with Radar Application***. Artech House, Norwood.

Bar-Shalom Y; and Blair W. (2000): ***Multitarget-Multisensor Tracking: Applications and Advances Volume 3***, Artech House, Norwood.

Nettleton E. (2003): ***Decentralised Architectures for Tracking and Navigation with Multiple Flight Vehicles***, PhD thesis, Australian Centre for Field Robotics, The University of Sydney.

Kim J. (2004): ***Autonomous Navigation for Airborne Applications***, PhD thesis, Australian Centre for Field Robotics, The University of Sydney.

Ridley M.; Nettleton E.; Sukkarieh S.; Durrant-Whyte H. (2002): ***Tracking in Decentralised Air-Ground Sensing Networks***, In International Conference on Information Fusion, July, Annapolis, Maryland, USA

Stone D. L.; and Barlow C. A.; and Corwin T. L. (1999): ***Bayesian Multiple Target Tracking***, Artech House.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 173-182

Journal of Global Positioning Systems
2004
© CPGPS

Performance Evaluation of Multiple Reference Station GPS RTK for a Medium Scale Network

T.H. Diep Dao, Paul Alves and Gérard Lachapelle

Department of Geomatics Engineering, University of Calgary, Canada
Tel: (1-403) 210 9796 Fax: (1-403) 2841980 Email: dthdao@gmail.com

Received: 15 November 2004 / Accepted: 3 February 2005

Abstract. Carrier phase-based differential GPS is commonly used for high accuracy RTK positioning because it effectively reduces the effects of spatially corrected errors such as orbital and atmospheric errors. The spatially correlated error reduction is a function of the correlated errors measured by the two receivers. Carrier phase-based single reference station (SRS) positioning is capable of providing cm accuracy for static positioning and dm for kinematic positioning under normal atmospheric conditions when the inter-antenna distance is less than approximately ten kilometres. However, under highly localized atmospheric activity, and/or with a longer inter-antenna distance, the residual differential error increases and the accuracy degrades. The University of Calgary MultiRef™ multiple reference station (MRS) approach uses a network of GPS reference station to model the atmospheric conditions over a geographic region to reduce correlated measurement errors. This approach uses a conditional least-squares adjustment to predict the errors in the network area. This study focuses on an evaluation of the MultiRef™ approach relative to the single reference station (SRS) approach in the observation, position and ambiguity domains. Long-term and short-term convergence accuracy tests are used to assess the effectiveness of the approach. The network used for this assessment is located in Southern Alberta, Canada. This is a medium scale network with baseline lengths ranging from 30 to 60 km. The results show a minor to significant improvement of the MRS method in all domains.

Key words: Multiple reference station GPS RTK, single reference station GPS RTK, least squares collocation, performance evaluation

1 Introduction

Single reference station (SRS) differential GPS RTK performs well under normal atmospheric conditions when the inter-antenna distances are less than ten kilometres, providing centimetre level accuracy under ideal conditions. However, under high atmospheric conditions or with longer inter-antenna distances, the position solution accuracy is degraded because of the decrease in the spatial correlation of errors, namely ionospheric, tropospheric and satellite orbit errors. This has led to the development of multiple reference station (MRS) differential approaches, which attempt to model the spatial correlated errors over a regional network and interpolate the corrections to rover positions (e.g. Lachapelle & Alves 2002). This paper evaluates the performance of the University of Calgary correction-based least-squares collocation MRS approach, namely MultiRefTM, relative to the traditional SRS approach.

Fully evaluating the performance of a MRS approach is a difficult task due to the numerous parameters that affect performance, the most important ones being network configuration, atmospheric activities, and processing options, if one assumes the use of high performance receivers and unobstructed satellite availability. Network features, such as the number of reference stations and inter-receiver distances, directly affect the performance of the MRS approach (e.g. Alves et al 2003). If the network scale is too big, it can be difficult to resolve the network ambiguities over long baselines and, as a result, the corrections may be unreliable. It is also essential that the rover be located within the region of the network. Alves et al (2003) have shown that a network of four reference stations surrounding the rover is an effective configuration and the addition of extra reference stations does not generally further improve the performance of the MRS approach. Different levels of atmospheric errors result in different levels of improvement. During quiet atmospheric conditions, the errors are fairly constant over

the network area and the SRS approach can perform very well. Under these conditions, the MRS approach may not yield much improvement. However, under more active atmospheric conditions associated with highly localized atmospheric activities, the MRS approach is expected to offer significant improvements because the errors are better modelled using a network of reference stations. The use of different processing options such as L1, dual frequency wide lane (WL) and dual frequency ionospheric free (IF) observables will lead to different results because of the specific advantages and disadvantages of each individual combination.

The focus of the analysis herein is on the improvement of the MRS approach relative to the SRS approach in the measurement domain, the long-term position domain and the convergence performance under both quiet and fairly active ionospheric conditions. A specific network configuration is used, as described in the sequel. The next section summarizes the theory of the correction-based least-squares collocation MRS algorithm. The testing methodology is then presented, followed by the presentation and discussion of the results, and conclusions.

2 Correction-Based Least-Squares Collocation Algorithm

MRS algorithms are divided into two main categories, namely the correction-based and tightly coupled approaches. The correction-based approach uses observations obtained at the reference stations to estimate the spatially correlated network errors and then interpolates these "corrections" to the rover position. Numerous correction-based algorithms have been developed using different approaches to interpolate the corrections to the rover. These include the linear combination algorithm (Han & Rizos 1996), the linear interpolation algorithms (Gao et al 1997, Wanninger 1995), the Partial Derivative algorithm (Wübbena 1996) and the least-squares collocation algorithm (Raquet 1998). Dai et al (2004) compared their performance and concluded that they are more or less equal. The major advantage of the correction-based approach is that, once the corrections are generated and applied to the carrier phase observables, existing standard single reference station algorithms and software can be used to process the corrected carrier phase observables. If this constraint is removed however, a tightly coupled approach that uses all observations obtained both at the reference stations and at the rover in one filter can yield superior results (Alves et al 2004).

The correction-based least-squares collocation approach, used in MultiRef™ optimally estimates the network corrections, based on the known coordinates of the reference stations, assuming that the network ambiguities are resolved (Raquet 1998, Alves et al 2003). It then uses the covariance properties of the errors to predict the estimated reference station double differential errors to the location of the rover with the condition of minimizing the sum of the differential error variances (Raquet 1998). The correlated errors are spatially modelled over the network region while the uncorrelated errors are filtered out. A stochastic ionospheric modeling is used to estimate the dual frequency slant ionospheric delays. The correction vector $\delta\hat{\mathrm{l}}$ at each reference station, and that at the approximated rover position, denoted as $\delta\hat{\mathrm{l}}_{\mathrm{r}}$, are computed as (Raquet 1998)

$$\delta\hat{\mathrm{l}} = \mathrm{C_l B^T (BC_l B^T)^{-1} (Bl - \Delta\nabla N)} \qquad (1)$$

$$\delta\hat{\mathrm{l}}_{\mathrm{r}} = \mathrm{C_{l_r,l} B^T (BC_l B^T)^{-1} (Bl - \Delta\nabla N)} \qquad (2)$$

where $\mathrm{C_l}$ is the covariance matrix of the reference station observations, $\mathrm{C_{l_r,l}}$ is the covariance matrix between the rover observations and the reference station observations, $\mathrm{B}\,(=\frac{\partial\Delta\nabla \mathrm{l}}{\partial \mathrm{l}})$ is the double difference matrix, l is the vector of measurement-minus-true range observables computed from the true coordinates of the reference stations, λ is the carrier wavelength, and $\Delta\nabla\mathrm{N}$ is the vector of double difference ambiguities between reference stations.

The signal covariance function used to define the covariance matrices represents the stochastic behaviour of the correlated errors that affect the measurements. It therefore theoretically plays an important role on the effectiveness of the MRS approach. Ideally, the covariance function coefficients should be estimated adaptively using real-time data (Fortes 2002, Alves 2004). However, previous results have shown that the MultiRefTM corrections are not very sensitive to the covariance function itself under a medium level of ionospheric activity (Fortes 2002). The covariance function used herein is a function of satellite elevations and inter-reference receiver distances.

The corrections are applied to the observations of one reference station, which is called the primary reference station, in the form of single difference between the station and the rover. These corrected observations are then used at the rover in single reference station mode.

3 Testing Methodology

The MultiRefTM software was evaluated in post mission (but assuming real-time operation) using data collected with five stations of the Southern Alberta Network (SAN) located in Southern Alberta. The network configuration is shown in Figure 1. Stations AIRD, COCH, STRA and BLDM acted as reference stations. In the tests described herein, the UOFC station located in the middle of the network acted as rover. This is a medium scale network with baseline lengths ranging from 30 to 60 kilometres. The AIRD station, located 24 kilometres away from UOFC, was chosen as the primary reference station. Two dual-frequency 24-hour data sets with a 1 second data rate collected on 24 May 2004 and 6 April 2004 were used for testing. The ionosphere was normal and relatively active, respectively, during these two days, with a double difference effect of up to 5 ppm on the active day.

The processing includes two steps: The first step is running the MultiRefTM software using observations collected from the above four reference stations to generate network corrections. As the result, single difference corrections for observations obtained at AIRD are estimated using equations (1) and (2). In the second step, the corrected AIRD observations, along with the raw UOFC observations, are used in a single baseline processing to estimate the UOFC position solutions, the so-called MRS solutions. It is assumed that MultiRefTM requires a certain time, e.g. 2 hours in these tests, for initialization. The uncorrected AIRD observations are also used in parallel to obtain the traditional SRS position solutions for comparisons. An external commercial software package, GrafNav Version 7.01, developed by Waypoint Consulting Inc. is used for the single baseline processing to provide independence. The software is capable of epoch-by-epoch carrier phase based differential processing using single frequency L1 and dual frequency observations. Ionospheric-Free (IF) model is used with dual frequency. In order to obtain the IF ambiguities, the Wide Lane (WL) ambiguities are resolved first and followed by the L1 and then L2 ambiguities. The processing options were setup to attempt resolving ambiguities after 11.6 minutes using single frequency L1 observations and after 4.6 minutes using dual frequency observations. The ionosphere model for single frequency processing option, which is the satellite broadcast Klobuchar model, was not used. A 15 degrees elevation cut off was used.

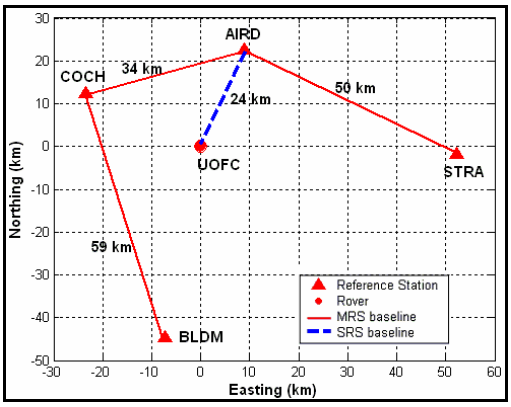


Figure 1. Minimal Southern Alberta Network (MSAN) Configuration

4 Results and Analysis

4.1 Double Difference Network Corrections

In order to obtain an approximate measure of the ionospheric activity in the region on these days, the local ionospheric K values are shown in Table 1 for three-hour intervals through out the 24-hour data sets. These values are calculated based on observations of the magnetic field fluctuations obtained at the MEA magnetometer station, which is located approximately 300 kilometres away from the UOFC station. Local K values theoretically range from 0 (quiet) to 9 (extreme). A more active ionosphere was observed during night-time, from 17:00 to 08:00 local time (00:00 - 15:00 UTC) than during daytime, from 08:00 to 17:00 local time (15:00 – 24:00 UTC) on both days. On 24 May 2004, the ionosphere was normal during night-time. Local K values of 3 to 4 were observed. On 6 April 2004, local K values of 5 to 6 were observed, indicating a more active ionosphere. A quiet ionospheric condition was experienced during daytime on both days with local K values of 2 to 3.

Double difference network corrections for different combined observables are shown in Figures 2 and 3 for 24 May 2004 and 6 April 2004, respectively. These are also the estimated double difference errors between AIRD and UOFC. The RMS values are shown in red separately for the active and quiet ionospheric periods of these two days. For 24 May 2004, the RMS double difference corrections on L1, L2 and WL observables during the active ionosphere period are 3.3 cm (1.4 ppm), 5.4 cm (2.2 ppm) and 4.5 cm (1.8 ppm), respectively. Over the baseline length of 24 km, these values are reasonable. The GF (Geometric-Free) corrections are 2.3 centimetres representing ionospheric errors of approximately 1 ppm. These corrections are smaller during the quiet ionosphere period but not significantly.

The IF (Ionospheric-Free) corrections are small and constant throughout the day, with a magnitude of 1 cm. These corrections are mainly due to the tropospheric residuals and noise.

Table 1: The local K values obtained at MEA geomagnetic station located in Edmonton (approximately 300 km away from UOFC station)

Hour of day (UTC time)	0 - 3	3 - 6	6 - 9	9 - 12	12 - 15	15 - 18	18 - 21	21 - 24
May 24, 2004	3	1	4	3	4	2	2	2
April 6, 2004	6	4	5	5	5	2	2	3

The network corrections are larger on 6 April 2004. This is expected due to the more active ionosphere as discussed. The estimated RMS value, for all errors on L1 observations, is 9.8 cm or 4 ppm over the 24 km baseline. The GF corrections are 6.6 cm, equivalent to 3 ppm over 24 km baseline. These are mainly caused by the ionosphere, which is fairly active. During daytime (quiet ionosphere), the corrections are again smaller than the ones obtained during night-time (more active ionosphere) but are larger than the ones obtained during daytime on 24 May 2004. The correction variation correlates well with the variation in the local K indices. The IF corrections, with a magnitude of approximately 1.1 cm for the entire day, show that the tropospheric residuals were consistent.

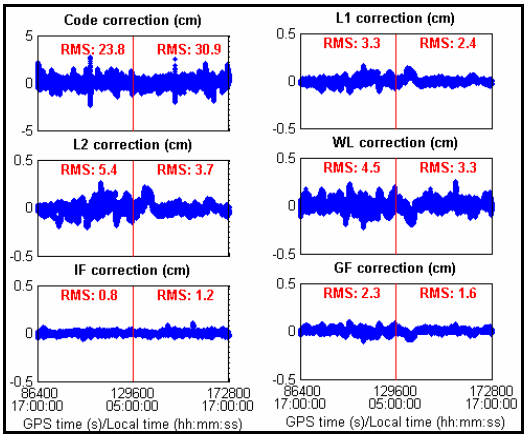


Figure 2. Double difference network corrections for all satellites pairs for 24 May 2004

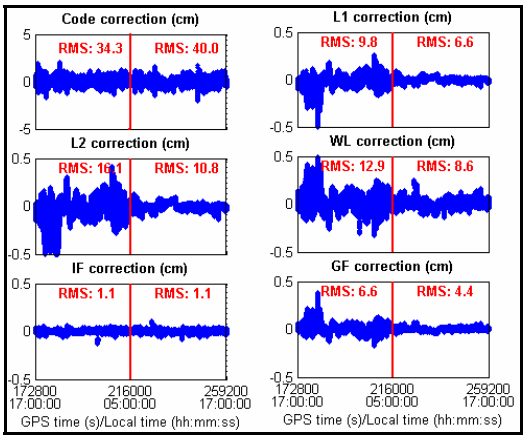


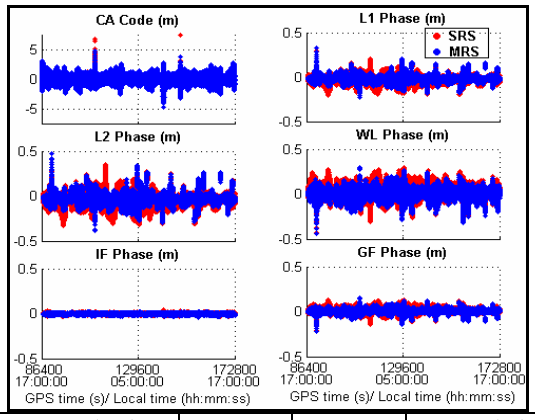
Figure 3. Double difference network corrections for all satellites pairs for 6 April 2004

The network ambiguities were resolved approximately 90 % of the time on both days. This suggests that the ionospheric model was effective in estimating the double difference slant ionospheric effects even under fairly active conditions. The network corrections can be considered reliable with such high percentage of fixed ambiguities.

4.2 MRS Improvement in Observation Domain

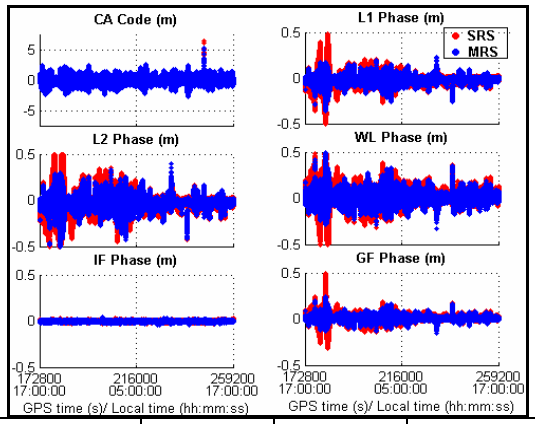
Figures 4 and 5 show the AIRD-UOCF double difference observable misclosures for 24 May 2004 and 6 April 2004, respectively. Both AIRD station and UOFC station were fixed to their true coordinates for this analysis. The misclosures calculated using the raw uncorrected AIRD observations (SRS) are shown in red while the ones using the corrected AIRD observations (MRS) are shown in blue. A table at the bottom of each figure gives a summary of the statistics.

The misclosures are generally larger for 4 April 2004. This is due to the higher ionospheric error experienced on that day. The MRS approach yields improvement relative to SRS approach with the use of L1, L2, WL and GF observables for both data sets. Higher level of improvement is observed under higher ionospheric condition. For example, MRS approach reduced the RMS of the misclosures by 45 % for 6 April (high ionosphere) compared to 35 % for 24 May 2004 (normal ionosphere). However, MRS does not yield any improvement for IF observables under either condition. This is because the largest error, ionosphere, is eliminated in this case while the double difference troposphere residuals and satellite orbit error are small over the distance of 24 km.



RMS (cm)	SRS	MRS	Impr. (%)
C/A Code	44.8	44.7	0
L1 Phase	3.5	2.4	31
L2 Phase	5.7	3.8	33
WL	4.5	2.9	36
IF	0.4	0.4	0
GF	2.2	1.4	36

Figure 4. Double difference residuals for all satellites pairs - 24 May 2004

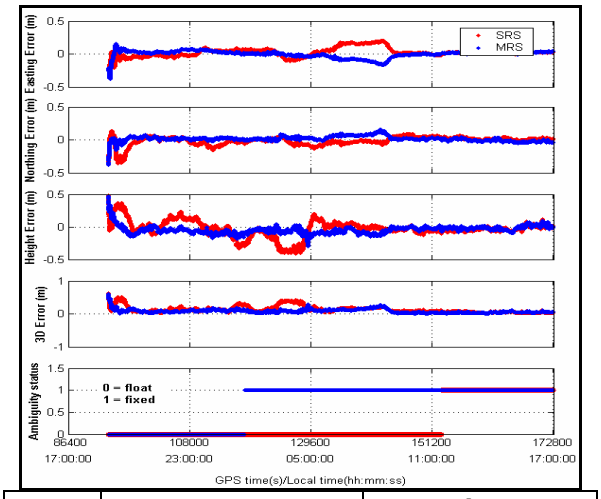


RMS (cm)	SRS	MRS	Impr. (%)
C/A Code	43.4	43.3	0
L1 Phase	4.2	2.5	40
L2 Phase	6.9	3.8	45
WL	5.5	3.0	45
IF	0.3	0.4	-33
GF	2.7	1.5	44

Figure 5. Double difference residuals for all satellites pairs - 6 April 2004

4.3 Long-term Position Domain Improvement with MRS

The single baseline AIRD-UOFC was processed to estimate epoch-by-epoch UOFC position solutions using the corrected and uncorrected AIRD observations. The UOFC position errors and the ambiguity status using L1 observations for 24 May 2004 are shown in Figure 6. The results with IF observables are shown in Figure 7. The SRS solutions are shown in red and the MRS solutions in blue. Statistics are given at the bottom of each figure for both converging and steady state position errors.



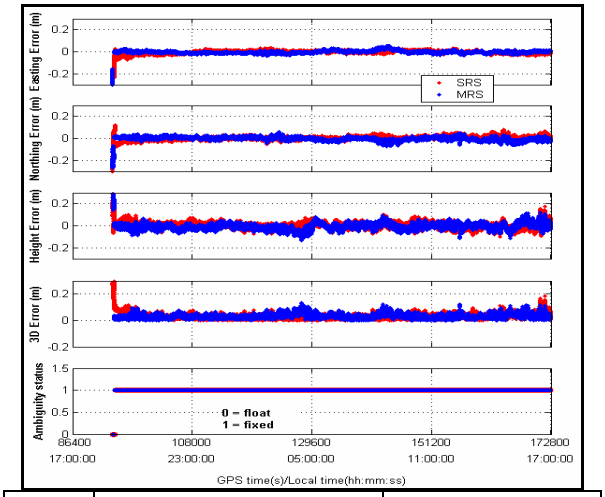
RMS (cm)	During Convergence Position – 2 hours				Steady State Position			
	E	N	H	3D	E	N	H	3 D
SRS	8	16	18	26	7	5	12	15
MRS	9	6	11	15	5	4	7	9
% Impr	-13	63	39	42	29	20	42	40

Figure 6. Position Accuracy in L1 mode after two hour network initialization - 24 May 2004

The L1 position solutions are affected by the ionospheric errors for both MRS and SRS cases. Compared to the SRS approach, the MRS approach yields a significant 11 cm improvement, equivalent to 42 %, in the 3D converging position accuracy during the first two hours. The improvement is 6 cm, equivalent to 40 %, in the 3D position accuracy after convergence. This shows that the MRS approach effectively reduces the differential ionospheric errors compared to the SRS approach under normal atmospheric conditions. As a result, the L1 ambiguities are resolved faster in the MRS case.

The ionospheric-free (IF) position solutions shown in Figure 7 are only affected by the tropospheric residuals, multipath and noise. The L1 and L2 ambiguities are therefore resolved very well in this case for both approaches. The MRS approach performs slightly worse by 2 to 3 cm during convergence. Both approaches offer a similar 3D position accuracy of 3 cm after convergence. This shows that, under quiet or normal atmospheric conditions, the MRS and SRS approaches yield more or less the same accuracy using IF observables, at least for this medium scale network. This is not surprising because

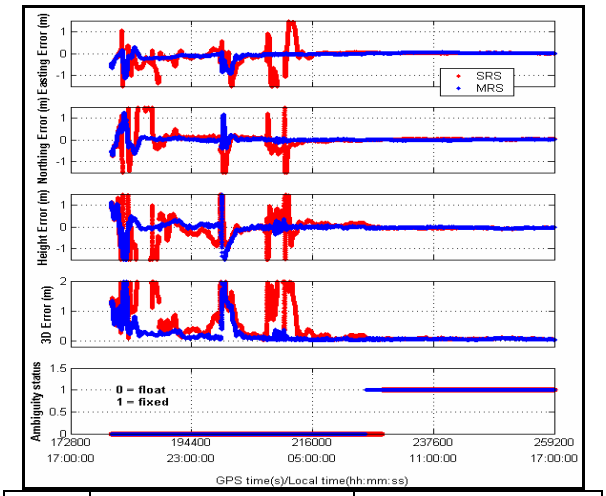
there is no ionospheric impact while the tropospheric error residuals are small in this case.



RMS (cm)	During Convergence Position – 1 hours				Steady State Position RMS (cm)			
	E	N	H	3D	E	N	H	3D
SRS	8	6	7	12	1	1	3	3
MRS	10	5	9	15	1	2	3	3
% Impr	-25	17	-29	-25	No significant difference			

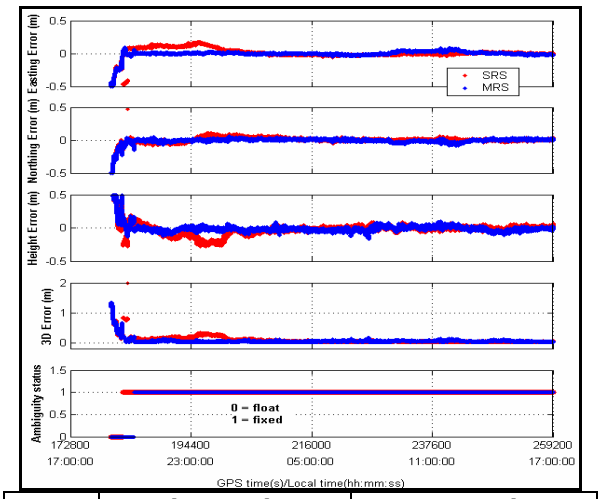
Figure 7. Position Accuracy in IF mode after 2 hour network initialization - 24 May 2004

Similar analysis was carried out for the 6 April 2004 data set. The UOFC position errors and the ambiguity status using L1 observations are shown in Figure 8. The statistics are provided for the first 15 hours of the day during which the ionosphere was active and for the last 7 hours of the day during which the ionosphere was quiet. The high level of ionospheric activity significantly degrades the L1 SRS position solution accuracy to 150 cm. The MRS approach yields significant improvements during this period, leading to a 3D position solution accuracy of 51 cm, an improvement of 66%. However these results still show that, during a relatively high level of ionospheric activity, the spatial decorrelation of the ionospheric effect is relatively rapid and the use of a medium scale multiple reference network can only improve the SRS method by a certain amount. The ambiguities cannot be resolved for either cases and the position solutions are estimated using the float ambiguities. During the last 7 hours of the day (quiet ionosphere), very accurate position solutions are obtained with both SRS and MRS approaches. The 3D position solution errors are less than 10 cm. The MRS approach shows a small improvement of 1cm centimetre during this period.



Posit-ion RMS (cm)	High Ionospheric Activity – First 15 hours				Low Ionospheric Activity – Last 7 hours			
	E	N	H	3D	E	N	H	3D
SRS	65	88	102	150	2	3	6	7
MRS	21	18	43	51	3	2	5	6
% Impr	68	80	58	66	-50	33	17	14

Figure 8. Position Accuracy in L1 mode after two hour network initialization - 6 April 2004



Positi-on RMS (cm)	High Ionospheric Activity –First 15 hours				Low Ionospheric Activity – Last 7 hours			
	E	N	H	3D	E	N	H	3D
SRS	11	11	13	20	1	1	2	3
MRS	7	7	11	15	3	2	3	5
% Impr	36	36	15	25	Not significant			

Figure 9. Position Accuracy in IF mode after 2 hour network initialization for 6 April 2004

Figure 9 show the UOFC position errors when using IF observables collected on 6 April 04. The ionospheric error is, again, eliminated in this case, resulting in a considerable improvement in position solution accuracy compared to the L1 solution. During the active ionospheric period, the SRS approach yields a 3D position accuracy of 20 cm. These statistics are calculated including the convergence period. The MRS approach yields an improvement of 15 % to 36 % for all easting, northing and height solutions, leading to a 5 cm (25 %) improvement in the 3D position solution accuracy. During the quiet ionospheric period, very accurate position solutions are obtained for both approaches and again, there is no significant difference between the two approaches.

4.4 MRS Improvement in Solution Convergence

Each of the 24 hour data set was divided into 24 1 hour segments, which were processed independently with GrafNav using L1 observations for both the MRS and the SRS approaches. MultiRefTM was setup to generate network corrections from the beginning to the end of each 24 hour data set without being reset. The 3D UOFC position errors for the 24 segments of 24 May 2004 are presented in Figure 10. The SRS solutions are in red and the MRS solutions in blue. For most of the segments, the MRS approach convergences faster and provides more accurate converging position solution accuracy. The improvement is noticeable during the periods of high ionospheric activity. Under quiet ionospheric conditions, there is an inconsistency in MRS improvement. For example, during the periods of 11:00-12:00 and 12:00-13:00 (Hours 19 and 20), the MRS approach performs worse than the SRS approach. Conversely, MRS yields better results during periods of 07:00-08:00 and 08:00-09:00 (Hours 15 and 16). However, the differences are not very significant. Moreover, the low level of ionospheric errors makes the ambiguities easier to resolve when using the SRS approach.

Figure 11 shows the averaged 3D position errors and averaged position component errors for the 24 segments. The MRS approach offers a significant improvement of 50 % on average for the L1 position solution accuracy during convergence. Figure 12 shows the maximum 3D position errors and the maximum errors in the three position components of the 24 segments, respectively. Improvement is observed in reducing the maximum position solution errors during the convergence time when the MRS approach is used.

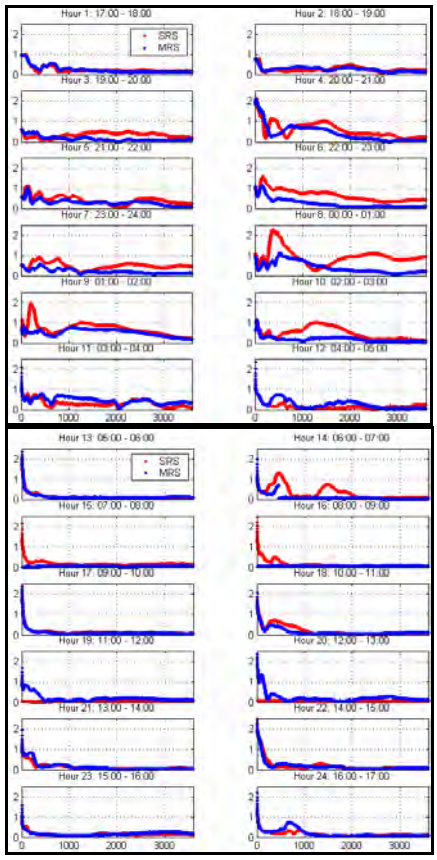


Figure 10. Convergence Analysis: 3D position errors in L1 mode for 24 1-hour data segments - May 24, 2004

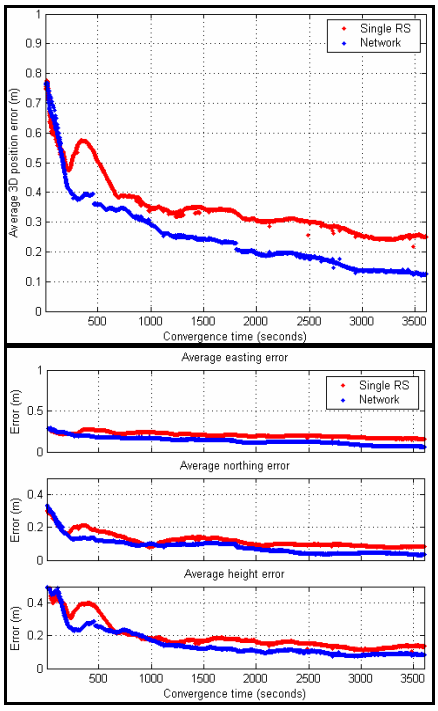


Figure 11. Convergence Analysis: Average 3D position component errors in L1 mode of 24 1-hour data segments - May 24, 2004

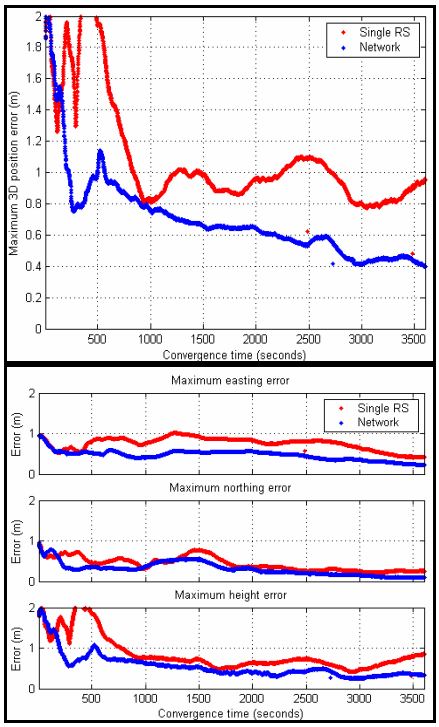


Figure 12. Convergence Analysis: Maximum 3D and position component errors in L1 mode for 24 1-hour data periods - May 24, 2004

A similar convergence analysis was carried out using the data collected on 6 April 2004. The results are presented in Figures 13 to 15 in a similar way to the ones for 24 May 2004. Compared to 24 May 2004, the ionosphere was more active for the first 15 segments. As a result, a very poor performance in convergence is obtained for the SRS approach. The MRS approach yields much faster convergence during these segments. Under the quiet ionospheric period, the MRS still results in better convergence for most of the segments in this case. On average, approximately 60 % to 70 % improvement is shown with the use of MRS approach. For most of the segments, the ambiguities cannot be resolved after one hour. Comparisons between the SRS and MRS solutions in maximum position solution errors during convergence are presented in Figure 15. The SRS approach suffers a maximum error of up to several metres occurred during period 19:00-20:00, due to the ionospheric error. The peak is removed when the MRS approach is used, making the latter more reliable.

Corrections generated during network initialization were used in the first segment for each data set. Interestingly, there is not much impact observed in both cases and the MRS approach yields improvement from the early beginning of the two first segments. Based on these limited results, it would appear that background network ambiguity convergence is not a major issue.

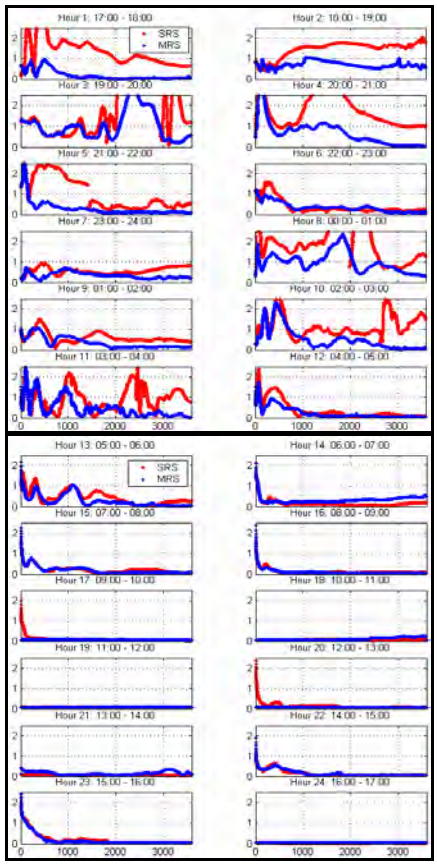


Figure 13. Convergence Analysis: 3D position error in L1 mode for 24 1-hour data periods -April 6, 2004

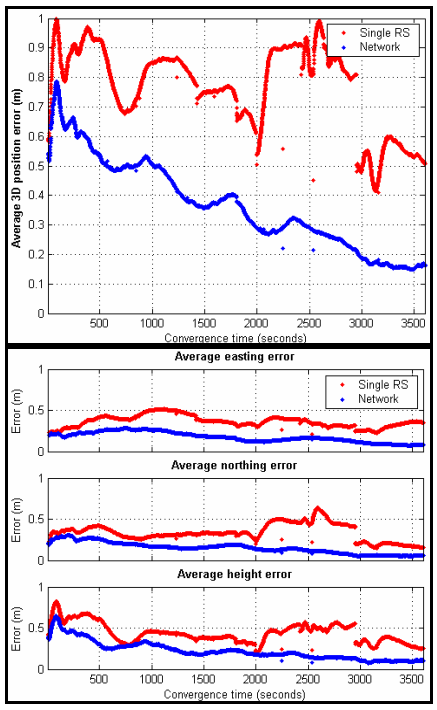


Figure 14. Convergence Analysis: Average 3D and position component errors in L1 mode of 24 1 hour data periods - April 6, 2004

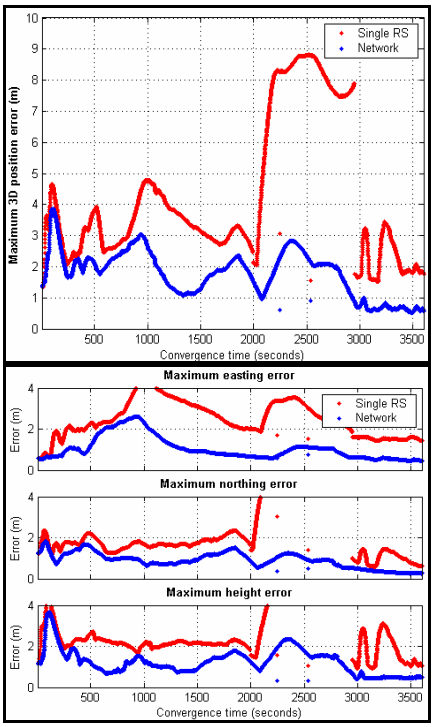


Figure 15. Convergence Analysis: Maximum 3D and position component errors in L1 mode for 24 1-hour data periods - April 6, 2004

The data sets used here have also been used by Alves and Lachapelle (2004) to compare the performance of three differential RTK approaches, namely SRS, MRS correction-based least-squares collocation and MRS tightly coupled approaches. The entire processing of the network and rover data was carried out using MultiRef™. Under low ionospheric conditions (24 May 2004) the position solution accuracy obtained using the MRS correction-based approach is the same as reported here. Under higher ionospheric conditions, the position accuracy is better by only a few centimetres. This shows full compatibility between the two evaluation approaches.

5 Conclusions

The MultiRefTM correction-based approach yields minor to significants improvements relative to the traditional single reference station approach in the observation, position and ambiguity resolution domains when using a medium scale network under the atmospheric conditions reported for these data sets. In the position domain, the MRS approach not only offers more accurate position solutions after convergence but also faster and more accurate solutions during convergence. The level of improvement, however, depends very much on the magnitude of the atmospheric errors and on the selection of observables. The MRS approach yields the most improvement under active ionospheric conditions using L1 observables due to its capability to model the spatially correlated errors. However, this advantage is reduced in this case when ionospheric-free observables are used or when the atmospheric errors are constant over the region. It remains to be seen as to whether the MRS method would yield significantly better results for a larger scale network when using ionospheric-free observables. More studies on the impact of using different network configurations and of the network initialization will be carried out in the near future.

References

Alves P, Ahn YW and Lachapelle G (2003): ***The Effects of Network Geometry on Network RTK Using Simulated GPS Data***, CD-ROM Proceedings of ION GPS/GNSS 03, Institute of Navigation, Portland, Oregon, USA, 11 pages.

Alves P, and Lachapelle G (2004): ***A Comparison of Single Reference Station, Correction-Based Multiple Reference Station and Tightly Coupled Methods using Stochastic Ionospheric Modelling***, Proceedings of the International Symposium on GPS/GNSS, Sydney, Australia (6-8 December).

Alves P (2004): ***Development of two novel carrier phase-based methods for multiple Reference station positioning***, PhD Thesis, published as Report No. 20203, Department of Geomatics Engineering, The University of Calgary.

Alves P, Lachapelle G and Cannon ME (2004): ***In-Receiver Multiple Reference Station RTK Solution***, CD-ROM Proceedings of ION GPS/GNSS 04, Institute of Navigation, Long Beach, California, USA, 9 pages.

Dai L, Wang J., Han S. and Rizos C. (2001): ***A Study on GPS/GLONASS Multiple Reference Station Techniques for Precise Real-Time Carrier Phase-Based Positioning***, Proceedings of ION GPS/GNSS 01, Institute of Navigation, Salt Lake City, UT, USA, pp 392-403.

Dai L, Han S, Wang J and Rizos C (2004): ***Comparison of Interpolation Algorithms in Network-Based GPS Techniques***, Navigation, 50(4), 277 – 293.

Fortes LPS (2002): ***Optimising the Use of GPS Multi-Reference Stations for Kinematic Positioning***, PhD Thesis, UCGE Reports Number 20158, Department of Geomatics Engineering, University of Calgary, Canada.

Gao Y, Li Z and McLellan J (1997): ***Carrier phase based regional area differential GPS for decimeter-level positioning and navigation***, Proceedings of ION GPS-97, Institute of Navigation, Kansas City, Missouri, USA, pp 1305-1313.

Han S and Rizos C (1996): ***GPS Network Design and Error Mitigation for Real-Time Continuous Array Monitoring Systems***, Proceedings of ION GPS-96, Institute of Navigation, Kansas City, Missouri, USA, pp 1827-1836.

Lachapelle G and Alves P (2002): ***Multiple Reference Station Approach: Overview and Current Research***, Invited Contribution, Expert Forum. Journal of Global positioning Systems, 1(2), 133-136.

Raquet J (1998: ***Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers,*** PhD Thesis, UCGE Report Number 20116, Department of Geomatics Engineering, University of Calgary, Canada.

Waypoint Consulting Inc. (2004): ***GrafNav/GrafNet, GrafNav Lite, GrafMov, Inertial Explorer for Window 95TM, 98TM, 2000TM, XPTM & NTTM Operating Manual***,Version 7.0. Canada

Wanninger L (1995): ***Improved Ambiguity Resolution by Regional Di®erential Modeling of the Ionosphere***, Proceedings of ION GPS-95, Institute of Navigation, Palm Springs, CA, USA, pp 55-62.

Wübbena C, Bagge A, Seeber G, Volker B, and Hankemmeier P, (1996): ***Reducing Distance Dependent Errors for Real -time Precise DGPS Applications by Establishing Reference Station Netwroks***, Proceedings of ION GPS-96, Institute of Navigation, Kansas City, Missouri, USA, pp 1845-1852.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 183-190

Journal of Global Positioning Systems
2004

The Application of Integrated GPS and Dead Reckoning Positioning in Automotive Intelligent Navigation System

Qingquan Li

Wuhan University Luoyu Road 129, Wuhan,Hubei, P.R.China
e-mail: qqli@whu.edu.cn Tel: + 86(27)6875651; Fax: +86(27)68756661

Zhixiang Fang

Wuhan University Luoyu Road 129, Wuhan,Hubei, P.R.China
e-mail: fang_zhi_xiang@163.com Tel: + 86(27)68778222; Fax: +86(27)68778222

Hanwu Li

Wuhan University Luoyu Road 129, Wuhan,Hubei, P.R.China
e-mail: lihw@whu.edu.cn Tel: + 86(27)68771974; Fax: +86(27)68778222

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. The applications of Global Positioning System (GPS) are increasingly widespread in China. GPS positioning is more and more popular. Especially, the automotive navigation system which relies on GPS and Dead Reckoning technology is developing quickly because of the anticipated huge future market in China. In this paper a practical combined positioning model of GPS/DR is put forward. This model designed for automotive navigation system makes use of a Kalman filter to improve position and map matching veracity by means of filtering the raw GPS and DR signals. In practical examples, the validity of the model is illustrated. Several experiments and their results of integrated GPS/DR positioning in automotive navigation system will show that a Kalman Filter based on integrated GPS/DR position is necessary, feasible and efficient for automotive navigation application. Certainly, this combined positioning model, similar to other models, can not resolve all situation issues. In the paper, the applicable principles of the model are given and the advantages and disadvantages of this model are compared with other positioning models. Finally, suggestions are given for further improving integrated GPS/DR system performance, and the application aspects of integrated GPS/DR technology in the automotive navigation system are summarized.

Key words: GPS; Dead Reckoning; Kalman Filter; Automotive Navigation System

1 Introduction

With the increasing popularity of automotive consumption in China, the intelligent automotive navigation system is more and more popular in daily life and provides convenient guidance for driving in large cities. Generally, the system relies on GPS in positioning. The real-time positioning information which is provided by GPS is very important for the phonetic broadcasting and map display of the automotive navigation system. However, in automotive navigation there are a lot of factors which will restrict and influence the positioning results which solely depend on GPS:

(1). Because of edifices and skyscrapers, built-up streets in the city, the GPS signals are so weak that the positioning errors are too large to be tolerant.

(2). The multi-path effects of GPS signals are very serious in the city, which cause the computative error of GPS positioning to be extremely large.

(3). At the crossroads, trestles, tunnels and underpasses in the city, the GPS signals are too weak to be tracked.

The above factors limit the continuity of navigation information provided by the vehicle based GPS system, which add difficulties to real-time vehicle navigation. In order to resolve this problem, many researchers have put forward various solutions. For example, Fang et al.

(1996, 1998, 1999) and Wang (1997) put forward the Kalman Filter model for GPS/odometer/magnetic compass/DR integration. Wu (1997, 1998), Zheng et al. (1999) and Xiong et al. (1997) have brought forward the Kalman Filter model for GPS/odometer/Gyro/DR integration, while Liu et al. (1995) and Chen et al.(1997) have proposed the Kalman Filter model for GPS/INS combination. Borenstein et al. (1994) have done the similar researches. Commonly they placed emphasis on the requirements of real-time calculation and the choices of combined coefficients and dimensions at the stage of model design. In the process of developing automotive navigation products, an integrated positioning model is developed to improve position accuracy.

2 The Integrated GPS/DR/MM Positioning Model

In this paper an integrated GPS/DR/MM positioning model is put forward. This model is used to overcome the limitation that the GPS positioning is null or weak in the navigation system, when the vehicle travels through the downtown, tunnel, underpass and crossroad etc. So the navigation system model will provide the most reasonable intelligent navigation services for users

2.1The Integrated GPS/DR Positioning Model

The basic principles of the integrated GPS/DR/MM positioning system are that the integrated GPS and DR positioning system provides basic positioning information, which will be corrected with the assistance of map matching to improve the accuracy and reliability of integrated positioning as much as possible. In many literatures the positioning model has been discussed (e.g., Fang, 1998; Fang et al., 1999; Wu et al., 1999; Fang et al., 1999; Wu, 1999; Xiong, 1997; Liu, 1995; Chen, 1997). The integrated GPS/DR positioning model is discussed and used more (Fang, 1998; Fang et al., 1999; Wu et al., 1999; Fang et al., 1999].

The states of the integrated GPS/DR navigation System are given as (Fang, 1998).

$$X = [e, V_e, a_e, n, V_n, a_n, \varepsilon, \psi]^T \qquad (1)$$

Here: e and n are the eastward and northward location components of vehicles, V_e and V_n are the eastward and northward components of velocity, a_e and a_n are the eastward and northward components of acceleration respectively, ε is the drift errors of velocity gyro and ψ is the calibrated coefficient of the odometer.

According to the definition of the state variable, the initiative variable can be defined as follows:

$$\overset{\circ}{X} = [V_e, a_e, \overset{\circ}{a}_e, V_n, a_n, \overset{\circ}{a}_n \ \overset{\circ}{\varepsilon}, \overset{\circ}{\psi}]^T \qquad (2)$$

Then we can process the data with the following statistical equations:

$$\overset{\circ}{X} = \mathrm{FX} + \mathrm{U} + \mathrm{W}(t) \qquad (3)$$

Here: the formula can be showed as:

$$\begin{pmatrix} V_e \\ a_e \\ \overset{\circ}{a}_e \\ V_n \\ a_n \\ \overset{\circ}{a}_n \\ \overset{\circ}{\varepsilon} \\ \overset{\circ}{\psi} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau_{ae}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{an}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{\varepsilon}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} e \\ V_e \\ a_e \\ n \\ V_n \\ a_n \\ \varepsilon \\ \psi \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{\overline{a}_e}{\tau_{ae}} \\ 0 \\ 0 \\ \frac{\overline{a}_n}{\tau_{an}} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ w_{ae} \\ 0 \\ 0 \\ w_{an} \\ w_{\varepsilon} \\ w_{\psi} \end{pmatrix} \qquad (4)$$

w_{ae} , w_{an} , w_{ε} and w_{ψ} are the white noises of $\left(o, \sigma_{ae}^2\right), \left(o, \sigma_{an}^2\right), \left(o, \sigma_{\varepsilon}^2\right), \left(o, \sigma_{\psi}^2\right)$ respectively; τ_{ae} and τ_{an} are the correlated time constants of vehicle acceleration change rates; $\overline{a}_e$, $\overline{a}_n$ are the average values of eastward and northward acceleration respectively, τ_{ε} is the equivalent time constant in first order Markov process of the velocity gyro drift.

2.2The Integrated GPS/DR/MM Positioning Model

The basic idea of the model is to get preliminary position information according to the GPS/DR positioning model, then extract possible information of all road sections from an existing digital map, including road section direction sets, distance sets, road section knots and lines sets, etc, and project the preliminary results onto the most probable line. Afterwards, the rationality of the results is evaluated according to certain evaluation function and corrected statistical equation. The evaluation result is likely to be:

$$\partial i = \begin{cases} -1 : Nomal \\ 0 : NULL \\ 1 : Exact \\ 2 : Depature \\ 3 : OnRoadWay \end{cases} \qquad \{\partial i \mid i = 0, -1, 1, 2, 3\} \qquad (5)$$

Here, -1-- maintain the original state, 0--null, 1--correct map matching, 2--departure from the navigation path, 3--running normally on the right navigation path.

Basic information for map matching, such as current road section location and orientation, the matching precision in the normal distribution and etc, are supplemented into the above statistical equation for forming Kalman Filter Equation. A maximum likelihood estimate of this equation is used as the positioning value so that the positioning rationality is improved. In the following, we lay emphasis on the formation of the model.

2.3 Model Formation

According to the conception of the integrated GPS/DR/MM navigation system, the observations of the integrated vehicle GPS/DR/MM navigation system should include:(1) the GPS positioning location (λ, l),(2) the change rate of vehicle course and angular velocity ω (3) the distance S of the vehicle odometer within the adopted period T, (4) the distance (Ds) on the road section that vehicle runs, the orientation (θ) of current road section, and the last matching result components D_X , D_Y and confidence k (%).

Therefore, the observation component of the system can be defined as:

$$\lambda = e + v_1 \quad (6)$$

$$l = n + v_2 \quad (7)$$

$$\omega = \frac{\partial}{\partial t}\left[tg^{-1}\left(\frac{v_e}{v_n}\right)\right] + \varepsilon + \varepsilon_\omega = \frac{v_n a_e - v_e a_n}{v_e^2 + v_n^2} + \varepsilon + \varepsilon_\omega \quad (8)$$

$$s = \psi T\sqrt{v_e^2 + v_n^2} + \varepsilon_s \quad (9)$$

$$Ds = Ds^0 + \mathrm{S} \quad (10)$$

$$D = D + S * \cos(\theta) \quad (11)$$

$$D_y = D_y{}^0 + S * \sin(\theta) \quad (12)$$

$$R_k = \max(K_{rdi} \mid 0 < i < n) = \min(D^k{}_{rdi} * P) \quad (13)$$

R_k refers to the matching road result with the confidence K, K_{rdi} is the reliable results of the i-th alternative road near the location point, D_{rdi} is the distance from preliminary position to the i-th road (point L(D_X , D_Y)) and P is the possible area to the i-th road.

The determination rules for value P are shown as:

P = 0: when the road is inaccessible to the vehicle, P is set as 0

P = 1/n: when vehicle is on the road but the distance between the vehicle and the exit of the road is in the tolerance range, P is set as the probability of subsequent road.

P = 1: when the vehicle is on the road for sure

Here we should consider the following factors when determining the positioning weight P:

(1) Use the direction of the road as the initial direction value of the computation;

(2) Use the first matching point on the road as the initial point of reckoning location to calculate the distance running on the road;

(3) Estimate the varieties of road direction and gyro direction to determine the scope of angle change ahead, which will be easy to pre-process the signal of gyro;

(4) P is set as 1 when we can be sure that the vehicle has entered a road section; and when the distance between the vehicle and the end of the road is in tolerance range (for example 50 meters), it is necessary to extract all the alternative roads in order to judge which road the vehicle will run on. P is set as the reciprocal of the number of the alternative roads;

(5) The weights of all neighbouring roads (main stem, auxiliary road, circle etc) can be set 0 when we are sure that the vehicle has entered a road section.

So the error equation can be written as:

$$Z(t) = \begin{pmatrix} \lambda \\ l \\ \omega \\ s \\ Ds \\ D_x \\ D_y \\ R_k \end{pmatrix} = \begin{pmatrix} e \\ n \\ \frac{v_n a_e - v_e a_n}{v_e^2 + v_n^2} + \varepsilon \\ \psi T\sqrt{v_e^2 + v_n^2} \\ 1 \\ S * \cos(\theta) \\ S * \sin(\theta) \\ min(D^k{}_{rdi} * P) \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \varepsilon_\omega \\ \varepsilon_s \\ S \\ D_X{}^0 \\ D_y{}^0 \\ \partial i \end{pmatrix} = h[t, X(t)] + V(t) \quad (14)$$

v_1 and v_2 are the noises of GPS positioning observations measurement, $(o, \sigma_1^2), (o, \sigma_2^2)$ are Gaussian white noises, ε is the first order Markov process component in velocity gyro drift which is the component of Gaussian white noises ($0, \sigma_\omega^2$) ; ε_s is the observation noise of the odometer output, it is $(0, \sigma_s^2)$ Gaussian white noise.

Therefore we can establish the model step by step according to deduction of a Kalman Filter:

$$\hat{X}(k/(k-1)) = \varphi_1(k/(k-1)) \quad (15)$$

$$\hat{X}(k) = \hat{X}(k/(k-1)) + K(k)[Z(k) - h[k, \hat{X}(k/(k-1))]] \quad (16)$$

$$P(K/K-1) = \varphi(K/(K-1))P(K-1)\varphi^T(K/(K-1)) \quad (17)$$

$$K(k) = P(k/(K-1))H^T(K)[H(K)P(k/(K-1))H^T(K) + R(K)]^{-1} \quad (18)$$

$$P(K) = [I - K(k)H(k)]P(k/(K-1)) \quad (19)$$

$$H(K) = \frac{\partial h[K, X(K(K-1))]}{\partial X^T(K(K-1))} \quad (20)$$

3 Combined Positioning Algorithm

From the above analysis, an extended combined positioning algorithm is put forward in this section. In this algorithm, different weights are assigned to different factors: GPS subsystem is corresponding to coefficient of information β_1, DR sub system is coefficient β_2, and map information is coefficient β_3 ($\beta_1+\beta_2+\beta_3=1$) . So the combined positioning Kalman Filter can be illustrated as Fig.1:

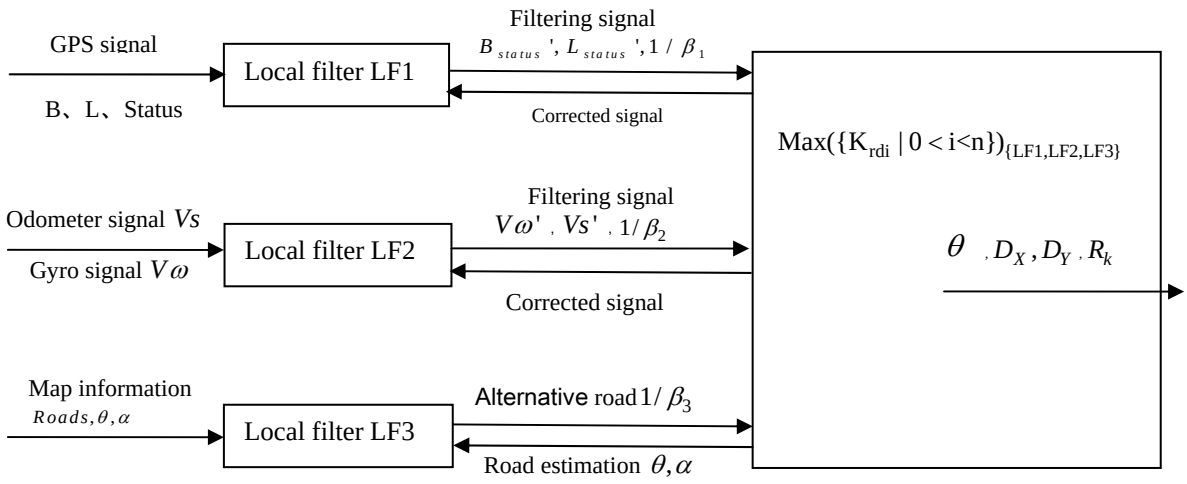


Fig 1. The Federated Kalman Filter structure of the integrated GPS/DR/MM navigation system

Local filter 1 (LF1) in Fig.1 is a standard Kalman Filter. Its dynamic equation is:

$$\begin{aligned} X_1(k+1) &= \phi(k)X_1(k) + U1(k) + W1(k) \\ Z_1(k) &= H_1(k)X_1(k) + V_1(k) \end{aligned} \quad (21)$$

Local filter 2 (LF2) in Fig.1 is the corresponding nonlinear Kalman Filter of DR system. Its dynamic equation is:

$$\begin{aligned} X_2(k+1) &= \phi(k)X_2(k) + U2(k) + W2(k) \\ Z_2(k) &= h[k, X_2(k)] + V_2(k) \end{aligned} \quad (22)$$

In the same way, we can get local filter 3(LF3)

$$\begin{aligned} X_3(k+1) &= \phi(k)X_3(k) + U3(k) + W3(k) \\ Z_3(k) &= h[k, X_3(k)] + V_3(k) \end{aligned} \quad (23)$$

So the whole filter algorithm is:

$$X_i(k) = P_i(k)[H_i^T(k)R_i^{-1}(k)Z_i(k) + P_i^{-1}(k/k-1)X_i(k/k-1)$$

$$P_i^{-1}(k) = P_i^{-1}(k/k-1) + H_i^T(k)R_i^{-1}(k)H_i(k)$$

$$X_i(k/k-1) = \phi_i(k-1)X_i(k-1)$$

$$P_i(k/k-1) = \phi(k-1)P_i(k-1)\phi^T(k-1) + Q_i(k-1) \quad , \quad i \in [1,3]$$

And the synthesized optimization of the whole system state can be defined as:

$$X\ (k) = P\ (k)[\beta_1 P_1^{-1}(k)X_1(k) + \beta_2 P_2^{-1}(k)X_2(k) + \beta_3 P_3^{-1}(k)X_3(k)] \quad (24)$$

$$P^{-1}(k) = \beta_1 P_1^{-1}(k) + \beta_2 P_2^{-1}(k) + \beta_3 P_3^{-1}(k) \quad (25)$$

$$Q^{-1}(k) = \beta_1 Q_1^{-1}(k) + \beta_2 Q_2^{-1}(k) + \beta_3 Q_3^{-1}(k) \quad (26)$$

4 Combined Position Experiment

A close region in a city of China is chosen as the testing environment, see Fig.2. The corresponding hardware configuration is shown in table 1:

Figure 2. Testing Route

Tab. 1 List of hardware collocations for test.

Item	Configuration
main frequency	166MHz
chip	SH4
memory	64M
CF card	128M
GPS	Gamin 15
GYRO	Matsushita
Vehicle	Buick commercial

Tab. 2 The Parameters of Positioning Equipments

item	Parameter
GPS	
Positioning precision	<50m Normally
Startup time	20s
GYRO	
Sensitivity	[22.5,27.5]mV/($^\circ .s^{-1}$)
Sensitivity drift	[-5.0% ,5.0%]
Zero Point Drift	[-10% ,10%]
Linearity	[-5.0% ,5.0%]
Working temperature	[-40,85]$^\circ C$

Parameters related to positioning equipment are shown in Table 2. The positioning state of GPS receiver is checked here. Fig. 3 shows the drift track of GPS positioning at a given position over a long period of time; Fig. 4 is the satellite distribution. In Fig. 3, we can see that the positioning state of GPS is not satisfactory, and there are noticeable drifts. The satellite visibility of the GPS module is satisfactory. In most cases, it allows for simultaneous visibility of 4 satellites with distinct signals. Thereby after filtering processing, GPS Observation can meet the requirements for dynamic vehicle navigation.

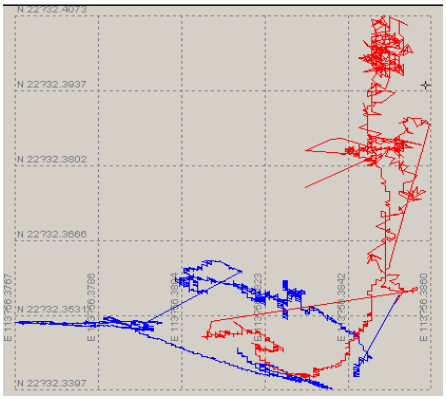


Fig 3. GPS Positioning Drift

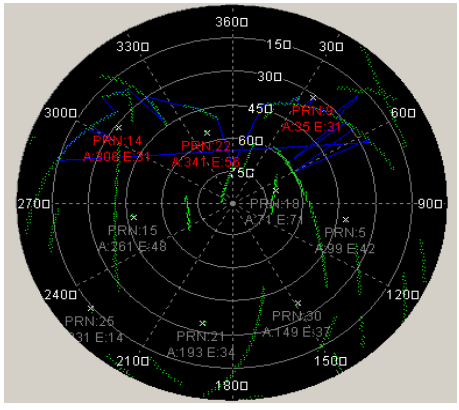


Fig. 4 Satellite Distribution

Figs.5-8 illustrate the deviation between recorded track and road feature points in different cases. Whether based on standalone GPS positioning or the integrated GPS/DR/MM positioning, accuracy of positioning in X-axis or Y-axis direction is not significant. But the deviations of positioning based on the two different methods are distinct. The positioning error of points can be greatly improved after GPS/DR/MM matching (normally the method can make the accuracy increase by 30%). In rare cases there will be large matching errors. This is mainly because, in the crossroads, there are several choices, and various values can be obtained with the map matching algorithm. According to the optimization rules, the most possible road is chosen based only on the signal value matching.

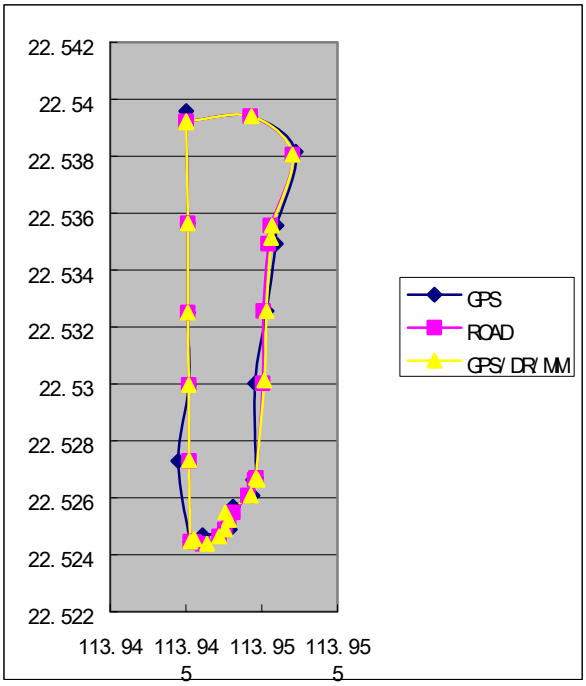


Fig 5. Trajectory

The navigation test program of integrated GPS/DR/MM positioning model is realized in this paper based on

WinCE operation systems, and in the experiment the vehicle moving tracks in various cases are recorded. It is obvious that in comparatively straight places whether single GPS positioning or integrated GPS/DR/MM positioning is adopted we can always get well matched results. In the curved road section, however, there will be distinct deviations if we use standalone GPS positioning model which can be greatly improved with integrated GPS/DR/MM model.

5 Conclusions

To solve the positioning problem in the vehicle navigation, a universal model of integrated GPS/DR/MM positioning is discussed in this paper. In the cities, it is important to improve GPS positioning precision by the model in the cases when GPS positioning signals are too weak, such as in the downtown areas, tunnels, underpasses, crossroads, etc. Obviously, forming the model will take up computing resources of the system more or less and the interaction with navigation digital map is frequent. The setting of the parameters of this model depends on experience to some extent, so appropriate parameters can be determined only after many trials. These parameters are quite different in digital map databases with different precisions. In the future study, we will consider the design and application of an adaptive positioning model

Fig 6. Deviation in X-axis

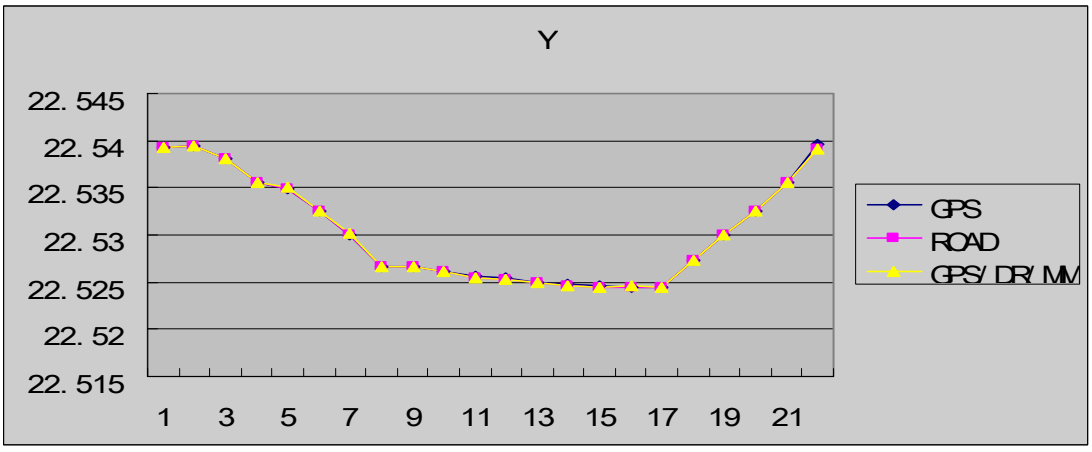


Fig 7. Deviation in Y-axis

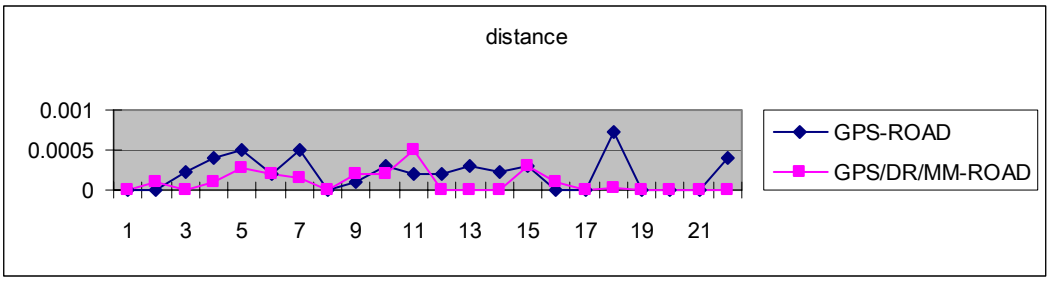


Fig 8. Deviation of positioning

References

Leonhardi, A., Nicu, C., Rothermel, K., (2001).***A Map-based Dead-reckoning Protocol for Updating Location Information***. Technical Report.2001/09. Institute of Parallel and Distributed High-Performance Systems (IPVR),Department of Computer Science,University of Stuttgart, Breitwiesenstr. 20–22, D-70565, Stuttgart, Germany

Fang, J., Shen, G., Wan, D., (1998). ***Establishment of an Adaptive Extended Kalman Filter Model for a GPS/DR Integrated Navigation System***.Control Theory and Applications, Vol.15, No.3, Jun.. 385-390.

Fang, J., Shen, G., Wan, D., (1999). ***Study of GPS/DR Integrated Navigation System for Urban Vehicle.*** China Journal of Highway and Transport.Vol.12, No.1, Jan..84-89.

Fang, J., Shen, G., Wan, D., (1999). ***An Adaptive Federated KalmanFilter Model for GPS/DR/Map Matching Integrated Navigation System in LandVehicle***. Control and Decision. Vol.14, No.5, Sept.. 448-452.

Fang, J., Wan, D., (1996). ***Research of Integrated Navigation Systems withGPS for Vehicle Navigation***. Journal of Southeast University. Vol.26, No.3, May. 96-102.

Wu, Q., Wan, D., Wang, Q., (1999). ***A New Kind of Filter Algorithms ofIntegrated Navigation System for Vehicle***. Journal of Chinese inertial technology.Vol.7, No.2, Jun.,.22-24.

Wu, Q., Wan, D., Xu, X., (1997). ***Design of GPS/DR Integrated Navigation System for Vehicle and Filter Algorithms.*** Journal of Southeast, Vol.27, No.2, Mar.. 55-59.

Wu, Q., Wan, D., (1998) ***An Integrated GPS/Velocity Gyro/Odometer Vehicle Navigation.*** Industrial Instrumentation & Automation. No.3,.pp.10-12.

Wang, Y., (1997). ***Simulation of Positioning Accuracy with GPS/DR/GIS IntegratedNavigation System for Vehicle***. Journal of WuYi University (natural science). Vol.11, No.4. 42-48.

Xiong, S., Jin, T., (1997). ***An Implementation of Dead-Reckoning for a Vehicle GPSNavigation System.*** Journal of Nanchang University. Vol.21, No.1,Mar. 52-57.

Zheng, P., (1999). ***Study of GPS/DR Integrated Navigation System for Vehicle***.Journal of Beijing University of Aeronautics and Astronautics, Vol.25, No.5, Oct. 513-516.

Liu, J., Yuan, X., (1995). ***GPS/INS Integrated Navigation Developing System***.Journal of Chinese Inertial Technology. Vol.3, No.1. 1-6.

Chen, Y., (1997). ***Parallel Implementation of Kalman Filter for Integrated INS/GPS Navigation System***. Control Theory and Applications. Vol.14, No.2, Apr. 219-223.

Borenstein, J., (1994) ***Internal Correction of Dead-reckoning Errors with the SmartEncoder Trailer.*** International Conference on Intelligent Robots and Systems (lROS'94)-Advanced Robotic Systems and the Real World. Munich, Germany, September12-16, 1994, 127-134.

Borenstein, J., Everet, H.R., Feng, L., Wehe, D., ***Mobile Robot Positioning & Sensors and Techniques.*** Invited Paper for the Journal of Robotic Systems, Special Issue on Mobile Robots. Vol.14, No.4. 231 – 249.

Borenstein, J., Feng, L., (1996) Gyrodometry: ***A New Method for Combining Data from Gyros and Odometry in Mobile Robots.*** Proceedings of the 1996 IEEE International Conference on Robotics and Automation, Minneapolis, Apr. 22-28, 423-428.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 191-199

Journal of Global Positioning Systems
2004
© CPGPS

The Advantage of an Integrated RTK-GPS System in Monitoring Structural Deformation

Xaiojing Li

Schools of Electrical Engineering & Telecommunications, School of Surveying & Spatial Information Systems, The University of New South Wales, Sydney NSW 2052, Australia
Email: xj.li@unsw.edu.au; Tel: 61-2-9385 5534, Fax: 61-2-9385 5388

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Monitoring structural response induced by severe loadings such as typhoon is an efficient way to mitigate or prevent damage. Because the measured signal can be used to activate an alarm system to evacuate people from an endangered building, or to drive a control system to suppress typhoon excited vibrations so as to protect the integrity of the structure. A 108m tall tower in Tokyo has been monitored by an integrated system combining RTK-GPS and accelerometers. Data collected by the multi-sensor system have been analysed and compared to the original finite element modeling (FEM) result for structural deformation monitoring studies. Especially, the short time Fast Fourier Transform (FFT) analysis results have shown that the time-frequency relation does give us almost instantaneous frequency response during a typhoon event. In this paper the feasibility of integrating advanced sensing technologies such as RTK-GPS with traditional accelerometer sensors, for structural vibration response and deformation monitoring under severe loading conditions, is discussed. The redundancy within the integrated system has shown robust quality assurance.

Key words: vibration, displacement, deformation, integrity, structural response

1. Introduction

Civil engineering structures are typically designed based on the principles of material and structural mechanics. Finite element model (FEM) analysis and wind tunnel tests of scaled models are often carried out to assist structural design (see, e.g., Penman et al, 1999). However, loading conditions in the real world are always much more complicated than can be imagined, and hence key man-made structures must be monitored to ensure that they maintain integrity of design, construction and operation.

In general, until recently, monitoring the dynamic response of civil structures for the purpose of assessment of damage has relied on measurements from accelerometer sensors deployed on the structure. Studies conducted on such data records have been useful in assessing structural design procedures, improving building codes and correlating the response of the structure with the damage caused. However, a double integration process is required to arrive at the relative displacements, and there is no way to recover the static or quasi-static displacement from the acceleration. Also it is difficult to overcome drift natural to the accelerometer. Therefore, it has been proposed to integrate an accelerometer sensor with GPS and other sensors in order to best utilize the advantageous properties of each.

In contrast to accelerometers, GPS can measure directly the position coordinates (Parkinson & Spilker, 1996), hence providing an opportunity to monitor, in real-time and full scale, the dynamic characteristics of the structure to which the GPS antennas are attached. In order to achieve cm-level accuracy, the standard mode of precise differential GPS positioning locates one reference receiver at a base station whose three dimensional coordinates are known, so that the second receiver's coordinates are determined relative to this reference receiver. Preliminary studies have proven the technical feasibility of using GPS to monitor dynamic structural deformation due to winds, traffic, earthquakes and similar loading events in, e.g., the UK (Ashkenazi and Roberts, 1997), USA (Kilpatrick et al, 2003), Singapore (Brownjohn et al, 1998) and Japan (Tamura et al, 2002). Although GPS offers real-time solutions, it has its own limitations. For example, GPS can only sample at a rate of up to 20Hz (Trimble, 2003), although higher rates may be possible in the future.

The paper is organized as follows. The second part presents an analysis of data collected using GPS, accelerometer and anemometer sensors while monitoring the responses of a 108m tall steel tower during Typhoon No. 21 on the first of October 2002. The third part discusses the necessity of integration of GPS and accelerometer sensors; and finally the paper concludes with a summary of the research findings.

2. Deformations Monitored by RTK-GPS and Accelerometer

2.1 The field monitoring system

As part of the collaborative research between the Tokyo Polytechnic University and the UNSW, a combined RTK-GPS and accelerometer system has been deployed on a 108m steel tower in Tokyo owned by the Japan Urban Development Corporation, in order to monitor the tower's deformation on a continuous basis. At the top of the tower, a GPS antenna together with accelerometers and an anemometer were installed. Another GPS antenna was setup on the top of a 16m high rigid building, as a reference point 110m away from the tower. In addition, strain gauges were set in the foundation of the tower to measure member stresses. Figure 1 illustrates the experimental setup. Note the local/monitoring coordinate system has been established so that X is East, Y is North and Z is pointing to the zenith.

The RTK-GPS and accelerometer data were recorded at 10Hz and 20Hz sampling rates respectively. And the data was collected from 19:00 to 21:00 Japan Standard Time (JST) during Typhoon No. 21 on 1 October 2002 (Tamura et al., 2002). Because the weather was cloudy and rainy the solar heating effect on the steel tower will be ignored in the analysis of later sections.

It is well known that light, flexible buildings are more favorable for resisting seismic force, while heavy, stiff buildings are more favorable for resisting wind force. Tall buildings in Japan have to satisfy these two opposite design criteria, and this is one of the most difficult design issues for tall buildings in Japan. The focus in this paper will be only on the effects of typhoon on the 108 m tower.

2.2 Typhoon data set

The overall plots of time series of the RTK-GPS measured displacements in X, Y and Z directions are shown in Figure 2. Measurements from the accelerometers (X and Y directions only) are given in Figure 4. From the least-squares polynomial fitting (blue lines) in Figure 2 the maximum displacements in the X and Y directions are around 5.8cm and -4.5cm respectively, indicating significant static and quasi-static movements. However in the Z direction the signal seems to bounce between 2cm and -2cm. This confirms the less accurate characteristics of the vertical measurement by using GPS. Figure 3 are the recorded wind speed and direction time series. The higher speed corresponds to larger displacement and stronger acceleration. The changes of the wind direction agreed with the signs of displacements in X and Y direction measured by RTK-GPS. For example, at 20:00 the wind was NW (Figure 9). Therefore, the displacements on X and Y directions should be positive and negative respectively and this is exactly what the GPS has recorded (Figure 2).

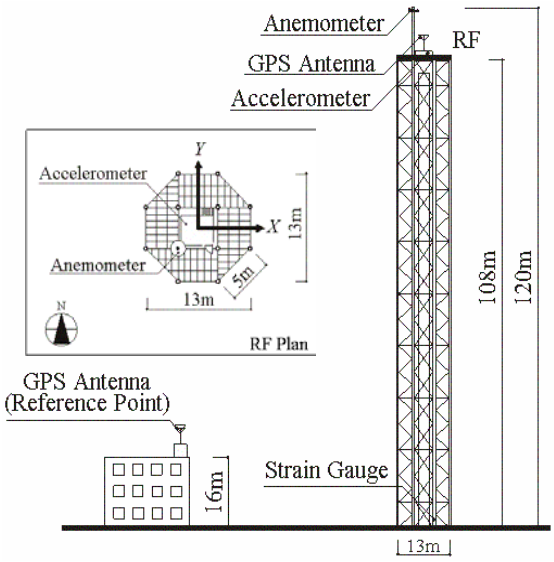


Figure 1 The 108m steel tower used for deformation monitoring study.

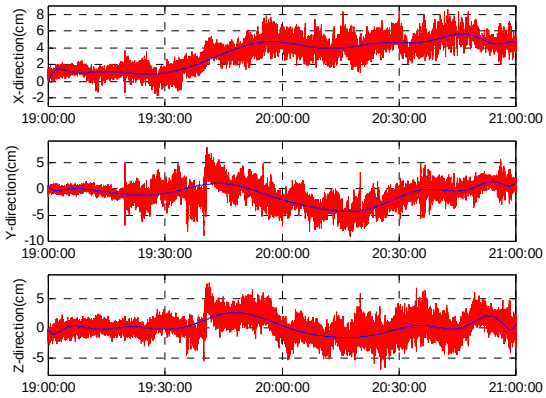


Figure 2 Displacement measured by RTK-GPS.

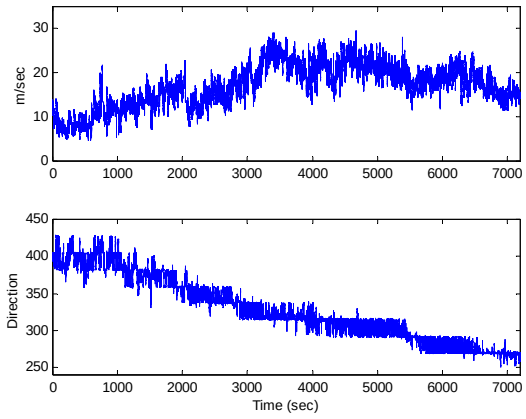


Figure 3 Wind speed and direction time series.

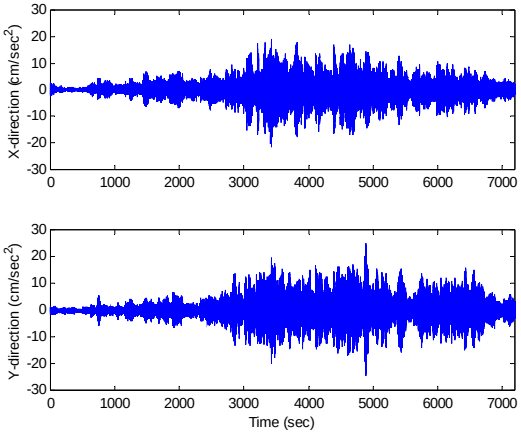


Figure 4 Acceleration measured by accelerometer.

2.3 Signals obtained from the typhoon data set

According to a previous study on the basic characteristics and applicability of RTK-GPS (Tamura et al, 2002), GPS results seem to follow closely the actual displacement under conditions when the vibration frequency is lower than 2Hz, and the vibration amplitude is larger than 2cm. Consequently it is possible to assess the accuracy of recorded data by analysing the spectrum of the wind-induced vibration of the tower. Using the FFT and zooming in the RTK-GPS spectrum, it is obvious that there is a 0.57Hz component in the Y direction (blue peak in Figure 5). However, in the X direction (red) there are two peaks at 0.54Hz and 0.61Hz. In the 0–0.20Hz range, the static and quasi-static responses can be clearly distinguished by comparing with the wind speed fluctuation spectrum Figure 6. And the noises such as multipath are contributing to all three directions coupled in the 0.05-0.2Hz range.

Figure 7 is a zoom-in of the accelerometer spectrum. It appears to be very clean at the lower frequency end (0 to 0.20Hz). It is obvious also that there is a 0.57Hz component in the Y direction (green peak). However, in the X direction (blue peaks) there are two peaks at 0.54Hz and 0.61Hz by further zooming in, which are identical to what RTK-GPS has picked up. In addition, there are less significant peaks around 1.86Hz and 2.16Hz for both the X and Y directions, while at 1.86Hz the X direction is much stronger than the Y. In the meantime, the overall FFT spectrum gives a peak at 4.56Hz as well (Figure 23).

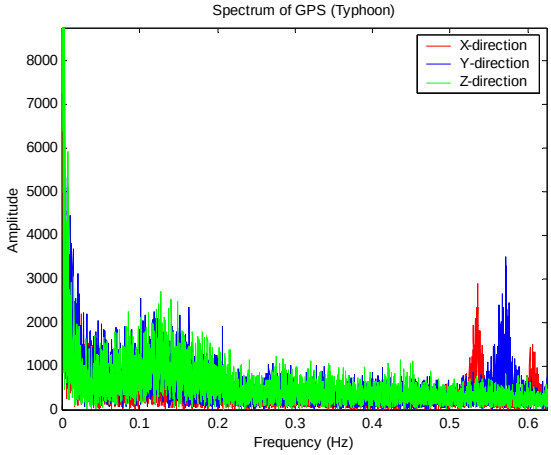


Figure5 Zoom-in of the FFT spectrum of RTK-GPS.

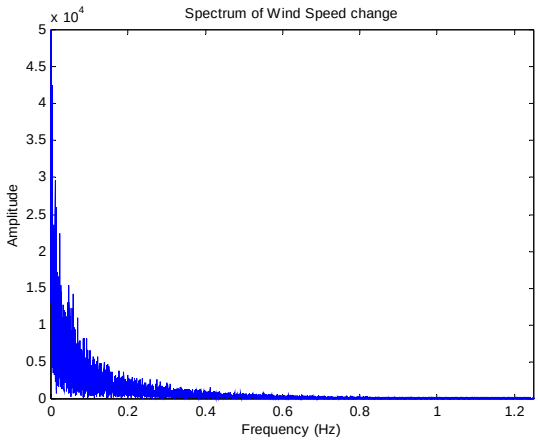


Figure 6 The wind speed fluctuation spectrum.

According to the anemometer recordings, the wind direction was mostly North and North-West during the two hour experiment. Figure 8 gives the mean direction every 51.2 seconds. North corresponds to the Y-direction of the monitoring system. Hence, it is useful to focus on the Y direction (blue in Figure 5).

It is clear that the spectrums for both the accelerometer and RTK-GPS have the dominant frequency of 0.57Hz, indicating that it is the lowest natural frequency of the steel tower. This result agrees with Tamura et al (2002),

using the power spectral density analysis and the random decrement (RD) technique through other typhoon event recording. FEM and frequency domain decomposition (FDD) analysis has also confirmed that 0.57Hz is the first mode natural frequency of the tower, 2.16Hz and 4.56Hz are the 2nd mode and 3rd mode respectively (Figure 9, after Yoshida et al, 2003). Although it is not clear what are the driving factors for the other peaks, there is no doubt that the GPS and accelerometer are complementary (by comparing Figures 5 and 7).

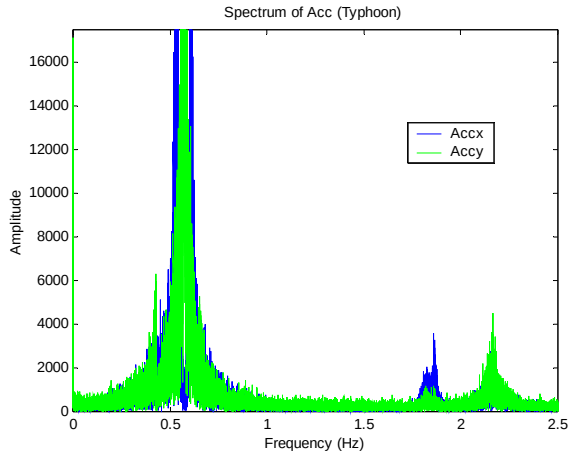


Figure 7 Zoom-in of the FFT spectrum of accelerometer.

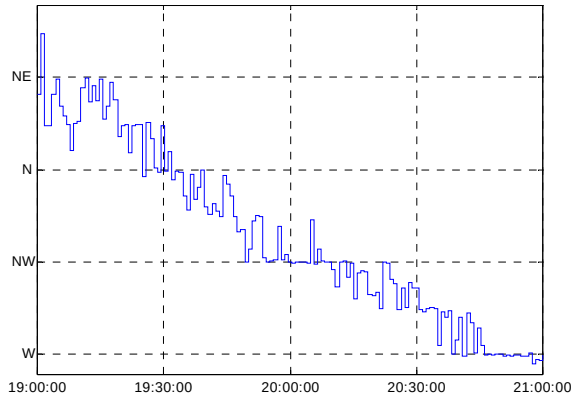


Figure 8 Wind direction change in 51.2s mean.

2.4 Wind-induced response of the tower: comparison between RTK-GPS and accelerometer

Wind-induced response of a structure generally consists of three components: a static component due to mean wind force; a quasi-static component caused by the low frequency wind force fluctuations; and a resonant component caused by the wind force fluctuation near the structure's first mode natural frequency (Tamura, 2003). Can the integrated GPS and accelerometer system provide all of the required information? The answer is probably "yes". Yet no single sensor can do it alone. From the signal analysis in Section 2.3, the GPS sensor gives more information at the low frequency end while the accelerometer sensor gives more at the high frequency end. This means that both of them lose information which is very important for civil engineers. The following analysis will study this problem in more detail.

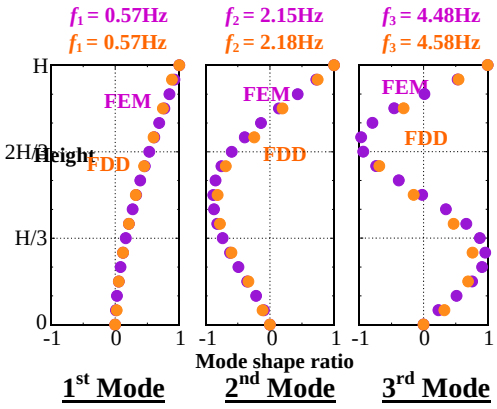


Figure 9 Modes shape obtained by FEM and FDD.

First, a double differentiation procedure is applied to a segment of RTK-GPS measured displacements data (both X and Y directions) of 100 second duration in order to convert them into acceleration. Figures 10 and 12 show the results of acceleration derived from RTK-GPS measured displacements, compared to accelerometer-measured accelerations. In the figures, the upper plots are the GPS displacements; the middle plots are the accelerations derived from the GPS measurements; and the bottom plots are the accelerometer-measured accelerations. Comparing the time series in the middle and bottom plots, it can be seen that they agree very well with each other, although the accelerometer does reveal more high frequency details, as indicated from the FFT spectrum analysis previously.

The reverse data processing is also possible. The same segments of RTK-GPS data (Figure10 and 12) are compared with the displacements derived from corresponding accelerometer measurements through a double integration process. In Figures 11 and 13, the red and blue are displacements measured by GPS and derived from accelerometer data respectively. The green and pink lines are the results of polynomial fitting. From the results shown in Figures 11 and 13, it can be seen that the vibrations (dynamic components) can be related to each other. But it is very hard to compare them quantitatively because the static and quasi-static displacements are missing from the accelerometer derived results, since there is no way to determine the integration constant.

Note that no intentional offset is used when plotting the two results.

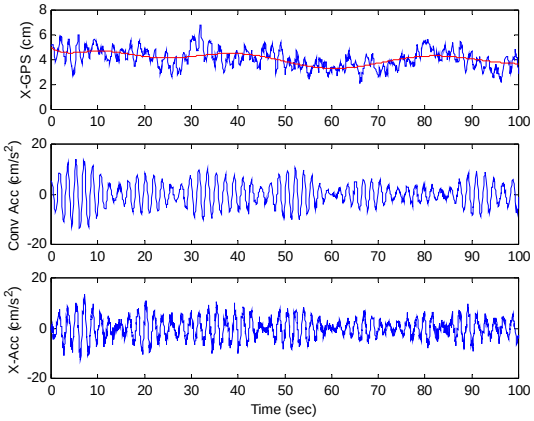


Figure 10 RTK-GPS derived and accelerometer-measured accelerations (X-direction).

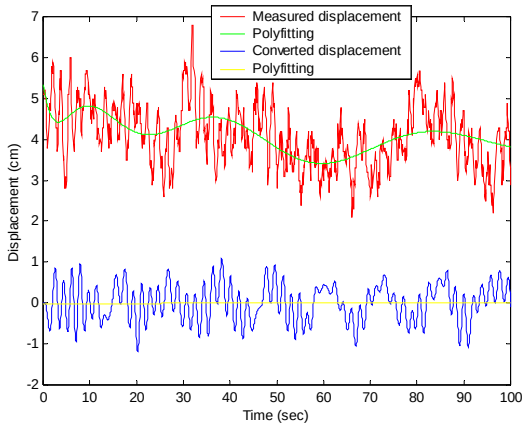


Figure 11 RTK-GPS measured and accelerometer derived displacements (X-direction).

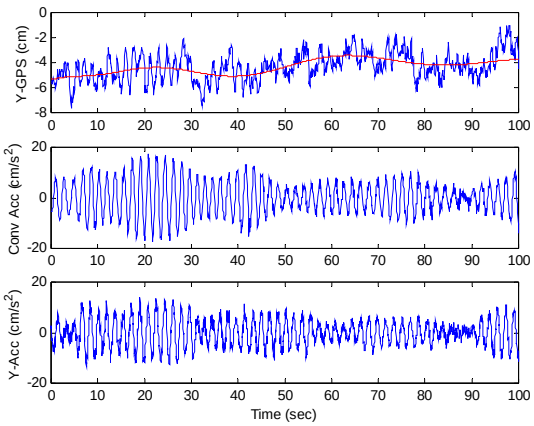


Figure 12 RTK-GPS derived and accelerometer-measured accelerations (Y-direction).

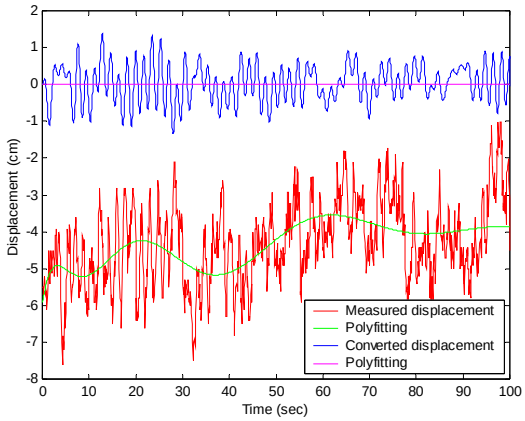


Figure 13 RTK-GPS measured and accelerometer derived displacements (Y-direction).

2.5 Time-Frequency analysis

The results of Section 2.3 are obtained by using the global FFT, whose spectrum represents the frequency composition of the WHOLE time series. It tells the relative strength (amplitude) of the frequency components, but does NOT tell when a particular frequency component occurs or takes over as the most significant frequency. It might be very useful to assess the response of or damage to the structure if it can be detected when a particular frequency component becomes dominant. A way of determining this frequency-time relationship is to apply the FFT to a short segment of the time series each time, or the so-called short time FFT analysis.

Consider the measured digital signal of $x[n]$ as a time series with respect to t. The spectrum $X[f_a]$ obtained by using Fourier transform of a discrete signal (DTFT) is a function of t, which can be derived as follows. First, Fourier transform of a discrete signal (Ambikairajah, 2003) is:

$$X[\theta] = \sum_{n=-\infty}^{\infty} x[n]e^{-jn\theta}, \quad -\pi \le \theta \le \pi \tag{1}$$

Where, $\theta = \omega T$ is the relative frequency, T is sampling period, ω is the frequency in radians.

Considering f_a is the frequency of the signal, and $\omega = 2\pi f_a$. The sampling frequency is determined as $f_s = \frac{1}{T}$. Therefore, θ can be given by:

$$\theta = 2\pi \frac{f_a}{f_s} \tag{2}$$

Then, the time length t of the measurements can be calculated by using the sampling period T multiple the numbers of samples n.

$$t = nT, \quad 0 \le n \le N \tag{3}$$

Substituting Eq. (1) with (2) and (3), Eq. (1) can be rewritten as:

$$X[f_a] = \sum_{n=0}^{N} x[n] e^{-jn2\pi \frac{f_a}{f_s}} = \sum_{n=0}^{N} x[n] e^{-j2\pi f_a nT} = \sum_{n=0}^{N} x[n] e^{-j2\pi f_a t} \tag{4}$$

Hence, from Eq. (4) we can see that the frequency f_a of the digital signal $x[n]$ is a function of time t. When t approaches zero, the output frequency from DTFT becomes as the instantaneous frequency. In other words, by applying DTFT on small samples of measurements, the frequency- time relationship can be established. The instantaneous frequency output can be expressed as:

$$X[f_a] = \sum_{n=N_1}^{N_2} x[n] e^{-j2\pi f_a t}, \Delta n = N_2 - N_1 \ \& \ t = \Delta nT \tag{5}$$

By taking the samples Δn in power of 2, the efficient algorithm for the short time FFT analysis is obtained.

We can now use Eq. (5) to analyze the tower's frequency response with respect to time during this typhoon event. In order to represent the typical frequency which appeared in the FFT spectrum, it is important to select appropriate number of samples of measurements to be analysed each time, and only the maximum peak to be taken into account. Figures 14 and 15 are the results for GPS measurements. A total of 512 samples were processed each time. Hence the time length is 51.2s. It can be seen clearly by comparing Figure 16 that when the wind speed was over 20m/s, the tower's first mode natural frequency 0.57Hz became dominant in Y direction. The peaks of 0.61Hz and 0.54Hz appeared in X direction at different instants of time. Figure 17 is the wind speed measurement analyzed by the short time FFT, given us that the wind force fluctuation between 0.02-0.05 Hz mostly. This can be used to explain the static and quasi-static components of the wind induced response of the tower.

The time-frequency results from the accelerometer measurements are shown in Figures 18, 19, 20 and 21. Figures 18 and 19 are analyzed in window length of 128 samples. Therefore the frequency with maximum amplitude will be registered every 6.4 seconds. In Figure 18, the two frequencies of 2.16 Hz and 1.86Hz which have been picked up by using global FFT analysis in X direction appeared at different instants of time as well. It could be caused by the change of wind speed and direction. The frequency of 0.57 Hz is the central of the appeared frequency bandwidth.

Figures 20 and 21 are processed with a window length of 256 samples. These two figures give us more clearly frequency boundary of the first mode natural vibration of the tower. Furthermore, it looks like that the tower responded with the higher mode vibration of small displacement to show that it is almost still in the weak wind, i.e. when the wind speed is lower than 10m/s.

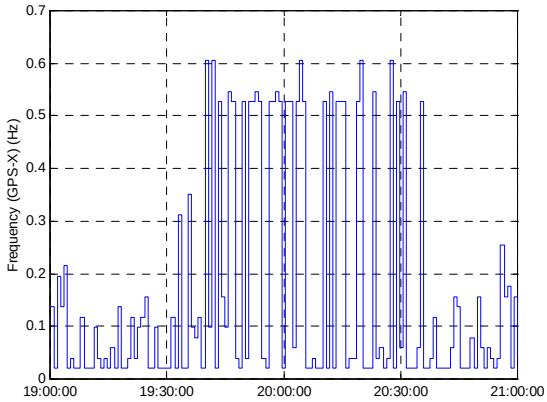


Figure 14 Time-Frequency for GPS x-dir analysis.

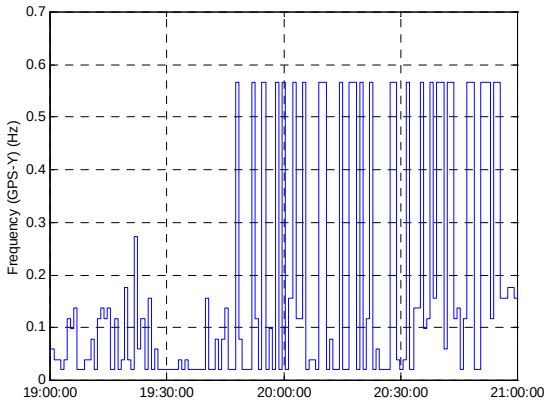


Figure 15 Time-Frequency for GPS y-dir analysis.

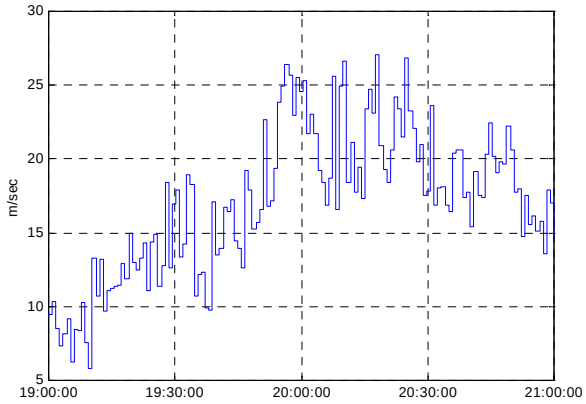


Figure 16 Mean wind speed every 51.2 seconds.

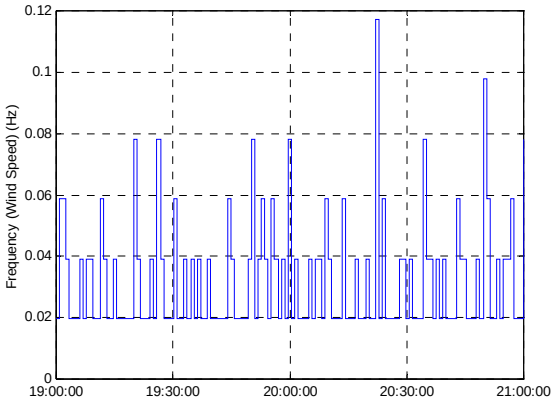


Figure 17 Time-Frequency for wind speed fluctuation.

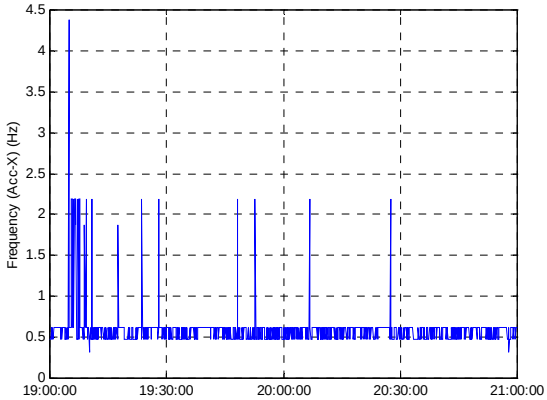


Figure 18 Time-Frequency analysis for Acc x-dir with a window of 6.4s.

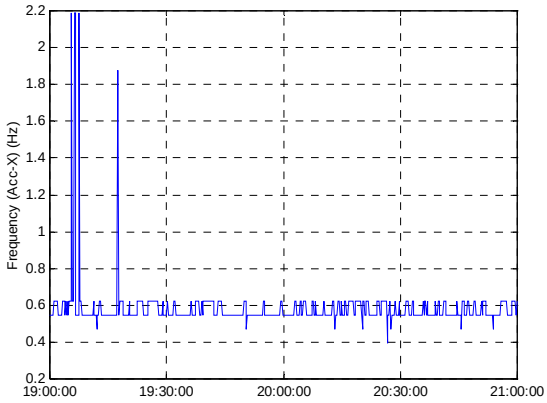


Figure 20 Time-Frequency analysis for Acc x-dir with a window of 12.8s.

All these results agree well with the previous global FFT analysis, but with the timing of events/signals, which enables cross-correlation with wind speed and direction data.

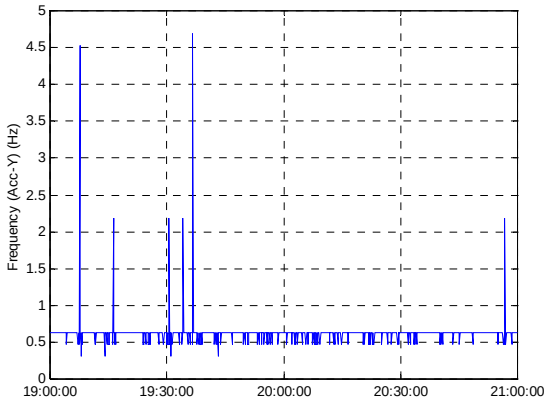


Figure 19 Time-Frequency analysis for Acc y-dir with a window of 6.4s.

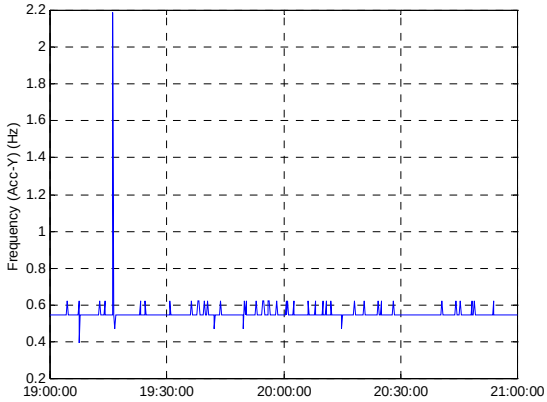


Figure 21 Time-Frequency analysis for Acc y-dir with a window of 12.8s.

It can be seen that in general GPS can pick up signals at the low frequency end (0-0.2Hz), probably contaminated by GPS-specific noise such multipath, while it is easier for the accelerometer to record high frequency signals (2Hz and above). *Therefore, the two sensors are complementary.* On the other hand, the two sensors do have some overlapping capability in the band between 0.2 – 2Hz, i.e., whatever is picked up by the accelerometer will also be picked up by the RTK-GPS. *This overlap provides redundancy within the integrated system which enables robust quality assurance* (i.e., double integration and differentiation as shown in Section 2) and the overlapping band is likely to increase with the advancement of GPS receiver technology.

Note currently the RTK-GPS data are collected using the single-base RTK approach. It is possible to enhance the RTK performance by employing network RTK, especially applicable in Japan considering the available nation-wide, very dense, and continuous GPS (CGPS) network.

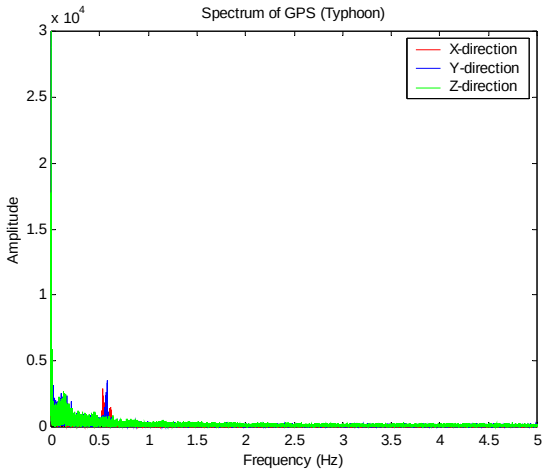


Figure 22 Overall FFT spectrums of GPS measurements during the typhoon.

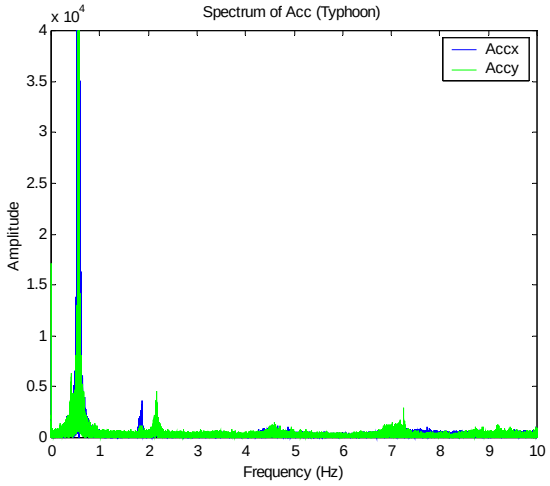


Figure 23 Overall FFT spectrums of accelerometer measurements during the typhoon.

The limitation of single-base GPS-RTK is the distance between the reference receiver and the user receiver so as to ensure that distance-dependent biases such as orbit error, and ionospheric and tropospheric signal refraction, can cancel. This has restricted the inter-receiver distance to 10km or less if very rapid ambiguity resolution is desired (i.e. less than a few seconds). Moreover, in dense urban environments the identification of a site for a reference station can be especially challenging. The reference station must have a clear view of the sky in order to track the satellites, be in close proximity to the user to minimize baseline separation errors, be relatively multipath free, and be on a stable platform. In most cases, rigid structures are low rise and often overshadowed by taller neighbouring buildings, making their ability to track satellites highly variable.

In network-RTK there is a data processing 'engine' with the capability to resolve the integer ambiguities between the static reference receivers that make up the CGPS network. The 'engine' is capable of handling data from receivers 50-100km apart, operates in real-time, instantaneously for all satellites at elevation cut-off angles down to a couple of degrees (even with high noise data that are vulnerable to a higher multipath disturbance). The network-RTK correction messages can then be generated.

Because there will be several reference stations available in the 100km radius of the user receiver, there are several advantages of network-RTK over the standard single-base RTK configuration (Rizos, 2002):

- No need for the user to operate his own reference station (cost reduction).
- Elimination of orbit bias and ionosphere delay.
- Better reduction of troposphere delay, multipath disturbance and observation noise.
- RTK can be extended to what might be considered 'medium-range' baselines (up100km).
- Low-cost single-frequency receivers can be used for RTK (further cost reduction).
- Improve the accuracy, reliability, integrity, productivity and capacity of GPS positioning.

Therefore, network-based GPS-RTK has been proposed as a key component for the future integrated system. It will be tested on the 108m steel tower currently monitored using the single-base GPS-RTK.

4. Concluding Remarks

GPS and accelerometer sensors have been installed on a 108m tall steel tower and data have been collected at 10Hz and 20Hz during a typhoon on 1 October 2002. The wind induced deformation has been analysed in both the time and frequency domains. In the frequency domain, both the GPS and accelerometer results show strong peaks at 0.57Hz, which agrees with previous studies by using different methods, although GPS measurements are noisy in the low frequency end and accelerometer measurements at the high frequency end. Measurements have been converted to displacement (in the case of accelerometer) and acceleration (in the case of GPS) through double integration and double differentiation respectively, for the purpose of direct comparison. The results agree with each other very well, except that the static component is missing from the accelerometer derived results.

Measurements from GPS, accelerometer and anemometer have been analysed in detail using short-time FFT. The

spectrum features have been well explained using such an approach.

The benefits of an integrated system of GPS and accelerometer sensors to monitor the structural deformation have been highlighted.

Acknowledgements

The author wishes to thanks several members of the Satellite Navigation And Positioning Group (SNAP) and the Photonics and Optical Communications Group (POCG) at UNSW for useful discussions and help.

Special thanks to the UNSW Faculty of Engineering for the scholarship supporting the author's PhD studies under the joint supervision of the Schools of Surveying and Spatial Information Systems, and Electrical Engineering and Telecommunications. This work would not have been possible without this joint supervision. The research is also sponsored by a Faculty Research Grant of the UNSW.

The collaboration of Professor Yukio Tamura and Mr Akihito Yoshida from the Tokyo Polytechnic University has been invaluable for these studies.

The author has benefited enormously from Associate Professor Ambikairajah's course on Digital signal processing.

Reference

Ambikairajah, E., 2003. ***Digital signal processing Lecture notes***. School of Electrical Engineering & Telecommunications, the University of New South Wales, Australia.

Ashkenazi, V., & Roberts, G.W., 1997. ***Experimental monitoring of the Humber Bridge using GPS***, *Proc. of Institution of Civil Engineers*, 120, 177-182.

Brownjohn, J., Pan, T.C., Mita, A., & Chow, K.F., 1998. ***Dynamic and static response of Republic Plaza***, *Journal of the Institution of Engineers Singapore*, 38(2) 35-41.

Kilpatrick, J., Kijewski, T., Williams, T., Kwon, D.K., Young, B., Abdelrazaq, A., Galsworthy, J., Morrish, D., Isyumov, N., & Kareem, A., 2003. ***Full scale validation of the predicted response of tall buildings: preliminary results of the Chicago monitoring project***, *Proc. 11th Intern. Conf. on Wind Engineering*, June 2-5, Lubbock, Texas, USA.

Parkinson, B., & Spilker, J., 1996. ***Global Positioning System: Theory and Applications***, Washington, DC: American Institute of Aeronautics & Astronautics, 793pp.

Penman, A., Saxena, K., & Sharma, V., 1999. ***Instrumentation, Monitoring and Surveillance****: Embankment Dams*, A.A. Balkema: Rotterdam, 288pp.

Rizos, C., 2002. ***Network RTK research and implementation: A geodetic perspective***, *Journal of Global Positioning Systems*, 1(2), 144-150.

Tamura, Y., Matsui, M., Pagnini, L.-C., Ishibashi, R., & Yoshida, A., 2002. ***Measurement of wind-induced response of buildings using RTK-GPS***, *J. Wind Eng. and Industrial Aerodynamics*, 90, 1783-1793.

Tamura, Y., 2003. ***Design issues for tall buildings from acceleration to damping.***, 11th International Conference on Wind Engineering, June 2-5, Lubbock, Texas, USA.

Trimble, 2003. http://www.trimble.com/ms860.html

Yoshida, A., Tamura, Y., Matsui, M., & Ishibashi, S., 2003. ***Integrity monitoring of builgings by hybrid use of RTK-GPS and FEM analysis***, 1st International Conference on Structural Health Monitoring and Intelligent Infrastructure, November 13-15, Tokyo, Japan.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 200-207

Journal of Global Positioning Systems
2004
© CPGPS

Secure Tracking using Trusted GNSS Receivers and Galileo Authentication Services

Oscar Pozzobon[1], Chris Wullems[1], Kurt Kubik[2]

[1] Qascom, Via O.Marinali 87, 36061 Bassano del Grappa (VI), Italy
e-mail: o.pozzobon@qascom.com, c.wullems@qascom.com; Tel: + 39 0424-525-473; Fax: +39 0424-527-800
[2] University of Queensland, Brisbane QLD, Australia
e-mail: kubik@itee.uq.edu.au; Tel: +61 7-3365-8328; Fax: +61 7-3365-4999

Received: 08 December 2004 / Accepted: 03 February 2005

Abstract. This paper describes a secure framework for tracking applications that use the Galileo signal authentication services. First a number of limitations that affect the trust of critical tracking applications, even in presence of authenticated GNSS signals, are detailed. Requirements for secure tracking are then introduced; detailing how the integrity characteristics of the Galileo authentication could enhance the security of active tracking applications. This paper concludes with a discussion of our existing tracking technology using a Siemens TC45 GSM/GPRS module and future development utilizing our previously proposed trusted GNSS receiver.

Key words: Galileo, GPS, security, authentication, privacy, tracking

1 Introduction

In recent years there has been an increasing presence of real time tracking applications in the GNSS (Global Navigation Satellite Systems) market. It is possibly one of the fastest growing areas in GNSS, however, fast growth brings with it significant drawbacks such as the lack of standardization and security support for applications that require security for revenue protection or even safety critical services.

The emergence of low cost GSM GPRS (General Packet Radio Service) providing wireless communications throughout Europe has fuelled rapid growth in the development of embedded, application specific platforms, which provides low-cost telematic solutions that are easy to integrate. This rapid growth has happened at the expense of a suitable security framework for development of safety-critical or financially-critical applications. To date there are few telematic solutions on the market that even offer security services. Most rely on the flawed encryption and key establishment protocols of GSM. One of the most significant issues with GSM encryption is the lack of support for cryptographic integrity protection.

The increase of tracking applications has resulted in the emergence of a number of new applications including GPS based road toll payments, GPS based insurance policy, location-based access control, finance and tracking for security. A number of these applications will be introduced in the following sections.

1.1 GNSS based road toll collection

Future road toll collection systems are planned to be based on GNSS technology in order to reduce infrastructure costs and to achieve region based tax collection facilitating regional tolling variations such as pollution-tax for highly polluted areas. As taxes are calculated based on location data from a GNSS receiver, it is imperative that location data is trusted. An additional concern is management of privacy in these technologies. This issue is to a large extent sill unresolved, as location data obtained by operators of these schemes are hardly controlled by the user.

In mid-2003, the EC (European Commission) adopted a proposal (COM(2003) 488) to align the national systems of road tolling for heavy-goods vehicles in Europe. This directive does not impose a particular technical solution, but another EC communication (COM(2003) 123) proposes the use of GNSS positioning and GSM/GPRS mobile communications for new electronic toll systems from 2008 onward (for all vehicle types). In addition, it proposes that migration from legacy microwave systems

to GNSS / mobile communications should be complete by 2012. Authentication, availability and integrity of the location data will be the main technical problem for revenue protection of toll operators. From the user perspective, privacy continues to be the biggest concern, which will demand innovative solutions for the management of location data privacy.

1.2 GNSS based insurance policy

GNSS tracking systems can be applied to insurance in obtaining time usage and location travelled by insured vehicles. This will permit insurance companies to create pay per usage policies. For some users this would result in cost reduction, as the policy cost can be estimated by vehicle usage as total kilometres, risks of travelled regions, average speed, time usage and time based risk. The use of this type of information in insurance policies continues to be an active area of research. A number of experiments have been conducted by Norwich Union and IBM in this area to test the user responses and system reliability.

1.3 GNSS based aircraft tracking

The Future Air Navigation System (FANS) is an example of aircraft tracking using a system facilitating a free flight. In this technology, FANS is responsible for communications, navigation and surveillance. The Flight Management Computer (FMC) uses GPS, inertial measurements, air data, and other navigation radios if available to facilitate surveillance by a traffic management (ATM) center, such that aircraft can be tracked at all times. The security of this system is important where there is an absence of radar systems to verify the reported location of the aircraft. Security of the data transmitted to the ATM is particularly critical in this situation, as it can assist in preventing intentional location spoofing.

1.4 GNSS based access control and auditing

Location can be used for the enhancement of access control systems and auditing. There are many applications where location context information can supplement existing security. Such applications would utilize location from GNSS systems in providing location-based audit trails, or access control policy where resources are granted or denied based on the location of a user. Security and trust of location is particularly important in this type of application, as insecure location acquisition would not only result in serious security breaches, but a false sense of security.

2 Limitations of existing technologies

There a number of security issues with tracking systems which utilize existing GNSS and telematic devices. These issues limit the potential for development of critical applications. The most significant limitations are:

1. Lack of signal authentication: There is no civil method to authenticate the GPS signal;
2. Lack of framework or standardized methodology to verify the integrity of a device and assess its security;
3. Lack of standardized telematic protocols that provide communications security; and
4. Significant privacy issues, such that it is difficult to obtain the location and preserve privacy at the same time.

The following subsections discuss these limitations in terms of signal authentication, device integrity and telematic security.

2.1 Signal authentication

There have been a number of recent efforts to quantify the extent of vulnerabilities and limitations the GPS civil signal imposes on civil applications in the presence of malicious attacks. Perhaps the most prominent vulnerability analysis was the report on the vulnerabilities of GPS in transportation, performed by the Volpe center for the US Department of Transportation (Volpe 2001).

As the GPS civil signal is not authenticated, it is possible to simulate it. In recent years simulators have become readily available, such that a GPS simulator can be hired relatively cheaply and can be connected to the antenna of a GPS receiver in a vehicular tracking module for example. Because of the possibility of signal simulation, the current generation of GPS tracking modules poses a potential security risk for use in safety or financially critical applications, such as the tracking of hazardous materials.

The U.S. Senate has recently approved a measure providing funds for the Transportation Security Administration (TSA) to develop measures for tracking trucks carrying hazardous materials (HAZMAT). As stated by (Gibbons 2004) More than 800,000 shipments of hazardous materials take place in the United States every day, including flammable fuel products, potentially explosive fertilizers, and volatile chemicals. GPS signal authentication is necessary for secure tracking in order to prevent a hijacker from simply spoofing the reported location.

The risk of signal simulation attacks is significantly reduced where the cost of hiring a signal simulator

outweighs the potential cost savings in defeating a tracking system. An example of this is performing a simulation attack in order to avoid the payment of road tolls or to cheat an insurance company by falsely reporting the number of kilometers traveled.

In addition to GPS signal spoofing, there is the potential for spoofing of augmentation data. The most widely used augmentation systems are satellite based augmentation systems including the European Geostationary Navigation Overlay Service (EGNOS) and the American Wide Area Augmentation System (WAAS). Both these augmentation systems do not provide authentication or cryptographic integrity protection. Spoofing of the correction data provided by these systems can introduce small but significant errors, which may be problematic where a few meters of error is critical.

2.2 Device integrity

A typical tracking device is composed of a GPS module and a communications module such as a GPRS modem or radio modem. Some tracking devices additionally contain a microprocessor with software to process the data from the GPS module. There are a number of limitations of current GPS modules:

- There is no cryptographic authentication or integrity protection of NMEA position, time or velocity data sourced from the GPS module; and
- The NMEA location data from the GPS module can be trivially simulated.

In addition to the limitations of GPS modules, the lack of authentication and end-to-end communications security between the tracking device and the telematic server can exacerbate the possible attacks that can be performed, such as masquerading as the tracking device to spoof the location.

2.3 Privacy

The growth of GNSS tracking applications has also created significant privacy concerns. The first techniques for managing privacy where proposed by (Spreitzer and Theimer 1993). These techniques were based on a location broker residing in the middleware layer. In recent times, considerable research has been conducted in the specification of protocols and policy representations in the context of a cellular location. For tracking applications, privacy is an issue for both single purpose tracking devices and multipurpose ones. For single purpose tracking, a user may be concerned with the usage of the location data by the destination application. Where a tracking device is used in a multi-application context, such as a device supporting both toll-payment and insurance tracking applications, the user must be able to configure privacy policies such that only the absolute minimum required location data for a given application is provided.

The user in effect should be able to adjust the accuracy of the location observations depending on the intended use and identity of the recipient, and the compliance with user's privacy policy. The most prominent effort to create location privacy standards outside the cellular domain is seen in the Geographic Location/Privacy (Geopriv[1]) working group of the Internet Engineering Task Force (IETF). This group has developed a number of draft standards for representation of privacy policy data and protocols for management of location privacy.

3 Secure tracking using Galileo

There are a number of requirements for security and privacy in GNSS tracking applications. The following subsections discuss these requirements and possible implementations in terms of satellite navigation system security, tracking device security, communications security to the telematic server, and privacy.

3.1 Signal authentication

Signal authentication of satellite navigation systems is required in order to ensure that the source of the satellite signaling is not from a simulator, but is genuine. There are a number of existing and proposed signal authentication methods (Hein and al 2002), (Scott 2003) which are summarized below:

- *Signal Authentication through Authentication Navigation Messages (ANM):* The ANMs would include a digital signature authenticating the other navigation messages that contain data including ephemeris and almanac data. Using the digital signature, the certified receiver is able to authenticate the source of messages and verify their integrity. These authentication messages are created on the ground and transmitted to the satellites for broadcast. This method has a security limitation, in that the messages can be acquired by a certified receiver and modulated over a simulated signal in order to spoof the Galileo signal. This would require functionality that is not commonly found in commercial signal simulators, and would require the operation to be performed within a very small time window. Documents from the Galileo design consolidation indicates that the Galileo

[1] Refer to http://ecotroph.net/geopriv

Open Service may support this type of signal authentication (Galilei 2003).

- *Signal Authentication through Spread Spectrum Security Codes (SSSC) (Scott 2003):* SSSCs are synchronous cipher streams seeded by an unsent digital signature from an Authentication Navigation Message, interleaved with normal spreading sequences. The receiver stores A/D samples and once the digital signature is received, it is able to generate the security spreading code reference signal and correlate it with the stored samples. If the SSSC is detected at the correct power level, the signal is authenticated. This technique has the innovative advantage that permits authentication in a signal open to the public without the difficulties of key distribution; however it has the limitation that the spoofing detection is proportional to the antenna gain and that the authentication verification is not immediate. A more secure type of authentication based on SSSCs is also proposed by (Scott 2003), utilizing a Civil Antispoof Security Module (CASM) with a preloaded Red Key and the authentication navigation message for seeding of the cipher stream generator. This type of signal authentication does not have the drawbacks of the public SSSC version.
- *Signal Authentication through Spreading Code Encryption (SCE):* Spreading code encryption is one of the oldest signal authentication techniques, currently used by the GPS P(Y) code, an exclusively military service, and is projected to provide authentication of the Galileo CS and PRS signals. As the spreading code is secret, without knowledge of the spreading code, signal access is denied. For this reason, the spoofer cannot simulate the signal, and hence authentication of the signal is achieved when the user possesses the correct spreading code. In GPS' P(Y) code, the P code is publicly known, and the secret spreading code is obtained using P code with a Red Key, or a Black Key and the Selective Availability Anti-spoofing Module (SAASM) (Callaghan and Fruehauf 2003). The Black Key is the Red Key encrypted with the public key of a given SAASM. , allowing the Red Key to be decrypted inside the tamper-resistant SAASM which contains its private key. The Black Keying infrastructure allows for electronic key distribution and does not compromise the classified Red Key.

Civil signal authentication is a challenge for next generation satellite systems. As detailed above, there are a range of different strength security solutions and proposals. The suitability of a particular signal authentication mechanism is dependant on the cost to defeat the mechanism; balanced against the possible gain should the mechanism be successfully defeated.

The Galileo signals and corresponding authentication schemes to date have not been decided. Based on generally available information, it is evident that Galileo will provide a number of different services, the following of which are projected to provide signal authentication: (Hein and al 2002)

- *Open Service (OS):* Based on available literature (Galilei 2003), encryption may be provided to the open service on the E5b-I data channel. This service will provide authentication through Authentication Navigation Messages (ANM);
- *Safety of Life Service (SOL):* This service will provide satellite and signal integrity messages and Authentication Navigation Messages (ANM);
- *Commercial Service (CS):* This service will provide access restriction and authentication through spreading code and data encryption (SCE); and
- *Public Regulated Service (PRS):* This service will provide access restriction and authentication through spreading code and data encryption (SCE).

Of particular interest to consumer applications such as insurance tracking and toll collection is the signal authentication provided on the Open Service. Fig. 1 illustrates a candidate navigation message authentication scheme as detailed in the Galilei Project Galileo Design Consolidation (Galilei 2003).

The authenticated navigation messages would be created by the Ground Control Centre and up-linked to the satellites. In theory a public key certificate certified by the Galileo certification authority would be included in the navigation messages, and could be verified by the Galileo certification authority certificate stored on a certified receiver. Once the public key is verified, the receiver would be able to verify the signature included in the navigation messages, and hence authenticate the source of the navigation messages, and implicitly the integrity of the messages.

For applications with greater security requirements, the CS and PRS signals will provide signal authentication through encrypted ranging codes. While there is no literature on the key distribution schemes, it can be assumed that that the implementation would be similar to the declassified Black Key distribution framework used

with the P(Y) code of GPS. The key distribution and key storage problem in this scenario are similar to the Selective Availability Anti Spoofing Module (SAASM) (Callaghan and Fruehauf 2003), used in military applications.

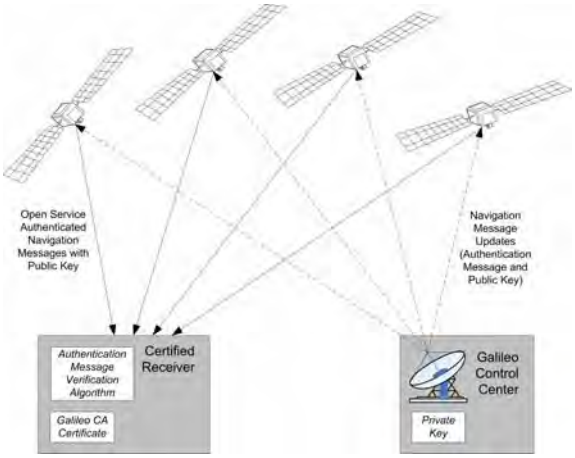


Fig. 1 Navigation authentication message

3.2 Receiver security for tracking applications

We have previously proposed the design for a trusted civil receiver in (Pozzobon et al. 2004). In this architecture a tamper resistant receiver uses public key cryptography to assure the chain of trust to the application, and provide authentication and cryptographic integrity of the data to the application. The tamper resistant receiver acquires and authenticates the signal, calculates the location and creates data authentication messages containing a digital signature of the location data, signal state and tamper-resistance state using the receiver private key. The data is sent to a telematic server via a wireless communications service such as GPRS.

The data is send using an extension to the National Marine Electronic Association (NMEA) protocol 0183 for GPS data navigation we have proposed (Pozzobon et al. 2004), which provides authenticated position, time and velocity information as well as signal state and tamper-resistance state of the receiver.

As integrity verification is a computation embedded in the device, the computation must be trusted. The device must implement all the necessary integrity verification algorithms and transmit to the telematic server any information regarding status, availability, integrity of the signal, and the results from the verification operations. Technology such as the Trusted Computing Platform (TCP)[2] can be used to build trusted systems, that is, systems where the application running can be trusted, with the assurance that it has not been compromised, or modified by an attacker.

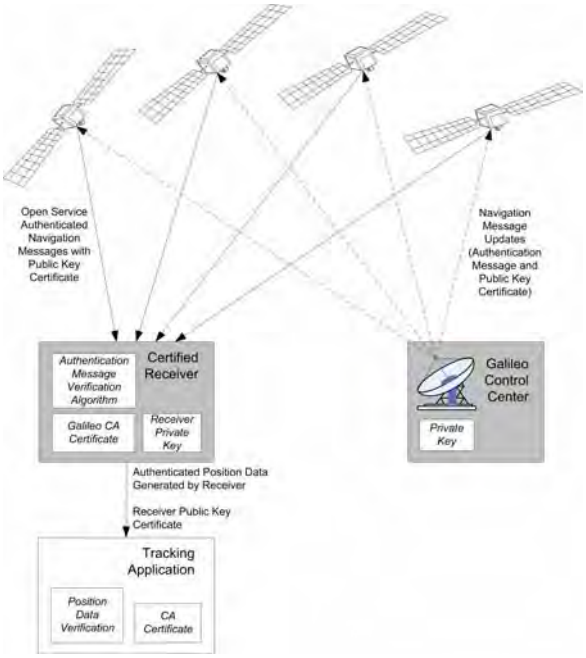


Fig. 2 Tracking with a certified receiver

The concept behind this technology is to build a computer with trusted building blocks (TBB). In TBB, the core root of trust for measurement (CRTM) is the Basic Input/Output system (BIOS) of the computer. The CRTM uses the trusted platform module (TPM) for cryptography operations (storage of keys, encryption) in order to trust the system boot[3] and verify the integrity of the subsequently executed applications. The whole "chain of trust" is based on public key infrastructure (PKI), RSA and 3DES algorithms.

3.3 Location privacy

As detailed in Section 2.3, there are a set of emerging standards for location privacy developed by the Geopriv working group of the IETF. A high level diagram of the interactions between architecture components in using Geopriv protocols for management of privacy between a GNSS tracking device and telematic server is illustrated in Fig. 3. The components of the architecture include:

- *GNSS tracking device:* The proposed tamper resistant GNSS device;
- *The Telematic Server:* The server that manages location and communications;

[2] Refer to https://www.trustedcomputinggroup.org/

[3] The procedure that starts all the necessary process of an operating system

- *Rule Holder:* The entity that provides the rules associated with a particular target for the distribution of location information; and
- *Rule Maker:* The authority that creates rules governing access to location information for a target.

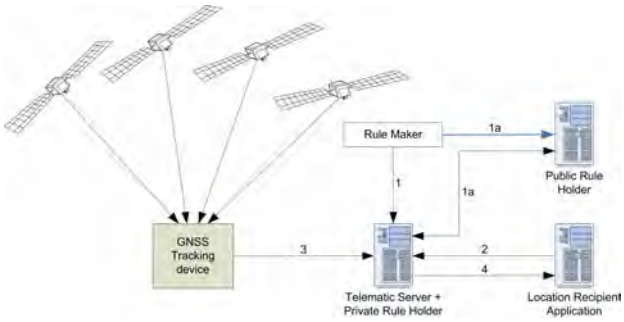


Fig. 3 Location privacy using Geopriv protocols

Fig. 3 additionally illustrated the process of privacy rule upload on the telematic server and how privacy is protected in the location acquisition process. This process, which complies with the RFC3693 requirements, consists of the following steps:

- *Rule Transfer:* The Rule Maker sends a Rule to the Telematic Server containing the privacy information;
- *(1a) Signed Rule:* the Rule Maker may write a Rule and place it in a Public Rule Holder as an alternative. The Telematic Server can access the Public Rule Holder to read the signed Rules;
- *Location Information Request:* The Location Recipient Application requests location information to the Telematic Server.
- *Locate:* The Telematic Server is either continuously receiving the location data from the GNSS tracking device or can request updates on the location. The communication is encrypted; and
- *Filtered Location Information:* The Telematic Server sends the location information to the Location Recipient Application. The information may be filtered in order to comply with the privacy policy and rules defined by the rule maker.

4 Secure tracking using current and emerging technology

The following subsections detail current and emerging technologies that have been developed or are currently under development.

4.1 Current tracking devices developed by Qascom

A number of tracking solutions have been developed using commercially available hardware such as the Siemens TC454 GSM module and SiRF5 GPS Receiver. Fig. 4 illustrates the components in the Qascom tracking blackbox. The location information is processed by a Java applet loaded into the flash memory of the TC45 GSM module. The java applet is responsible for processing GPS location, velocity and time data. The information is then processed according to the requirements of the application, such as insurance.

The applet signs the resulting information destined for the telematic server by invoking a sign function using the SIM Toolkit interface[6] of the SIM card. This functionality requires a SIM card with support for public key operations accessible through the SIM toolkit interface. The signature of the data provides cryptographic integrity protection of the data as well as authenticating the source of the data. The resulting information and signature are sent to the telematic server using one of two supported communication modes: SMS (Short Messaging Service) messaging or GPRS (General Packet Radio Service).

To ensure the privacy of the data communicated to the telematic server, appropriate security must be used. Where SMS is chosen as the mode of communication, no session cryptography is used. This is for three reasons:

- SMS is a point-to-point delivery service where the end point is another device on the GSM network. It is assumed that a destination on the GSM network cannot easily be spoofed. The tracking device can be authenticated using the source MSISDN (Mobile Station International ISDN Number);
- GSM provides data encipherment using the A5 algorithm from the mobile station to the base stations. The core network is assumed to be protected from intruders. Flaws in the encryption and key establishment protocols of GSM are overcome by the use of new protocols in 3G deployments; and
- A significant number of messages would be required for authentication and establishment of keys, resulting in a communications protocol that is too costly.

In this mode, the telematic server must have a mobile station present on the GSM network.

[4] Refer to http://www.siemens-mobile.com
[5] Refer to http://www.sirf.com/
[6] SIM Application Toolkit (SAT) is defined in GSM 11.14 standard for 2G networks, and 3GPP 31.111 for 3G networks.

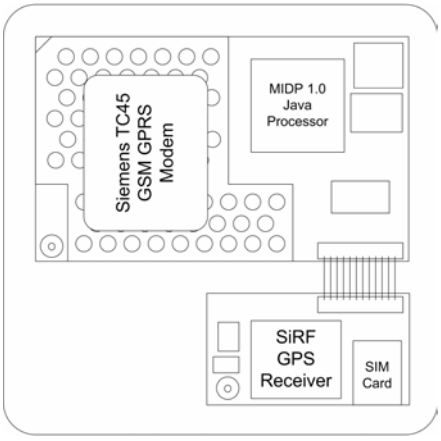


Fig. 4 Qascom GPS/GSM tracking box

Where GPRS is chosen as the mode of communications, the processed data and signature must pass through the GPRS Gateway (GGSN) and over the Internet to the telematic server as illustrated in Fig. 5. For this reason, the tracking device must be able to authenticate the destination telematic server, and establish session keys with this entity in order to transfer the processed data. This process is facilitated through a small implementation of SSL (Secure Sockets Layer) in the Java applet.

This implementation of SSL uses the SIM card to verify the public key certificate of the telematic server and perform public key operations for the exchange of keys. Computation of the pre-master-secret is performed in the Java applet, using the real-time clock in the GSM module for generation of random numbers. The pre-master-secret generated by the applet is then encrypted in the SIM card using the telematic server's public key. The master secret is generated from the pre-master secret as specified in (Dierks and Allen 1999). The master-secret is the key used for encrypting the session between the device and the telematic server. The processed data and signature can then be transmitted over the established security context. This mode of communications is significantly more secure than the SMS mode.

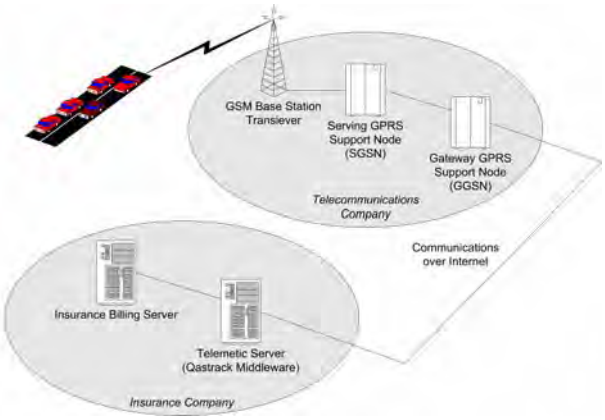


Fig. 5 GNSS tracking using GPRS

While the current solution provides protection from malicious attacks on the communications between the tracking device and the telematic server, the solution does not prevent an attacker from loading malicious firmware (Java applet) onto the tracking device. As the device has no mechanism to authenticate the firmware, the malicious software could generate data that minimizes the amount payable to an insurance company, for example. The software would still be able to use the cryptographic functionality of the SIM card, such that the telematic server would be unaware of such malicious activity.

In addition, the current GPS implementation provides no signal authentication or protection from spoofing. Thus, it is possible to attach a GPS signal simulator to the antenna and spoof the location.

4.2 Next generation devices being developed by Qascom

We are currently involved in development of a next generation tracking device that will fulfill the security requirements of applications such as insurance tracking. The proposed tracking device will contain a trusted GNSS receiver, which will be developed by a consortium of research institutions and companies including Qascom in Europe. The trusted GNSS receiver is described by Pozzobon, Wullems and Kubik in (Pozzobon et al. 2004). The proposed tracking device will contain a general purpose processor with tamper-resistant key storage and a cryptographic coprocessor (trusted platform module). In addition to supporting secure communications over GPRS, the tracking device will also provide the facility for authentication of firmware as well as support for trusted GNSS positioning. Fig. 6 illustrates the secure tracking box.

The verification of the data provided by the trusted GNSS receiver is performed by the authenticated software. This verification initially involves verification of the public key of the trusted GNSS receiver, before being able to verify the digital signatures contained within the authentication data.

This software not only verifies the GNSS position, time and velocity data, it additionally processes the data as required by the application. The resulting data to be sent to a telematic server contains the processed GNSS data, signal state and device compliance reports obtained from the trusted GNSS receiver, the public key certificate of the tracking device, and a digital signature of this data. The digital signature is calculated by the cryptographic co-processor using the private key stored in the tamper-resistant, secure key storage. The telematic server first must verify the public key certificate of the tracking device, after which it is able to verify the data received.

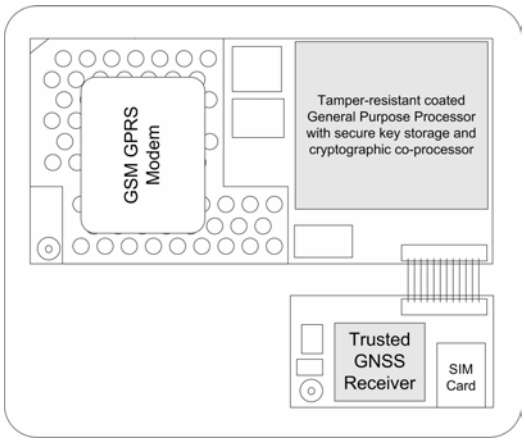


Fig. 6 Qascom secure tracking box

5 Conclusions

This paper has described a secure framework for tracking applications that use the Galileo Authentication Services. Requirements for secure tracking in both consumer and critical applications were introduced, detailing how the new signal characteristics of Galileo can be used to enhance the security of tracking applications. Requirements in terms of authenticated signaling, device security and location privacy were introduced. This paper concluded with a discussion of both existing and future tracking device developments and detailed the strategies used to mitigate the security issues in the presence of the Galileo security differentiators.

References

Callaghan, S., and Fruehauf, H. (2003). ***SAASM and Direct P(Y) Signal Acquisition***, The Journal of Defense Software Engineering, 16(6), 12-16.

Dierks, T., and Allen, C. (1999). ***The TLS Protocol Version 1.0,*** Network Working Group, Internet Engineering Task Force, Request for Comments 2246.

Galilei. (2003). ***The Galilei Project: GALILEO Design Consolidation,*** European Commission.

Gibbons, G. (2004). ***HazMat Keeps on Truckin',*** *GPSWorld* (October), 6.

Hein, G. W. et al. (2002). ***Status of Galileo Frequency and Signal Design.*** Brussels.

Pozzobon, O., Wullems, C., and Kubik, K. ***Requirements for Enhancing Trust, Security and Integrity of GNSS Location Services,*** Institute of Navigation (ION), 60th annual meeting, Dayton, OH, USA.

Scott, L. ***Anti-Spoofing & Authenticated Signal Architectures for Civil Navigation Systems,*** *ION GPS, GNSS 2003*, Portland, OR.

Spreitzer, M., and Theimer, M. ***Providing Location Information in a Ubiquitous Computing Environment.*** *Fourteenth ACM Symposium on Operating System Principles*, 270–283.

Volpe, J. A. (2001). ***Vulnerability Assessment of the Transportation Infrastructure Relying on the Global Positioning System.***

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 208-217

Journal of Global Positioning Systems
2004
© CPGPS

NAVIO – A Navigation and Guidance Service for Pedestrians

Günther Retscher and Michael Thienelt

Institute of Geodesy and Geophysics, Vienna University of Technology, Gusshausstrasse 27-29, A-1040 Wien, Austria
e-mail: gretsch@pop.tuwien.ac.at Tel: +43 1 58801 12847; Fax: +43 1 58801 12894

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. In the research project NAVIO (Pedestrian Navigation Systems in Combined Indoor/Outdoor Environements) at our University we are working on the improvement of navigation services for pedestrians. Thereby we are mainly focusing on the information aspect of location-based services, i.e., on the user's task at hand and the support of the user's decisions by information provided by such a service. Specifications will allow us to select appropriate sensor data and to integrate data when and where needed, to propose context-dependent routes fitting to partly conflicting interests and goals as well as to select appropriate communication methods in terms of supporting the user guidance by various multimedia cartography forms. These taks are addressed in the project in three different work packages, i.e., the first on "Integrated positioning", the second on "Pedestrian route modeling" and the third on "Multimedia route communication". In this paper we will concentrate on the research work and findings in the first work package. For continuous positioning of a pedestrian suitable location technologies include GNSS and indoor location techniques, cellular phone positioning, dead reckoning sensors (e.g. magnetic compass, gyro and accelerometers) for measurement of heading and travelled distance as well as barometric pressure sensors for height determination. The integration of these sensors in a modern multi-sensor system can be performed using an adapted Kalman filter. To test and to demonstrate our approach, we take a use case scenario into account, i.e., the guidance of visitors to departments of the Vienna University of Techology. The results of simulation studies and practical tests could confirm that such a service can achieve a high level of performance for the guidance of a pedestrian in urban areas and mixed indoor and outdoor environments.

Key words: Pedestrian navigation, Location-based services, Multi-sensor systems, Integrated positioning, Indoor location

1 Introduction

Pedestrian navigation services require continuous positioning and tracking of a mobile user with a certain positioning accuracy and reliability. Especially navigating in urban environments and mixed indoor and outdoor areas is a very challenging task as pedestrians move in spaces where no one of the known location methods works continuously in standalone mode. A solution of the problem can only be found if different location technologies are combined in the sense of a modern multi-sensor system. In this paper suitable location technologies for pedestrian navigation are identified and investigated. These technologies include GNSS and indoor location services as well as cellular phone positioning for absolute position determination; dead reckoning sensors (e.g. magnetic compass, gyros and accelerometers) for measurement of orientation and travelled distance from a known start position as well as barometric pressure sensors for height determination. For location determination of a pedestrian in a multi-storey building the use of WLAN (Wireless Local Area Networks) is investigated. To achieve an integrated positioning determination with other sensors and a seamless transition between indoor and outdoor areas, a multi-sensor fusion model based on an extended Kalman filter approach is employed. Finally, in a practical use case scenario the guidance of a pedestrian from public transport stops to our Department of the Vienna University of Technology is investigated. The results of this study showed that such a pedestrian navigation service can achieve a high level of performance.

2 Integrated positioning in navigation services

A reliable pedestrian navigation services requires the determination of the current user's position using different sensors that are integrated into the system design. In the work package "Integrated positioning" of

the research project NAVIO (Pedestrian **Nav**igation Systems in Combined **I**ndoor/**O**utdoor Environements) the following challenging tasks are addressed:

- The capability to track the movements of a pedestrian in real-time using different suitable location sensors and to obtain an optimal estimate of the current user's position.
- The possibility to locate the user in 3 dimensions with high precision (that includes to be able to determine the correct floor of the user in a multi-storey building).
- The capability to achieve a seamless transition for continuous positioning determination between indoor and outdoor areas.

Thereby a navigation support must be able to provide location, orientation and movement of the user as well as related geographic information matching well with the real world situation experienced by pedestrians. The integration of the sensors in a modern multi-sensor system can be performed using a Kalman filter as this algorithm is particularly suited for real-time evaluation. In the following the state-of-the-art in mobile positioning is discussed and suitable sensors for integrated positioning in a pedestrian navigation service are identified.

2.1 State-of-the-art in mobile positioning

Satellite-positioning technologies (GNSS) are employed most commonly for outdoor navigation. Then the achievable positioning accuracies of the navigation system are on the few meters to 10 m level in standalone mode or sub-meter to a few meter level in differential mode (e.g. DGPS). If an insufficient number of satellites is available for a short period of time due to obstructions, then in a conventional approach observations of additional sensors are employed to bridge the loss of lock of satellite signals. For pedestrian navigation, sensors such as a low-cost attitude sensor (digital compass) giving the orientation and heading of the person being navigated and a digital step counter or accelerometers for travel distance measurements can be employed. Using these sensors, however, only relative position determination from a known start position (also referred to as Dead Reckoning DR) is possible and the achievable accuracy depends on the type of movement tracking sensors used and the position prediction algorithm adopted.

For indoor positioning different techniques have been developed. They offer either absolute or relative positioning capabilities. Some of them are based on short-range or mid-range technologies using sensors such as transponders or beacons installed in the building (see e.g. Klinec and Volz, 2000; Pahlavan *et al.*, 2002). An example are the so-called Local Positioning Systems (LPS) that have an operation principle similar to GNSS. The LPS systems claim to achieve a distance measurement accuracy of about 0.3 to 1 m (see e.g. Werb and Lanz, 2000; Sypniewski, 2000), but no details are given on the test results and the achievable accuracy on position fixing. Other indoor positioning systems include so-called Active Badge or Active Bats Systems (Hightower and Boriello, 2001). These systems are mainly employed for the location of people and finding things in buildings. Also WLAN (Wireless Local Area Networks) can be employed for location determination. In this case, the signal strength of the radio signals from at least one WLAN access point installed in the building is measured. The location fix is then obtained with triangulation using measurements to several access points or through comparison with in database stored signal strength values from calibrated points (this method is also referred to as fingerprinting). Further information can be found in e.g. (Bastisch *et al.*, 2003; Beal, 2003; Imst, 2004; Retscher 2004b). As the indoor radio channel suffers from severe multipath propagation and heavy shadow fading, the fingerprint method provides higher accuracies than triangulation. It is reported that positioning accuracies of about 1 to 3 m could be obtained in a test office building using the fingerprint WLAN positioning method (Imst, 2004). Another alternative in indoor geolocation applications is the use of ultra wideband (UWB) systems, which exploit bandwidths in excess of 1 GHz, to measure accurate time of arrival (ToA) of the received signals for estimation of distance (Pahlavan *et al.*, 2002). With results of propagation measurement in a typical modern office building, it has been shown that the UWB signal does not suffer multipath fading (Win and Scholtz, 1998), which is desirable for accurate ToA estimation in indoor areas. The main disadvantage, however, is the possible interference of UWB devices with the GPS system. Also Bluetooth, which has been originally developed for short range wireless communication, can be employed for locating mobile devices in a certain cell area that is represented by the range of the device which is typically less than 10 m. It can be employed for location determination using active landmarks. Locating the user on the correct floor of a multistory building is another challenging task. For more accurate determination of the user's position in vertical dimension an improvement can be achieved employing a barometric pressure sensor or digital altimeter additionally (see Retscher and Skolaut, 2003).

As an alternative for location determination in indoor and outdoor environments, mobile positioning services using cellular phones can be employed. Apart from describing the location of the user using the cell of the wireless network, more advanced positioning methods have been developed. Most of them are based on classical terrestrial

navigation methods where at least two observations are required to obtain a 2-D position fix (see e.g. Balbach, 2000; CPS, 2001; Drane *et al.*, 1998; Hein *et al.*, 2000; Retscher, 2002). The achievable positioning accuracy thereby depends mainly on the location method and type of wireless network (GSM, W-CDMA, UMTS). As advanced and more accurate methods, such as the E-OTD (Enhanced Observed Time Difference) method, require modification of the network as well as installation of additional hardware in the network and reference stations which are called LMU's (Location Measurement Units), they have not been widely deployed yet. Recent developments have therefore been concentrated on the reduction of network modification. The so-called Matrix method (see Duffett-Smith and Craig, 2004) does not need any additional hardware in the network apart from a SMLC (Serving Mobile Location Centre) where the location determination of the mobile handset is performed. Using this method positioning accuracies of 50 to 100 m at the 67 % reliability level can be achieved in the GSM network.

2.2 Suitable sensors for pedestrian navigation services

Suitable sensors and location techniques for pedestrian navigation have been identified at the start of the project NAVIO. Table 1 gives an overview about the positioning methods and the sensors that will be employed in our project. For absolute position determination primarily GNSS is employed. In the case of no GNSS availability it can be replaced by location techniques using cellular phones or indoor positioning systems (e.g. WLAN positioning). Apart from this sensors, relative DR (Dead Reckoning) sensors are employed for the observation of the travelled distance (from velocity and acceleration measurements), direction of motion or heading and height difference. The observables as well as their accuracies are summarized in Table 1.

2.3 Integrated positioning using a multi-sensor fusion model

An integrated position determination, using observations of all available sensors, however, is not performed in most common navigation systems. In vehicle navigation systems for instance the resulting trajectory is determined mainly based on the dead reckoning observations; GNSS is used for updating and resolving the systematic error growth of the DR observations. For guidance of a pedestrian in 3-D space and updating of his route, continuous position determination is required with

Method	Sensor	Observations	Accuracy
GNSS	e.g. Garmin GPS 35 DGPS	y, x, z	6-10 m 1-4 m
Indoor Positioning	e.g. WLAN Positioning IMST ipos	y, x, z	1-3 m
Cellular Phone Positioning	GSM (e.g. Matrix method)	y, x	50-100 m
Dead Reckoning	e.g. PointResearch DRM-III Dead Reckoning Module	y, x z φ	20-50 m per 1 km 3 m 1°
Direction of Motion (Heading)	e.g. Honeywell Digital Compass Module HMR 3000	φ	0.5°
Acceleration	e.g. Crossbow Accelerometer CXTD02	a_{tan}, a_{rad}, a_z	> 0.03 ms^{-2}
Velocity from GNSS	e.g. Garmin GPS 35	v_y, v_x v_z	~ 0,05 m^{-1} ~ 0,2 m^{-1}
Barometer	e.g. Vaisala Pressure sensor PTB220A	z	1-3 m

Tab. 1 Sensors for pedestrian navigation services with their observables and accuracies

(Garmin, 2004; Imst, 2004; Duffett-Smith and Craig, 2004; PointResearch, 2004; Honeywell, 2004; Crossbow, 2004; Vaisala, 2004)

where y, x, z are the 3-D coordinates of the current position, v_y, v_x, v_z are the 3-D velocities, φ is the direction of motion (heading) in the ground plane xy, a_{tan} is the tangential acceleration and a_{rad} is the radial acceleration in the ground plane xy, a_z is the acceleration in height (z coordinate)

positioning accuracies on the few meter level or even higher, especially for navigation in multi-storey buildings in vertical dimension (height) as the user must be located on the correct floor. The specialized research hypothesis of this work package in the project NAVIO is that a mathematical model for integrated positioning can be developed that provides the user with a continuous navigation support. Therefore appropriate location sensors have to be combined and integrated using a new multi-sensor fusion model. A Kalman filter approach is particular suited for the integration and sensor fusion in real-time. Extending basic filter concepts, a Kalman filter approach which integrates all observations from the different sensors will be developed. The model must be able to make full use of all available single observations of the sensors at a certain time to obtain an optimal estimate of the current user state (i.e., position, orientation and motion). For further information on the multi-sensor fusion model the reader is referred to Retscher and Mok (2004) and Retscher (2004b).

3. Practical sensor tests

Practical tests in our research project are carried out for the guidance of visitors of the Vienna University of Technology to certain offices in different buildings or to certain persons. Thereby we assume that the visitor employs a pedestrian navigation system using different sensors that perform an integrated positioning. Start points are nearby public transport stops, e.g. underground stop Karlsplatz in the center of Vienna or railway station Südbahnhof near our university. In the following, results of satellite positioning and first test measurements with the Dead Reckoning Module DRM III from Point Research are presented.

3.1 Satellite positioning test results

Figure 1 shows the GPS measurements for the path from the underground stop Karlsplatz near our University to our Institute building of the Vienna University of Technology in the Gusshausstrasse located in the fourth district of the city of Vienna using two different GPS receivers, i.e., the Trimble GPS Pathfinder Pocket and the Garmin eTrex. The length of the path is approximately 500 m and it starts at the exit of underground station Karlsplatz where open skys provide free satellite visibility. Then the pedestrian walks through a park (i.e., the Resselpark) with trees where satellite signals are frequently blocked over short periods. Both GPS receivers, however, are able to determine the track of the pedestrian with a reasonable positioning accuracy. It can be seen, that the track of the Garmin eTrex receiver is much smoother as he performs some filtering or smoothing to estimate the receiver track compared to the Trimble GPS Pathfinder Pocket which provides the original GPS single point positions. After leaving the park, the path continues in a narrow street (i.e., Karlsgasse) onwards to our Institute building where 5-storey buildings with heights of typically 20 m cause obstructions of the satellite signals. The measurement result from the Garmin eTrex shows an increasing deviation from the true pedestrian path where the maximum deviation in the range of 13 m is reached at the intersection of Karlsgasse with Frankenberggasse. Then the position changes quite significantly as more GPS satellites become available. In the following, the positions show again a drift from the true path. As the Trimble GPS Pathfinder Pocket receiver does not apply any filtering or smoothing, the positions in the Karlsgasse are much more scattered than with the Garmin eTrex. Maximum deviations from the true path of the pedestrian of up to 25 m are reached and in some parts no position determination is possible. This gaps have to be bridged using dead reckoning observations. At the intersection of Karlsgasse with Gusshausstrasse enough satellites are visible for positioning and the path ends in front of the building where our Institute is located.

3.2 Dead reckoning test results

For the measurements the Dead Reckoning Module DRM III from Point Research (PointResearch, 2004) was employed. The system is a self contained navigation unit where GPS is not required for operation. It provides independent position information based on the user's stride and pace count, magnetic north and barometric altitude. The module is designed to self-calibrate when used in conjunction with an appropriate GPS receiver, and can produce reliable position data during GPS outages. The system consists of an integrated 12 channel GPS receiver, antenna, digital compass, pedometer and altimeter. The module is clipped onto the user's belt in the middle of the back and the GPS antenna may be attached to a hat. Firmware converts the sensor signals to appropriate discrete parameters, calculates compass azimuth, detects footsteps, calculates altitude and performs dead reckoning position calculation. A Kalman filter algorithm is used to combine dead reckoning position with GPS position to obtain an optimum estimate for the current user's position and track. With the dead reckoning module and GPS integrated together, a clear view of the sky is only required for obtaining the initial position fix. The fix must produce an estimated position error of 100 m or less to begin initialization. Subsequent fixes use both dead reckoning and GPS data, so obstructed satellites are not as critical as in a GPS only configuration. The Kalman filter continuously updates calbration factors for stride length and compass mounting

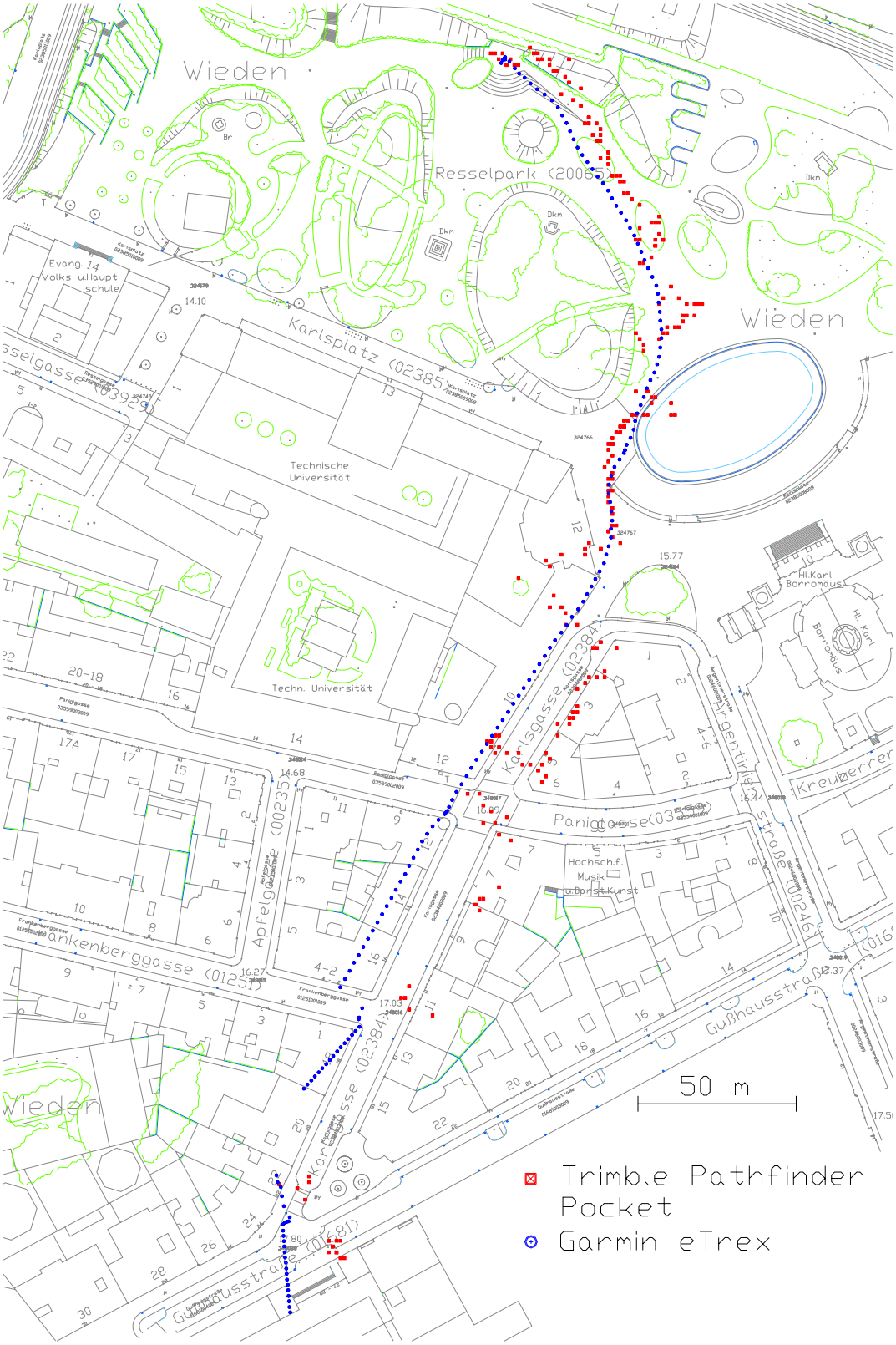


Fig. 1 GPS measurements for the pedestrian path from the underground station Karlsplatz to our office building of the Vienna University of Technology

offset. The GPS position error must be less than 30 m before GPS data will be used by the Kalman filter, and the first such fix will also initialize the module's latitude and longitude. Subsequently, the filter will use any GPS position fix with an estimated position error of 100 m or less, adjusting stride, body offset, northing, easting, latitude and longitude continually.

First of all the Dead Reckoning Module DRM III was tested in open area with GPS satellite visibility. As test site pedestrian paths in the park of Schönbrunn Palace in Vienna have been chosen. Figure 2 shows the trajectory of the pedestrian as well as GPS and two different dead reckoning measurements. For the dead reckoning measurements the GPS positioning and the calbration of the stride length using the Kalman filter algorithm was deactivated. Without using the filter, GPS measurements are not employed to calibrate the stride length and the dead reckoning module uses the preset value of 800 mm for the stride length. The heading of the user is determined from measurements of a digital compass and a gyro. For the first dead reckoning measurement (No. 1) shown in Figure 2 both sensors are employed to obtain the heading, for the second measurement (No. 2) only the observations of the compass are employed. This results in larger deviations from the trajectory for the second dead reckoning measurements; they range from 17 m over a distance of 150 m and 29 m over 200 m. For the first dead reckoning measurement the deviations from the trajectory are in the range of 7 m over a distance of 150 m and 20 m over 200 m. It can therefore be recommended that a combination of compass and gyro measurements are employed for heading observation. An improvement of the dead reckoning measurements can only be achieved if the calibration of the stride length is employed to obtain a better estimate for the travelled distance. For comparison, GPS measurements from the internal receiver of the DRM III module are shown in Figure 2 which reach a maximum deviation of 7 m from the trajectory.

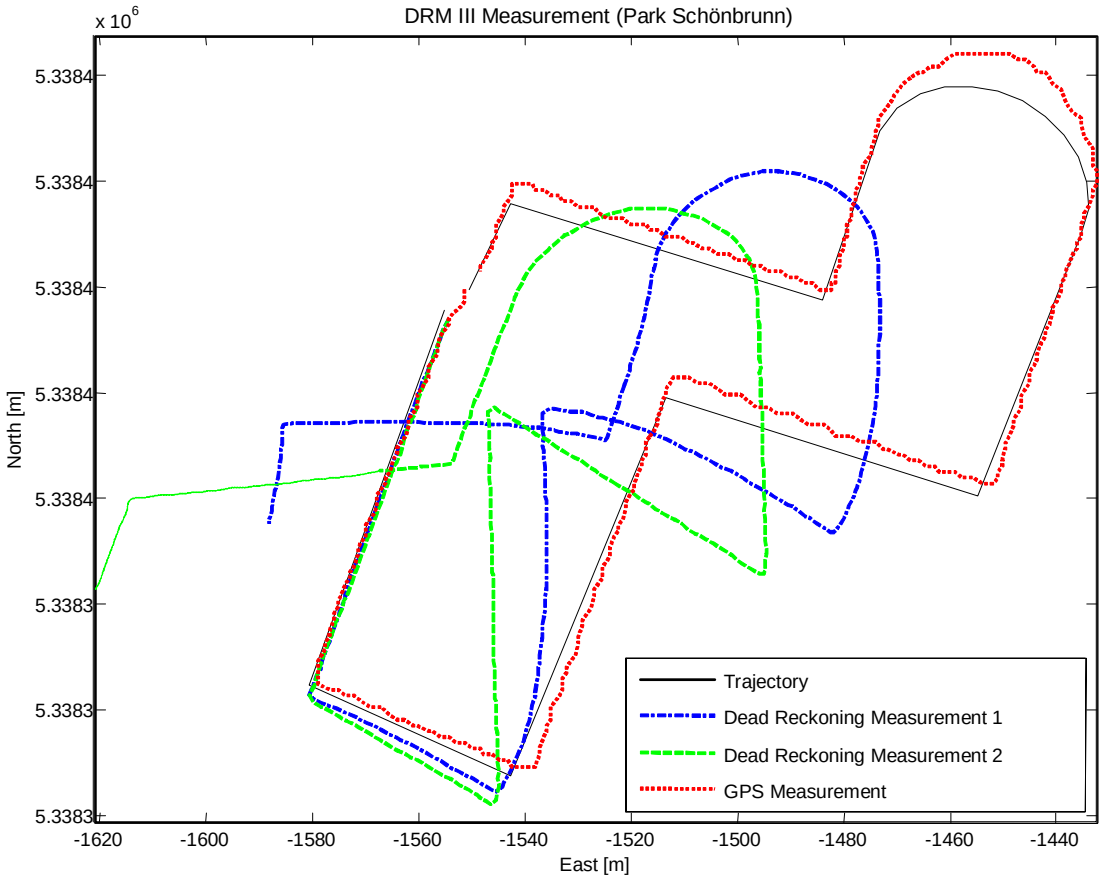


Fig. 2 Test measurements with the Dead Reckoning Module DRM III in the park of Schönbrunn Palace in Vienna

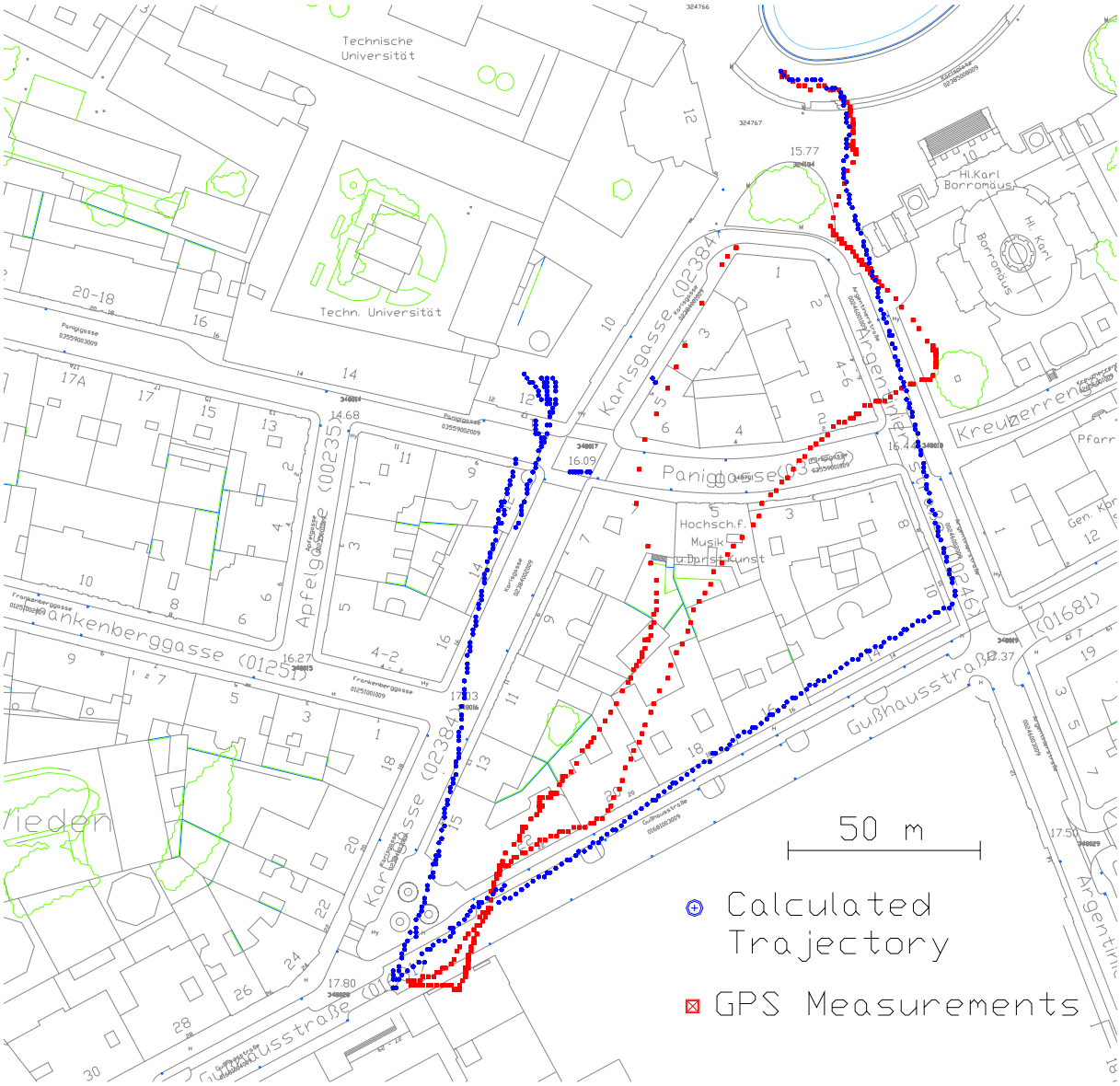


Fig. 3 Test measurements with the Dead Reckoning Module DRM III along a closed loop on narrow streets in the city Vienna

Figure 3 shows the measurement results of the DRM III dead reckoning module and GPS single point positions on a closed loop in the city of Vienna starting from the Resselpark about 160 m along Argentinerstrasse, then 200 m along Gusshaustrasse and 290 m along Karlsgasse back to the start point. The total length of the path is around 550 m. The streets are quite narrow with 5-storey buildings with an average heigth over 20 m causing frequent obstructions of the satellite signals and high GDOP values. As can be seen in Figure 3, the GPS only measurements are quite far away from the true path of the pedestrian along most parts of the track and a reliable match to the correct street would not always be possible. Using the position estimates of the dead reckoning module the resulting trajectory follows the pedestrian track along the most part of path and the deviations are only in the range of a few meters. Due to the large errors of the GPS postions, however, the calbration algorithm of the DRM III fails at the end and the resulting trajectory cannot follow the track of the pedestrian any more. Figure 4 shows the calbration of the stride length in the Kalman filter and it can be seen that the stride length gets smaller and smaller until it reaches nearly 500 mm which is not a matter of fact. In this case, it seems that the weighting of GPS positioning is too high in the Kalman filter calibration process for the stride length.

In our project, an improvement of the current position estimate of a pedestrian using observations of different dead reckoning sensors in combination with GPS and other absolute postioning techniques (see Table 1) should

be achieved using a new multi-sensor fusion model based on an extended Kalman filter approach. Further information about this approach can be found in Retscher and Mok (2004) and Retscher (2004b).

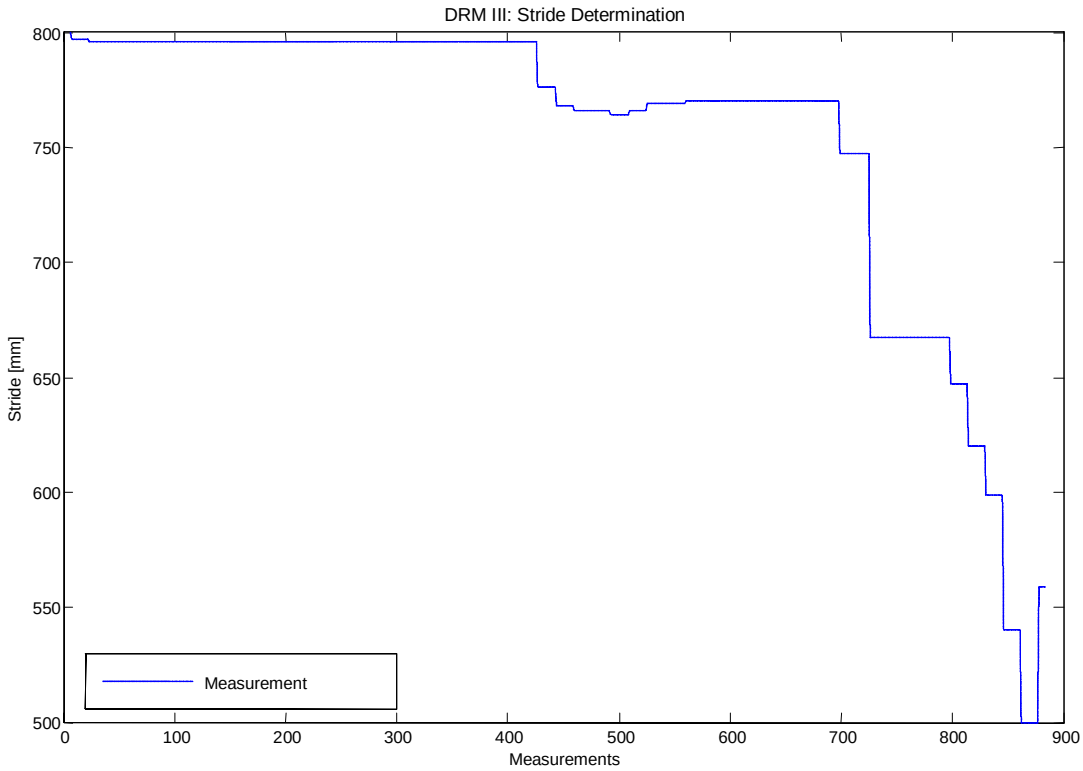


Fig. 4 Calibration of the stride length using GPS observations in the Kalman filter of the DRM III Dead Reckoning Module

4. Conclusions

In the NAVIO project major aspects being important when concepting a pedestrian navigation service are investigated, i.e., integrated positioning, multi-criteria route planning, and multimedia route communication (see Gartner *et al.*, 2004a and b). As a result, a specific pedestrian navigation service as use case will derive the requirements on positioning, route planning, and communication. A prototype of the service will guide visitors of the Vienna University of Technology to departments and persons. Practical tests will allow us to evaluate and demonstrate the usability of the service, and thus, prove the projects attempts.

With the work package "Integrated Positioning" of the project we will contribute to the integration of location sensors in the sense of a multi-sensor system to achieve a continuous positioning of the user of the service and a seamless transition between indoor and outdoor areas. Suitable sensors and location methods have been identified and the basic concept of a multi-sensor fusion model for integrated positioning has been developed (Retscher and Mok, 2004; Retscher 2004b). Special emphasis has been given on the location determination and navigation of a pedestrain in a multi-storey building. Currently we are investigating the use of WLAN positioning for location determination in indoor areas.

The second work package of the project NAVIO on "Pedestrian route modeling" is dealing with the ontological modelling of navigation tasks, deriving well founded criteria and optimization strategies in route selection; and the third work package on "Multimedia route communication" is working on models for context-dependent communication modes of route information. In general, it can be said that the results of the project NAVIO will contribute to the improvement of modern (pedestrian) navigation services.

Acknowledgements

The research work presented in this paper is supported by the FWF Project NAVIO of the Austrian Science Fund

(Fonds zur Förderung wissenschaftlicher Forschung) (Project No. P16277-N04).

References

Balbach O. (2000) ***UMTS – Competing Navigation System and Supplemental Communication System to GNSS***, in: Papers presented at ION GPS Meeting 2000, Salt Lake City, Utah, U.S.A., September 19-22, 2000.

Bastisch H., S. Häfker, J. Krause, S. Wilking, B. Haje (2003) ***Projekt NOM@D - Location Based Services***, Presentation, Communication Networks, University Bremen, Germany, http://www.comnets.uni-bremen.de/itg/itgfg521/aktuelles/fg-treffen-230103/20030120-LBS-Allgemein-v2.pdf, Last access 07.2004 (German).

Beal J. R. (2003) ***Contextual Geolocation - A Specialized Application for Improving Indoor Location Awareness in Wireless Local Area Networks***, in: Papers presented at the Midwest Instruction and Computing Symposium MICS2003, The College of St. Scholastica, Duluth, Minnesota, USA, http://www.css.edu/depts/cis/mics_2003/MICS2003_Papers/Beal.PDF, Last access 07.2004.

CGALIES (2002) ***Co-ordination Group on Access to Location Information by Emergency Services***, Work Package 1 Report, 2002, http://www.telematica.de/cgalies/, Last access 07.2004.

Crossbow (2004) ***CXTD Digital Tilt and Acceleration Sensor***, Product Information, Crossbow, USA, http://www.xbow.com/Products/Product_pdf_files/Tilt_pdf/6020-0012-01_B_CXTD.pdf, Last access 07.2004.

CPS (Cambridge Positioning Services Ltd.) (2001) Cursor™ Mobile Location System, Product Information at http://www.cursor-system.com/, Last access 07.2004.

Drane Ch., M. Macnaughtan, C. Scott (1998) ***Positioning GSM Telephones***, IEEE Communications Magazine, April 1998.

Duffett-Smith P. J., J. Craig (2004) ***Matrix, and Enhanced Satellite Positioning***, Invited paper presented at the 5th IEE International Conference on 3G Mobile Communication Technologies, 18-20 October 2004, Savoy Place, London, UK, 4 pgs.

Garmin (2004) ***Garmin GPS Products***, Garmin Ltd., USA, http://www.garmin.com/products/, Last access 07.2004.

Gartner G., A. Frank, G. Retscher (2004a) ***Pedestrian Navigation System for Mixed Indoor/Outdoor Environments***, in: Gartner G. (Ed.): Geowissenschaftliche Mitteilungen, Schriftenreihe der Studienrichtung Vermessungswesen und Geoinformation, TU Wien, Heft 66, Papers presented at the 2nd Symposium on Location Based Services and Telecartography, January 28-29, 2004, Vienna, Austria, pp. 161-167.

Gartner G., A. Frank, G. Retscher (2004b) ***Pedestrian Navigation System in Mixed Indoor/Outdoor Environment – The NAVIO Project***, in: Schrenk M. (Ed.): CORP 2004 and Geomultimedia04. Proceedings of the CORP 2004 and Geomultimedia04 Symposium, February 24-27, 2004, Vienna, Austria, pp. 165-171, http://corp.mmp.kosnet.com/CORP_CD_2004/archiv/papers/CORP2004_GARTNER_ FRANK_RETSCHER.PDF, Last access 07.2004.

Hein G., B. Eissfeller, V. Oehler, J. O. Winkel (2000) ***Synergies Between Satellite Navigation and Location Services of Terrestrial Mobile Communication***, in: Papers presented at ION GPS Meeting 2000, Salt Lake City, Utah, U.S.A., September 19-22, 2000.

Hightower J., G. Borriello (2001) ***Location Systems for Ubiquitous Computing***, Computer, Vol. 34, No. 8, IEEE Computer Society Press, August 2001, pp. 57-66.

Honeywell (2004) ***HMR 3000 Digital Compass Module, User's Guide***, Honeywell International Inc., USA, http://www.ssec.honeywell.com/magnetic/datasheets/hmr3000_manual.pdf, Last access 07.2004.

Imst (2004) ***Indoor Locating – Imst ipos***, Project c21, Presentation, IMST GmbH, Carl-Friedrich-Gauß-Str. 2, D-47475 Kamp-Lintfort, Germany.

Klinec D., S. Volz (2000) ***NEXUS – Positioning and Communication Environment for Spatially Aware Applications***, in: Papers presented at ISPRS 2000, Amsterdam, The Netherlands, Part B2, pp. 324-330, see also: http://www.nexus.uni-stuttgart.de/, Last access 07.2004.

Pahlavan K., X. Li, J.-P. Mäkelä (2002) ***Indoor Geolocation Science and Technolgy***, IEEE Communications Magazine, February 2002, pp. 112-118.

PointResearch (2004) ***DRM-III Dead Reckoning Module - Engineering Development Tools***, PointResearch Corporataion, USA, http://www.pointresearch.com/drm_eval.htm, Last access 07.2004.

Retscher G. (2002) ***Diskussion der Leistungsmerkmale von Systemen zur Positionsbestimmung mit Mobiltelefonen als Basis für Location Based Services***, in: Papers presented at the Symposium on Telekartographie, January 28-29, 2002, TU Wien, Austria, published as Geowissenschaftliche Mitteilungen, No. 58, Institute of Cartography, Vienna University of Technology, Austria, pp. 41-58 (German).

Retscher G., G. Skolaut (2003) ***Untersuchung von Messsensoren zum Einsatz in Navigationssystemen für Fußgänger***, Zeitschrift für Geodäsie, Geoinformation und Landmanagement ZfV, No. 2, pp. 118-129 (German).

Retscher G. (2004a) ***Multi-sensor Systems for Pedestrian Navigation and Guidance Services***, in: Papers presented at the 4th Symposium on Mobile Mapping Technology, March 29-31, 2004, Kunming, China, CD-Rom Proceedings, 7 pgs.

Retscher G. (2004b) ***Multi-sensor Systems for Pedestrian Navigation***, in: Papers presented at the ION GNSS 2004 Conference, September 21-24, 2004, Long Beach, California, USA, CD-Rom Proceedings, 12 pgs.

Retscher G. (2004c) ***Pedestrian Navigation Systems and Location-based Services***, Invited paper presented at the 5th

IEE International Conference on 3G Mobile Communication Technologies, 18-20 October 2004, Savoy Place, London, UK, 5 pgs.

Retscher G., E. Mok (2004) ***Sensor Fusion and Integration using an Adapted Kalman Filter Approach for Modern Navigation Systems***, Survey Review, Vol. 37, No. 292 April 2004.

Sypniewski J. (2000) ***The DSP Algorithm for Locally Deployable RF Tracking System***, in: Papers presented at International Conference on Signal Processing with Applications, Orlando, October 2000, http://www.syptech.com/publications/publications.html, Last access 07.2004.

Vaisala (2004) ***PTB220 Digital Barometer***, Vaisala, Finland, http://www.vaisala.com/ DynaGen_Attachments/Att2468/PTB220%20Brochure.pdf , Last access 07.2004.

Werb J., C. Lanz (2000) ***Designing a positioning system for finding things and people indoors***, http://www.rftechnologies.com/pinpoint/index.htm, Last access 07.2004.

Win M., R. Scholtz (1998) ***On the Performance of Ultra-Wide Bandwith Signals in Dense Multipath Environment***, IEEE Commun. Letters, Vol. 2, No. 2, Feb. 1998, pp. 51-53.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 218-225

Journal of Global Positioning Systems
2004

Improved atmospheric modelling for large scale high-precision positioning based on GNSS CORS networks in Australia

Craig Roberts[2], Kefei Zhang[1], Chris Rizos[2], Allison Kealy[3], Linlin Ge[2], Peter Ramm[4], Martin Hale[4], Doug Kinlyside[5] and Paul Harcombe[5]

School of Surveying and Spatial Information Systems, University of New South Wales
Email: c.roberts@unsw.edu.au , Tel: +61-2-9385 4464 Fax: +61-2-9313 7493

[1] School of Mathematical and Geospatial Sciences, RMIT University
[2] School of Surveying and Spatial Information Systems, UNSW
[3] Department of Geomatics, The University of Melbourne
[4] Spatial Information Infratructure, Dept of Sustainability and Environment, Victoria Government
[5] Department of Lands, NSW Government

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. This contribution describes a recent Australian Research Council (ARC) project funded under the ARC-Linkage Scheme. The research team comprises researchers from RMIT University, UNSW, University of Melbourne, Spatial Information Infrastructure and the Department of Lands, NSW. The aim of the project is to enhance the utility of continuously operating reference station (CORS) networks in the states of Victoria and New South Wales by developing improved atmospheric correction models to support high accuracy, real-time positioning even when the density of reference stations is insufficient for standard operational GPS techniques such as RTK ('real-time kinematic'). Many applications of Global Navigation Satellite System (GNSS) technology, such as surveying, mapping and precise navigation, require real-time positioning accuracies to centimetre levels. To support these applications, many countries are establishing dense CORS networks with stations, positioned typically a few tens of kilometres apart. However, for Australia with its large and sparsely populated landmass, such dense networks cannot be justified economically. This ARC project will investigate enhancements of sparse networks to maintain similar levels of accuracy as dense CORS networks. It will seek a better understanding and modelling of atmospheric conditions, currently a major limitation in the use of sparse networks for high accuracy techniques. This paper will describe the status of current developments in CORS network infrastructure in Australia, namely GPSnet in Victoria and SydNet in New South Wales. The major research components of the project will be outlined and the technical and practical challenges will be discussed, including some methodologies that will be investigated.

Key words: CORS, GNSS, GPS, network RTK.

1. Introduction

The standard mode of high accuracy *differential* positioning requires one Global Positioning System (GPS) receiver to be located at a "reference station" with known coordinates, while the second "user" receiver simultaneously tracks the same satellite signals. When the carrier phase measurements from the two receivers are combined and processed, the mobile user's receiver coordinates are determined relative to the reference receiver. This can be done in real-time, if the reference receiver data is transmitted to the user's receiver, even while the receiver is moving. The ultimate implementation of such a technique is known as "real-time kinematic" (RTK), and is capable of cm-level accuracy under certain constrained operational conditions (Rizos, 2002a). Most precise land positioning applications have required the user to set-up their own reference station (and if real-time operations were required, the data communication link between reference and user receiver as well). An alternative to this "do-it-yourself" approach is to take advantage of a CORS network of receivers. This is more convenient for users, as they are able to position using only a single receiver, relying instead on a service provider to operate the reference receivers for them. Users may have a variety of accuracy requirements, so CORS networks must be able to cater for the most demanding users, typically those

seeking cm-level accuracy. However, CORS networks also permit innovative network-based techniques to be used, and this is discussed later in this section.

High quality GPS reference stations have been established, in a sparse global network, since the late 1980s to support scientific applications such as tectonic/earthquake hazard research, geodetic reference system definition and maintenance, and atmospheric studies. These stations are typically located hundreds, or even thousands, of kilometres apart. Recently, networks of CORS are being established to support demanding navigation/positioning applications such as surveying and mapping, engineering machinery guidance, precision farming, harbour fleet management, rescue and emergency services, and structural monitoring. By improving the availability of reference station data for users that demand high positioning accuracy (and integrity), *in real-time*, the number and variety of applications can grow rapidly. CORS networks are therefore critical infrastructure enabling the basic utility of high precision positioning to a diverse range of users.

Regional CORS networks are currently being established in many countries as a foundation for the establishment of a Spatial Data Infrastructure (Brown et al, 2002; Zhang et al, 2001; Rizos & Han, 2003). The distribution and density of a CORS network is constrained by the establishment costs per reference station, the areal extent to be serviced and the positioning accuracy requirements of the most demanding users (it is assumed that if the highest accuracy requirements are satisfied, all less demanding requirements can be easily met). Ideally, a dense network of reference receivers would be established and able to satisfy cm-level user accuracy requirements. Existing CORS networks in, for example, Germany, UK and Japan are sufficiently dense to restrict the maximum baseline length between a user and nearby reference station to well under 40km, which is often sufficient for cm-level accuracy techniques based on a single reference station, using high quality (dual-frequency) receivers that permit rapid "ambiguity resolution" (AR). However, as the inter-receiver distance increases, the *residual* atmospheric biases (due to differential ionospheric and tropospheric delay of the GPS satellite signals) in the *double-differenced* GPS observables (the standard observation model used for carrier phase-based positioning – Dedes & Rizos, 2002) increases, making AR more difficult (and even impossible using current rapid positioning techniques). Hence this distance constraint for rapid AR makes accurate positioning with respect to sparse CORS networks problematic. *This has profound ramifications in Australia due to its large areal extent and relatively sparse population*.

The aim of this project is to enhance the utility of CORS networks in Australia, by developing new GPS atmospheric correction models to support cm-level, real-time positioning even when the density of reference stations is insufficient for standard operational techniques.

2. Status of CORS networks in Australia

In the early 1990s, the Australian Regional GPS Network was established to support geodetic applications, and now consists of 16 reference stations on mainland Australia (with an average spacing of one thousand kilometres).

A CORS network to support surveying, mapping and high-end navigation users was first established in 1994 when Victoria's *GPSnet* was deployed (Hale, 2000). It has since grown to 19 reference stations with a regional and rural network coverage and a denser Melbourne and Environs network providing online GPS data access (www.land.vic.gov.au/GPSnet). Users can combine these files with GPS data collected across Victoria, for post-processing to obtain cm-level accuracy position results (see Figure 1). *To date, this is the only state-wide CORS network in Australia*.

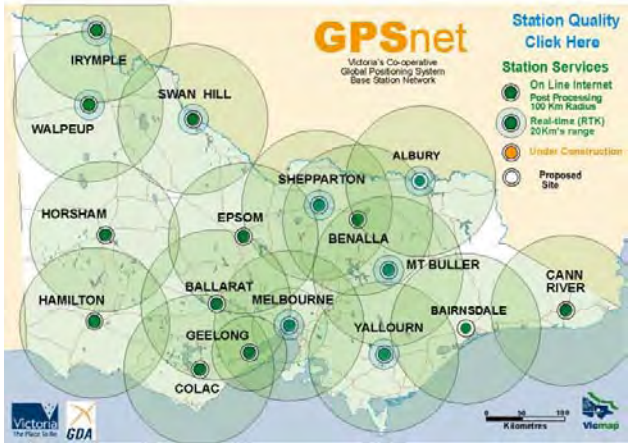


Figure 1 Current Status of Victoria's Rural and Regional GPSnet network infrastructure

CORS networks around the world will be progressively upgraded to provide a real-time capability for high-end users (Rizos, 2002b). This is also true of Victoria's *GPSnet*, with eight stations currently broadcasting (single-reference) RTK corrections to local users. However, the distance-dependent atmospheric biases referred to earlier, restrict the length of GPS baselines, thereby limiting the applications of RTK techniques to within 10-15km of a broadcasting reference station. As the average reference station spacing in *GPSnet* is of the order of 100km (except in the Melbourne area where the spacing decreases to about 50km), there remain large areas of Victoria where RTK cannot be used (Roberts et al, 2003).

Ibid (2003) proposes developing *GPSnet* from a "passive" CORS network, that is where users must download data from an archive site for post processing, to an "active" network (data is broadcast to users) in a three stage process. The first stage is already evolving and entails simply broadcasting base station data via UHF/VHF radio as standalone RTK units to local users within a range of up to 20km.

Stage two is part of a new Location Based Services (LBS) Positioning System project currently underway and being implemented by Spatial Information Infrastructure of the Department of Sustainability and Environment. Six well distributed base stations will provide DGPS correction messages to users across Victoria thereby providing better than 0.5m horizontal positioning in real-time (see Figure 2). This service is planned to go live wirelessly to GPSnet cooperative hosts, contributors and partners in December 2004 and to other users by June 2005.

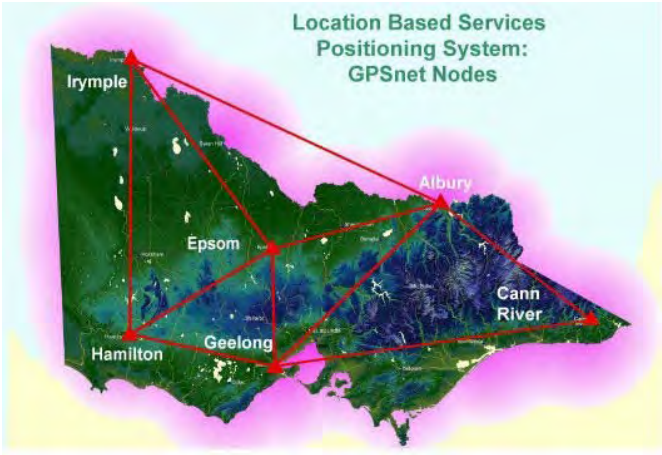


Figure 2 Victoria's GPSnet with LBS Positioning System stations generating a networked, statewide DGPS correction message which is then broadcast to users

The most challenging aspect of implementing such an Internet-based DGPS system is connecting base stations and the GPSnet Central Server Cluster in a manner that is technically and economically sustainable. Access for users will be via mobile Internet ie GPRS via Bluetooth™ to a GPS device. There will be up to 10 different types of user selectable solution types from DGPS RTCM through to networked CMR (Compact Measurement Record – Trimble Proprietary message format) and FKP emulation (surface correction parameter – FKP (translated from German)). The important aspect of this service is that the user makes the selection (Millner, 2004).

Ultimately by mid 2005 an estimated 21 *GPSnet* base stations will able to be connected to the GPSnet Central Server Cluster. Having stations connected to an active network will ease the transition to stage three: cm-level positioning across Victoria using full network RTK correction generation.

Since the mid-1990s much research has been invested in developing reference network techniques to *map* regional atmospheric conditions (and other unmodelled double-differenced biases). These network-based techniques effectively extend the baseline length to support *medium-range* RTK surveying, even up to 70km or more in mid-latitude areas (Rizos & Han, 2003). The distribution of *GPSnet* reference stations provides a suitable infrastructure to investigate network-based RTK over a significant regional extent.

Figure 3 Proposed coverage map for the SunPOZ network in SE Queensland

First generation commercial network-based RTK systems have been available over the last few years. A so-called *Virtual Reference Station* (VRS) test-bed network has been operating for several years in the Ipswich region, south of Brisbane (see Figure 3). This network, known as *SunPOZ*, uses Trimble's proprietary VRS™ software (and their hardware) to provide cm-level positioning in the region. However, this CORS infrastructure is not challenging the constraints that affect RTK, as the longest distance between reference and user receivers is of the order of only 30km (Higgins, 2002).

The latest CORS infrastructure in Australia is *SydNet*, currently being deployed in the Sydney basin area (see Figure 4). *SydNet* will consist of 10-15 reference stations

and will service the eastern seaboard from Wollongong to Newcastle (Rizos et al, 2003). Like the *SunPOZ* VRS network, *SydNet* will be a CORS network designed from the very beginning to support real-time operations, in contrast to *GPSnet*, which, due to its initial development from 1994, is evolving toward real-time operations. However, the high density of *SydNet* stations is only intended to support research into new network-based GPS positioning techniques. This is possible because of the ability to "thin" the CORS network, and to use some *SydNet* stations as *simulated* user receivers for check purposes. Hence, although *SydNet* is an ideal test network for the proposed research, it is not a blueprint for a state-wide CORS network. The extension of the GPS network across the state of New South Wales would be based on receiver spacings similar to those of Victoria's *GPSnet* (or even greater).

Figure 4 Current *SydNet* configuration

SydNet will provide network generated RTK corrections using the in-house software developed in cooperation with Nanyang Technical University (NTU), Singapore and currently in operation on the Singapore Integrated Multiple Reference Station Network (SIMRSN) (Goh, 2002). Corrections will be archived on an Internet server housed at the Network Control Centre (NCC) hosted by ac3 (host company) at the Australian Technology Park, Redfern. This is a secure site guaranteeing reliable operations for users. The base stations are located at the state government Rail Corporation owned train stations and linked to the NCC via the existing optical fibre network owned by the telecommunications provider Argus. In the first instance, users will access the correction messages from the Internet via a GPRS connection using CDMA, 3G or W-LAN. Other methods of broadcasting correction messages to users will be investigated as *SydNet* evolves. An initial seven station subset of *SydNet* will be launched at the GNSS2004 conference in Sydney.

GPSnet, *SunPOZ* and *SydNet* all provide a public utility function supported by their respective state government departments (Department of Sustainability and Environment, Dept of Natural Resources, Mines & Energy, Qld and Dept of Lands, NSW respectively). However only *SydNet* (and to a lesser extent *GPSnet*) will perform a dual function as a research test-bed and public utility. *SunPOZ* is a proprietary system and relies on the manufacturer (Trimble) for further advances and enhancements.

3. Motivation and Aims for Research Project

CORS networks have already been established in many countries, and these are likely to be upgraded to provide cm-level positioning accuracy to high-end users, even in real-time. One challenge is to make the spacing of reference receivers as great as possible, without sacrificing user accuracy and reliability. Most CORS networks currently offering network-RTK services have reference station spacings of a few tens of kilometres at most. Internationally, research is being undertaken to extend the station spacing, primarily through improved atmospheric (ionospheric and tropospheric) bias modelling. Achieving cm-level positioning accuracy over distances of 70km or more has not yet been verified on an operational basis (though university studies have demonstrated promising results for short trials with optimised research software). Using the commercially available Trimble VRS™ system, Higgins (2002) and Retscher (2002) have reported horizontal positioning accuracy of the order of ±5cm, for baselines up to 35km in length.

Other challenges include improving the stochastic models for network-based positioning, extending the data modelling algorithms to new satellite signals [i.e. "GPS modernization" (McDonald, 2002); and GALILEO (Galileo, 2003)], determining which is the "best" correction message formats, and outputting the bias modelling results in appropriate forms for non-positioning users. Australia has already made significant contributions to CORS research since the late-1990s (see Rizos & Han, 2003, for an overview), and with access to Victoria's *GPSnet* and NSW's *SydNet* further improvements to real-time network positioning techniques can be expected. The fundamental research facilities to be used in this project represent a significant infrastructure investment by the two government partners, Department of Sustainability and Environment and Dept of Lands, NSW. The research team has identified several research areas, primarily concerned with improving or extending current GPS atmospheric bias models to;

- Increase the inter-receiver distances for real-time (or post-mission), rapid GPS carrier phase-based positioning.
- Take advantage of the extra signals to be transmitted by the "modernized" GPS constellation (first satellites to be launched in 2005), and eventually the new GALILEO system (planned deployment to commence later in the decade).
- Permit more realistic network stochastic models to be provided to users, hence improving the quality or integrity information associated with high accuracy GPS positioning results.
- Develop tropospheric signal bias "maps" as a by-product of CORS data processing, so as to, for example, correct differential Interferometric Synthetic Aperture Radar (InSAR) results of ground subsidence/deformation.
- Demonstrate the feasibility of a CORS network to monitor the ionospheric and tropospheric "weather" conditions, so contributing to synoptic atmospheric studies.

4. Research Road Map

To achieve the aims of this project the following distinct research tasks have been identified:

1. Spatio-temporal models for ionospheric and tropospheric bias variability for incorporation into "second generation" network-RTK systems.
2. Validation of the aforementioned models, and techniques for interpolation of atmospheric biases at a user's location.
3. Stochastic models appropriate for network-based positioning (for real-time and post-mission modes of operation).
4. Extension of bias modelling for the additional signal frequencies to be transmitted by future Global Navigation Satellite Systems.
5. Investigations to determine the appropriate message format for transmitting such model information to users in real-time.
6. Tropospheric bias model products for non-positioning applications, such as for meteorological and remote sensing studies.

These research tasks are expanded as follows.

4.1 Models of Ionospheric and Tropospheric Bias Variability

Amongst the issues to be investigated for research activity into spatio-temporal models for ionospheric and tropospheric signal delay are:

- The applicability of current global International GPS Service (IGS) ionosphere and troposphere models;
- The applicability of ionospheric models derived from the U.S. Wide Area Augmentation System (WAAS);
- The variability of ionospheric delay determined using the dual-frequency instrumentation of GPSnet; and
- The use of GPSnet and SydNet data to estimate the variability of the tropospheric delay (primarily wet component) from long time series of GPS data.

The outcome will be improved atmospheric bias models appropriate for network-based positioning, as well as other non-positioning applications. To support this, the Department of Sustainability and Environment, Government of Victoria is currently making a significant capital investment in infrastructure for a real-time, network-positioning, wireless upgrade to *GPSnet*, through which these models can be implemented. This will have a critical impact on how state-wide CORS networks can be designed and implemented, in Victoria, NSW and other Australian states.

4.2 Validation of Atmospheric Models for Network-RTK

Validation of the spatio-temporal models for ionospheric and tropospheric signal delay developed in Sect 4.1, as well as testing and developing interpolation algorithms will encompass such activities as:

- The use of GPSnet to validate the ionospheric models using data from extra dual-frequency GPS receivers placed at interpolation (check) points;
- The use of SydNet to validate the tropospheric models using data from selected interpolation (check) points (as the SydNet will be dense enough to model horizontal tropospheric inhomogenieties), and
- The investigation of appropriate interpolation techniques (such as those identified by Dai et al, 2003).

The outcome will be validated atmospheric bias models appropriate for network-based positioning, which can be tested within network-RTK software such as that currently being implemented in *SydNet*. (*SydNet* will use the network-based positioning algorithms jointly developed by UNSW and Nanyang Technological University, and implemented in the Singapore Integrated Multiple Reference Station Network – Chen et al, 2000; Rizos, 2002b).

4.3 Stochastic Models for Network-Based Positioning

Research activity into stochastic modelling will investigate:

- The impact of neglecting rigorously-derived stochastic information on both network data processing and user position determination;
- Appropriate simplified stochastic models; and
- The "packaging" of such stochastic information into suitable network-RTK messages.

The outcome will be proposals for how such improved stochastic modelling can be implemented within operational network-RTK software, with initial tests to be carried out on *SydNet* with the modified network-RTK software described by Chen et al (2000). This work will be based around some initial research in this area reported by Musa et al (2003).

4.4 Atmospheric Bias Modelling for Next Generation GNSS

This work will concentrate primarily on research activity into ionospheric delay modelling for next generation GNSS, and will investigate:

- The combinations of triple-frequency signals from the modernized GPS's L1, L2, L5, and GALILEO's multiple-frequencies that are best suited for determining ionospheric bias information, and
- The appropriate "transition" arrangements that can be made to continue operations in current (and upgraded) CORS networks as new satellites are launched.

Much of this research will be undertaken using simulated data generated using simple error models as implemented in the Bernese geodetic software, Matlab-type tools as well as via signal simulators. The outcome will be proposals for how network-based positioning will "evolve" as the new GNSS satellites are launched from 2005 through to the end of this decade.

4.5 Algorithms for Network-RTK Message Formats

This task will investigate how the network-derived correction data will be "packaged" for users. There is considerable debate concerning the RTCM-type messages for network-RTK (Rizos & Han, 2003), and there are several candidates. Amongst the formats to be investigated are:

- The Virtual Reference Station (VRS)-based approach of user-customised reference station data (as implemented within Trimble's VRS™ system), that utilises the user's location;
- The Virtual Reference Cell (VRC) approach that utilises the user's approximate location (Retscher, 2002);
- The "Area Correction Parameter" (surface correction parameter - FKP) approach sends out the entire correction model (Wübbena & Bagge, 2002), permitting the user to then perform the bias interpolation themselves; and
- Sending all raw data to the user, and allowing them to select subsets of the data for integrity checking.

The outcome will be recommendations to the "best" approach, as well as a suite of algorithms from which the industry-endorsed format can be implemented within the network-RTK user software.

4.6 Tropospheric Models for Non-Positioning Applications

Research activity into tropospheric model products will investigate:

- Areal models of tropospheric delay correction (for the wet and dry components) for differential InSAR (DInSAR), and
- Time-sliced (synoptic) areal models for possible assimilation into meteorological systems.

The outcome will be a methodology for mapping the tropospheric delay from CORS networks, so that it can be used to correct DInSAR-derived ground subsidence maps described in Janssen et al (2004). Tests of the tropospheric delay maps derived from GPSnet data on DInSAR results for areas of the Gippsland that are undergoing subsidence due to offshore oil extraction. Research is currently underway in this area. The effect of neglecting differential tropospheric signal delay in DInSAR results is being assessed (Ge et al, 2001).

3. Concluding Remarks

CORS networks in other countries are sufficiently dense to restrict the maximum baseline length between a user and nearby reference station to well under 40km. Under typical operational conditions this is usually adequate for cm-level positioning accuracy. Due to Australia's large areal extent and relatively sparse population, the number of reference stations required for a state-wide network using existing techniques would therefore be uneconomical. One option would be to, rather than offer a state-wide service, provide the necessary station density within regions with high population densities. However, in Australia some of the most important users of satellite positioning technology are in precision agriculture and mining applications, and require infrastructure across large rural areas. This research is therefore of national significance as it will provide an infrastructure that benefits activities for rural and regional communities. To service these users adequately the state governments need to implement the outcomes proposed in this research project in order to capitalise on sparse CORS infrastructure.

The primary outcome of the proposed research will be enhanced CORS operational algorithms and procedures, extending the capabilities of current commercially available network-based positioning systems such as VRS. Such a refinement will lead to "second generation" CORS networks, with sparser spacing of reference stations than current "first generation" VRS systems, without sacrificing performance (accuracy, time-to-AR, reliability, availability and integrity).

Acknowledgements: The authors would like to gratefully acknowledge the research funds granted for this project through the Australian Research Council Linkage project LP0455170.

References

Brown, N., Kealy, A. & Williamson, I. (2002***) Issues related to GPS phase observations for improved quality estimation***, The Journal of Cartography, Australia, 31(2), 143-151.

Chen, X., Han, S., Rizos, C. & Goh, P.C. (2000) ***Improving real-time positioning efficiency using the Singapore Integrated Multiple Reference Station Network (SIMRSN)***, 13th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Salt Lake City, Utah, 19-22 September, 9-18.

Dai, L., Han, S., Wang, J., Rizos, C. (2002) ***Comparison of interpolation algorithms in network-based GPS techniques***, Journal of Navigation, 50(4), pp 277 – 293.

Dedes, G. & Rizos, C. (2002) ***GPS positioning models for single point and baseline solutions***, Chapter 9 in Manual of Geospatial Science and Technology, J. Bossler, J. Jenson, R. McMaster & C. Rizos (eds.), Taylor & Francis Inc., ISBN 0-7484-0924-6, 114-126.

Galileo (2003) http://europa.eu.int/comm/dgs/energy_transport/galileo/index_en.htm

Ge, L., Chen, H.Y., Han, S. & Rizos, C. (2001***) Integrated GPS and Interferometric SAR techniques for highly dense crustal deformation monitoring***, 14th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Salt Lake City, Utah, 11-14 September, 2552-2563.

Goh, P.C. (2002) ***Development of an Integrated Multiple Base Station Infrastructure to support Concurrent High Precision Differential Positioning Applications***, Nanyang Technical University, Internal Report.

GPSnet (2004) http://land.vic.gov.au/GPSnet

Hale, M. (2000) ***GPSnet: The next decade***, paper presented at the Survey Conference of the Institution of Surveyors Victoria and the Institution of Engineering Surveyors Australia, Victoria Division, 27 May.

Higgins, M. (2002) ***Australia's changing surveying infrastructure from marks in the ground to virtual reference stations***, FIG XXII International Congress, Washington, 19-26 April, TS5.6.

Janssen, V., Ge, L. & Rizos, C. (2004) ***Tropospheric corrections to SAR interferometry from GPS observations***, GPS Solutions, 8(3), 140 – 151.

McDonald, K. (2002) ***The Modernization of GPS: Plans, new capabilities and the future relationship to Galileo***, Journal of Global Positioning Systems, 1(1), 1-17.

Millner, J. (2004) ***GPSnet Location Based Services Positioning System Project Manager***, Personal communication

Musa, T.A., Wang, J., Rizos, C. & Satirapod, C. (2003) ***Stochastic modelling for network-based GPS positioning***, 6th Int. Symp. on Satellite Navigation Technology Including Mobile Positioning & Location Services, Melbourne, Australia, 22-25 July, CD-ROM proc., paper 41.

Retscher, G. (2002) ***Accuracy performance of virtual reference station (VRS) networks***, Journal of Global Positioning Systems, 1(1), 40-47.

Rizos, C. (2002a) ***Making sense of the GPS techniques***, Chapter 11 in *Manual of Geospatial Science and Technology*, J. Bossler, J. Jenson, R. McMaster & C. Rizos (eds.), Taylor & Francis Inc., ISBN 0-7484-0924-6, 146-161.

Rizos, C. (2002b) Network ***RTK research and implementation: A geodetic perspective***, Journal of Global Positioning Systems, 1(2), 144-150.

Rizos, C., Kinlyside, D., Yan, T., Omar, S. & Musa, T.A. (2003) ***Implementing network RTK: The SydNet CORS infrastructure***, 6th Int. Symp. on Satellite Navigation Technology Including Mobile Positioning & Location Services, Melbourne, Australia, 22-25 July, CD-ROM proc., paper 50.

Rizos, C. & Han, S. (2003) ***Reference station network based RTK systems - Concepts & progress***, Wuhan University Journal of Nature Sciences, 8(2B), 566-574.

Roberts, C., Zhang, K., Hale, M. & Millner, J. (2003) ***Towards real-time network-based RTK corrections for Victoria's GPSnet***, 6th Int. Symp. on Satellite Navigation Technology Including Mobile Positioning & Location Services, Melbourne, Australia, 22-25 July.

Wübbena, G. & Bagge, A. (2002) ***RTCM Message Type 59–FKP for transmission of FKP, Version 1.0***, Geo++ White Paper Nr. 2002.01, www.geopp.de/download/geopp-rtcm-fkp59.pdf

Zhang, K., Talbot, N., Hale, M. & Millner, J. (2001) ***Victorian high precision permanent GPS tracking network system***, 14th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Salt Lake City, Utah, 11-14 September, 3077-3085.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 226-231

Journal of Global Positioning Systems
2004
© CPGPS

High Frequency Deflection Monitoring of Bridges by GPS

Gethin W Roberts, Emily Cosser, Xiaolin Meng, Alan Dodson

IESSG, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK
e-mail: gethin.roberts@nottingham.ac.uk Tel: + 44 115 9513933; Fax: +44 115 9513881

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. The use of GPS for the deflection and deformation monitoring of structures has been under investigation for a number of years. Previous work has shown that GPS not only measures the magnitude of the deflection of the structure, but also it is able to measure the frequency of the movement. Both sets of information are useful for structural engineers when assessing the condition of the structure as well as evaluating whether Finite Element (FE) models of such structures are indeed correct. GPS has the advantage of resulting in an absolute 3-D position, with a very precise corresponding time tag. However, until recently, the maximum data rate was typically 10-20 Hz, meaning that the maximum detectable frequency was about 5-10 Hz. GPS also has the disadvantage of multipath and cycle clips, and the height component's accuracy is typically 2 – 3 times worse than that of plan. Previous work at the IESSG has included the integration of RTK GPS, gathering data at a rate of up to 10Hz, with that of data from an accelerometer, typically gathering data at up to 200 Hz. Accelerometers tend to drift over time, and can not detect low vibration frequencies, but the acceleration data can be double integrated resulting in changes in positions. The integration of GPS and accelerometers can help to overcome each others' shortfalls. This paper investigates the use of high rate carrier phase GPS receivers for deflection monitoring of structures. Such receivers include the Javad JNS100, capable of gathering data at up to 100 Hz. Static trials have been conducted to investigate the precision of such a receiver, as well as the potential applications of such a high data rate. Trials were carried out in a controlled environment and actual bridge monitoring, and comparisons made with a Leica SR510 receiver.

Key words: Bridge Deflection Monitoring, Data Sampling Rate, RTK GPS, Data Processing.

1 Introduction

In 2001, The University of Nottingham was awarded a three year grant from the UK's Engineering and Physical Sciences Research Council (EPSRC). The overall objective of this project is the creation of a system employing advanced computational tools coupled with GPS and accelerometer sensors able to remotely monitor the health of operational bridges without on-site inspection. For the field measurements being able to validate a computational model, such as a Finite Element (FE) model, number and locations of sensors, sampling rate and positioning accuracy are the indexes needed to be considered. The first factor is normally determined by the civil engineers, according to the size and design of the monitored structure. The second and third indexes are also related to the size and design of the structure but should be decided by the surveying engineers through choosing appropriate surveying instruments. For a small bridge of several tens of metres in length, the amplitude and vibration frequency of the vertical movements can be a couple of millimetres and tens of Hz. For a large bridge, such as the Humber Bridge, the amplitude and vibration frequency in the vertical direction can be of the order of up to a meter and tenth of Hz, respectively (Meng et al. 2003). To date the highest GPS data rate used in experiments has been 10 and 20 Hz, which means that only bridge dynamics of lower than 10 Hz could be detected, taking other error budgets into account.

To overcome the abovementioned shortfall, integration of GPS with triaxial accelerometers has been investigated in a post-processing way. This approach can significantly expand the valid measurable frequency to higher than 100 Hz. However, the absolute positioning fixes have to be provided by the GPS solutions and there is lack of real-time data transmission approach and relevant algorithm to integrate data from two kinds of sensors.

Two JNS100 GPS OEM boards were recently purchased from Javad Navigation Systems (Javad Navigation

Systems, 2004), which are able to output raw data and positions 100 times a second without interpolation (Figure 1). These GPS receivers can be used to measure bridge movements and also identify frequency dynamics not higher than 50 Hz, due to Nyquist theorem.

Figure 1 The JNS100 OEM board GPS receiver

This paper investigates the use of these high rate code/carrier phase GPS receivers for deflection monitoring of structures. Zero baseline, short baseline and kinematic trials have been conducted to assess the precision of such a receiver, as well as the potential applications of such a high data rate. These trials were carried out in a controlled environment as well as for a real bridge monitoring, and comparisons are made with Leica SR510 single frequency GPS receivers gathering data at a sampling rate of 10 Hz.

2 Evaluation of receivers' noise levels in static status: zero baseline (ZBL) and short baseline (SBL) tests

The raw code and carrier phase data are output from the receiver to a laptop and recorded using software called PCView. The raw data is automatically converted to Rinex format for post-processing. When the receiver outputs data at 100 Hz there were data overrun problems first on the serial port and then also on the USB port. Due to this the data collected for this paper was only recorded at a 50 Hz data rate, which is still fast enough to measure much higher frequency structural dynamics than has ever been possible with GPS before. The data overrun problems are currently being investigated and using these receivers at 100 Hz data rate will be the subject of future papers.

The JNS100 receivers record code and carrier phase data, only on the L1 frequency. Software for processing single frequency data in the context of bridge monitoring has been developed by the authors. This software, called Kinpos, was used to process all the GPS data for this paper. For more information about the software development please see Cosser et al. (2004) or Cosser (2004).

2.1 Zero baseline test

Two separate zero baseline trials were conducted on two consecutive days with the JNS100 receivers used on the first day and the Leica SR510 single frequency receivers used on the next. The receivers recorded at the same times on the two days, but offset by 4 minutes, so that they would be recording data with the same satellite geometry. On both days the two receivers of same type were connected by a splitter to the same antenna, a Leica AT503 choke ring antenna which was located on the roof of the IESSG building. The aim was to compare the data from the Leica receivers and JNS100 receivers under similar observation conditions. The Leica dual and single frequency GPS receivers had been used during many bridge trials in the past and their applicability to bridge monitoring is known.

The JNS100 receivers were always set up to record at a 50 Hz data rate for all the trials outlined in this paper. In the Kinpos software the data was then processed at a 50 Hz data rate and also resampled before processing to 10 Hz so that it could be directly compared to the Leica data. The standard deviation of the JNS100 coordinates appears greater for the 10 Hz data than for the 50 Hz data. In each case the spread of the data is the same, but a lower standard deviation is recorded for the 50 Hz data as there are more sample points.

The processed coordinates in WGS84 are then converted to those in a local coordinate system. The standard deviations of the east, north and vertical components for the Leica and JNS100 receivers can be seen in Table 1. For a fairer comparison the Leica data are compared only with the resampled 10 Hz JNS100 data. It can be seen that the Leica data has a lower standard deviation in every component when compared to the JNS100, with the largest difference being seen in the vertical direction. Figure 2 shows the time series of vertical coordinate error of the Leica and JNS100 data at 10 Hz, while comparing to the true coordinates. It is clear from this graph and from Table 1 that the Leica receiver has a smaller spread of coordinates in the vertical direction. This implies that there is a better resolution of the carrier phase by the Leica receivers. However, the noise levels caused by the receivers are less than 1 cm for all three directions and both types of GPS receivers, demonstrating the appropriateness of these receivers for high precision applications.

Table 1. Standard Deviations of JNS100 and Leica Receivers from ZBL

	Standard Deviations (m)		
	East	North	Vertical
JNS100 (50 Hz)	0.0018	0.0023	0.0034
JNS100 (10 Hz)	0.0019	0.0021	0.0041
Leica (10 Hz)	0.0013	0.0017	0.0029

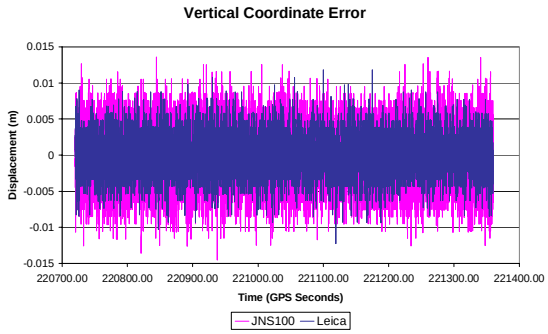


Figure 2. Time Series of Vertical Error for JNS100 and Leica Receivers

2.2 Short baseline tests

A short baseline test is a truer representation of survey conditions and so the performance of the receivers in practice can be better assessed. Atmospheric errors and clocks are still mitigated, but multipath is now present in the solution.

A short baseline trial was conducted on The University of Nottingham campus during July 2004. Two AT503 antennas were positioned on two established points, the coordinates of which were known from previous static surveys. The two points were roughly 50 metres apart. At each end of the baseline, a JNS100 receiver and a Leica SR510 receiver were connected by a splitter to the same antenna, meaning that the baselines measured by each receiver combination were the same.

The baselines for this trial were processed in Kinpos and the results can be seen in Table 2 and Figure 3. It can be seen from Table 2 that once again the standard deviations in all three components are lower for the Leica receivers, the largest difference being in the east component, at 1.2mm, demonstrating slightly higher multipath in East-West direction. Figure 3 shows the time series of vertical coordinate error for the Leica receivers and the JNS receivers at 10 Hz. The systematic bias of multipath is now visible within the data and follows the same pattern with slightly different amplitudes for both receiver pairs.

From Figure 3, it can be found that to improve the positioning precision, multipath need to be mitigated either using appropriate mitigation algorithm or through an internal filter of the receiver hardware and a choke ring antenna. Dodson et al. (2001) investigated the use of an adaptive filtering technique for reducing the impact of multipath for structural deformation monitoring.

Table 2. Standard Deviations of JNS100 and Leica Receivers from SBL

	Standard Deviation (m)		
	East	North	Vertical
JNS100 (50 Hz)	0.0037	0.0056	0.0064
JNS100 (10 Hz)	0.0037	0.0056	0.0067
Leica (10 Hz)	0.0025	0.0050	0.0057

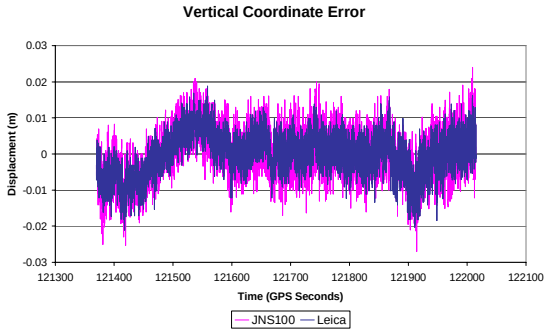


Figure 3. Time Series of Vertical Error for JNS100 and Leica Receivers

3. Evaluation of receivers' noise level in dynamic status: platform and bridge trials

3.1 Platform test

To test the potential of the JNS100 receivers in a dynamic environment, a platform was set up on The University of Nottingham campus (Figure 4). A wooden frame was suspended from a tall tripod by means of a bungee cord, which allowed free oscillation of the platform. The reference receiver was located approximately 10 metres away from the test rig, where an AT503 antenna was connected via a splitter to the Leica SR510 and JNS100 receivers. An AT502 navigation antenna was mounted on the test rig, which was then, via a splitter, connected to the JNS100 and Leica SR510 receivers.

Using the test rig, two different trials were conducted. For the first test, the platform was in rotation either held still or disturbed from its resting position by someone forcing the platform to move up and down. For the second trial, the platform was just left to swing.

The first trial was conducted over a 10 minute time period, where the bungee platform was held still for two minutes and then made to oscillate for 2 minutes and so on in rotation. The results for this trial for the JNS receiver measuring at 50 Hz and resampled at 10 Hz, and for the Leica receiver measuring at 10 Hz can be seen in Figure 5. The amplitude of oscillation of the bungee platform was measured as between 15 and 20 cm by both GPS receivers. The JNS receiver has a period within the last two minutes where there are a number of jumps within the time series, which are caused by undetected cycle slips. Apart from these jumps the measured

displacement is very similar for both receivers. This demonstrates the capability of the JNS receivers to measure in a dynamic environment.

Figure 4 Platform for a Dynamic Test Using JNS100 and Leica receivers

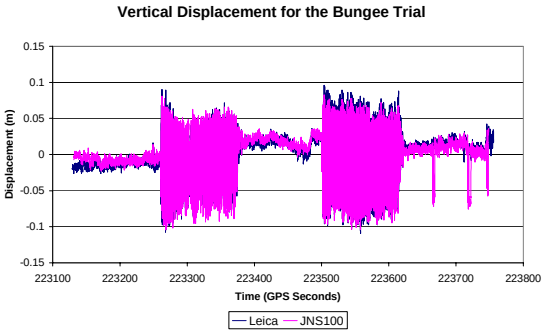


Figure 5 Time Series of the Vertical Displacement by JNS100 and Leica receivers

In the second platform trial the bungee was just left to swing with the wind. The results for this trial for the east, north and vertical coordinates measured by the Leica and JNS100 receivers can be seen in Table 3. In the trials, the sampling rates for JNS100 and Leica receivers were set to 50 Hz and 10 Hz, respectively. For this trial the results for both types of receiver match well, with the standard deviations in the vertical and north component actually being slightly better for the JNS100 receiver. Figure 6 shows that the multipath characteristics displayed by both receiver solutions, in the vertical direction, are the same. This is an encouraging result for the JNS receiver, showing that in this dynamic environment they can measure to the same degree of precision as survey grade GPS receivers.

Table 3. Standard Deviations of JNS100 and Leica Receivers from Platform Test

	Standard Deviations (m)		
	East	North	Height
JNS100 (50 Hz)	0.0074	0.0078	0.0113
JNS100 (10 Hz)	0.0074	0.0078	0.0115
Leica (10 Hz)	0.0074	0.0079	0.0118

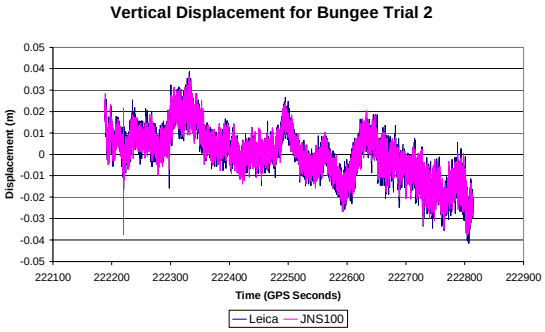


Figure 6 Time Series of the Vertical Movement by JNS100 and Leica receivers

4 Bridge trial

A GPS and accelerometer bridge trial was conducted on the Wilford Suspension Footbridge in Nottingham, over two days in July 2004 (6th and 7th). This bridge has been the focus of many trials conducted by The University of Nottingham, due to its proximity and relatively large amplitude movements. For more information on previous trials conducted on the Wilford Bridge, see for example Roberts et al. (2001) The purpose of this trial was to analyse the performance of the JNS100 receivers in a bridge environment.

In this trial, one reference station was set up on the bank of the river, on a point whose coordinates were well established from previous trials (Figure 7). The rover receiver was located at the mid span of the bridge, where the most movement is expected (Figure 8). At both locations an AT503 antenna was connected via a splitter to both the JNS100 and Leica single frequency receivers. A number of sessions of data were collected on each day, a selection of which will be analysed.

The GPS results for first session on the 7th July, which was the second day of the trial, can be seen in Table 4. It contains the standard deviations of the east, north and vertical components for the JNS100 and Leica receivers. For this particular session, the JNS100 receivers actually performed better than the Leica in all three component directions, the largest difference being seen in the north direction. Both receivers were seeing exactly the same satellites. The difference in standard deviations in the vertical direction was very small as the same multipath

pattern can be seen in both times series (Figure 9). For all the sessions during the bridge trial, the results from the JNS100 and Leica receivers were very similar. In some cases the JNS100 was slightly more accurate that the Leica and in some cases this was the other way around. The difference between the two receivers in all cases was very small, showing that in the bridge environment the performance of the JNS100 is comparable with the Leica receivers, even at a much higher sampling rate.

Figure 7. Reference station

Figure 8. Rover station

Table 4. Standard Deviations of JNS100 and Leica Receivers from Bridge Trial

	Standard Deviation		
	East	North	Vertical
JNS100 (50 Hz)	0.0025	0.0029	0.0043
JNS100 (10 Hz)	0.0025	0.0029	0.0045
Leica (10 Hz)	0.0027	0.0036	0.0046

Also, in the bridge trials the GPS results are compared to a closely located triaxial accelerometer measuring at 50 Hz as well. The periods of the largest movement seen in Figure 10 correspond to times in which people on the bridge jumped up and down in unison 'forcing' the bridge to move and then left to oscillate at its natural frequency. In this graph the forced movement is apparent in both the accelerometer and JNS100 data. When the forced movement stops the accelerometer displays a sinusoidal decay, which is movement at the bridge's natural frequency. This sinusoidal decay is not clear in the GPS data as it is masked by the noise. However, frequency analysis reveals that that this sinusoidal pattern is still present in the GPS data even though it cannot be discerned by the eye (Meng et al., 2004).

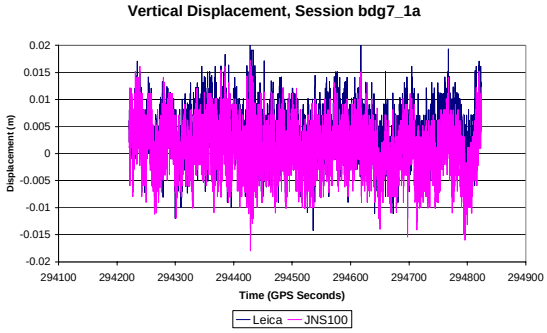


Figure 9. Vertical displacement by JNS100 Resampled to 10 Hz and the Leica receivers

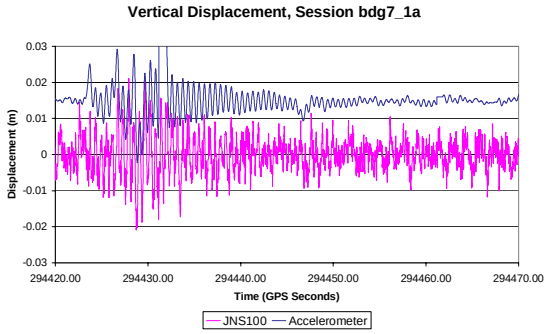


Figure 10. JNS100 and accelerometer displacement in the vertical direction both recorded at 50 Hz. The graph focuses on a time where there was the largest movement on the bridge. The accelerometer data is offset by 0.015m

4 Conclusions

This paper has outlined the preliminary work conducted with the JNS100 receivers. Zero baseline and static short baseline trials have been conducted to assess the precision of the receivers compared to known high quality survey grade receivers (Leica system 500 single frequency receivers). The results showed that the Leica receivers performed slightly better than the JNS100 in the static trials, but the difference was small. The JNS100 receivers do have a high precision carrier phase observables.

Kinematic trials were performed on a bungee test rig and also on a bridge. In a dynamic situation the JNS100 receivers performed as well as the Leica receivers. The JNS100 results, measured at 50 Hz, were also compared

to those from a closely located triaxial accelerometer measuring at the same data rate.

JNS100 bridge trial results compared well to the accelerometer findings, when identifying the periods of largest movement. Most movement on the bridge was masked by the GPS noise, but periods where large displacements occurred could be discerned.

Acknowledgements

This project has been funded by the UK's Engineering and Physical Sciences Research Council (EPSRC) in conjunction with Cranfield University (ref no GR/R28218/01). The authors would like to thank Louise Arrowsmith who took part in the field work whilst working towards her MSc.

References

Cosser E (2004) **Bridge Deflection Monitoring and Frequency Identification with Single Frequency GPS Receivers**, ION GPS 2004, The 17th Technical Meeting of the Satellite Division of the Institute of Navigation, 21-24 September, Long Beach, California.

Cosser E, Roberts G W, Meng X and Dodson A H (2004) **Single Frequency GPS for Bridge Deflection Monitoring: Progress and Results**, 1st FIG International Symposium on Engineering Surveys for Construction Works and Structural Engineering, 28 June-1 July, Nottingham, England.

Dodson A H, Meng X and Roberts G W (2001) **Adaptive Methods for Multipath Mitigation and Its Applications for Structural Deflection Monitoring**, International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS 2001), 5-8 June, Banff, Canada.

Javad Navigation Systems (2004) **OEM boards**, [online]. Available at: <http://www.javad.com/> [Accessed 12 August 2004].

Langley R (1997) **GPS Receiver System Noise**, GPS World 8(6): 40-45.

Meng X (2002) **Real-time Deformation Monitoring of Bridges Using GPS/Accelerometers**, PhD thesis, The University of Nottingham, Nottingham, UK.

Meng X, Meo M, Roberts G W, Dodson A H, Cosser E, Iuliano E and Morris A (2003) **Validating GPS Based Bridge Deformation Monitoring with Finite Element Model**, GNSS 2003, 22-25 April, Graz, Austria.

Meng X, Roberts G W, Dodson A H, Andreotti M, Cosser E and Meo M (2004) **Development of a Prototype Remote Structural Health Monitoring System (RSHMS)**, 1st FIG International Symposium on Engineering Surveys for Construction Works and Structural Engineering, 28 June-1 July, Nottingham, England.

Roberts G W, Meng X and Dodson A H (2001) **The Use of Kinematic GPS and Triaxial Accelerometers to Monitor the Deflections of Large Bridges**, 10th International Symposium on Deformation Measurement, FIG, 19-22 March, California, USA.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 232-241

Journal of Global Positioning Systems
2004

A Performance Analysis of Future Global Navigation Satellite Systems

Cedric Seynat[1], Allison Kealy[2], Kefei Zhang[3]

[1]GPSat Systems Australia, Suite 1/22 Aberdeen Road, McLeod, Victoria,
email: cedric.seynat@gpsatsys.com.au, Tel: +61 (0)3 9455 0041 Fax: +61 (0)3 9455 0042

[2]Department of Geomatics, The University Of Melbourne, Victoria, Australia
email: akealy@unimelb.edu.au, Tel: +61 (0)3 8344 6804 Fax: + 61 (0)3 9347 2916

[3]School of Mathematical and Geospatial Sciences, RMIT University, Victoria, Australia
email: kefei.zhang@rmit.edu.au, Tel: +61 (0)3 9925 3272

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. For an increasing number of applications, the performance characteristics of current generation Global Navigation Satellite Systems (GNSS) cannot meet full availability, accuracy, reliability, integrity and vulnerability requirements. It is anticipated however that around 2010 the next generation of GNSS will offer around one hundred satellites for positioning and navigation. This includes constellations from the US modernised Global Positioning System, the Russian Glonass, the European Galileo, the Japanese Quasi-Zenith Satellite System and the Chinese Beidou. It is predicted that the performance characteristics of GNSS will be significantly improved. To maximise the potential utility offered by this integrated infrastructure, this paper presents an approach adopted in Australia to quantify the performance improvements that will be available in the future. It presents the design of a GNSS simulation toolkit developed in Australia and the performance expectations of future GNSS for a number of important applications within the Asia Pacific region. In quantifying the improvement in performance realised by combined systems, this paper proposes a practical approach to facilitate the development of innovative applications based on future GNSS.

Key words: GPS, Galileo, GNSS, Simulation

1 Introduction

Currently, there are only two satellite navigation systems in operation, the Global Positioning System (GPS), and the Russian equivalent Glonass. The GPS signal is free but its availability is not guaranteed and currently most users are prepared to accept this risk. However, as satellite navigation becomes a vital technology across a number of critical industrial sectors, the prospect of, for example, a nation's transport infrastructure becoming dependent on this technology is a strategic risk that most industrial countries are not willing to accept. This argument initiated the Galileo programme in Europe. Galileo is a Global Navigation Satellite System (GNSS) and 30 satellites orbiting the Earth at an altitude of 23,616km (three spares) will transmit a navigation signal that can be received almost anywhere, and from which a receiver can determine its position and time. Unlike GPS, Galileo will also offer a guarantee of service to users who are willing to pay for it (e.g. commercial service – CS, and Public Regulated Service PRS) in addition to a free signal similar to that of GPS (Open Service - OS and Safety of Life service - SoL). Galileo will be available to the public in 2010 (European Commission, 2003).

Despite many technical differences between these three GNSS systems, the commonality of the carrier frequencies they use creates the potential for the future development of an interoperable GNSS receiver, as illustrated in Table 1, which compares the services available and associated signals both now and at around 2015. Aside from these three core GNSS infrastructures, a number of additional space-based navigation systems are also under development through various national programmes. Japan is currently developing the Quasi-Zenith Satellite System (QZSS), with three satellites placed in a special orbit that maximises coverage over Japan. The QZSS will complement other existing GNSS systems over Japan, but at the same time, these satellites will also be available over Australia and the South East Asian region (Petrovski 2003). In China, the Beidou

navigation system is also being developed. Current Beidou satellite navigation and positioning system consists two geosynchronous satellites based on the DFH-3 bus. There shall be four satellites, two operational and two backups upon completion of the system (Chinese Defence Today, 2004). In addition, many augmentations to GNSS are either under development, or have been commissioned, or are already in use: Space Based Augmentation Systems (SBAS) have been built (or are being deployed) by the US, Europe, India, Japan, and China. Ground-Based Augmentation Systems (GBAS) offer tremendous performance benefits to the aviation sector and have led to the development of the American Wide Area Augmentation System (WAAS), Local Area Augmentation Systems (LAAS) and the Australian Ground-Based Regional Augmentation System (GRAS).

This brief overview of current and future navigation infrastructures illustrates the huge potential that exists for future navigation and positioning applications. The vast majority of the world will be users of these existing systems. The fundamental questions then are: "Which system or systems should a country use?"; "How to choose a combination of the systems?"; "What are the benefits and respective merits of those systems?". There is no simple answer to these questions, as the best solution will undoubtedly depend on the targeted application, which has its own requirements in terms of accuracy, reliability, robustness, cost and other application-specific criteria. What can be provided, however, is a means whereby parameters that describe these performance requirements can be computed.

High-accuracy software simulations are a cost-effective and precise approach of determining the performance characteristics attainable from the future GNSS, and have been recognised as an appropriate pre-development tool for satellite navigation systems and applications in Japan (Petrovski 2003) and Europe (Seynat 2003). The technical benefits of this approach lie in the fact that the simulations are reproducible and totally controlled, and parameters can be changed individually if necessary for an in-depth understanding of the underlying effects.

This paper introduces a simulation toolkit developed to conduct a qualitative assessment of the performance characteristics of the future GNSS infrastructure. The design and development of the simulator architecture as well as the models describing all influencing effects on the performance of GNSS are presented. Finally, representative results over Australia are demonstrated and future developments are outline.

Table 1: Current and future GPS, Glonass and Galileo services and signals for 2004 and planned for 2015

	GPS		Glonass		GALILEO	
Services	**2004**	**2015**	**2004**	**2015**	**2004**	**2015**
Basic Positioning (unencrypted)	**SPS** ✓ L1 CA	**SPS** ✓ L1 CA ✓ L2C ✓ L5	**SP** ✓ L1	**SP** ✓ L1 ✓ L2 ✓ 3rd Signal		**OS** ✓ L1 ✓ E5a ✓ E5a
Integrity/safety (unencrypted)				Integrity message		**SoL** ✓ L1 ✓ E5b ✓ E5a
Commercial/value-added (encrypted)						**CS** E6
Security/military (unencrypted)	**PPS** ✓ L1 P(Y) ✓ L2 P(Y)	**PPS** ✓ L1 P(Y) ✓ L2 P(Y) ✓ L1 M ✓ L2 M	**HP** ✓ L1 ✓ L2	**HP** ✓ L1 ✓ L2 ✓ Unknown		**PRS** ✓ L1 ✓ E6
SPS-standard positioning service, **PPS** – precise position service, **SP** – standard precision, **HP**-high precision, **OS** – open service, **SoL** – safety of life service, **CS** – commercial service, **PRS** – public regulated service						

2 Simulation technology and methodology

Current and future users of navigation technologies need to understand and quantify the performance they can expect from the candidate systems, used individually or in combinations. The GNSS Simulation Tool (GST) developed in this research aims to provide a toolkit that reproduces the performance behaviour of existing and planned GNSS, in order to support the development of

next generation navigation applications. The main objectives of the GST are:

- ✓ To provide an accurate, independent tool capable of analysing customised application scenarios, based on realistic models of all effects relevant to the performance of an application of the satellite navigation technology;
- ✓ To produce precise technical data on the performance of navigation applications in the form of predefined or custom-defined navigation scenarios;
- ✓ To generate simulated navigation data as a real-world navigation receiver would capture, in a format familiar to application developers, in support of algorithm development and testing; and
- ✓ To allow external programs and file formats to connect to the GST, in order to maximise the re-use of existing expertise and data sharing.

To achieve these objectives, at the core of the GST is a set of models relevant to the description of navigation systems. These models are designed to be fully configurable by the user. The GST also has a set of data analysis tools and external file readers that either initialise some of the GST model parameters or fully replace a model, depending on the specific file used. In addition, the GST can also be configured to use models developed externally by third parties. These external models can take the form of a Windows DLL or an executable (".exe") file.

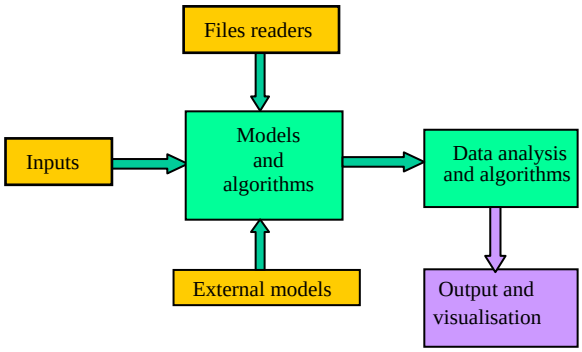


Figure 1: Functional diagram of the GST software architecture

3.1 Space Segment

Orbits: The GST computation of the satellites coordinates is based on the standard Keplerian orbit parameters. Satellite positions in the GST can also be imported from actual data provided from real GPS satellites, such as those provided on a daily basis by the International GPS Service (IGS, 2004). The orbit files are generated using the standard sp3 format, and the GST contains a sp3 file reader. The values from the sp3 files can be input in the GST in place of the computed Keplerian coordinates. Alternatively, the GST can use orbital parameters of the GPS constellation, as provided in the YUMA format from the Internet (USCGNC, 2004).

Signal-in-space: The satellite signal is defined in the GST by a set of signal characteristics:

- ✓ The frequency f of the carrier;
- ✓ The modulation scheme: Binary Phase Shift Keying (BPSK) or Binary Offset Carrier (BOC); and
- ✓ In case of the BOC modulation, the two integers *n* and *m*, multiples of the base frequency 1.023MHz, defining the sub-carrier frequency fs and the code rate fc,

Navigation message: The navigation message sent by the satellite is not modelled in the GST. This means that the satellite ephemeris and clock are not transmitted to the receiver model and therefore not used in the position computation of the receiver. However the ephemeris error is represented in the GST as an error of δ_{eph}.

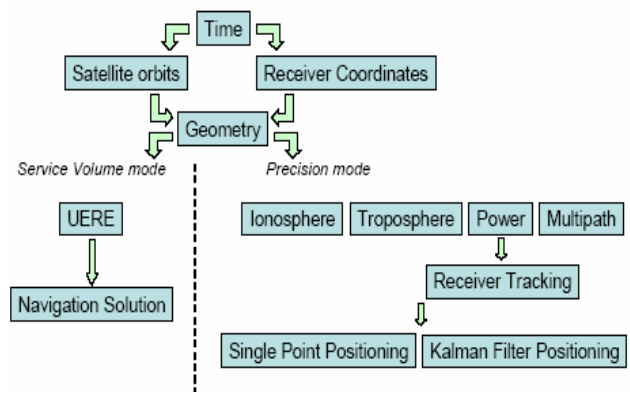


Figure 2: Functional diagram of the GST software architecture

3.2 Environment models

The term 'environment' refers here to the propagation medium through which the satellite signals travel before reaching the receiver antenna. The main effects considered are the transmission delay due to ionospheric effects, tropospheric effects and multipath.

Ionosphere: The transmission delay inferred by the ionosphere is based on the knowledge of the total electron content (TEC) along the transmission path of the signal (Parkinson 1996). In order to obtain an accurate estimate of the TEC, two approaches have been implemented in the GST:

- ✓ The first implementation uses a global ionospheric model, NeQuick (Leitinger 1996), as an external model plugged-in the GST architecture.
- ✓ The second implementation uses measured data of the vertical Total Electron Content (vTEC) regularly available from public internet sources and published in the Ionospheric Map Exchange (IONEX) format.

The two implementations described above for modelling of the ionospheric effect on the signal transmission are complementary rather than exclusive. While TEC maps provided by the IONEX files are more representative of the actual situation that one wants to simulate with the GST on a specific date, time, and location, the NeQuick model, is more generic and can be used to simulate a data and time set in the future.

Troposphere: The signal transmission delay due to tropospheric effects is modelled using the refraction index of the troposphere (Di Giovanni 1990). The Hopfield model is used to estimate the refraction index (Hofmann 1998). The model is dependent on the estimates of local temperature, atmospheric pressure and water vapour partial pressure. In the GST these values can be set manually or alternatively, the GST can read meteorological data collected at ground reference stations and output in the Receiver Independent Exchange (RINEX) format (Gurtner 2002).

Multipath: Multipath is highly dependent on the local environment surrounding the navigation receiver. Designed to be a generic tool, the GST models the multipath effect as its end effect on the measured range. Multipath causes a range measurement error, which can be isolated from other error sources in actual receiver measurements. The model used in the GST is an empirical model based on observations of time series of multipath range error (Parkinson 1996). The range error caused by multipath, δ_{mult} , is modelled as follows:

$$\delta_{mult}(t) = b_{mult} + a_{mult} n_c(t) \cos\left(k\eta(t)\right) \qquad (1)$$

where b_{mult} is the multipath bias error, a_{mult} is the amplitude of the multipath error, η is the satellite elevation angle at the receiver location, k is a factor required to adjust the elevation dependence, and $n_c(t)$ is correlated Gaussian noise.

The model in Equation (1) is appropriate to represent the range error caused by the multipath for several reasons. First, the elevation dependence is modelled in such a way that satellite at a low elevation causes higher multipath errors, as it is observed from actual measurements. The factor "k" allows adjusting the peak of this elevation dependence. Secondly, the use of correlated Gaussian noise also represents effects seen in actual observations. Typical correlation times observed in multipath effects are of the order of a few minutes. To simplify the tuning of this model, the GST proposes several sets of default values corresponding to low, medium, and strong multipath environments. Expert users are also allowed to change parameters separately if they wish.

3.3 User Segment

Receiver coordinates: Receiver coordinates are input in the GST as a latitude, longitude, and height in WGS-84.

Masking angle: The GST allows the definition of a customised masking profile to represent the specific situation at a particular location.

Receiver carrier-to-noise ratio: The total signal carrier-to-noise ratio is a measure of the signal quality and influences the tracking error. It is modelled as the sum of the signal-to-noise ratio where no jamming is present, plus the jammer-to noise ratio.

Receiver tracking error: The modelling of the accuracy of the range measurement in the GST assumes a Non-Coherent Early-Late Processing. Two models are implemented, one for the GPS BPSK modulation scheme (Kaplan 1996) and one for the Galileo BOC scheme (Betz 2000).

Receiver position: The receiver model computes an estimation of its position based on measured range between itself and each satellite in view. In the current implementation of the GST, the measured range is modelled as the true satellite-to-receiver range plus the sum of the errors described in the previous sections. Then the position is estimated using a weighted least square algorithm.

User Equivalent Range Error: The models described in the previous sections account precisely for the different error sources affecting the navigation solution. An alternative, less accurate, modelling of these errors is commonly used when simulations need to be run for long simulated periods and over large geographical areas. In this alternative method, the errors are grouped in a User Equivalent Range Error (UERE). In the GST the UERE for each satellite is a function of its elevation. Published UERE values for GPS and Galileo were used (Shaw 2000, Benedicto 2000).

3.4 GST predefined figures of merits

To describe the output of the GST a number of standard performance parameters are computed.

- ✓ *Availability of navigation solution*: The receiver can make an estimate of its position and clock bias when at least 4 satellites are in view. The instantaneous availability of the navigation solution $a^i_{nav}(t)$ is a flag set to 1 if at least four satellites are visible at a time t, and set to 0 otherwise: The availability of the navigation solution in a time window is the percentage of the time, in the time window, when $a^i_{nav}(t) = 1$.

- ✓ *Continuity of navigation solution:* The continuity of the navigation solution is a quantitative estimate whether the navigation solution can be computed without interruption. The instantaneous continuity of the navigation solution, $c^i_{nav}(t)$, is a flag set to 1 if the navigation solution is available at the time t and time t-DT, and set to 0 otherwise. The continuity of the navigation solution in a time window is the percentage of the time in the time window when $c^i_{nav}(t)=1$.
- ✓ *Availability of Accuracy*: The availability of the accuracy expresses whether the position estimate made by the receiver is accurate enough to be used for a particular application. Different applications have different accuracy requirements, and if the position computed by the receiver is not accurate enough for a given application to be successfully carried out, then satellite navigation is not appropriate. The instantaneous availability of accuracy, $a^i_{acc}(t)$, is a flag set to 1 if the positioning error of the receiver is less than a user-defined threshold, and set to 0 otherwise. The availability of the accuracy in a time window is the percentage of the time when $a^i_{acc}(t)=1$.
- ✓ *Continuity of accuracy*: The continuity of the accuracy is equivalent to the continuity of navigation solution defined previously, using the availability of accuracy to estimate continuity.
- ✓ *Dilution of precision*: The Dilution Of Precision (DOP) is a well-known indicator of the geometry of the satellites in view from a receiver.
- ✓ *Position error:* The instantaneous position error is the scalar difference between the receiver true position and the position estimated by the single point positioning algorithm of the receiver model.

3.5 Model Validation

Each of the models above has been carefully tested individually. Also, the GST results have been validated by independent sources. It is beyond the scope of this paper to present a detailed validation report of the models. Validation has been carried out in the following areas:

- ✓ Receiver and satellite geometry: coordinates, line-of-sight vectors, and DOP values were compared against those obtained from a separate software tool (Satellite ToolKit)
- ✓ Receiver RAIM availability figures were compared against a service volume simulator developed in Europe (Loizou 2002).
- ✓ World and Europe performance maps published in (Lachapelle 2002) and (Leonard 2003) were successfully reproduced with the GST
- ✓ The simulated receiver pseudorange and carrier measurements were input in the GPS/Glonass/SBAS data Toolkit Teqc (UNAVCO 2004) and successfully tested for quality.

Other comparisons have been conducted and further investigation is currently being conducted by the authors.

4 Simulation Results

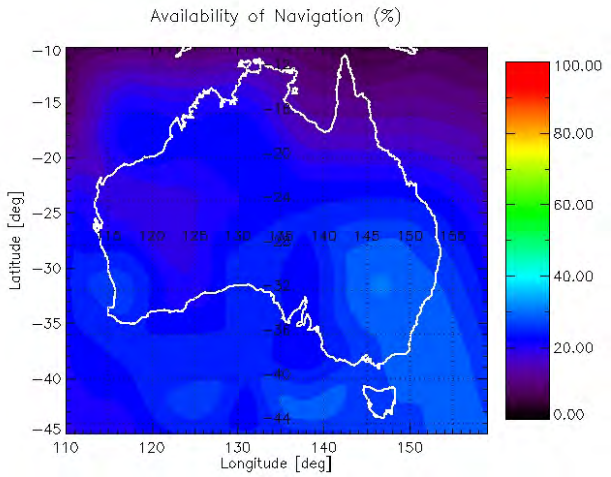
This section presents a sample of the typical simulation results obtained from the GST spanning across the entire suite of the models described earlier. The aim of this section is to illustrate the capabilities of the GST in simulations focused on Australia. The results presented here cover a wide range of topics and this will be the first investigation undertaken in Australia.

4.1 Availability of Navigation Solution over Australia

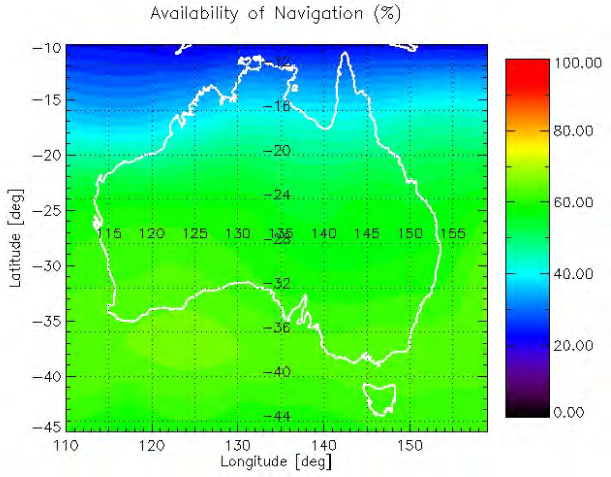
The global availability of the navigation solution over Australia is dependent upon the number of satellites in view and the mask angle. Simulations were run at a 40° mask angle for the Galileo constellation only, the GPS constellation only, and for the combined GPS+Galileo constellation. Simulations of 24 hours were run.

At a 40° mask angle, both the GPS and Galileo constellation when used alone offers reduced availability of the navigation solution. The GPS constellation is unusable most of the time (about 20% availability on average), and the Galileo constellation offers about 60% availability almost everywhere except in the northernmost region of the country.

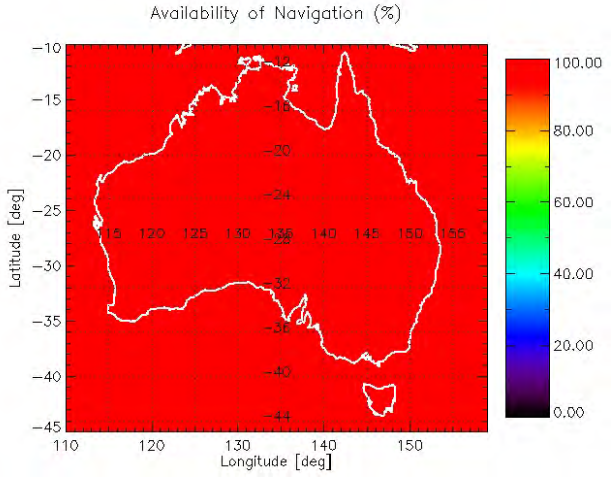
The clear latitude dependence shown in Figure 3(b) is due to the spherical symmetry of the Galileo constellation with respect to the Earth's centre. The advantage of using a combined GPS and Galileo constellation is clearly shown on Figure 3, as the availability of the navigation solution remains 100% globally.



(a) GPS

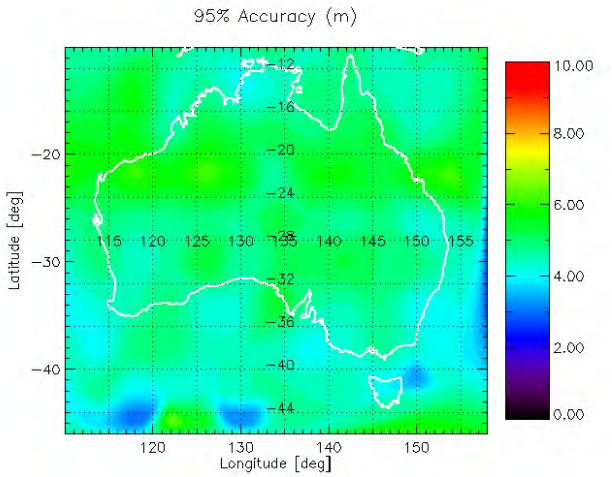


(b) Galileo

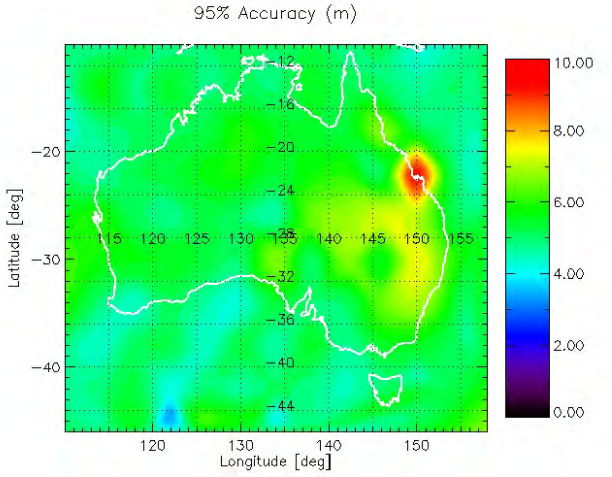


(c) Galileo+GPS

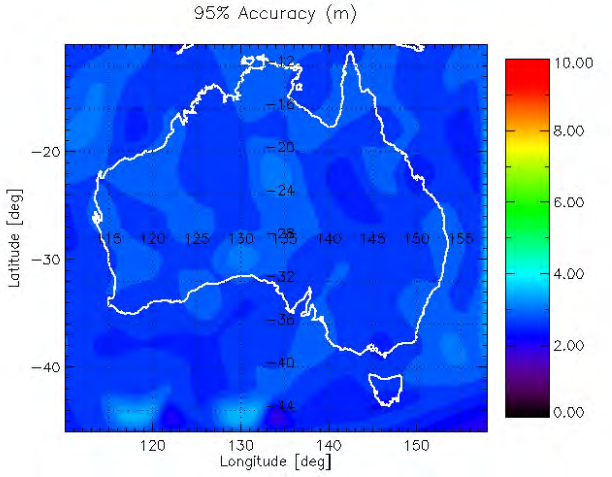
Figure 3 - Availability of Navigation Solution over Australia for a 40° mask angle



(a) Galileo



(b) GPS



(c) Galileo+GPS

Figure 4 - 95% Percentile accuracy for a dual-frequency (a) L1/E5a Galileo receiver, (b) L1/L5 GPS receiver and (c) combined Galileo/GPS receiver at the same 2 frequencies

4.2 Positioning Accuracy

The positioning accuracy presented in this section was obtained by using the GST in Service Volume mode. The UERE budget used for this purpose was a dual frequency L1/E5a Galileo receiver and dual frequency L1/L5 GPS receiver. Of course such receivers do not exist yet but their expected performance was presented in (Shaw, 2000; Benedicto, 2000) and are used here. A constant mask angle of 20° was used, for a simulation time of 24 hours.

Figure 4 shows that the 95% accuracy to be expected from Galileo or GPS dual-frequency receivers is of the same order of magnitude (about 6m across the country). The combined performance of Galileo and GPS is expected to be of about 2-3m, which is a significant improvement compared to each individual system

The effect of using a combined GPS/Galileo receiver is also to remove geographical disparity across the country. The time traces of positioning accuracy (not displayed here) are also more consistent, with less fluctuation. This effect of combined use is as important as the absolute gain in accuracy itself. Users using Galileo/GPS receivers can expect a consistent performance of their system, which may reduce the need for costly augmentations in areas and at times of the day where individual systems fail to deliver the required performance level.

4.3 Impact of QZSS on navigation performance in Australia

The QZSS has the primary objective of augmenting GPS over Japan so that satellite availability in dense urban areas like Tokyo remains high. The orbit of the QZSS satellites is such that there will always be at least one, and often 2 satellites in view and at high elevation from receivers located in Japan. However, the QZSS satellites also pass above Australia, and for that reason they offer the potential of augmenting GPS in Australia as well. This potential is assessed here with the GST.

The design of the QZSS constellation orbits is presented by Petrovski (2003), with several options being considered including an "8-shape" orbit used in subsequent GST simulation shown here. The QZSS satellites will pass above or near Australia. It is therefore possible to envisage that those satellites can be used in Australia, in the same way as they will be used in Japan, i.e. as an augmentation to GPS. The number of visible QZSS satellites and their elevation is presented here for four Australian state capitals, namely Melbourne, Sydney, Darwin and Perth. These locations have been chosen because they are widespread across the country and because the benefits of using QZSS in urban areas will be illustrated later, particularly in low visibility situations that can occur in Melbourne or Sydney.

Figure 6 indicates that QZSS satellites in an "8-shape" orbit will be well visible from Australia. In Melbourne and Sydney, at least one satellite will be visible above 60° at all times. In Darwin, all 3 satellites will be visible at all times from locations with mask angles lower than 25. The figures from Perth are also excellent.

Figure 5 shows that QZSS is particularly relevant to Australian GNSS users. In areas of reduced visibility, the availability of a single additional satellite in all times can make a difference between getting a position fix or not. At least one QZSS satellite will be almost always in view at any time, even in urban canyons. With this preliminary result based on geometry, it appears that QZSS usage in Australia should be investigated further, at both technical and political levels. To illustrate the navigation improvement resulted from the combined use of GPS and QZSS in Australia, the simulation results in an urban area is presented in Figure 6, for the location of Melbourne, and a mask profile representing an urban canyon.

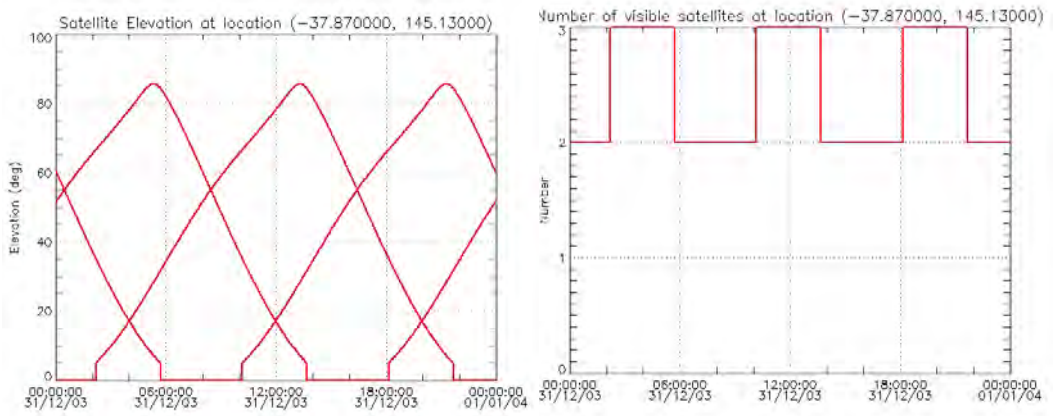
Figure 6 shows the availability and continuity of the accuracy in an urban environment at the location of Melbourne for the GPS only and combined GPS/QZSS cases. The simulation uses a single frequency receiver and the plot displayed here is relevant to mass market applications, such as personal mobility. The gain in availability and continuity of accuracy is clear from these plots, although for an accuracy threshold of 12m, 100% availability and continuity is not achieved at all times even with combined QZSS+GPS receivers. Additional systems, such as the addition of Galileo or pseudolites, may be required depending on the application considered. Investigation into pseudolite networks in urban environments is a potential field of study that can be made with the GST.

Overall, the results presented in the previous three sections show the significant improvement in navigation performance to be expected from a receiver capable of tracking several independent systems. More than in positioning accuracy, the improvement is in the consistency of the performance and therefore in the reliability of the application from the end-user perspective. This is an important consideration from a marketing perspective, as users are more likely to adopt the new technology if its reliability is certified

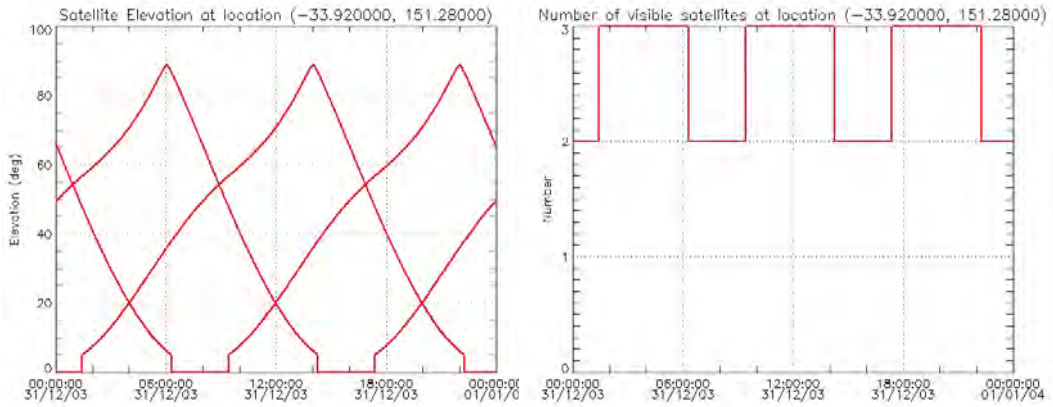
4.4 Raw data generation for post-processing by third party software

Based on its accurate models and its end-to-end capability, the GST can be used to generate simulated measurement as they would be received by a real receiver. The GST uses the widely accepted RINEX 2.1 format to output simulated data. The GST generates the RINEX observation and navigation files. Part of the

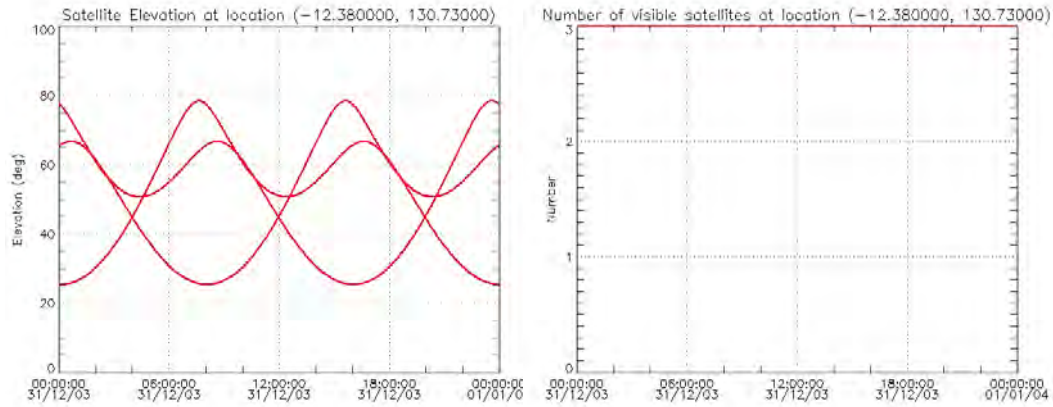
validation exercise outlined in Section 3.5 was to produce RINEX files and input them in the TEQC software for quality check. More generally, the availability of simulated Galileo and GPS data opens the possibility for research organisations to start testing algorithms and applications early, in readiness for the future availability of the hardware. This approach saves development and testing effort and can be achieved using the GST. Such capabilities will be illustrated in further publications.



(a) Melbourne



(b)Sydney



(c)Darwin

Figure 5 “8-shape” QZSS satellite elevation (left) and number of visible satellites (right) at Melbourne, Sydney and Darwin in Australia

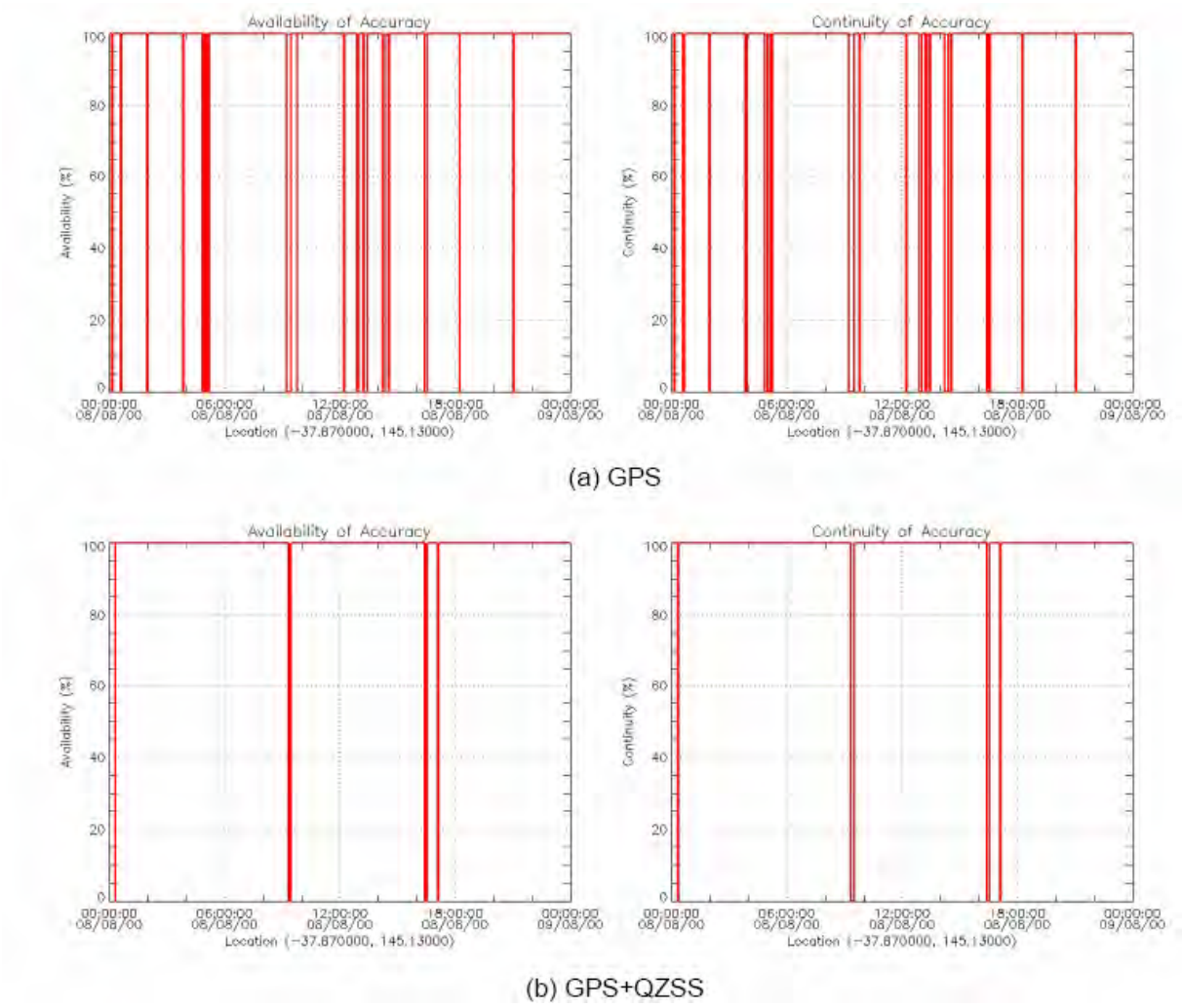


Figure 6 Availability of Accuracy and Continuity of Accuracy from a receiver in an urban area at the location of Melbourne, obtained from an accuracy threshold of 12m

5 Conclusions

This paper presents the GNSS Simulation Tool, a software tool of navigation systems that aims to provide a means for research organisations and industry to develop navigation applications using the current and emerging technologies.

The reason to build such a tool today originates from the fact that navigation technology, and especially satellite navigation, is now at a major crossroad: new satellite systems are being built in Europe, Japan, India, China and the United States. Current systems are being maintained, augmented and rapidly improved, in Russia and the United States. Adding to the advent of these new technologies is the booming need in today's society for positioning information. It is the dual growth of the technology and the market that makes the future of navigation multidisciplinary and challenging. The current GST itself cannot handle fully the complexity and variety of the challenges to come in the near future. However, it provides a low-cost, flexible tool to provide specific answers to a number of performance related questions.

The potential areas of further development for the GST are numerous. One of the most promising applications is the generation of raw pseudorange and carrier phase measurements, formatted in RINEX 2.1, to be directly usable for algorithm development and application certification. The capability to generate RINEX files is already presented in the GST, but further models and validation can be added, such as the possibility to simulate satellite outages, or the random introduction of corrupt measurements, that a real receiver would experience. Also, the Galileo RINEX format, when it

becomes available, can be easily incorporated into the GST.

For application development, the introduction of pseudolites in the simulation would be a useful addition to the tool. From the modelling perspective, ionosphere scintillation, troposphere irregularities (e.g. rainfall), clock errors, and refined receiver algorithms are other interesting developments currently being considered.

The simulation results presented here are just a subset of the entire suite of outputs that the GST can provide. They have shown the quantitative benefits of complementing GPS with another system such as Galileo. Future availability of complementary systems is particularly relevant to the South-East Asia regions and countries such as Australia are in a privileged position to develop innovative applications based on an interoperable use of future GNSS.

6 Acknowledgements

GST validation uses the service volume simulation tool developed by VEGA Plc and kindly provided by John Loizou. John is thanked for providing the tool and his insightful comments on validation and simulation. Partial financial support from Victorian Partnership for Advanced Computing (VPAC) and Corporative Research Centre for Micro-technology endorsed to A/Prof Zhang is highly acknowledged.

References

Barnes J, Rizos C, Wang J, Small D, Voigt G, and Gambale N (2004) ***High precision indoor and outdoor positioning using LocataNet,*** Journal of Global Positioning Systems, 2(2):73-82

Bartone CG (1996) ***Advanced pseudolite for dual-use precision approach applications,*** Proceedings of 9th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Kansas City, Missouri, 17-20 Sept., 95-105

Benedicto J., Dinwiddy S., Gatti G., Lucas R., Lugert M., ***Galileo: Satellite System Design and Technology Developments,*** European Space Agency, November 2000, http://ravel.esrin.esa.it/docs/galileo_world_paper_Dec_2000.pdf.

Chinese Defence Today (2004) BD-1 Navigation Satellite, http://www.sinodefence.com/space/ spacecraft/bd1.asp.

Di Giovanni G., Radicella S., ***An analytical model of the electron density profile in the ionosphere,*** Advances in Space Research, Volume 10, 1990.

European Comission (2003), ***The Galilei Project,*** Galileo Design Consolidation, http://europa.eu.int/comm/dgs/energy_transport/galileo/doc/galilei_brochure.pdf.

Featherstone WE, Kirby JF, Zhang KF, Kearsley AHW, Gilliland JR (1997) ***The quest for a new Australian gravimetric geoid***, in: Segawa J, Fujimoto H, Okubo S (eds), Gravity, Geoid and Marine Geodesy, Springer, Berlin, 581-588

Gurtner W., ***RINEX: The Receiver Independent Exchange Format Version 2.10,*** ftp//igscb.jpl.nasa.gov/igscb/data/format/ (2002)

Hofmann-Wellenhof B., Lichtenegger H., Collins J., ***GPS Theory and Practice,*** Springer-Verlag, 1998.

IGS website: http://igscb.jpl.nasa.gov (last visited November 2004)

Kaplan E., ***Understanding GPS Principles,*** Artech House, 1996.

Lachapelle G., Cannon M., O'Keefe K., Alves P**.,** ***How will Galileo Improve Positioning Performance?*** GPS World, September 2002, 38-48.

Leitinger R., Radicella S.(1996), ***NeQuick Ionospheric Model, Software Documentation,*** http://www.itu.int/ITUR/software/study-groups/rsg3/databanks/ionosph/Rec531/NeQuick_ITUR_software.pdf**.**

Leonard A., Krag H., Lachapelle G., O'Keefe K., Huth C., Seynat C., ***Impact of GPS and Galileo Orbital Plane Drifts on Interoperability Performance Parameters,*** GNSS 2003 Conference, Graz, Austria, 22-25 April 2003.

Loizou J. ***Service Volume Tool for RAIM computation,*** Private Communication**.**

Parkinson B., ***Spilker J. (Editors), Global Positioning System: Theory and Applications,*** Volume 1, Progress in Astronautics and Aeronautics Volume 163, 1996.

Petrovski I. (2003)**,** ***QZSS - Japan's New Integrated Communication and Positioning Service for Mobile Users,*** GPS world, June 2003 issue.

Seynat C., Pidgeon A., and Loizou J. (2003), ***Combined use of Galileo and GPS: towards innovative navigation research,*** Proceedings of SatNav2003, the 6th International Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services, Melbourne, Australia, 22-25 July 2003

Shaw M., Sandhoo K., Turner D., ***GPS Modernization,*** Proceedings of The Royal Institute of Navigation GNSS-2000, Edinburgh, Scotland, May 2000.

UNAVCO Facility, Boulder, Colorado, Teqc homepage, http://www.unavco.org/facility/software/teqc/teqc.html (last visited November 2004)

United States Coast Guard Navigation Center, GPS Almanac Information, http://www.navcen.uscg.gov/gps/almanacs.htm (last visited November 2004)

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 242-250

Journal of Global Positioning Systems
2004
© CPGPS

GNSS Coordination at the National Level: the Australian Experience

Donald H Sinnott

Department of Transport and Regional Services, Canberra ACT, AUSTRALIA
e-mail: don.sinnott@ieee.org Tel: + 61 8 8264 4515

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. In May 2000 the Australian Minister for Transport and Regional Services, advised by his Department, established a non-executive stakeholder body, the Australian GNSS Coordination Committee (AGCC), with terms of reference aimed at national coordination of GNSS application. This initiative responded principally to perceptions of potential for economies and efficiencies from national-level standardising and investment-sharing of equipment and services, especially in GNSS infrastructure and augmentation. In the event, in its first three years the AGCC was little able to exert significant influence in such market-driven areas. Rather, it successfully developed for government endorsement, in August 2002, a wide-ranging national GNSS policy and also addressed priority applications issues concerning GNSS jamming and interference, spectrum licensing, legal positioning/timing matters, and national and international connections, including with GPS and Galileo program management. Following a performance review in 2003 the AGCC's mandate was extended to 2006, with revised terms of reference. This paper critically examines the experience of the AGCC in national-level coordination of GNSS application. As in many countries, Australia does not control sources of GNSS signals and applications are pervasive within a free-market economy. No single government agency or industry sector has general GNSS control or policy mandate. The degree to which, in this environment, a non-executive body like the AGCC can be effective in its role is discussed. The experience and future plans of the AGCC reported in this paper raise topics of relevance not only for Australia but for other countries as well that seek a degree of national coordination and efficiency in GNSS application

Key words: Coordination, Australia, national, government, jamming.

1 Introduction

The Australian Global Navigation Satellite Systems coordination Committee (AGCC) was established for an initial three-year term by the Deputy Prime Minister and Minister for Transport and Regional Services in May 2000. Its establishment came from work of an interim committee that researched the value of having such a body of GNSS stakeholders able to advise government on cross-sector issues relating to the uptake and application of GNSS in Australia.

There were multiple motivations for the establishment of the AGCC and these have mostly proven enduring over its initial term and into what is now the first year of its second term. As is outlined below, the AGCC has found more tasks than it has resources to address and new areas for consideration continue to be added to the AGCC's interest-list, with current considerations such as Galileo, GPS modernisation and the relentless development of GNSS applications technology raising a series of issues of national importance to Australia.

The most visible outcome of the AGCC's first term was the drafting of a significant national policy document, *Positioning for the Future*, (AGCC, 2002) which the Minister released in August 2002. The policy document was developed by a team established by the AGCC that canvassed all major stakeholder groups and Federal government Departments for concurrence before seeking Ministerial endorsement and release. It is a challenging and forward-looking document.

A description of the first three year's activity of the AGCC, structured against the major headings of *Positioning for the Future*, was given at SatNav2003 (Sinnott, 2003) and will not be repeated in detail here. Suffice it to say that, in addition to the focus provided by development of *Positioning for the Future*, specific issues of GNNS jamming and interference, spectrum licensing, legal issues attached to GNSS reliance for position and timing and international programs in GPS and Galileo were major themes. Some commentary under these and

other headings follows in subsequent sections of the present paper.

At the time the SatNav2003 paper was prepared a formal external review of the AGCC was under way but its findings were not available. The review was released in mid-2003 so now, a year on, it is timely to reflect on both internal and external perceptions of the AGCC's performance and its potential for the future.

2 Maturing of the AGCC's terms of reference

The initial terms of reference for the AGCC have been revised to some extent as a result of the review of 2003. To some extent the revisions were a reflection of experience and an assessment of what is reasonably possible to expect from such a group. But the real underpinning for any terms of reference must come from the *why* and *what* questions – *why* is GNSS important to Australia and *what* can a non-executive advisory and consultant body deliver?

2.1 The GNSS market environment in Australia

A key factor on which the existence of the AGCC was, and is, predicated is expectation of sustainment and growth in demand for GNSS services in Australia. This demand is in turn driven by important economic activity for which GNSS is a crucial infrastructure element.

There are no readily accessible figures for market penetration or projections for growth of GNSS in Australia but some international comparisons are informative. Based on figures developed by DSTL UK, global sales of GPS-based products grew at a compounded annual rate of over 30% from 1996 to 2003. This growth projection is the economic driver for European investment in Galileo.

Notably, the European market assessment is that personal location services – mobile phones and in-car telematics – will be a major component of GNSS growth projections. In a recent European market analysis (Styles *et al*, 2003) gross annual product revenues from GPS/Galileo enabled mobile phones were estimated to grow from €23B in 2010 to €92B in 2020. In-car systems, currently a negligible market, are estimated to deliver gross annual product revenues of €40B by 2020. Location-base services of these types comprise a sector largely unexploited as yet in Australia. Past experience has been that Australians are fast adopters of new technology so modelling rates of penetration of new GNSS applications on those of the US and Europe is broadly valid.

With the appearance of Galileo, and an increasingly competitive market place for augmentation services, some of the constraining features associated with current GPS services will be set aside, thereby encouraging further GNSS applications. Integrated GPS/Galileo sets will have access to a combined constellation of over 50 satellites, providing high levels of accessibility, precision and integrity as well as avoiding many of the current issues attached to multi-path and restricted angles of sight attending operations in urban environments. Advances in GNSS user equipment and, potentially, higher power GNSS transmissions will allow more reliable use of GNSS in shadowed and indoor environments.

As projected globally, new markets in Australia in personal location-based services can be expected, including applications that rely on a higher tolerance by the public for personal surveillance and localisation that may be sourced in community concerns about international terrorism. The penetration of RFID tagging for freight and inventory control will, far from displacing GNSS location logging, work in synergy with GNSS technology where the inherent short-range nature of RFID logging can be creatively augmented with the long-range capability of GNSS/GSM position logging. And while there have been some initial Australian moves to exploit GNSS technologies to monitor maritime and land transport vehicles (for example, the VicRoads Intelligent Access Program, as described by Koniditsiotis (2003)) there remains far more scope for potential exploitation of such technologies in Australian transport applications, such as for tolling (Kallweit, 2003) and traffic congestion management.

Further, the Australian market is as open as any in the world so it is not expected that growth rates will be impacted negatively by government controls: the introduction and application of new technology is essentially left to the free market. Australian federal governments of any persuasion are unlikely to intervene to constrain or manage penetration of a new technology for reasons other than equity (as in guarantees for a level of services judged to be a universal entitlement, such as is the case for telecommunications) or civil rights and privacy (as in prohibiting interception or jamming, again by analogy with telecommunications).

There are two areas in which legislative actions may work to encourage GNSS penetration in Australia.

- It is likely that, eventually, Australia may introduce requirements on mobile communications carriers to provide a means of localisation of emergency calls made from mobile phones. In the US the E911 mandate and in Europe the somewhat softer provisions of the E112 reporting requirements are already driving industry to develop localisation techniques, predominantly GNSS-based, and single-chip GNSS receivers, costing less than $US10 in bulk, for embedding in new generation handsets.
- A different type of intervention mirrors that of many governments of capitalist economies, which promote

competition through means such as anti-trust legislation and competition policy (as in Australia's Trade Practice Act). It can be expected that Australia's competition policies will remain substantially unchanged and will continue to be enforced with vigour. This is likely to have a positive rather than constraining impact on growth of GNSS applications.

However, it has also been pointed out (Perez, 2002) that, historically, states act to standardise and consolidate commercial practices in a technology regime once it has encountered a mid-life turning point where rebalancing of individual and social interests within capitalism is called for. This is arguably typical of the current state of the ICT revolution in which GNSS development is established. On this basis, albeit circumstantial, the possibility for some government involvement in regulating or controlling aspects of GNSS application, and thereby impacting market projections, cannot be ruled out.

The conclusion of this sketch of the future is that Australia can expect markets for GNSS technologies to expand at a rate close to 30% pa for the next decade in a free-market environment. While the Federal government is unlikely to be drawn to intervene in this market-driven growth there may be some areas where equity, security or privacy issues, among others, could drive some legislative intervention. It is in such areas where government recourse to a body such as the AGCC, from whence independent and balanced cross-sectoral advice can be sourced, is likely to be critically important.

2.2 What can the AGCC reasonably be expected to achieve?

Before passing on to more specific issues it is important to underline what a non-executive body such as is the AGCC can realistically be expected to achieve. Australia does not control sources of GNSS signals and has no realistic ambition to do so. It relies on the provision of such signals by major countries and power blocks and seeks to encourage applications based on this technology that return national benefit within a free-market economy. No single portfolio agency or industry sector has control over GNSS applications or a specific policy mandate for it in Australia (except that Defence has an agreement with the US on military aspects of GPS and there are GNSS implications in international aviation and maritime navigation agreements to which Australia is party). Accordingly, a body such as the AGCC has a quite limited locus of control and must see itself as primarily a consultative and advisory body. Nevertheless, as is set out in what follows, it can fulfil a valuable role if its advice is sought and valued on the basis of the informed base from which it comes.

2.3 The AGCC business case – initial terms of reference

In September 1999 an interim AGCC prepared a business case for formal establishment of the AGCC. The interim committee noted the fact that GNSS was in extensive use in a wide range of sectors and, in particular, was finding increasing application in multi-modal transport applications in Australia. It pointed to the growth of national and regional based differential GPS networks, created by governments and their agencies as well as by the private sector and industry throughout Australia, and saw some national cost-benefit possible by coordination of these burgeoning differential GPS systems.

The business case was much influenced by particular case studies that focussed on the wasteful proliferation of real-time differential GPS services. But it also saw benefits, less well defined in economic terms, coming from having an expert forum in the AGCC that could advise government pro-actively on GNSS application areas where government should play a role in policy and standards frameworks. Radio frequency spectrum matters were a particular case in point.

One weakness of the business case is that in most broad sectors of GNSS application, such as transport, the application of GNSS is against quite clear sectoral objectives and it is not clear to those involved in such sectors that there is benefit from wider coordination. Thus, for example, while there was, and is, strong evidence that there are substantial benefits in road transport from greater application of GNSS technologies it is less clear that a more widely-based coordination body advising government would materially improve outcomes in this one sector. There are already bodies like the Australian Transport Council and Intelligent Transport Systems Australia (ITS) that can be expected to include in their purview relevant aspects of GNSS application to transport. Similarly, the Department of Defence might see itself as internally self-sufficient in managing defence GNSS applications; its attitude to the AGCC might well be more informed by the economy it presents by allowing a single point of contact with key representatives of civil users of a military system, GPS, than by any expectation of increased defence efficiency in GNSS application.

This background is encompassed in the forward-looking statement made in the business case when leading into describing a proposed work program. It reads as follows.

The proposed Australian GNSS Coordination Committee (AGCC) will be a proactive Committee identifying areas of GNSS use which can be managed better through a nationally coordinated approach. It presents an ideal forum for exchange of information with an international flow-on. The Committee would provide information to, and liaise with, Australian representatives on

international bodies, and would seek information exchange from similar bodies. The Committee would also look at gaps in the use of GNSS between the different transport modes and how they can be overcome. The range of skills and backgrounds which the members would bring to the Committee would be compounded in impact by the networking opportunities made possible.

2.4 The AGCC review – refining the terms of reference

Establishment of the AGCC in 2000 was for an initial term of three years so in 2003 an external review was carried out. The review noted that the environment in which the AGCC worked had changed considerably from that foreseen in the initial business case. In particular, just before the AGCC's first meeting, selective availability was removed from the GPS standard positioning system signal, making many metric-precision differential GPS systems redundant in terms of precision improvement. With this was removed a major plank of the initial "national efficiency" coordination role foreseen for the AGCC. Conversely, the maturing plans of the European Union for Galileo, barely hinted at in the 1999 business case, posed a new set of challenges in developing a national position on GNSS in Australia.

The review noted that the AGCC had struggled to achieve outcomes in some areas. These included its inability to influence GNSS infrastructure, including augmentation systems, standards, protocols and receiver technologies. This in undoubtedly a true assessment and is is now accepted in the new terms of reference that these are not areas in which a major AGCC impact will be felt.

On the positive side, the review noted the substantial outcome represented by release by the Federal Government of a wide-ranging policy statement on GNSS, *Positioning for the Future*, which had been developed by the AGCC. Other areas in which the Committee's work was deemed effective was in interference and jamming, legal issues, security for the GNSS spectrum and information provision to the GNSS user community. Each of these areas is addressed in following sections of this paper.

The net assessment of the AGCC was sufficiently positive for the review to recommend a further three-year term for its activity. Proposed amended terms of reference were then developed by officials of DOTARS and the AGCC Chair, noting the assessments and recommendations of the review, the experience of the AGCC's first term and the resources currently accessible to the AGCC. These were put to the Minister with the review recommendation for a further three-year term. The Minister agreed with the recommendation for extending the AGCC's term to 2006, reappointed the Chair and has formalised current terms of reference.

It is instructive to compare and contrast the original terms of reference with those now in place, as the shift in emphasis is in part the product of three years experience in what is reasonable and feasible for such a body to address. In particular, other countries in a similar position to Australia – ie dependent on GNSS-derived services but not having any vesting in space assets or GNSS control – might gain some benefit from noting the evolution in the AGCC terms of reference.

The following shows a comparison of the body of the original (2000) and current (2004) terms of reference, subdivided into clauses of convenience, which requires some reordering of paragraphs to allow comparison. (Some minor amendment to grammar, shown in square brackets, of the 2000 version has also been effected to allow more ready comparison, but without altering the sense.)

Preamble to original (2000) Terms of Reference:
The AGCC will [function by] …

Preamble to current (2004) Terms of Reference:
The AGCC is the national advisory body to Minister for Transport and Regional Services on issues relevant to Global Navigation Satellite Systems (GNSS). In developing advice, the Committee consults with Australian GNSS stakeholder communities, and is informed by linkages with international GNSS providers and authorities. The Committee provides information on GNSS developments to Australian stakeholders, and encourages the take-up of GNSS applications.
The Committee has a role in …

Clause 1, 2000 version:
consider[ing] and develop[ing] mechanisms to coordinate all aspects of GNSS on land, sea, and in the air, including:
- development and maintenance of a national strategic policy towards GNSS;
- coordination of national infrastructure development taking advantage of economies arising from multiple use of common systems; and
- recommendation of direction for preferred GNSS standards and protocols for use in Australia;

Clause 1, 2004 version:
- developing and facilitating national GNSS policy;
- harmonising GNSS standards and protocols to achieve the potential economic and social benefits from applications;

Clause 2, 2000 version:
promot[ing] the safe and effective utilisation and development of GNSS in Australia, including through:
- promotion of GNSS user education and effective information dissemination;

- an integrated approach to establishing mechanisms to minimise the effects of interference to GNSS; and
- investigation of specific GNSS related issues through the initiation and management of studies;

Clause 2, 2004 version:
promoting the efficient and effective development of national GNSS infrastructure;

Clause 3, 2000 version:
coordinat[ing] national security issues and assist[ing] in keeping users aware of GNSS developments and security in all transport modes and relevant areas and provid[ing] a forum for an exchange of information on receiver technology and applications;

Clause 3, 2004 version:
- coordinating advice on national security issues as they impact on GNSS applications by Australia;
- promoting the further penetration and application of GNSS technology to all modes of transport where efficiency, safety, environmental and other benefits can be realised;

Clause 4, 2000 version:
Coordinate[ing] the application of augmentation systems, particularly the provision of new augmentation systems, taking advantage of economies available through sharing common systems;

Clause 4, 2004 version;
deleted

Clause 5, 2000 version:
coordinat[ing] and influenc[ing] national and international use of GNSS through existing radiocommunications fora in Australia; and

Clause 5, 2004 version:
promoting and protecting GNSS spectrum management interests and issues in national and international radiocommunication fora;

Clause 6, 2000 version:
coordinat[ing] the national use of GNSS in other relevant applications.

Clause 6, 2004 version:
- maintaining an informed understanding on the provision and application of GNSS nationally and internationally, through liaison with relevant authorities;
- providing advice on GNSS matters requested by the Minister for Transport and Regional Services and undertaking specific tasking referred to it by the Minister.

What is clear in this clause-by-clause comparison is that the current terms of reference emphasise the participative, consultative and advisory nature of AGCC operations rather than promoting an unrealistic expectation that the Committee will have the resources and authority to intervene in market-driven areas in a significant way. There is now a concentration on ends and an avoidance of prescription of means. In particular, the focus on rationalising augmentation systems, which played such a major part in initial thinking about AGCC activity, is significantly de-emphasised. All these changes reflect the reality of the first three year's operation of the AGCC and provide some useful guidance to any other nation contemplating a group similar in function to the AGCC.

What does carry over from the original to the current terms of reference is a recognition of the importance of networking through AGCC membership. There is no other interactive forum in which GNSS stakeholder representatives meet to compare notes on a technology which is pervasive. In day-to-day operations there is little reason for the geodesic community to touch base with aviation services or defence, for example, yet experience has shown that the AGCC's linking thread of GNSS has allowed valuable synergies and mutual benefits to be realised for stakeholders from these sectors, and others.

3 Significant AGCC outcomes

In its activity to date the AGCC has addressed a number of significant issues and delivered some important outcomes. A selection is described in what follows.

3.1 Licensing of GPS spectrum

The radiofrequency spectrum is managed in Australia under the *Radiocommunications Act 1992*. The Australian Communications Authority (ACA), which is established by the Act, is responsible for planning, licensing and technical standards for use of the spectrum. Although GPS signals are widely used in Australia there seemed, at the time of the AGCC's inception, no assurance that the spectrum used for GPS transmissions would be protected and that, indeed, nothing stood in the way of the ACA allowing other use of this spectrum in accordance with its charter.

There were several options under the *Radiocommunications Act* by which GNSS signals could be protected. After study by the AGCC, including referral to its Legal Issues working group and ACA advisers, licensing of the primary in-space transmissions has been achieved through having the Department of Defence hold a space licence (a form of apparatus licence) and covering receivers through a class licence. These arrangements are now in place so that Australian users of GPS services may rest assured that they have free and unhindered access to GPS signals and that legal remedies exist to prohibit incursions into this spectrum.

It might be considered anomalous that the Department of Defence should, at significant annual cost to it, hold a

licence for civil GPS services when its access to the military GPS signals is protected in other ways. While Defence does have some requirement to access civil signals its holding of the civil GPS signal licence is primarily for benefit of the rest of the Australian user community. For its part, ACA has a charter to charge annual fees for spectrum licensing so the current arrangement with Defence is no different from any other commercial licensing arrangement. The AGCC has reflected that, from the perspective of the Australian taxpayer, it must be considered curious that two agencies they fund are committing a good deal of administrative time and energy to exchanging taxpayer money annually to achieve such an obvious and long-term community need. One might also have some concerns that, under funding pressures, Defence might at some future time opt to cease holding the licence and thereby leave the spectrum unprotected. There seems a good case for some minor legislative change to avoid this current unwieldy and essentially unsatisfactory fix by an amendment to the *Radiocommunications Act* and this has been tabled as a recommendation from the AGCC's Legal Issues Working Group.

3.2 Ban on GNSS jammers

The AGCC had been concerned that, while it is an offence in Australia to operate an unlicensed transmitter, such as might be used to attempt to jam or interfere with GNSS signals, it has not until recently been an offence to supply, or possess for the purposes of supply, such a device. With jamming devices advertised for sale (albeit at a significant price) on the web, and circuit designs and instructions sufficient for a trained technician to construct such a device similarly available, the AGCC consulted with the ACA to determine an appropriate regulatory response.

The AGCC noted the growing reliance our community has on GNSS signals for position-fixing, navigation and timing so that jamming and interference to GNSS signals poses a significant public safety and security risk. The AGCC agreed with the ACA that there appeared to be no legitimate radiocommunications use for a GNSS jammer and that, so long as legitimate Defence use was permitted, for which an exemption exists, a legislative approach was appropriate.

A public consultation process was initiated by ACA in August 2003 and followed through to the enactment of legislation. The outcome, as announced on 1 September 2004 by the ACA (Australian Communications Authority, 2004), is that devices that can be used to jam GNSS are now prohibited under section 190 of the *Radiocommunications Act 1992.* The impact of the prohibition is that any person who supplies, or possesses a jamming device for the purpose of supply, can be prosecuted under the Act. Penalties range from fines of up to $165, 000, to imprisonment. The ACA has publicised the ban to increase public and supplier awareness of GNSS jamming device prohibitions in the hope that widespread awareness will minimise the need for regulatory action after the event. Of course it may be said that such legislation only deters those who abide by the law and that legislation by itself does not remove the threat of GNSS jamming. This is undoubtedly true but it is equally true that the legislation *ups the ante* in a major way for those who might seek to disrupt a significant element of our national infrastructure.

3.3 Evidentiary use of GNSS signals

On occasion GPS positioning data is used for evidence purposes in the courts, and there have been calls for an independent authority to verify the performance of the GNSS signal at a given point in time. It is expected that, as awareness grows of the application of GNSS signals to "prove" location and time in court proceedings, there will be more emphasis on this matter.

The AGCC's Legal Issues Working Group has investigated the way in which GPS position and time is derived and how it relates to the GPS data collected and processed by Geoscience Australia and to coordinated universal time (UTC) as measured by the National Measurement Institute in Australia. The group investigated whether the data collected and the tracing of GPS time in this way was sufficient for the purposes of evidence in Australian courts for signal verification.

The group concluded that the National Measurement Institute and Geoscience Australia record and maintain sufficient data to support the accurate and reliable calculation of timing and position data, and that these records may provide appropriate evidence for the verification of GPS signals. Geoscience Australia also collects data that may provide some assistance to verify GLONASS signals if required. This advice is now listed on the AGCC web site.

3.4 Vulnerability

GNSS signals are extremely weak and their reception may easily be interfered with accidentally or intentionally, leading to erroneous time or position being derived by users or service being denied. Sources of unintentional interference include spurious emissions from other electrical/electronic equipment operating at different frequencies than the GNSS signal, but producing some signal power in the GNSS frequency bands (e.g. harmonics, bandwidth spill-over). Spectrum Management Regulations managed by the Australian Communications Authority (ACA) allow for the

enforcement of sanctions against users of equipment producing the unintentional interference to alleviate the problem. This can be achieved in a number of ways including electromagnetic shielding and/or RF filtering.

As noted in section 3.2 above, as a result of AGCC actions it is now illegal to own, operate or supply a device intended to jam GNSS transmissions. However, the possibility still exists for mischievous or malicious persons to ignore legal sanctions and seek to jam GNSS. Sources of intentional interference include dedicated barrage jammers designed to stop the GNSS receiver from forming a navigation solution, and dedicated deceptive jammers designed to cause the GNSS receiver to form an incorrect navigation solution. Under conditions of military hostility it must be expected that forces will seek to limit use of adversary GNSS, so both sides would have GNSS jamming capabilities and means for protecting their own use of GNSS. This military discipline, NAVWAR, is a specialist area which is not touched on further in this paper.

The risk of a civil GNSS receiver failing due to the presence of RF interference depends significantly on the scenario in which the receiver is operating. The effect of the loss of GNSS is again scenario-specific, but could range from limited impact (for example, a GNSS-guided agricultural machine loses its guidance, requiring the operator to navigate manually) to a more significant economic loss (for example, banking transactions cease because of time-synchronisation problems) and potentially loss of life (for example, a bushwalker relying solely on GNSS for navigation becomes lost when his receiver fails).

Up until now, there have been relatively few reported incidents of operations being significantly affected by interference-related failures. It must be noted, however, that interference, whether intentional or unintentional, can only be successfully mitigated if the interference can be recognised. Systems that rely on GNSS for time information specifically can be more easily deceptively jammed and it is often much more difficult to identify that deceptive jamming is occurring. Such GNSS-time dependent systems include cell-phone base stations.

Failure of a GNSS receiver to produce navigation and timing information does not, of itself, cause damage to a system or operation relying on that information. Systems and operations that rely on GNSS for information can be designed to operate for periods of time when the GNSS signal is unavailable and it is important to promote wider awareness of the need for critical navigation and timing systems to have adequate mitigation procedures in place, rather than rely totally on a GNSS system that, while extremely reliable, represents a potential single point of system failure.

It follows that the key to mitigating the effect of interference to GNSS receivers is to educate the user to put in place adequate backup systems and/or procedures that either: (i) prevents the interference from occurring; or (ii) allows the system/operation to continue with safety in the absence of GNSS information.

The AGCC has sought to promote, and be involved in, such educative processes. It is pleasing to note that the Attorney General's Department and other agencies addressing critical infrastructure protection have now included GNSS as an element of this infrastructure, although it is rarely seen as such by the general populace.

3.5 Factual public information

The AGCC has a useful web site, www.agcc.gov.au, that is quite heavily used to source factual information, including reliable links to other GNSS web sites, GNSS course information and breaking GNSS news. Suggestions for improvements are always welcome.

Although it is over four years since the US removed selective availability from the GPS transmissions there are still sporadic media outbreaks of rumours of its reimposition. All such rumours have been shown to be false and the AGCC has responded to the more serious media misrepresentations by using its contacts in US agencies that manage GPS to issue authoritative statements refuting ill-founded claims.

It is noted that US GPS policy is currently under review and a statement is expected at any time. It is expected that, because the US is a the nation most dependent on civil use of GPS, this will only serve to fortify the long-term intent to make GPS maximally useful for the civil community as well as for the military purposes for which it was originally designed.

3.6 Australia's relationship with Galileo

The AGCC has been active in seeking to build bridges with the EC and to promote dialogue with the intent of defining the parameters for an Australia-EU relationship on Galileo. There were AGCC/EC discussions in Brussels in April 2002 and there has been ongoing exchange of correspondence. Clearly any formal EU/Australia relationship specifically on Galileo would have to be established and managed through government channels and no government position has yet been adopted.

There is, however, a general framework for EU/Australia cooperation in which Galileo gets some coverage. In April 2003, the Australian Department of Foreign Affairs and Trade entered into an overarching *agenda for cooperation* with the EU that includes provision for

development of arrangements to enable cooperation associated with the Galileo Satellite Navigation project. This includes a framework for ongoing cooperation with the Galileo Joint Undertaking in the following areas.

- on-ground infrastructure in Australia;
- the potential for industrial cooperation
- scientific and commercial Galileo application;
- associated research and development;
- co-operative research in the field of the radiofrequency spectrum, including research into mitigation of signal interference; and
- standards

The AGCC has been monitoring carefully the development of the agreement between the EU and USA on Galileo/GPS frequency access and other matters and has been briefed by the US State Department. The conclusion of this agreement in June 2004 is seen as assisting Australia in achieving some of what was intended from the EU/Australia *agenda for cooperation*.

Prompted by AGCC advice that there are potentially very significant benefits for Australia in some closer relationship with Galileo that is likely from the general *agenda for cooperation* DOTARS officials have moved to establish an Interdepartmental Committee (IDC) to progress this matter. The IDC has met once (September 2004), confirmed that a whole-of-government approach would be appropriate to progressing any plans for an EU/Australia relationship on Galileo, and has determined areas in which further information is to be sought.

The AGCC will seek to continue to advise the IDC in accordance with its advisory charter. The AGCC sees opportunities for Australia to secure major benefits from the greatly expanded position fixing and timing capabilities that will be made available when, with Galileo fully deployed, there are over 50 GNSS satellites in orbit. One issue implied within the *agenda for cooperation,* on which the AGCC has sought to gain more widespread acceptance, is the value to Australia, as well as to the Galileo consortium, of having Galileo reference stations on the Australian landmass, desirably collocated with existing GPS reference facilities.

3.7 Privacy issues

Privacy and civil surveillance in this country falls under a number of legislative provisions of the Commonwealth or States and Territories. Legislation tends to be reactive and is primarily targeted at regulating the collection of individuals' information by government agencies, with a tacit understanding that the private sector will be largely self-regulating. The impact of new and emerging technologies has yet to be felt in many areas of legislation. In particular, legislation covering the individual's right to privacy of location, as opposed to more traditional privacy rights over such things such as medical records and telecommunications, is quite immature. Although GNSS of itself cannot be held to infringe, or potentially infringe, on any such privacy rights, clearly when allied with logging or mobile communications technology it can enable infringement of such rights.

As location-based services become more ubiquitous there will be a growth in concern about individuals' right to privacy over their location and movement history. The aggregation of location information with other data, though each element of the aggregation may be relatively benign, clearly opens opportunity for serious abuse of individual privacy rights. The AGCC has recently completed a study on this matter that concluded that present privacy legislation is adequate at the current state of technology. It has alerted the Privacy Commission to its findings. It is expected that the AGCC will need to review its position on this matter at some future date as further GNSS-based location technologies are considered or adopted.

4 The future of the AGCC

Because the AGCC operates in an area of policy and practice which is pervasive but which is no one sector's or government portfolio agency's preserve it is difficult to sustain a case for resourcing its activity. The Department of Transport and Regional Services has provided secretariat services, including hosting the web site, without which the committee could not function and stakeholder agencies from which members are drawn have supported their representatives' attendance at meetings and some associated additional time commitments. However, it has been difficult to secure membership from major GNSS-user sectors like agriculture and SMEs. In its present phase of operation, with extremely tight funds, the AGCC is unable to cover costs of member travel or attendance. Although the facts speak very clearly otherwise, there is a risk that in this environment members may feel that their efforts are underappreciated as well as unrewarded and may lose interest in continuing to serve.

In the present funding climate the AGCC can operate only to the end of its current term, ending mid-2006. If it is to continue beyond this date, a major task for the next two years will be to continue to grow its support base and credibility among its stakeholder community.

5 Conclusion

The AGCC has, within its terms of reference and operating as what amounts to a volunteer committee, delivered excellent outcomes over its four-year life. The

paper has outlined some of the key areas of its work during this period, pointing to the major challenges that remain in the future. These come from the explosive growth of GNSS applications, the entry of Galileo into the GNSS scene and GPS modernisation, the potential for uncritical reliance on GNSS to expose vulnerabilities, and a potential threat of legal and related issues in the area of privacy.

References

AGCC (2002) *Positioning for the Future – Australia's Satellite Navigation Strategic Policy*, Australian GNSS Coordination Committee, ISBN 0-642-50214-5, August.

Australian Communications Authority (2004) *Ban on RNSS jammers*, Media Release 73, September, http://www.aca.gov.au/aca_home/publications/reports/info/prohibit.htm

Kallweit T (2003) *Exacting a toll: GPS, Microwaves Precise Swiss System*, GPS World, June, 16-22

Koniditsiotis C (2003) *Intelligent Access Program (IAP) – Using Satellite Based Tracking to Monitor Heavy Freight Vehicle Compliance*, Paper 93, Proceedings of SatNav 2003, Melbourne Australia, 22-25 July (CDROM)

Perez C (2002) *Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages*, Edward Elgar Publishing Ltd, Cheltenham UK, 52

Sinnott DH (2003) *The role and activity of the Australian GNSS Coordination Committee (AGCC)*, Paper 53, Proceedings of SatNav 2003, Melbourne Australia, 22-25 July (CDROM)

Styles JS, Costa N, Jenkins B (2003) *In the driver's seat: Location based services power GPS/Galileo market growth*, GPS World, October, 34-40.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 251-258

Journal of Global Positioning Systems
2004
© CPGPS

Performance Evaluation of the Wide Area Augmentation System for Ionospheric Storm Events

S. Skone, R. Yousuf and A. Coster

Department of Geomatics Engineering, University of Calgary, 2500 University Dr. N.W., T2N 1N4 Calgary, Canada
e-mail: sskone@geomatics.ucalgary.ca Tel: + 01-403-220-7589; Fax: +01-403-284-1980

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. One of the greatest challenges in developing accurate and reliable satellite-based augmentation systems (SBAS) is modeling of ionospheric effects. Wide area GPS networks are generally sparse (station spacings of 500-1000 km), and ionosphere models can suffer degraded performance in regions where large spatial gradients in total electron content (TEC) exist. Of particular concern for Wide Area Augmentation System (WAAS) users is the feature called storm enhanced density, which is associated with large TEC gradients at mid-latitudes. This effect is a significant source of error in the WAAS correction models. The Canadian GPS Network for Ionosphere Monitoring (CANGIM) consists of three GPS reference stations in western Canada, augmented by two additional sites in the northern United States. In addition to measures of ionospheric activity, WAAS messages are collected continuously at these sites and decoded (post-mission) at University of Calgary. Localization schemes have been developed to compute WAAS ionosphere corrections for any location in North America. In this paper, performance of the broadcast WAAS ionosphere model is quantified through comparison with truth data from over 400 GPS reference stations in North America. WAAS ionosphere model accuracies throughout North America are evaluated for intense storm events, and compared with WAAS Grid Ionosphere Vertical Error (GIVE) bounds. Limitations in the WAAS ionosphere model are identified for enhanced ionospheric activity and, in particular, the storm enhanced density phenomenon.

Key words: Ionosphere, WADGPS, WAAS, GPS, positioning, geomagnetic storm

1 Background

1.1 Storm enhanced density

Storm enhanced density (SED) was originally recognized in the early 1990's with the Millstone incoherent scatter (IS) radar (Foster et al., 2002; Foster and Vo, 2002) and has been studied in detail with satellite data from the DMSP and IMAGE satellites, and with TEC data collected from multiple GPS receivers located across the US and Canada (Coster et al., 2003a; Coster et al., 2003b). Analysis of the GPS TEC data shows that during geomagnetic disturbances, ionospheric plasma is transported from lower latitudes to higher latitudes, redistributing plasma across latitude and local time.

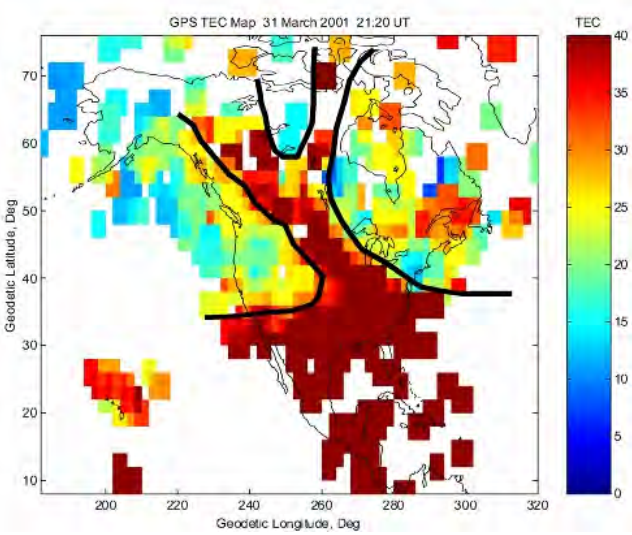


Fig. 1 An example of storm enhanced density over North America during a geomagnetic storm event March 31, 2001

SED can be described as a narrow plume of greatly enhanced TEC values (~100 TEC units). A typical SED plume would extend from the U.S. New England coast across the Great Lakes region and into central Canada. An example is shown in Figure 1. Over a period of several hours, this plume of enhanced TEC moves west across North America. Since the ionospheric TEC is generally dependent on local time, it is expected that similar effects would be observed over Europe. This, however, has not yet been fully verified. This paper demonstrates the impact of this feature on WAAS ionosphere model accuracies in North America.

1.2 Wide area ionosphere model

The presence of large ionospheric gradients associated with SED is a concern for wide area ionosphere modeling. In such models, slant line-of-sight TEC observations are derived from a number of permanent dual-frequency GPS reference stations. Typically six to nine TEC measurements are available at each station at a given epoch. These observations are generally mapped to an ionospheric shell at a height of 350-400 km (where the majority of electron density is concentrated) in order to reduce the model to two dimensions (latitude and local time). Additionally, the slant TEC observations are normalized to vertical TEC in order to remove the slant path dependence and ensure consistency between all observations. The vertical TEC observations are combined to define a spatial and temporal model of ionospheric corrections over the given region (e.g. El-Arini et al., 1995). Satellite and receiver inter-frequency biases are also estimated in the ionosphere model solution.

The WAAS ionosphere corrections are derived from a permanent network of 25 reference stations in the United States. A grid of ionosphere corrections is defined at regular spacings in latitude and longitude (Altshuler et al., 2002). The grid spacing is 5° for geographic latitudes of 55° or less, 10° for geographic latitudes in the range 65°-75°, and 10° in latitude and 90° in longitude for the geographic latitude 85° (DoT, 1999). This grid is transmitted to single frequency users, where ionospheric corrections are interpolated from the grid values and applied to local observations.

All WAAS ionospheric corrections are associated with a grid ionosphere vertical error (GIVE) which is also transmitted to users. The discrete GIVE values (after translation to meters) range from 0.30 m to 45.0 m. Any larger values imply “not monitored”. It is crucial that GIVE values accurately bound the errors in ionosphere corrections for safety-critical aviation applications. Investigations have been conducted to reliably detect and bound WAAS ionosphere errors during periods of increased ionospheric activity (Walter et al., 2000). Limitations have been identified for satellite-based augmentation systems at low latitudes (Doherty et al., 2002) and high latitudes (Skone and Cannon, 1999).

1.3 CANGIM

The Canadian GPS Network for Ionosphere Monitoring (CANGIM) currently consists of three primary stations in Western Canada: Calgary (51.08°N, 114.13°W), Athabasca (54.72°N, 113.31°W) and Yellowknife (62.48°N, 114.48°W). These stations allow latitude profiling of the mid- to high-latitude ionosphere. Additional sites have recently been installed at MIT Haystack Observatory, near Boston Massachusetts, with a second site in northern Minnesota. Future plans include the addition of three more reference sites to improve coverage in the western Canadian sector, with stations in Whitehorse, Churchill and Inuvik (Figure 2). Further sites will be installed in eastern Canada for extended coverage of the North American region.

The CANGIM sites are equipped with NovAtel Modulated Precision Clock (MPC) receivers (NovAtel, 2002) and NovAtel 600 antennas. These units contain dual-frequency Euro4 cards with an internal integrated PC and precise oscillator. A comprehensive user-interface command structure allows direct access via modem or internet connection. The CANGIM receivers have specialized firmware which provides scintillation parameters extracted from 50 Hz L1 phase observations, in addition to raw GPS code and phase observations, rates of change of TEC, absolute TEC values, and WAAS messages. The firmware version (scintw) was developed by A.J. Systems (Van Dierendonck et al., 1996).

The CANGIM network does not currently operate in real-time. Communication infrastructure is in place, however, to allow real-time streaming of data to a central processing facility at University of Calgary for future research and development purposes. Algorithms have been developed (El-Gizawy et al., 2002) to monitor the level of ionospheric activity, and associated impact on GPS users, for near real-time applications. After several hardware upgrades, the three primary stations have been operating reliably since May 2003. WAAS messages from the CANGIM are currently archived at University of Calgary. In this paper, analyses focus on post-processing of the WAAS ionosphere data from CANGIM.

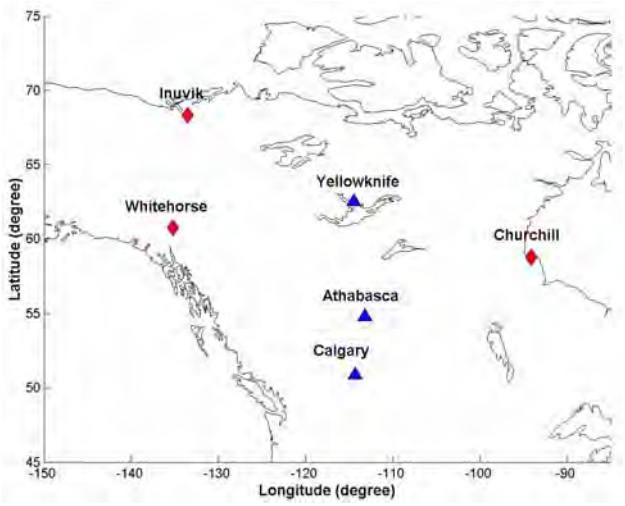


Fig. 2 Site locations of the installed MPC units (triangles) and additional proposed site locations (red diamonds)

2 Geomagnetic storm events

Extreme geomagnetic storms were observed during a two-month period in 2003. Activity commenced with one of the most severe storms of the past 15 years, in late October 2003. A major solar flare developed at approximately 11:00 UT on October 28. A severe geomagnetic storm commenced in the Earth's environment at 06:00 UT on October 29. Activity continued for several days, with further coronal mass ejections at approximately 21:00 UT October 29 and 16:00 UT October 30. Enhanced solar emissions were again observed in late November 2003, with an extended geomagnetic storm event on November 20, 2003.

The level of global ionospheric activity during these events can be quantified using the conventional space weather index Kp. This index is based on observations of magnetic field fluctuations at ground-based magnetometer stations (periods of enhanced ionospheric activity being characterized by strong electric currents which are observed as magnetic field perturbations at the Earth's surface). The planetary Kp index is derived from measurements of magnetic field variations at thirteen global stations at (approximately) equally spaced longitudes. This index is derived at three-hourly intervals and values range from 0 (quiet) to 9 (extreme). Such indices provide an approximate measure of global ionospheric activity at higher latitudes.

Figure 3 shows the Kp index for the storm period October 29-31, 2003. Kp values of 9 were observed on October 29 and 30. This indicates severe storm events for extended periods on both days. Communications were disrupted for commercial aircraft operating in polar regions, and satellite instruments were shut down to mitigate the impact of enhanced radiation in the space environment. Aurora were observed at mid-latitudes in the United States. Development of strong SED was also observed in North America.

Figure 4 shows Kp indices for the period November 19-21, 2003. Values greater than 7 were observed for an 18-hour period commencing 09:00 UT November 20, with a Kp index of 9 during the interval 18:00-21:00 UT. This indicates high ionospheric activity for an extended period on November 20, with severe storm effects in the evening hours (UT). Aurora were again observed at mid-latitudes, in the United States and Europe, with SED developing in North America – starting at 18:00 UT and continuing through 22:00 UT. The impact of these storm effects (October and November 2003) is evaluated in terms of WAAS ionosphere model accuracies in Section 4.

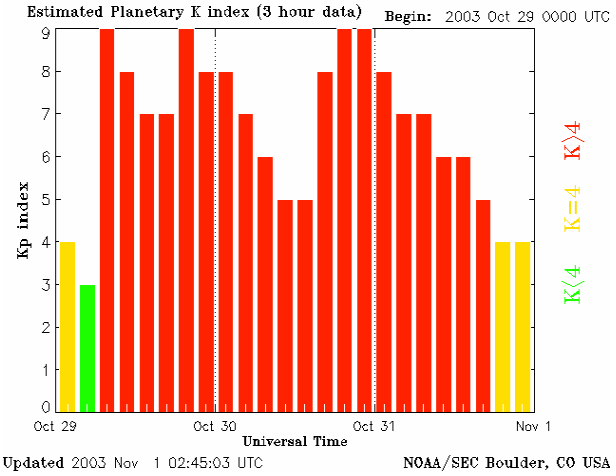


Fig. 3 Kp values for October 29-31, 2003 (NOAA SEC)

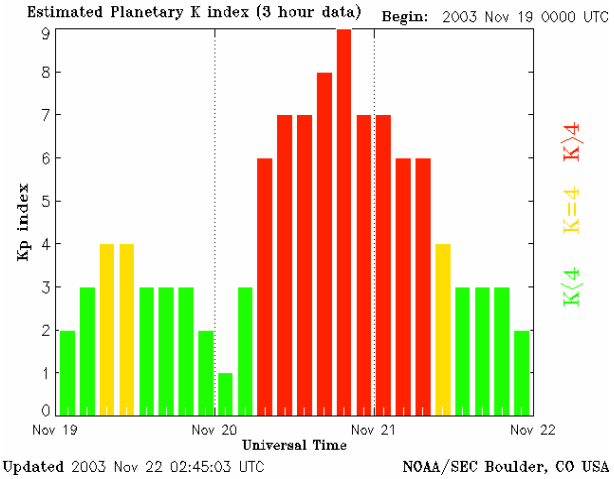


Fig. 4 Kp values for November 19-21, 2003 (NOAA SEC)

3 GPS TEC maps

SED was observed in maps of the vertical TEC derived for North America for both the October and November 2003 events. These maps were produced by combining

dual-frequency data from more than 400 ground-based GPS receivers around the United States and Canada. The data used to produce these maps included observations from both the International GPS Service (IGS) (http://igscb.jpl.nasa.gov/) and the Continuously Operating Reference Stations (CORS). The CORS network is coordinated by the U.S. National Geodetic Survey (http://www.ngs.noaa.gov/CORS/). The data were accessed via publicly available data archives on the World Wide Web ftp://cddisa.gsfc.nasa.gov/gps/ and http://sopac.ucsd.edu.

The line-of-sight TEC values were converted to vertical TEC values using a simple mapping function, and associated to an ionospheric pierce point latitude and longitude, assuming a peak ionospheric height of 400 km. GPS satellite and receiver inter-frequency biases were estimated and removed from the data. To minimize mapping function errors in modeling the vertical TEC, the elevation angle was restricted to be greater than 30 degrees. The satellite or SV biases used were determined by the JPL modeling efforts (Wilson et al., 1999). The receiver biases were determined by an in-house process. Maps of TEC were prepared at 5-minute intervals to be consistent with WAAS ionosphere grid update intervals. The vertical TEC has been binned in 2° x 2° latitude/longitude bins. No smoothing is used; the high level of detail is due primarily to the persistence of a well-organized connected structure and to the large quantity of data processed.

Large gradients in the TEC distribution are observed over the continental United States into Canada during the October 2003 event. A plume of TEC associated with SED is apparent at 22:04 UT on October 29 (Figure 5). The plume of SED has moved north and westward over a period of several hours, leading to the increased TEC values and gradients pictured in Figure 5. While observations are sparse at the high latitudes, it appears that the SED feature may extend over the polar region into Europe.

Similar to the October event, the presence of SED is clearly observed over North America on November 20, 2003. Figure 6 shows the distribution of TEC over North America at 19:00 UT and 21:00 UT. The plume of SED has moved westward during the two-hour period, with a narrow region of enhanced electron content through the eastern United States into western Canada. The evolution of this feature in global maps of TEC demonstrates that the electrons observed at 19:00 UT flow over the pole, from North America into Europe.

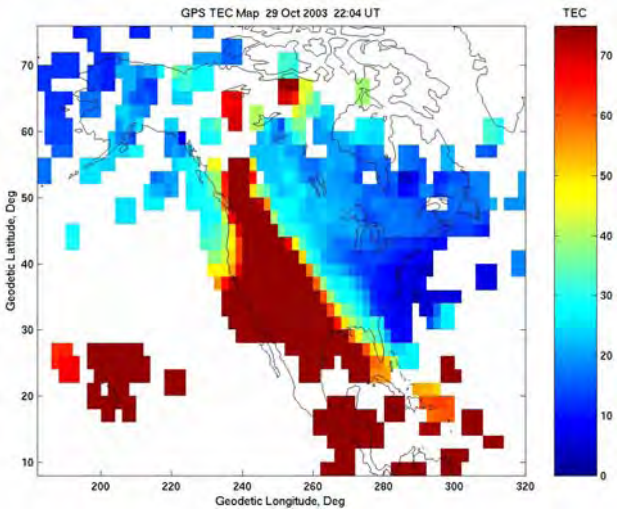


Fig. 5 GPS TEC map derived from 400+ reference stations, 22:04 UT October 29, 2003

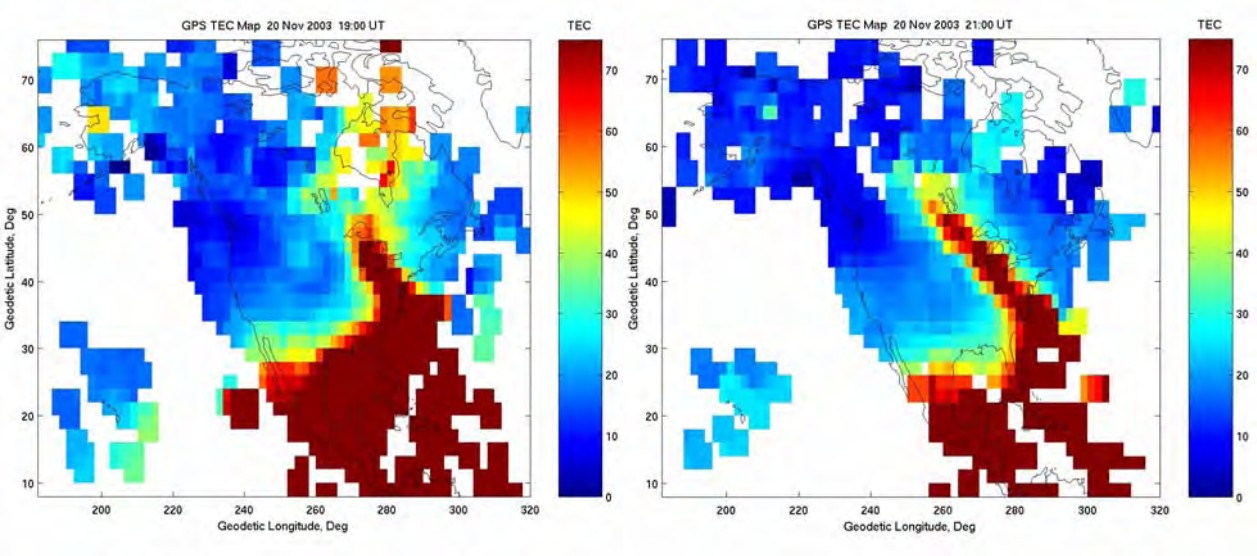


Fig. 6 GPS TEC maps derived from 400+ reference stations, for 19:00 and 21:00 UT, November 20, 2003

4 Results and analysis

Accuracies of the WAAS ionospheric corrections are analysed in the region of SED during the October and November 2003 storm events. To evaluate the performance of WAAS ionosphere models during these events, maps of the vertical TEC predicted by WAAS are compared with high-resolution TEC maps produced from the more than 400 GPS receivers in North America (e.g. Figures 5 and 6). As stated previously, the WAAS map is produced using dual-frequency observations from 25 reference stations. It is expected that this sparse network, as compared with the 400+ stations used to produce Figures 5 and 6, may not reflect some of the localized features associated with the SED. The intention is to evaluate 1) the extent to which the WAAS ionosphere model captures the SED feature, and 2) the extent to which GIVE values bound the observed WAAS model errors.

4.1 October 29, 2003

The WAAS map for 22:00 UT on October 29 is shown in Figure 7. This figure can be compared directly with Figure 5. The SED feature in the western United States is captured, to a large extent, in the WAAS ionosphere map. The magnitude of TEC values is decreased at latitudes above 45°, however, as compared with Figure 5. This is likely due to the spatial distribution of WAAS reference stations, in which there is no direct coverage of Canadian latitudes (although there is coverage at higher latitudes on the west coast of North America – in Alaska). With the lack of observations in the latitude range 45°-55°, the SED feature is not fully resolved in the WAAS ionosphere model. Discrepancies in relative spatial variations within a given map are attributed to the different spatial resolution achieved for the two networks used in the processing.

It is assumed here that the high-resolution map (Figure 5) represents "truth". Figure 8 shows the difference between the high-resolution TEC map and the WAAS TEC map in Figure 7, converted to L1 range delay. Any bias between the two TEC maps has been removed in generating Figure 8 - such that the mean spatial value in Figure 8 is 0 m. Larger differences are observed in the northwest, in the region of SED. These larger positive errors indicate the extent to which the WAAS model underestimates the northward extent of the enhanced plume of TEC. Errors in ionospheric corrections are as large as 10 m at latitudes of 45°-48°.

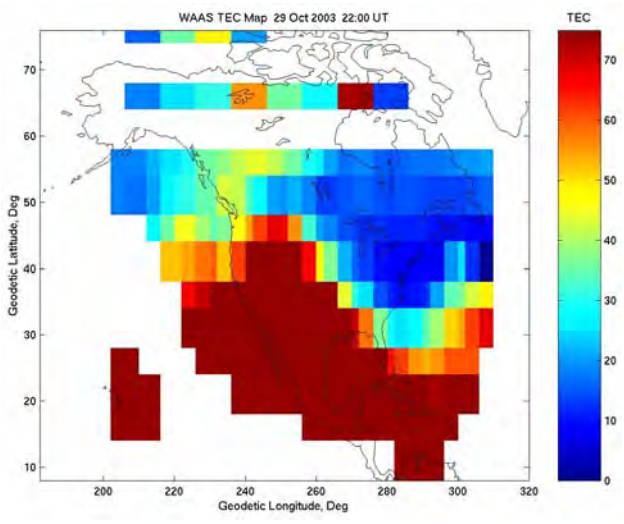


Fig. 7 TEC map produced from WAAS ionosphere grid at 22:00 UT on October 29, 2003

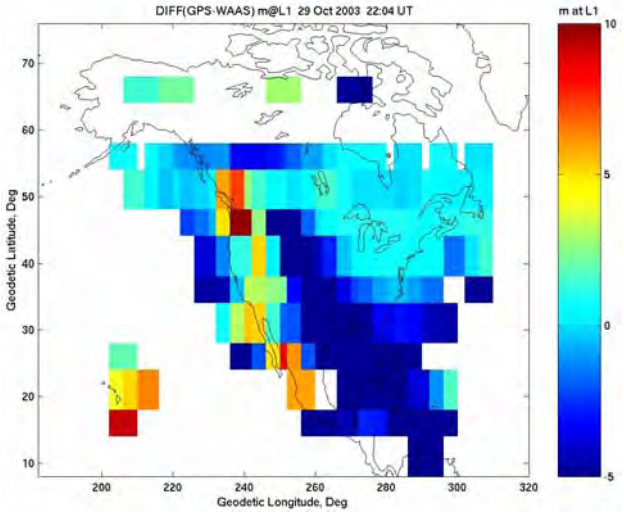


Fig. 8 Difference between GPS high-resolution TEC and WAAS TEC for 22:04 UT on October 29, 2003 (converted to L1 range delay)

In order to assess whether the WAAS ionosphere errors are bounded appropriately, the WAAS GIVE values are compared with the WAAS model errors plotted in Figure 8. The GIVE values are plotted in Figure 9. These values are valid for approximately the same time at which Figures 5, 7 and 8 are valid. GIVE values greater than 45 m are observed throughout most of North America, indicating that the ionosphere corrections are "not monitored" and users are denied ionospheric corrections over a large region. In the northeastern United States, GIVE values of 10-20 m are observed. These error bounds exceed the WAAS model errors computed in Figure 8 for this region – indicating that the GIVE values appropriately bound the observed ionospheric model errors during this event.

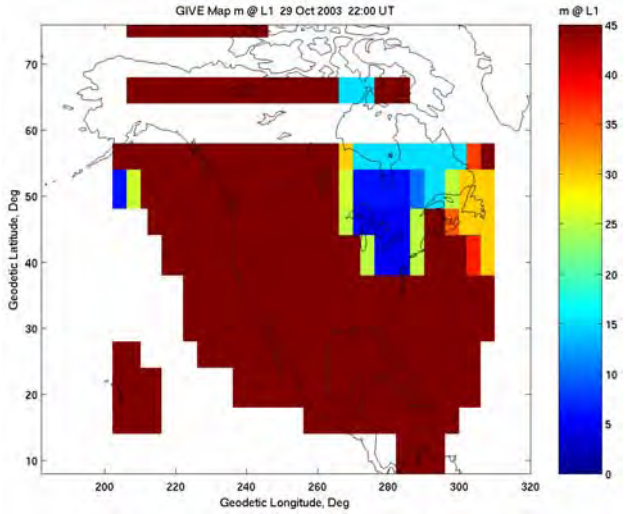


Fig. 9 GIVE values for the WAAS ionosphere grid at 22:00 UT on October 29, 2003

4.2 November 20, 2003

The WAAS maps for 19:00 and 21:00 UT on November 20 are shown in Figure 10. This figure can be compared directly with Figure 6. The SED feature developing in the eastern United States at 19:00 UT is captured at lower latitudes in the WAAS ionosphere map. Similar to the October storm event, TEC values decrease significantly at latitudes above 45°, as compared with the high-resolution "truth" maps (Figure 6). This is again due to the lack of WAAS reference stations in Canada, where WAAS observability is degraded at Canadian latitudes. The SED feature is not fully resolved in the WAAS ionosphere model, as compared with the fine-scale truth maps. The larger WAAS grid spacing (5° × 5°) does not allow for full resolution of the SED gradients.

Figure 11 shows the difference between the high-resolution truth TEC maps and the WAAS TEC maps in Figure 10, with the TEC values converted to L1 range delay. Again, any bias between the two TEC maps has been removed in generating Figure 11 - such that the mean spatial value in Figure 11 is 0 m. Larger differences are observed at the higher latitudes at 19:00 UT, where the region of SED has extended through Canada and over the pole. WAAS model errors as large as 10 m are also observed in the northern United States at 19:00 UT, where the WAAS model underestimates the SED enhancement in this region. At 21:00 UT, WAAS model errors of 5-10 m are observed in a narrow region extending northwest through the central United States, consistent with the plume of SED in Figure 6. This again reflects deficiencies in the WAAS model in resolving such localized features.

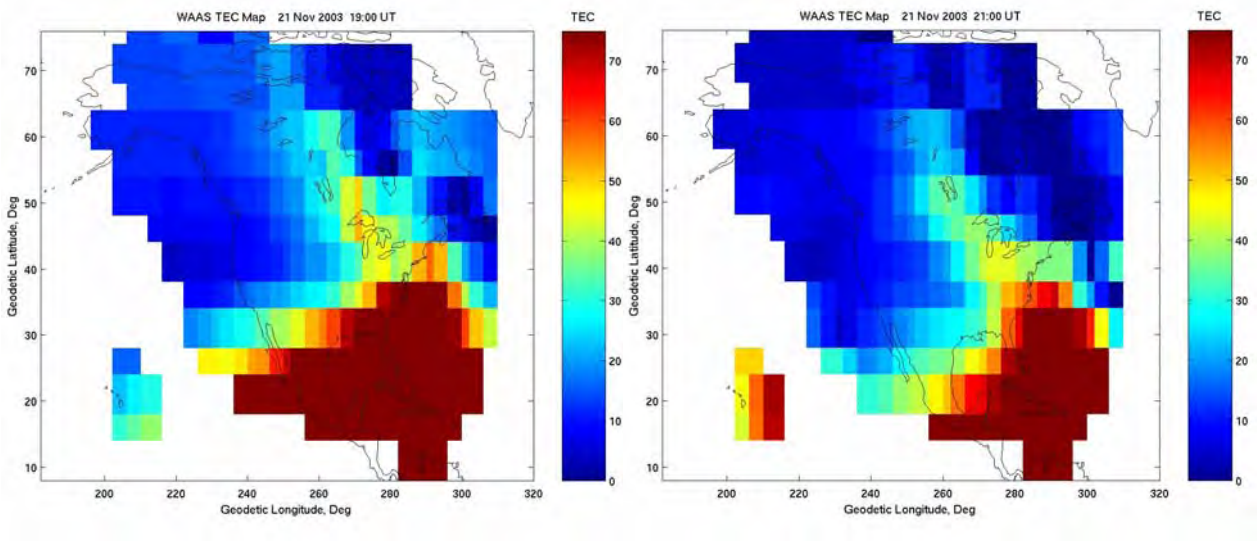


Fig. 10 TEC maps produced from WAAS ionosphere grid at 19:00 and 21:00 UT on November 20, 2003

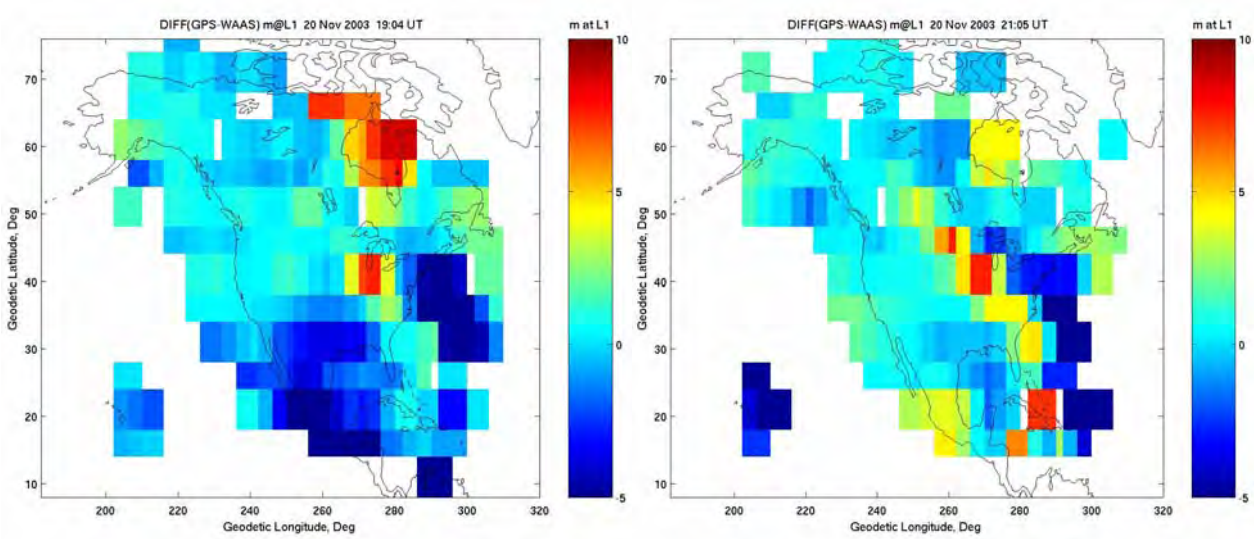


Fig. 11 Differences between GPS high-resolution TEC and WAAS TEC for 19:00 and 21:00 UT on November 20, 2003 (converted to L1 range delay)

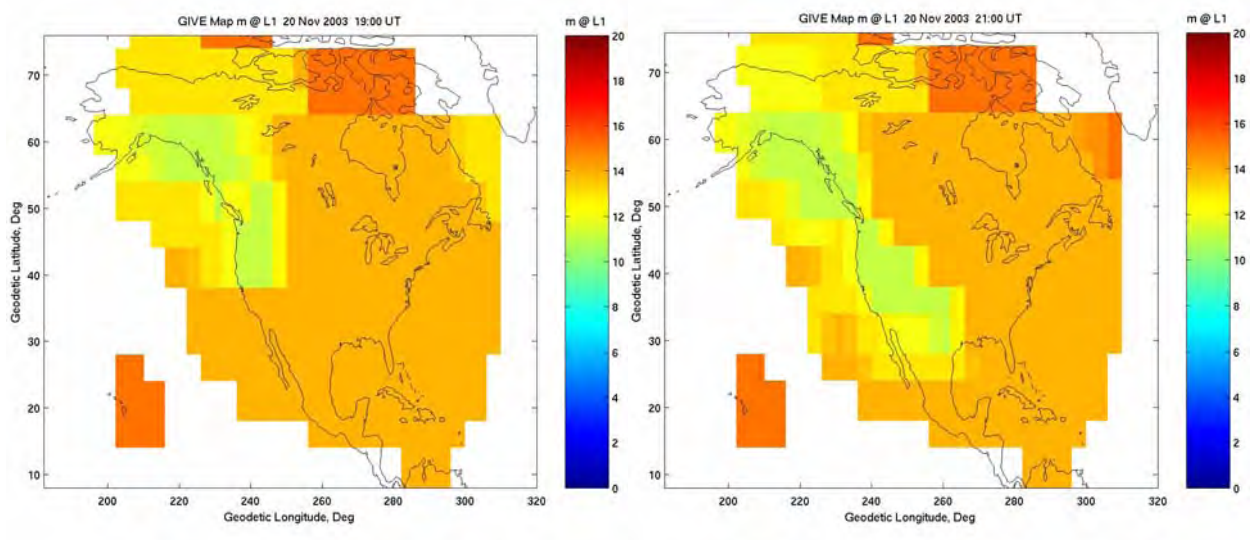


Fig. 12 GIVE values for the WAAS ionosphere grid at 19:00 and 21:00 UT on November 20, 2003

In order to assess whether the WAAS ionosphere errors are bounded appropriately, the WAAS GIVE values are plotted in Figure 12. These values are valid for approximately the same time at which Figures 6, 10 and 11 are valid. No GIVE values greater than 45 m are observed within North America during this event, such that the availability of WAAS ionospheric corrections is maintained throughout the storm period. The GIVE values are in the range 10-16 m throughout North America, and these bounds exceed the errors plotted in Figure 11 – indicating that the WAAS model errors are bounded in all regions during this event.

5 Conclusions

Severe geomagnetic storm events took place October 29-31, 2003 and November 20, 2003. The impact of these events was observed globally, with aurora in Europe and North America and the development of SED at mid- to high-latitudes. In this paper, investigations were conducted to evaluate performance of the WAAS ionosphere model during these storm events. High-resolution "truth" TEC maps were derived using dual-frequency observations from more than 400 GPS reference stations in North America.

SED was observed to develop during the late hours UT on October 29, 2003. Increased TEC values were observed at approximately 22:00 UT over the west coast of North America. A high-resolution TEC map was compared with the WAAS ionosphere values, and the WAAS model was observed to underestimate TEC at the higher latitudes extending into Canada. This is attributed primarily to the lack of WAAS reference stations (and,

therefore, ionosphere observations) at Canadian latitudes. WAAS model errors as large as 10 m were observed at geographic latitudes of 45°-48°. The GIVE values were flagged as "not monitored" throughout most of North America during this event, indicating that the ionospheric storm threat had been detected. While it is important to bound the ionosphere model errors reliably, the very large GIVE values significantly over-bound the errors and deny WAAS service throughout an extended region.

An extreme geomagnetic storm was also observed to develop during the late hours UT on November 20, 2003. A plume of enhanced electron content (SED) developed in the eastern United States at approximately 18:00 UT, extending northward into Canada, and moving westward over the central United States by 21:00 UT. High-resolution truth TEC maps were again compared with the WAAS ionosphere predictions, and the largest WAAS model errors were observed near the region of SED. The WAAS model was deficient in resolving fine-scale structure and gradients associated with the SED, and WAAS model errors as large as 10 m were observed in the United States. All WAAS ionosphere model errors were, however, bounded by the GIVE values for this event.

Acknowledgements

The authors wish to acknowledge the International GPS Service (IGS) and the Continuously Operating Reference Station (CORS) network, for raw GPS data used in the processing. Introductory material has been published previously in Proceedings of the European Navigation Conference GNSS 2004.

References

Altshuler E, Cormier D, and Go H (2002) ***Improvements to the WAAS ionospheric algorithms***, Proceedings of the ION GPS 2002, Portland, OR, 24-27 September

Coster A J, Foster J C, Erickson P, Sandel B, and Rich F (2003a) ***Monitoring space weather with GPS mapping techniques***, Proceedings of the ION National Technical Meeting 2003, Anaheim, CA, 22-24 January 22-24

Coster A J, Foster J C, and Erickson P J (2003b) ***Monitoring the ionosphere with GPS***, GPS World, 40-45

Doherty P H, Dehel T, Klobuchar J A, Delay S H, Datta-Barua S, de Paula E R, and Rodrigues F S (2002) ***Ionospheric effects on low-latitude space-based augmentation systems***, Proceedings of the ION GPS 2002, Portland, OR, September 24-29

DoT (1999) ***Federal Aviation Administration Specification Wide Area Augmentation System (WAAS)***, FAA-E-2892B, Change 1

El-Arini M B, Conker R, Albertson T, Reegan J K, Klobuchar J A, and Doherty P (1995) ***Comparison of real-time ionospheric algorithms for a GPS wide-area augmentation system (WAAS)***, NAVIGATION: Journal of the Institute of Navigation, 41(4):393-413

El-Gizawy M, and Skone S (2002) ***A Canadian ionospheric warning and alert system***, Proceedings of the ION GPS 2002, Portland, OR, 24-27 September

Foster J C, Erickson P J, Coster A J, Goldstein J, and Rich F J (2002) ***Ionospheric signatures of plasmaspheric tails***, Geophysical Research Letters, 10.1029/2002GL015067

Foster J C, and Vo H B (2002) ***Average characteristics and activity dependence of the subauroral polarization stream***, Journal of Geophysical Research, 0.1029/2002JA009409

NovAtel (2002) ***Modulated Precision Clock User Manual***

Skone S and Cannon M E (1999) ***Adapting the wide area ionospheric grid model for the auroral region***, Canadian Aeronautics and Space Journal, 45(3):236-244

Van Dierendonck A J, Hua Q, Fenton P, and Klobuchar J (1996) ***Commercial ionospheric scintillation monitoring receiver development and test results***, Proceedings of the ION 52nd Annual Meeting, Cambridge, MA, 19-21 June

Walter T, Hansen A, Blanch J, Enge P, Mannucci T, Pi X, Sparks L, Iijima B, El-Arini B, Lejeune R, Hagen M, Altshuler E, Fries R, and Chu A (2000) ***Robust detections of ionospheric irregularities***, Proceedings of the ION GPS 2000, Salt Lake City, UT, 19-22 September

Wilson B D, Yinger C H, Feess W A, and Shank C (1999) ***New and improved - the broadcast inter-frequency biases***, GPS World, 10(9):56-66

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 259-264

Journal of Global Positioning Systems
2004
© CPGPS

Japanese Regional Navigation Satellite System "The JRANS Concept"

Hideto (Duke) Takahashi

ITOCHU Corporation, 2-5-1, Kita Aoyama, Minato-ku, Tokyo 107-8077, Japan
Email: takahashi-hideto@itochu.co.jp; Tel: 81-3-3497-3114, Fax: 81-3-3497-2991

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Current Global Navigation Satellite Systems (GNSS) have enabled quality of life improvements and new business opportunities on an international scale. The range of applications and improvements includes a multitude of disciplines such as Agriculture, Transportation, Recreation, Public Safety and Security. Independent financial institutions have estimated the annual business market value for GNSS user equipment and related components in the billions of US dollars. The Government of Japan recognizes the importance of investing in GNSS today to establish a foundation for future quality of life improvements and business opportunities for current and future generations. The current Japanese program QZSS—Quasi Zenith Satellite System with 3 GPS-supplementary satellites—represents a bold step in the development of a Regional Navigation Satellite System (RNSS) for all of Asia. In January 2004, the Council for Science and Technology Policy (CSTP) in the Cabinet Office published a report regarding the future outlook of RNSS in Japan. This paper provides an update on QZSS progress, the real-world challenges and demands facing Japanese decision makers as reflected in the CSTP report, and a glimpse into the future options for the expansion of 3-satellite QZSS to a 7-satellite constellation system that can autonomously provide satellite-based position, velocity and time services, while preserving the reciprocity and compatibility with the GPS. In September 2004, CSTP published another report regarding space policy that states a long-term goal of the government to build the "autonomous" and "GPS-complementary" regional satellite navigation system in the future..

Key words: GPS, GNSS, QZSS, JRANS

1. Introduction

Japan, as well as many other countries, has been largely dependent upon the US owned and operated Global Positioning System (GPS) for obtaining space-based position, velocity and time services.

In the 1970s, the US Department of Defense began GPS development as a military force enhancer. In 1983, President Reagan offered GPS civil services to the world, free of direct charges, as a result of the KAL007 disaster. This global offer sparked widespread civil use of GPS and significant investment in civil GPS technologies, to include GPS civil augmenting satellites (e.g. US Wide Area Augmentation System (WAAS), European Geostationary Navigation Overlay System (EGNOS) and the Japanese Multi-functional Transport Satellite (MTSAT)) and GPS civil user equipment to support a broad range of applications from transportation to agriculture.

From an International Civil Aviation Organization (ICAO) perspective, ICAO document A32-19 establishes and affirms the fundamental legal principles governing the use of Global Navigation Satellite Systems (GNSS) by ICAO contracting States. A key theme of A32-19 is preservation of State authority and responsibility for the provision of air navigation services in their sovereign airspace. In other words, a State can approve GNSS for navigation use in its sovereign airspace, but the State retains authority and responsibility (and legal liability) for the use of GNSS for navigation. Specific State requirements are further clarified in principle number four of A32-19. This principle specifically states that: "Every State providing GNSS services…shall ensure the continuity, availability, integrity, accuracy and reliability of such services…including effective arrangements to minimize the operational impact of system malfunctions or failure and to achieve expeditious service recovery".

GPS is becoming a mainstay of everyday life in Japan; the Government of Japan (GOJ) is responsible for providing air navigation services in sovereign airspace; it

is difficult to receive GPS signals in Japan due to mountainous terrain, and tall buildings. These provided incentives for the GOJ to investigate the practicality of developing a standalone GPS-compatible system capable of independently satisfying position, velocity and time requirements for Japan and throughout the Asian region.

The remainder of this paper provides an overview of some of the issues to be addressed, options considered, decisions made, and highlights a future opportunity for the Asian region.

2. Real-World Issues

2.1 Cost:

A regional solution for meeting Japanese position, velocity and timing needs will require resources and long-term investment.

The US has invested billons of dollars to design, develop and implement a GPS "system". This investment goes well beyond building and launching satellites. The critical "brains" of GPS (the ground control segment with its Master Control Station, Monitoring Stations, Control Stations, Processors, Software, Communications, Security, etc.) assures the satellites perform and the "overall system" enables reliable position, velocity and time within SPS specifications.

Then, there is the upfront and recurring cost of things that are not generally seen, such as the papers and guidance directions provided by the scientific community; development of interfaces, compatibilities, documentation, standardization, upgrades, training, etc.

The other critical cost dynamic is sustaining a space-based system once it is operational. Replenishing satellites, coordinating launcher availability, schedules, risk mitigation (in the event of a launch failure), or an anomalous event causing the disruption of satellite signals. requires an enormous amount of planning, knowledge, investment and most importantly experience.

2.2 Time Scale

GPS is being used today; as you read this paper, millions of people throughout Japan and Asia are using GPS. This use is growing on a daily basis. If Japan is going to seriously consider fielding an independent GPS-compatible system, then the system is needed as soon as possible--not 10-15 years from now.

Once again, practical experience is critical to select the best systems and subsystems, take advantage of approved standards, documentation, international guidelines, and employ proven risk mitigation techniques, etc. The net result is that the time to field an independent GPS-compatible system should be reduced.

It is well known that the GPS signals enable useable services for a host of civil, as well as military applications. In particular, GPS is also being relied upon by the US and other countries as a foundation for aviation navigation services.

2.3 Equipment Compatibility

GPS user equipment has a well established manufacturing base. Those making GPS receivers and GPS "engines" have the knowledge, experience and wherewithal to provide useable equipment. If an independent GPS-compatible Japanese system is going to be embraced and provide benefits for all users, then user equipment must be readily available and affordable.

If the manufacturers have to "invent" new user equipment technology and components to accommodate a Japanese regional system, there is an established process to support the "invention" in Japan. Part of this process (particularly for aviation applications) involves the development of manufacturing standards and recommended practices. Standards and recommended practices can literally take years to develop and be accepted by the international community. Standards and recommended practices, associated research and development, testing, component design, interfaces, etc. require investment on the part of those manufacturers planning to produce user equipment. These costs will be recovered by those manufacturers producing the user equipment.

The net result of an "invention requirement" for GPS-compatible Japanese user equipment: getting user equipment to the marketplace will be delayed and equipment will likely not be readily available; and GPS-compatible Japanese user equipment costs will be increased--likely higher than today's GPS equipment.

2.4 Opportunities

GPS was fully funded by the US Department of Defense (DoD) and the US Department of Transport (DoT) to satisfy position, velocity and time requirements as well as "other" military missions. The point is that the fundamental GPS ‘bus’ being used by the US DoD can accommodate additional payloads besides the navigation payload.

Japanese GPS-compatible regional satellite system will not necessarily have the same requirements as the US DoD GPS satellites. This affords the opportunity to explore the inclusion of other payloads on-board the fundamental GPS 'bus'--such as communication and weather packages. These payloads could be used to generate revenues to offset the overall cost to the GOJ for the design, development, deployment, operation and maintenance of a Japanese GPS-compatible regional satellite system.

2.5 Ownership

Additional payloads and the opportunity to generate revenue, begs the question regarding who should own and operate a Japanese GPS-compatible regional satellite system.

The US government owns and operates GPS. However, a strong and practical argument can be made to allocate ownership and operation of a Japanese GPS-compatible regional satellite system to a commercial enterprise. This argument is essentially business based--upfront investment costs, revenue generation, return on investments and overall system sustainment. Counter arguments to the business based argument focus on States' responsibilities, and subsequent liabilities as well as continuity of the commercial enterprise. For instance, should a Japanese GPS-compatible regional satellite system be used to support aviation navigation, then principle number four of A32-19 (described earlier) applies.

However, it is quite possible that a cooperative government-industry arrangement could be established in order to accommodate Japanese government responsibilities for assuring GNSS services while offsetting government spending by taking advantage of commercial investments.

3. Options

Real-world issues as highlighted in the previous section, have established a path for the Japanese decision makers to follow when formulating an approach for developing a GPS-compatible Regional Satellite System capable of serving Japan as well as all of Asia. The significance and importance of these issues:

- Cost: initial and long-term commitments
- Time: GPS is here today--we need an independent and complementary system now
- Compatibility: GPS is being used today, more of the same is better and smarter
- Opportunities: GPS is a bus--it can carry other payloads
- Ownership: government, industry or both--it is a question.

were viewed differently by Japanese government and industry representatives. These differences of opinion had a positive effect because they resulted in several program options for a Japanese GPS-compatible regional satellite system. In the end, these options provided a basis for moving forward with a comprehensive Japanese program.

A brief overview of two viable options will be provided prior to discussing the current status of the Japanese program.

Option A: Japanese Regional Navigation Satellite System (JRANS)

In September 2000, the JRANS concept was developed by a Japanese industry partnership of ITOCHU Corporation, NEC Corporation and TOSHIBA Corporation. JRANS conceptual briefings were provided to several Japanese government representatives as well as US government and industry personnel working with GPS.

Based on the positive feedback from these initial briefings, ITOCHU and NEC TOSHIBA Space Systems, Ltd. ("NTSpace", a joint venture between NEC and TOSHIBA formed in April 2001 to merge their respective space business) continued working and discussing the JRANS concept with Japanese and US government and industry personnel. The JRANS concept and developmental approach was further refined to satisfy current and future operational requirements and assure full compatibility and interoperability with GPS. Highlights of the JRANS concept are:

- Fully complementary, interoperability and compatibility with GPS
- Capable of autonomous navigation and complementary / regional backup for GPS
- Satellite coverage will be regional (i.e. over Asia, see figure 1.)
- Free of direct user charges (like GPS)
- Private sector can participate and provide commercial services

The proposed JRANS program is a two-phase build-up of quasi-zenith (QZO), then another quasi-zenith and geostationary orbiting satellites (QZO and GEO), see figure 1.

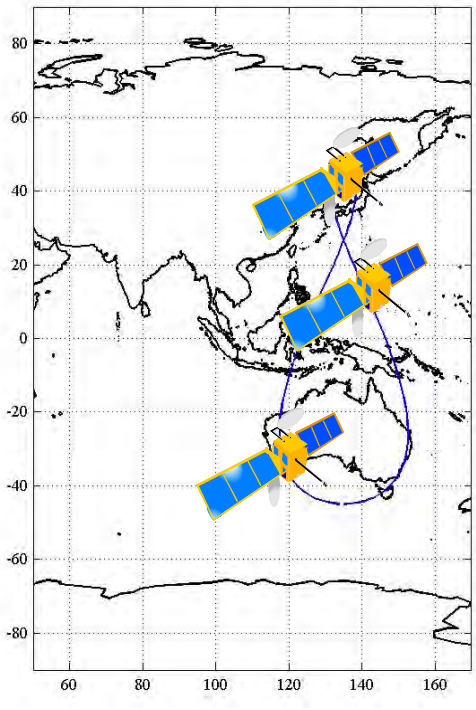


Fig. 1a JRANS: PHASE 1--three Satellites Quasi-Zenith Orbit

For a satellite navigation receiver to calculate a solution, four satellite signals in view with good geometry must be received to determine latitude, longitude, altitude and time. Satellite signals are “line-of sight” transmissions and can be easily blocked by high terrain, buildings, etc. This blocking of the signals is referred to as masking. Use of GPS in Japan can be difficult because of this masking situation. Natural geographic features, such as mountainous terrain and manmade features, such as tall buildings often render GPS services unavailable in the most critical of situations.

The advantage of a fully populated (7 satellite) constellation is having four satellites in view, at high mask angles, broadcasting GPS-type information, being controlled by Japan, and in coordination with the U.S. under the bilateral security treaty.

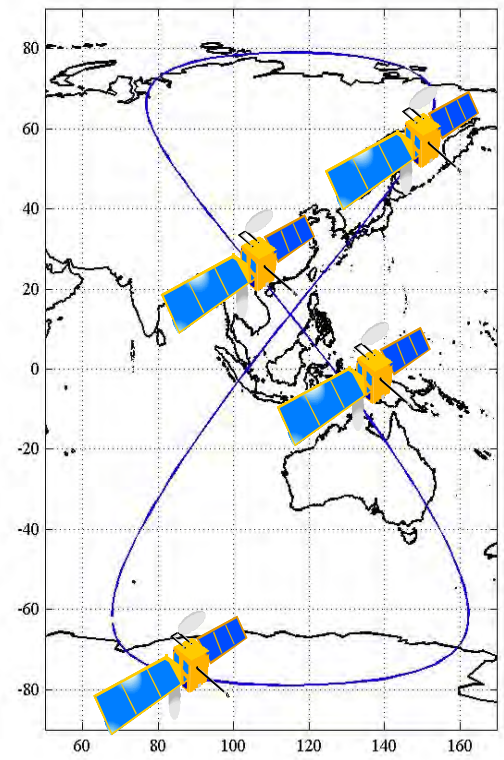


Fig 1b JRANS: PHASE 2—four Satellites QZO and GEO

Having four Japanese GPS-compatible satellites in view at higher elevations is particularly beneficial for those operating in mountainous areas and "urban canyons", as illustrated by Figure 2:

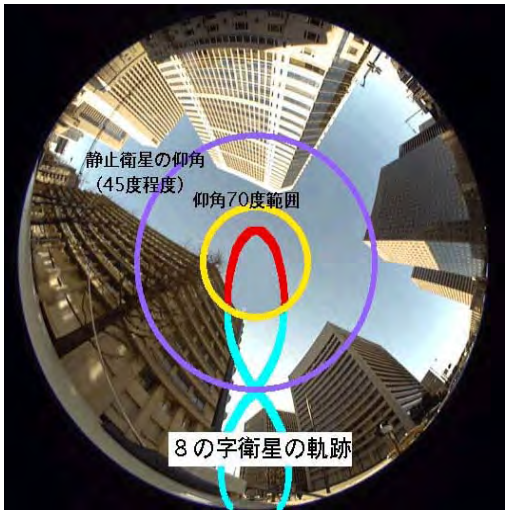


Fig 2: An example of masking in the urban canyon.

For GNSS users in Asia, there will also be a significant improvement in overall end state user performance as a result of better geometric dilution of precision (GDOP). Recall, GDOP is all geometric factors that degrade the accuracy of position fixes derived from externally referenced navigation systems.

Specifically, for a user in Tokyo Japan, using only GPS, the GDOP for a 10 degree mask angle is 2.39; the GDOP for a 30 degree mask angle is 6.88. That same user, at the same location, using GPS and JRANS together would have a GDOP of 1.80 at a 10 degree mask angle and a GDOP of 3.59 with a 30 degree mask angle. Given that a 30 degree mask angle is typical for Japanese urban areas (Figure 2), the improvement (from a GDOP of 6.88 with GPS only to 3.59 with GPS and JRANS) will improve the overall performance accuracy thereby increasing the utility of satellite-based position, velocity and time.

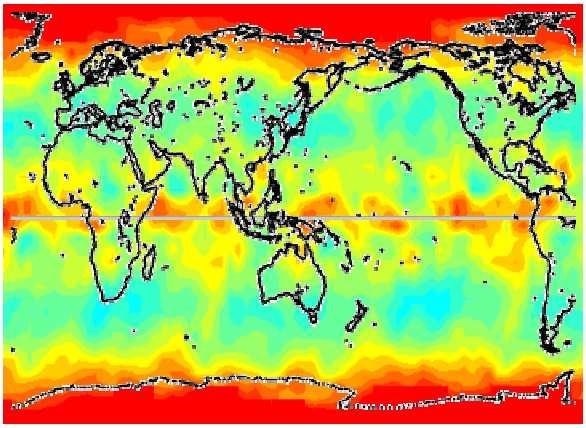


Figure 3: GDOP: GPS Only

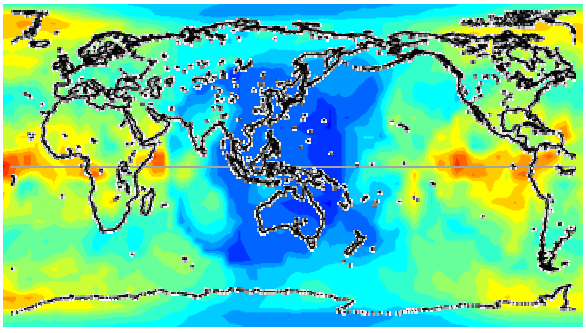
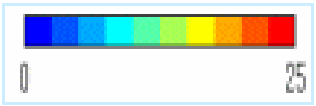


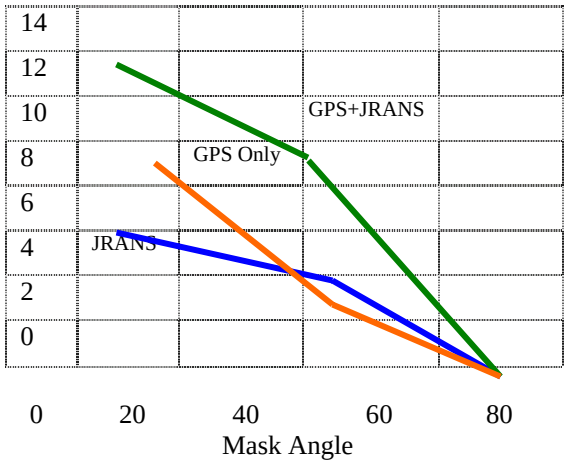
Figure 4: GDOP: GPS and JRANS



GDOP Scale

When one takes into consideration both JRANS signals and GPS signals, the combined benefits for an end state user--in terms of available satellite signals--are quite significant, as illustrated in Figure 5:

Figure 5: Satellites in View



Option B: Quasi-Zenith Satellite System (QZSS)

The QZSS option is actually the first phase of JRANS, as mentioned above, three satellites in quasi-zenith orbit (Figure 1).

In June 2002, the GOJ's Council for Science and Technology Policy of the Cabinet Office gave the go-ahead to begin working on QZSS research and development. The government role can be classified as research and development. The plan is to design and develop the first three QZSS satellites and the budget (US$52M in FY2003 and US$77M in FY2004) has been approved respectively. Several Japanese government agencies are involved in this first phase, to include:

- MEXT (Ministry of Education, Culture, Sports, Science and Technology): Experimental satellite positioning technology
- MIC (Ministry of Internal Affairs and Communications): Precise timing control and communication
- METI (Ministry of Economy, Trade and Industry): Key technologies for advanced satellite bus
- MLIT (Ministry of Land, Infrastructure and Transportation): High-accuracy DGPS augmentation system

A notional QZSS development schedule has been prepared and is expected to be used for planning and budgeting purposes. Overall, the GOJ is planning to invest approximately JPY50B (US$450M) in research and development funding for QZSS during the period 2003-2009. The notional schedule is:

- FY 2003: Definition Phase (US$52M budget approved)

- FY 2004: Research & Development Phase (US$77M budget approved)
- FY 2005-2008: Engineering & Manufacturing Phase
- FY 2009: 1st Satellite Launch

From a commercial perspective, and in addition to the basic navigation functions that will be fully compatible with GPS, QZSS can provide communication services, broadcasting services and differential GPS services. In November 2002, in response to these commercial opportunities, Japanese industry jointly formed the Advanced Satellite Business Corporation (ASBC) to conduct the feasibility study for determining the opportunities of using QZSS to provide commercial services, such as S-band communications and broadcasting for mobile users.

4. Current Status

As indicated in the previous section, the GOJ has committed resources to begin the design and development of a GPS-compatible Regional Satellite System capable of serving Japan as well as all of Asia. It is also evident that the design and development road will be lengthy without support from the US.

Both Japanese government and industry recognize the importance of working cooperatively with US government and industry. As early as April 2002, ITOCHU and NTSpace have been making public presentations in the US regarding Japanese planning efforts and opportunities for US-Japanese government and commercial cooperation.

From a government-to-government perspective, on 22 September, 1998 a US-Japan GPS partnership was forged when President William Clinton and Prime Minister Keizou Obuchi issued a joint statement regarding cooperation in the use of the GPS standard positioning service. Taking the statement, US-Japan GPS Plenary Meetings have been held to further harmonize joint activities. In addition, the Plenary agreed to form a Joint Technical Working Group to further the close cooperation between the US and Japan The Working Group goals are essentially to:

- Assure maximum QZSS interoperability with GPS
- Optimize the QZSS design to maximize GPS-QZSS performance in Asia
- Increase commercial opportunities for GPS-QZSS applications

5. Future Outlook for Asia

GPS reliance cannot be denied; neither can the significance of GPS for a broad spectrum of Asian users. In general, the aviation community appears to have a well established set of performance requirements for satellite-based navigation and are actively pursuing the development and implementation of civil augmentation systems. However, these "wide area coverage" augmentation systems:

- US: Wide Area Augmentation Systems (WAAS)
- Europe: European Geostationary Navigation Overlay System (EGNOS)
- Japan: MTSAT Satellite-base Augmentation System (MSAS)
- India: GPS/GLONASS and Geostationary Augmented Navigation (GAGAN)
- Australia: Ground-based Regional Augmentation System (GRAS)

All have one thing in common: the US GPS. If GPS signals "go away" then the utility of these augmentation systems will be close to zero.

The Japanese government and industry personnel have carefully studied the significance of GPS on our daily lives, considered critical enabling issues, such as cost, time, GPS compatibility, additional business opportunities, and public / private ownership.

A decision was made to move forward with a GPS-compatible regional system capable of providing independent, satellite-based position, velocity and time services while taking advantage of the broad range of benefits available from GPS signals and the current GPS industry..

For all of Asia, it is the most important fact that (in the near future) a combined US GPS and a standalone Japanese Regional Navigation Satellite System (RNSS) will provide a robust foundation for current and future generations of GPS users!

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 265-272

Journal of Global Positioning Systems
2004

Treatment of Biased Error Distributions in SBAS

Todd Walter, Juan Blanch, and Jason Rife

Stanford University, Stanford, CA 94305-4035

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. The original protection level equations for SBAS assumed that all actual error distributions could be easily overbounded by zero-mean gaussian distributions. However, several error sources have since been found that could lead to significant biases for specific users. The expectation is that over long periods of time and all users, the aggregate errors should have a very small mean. However, certain users, at specific times or locations, may have significant biases in their measured pseudoranges. One source of bias is signal deformations. Originally thought of as a failure mode, it is now recognized that geostationary satellites have a noticeably different signal than the GPS satellites (primarily due to their bandwidth limit). Recent results also show that the GPS satellites have measurable differences from satellite to satellite as well. The magnitude and sign of the biases depend on the user equipment and have been shown to have significant unit-to-unit variation. A biased distribution may be overbounded by a zero mean gaussian, provided the sigma value has been sufficiently increased. As the bias becomes larger, this inflation leads to a greater loss of availability than if the protection level equations had explicitly accounted for it. It is therefore important to find the smallest possible inflation to adequately bound the bias. This paper makes use of new overbounding methods to relate the required inflation to the bound.

Key words: SBAS, bias, overbound, integrity, CDF

1 Introduction

The SBAS signal specification (RTCA 2001) (ICAO 2000) defines how information is transmitted from the ground system to the user. Because this link uses a geostationary stationary satellite with a signal structure nearly identical to GPS, the data capacity is very limited (250 bits per second). Additionally, it was originally felt that all differential GPS error sources would be essentially zero mean (Walter *et al.* 1997). Consequently, the integrity information sent to the user contains no explicit provisions for protecting against biases. Instead users are sent protection factors that correspond to zero-mean error distributions. The users combine the received protection factors using their own local knowledge to calculate Protection Levels (PLs) that correspond to their position estimate. The broadcast protection factors must be sufficient such that any individual user has less than a one in ten million chance, for each approach, that their true position error exceeds the calculated PL. The ground system must guarantee these protection factors without knowing precisely where the users are, or which satellites they observe.

Unfortunately, it has recently been demonstrated that many sources of unobservable biases exist. Among these are nominal signal deformations (Phelts *et al.* 2004a) (Phelts *et al.* 2004b) and antenna group delay variations (Shallberg and Grabowski 2002). These sources create biases that are transparent to the ground system and may be unique to each user. The nominal signal deformations create biases that are dependent on the user's receiver RF filter and correlator spacing. Although it may be possible to restrict user designs or to calibrate the effect, manufacturing tolerances will still result in some non-negligible bias. The antenna group delay variations create repeatable biases in the raw GPS observables used by the ground network. These biases are always present in the measurements and thus are not easily determined. Consistent values across the antennas can lead to a biased satellite clock estimate for example. Calibration is also difficult due to limitations of anechoic chamber measurements and the effects of manufacturing variations. Regardless there still will be some residual bias effects that must be taken into account.

Recently it was noted that the two MHz bandwidth of the INMARSAT geostationary satellites (GEOs) creates a signal that is fundamentally different from the GPS

signals (Phelts *et al.* 2004a). This difference can create user specific biases on their psuedorange measurements. The SBAS signal specification currently has no provision for the user to remove their specific value. Consequently, this bias must be protected by the broadcast User Differential Range Error (UDRE) term for each GEO. Unfortunately, the narrowband GEO bias may be several meters. However, it is desirable to still send as small a UDRE as possible. This paper analyses the case of combining a few measurements with large biases together with many distributions with smaller biases. It provides a means for treating each distribution separately to minimize any increase on the good distributions. This analysis is based on the recently developed technique of excess mass bounding (Rife *et al.* 2004a) (Rife *et al.* 2004b) (Rife *et al.* 2004c).

2 Excess Mass Overbounding

A long-standing problem in SBAS and GBAS integrity analysis is overbounding, or providing an upper bound for a particular error distribution. The problem is even more challenging when combining many error sources together. There are two related issues at stake here: the first is practical, what is the true error distribution given a limited amount of data; the second is analytical, how to best represent and combine the error distributions. This paper will address the second part only. To broadcast the information to the user on a limited data channel, each error distribution must be represented in a simple functional form. However, this simplified form must predict at least as much mass in its tails as the true distribution. The error bounds predicted by the simplified form must be as large as the true bound. This concept is called overbounding.

Originally, the requirement to have increased mass at the tails of the distribution meant that the central part of the distribution would require decreased mass, as both distributions had the same total area under the curve. However, it was recently recognized that the overbounding distribution did not have to integrate to unity. Instead, it could have excess mass at both the tails and in the central core. This would result in a loss of performance, but it would allow the analysis to proceed.

A further requirement is that when the multiple error sources are combined together, there exists a method for combining the individual overbounds such that the convolution of errors is overbounded. This overbounding has to hold for all users for any geometry they might have.

2.1 Excess Mass PDF Bounding (EMP)

Given an actual error distribution, $g_a(y)$, we wish to select an overbounding PDF such that

$$g_o(y) \geq g_a(y) \quad \forall_y \tag{1}$$

This has the property that the overbounding distribution has excess mass, i.e., it integrates to a value greater than one.

$$\int_{y=-\infty}^{\infty} g_o(y)dy \geq 1 \tag{2}$$

The convolution of two excess mass overbounding distributions will overbound the convolution of the two original distributions, as all values are positive.

$$\begin{aligned} h_a(z) &= \int_{y=-\infty}^{\infty} f_a(z-y)g_a(y)dy \\ h_o(z) &= \int_{y=-\infty}^{\infty} f_o(z-y)g_o(y)dy \end{aligned} \tag{3}$$

has the property that

$$h_o(z) \geq h_a(z) \quad \forall_z \tag{4}$$

By induction, this can be extended to the general case of convolutions of multiple weighted distributions.

2.2 Excess Mass CDF Bounding (EMC)

One practical difficulty with PDF overbounding is that it requires the overbounding PDF to be everywhere larger than the actual PDF. However, an experimental PDF may have "spikes", or if the data is interpreted as being a delta function at each observed point, then it can require a very large amount of excess mass to overbound it at every point. Instead, it is preferable to integrate over regions to smooth out the actual PDF. It is possible to specify a weaker constraint that has the desirable features of EMP, and works well with real data. This latter process is called Excess Mass CDF (EMC) bounding and was developed in Rife *et al.* (2004b). One can establish left and right CDF bounds according to

$$\begin{aligned} G_L(x) &= \int_{-\infty}^{x} g_o(y)dy \\ G_R(x) &= 1 - \int_{x}^{\infty} g_o(y)dy \end{aligned} \tag{5}$$

where now the only requirement is that

$$G_R(x) \leq G_a(x) \leq G_L(x) \tag{6}$$

where $G_a(x)$ is the corresponding CDF of $g_a(y)$. This has the property of bounding the total mass at each tail. In

addition, it has been demonstrated by Rife *et al.* (2004b), that this property is maintained through convolution. Therefore, both EMP and EMC satisfy our requirements for predicting at least as large an error at the tail as the true distribution and maintaining this property through convolution. Further, neither method places constraints on the true underlying distribution beyond those specified in Equations (1) or (6). This is a nice feature as the actual distribution may not be symmetric or unimodal which were two requirements of the original overbounding analysis (DeCleene 2000).

3 Application to SBAS

To see how these overbounding concepts may benefit SBAS, one must understand the Vertical Protection Level (VPL) equation. The VPL equation specifies how the protection terms for each individual error component are to be combined to find the upper bound on the positioning error along the vertical axis. The WAAS VPL equation (see Appendix J of RTCA 2001) has the user combine a series of broadcast σ values and multiply them by a term, $K_{V,PA}$, corresponding to the expected probability for a unit-variance zero-mean gaussian (Walter *et al.* 1997) (RTCA 2001)

$$VPL = K_{V,PA}\sqrt{\sum_{i=1}^{N} s_{U,i}^2 \sigma_i^2} \tag{7}$$

where $s_{U,i}$ depends upon the user's geometry and σ_i are formed from overbounding sigmas broadcast to the user. Each individual σ_i is made up of four error terms

$$\sigma_i^2 = \sigma_{i,\,flt}^2 + \sigma_{i,\,UIRE}^2 + \sigma_{i,\,air}^2 + \sigma_{i,\,tropo}^2 \tag{8}$$

The first two terms are based on values broadcast to the user, the third term bounds the local aircraft's thermal and multipath error, and the final term is a standard value specified in the MOPS. The *flt* term stands for fast and long term corrections. It bounds the satellite clock and ephemeris error terms and is derived from broadcast UDRE and degradation parameters. The UIRE term stands for User Ionospheric Range Error and is based on the interpolated value of the individually broadcast Grid Ionospheric Vertical Error (GIVE) terms.

The requirement is that the VPL equation will bound the true error to the desired probability for any $s_{U,i}$

$$P\left(\left|\sum_{i=1}^{n} s_{U,i}\varepsilon_i\right| > VPL\right) \le P_{HMI} \quad \forall_{s_{U,i}} \tag{9}$$

where ε_i are errors drawn from error distributions $g_i(y)$, and P_{HMI} is the allowable probability of Hazardously Misleading Information (HMI). For this application, P_{HMI} is 10^{-7} per approach. The VPL equation does not directly allow for biases or excess mass distributions.

We can choose a zero-mean gaussian as the excess-mass overbounding distribution. For any real distribution $g_a(y)$, define an overbound such that

$$g_o(y) = K\mathcal{N}_y(0,\sigma_o) = \frac{K}{\sigma_o\sqrt{2\pi}} e^{-y^2/(2\sigma_o^2)} \tag{10}$$

K represents the excess mass of the distribution, that is

$$\int_{y=-\infty}^{\infty} g_o(y)dy = K \tag{11}$$

The next sections define methods for finding σ_o and K given what is known about the actual distributions.

3.1 Bounding Non-Zero Mean Gaussian Distributions

Assume for this section, that an actual bounding gaussian exists with some actual mean and sigma: μ_a, σ_a. The first goal is to find an overbounding zero-mean excess mass gaussian with sigma σ_o. The requirement for excess mass PDF bounding is that

$$K\mathcal{N}_y(0,\sigma_o) \ge \mathcal{N}_y(\mu_a,\sigma_a) \tag{12}$$

This can be rewritten as

$$\frac{K}{\sigma_o} e^{-\frac{y^2}{2\sigma_o^2}} \ge \frac{1}{\sigma_a} e^{-\frac{(y-\mu_a)^2}{2\sigma_a^2}} \tag{13}$$

Given values for σ_o, μ_a, and σ_a, this equation can be used to formulate a constraint on the minimum allowable value of K. This constraint is given by

$$K_{PDF\_\min} = \frac{\sigma_o}{\sigma_a} e^{\frac{\mu_a^2}{2(\sigma_o^2-\sigma_a^2)}} \tag{14}$$

Thus, for any σ_o such that $\sigma_o > \sigma_a$, there exists a minimum K that satisfies (14). Note that this is not specifying the minimum value of K across all possible values for σ_o, but rather merely the minimum value for a specific σ_o. This condition creates an overbounding PDF that is tangent to the actual PDF at one point and above it at all other points. Given a μ_a and σ_a, the next goal is to choose the best combination of K and σ_o. The best combination is the one that will ultimately minimize the VPL for the user. Therefore, one must look at the predicted error limit corresponding to P_{HMI}. If the bias and sigma information could be transmitted to the user, the error bound for N error sources would be given by

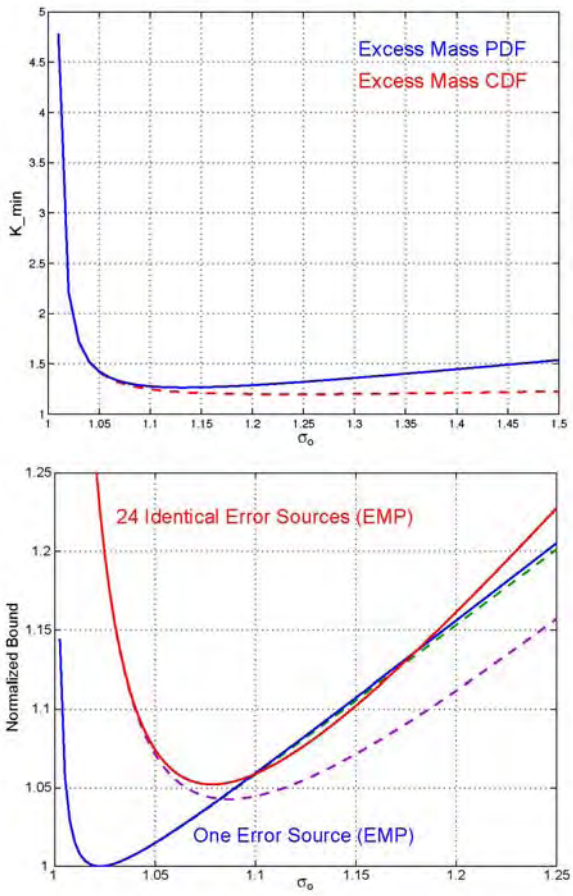


Figure 1 (a and b). The top figure (a) shows the minimum K value as a function of σ_o. The bottom figure (b) shows the normalized bound also as a function of σ_o.

$$\sum_{i=1}^{N}\left|S_{U,i}\cdot\mu_{a,i}\right|+A(P_{HMI})\sqrt{\sum_{i=1}^{N}S_{U,i}^2\cdot\sigma_{a,i}^2} \tag{15}$$

where, for a zero mean gaussian, A is related to the normal CDF and is given by

$$A(P_{HMI})=Q^{-1}(1-\frac{P_{HMI}}{2})=\sqrt{2}\,\mathrm{erfc}^{-1}(P_{HMI}) \tag{16}$$

The value $K_{V,PA}$ = 5.33 in the VPL equation corresponds to the Probability of HMI of 10^{-7} in (16). Equation (15) represents the lowest possible VPL that could be transmitted to the user and will be used as a reference point to judge later implementations.

The excess mass in the distributions must be taken into account. A K value of two indicates there is twice as much mass in the tails compared to a normal distribution (as well as in the core) and therefore errors beyond a certain value are twice as likely. This number directly scales the P_{HMI}. The zero-mean, excess mass bound is therefore given by

$$A\left(P_{HMI}/\prod_{i=1}^{N}K_i\right)\cdot\sqrt{\sum_{i=1}^{N}S_{U,i}^2\cdot\sigma_{o,i}^2} \tag{17}$$

where the P_{HMI} is lowered by the product of the K values. The K's and σ_o's can then be chosen to minimize this bound.

An example is provided in Figure 1. In this example the following non-dimensional values are set: σ_a = 1 and μ_a = -.25. Figure 1a shows the minimum K value as a function of σ_o. For small values of σ_o, K must be large to ensure bounding at the tail. As σ_o increases, K hits a minimum value and then starts to increase again. This increase is now because the larger σ_o value would otherwise fail to bound the core of the distribution. Figure 1b shows the bound from (17) normalized by the ideal bound (15) as a function of σ_o. The solid blue line in the figure corresponds to an individual distribution. For a single distribution, a small σ_o is the choice with a corresponding large K value. As can be seen, it is possible to pick values for σ_o and K that match the ideal lower bound. The solid red line corresponds to 24 identical distributions convolved together. Here the best choice is to accept a larger σ_o value with a correspondingly smaller K. This result is logical as K^N can grow quite rapidly while the RSS of the σ_o terms grow much more slowly.

The next goal is to use these overbounds within the confines of the VPL equation. As can be seen from (7) and (17), if the broadcast sigma, σ_B, is chosen such that

$$\sigma_B\geq\frac{A\left(P_{HMI}/\prod_{i=1}^{N}K_i\right)}{K_{V,PA}}\cdot\sigma_o \tag{18}$$

Then the VPL requirement will be met. Thus, it is not possible to safely transmit the σ_o values, but by inflating each by at least the amount in (18) one can find the safe broadcast values σ_B. A nice feature of this inflation is that it is independent of the details of the user's geometry ($s_{U,i}$). The penalty is that every distribution must be increased by at least this same amount, even the non-biased ones. Thus, the broadcast value can be found in this two step approach, first find an excess mass zero mean to bound the individual distributions, then uniformly inflate the first overbounding sigma, so, as shown above to find the broadcast value that will work in the VPL equation.

In the above example, σ_o was chosen to be 1.08 for 24 identical distributions with a corresponding K value of 1.3. Therefore, the broadcast value, σ_B, must be $A(10^{-7}/1.3^{24})/5.33=A(1.84\times10^{-10})/5.33=1.2$ times larger than σ_o. Thus, a broadcast value of σ_B = 1.3 is sufficient to bound 24 distributions with unity variance and an absolute mean value of 0.25.

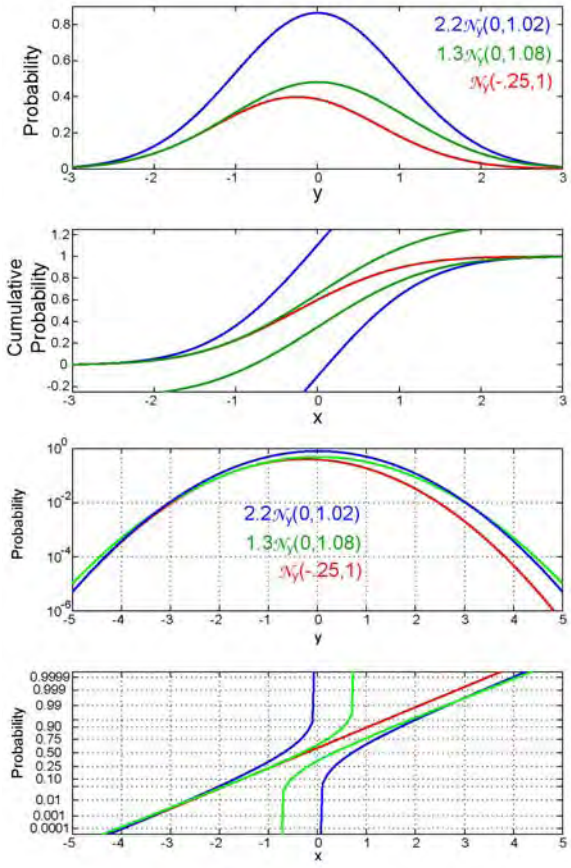


Figure 2 (a and b). The top two graphs (a) show the PDF (upper) and CDF (lower) for the two optimum cases in Figure 1b. The blue line corresponds to the optimal single error source bound while the green line is optimal for 24 identical distributions. The red line corresponds to the actual distribution. The bottom two graphs (b) are the same plots but on different scales to help see that the bounds (blue and green) always stay above the red on the top plots, and to the left and right on the bottom plots.

3.2 Applying Excess Mass CDF Bounding

The prior analysis used EMP bounding to derive appropriate bounds. The requirement for excess mass CDF bounding are analogous to (13) and can be written in terms of the complimentary error function as

$$\frac{K}{2}\operatorname{erfc}\left(-\frac{x}{\sqrt{2}\sigma_o}\right) \geq \frac{1}{2}\operatorname{erfc}\left(-\frac{x+|\mu_a|}{\sqrt{2}\sigma_a}\right) \quad \forall_x \tag{19}$$

Although there is not a convenient closed form expression for K, it can be found numerically from the maximum value of the ratio

$$K_{CDF\_\min} = \max\left(\frac{\operatorname{erfc}\left(-\frac{x+|\mu_a|}{\sqrt{2}\sigma_a}\right)}{\operatorname{erfc}\left(-\frac{x}{\sqrt{2}\sigma_o}\right)}\right) \tag{20}$$

Against biased gaussian distributions, K_{CDF_min} will always be smaller than K_{PDF_min}, although the difference will be small. All of the remaining equations above will work for EMC, except that a slightly smaller value for K may be used. Figures 1 and 2 show the comparison of the EMC vs. EMP. In Figure 1a, the EMC K value, shown with the dashed line, is slightly smaller. This improvement increases for larger values of σ_o. The main difference is that the growth at larger σ_o values is noticeably smaller. This is because the integration of the mass at the tails is nearly sufficient to cover the core distribution. In Figure 1b, the dashed lines represent the bound from (17) normalized by the ideal bound (15) as a function of σ_o for EMC. For the single source case (dashed green line) there is little difference, both can achieve the ideal bound (The CDF of the overbounds and the actual distributions are tangent at 5.33). However, for the 24 identical distribution case, there is a small improvement. Now the optimal value of σ_o is closer to 1.09 and the bound is only about 4% greater than ideal compared to 5% for EMP. Although EMC achieves lower bounds and is more practical with real data, the rest of this paper will continue to analyze EMP bounding. This is because of the closed form equation possible with EMP (14). It is worth remembering that this adds a small amount of conservatism in the analysis. If one applied EMC instead, one would have a small improvement. This results in a small margin against integrity.

3.3 Validating Broadcast Values

In the WAAS safety analysis, the algorithms were derived and then tested against real data. The data is used to validate the performance of the algorithms. Because the MOPS only allows discrete broadcast values for the UDRE (14 numeric values) and GIVE (15 numeric values), it is convenient to partition the data by σ_B. Thus, rather than trying to find a σ_B value given a real distribution, the important question becomes: Does the chosen σ_B value sufficiently bound the actual histogram?

Because the feared biases are completely transparent to the system, they are not present in the validation data. Thus, the observed histograms are all nearly zero-mean. The feared biases are bounded by separate analysis and then included in this methodology. One can write that the actual gaussian bounding parameters as fractions of the desired broadcast value, that is:

$$\sigma_a = \alpha \cdot \sigma_B, \quad \mu_a = \gamma \cdot \sigma_B \tag{21}$$

where α is between 0 and 1. The notation in this paper is slightly different from that in Rife *et al.* (2004c). This paper looks in terms of reduction from the broadcast value rather than inflation from the actual, thus $\alpha = 1/\Theta$.

From the histogram of data, it is possible to determine values of α and K that bound the data. However, γ will be unobserved as the system will have removed known biases. In Schempp and Rubin (2002) they determine values for α (labelled σ in their paper) for each of the observed UDRE and GIVE values. The observed biases are very small and can be treated as zero. The important question is given the observed α's how much margin is there for unobserved biases? The goal is to find the largest tolerable biases for given values of α.

It is useful to define σ_o also in terms of σ_B. From (18) it can be seen that another parameter, η, can be defined such that

$$\sigma_o = \eta \cdot \sigma_B,$$
$$\eta \equiv \frac{K_{V,PA}}{A\left(P_{HMI} / \prod_{i=1}^{N} K_i\right)} = \frac{K_{V,PA}}{\sqrt{2}\,\mathrm{erfc}^{-1}\left(P_{HMI} / \prod_{i=1}^{N} K_i\right)} \tag{22}$$

This parameter represents an upper bound on α. This parameter is related to the ξ parameter in Rife *et al.* (2004c) as $\xi = \eta/\alpha$. As the product of the K's increases, so must the margin between σ_o and σ_B. Figure 3 shows a plot of η versus the product of K's. Fortunately, it is a weak function and the product can grow quite large while decreasing η by only a small amount.

To find a bound on γ, Equation (14), can be rewritten as

$$K_{PDF\_\min} = \frac{\eta}{\alpha} e^{\frac{\gamma^2}{2\left(\eta^2 - \alpha^2\right)}} \tag{23}$$

and inverted to yield

$$\gamma_{\max} = \sqrt{2\left(\eta^2 - \alpha^2\right)\ln\left(K_{PDF\_\min}\frac{\alpha}{\eta}\right)} \tag{24}$$

However, K_{PDF_min} depends on the other parameters. An expression for it can be found by making a few other assumptions. First, the product of the K's can be expressed as a function of η, by inverting its definition in (22)

$$\prod_{i=1}^{N} K_i = \frac{P_{HMI}}{\mathrm{erfc}\left(\frac{K_{V,PA}}{\sqrt{2}\eta}\right)} \tag{25}$$

Next assume that the K product consists of two parts: a fixed allocation for nearly zero mean distributions (K_{other}), and a remainder that is evenly split among N satellites. This is the key piece of the approach as it allows for some satellites, such as the GEOs, to be treated differently than the others. Then K_{PDF_min} is given by

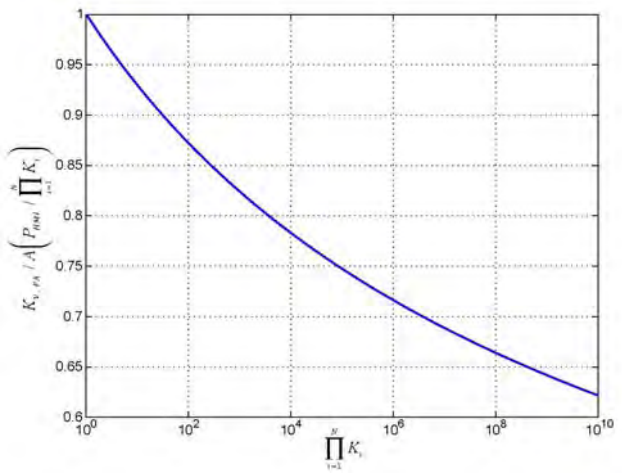


Figure 3. The value η is shown as a function of the product of the K's. As can be seen it is a slow to decrease with very large K values.

$$K_{PDF\_\min} = \left(\frac{P_{HMI}}{\mathrm{erfc}\left(\frac{K_{V,PA}}{\sqrt{2}\eta}\right)K_{other}}\right)^{1/N} \tag{26}$$

This value can be substituted into (24) to yield the final expression for the maximum tolerable bias

$$\gamma_{\max} = \sqrt{2\left(\eta^2 - \alpha^2\right)\ln\left(\left[\frac{P_{HMI}}{\mathrm{erfc}\left(\frac{K_{V,PA}}{\sqrt{2}\eta}\right)K_{other}}\right]^{1/N}\frac{\alpha}{\eta}\right)} \tag{27}$$

Since α is known from the histogram and P_{HMI}, $K_{V,PA}$ and K_{other} are fixed parameters, this function can be plotted versus η to determine the maximum value.

An example is now provided to show how this process may be used. WAAS has collected extensive data in support of its certification in 2003 (Raytheon 2002). This data has been examined in many ways, it has been found that the largest α_{GPS} for a particular σ_B is .65 (Schempp 2004) for GPS UDRE and GIVE values. The goal is to determine the largest allowable bias on the GEO satellites. For GPS and ionospheric error distributions, allow K values up 1.15 to account for small biases and other non-gaussian behaviour. Assuming a 12-channel receiver and two narrowband GEOs, allows 22 GPS or ionospheric induced errors and 2 GEO errors. The parameter K_{other} is set to $1.15^{22} = 21.64$. For the GEO the α_{GEO} is either 0.7 for a minimum UDRE value of 7.5 m or 0.35 for a minimum value of 15 m. Figure 4 shows the results for setting K_{other} = 21.64 and α_{GEO} = 0.7 (a) or α_{GEO} = 0.35 (b), and assuming one or two GEO satellites.

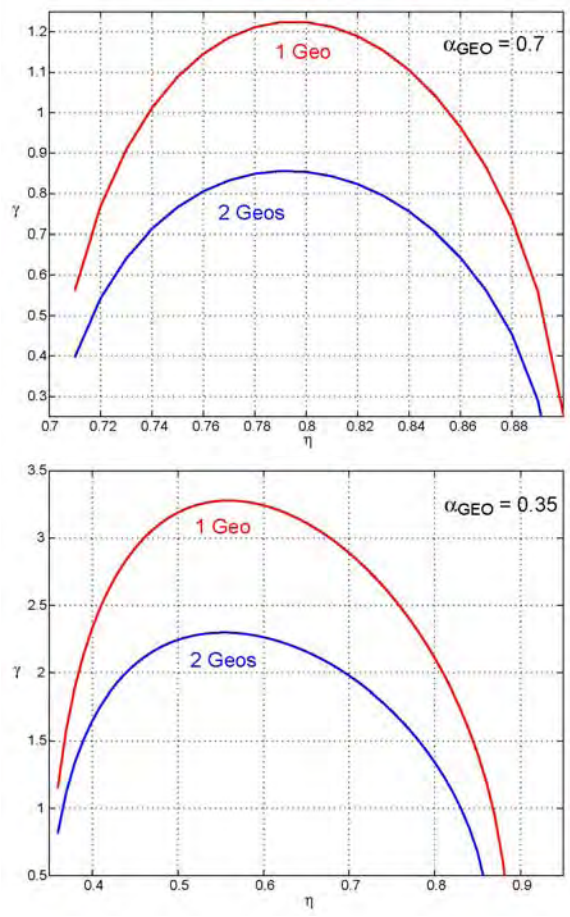


Figure 4 (a and b). The top figure (a) shows the maximum γ value as a function of η for two cases with α_{GEO} = 0.7. The bottom figure (b) shows the maximum γ value as a function of η for two cases with α_{GEO} = 0.35.

As can be seen from the figures, larger biases can be tolerated for smaller α_{GEO} values and for fewer numbers of geostationary satellites. For the UDRE of 7.5 m (σ_B = 2.28 m) the maximum γ value occurs near η = 0.8 and is approximately 1.2 (μ = 2.74 m) for one satellite, or 0.85 (μ = 1.94 m) for two. For the UDRE floor of 15 m (σ_B = 4.56 m) the maximum γ value occurs near η = 0.55 and is approximately 3.3 (μ = 15 m) for one satellite, or 2.3 (μ = 10.49 m) for two. However, in the latter example, η cannot be allowed to go as low as 0.55 as α_{GPS} must be above 0.65. This additional constraint from the GPS satellites implies that the lowest allowable η is about 0.7. Thus, the true maximum biases for the UDRE floor of 15 m are approximately 2.9 (μ = 13.22 m) for one satellite or 2.0 (μ = 9.12 m) for two.

4. Conclusions

This paper took the concepts of excess mass bounding and applied them specifically to the case when a few biased error distributions are combined with many well-behaved ones. Specifically the case for WAAS with two narrowband GEOs was examined. It was shown that a methodology exists for determining the maximum tolerable bias under certain constraints. This methodology shows that based on the WAAS validation data, biases as large as 2.74 m can be tolerated for a GEO UDRE value of 7.5 m, or as large as 13.22 m for a GEO UDRE value of 15 m. These value decrease to 1.94 m and 9.12 when two GEOs may be visible.

Excess mass bounding provides a methodology for formally accounting for biases that may be present. This methodology is fully compatible with the WAAS MOPS VPL equation that is based on zero-mean gaussian error distributions. CDF bounding in particular will work with a wide variety of actual distributions that are non-zero mean and non-gaussian. The methodology is also very flexible in allowing certain distributions to be treated differently from the others. Each distribution may be bounded with unique K and α values. This method allows for the accommodation of larger biases than allowed by other methods.

Acknowledgements:

The authors would like to acknowledge the funding from the FAA satellite navigation product team..

References

DeCleene B (2000) ***Defining pseudorange integrity – overbounding,*** Proceedings of ION GPS 2000, Salt Lake City, Utah, September

ICAO (2000) ***Annex 10 — Aeronautical Telecommunications.Volume I (Radio Navigation Aids),*** ICAO publication November 2000

Phelts RE, Walter T, Enge P, Akos DM, Shallberg K, and Morrissey T (2004a) ***Range biases on the WAAS geostationary satellites,*** Proceedings of the ION National Technical Meeting, San Diego, California, 26-28 January 2004

Phelts RE (2004b) ***Nominal signal deformation: limits on GPS range accuracy,*** Proceedings of the 2004 International Symposium on GPS/GNSS, Sydney, Australia, 6-8 December

Raytheon (2002) ***Algorithm contribution to HMI for the wide area augmentation system,*** Raytheon Company Unpublished Work.

Rife J, Pullen S, Pervan B, and Enge P (2004a) ***Paired overbounding and application to GPS augmentation,*** Proceedings IEEE Position, Location and Navigation Symposium, Monterey, California, April

Rife J, Pullen S, Pervan B, and Enge P (2004b) ***Core overbounding and its implications for LAAS integrity,*** Proceedings of ION GNSS 2004, Long Beach, California, September

Rife J, Walter T, and Blanch J, (2004c) ***Overbounding SBAS and GBAS error distributions with excess-mass functions,*** Proceedings of the 2004 International Symposium on GPS/GNSS, Sydney, Australia, 6-8 December

RTCA (2001) ***Minimum operational performance standards for global positioning system/wide area augmentation system airborne equipment,*** RTCA publication DO-229C.

Schempp TR and Rubin AL (2002) ***An application of Gaussian overbounding for the WAAS fault free error analysis,*** Proceedings of ION GPS 2002, Portland, Oregon, 24-27 September

Schempp TR (2004) ***A range domain bound on the GEO unobservable bias,*** Raytheon Company unpublished work, June 2004

Shallberg K and Grabowski J (2002) ***Considerations for characterizing antenna induced range errors,*** Proceedings of ION GPS 2002, Portland, Oregon, 24-27 September

Walter T, Enge P, and Hansen A (1997) ***A proposed integrity equation for WAAS MOPS,*** Proceedings of ION GPS 97, Kansas City, Missouri, 16-19 September

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 273-279

Journal of Global Positioning Systems
2004
© CPGPS

A Step, Stride and Heading Determination for the Pedestrian Navigation System

Jeong Won Kim

GNSS Technology Research Center, Chungnam National University, Korea
e-mail: kimjw@cnu.ac.kr Tel: + 82-42-821-7709 ; Fax: +82-42-823-5436

Han Jin Jang

GNSS Technology Research Center, Chungnam National University, Korea
e-mail: handoo01@cnu.ac.kr Tel: + 82-42-821-7709 ; Fax: +82-42-823-5436

Dong-Hwan Hwang

GNSS Technology Research Center, Chungnam National University, Korea
e-mail: dhhwang@cnu.ac.kr Tel: + 82-42-821-5670 ; Fax: +82-42-823-5436

Chansik Park

School of Electrical and Computer Engineering, Chungbuk National University, Korea
e-mail: chansp@cbucc.chungbuk.ac.kr Tel: + 82-43-261-3259 ; Fax: +82-43-268-2386

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Recently, several simple and cost-effective pedestrian navigation systems (PNS) have been introduced. These systems utilized accelerometers and gyros in order to determine step, stride and heading. The performance of the PNS depends on not only the accuracy of the sensors but also the measurement processing methods. In most PNS, a vertical impact is measured to detect a step. A step is counted when the measured vertical impact is larger than the given threshold. The numbers of steps are miscounted sometimes since the vertical impacts are not correctly measured due to inclination of the foot. Because the stride is not constant and changes with speed, the step length parameter must be determined continuously during the walk in order to get the accurate travelled distance. Also, to get the accurate heading, it is required to overcome drawbacks of low grade gyro and magnetic compass. This paper proposes new step, stride and heading determination methods for the pedestrian navigation system: A new reliable step determination method based on pattern recognition is proposed from the analysis of the vertical and horizontal acceleration of the foot during one step of the walking. A simple and robust stride determination method is also obtained by analysing the relationship between stride, step period and acceleration. Furthermore, a new integration method of gyroscope and magnetic compass gives a reliable heading. The walking test is preformed using the implemented system consists of a 1-axis accelerometer, a 1-axis gyroscope, a magnetic compass and 16-bit microprocessor. The results of walking test confirmed the proposed method.

Key words: Pedestrian navigation system, Step detection, Stride determination, Heading determination

1 Introduction

Pedestrian navigation system(PNS) provides velocity and position of a person and can be applied to many other areas such as E-911 service, location based services (LBS), tourism, rescue, military infantry, medical studies, leisure, and navigation for the blind. In PNS, it is necessary to locate the position of the user in any time and any environment. Even GPS is useful personal navigation system, its availability is significantly reduced when a signal is blocked. Also ultra wide band (UWB) and radio frequency identification (RFID) techniques are introduced for personal navigation, but these systems

require dense infrastructure. For these reasons, a self-contained navigation system based on a dead reckoning (DR) principle is of interest (B. Merminod et al, 2002). To locate the position of the PNS user, distance and heading from a known origin have to be measured with an acceptable level of accuracy. In PNS, an accelerometer is used to count the number of steps, which is combined with the stride to obtain the travelled distance. In addition, a magnetic compass or gyroscope is used as a heading sensor.

The stride and step are important parameters for PNS dead reckoning algorithm. Many methods have been suggested to detect a step. One such method is to detect the peaks of vertical acceleration, which correspond to the step occurrences because the vertical acceleration is generated by vertical impact when the foot hits the ground. If the vertical impact is larger than given threshold, it is considered as a step. Since the pattern of impact signal depends on type of movement (going up or down stairs, crawling, running etc.) and type of ground over which the person walks (hard or soft surface, sand), the determination of threshold is not so easy for reliable step detection (Ladetto and Merminod, 2002). This paper proposes reliable step detection method based on pattern recognition. From the analysis of the vertical and horizontal acceleration of the foot during one step of the walking, the signal pattern of walking behaviours is obtained.

The stride of the walker in PNS is a scale factor in a dead reckoning algorithm. Unlike a scale factor of an odometer in a car navigation system, the stride in PNS is a time-varying parameter (Mar and Leu, 1996). The predetermined stride cannot be used effectively for the distance measurement because the strides of the walker are different according to the human parameters. The stride depends on several factors such as walking velocity, step frequency and height of walker etc. As the stride is not a constant and can change with speed, the step length parameter must be determined continuously during the walk to increase the precision. It is suggested that the stride could be estimated online based on a linear relationship between the measured step frequency and the stride (Levi and Judd, 1996). A real-time step calibration algorithm using a Kalman filter with GPS positioning measurement was also proposed (Jirawimut et al,2003). In this paper, we analyse a relationship between stride, step period and acceleration to obtain simple and robust method of stride determination. A real time online estimation is possible by using only 1-aixis accelerometer.

The combination of gyroscope and magnetic compass has already been applied in car navigation (Mar and Leu, 1996) and it might be a very useful heading sensor for pedestrian navigation system. However low cost sensor has important drawbacks: A low cost gyro has large bias and drift error. The magnetic disturbances can be induced fatal compass error. Moreover the error is occurred by an oscillation of human body in a walking behaviour. In this paper, a gyro and a magnetic compass are integrated using Kalman filter for reliable heading of pedestrian.

To evaluate the performance of the proposed methods, actual walking test in the indoor environment is conducted. The equipment of walking test is implemented using a 1-aixs accelerometer, a 1-axis gyroscope and a magnetic compass. It consists of two parts: a sensor module and a navigation computer module. The sensor module is attached on the ankle. The step number and stride is computed using the output of the accelerometer on the sensor module. And walking direction is obtained from the gyro and magnetic compass module. The experiments show the very promising results: less than 1% step detection error, less than 5% travelled distance error and less than 5% heading error.

2 Step detection

2.1 Step behaviour analysis of pedestrian

A cycle of human walking is composed of two phases: standing and walking phase. The step detection means a recognition of walking phase. The walking phase is divided into a swing phase and a heel-touch-down phase. Each phase is shown in figure 1.

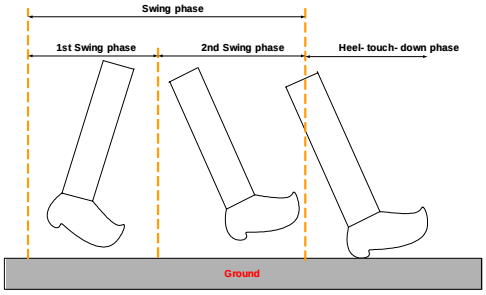


Fig. 1 A walking behaviour

In the 1st swing phase, the foot is located on behind of gravity centre of human body. And the foot is located on front of gravity centre of human body in the 2nd swing phase. The foot accelerated during swing phase. The acceleration is composed of vertical and horizontal components as shown in figure 2, where a, h, g means horizontal acceleration, vertical acceleration and gravity force, respectively.

Figure 3 and 4 show motion of leg in the 1st swing phase and the 2nd swing phase respectively.

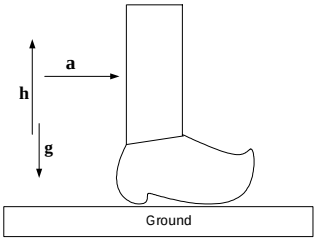


Fig. 2 Leg of walker

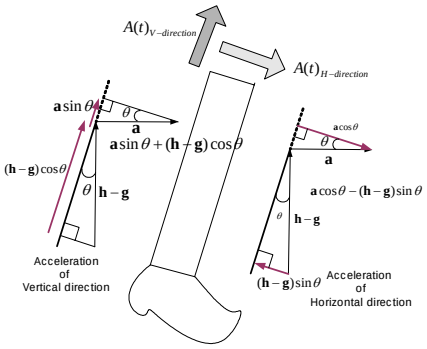


Fig. 3 1st Swing phase

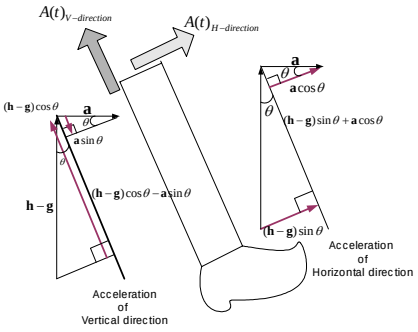


Fig. 4 2nd Swing phase

The horizontal direction acceleration and vertical direction acceleration during the swing phase is denoted in equation 1, where $\theta(t)$ is inclination angle of the leg at time *t*.

$$A(t)_{H-direction} = (h-g)\sin\theta(t) + a\cos\theta(t)$$
$$A(t)_{V-direction} = (h-g)\cos\theta(t) - a\sin\theta(t) \quad (1)$$

In many researches, a step is declared when the measured $A(t)_{H-direction}$ or $A(t)_{V-direction}$ is larger than the threshold. However since the $\theta(t)$ depend on characteristics of walking which is different from each person, it is hard to determine the exact value of threshold of $A(t)_{H-direction}$ or $A(t)_{V-direction}$. The step number is miscounted when wrongly predetermined threshold is applied. By using the signal pattern of acceleration, this problem can be solved. Typical signal pattern of acceleration is obtained from the computer simulation. We adopted common assumptions that a typical inclination of leg was within the limit of 30 degree ~ 50 degree and a, h have a range of 0.8 ~ 2.3g and 0.6 ~ 2.0g. The pattern of acceleration signal in figure 5 is obtained from 625 times simulations.

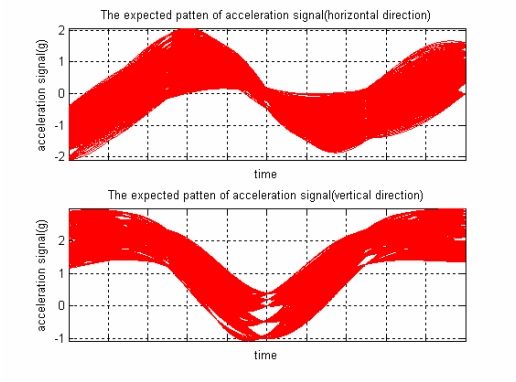


Fig. 5 Pattern of horizontal and vertical acceleration signal

Figure 5 shows the typical pattern of acceleration signal on the swing phase. The acceleration of horizontal direction has 1 positive peak and 1 negative peak in swing phase while the acceleration of vertical direction has 1 negative peak only.

The heel-touch-down phase follows the swing phase. A heel-touch-down is impact motion which hits the ground. In heel-touch-down phase, a heel hits the ground at first. And then a sole of foot and toe contact with the ground. When the foot hits the ground, the ground repulses the foot. At this time, impact force acts on the foot. The figure 6 shows the heel-touch-down phase.

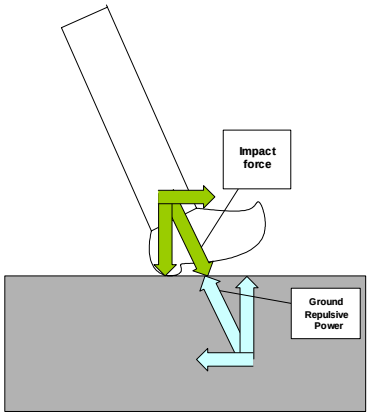


Fig. 6 Heel-touch-down phase

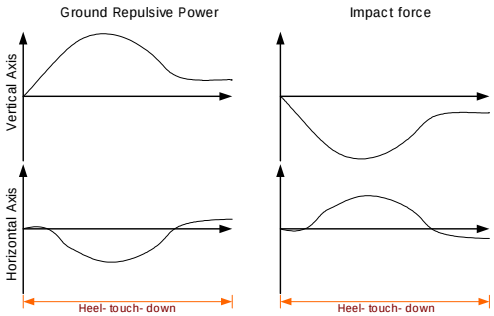


Fig. 7 The typical pattern of signal in heel-touch-down phase

Figure 7 shows typical repulsive and impact force patterns during the heel-touch-down phase.

By combining the swing phase and heel-touch-down phase in the figures 5 and 7, we obtain the signal pattern of one walking cycle. Figure 8 and 9 show entire signal pattern of the walking phase.

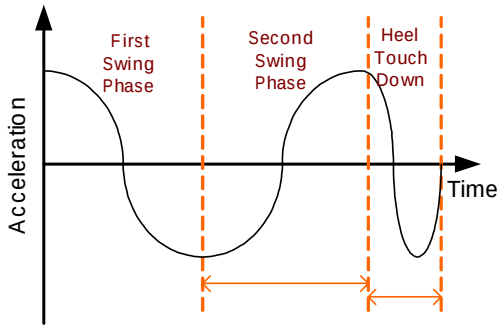


Fig. 8 Vertical acceleration signal pattern in walking phase

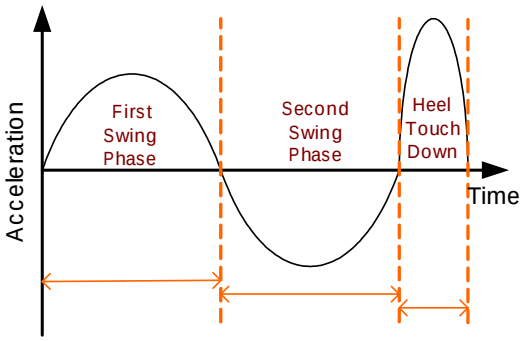


Fig. 9 Horizontal acceleration signal pattern walking phase

It is expected intuitively that the period of heel-touch-down phase is much shorter than the period of swing phase. The figure 10 shows a real horizontal acceleration signal in one step. It coincides with the signal pattern model in figure 9.

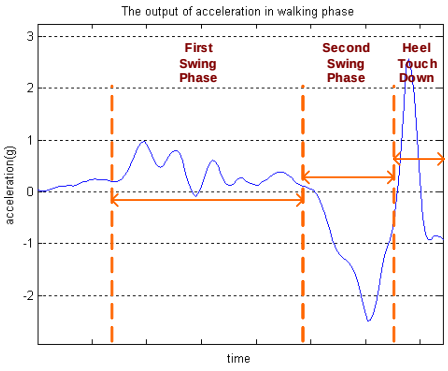


Fig. 10 Real horizontal acceleration signal

2.2 Step detection method

To discriminate one cycle of walking behaviour, the signal pattern of swing phase and heel-touch-down phase in figure 8 and 9 is adopted. The accelerometer measures the signal which is caused by walking behaviour. The step number is counted when all three phases (1st swing, 2nd swing and heel-touch-down phase) are detected. This method reduces step misdetection probability and increase reliability. Recognizing swing and heel-touch-down pattern using sequential multi-threshold gives a robust and reliable step detection. Also the method can reduce misdetection probability of non-walking behaviour such as sitting, turning, kicking and jumping etc. The detail detection algorithm is given in figure 11.

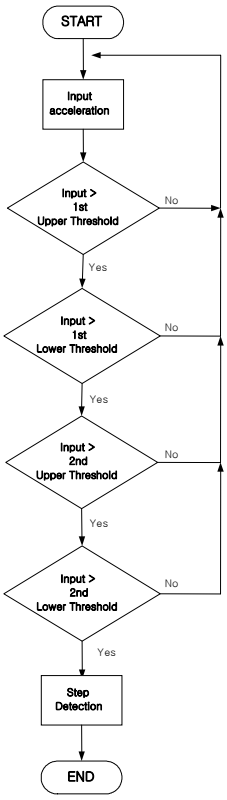


Fig. 11 Flow chart of step detection

3 Stride determination

Because the stride is not a constant value and changes with speed, the stride parameter must be determined continuously during the walk to increase its precision. The stride relates on walking speed, walking frequency and acceleration magnitude. In typical human walking behaviour, a period of one step becomes shorter, a stride becomes larger and the vertical impact becomes bigger as the walking speed increases. The relation between stride, period of one step and acceleration is established thru the

actual walking test. Figure 12 show test result of two type strides: 60 cm and 80 cm stride.

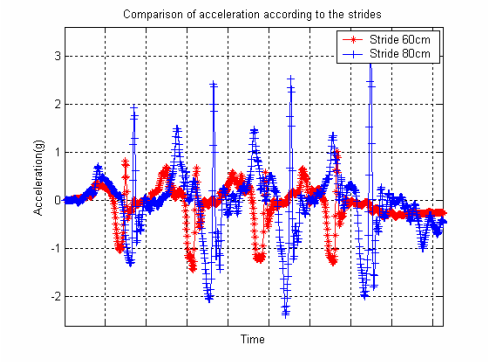


Fig. 12 The acceleration signal of 60 cm and 80 cm stride

The tester walks with the fixed stride using ground marks. In the figure, the relation between the acceleration and stride is clearly shown. The tables 1 and 2 show relation between acceleration and one step time in this test. The longer stride induces the bigger acceleration. However a difference of one step time is hard to apply stride determination because of small difference in measurements.

Tab. 1 The mean of acceleration absolute value

Stride	Mean value (g)
60cm	0.2882
80cm	0.5549

Tab. 1 The period of one step

Stride	Mean of time (sec.)
60cm	0.675
80cm	0.662

Equation 2 is the experimental equation obtained from several walking tests, where means the measured acceleration and represents the number of sample in one cycle of walking. The equation represents the relation between measured acceleration and stride. It is used for online estimation of the stride.

$$Stride(m) = 0.98 \times \sqrt[3]{\frac{\sum_{k=1}^{N} |A_k|}{N}} \qquad (2)$$

4 Heading determination

The gyroscope and magnetic compass is widely used to determine heading. The characteristics of two sensors are summarized in Table 3, where advantage of one sensor is disadvantage of the other.

Tab. 3 Comparison between compass and gyroscope

	Advantage	Disadvantage
Magnetic compass	-absolute azimuth -long term stable accuracy	unpredictable external disturbances
Gyro scope	-no external disturbances -short term accuracy	relative azimuth drift

From table 3, an optimal and reliable system might be expected by integrating the gyroscopes with the magnetic compass. In the integrated system, the gyroscope can correct the magnetic disturbances, at the same time the compass can determine and compensate the bias of the gyros and the initial orientation. The combination of gyroscope and magnetic compass has already been applied in the car navigation system. The integration method of the gyroscope and the magnetic compass used in this paper is given in Figure 13.

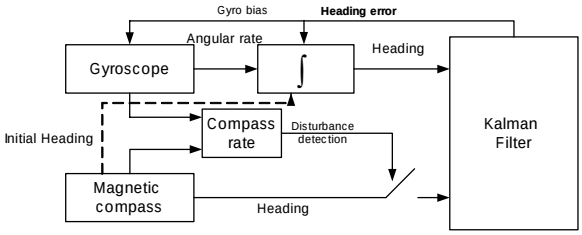


Fig. 13 Scheme for an integration of gyroscope and magnetic compass

When the pedestrian is walking, the influence of magnetic disturbance sources changes unpredictably, creating a error in the compass heading. This error degrades the performance of integration system. The impact of error can be reduced by detecting the disturbance. The error can be observed via the angular rate of compass heading:

$$\omega_{compass} = \frac{\psi_{compass}(t_k + \Delta t) - \psi_{compass}(t_k)}{\Delta t} \qquad (3)$$

where ω is angular rate, ψ is heading and Δt is the time interval. The disturbance can be detected when a difference of compass angular rate $\omega_{compass}$ and gyroscope angular rate ω_{gyro} is larger than given threshold. The compass measurement is ignored. The states of Kalman filter are heading error and sensor error (gyro bias).

5 Experiments

In order to evaluate the performance of the proposed method, the actual walking test is done. The tester is a male aged 26 with 175cm height. The experiments are done at the 4th floor hallway of the engineering building, Chungnam National University, Daejeon, Korea. In the experiments, walking distance determination and heading determination are carried out separately.

5.1 Experimental setup

Figure 14 shows the experimental equipments.

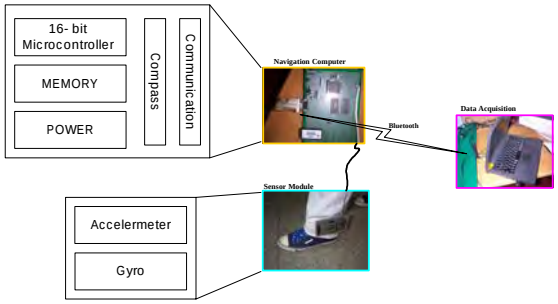


Fig. 14 Experimental equipment

The experimental equipments consist of the sensing module, the navigation computer and a data acquisition system (notebook computer). The body-worn sensing module consists of a 16-bit microcontroller, a MEMS accelerometer (ADXL105, Analog device Inc.), a gyroscope (MEMS DMU, Crossbow Inc.), a low-cost digital magnetic compass sensor (CMPS03, ROBOT Electronics Inc.) and other electrical parts (RS-232 converter, DC-DC converter, 9V battery, Bluetooth modem). The sensor module is attached on the ankle with horizontal direction as shown figure 14.

5.2 Experiment of walking distance determination

To evaluate performance of the step detection and stride determination algorithm, the tester was asked to walk for pre-determined path (74.2m and 145.6m straight path). Figure 15 shows the output of accelerometer.

The true step number of first test is 100 steps and second test 200 steps. The stride is determined using equation 2. The figure 16 and 17 show the strides of left leg.

The mean of estimated stride is obtained as 76.1 cm and 75.9 cm respectively. Table 4 shows result of walking test in detail.

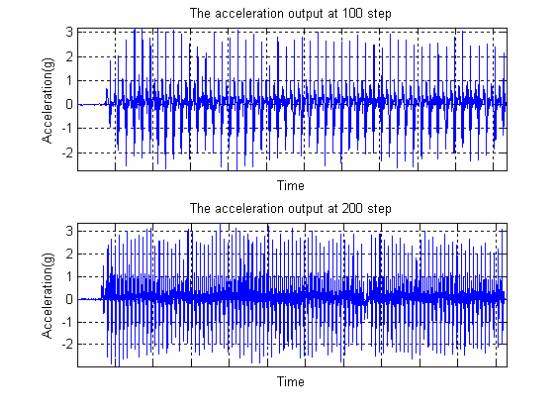


Fig. 15 The output signal of accelerometer

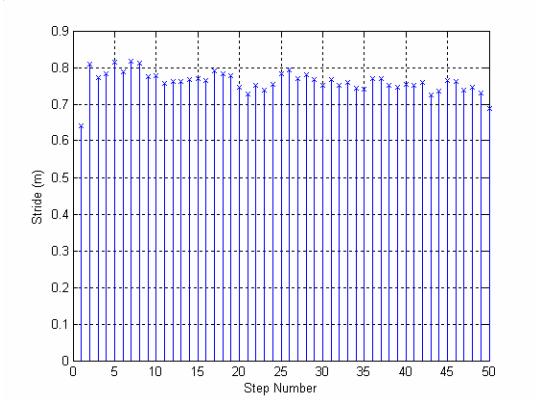


Fig. 16 Estimated stride in 1st test

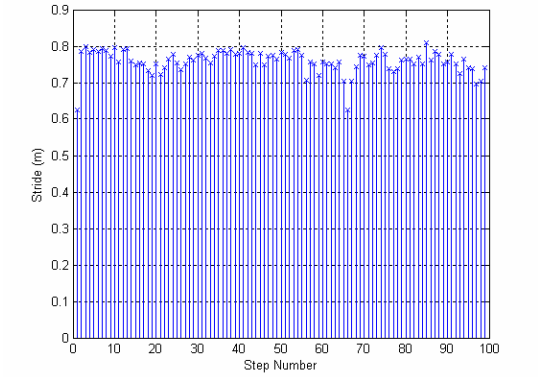


Fig. 17 Estimated stride in 2nd test

In the 1st test, the proposed method count step number without loss, while the 2 step detection is lost in 2nd test. The 2 step loss is happened in the last 199th and 200th step where the walking pattern is abruptly changing. The walking distance error is obtained 2.5m, 6.1m respectively. The travelled distance with less than 5% error is obtained. These results verify that the proposed method can measure accurate step numbers and distance.

Tab. 4 The measured walking distance

Actual walking behavior			Measured step number	Measured walking distance
1st test	Step number	100 step	100 step	76.728 m
	Walking distance	74.2m		
2nd test	Step number	200 step	198 step	151.674 m
	Walking distance	145.6m		

5.3 Experiment of heading determination

For heading determination test, the tester walks a straight path of north direction. Figure 18 is result of heading determination test.

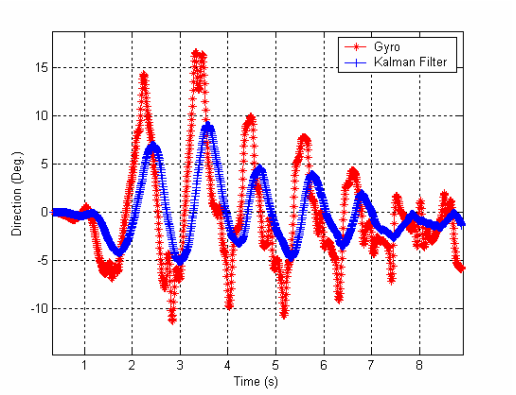


Fig. 18 Estimated heading by gyro and by integration

In the figure, the heading of stand-alone gyro shows oscillatory errors due to the body motion. Kalman filter in the integrated system reduces these errors. The experiments show that the heading of pedestrian can be determined with accuracy of 5 degree.

6 Conclusions and outlook

This paper proposes methods to estimate the PNS DR parameters: step, stride and heading. For accurate step detection, we analyse the vertical and horizontal acceleration of the foot during one step of the walking. With this analysis, a new step determination based on the pattern recognition is proposed and the step number can be counted accurately. The relationship between stride and acceleration is derived from actual test. An efficient stride determination method where the stride can be estimated online, so that the user does not need to specify his/her stride, is proposed. The integration scheme of the gyro and magnetic compass is proposed for error compensation of gyro and disturbance rejection of magnetic compass. The experiments using the actual walking tests in indoor shows that the proposed method gives less than 1% step, 5% travelled distance and 5% heading errors. It is expected that the proposed PNS will be very useful navigation system for pedestrian navigation.

References

Gabaglio, V., (2003): ***GPS/INS Integration for Pedestrian Navigation*** Ph. D. dissertation. Institute of Geomatics of the Swiss Federal Institute of Technologye in Lausanne.

Mar, J., (1996): ***Simulations of the positioning accuracy of integrated vehicular navigation systems.*** In: J.-H. Leu(Eds.): Proc. Inst. Elect. Eng. Radar, Sonar Navigation, vol. 143, Apr., 121–128.

Ladetto, Q., (2002): ***In Step with INS.*** In: B. Merminod (Eds.):GPS WORLD magazine, 30-38

Quentin Ladetto (2000): ***On foot navigation: continuous step calibration using both complementary recursive prediction and adaptive Kalman filtering.*** Proceedings of ION GPS 2000, 1735~1740.

Jirawimut, R., (2003): ***A Method for Dead Reckoning Parameter Correction in Pedestrian Navigation System.*** In: P. Ptasinski; V. Garaj; F. Cecelja; W.Balachandran (Eds.):IEEE Transactions on Instrumentation and Measurement , Vol. 52, No.1.

Levi, R. W., (1996): ***Dead Reckoning Navigational System using Accelerometer to Measure Foot Impacts.*** In: T. Judd,(Eds.): United States Patent No. 5,583,776

Lee, S.-W., (2001): ***Recognition of Walking Behaviours for Pedestrian Navigation.*** In: K. Mase,(Eds.): Proc. 2001 IEEE Int'l Conf. Control Applications (CCA 01), IEEE Control Systems Soc., Piscataway, N.J., 1152–1155.

Gabaglio, V., (1999): ***Real-time calibration of length of steps with GPS and accelerometers.*** In: B. Merminod (Eds.):Proceeding of GNSS 99, 599–605.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 280-289

Journal of Global Positioning Systems
2004
© CPGPS

GNSS for sports – sailing and rowing perspectives

K.Zhang, R. Deakin, R. Grenfell, Y. Li, J. Zhang, W.N. Cameron, D.M. Silcock

School of Mathematical and Geospatial Sciences, RMIT University, Melbourne, Australia
e-mail: Kefei.zhang@rmit.edu.au; Tel: + +61-3-99253272; Fax: +61-3-96632517

Received: 9 December 2004 / Accepted: 3 February 2005

Abstract. This paper introduces two sport-related projects conducted by the Satellite Positioning and Orientation Research Team (SPORT) at RMIT University – the speed sailing world record challenge and development of a smart GPS rower tracking system. In the first project, both traditional and contemporary surveying technologies are investigated to assist the Macquarie Speed Sailing Team to reliably record and subsequently claim a world speed sailing record. In the second project, an integrated rower tracking system has been developed in collaboration with other research partners and the system has been used prior to and during the Athens Olympic Games. Three Olympic rowing medals were won by Australia. The technology, research procedures and major developments are presented.

Key words: GNSS for sports, low-cost GPS, speed sailing, rowing

1 Introduction

The School of Mathematical and Geospatial Sciences RMIT University has been involved in a number of sports related research projects in the past few years. Two notable projects are: (1) the world sailing speed challenge and (2) Development of a miniaturised, high precision GPS rower monitoring and coaching system. The aim of first project was to identify the best technology to help the Macquarie Speed Sailing Team improve their sailing speed recording. The second project is funded by the Australian Cooperative Research Centre (CRC) for microTechnology, under Project 2.5 "Interface Technologies for Athlete Monitoring". An aim of Project 2.5 is to develop monitoring equipment that is essentially unobtrusive such that the athlete is virtually unaware of its presence in training and competition. This also includes investigation into the acquisition and interpretation of sport's data and its meaningful presentation to both coaches and athletes (CRC microTechnology, 2003).

1.1 World sailing speed record project

World Sailing Speed Records are awarded by the World Sailing Speed Record Council (WSSRC, 2002), an affiliated body of the International Sailing Federation (ISAF, 2002). Eligible records are average velocities over a 500-metre course by a yacht whose only method of propulsion is the natural action of the wind on the sail. A record will only be ratified if the attempt has been monitored by a commissioner appointed by the ISAF/WSSRC. The course may be defined by floats on the water or by transit posts on shore and a timed run is the difference between start and finish times recorded to the nearest one hundredth of a second. Speed, distance divided by time, is calculated to the nearest one hundredth of a knot with allowance made for the resolved component of any tidal stream and/or current flow on the course. A course is deemed unsuitable if the tidal flow and/or current exceed one knot.

This paper presents an analysis of both GPS and video methods of determining sailing velocity. Simulations of GPS position recording and video recording were made during a recent record attempt by mounting a GPS receiver and the video camera on the Macquarie team's support boat (a power boat) and making three runs along the marked course. Comparing the GPS-derived velocity with the video-derived velocity is a useful means of benchmarking GPS velocity against an approved ISAF/WSSRC technique. In addition, a Kalman filter is used to verify a simple method of determining approximations to instantaneous velocity from kinematic GPS.

1.2 CRC for micro technology rower tracking project

The five-year project, "Development of a miniaturised, high precision GPS rower monitoring and coaching system", initially investigated the feasibility of using GPS to aid inertial devices for Position, Velocity and Acceleration (PVA) determination in real-time. The positional information is then combined with other athlete physiological information and integrated into a dedicated electronic device to package, analyse and relay the information to the coach. The continuous monitoring of three-dimensional PVA of the rowing boat with a very high frequency and accuracy is achieved through an integration of GPS, an inertial navigation system with sensors for athlete physiological information, a data communication mechanism, and an interactive visualisation procedure.

The ability to measure and record athlete physiological information (eg. aerobic capacity, strength training and endurance performance) and positional information associated with athlete movement in real-time is critical in the process of athlete training and coaching. Blood oxygen (oxygen consumption), respiration, heart rates (myocardial and hemodynamic responses), velocity, acceleration/force, changes in direction and position, and many other factors are required in elite athlete training and coaching. The position, movement (i.e. velocity and direction) and force (i.e. acceleration) information plays an important role in effective analysis of the athlete performance, especially for rowers. For example, the stroke rate, force and synchronisation of athletes are critical for the performance of rowers in a competition.

Currently, the stroke information can only be measured in either well-controlled situations in dedicated sports laboratories or using simulation devices. Much of the equipment is either too heavy, expensive or obtrusive and multiple factors which are difficult to control have limited the use of sport-specific field testing. Reliable analysis of the stroke rate and stroke length in rowing has been a challenge for a long time due to the unavailability of the real scenario data, in particular high-precision PVA data (Larsson, 2003). Existing technologies used for this purpose include theoretical studies (eg. Zatsiorsky and Yakunin, 1991), video-footage procedure (Kleshnev, 2001), indoor tank procedure (Lin et al., 2003), computer modelling (Atkinson, 2003; Kirtley and Smith, 1996) and ergometer (Upson, 2003; Elliott et al., 2003). Therefore, smart real-time monitoring during training and competition to help elite athletes to improve their performance and avoid injuries is critical for both athletes and coaches.

This paper introduces an innovative "smart" GPS rover tracking system that is currently under development at SPORT, RMIT. This paper first outlines the main phases in a standard rowing stroke for a better understanding of the specific technique. The paper then briefly summarises the proposed research methodologies for the project and presents some selected results of recent tests. Promising results have been obtained using low-cost GPS. It has shown that low-cost, code-only GPS can provide great potential for rowing applications and the stroke signature captured from GPS is of great value for rowers and their coaches. A prototype system has been built and the system was extensively used by Australian rowers prior to and in the Lucerne Regatta and Athens Olympic Games where a number of gold medals were won.

2. World sailing speed records

World Sailing Speed Records (WSSRC, 2002) are established in sail area divisions:

- ✓ 10 Sq. m Class: up to and including 10 m^2;
- ✓ A Class: from 10 m^2 up to and including 150 square feet (13.94 m^2);
- ✓ B Class: from 150 square feet up to and including 235 square feet (21.83 m^2);
- ✓ C Class: from 235 square feet up to and including 300 square feet (27.87 m^2); and
- ✓ D Class: over 300 square feet.

The fastest of these class records also becomes the outright World Sailing Speed Record.

In 1993 at Shallow Inlet, Simon McKeon and Tim Daddo, sailing the triplanar wing sail yacht *Yellow Pages Endeavor*, captured the B, C and D Class WSSRs on the way to recording the highest speed ever attained by any craft under sail of 46.52 knots. They achieved these multiple records by adjusting the area of the wing sail between runs. *Yellow Pages Endeavour* was designed and constructed by Lindsay Cunningham and raced by a group of volunteers. In March 2002, the group, then known as the Macquarie World Speed Sailing Team attempted to raise their own record above 50 knots with a new yacht *Macquarie Innovation* (see Fig. 1). This yacht, an improvement on *Yellow Pages Endeavour,* is a solid wing sail attached by aerofoil sections to three small pontoons, one of which contains the skipper and crew. In the right wind and sea conditions, with the skipper steering and the crew trimming the wing sail, the crew capsule lifts clear of the water and the yacht rises slightly, planing on the pontoons. Small hydrofoils under the pontoons assist the hull planing and offer side force resistance as well as steering control from the front pontoon.

Both *Macquarie Innovation* and *Yellow Pages Endeavour* are developments of the speed potential exhibited by C Class catamarans, the yachts that contest the Little America's Cup – a challenge cup, like the America's Cup, where a sailing club challenges the cup holder (another sailing club) to a series of match races. McCrae Yacht Club on Port Phillip Bay, Victoria won the cup in 1985

with *Victoria-150* and defended it successfully until 1996. *Victoria-150*, designed and built by Lindsay Cunningham, was the first C Class to effectively use a multi-slotted aerofoil wing section as a sail (Landmark, 2002).

2.1 Time recording of the sailing record attempts

Timing of the record attempt by the Macquarie team employed a "clever, low-technology" solution developed by the team in previous attempts and now embodied in the ISAF rules. It uses a video camera mounted in the crew capsule and aimed to capture images of the transit posts onshore. The camera is capable of recording images at 25 frames per second. The transit posts, placed 8m apart in pairs, define start and finish lines of the 500m course (see Fig. 2). As the yacht passes the start line the video camera records an image of two starting posts in transit and some time later, crossing the finish line,

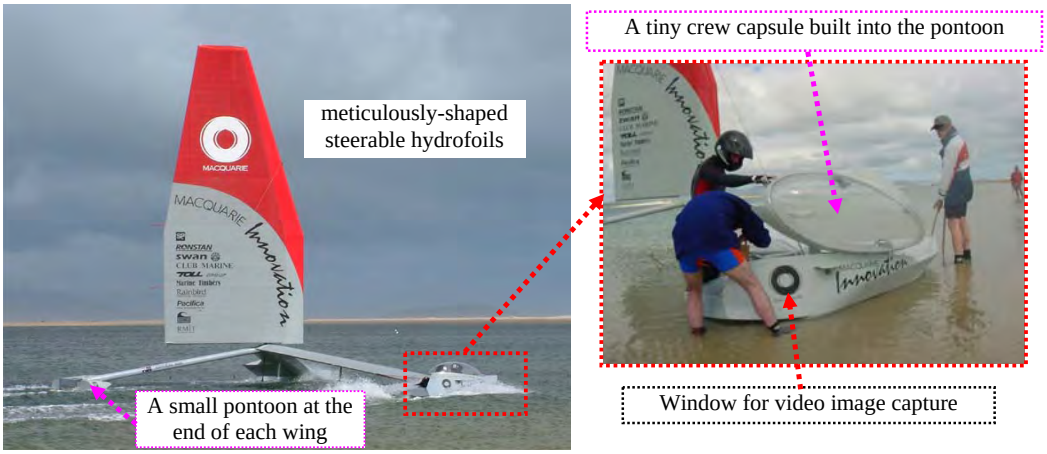


Fig. 1 Images of the Macquarie Innovation – a triplanar asymmetrical wing-sail yacht

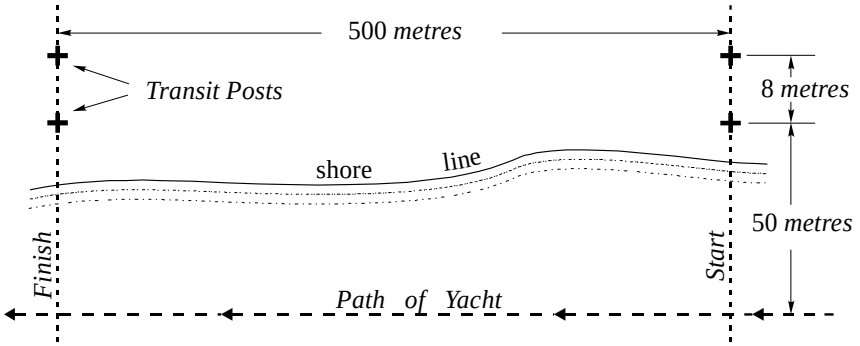


Fig. 2 Schematic plan of transit posts defining the course

(a) video image frame "*n*" just before crossing finish line

(b)transit post pair surveyed by both total station and GPS

(c) video image frame "*m*" just after crossing start line

Fig. 3 Diagrammatic view of the transit posts surveyed by both conventional and GPS methods (b) and video image frames "*m*" and "*n*" at the start (c) and finish (a)

records another image of the finish posts in transit (see Fig. 3).

After a run on the course, the video is reviewed frame by frame, noting the frame numbers *m* and *n* showing the start and finish transits. Subtracting frame numbers and multiplying by the frame rate gives the elapsed time, which divided into the distance yields the average speed.

2.2 GPS and the transit post recording procedure

This timing technique, video camera and transit posts described above, presents some difficulties and restrictions for competing teams. These include:

- The positioning of the transit posts is critical; not only must they produce parallel transit lines but these lines must be at least 500m apart and approximately perpendicular to the path of the yacht. This requirement necessitates a survey to mark the positions of the posts and a plan endorsed by a surveyor to satisfy the ISAF/WSSRC commissioner.
- The yacht is restricted to sailing close to the shore (and transit posts) so that video images of transits are clear and distinct.
- Since the yacht must travel relatively close to the transit posts and at a great speed, there is some danger to the crew if the yacht becomes uncontrollable and collides with the transit posts.
- The direction of the shore (and hence the course) limits the allowable wind direction since a yacht's maximum potential speed is restricted to a narrow range of wind angles from the direction of travel.
- The video camera must be calibrated to ensure an accurate frame rate.
- Reviewing the video images is a time consuming process and is subject to human error (errors in visually interpreting the actual transit).

Recognising these restrictions, the Macquarie team approached the School of Mathematical and Geospatial Sciences, RMIT, with a proposal to investigate the use of on-board GPS as a more flexible means of determining sailing velocity and distance. GPS has the following attractions and possible advantages over the present method.

- GPS is a proven robust positioning technology, well documented in the surveying and geodetic literature. When used in *kinematic differential* mode, GPS is capable of determining positions at centimetre level precision at precise and regular time intervals as small as 0.1 sec.
- Kinematic Differential GPS positioning removes the restriction of marked courses. Positions can be determined at time intervals, say $\Delta t = 0.1 \sec$, independent of the yacht's sailing direction. Differences in position divided by time differences, yield velocity. In addition, simple velocity plots can be used to determine which section of a yacht's speed record attempt should be used to determine average velocity.

Simulations of GPS position recording and video recording were conducted during a recent record attempt by mounting a GPS receiver and the video camera on the Macquarie team's support powerboat and making three runs along the marked course. Table 1 shows a comparison of the GPS-derived velocity with the video-derived velocity which is a useful means of benchmarking GPS velocity against an approved ISAF/WSSRC technique. In addition, a Kalman filter is used to assess the precision of kinematic GPS positions and verify a simple method of determining approximations to instantaneous velocity from kinematic GPS. Comparing the average velocities derived from GPS observations with those derived from Video Camera observations shows that the average differences for the three courses are 0.017m/s for velocity and -0.026s for time respectively. It is therefore concluded that GPS is a viable alternative to the video camera technique.

3. Science of rowing and athlete tracking

Rowing is a highly developed, and becomes an increasingly popular, international sport. It combines a wonderful spectacle with a heated competition. Rowing races usually cover a distance of 2,000m in river, canal or lake-type competition environments in six lanes. To win the competition, atheltes have to qualify through four pre-determined rounds: the preliminary round (heats), the repeat round (repechages), the semi-finals and the finals. The "A" final determines the first six places and the runners-up; the "B" final determines the next six places (ie 7th to 12th positions). The number of rounds per event depends on the number of crews taking part.

The races are judged under the supervision of umpires, who are members of the Jury for the event. The Jury members are placed at various locations on and off the competition course, such as the starting line, where the races begin under the supervision of the aligner and the starter; along the course of the race in the competition lanes under the supervision of umpires; the finishing line with the finish-line umpire; the identity verification stage of the crews before their embarkation onto the boats; the weighing-in of the athletes; the weighing-in of boats; and, in general, in all areas directly related to the competition, the athletes and their equipment (Athens Olympics, 2004).

There are 14 different boat classes raced in Olympic rowing. These include eight sculling events in which two oars are used (see Fig. 4a), one in each hand, and six sweep-oared events in which the rowers use one oar with

Table 1 A comparison of velocity and time derived from kinematic GPS and video camera

Course	GPS		Video camera		Differences	
	Velocity (m/s)	Time (seconds)	Velocity (m/s)	Time (seconds)	Velocity (m/s)	Time (seconds)
A3	17.28	28.94	17.30	28.90	0.02	-0.04
A4	17.34	28.83	17.37	28.78	0.03	-0.05
A5	17.30	28.89	17.30	28.90	0.00	0.01

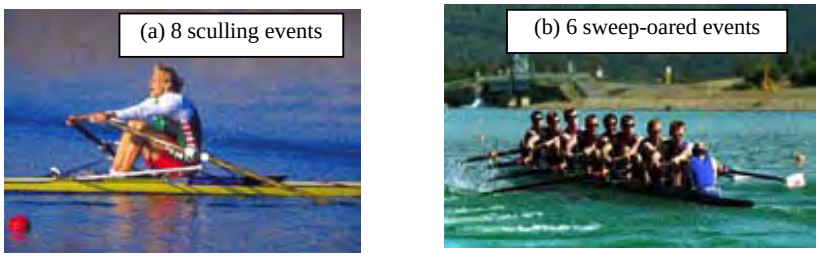


Fig. 4. Sculling (two oars used, one in each hand) and sweep-oared (one oar with both hands) scenarios in Olympic competition (Athens Olympics, 2004)

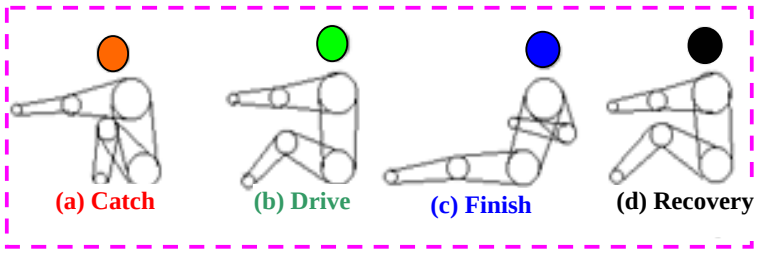


Fig. 5. Schematic diagrams showing the four main sequential phases (ie. a) catch, b) drive, c) finish and d) recovery) in a rowing stroke and the position of head, arm and legs of the rower)

both hands (see Fig. 4b). The sculling boat classes are the single, the double and the quadruple sculls with crews of one, two or four athletes respectively, as well as the lightweight double. The sweep row categories include the pair, the four, the lightweight four (for men only) and the eight with coxswain.

The Athens 2004 Olympic Games Rowing events were held at the Schinias Olympic Rowing and Canoeing Centre from 14 to 22 August 2004. A total of 550 athletes (358 men and 192 women) from all over the world took part in 14 rowing events. Forty-five Australian rowers (28 men and 17 women) took part in 11 rowing events.

3.1 Rowing stroke

A rowing stroke is a precise movement with rowers using their legs, back and arms to generate power. A stroke begins with the placing of the blade in the water and ends with the re-emergence of the blade from the water and positioning for another cycle. The rowing stroke can be divided into four main phases: catch, drive, finish and recovery (Mickelson and Hagerman, 1979) (see Fig. 5). These sequential phases must flow from and into each other to produce a continuous and fluid movement.

At the catch, the blade is placed into the water quickly with minimal disturbance to the boat. The rower's arms are extended outward, torso is tilted forward, and legs are compressed. A good catch produces a minimal amount of back and front splash and causes no check. The catches of all crews of a boat must be identical. Out of step catches (unsynchronisation) cause balance problems and reduce a boat's speed. The blade must be fully squared to the water at the catch (RowersWorld, 2003).

The boat gains its speed on the drive. In this portion of the stroke, the oarsman applies power to the oar with forces from arms, back and legs, and swings his torso away from the stern of the boat. The handle of the oar is pulled in a clean, powerful and levelled motion towards the bow of the boat with a constant force.

At the finish, the oarsman finishes applying power to the oar handle, removes the blade from the water sharply, and feathers the oar (rotate it by 90º) so that the blade becomes parallel to the surface of the water.

At the recovery, rowers are given a brief rest to prepare for the next stroke. The oarsman must slide towards the stern of the boat and prepare the blade for the next catch. Crews exhibit an approximate 2:1 ratio between the times spent on the recovery and the times spent on the drive. At the end of the recovery, the oar is gradually squared and prepared for the catch (*ibid*.).

Understanding which movements should occur in each phase of the stroke allows coaches to design effective conditioning programs and evaluate rowing performance effectively. Success in competitive rowing is achieved by taking the shortest time to complete a course (usually 2000m) which directly links to the average velocity of the boat. Acceleration is proportional to force since the boat is accelerated as it reacts with the sweeping arc of the oar. Three factors affecting boat velocity are: stroke power, stroke length and stroke rate. These factors are important determinants of rowing performance. The stroke power determines how fast the boat travels in a stroke, the length is associated with how far the boat travels in each stroke and the rate is the number of strokes rowed per minute (Seiler, 2003). Therefore the rower must achieve an optimal combination of high stroke power, long stroke length and high stroke rate.

Fig. 6 presents the stroke signals captured using geodetic-type GPS receivers. It is demonstrated that the signals captured provide a clear picture of the rowing stroke phases as described above. In this particular stroke, the graph indicates that the rower has harmonised well in his stroke cycle by using appropriate time (1:2) in the catch and the drive.

3.2 Indoor training using ergometer

Fig. 7 shows a typical athlete training procedure in an indoor environment using an ergometer (O'Sullivan and O'Sullivan, 2001). Much of the equipment is either too heavy, expensive, obtrusive or unreliable. Therefore, smart real-time monitoring during training and competition to help elite athletes to improve their performance and avoid injuries is critical for both athletes and coaches. Any methodology that would improve the situation would not only bring benefits to the rower practice, but also to many other sport-related applications including both team sports and individual athletes (eg. Zhang et al., 2003).

A major step forward would be achieved by a comprehensive system that could be used to obtain physiological information of the athlete together with precise movement information through an independent platform which is a low-cost, low-maintenance, miniaturised and integrated sensor system. GPS has been identified as a key element to the success of the project. Such a system is clearly ambitious and will not be achieved in a single step. This project will provide a starting point towards this ultimate goal, through the combination of advanced global navigation satellite systems technology, smart wireless communication, on-line signal processing and GIS, to measure movement information in real-time to a high precision.

4. Development of GPS Based Prototype System

In many sports, it is desirable for movements to be repeated multiple times to obtain a consistent positive result. A greater understanding of sensor-based human performance measurement, such as the determination of a characteristic signature of the "perfect" movement, is required to analyse the performance of the athlete (Seiler, 2003). Therefore, the ability to measure and record positional information together with athlete physiological information in real-time is critical to the process of athlete training and coaching. Physiological information can be relatively easy to obtain using relevant detectors as described above. However, real-time high precision positioning of the athlete has been a challenging task (eg. Fyfe et al., 2001; Hutchings et al., 2000; Larsson, 2003).

A prototype rower tracking system has been developed. Fig. 8 outlines the system architecture and Fig. 9 presents the major phases of the development. Online calculation and a user friendly visualisation mechanism is also integrated into the system.

4.1 Low-cost code GPS

A number of critical factors contribute to the applicability of the GPS system to rowing practice. This includes the precision, cost, volume and weight of the system, and its integration with other sensors including accelerometers, communication mechanism, and a personal digital assistant (PDA). In spite of the “high-end” geodetic-type receiver providing very high precision results, its disadvantages are the requirements of differential operation, large on-line data processing power and establishment of an independent base station. The cost, volume and weight of this type of GPS receiver and its sophisticated operational procedure unavoidably preclude its practical use.

Because of the challenges of diverse applications and a broadening market, the low-cost GPS has evolved significantly during the past decade (Xiao et al., 2003).

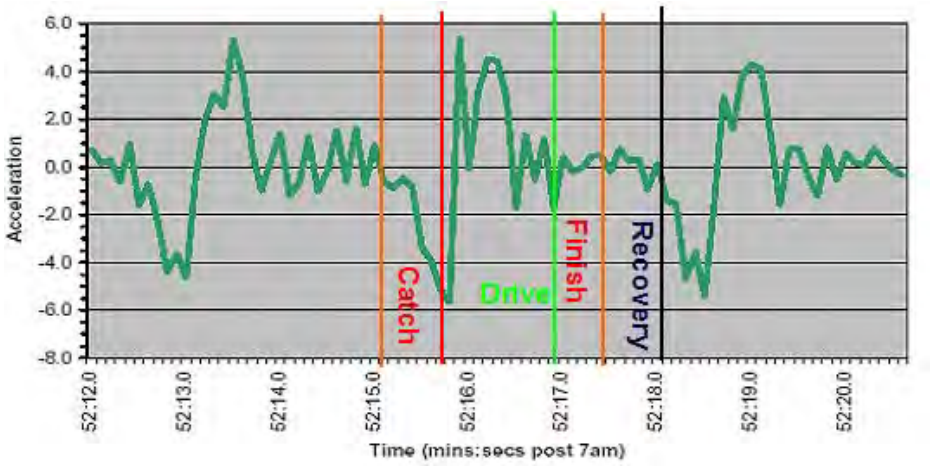


Fig. 6 Schematic rowing stroke signature captured from high precision GPS measurement (Trimble 5700, 10Hz)

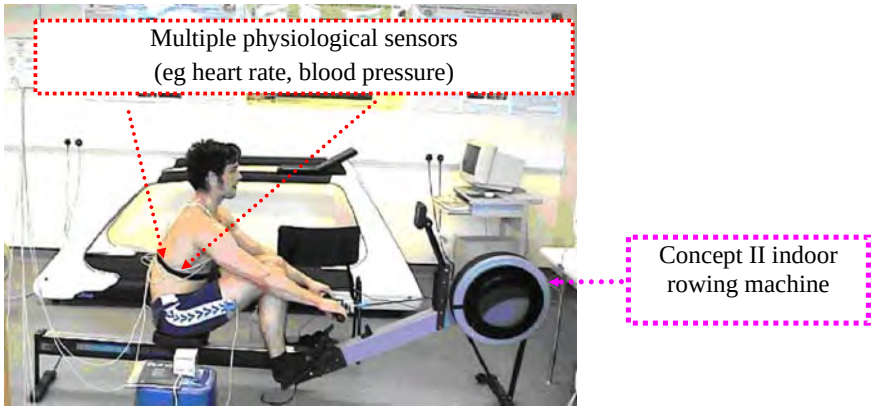


Fig. 7 Typical indoor training instrument (ie. ergometer) and athlete performance and information collection procedure using the concept II indoor device with various physiological sensors attached

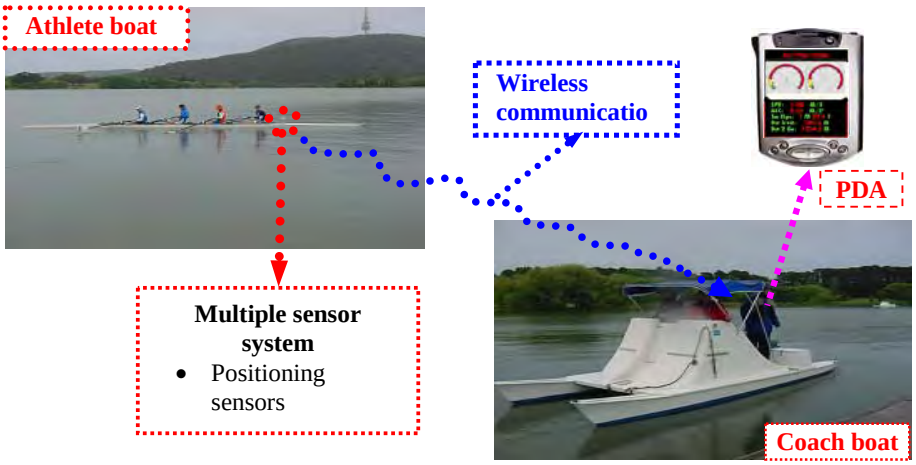


Fig. 8 Skeleton of the system structure of the "smart" rower tracking multi-sensor system with wireless communication and on-line processing functionalities.

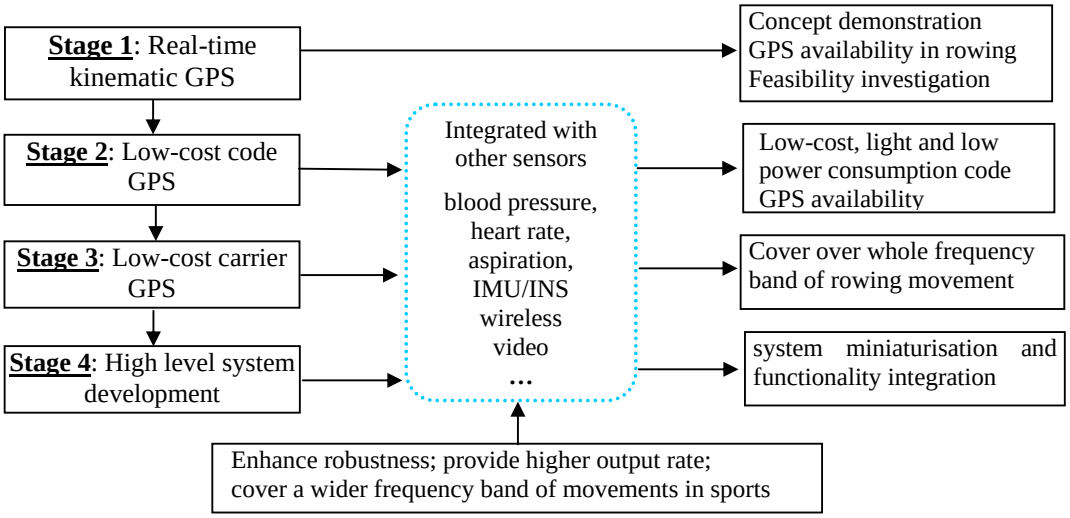


Fig. 9 Development roadmap of the prototype rowing tracking system

Increasingly miniaturised devices have been developed to be wearable and embedded in other devices. Significant new developments of the low-cost GPS units are listed below.

- The ability to process weak GPS signals,
- Combined functionality and integration with other systems such as cellphone or other wireless communication systems (eg. Bluetooth (2003), PDA, and Internet browser),
- Full compatibility incorporating GPS with other devices, and
- Reduction of power consumption, size and weight, and price.

Given these developments, the feasibility of the low-cost, code-only GPS receiver has been investigated and the performance of a low-cost receiver is presented in the next section. A code GPS receiver and a personal digital assistant (PDA) that form the first version of the prototype system are shown in Fig. 10.

4.2 Low-cost carrier GPS

A rower tracking system with the functionality to output PVT information at a rate of 10 Hz has been found necessary and essential in the rowing experiments. In order to develop such a system, a Canadian Marconi Company's (CMC) SuperStar II GPS OEM board is used to form the hardware basis of the system. The SuperStar II can provide PVT solution at a rate up to 5 Hz as well as raw measurements at a maximum rate of 10 Hz. The raw measurements include code phase, carrier phase and signal-to-noise ratio (SNR). Robust algorithms and associated software/firmware have been developed and the current prototype system configuration is shown in Fig. 11(a). The algorithms for PVA solution are developed using a number of special treatments for rowing-specific application. These algorithms have been evaluated through a number of static and kinematic trials as shown in section 5. More rowing trials are on the way and further refinement of algorithms is being undertaken.

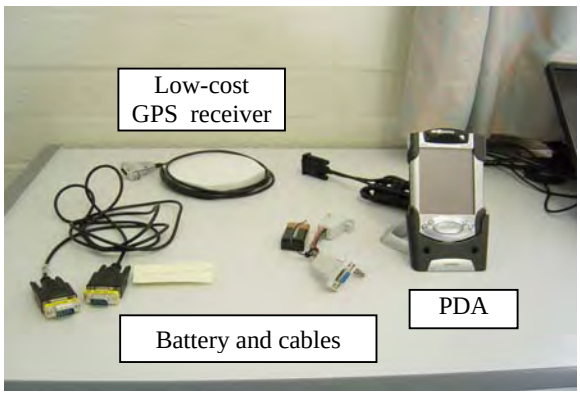


Fig. 10. Configuration of the low-cost code GPS (1Hz) system

5. Field experiments and results

A number of rowing trials have been conducted to assess the performance of both high-end and low-cost GPS systems and to identify potential problems in the river environment. RTK GPS has been proved to be able to provide high precision positioning in a lake environment.

However, there are a number of considerations: multipath effects, signal obstruction, satellite visibility, and obtrusion etc. Ideally the presence of any instrument should not cause direct visual or physical impact on the

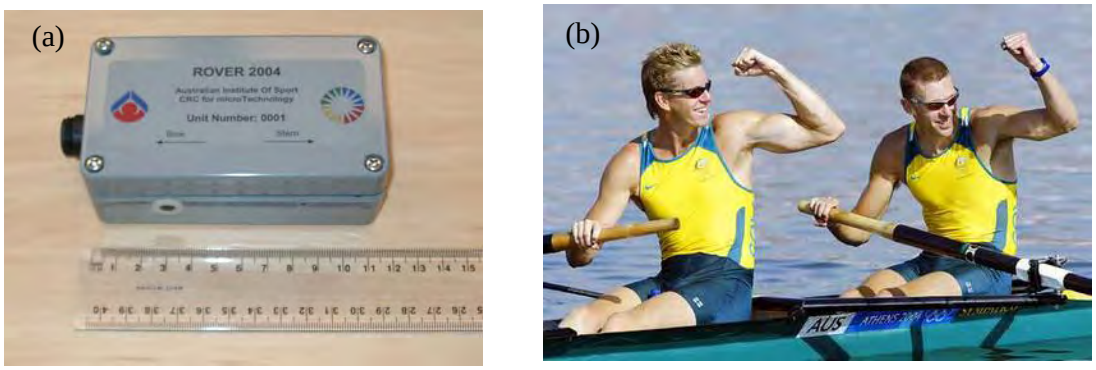


Fig. 11 A compact prototype rower tracking system (a) and photo of James Tomkins and Drew Ginn who won men's pair gold medal (b) in Athens Olympics

athlete. Therefore, the size and height of the antenna is a primary consideration.

The solution from geodetic-type GPS receivers acts as the reference while evaluating the solution of code-only GPS. The base station is located on the bank of a river about ~2km away from the course of the boat trial. The baseline solutions from each of the rowing antennas were processed independently from the base station using the post-processed kinematic technique.

If assuming that the accuracy of the position to one GPS rover is the same as to the other, then, from the simple (Least Squares Adjustment) error propagation law, the accuracy of the position of the kinematic GPS measurement (for a single baseline) can be estimated. A few millimetre accuracy of the antenna height was achieved in a three (consecutive) day trial (Zhang et al., 2003). Given the closeness of the antenna and the reflective nature to the water surface, the performance of the PPK GPS presents consistent results.

An important task of this project was to test the performance of a code GPS receiver which provides low-cost and light weight, less complicated operation, and less communication and storage requirements. Two procedures were used to test the performance of the low-cost GPS receiver: comparison with the high-end GPS receiver and zero motion test of the code GPS receiver (static performance) (Zhang et al., 2003).

Fig. 12 shows the mounting of the two types of GPS receiver (high-end carrier phase Trimble 5700 and low-cost code-only Rojone Genius 1 receivers). Note that the reflective nature of the water surface environment normally causes a high potential multipath effect which could contribute to a large amount of error (upto a few meters for the code GPS receiver). It was expected that elevating the antenna could potentially mitigate multipath effects. Contrary to expectations, the results indicate no significant multipath effects, which will provide guidance for antenna mounting in future trials.

Fig. 13 shows the boat trajectory of the trial for a ~1.5 hour "run". The continuity of the raw GPS measurements and solutions is evident. A close examination indicates that over 99.9% epochs have been resolved with a consistently high level of accuracy.

Fig. 12. RTK geodetic (carrier phase) and low-cost (code-only) GPS receivers mounted on the same boat

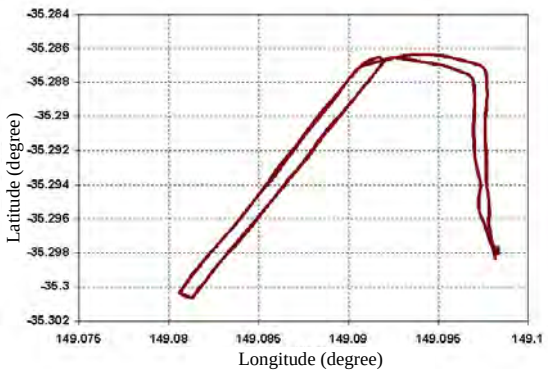


Fig. 13 Boat/rover trajectory obtained from code GPS

6. Conclusions

This paper presents two applications of the GNSS technology in sports: sailing and rowing. Determination of the sailing course at Shallow Inlet for the recent attempt on the World Sailing Speed Record and the feasibility of using GPS technology is described. It is shown that GPS is a superior technology to the current video camera procedure for speed recording. As a result of RMIT Satellite Positioning and Orientation Research Team (SPORT) findings, the Australian team has received World Speed Record Council approval to introduce the method of timing demonstrated in this paper for world record attempts.

This paper also outlines SPORT involvement in the recent development of the CRC for microTechnology Project 2.5. It is demonstrated that low-cost GPS receivers can provide high-accuracy velocity and acceleration information. It shows the feasibility of GPS technology to assist in elite rowing training. It has been demonstrated that the stroke signature captured from GPS is of great value for the investigation of, for example, the duration of the drive and recovery phases, the total time per stroke cycle, drive to recovery ratio, and the relationship of the hands and seat during the drive phase. A number of prototype rower tracking systems have been developed and used by Australian athletes prior to and during the Athens Olympics. Gold was struck twice as Australia came in first in the Men's Quad Skull. Australia snagged three rowing medals (gold, silver and bronze).

It has clearly been shown what can be achieved through "smart" integration of GPS and other advanced technologies for innovative sports applications.

Acknowledgements

The rower tracking project is funded under the CRC for microTechnology Project 2.5. The authors would like to thank Prof A. Hahn, Dr C. Gore and Dr A. Rice from the Department of Physiology, Australian Institute of Sports for their assistance, support and co-operation.

References

Athens Olympics (2004): ***Athens Olympic rowing***, http://www.athens2004.com/en/RowingRules.

Atkinson W. C. (2003): ***The ergometer and rowing compared.*** http://www.atkinsopht.com/row/erg/ergorowg.html.

Bluetooth (2003): ***The official bluetooth website.*** http://www.bluetooth.com/, accessed August.

CRC microTechnology (2003): ***Cooperative Research Centre for microTechnology.*** http://www.microtechnologycrc.com/CRC/crcmicro.nsf, accessed May 2003.

Elliott,B., Lyttle, A. and Burkett, O. (2003): ***The RowPerfect ergometer: A training aid for on-water single scull rowing***. Sports Biomechanics, 1(2).

Fyfe, K.R., Rooney, J.K. and Fyfe, K.W. (2001): ***Motion analysis system***. US Patent Office US6301964.

Hutchings, L.J., Gross, R. and Jarpe, S. (2000): ***System and method for measuring movement of objects***. United States Patent Office US6122960.

ISAF (2002): ***International Sailing Federation***. http://www.sailing.org/.

Kirtley, C. and Smith, R.A. (1996): ***Movement toolbox: a multimedia package for movement analysis.*** Proceedings of the First Australasian Biomechanics Conference, Sydney, Feb 1996. also available at http://www2.fhs.usyd.edu.au/ess/biomech/abc/proceed.html.

Kleshnev, V. (2001): ***Racing strategy in rowing during Sydney Olympic Games***. Australian Rowing 24(1):20-23.

Landmark (2002): ***A world record is in the wind – GPS helps time world speed-sailing attempts.*** Landmark, issue 14, December 2002.

Larsson, P. (2003): ***Global Positioning System and sport-specific testing***. Sports Med, 33(15): 1093-1101.

Lin, A., Mullins, R. and Pung, M. (2003): ***Applications of accelerometers in sports training.*** http://search1.analog .com/searchProxy.asp?queryText=sports&, accessed May.

Mickelson, T.C., and Hagerman, F.C. (1979): ***Functional anatomy of the rowing stroke.*** The Oarsman for May/June, 6-9.

O'Sullivan, F. and O'Sullivan, J. (2001): ***Modelling multivariate biomechanical measurements of spine motion during a rowing. exercise,*** http://homerweb. ucc.ie/.

RowersWorld (2003): ***About rowing.*** http://www.rowersworld.com/Content/About/2000_7_10n.php, accessed May 2003.

Seiler, S. (2003): ***Rowing: physiology and performance.*** http://home.hia.no/~stephens/rowing.htm, accessed May3.

Upson, D. E. (2003): ***A kinematic and temporal analysis of ergometer rowing and on-water rowing,*** MSc Thesis, Southern Connecticut State University, 36pp. also available at http://www.southernct.edu/departments/graduatestudies/fellows/upson_diane.pdf.

WSSRC (2002): ***World sailing speed record council.*** http://www.sailspeedrecords.com/ rulesinshore.html.

Xiao, B., Zhang, K., Grenfell, R. and Norton, T. (2003): ***Handheld GPS: Current status and further developments.*** Position Magazine, No.6, 58-59.

Zatsiorsky, V.M. and Yakunin, N. (1991): ***Mechanics and biomechanics of rowing: a review.*** International Journal of Sport Biomechanics, 7, 229-281.

Zhang, K., Li, Y. and Grenfell, R. (2003): ***Current development of a low-cost, high output rate, RTK GPS multisensor system for rowers***. Proc. of SatNav2003, 22– 25 July, Melbourne Australia, 14p.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 290-295

Journal of Global Positioning Systems
2004
© CPGPS

Convergence of Block Decorrelation Method for the Integer Ambiguity Fix

Samsung Lim* & Binh Quoc Tran**

*School of Surveying & Spatial Information Systems, UNSW, Sydney NSW 2052, Australia
Email: s.lim@unsw.edu.au Tel: +61-2-9385 4505 Fax: +61-2-9313 7493

**Department of Geography, Hanoi University of Sciences, VNU, 332 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam
Email: quocbinh@mail.ru

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Because of the integer-valued nature of carrier phase ambiguities, it is essential to fix the float estimates into integer values in order for high precision DGPS positioning. A decorrelation process is necessary to solve the problem since double-differenced ambiguities are highly correlated in general. In this paper, Block Decorrelation Method (BDM) is presented and tested for its convergence. BDM divides the variance-covariance matrix into four blocks and decorrelates them simultaneously. A number of randomly selected examples show that BDM is comparable to the existing decorrelation algorithm, however its speed of convergence is relatively faster due to the computations performed on small blocks.

Key words: Carrier Phase, Decorrelation, Double-Differenced, Integer Ambiguity, Variance-Covariance

1 Introduction

A float estimate for an initial ambiguity of GPS carrier phase measurements can be obtained by the ordinary least squares technique. However, it is essential to fix the integer value in order to achieve high precision positioning. The problem of integer ambiguities is equivalent to the minimization of the following objective function (Teunissen, 1998):

$$\min S(a) = (\hat{a} - a)^T Q_{\hat{a}}^{-1} (\hat{a} - a) \text{ with } a \in Z^n \qquad (1)$$

where $\hat{a}$ is a vector of n float values of double-differenced ambiguities, which is obtained by the least squares estimation with respect to the corresponding variance-covariance matrix $Q_{\hat{a}}$, a is a vector of n unknown integer values of ambiguities, and Z^n is the n-dimensional integer space.

Searching for the minimum for $S(a)$ is difficult because it involves the discrete parameter a. In practice, Equation (1) is usually solved by a discrete search strategy. The idea is that the search space Z^n can be replaced by a small subset or *ambiguity search space* bounded by hyper-ellipsoid:

$$(\hat{a} - a)^T Q_{\hat{a}}^{-1} (\hat{a} - a) \le \chi^2 \qquad (2)$$

where χ^2 is a suitably chosen constant which ensures that the ellipsoid contains at least one integer vector a. The variance-covariance matrix $Q_{\hat{a}}$ affects overall the geometry of elongation and rotation of the search ellipsoid and χ^2 determines its size. Consequently, $Q_{\hat{a}}$ directly affects the effectiveness of the search process. In an ideal case, $Q_{\hat{a}}$ is diagonal and hence a can be obtained simply by rounding the float solution $\hat{a}$ to nearest integer values.

Double-differenced ambiguities are highly correlated, and consequently, $Q_{\hat{a}}$ is far from diagonal. It means that the search ellipsoid is greatly elongated and its principal axes are misaligned with the grids axes (Teunissen 1998, De Jonge 1996). To speed up the search process, $Q_{\hat{a}}$ needs to be "decorrelated". This can be done by using the linear transformation $z = Z^T a$ (Teunissen 1998). Equation (1) is now equivalent to:

$$\begin{aligned} &\min S(z) = (\hat{z} - z)^T Q_{\hat{z}}^{-1} (\hat{z} - z) \\ &\text{with } \hat{z} = Z^T \hat{a},\ z = Z^T a,\ Q_{\hat{z}} = Z^T Q_{\hat{a}} Z \end{aligned} \qquad (3)$$

In order to preserve the nature of a in z, the transformation matrix Z must satisfy two conditions:

- C1: Elements of Z must be integers;
- C2: Elements of Z^{-1} must be integers too. It is equivalent to $\det[Z] = \pm 1$.

Any permutation matrix or triangular integer matrix with ± 1 in its diagonal meets both conditions C1 and C2. The transformation matrix Z^T will be chosen so that $Q_{\hat{z}}$ become near-diagonal and its condition number c become near 1. In general, a diagonal matrix $Q_{\hat{z}}$ with condition number c=1 is impossible because of the two conditions C1 and C2 above.

There may exist several methods of devising Z^T. A group of methods based on Gauss transformation algorithm (Strang 1997, Teunissen 1998, Xu 2000) is at hand. To decorrelate the element $(Q_{\hat{a}})_{ij}$, Z_k^T can be constructed as an identity matrix except one nonzero element at row i and column j:

$$\begin{aligned} &\left(Z_k^T\right)_{ij} = -\left[\left(Q_{z,k}\right)_{ij} / \left(Q_{zk}\right)_{jj}\right]^{\text{int}} \\ &Q_{z,k} = Z_{k-1}^T Q_{z,k-1} Z_{k-1}, \quad Q_{z,0} = Q_{\hat{a}}, \quad Z^T = Z_h^T Z_{h-1}^T \ldots . Z_1^T \end{aligned} \tag{4}$$

where the operator $[\cdot]^{\text{int}}$ denotes rounding to the nearest integer. The procedure consists of many steps until the last matrix Z_k^T becomes an identity matrix. The transformation matrix Z^T is a product of matrices Z_k^T, $(k = 1,\ldots,h)$. Although Gauss transformation algorithm usually decorrelates $Q_{\hat{z}}$ well, its convergence is slow because each element of $Q_{\hat{a}}$ should be decorrelated separately.

Another group of methods are based on the factorization of the original matrix $Q_{\hat{a}}$:

$$Q_{\hat{a}} = K^T D K \tag{5}$$

where D is a diagonal or a "near diagonal" matrix. Processing float-valued K^T properly, one can obtain integer-valued transformation matrix Z^T satisfying C1 and C2 so that

$$Q_{\hat{z}} = Z^T Q_{\hat{a}} Z \tag{6}$$

where $Q_{\hat{z}}$ is an almost diagonal matrix. The most popular method for factorizing $Q_{\hat{a}}$ is Cholesky's factorization (De Jonge 1996, Xu 2000):

$$Q_{\hat{a}} = LDL^T \quad \text{or} \quad Q_{\hat{a}} = U^T D U \tag{7}$$

where L and U are lower- and upper-triangular matrices with diagonal elements 1, respectively, and D is a diagonal matrix. Note that D is an ideal form of $Q_{\hat{a}}$.

The simplest way to construct Z^T is to round each element of L or U^T to nearest integers, that is,

$$Z^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ [L_{21}]^{\text{int}} & 1 & 0 & 0 \\ .... & .... & .... & ... \\ [L_{n1}]^{\text{int}} & ... & [L_{n,n-1}]^{\text{int}} & 1 \end{bmatrix}^{-1} \quad \text{or}$$

$$Z^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ [U_{21}]^{\text{int}} & 1 & 0 & 0 \\ .... & .... & .... & ... \\ [U_{n1}]^{\text{int}} & ... & [U_{n,n-1}]^{\text{int}} & 1 \end{bmatrix}^{-T} \tag{8}$$

For the factorization of $Q_{\hat{a}}$, one can also use Gram-Schmidt orthogonalization process (Grapharend 2000, Xu 2000):

$$Q_{\hat{a}} = V^T V = Z^{-T} O^T O Z^{-1} = Z^{-T} Q_{\hat{Z}} Z^{-1} \tag{9}$$

where O is an almost orthogonal matrix, and consequently, $Q_{\hat{z}}$ is almost diagonal.

Methods based on factorizations are usually faster than methods based on the Gauss transformation. However, due to the fact that they deal with the original matrix $Q_{\hat{a}}$ indirectly via K^T, their results may be relatively worse. Also, some of the methods still experience difficulty with convergence of iteration process (Xu, 2000).

In this paper, a new method for integer decorrelation of variance-covariance matrix $Q_{\hat{a}}$, which is faster than the method based on the Gauss transformation, will be presented. This method deals with the original matrix $Q_{\hat{a}}$ directly, but unlike the Gauss transformation, it divides $Q_{\hat{a}}$ into 4 small blocks and decorrelates elements in each block simultaneously. Therefore, the new method will be named as "Block Decorrelation Method" (BDM) hereafter.

2 Block Decorrelation Method

Consider the following matrix multiplication:

$$Z_k^T Q_{\hat{a}} Z_k = \begin{bmatrix} I^k & 0 & 0 \\ x_k^T & 1 & 0 \\ 0 & 0 & I^{n-k-1} \end{bmatrix}$$

$$* \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12}^T & q_{22} & q_{23} \\ q_{13}^T & q_{23}^T & q_{33} \end{bmatrix} * \begin{bmatrix} I^k & x_k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I^{n-k-1} \end{bmatrix}$$

$$= \begin{bmatrix} q_{11} & s_k & q_{13} \\ s_k^T & p_k & t_k^T \\ q_{13}^T & t_k & q_{33} \end{bmatrix} \quad (10)$$

$s_k = q_{11} x_k + q_{12}$,

$t_k = q_{13}^T x + q_{23}^T$,

$p_k = (x_k^T q_{11} + q_{12}^T) x_k + x_k^T q_{12} + q_{22} = s_k^T x_k + x_k^T q_{12} + q_{22}$

where I^m is an identity matrix of size *m*; the symmetric positive-definite matrix $Q_{\hat{a}}$ of size *n*n* is divided into 3 by 3 blocks because of the pre-multiplication by Z_k^T and post-multiplication by Z_k. Note that p_k is a scalar, but s_k, t_k are vectors. In Equation (10), it is clear that:

If the elements of x_k are integers, then Z_k^T is an admissible transformation matrix that satisfies C1 and C2.

The upper-left q_{11} and lower-right q_{33} are intact after the multiplications. The blocks q_{13} and q_{13}^T do not change too.

Since $Q_{\hat{a}}$ is symmetric positive-definite, q_{11} is invertible. Consequently the off-diagonal blocks s_k and s_k^T will be zero if x_k satisfies the following.

$$s_k = q_{11} \hat{x}_k + q_{12} = 0 \text{ or } \hat{x}_k = -q_{11}^{-1} q_{12} \quad (11)$$

Due to the condition C1, x_k cannot be a float solution of (11) and s_k, s_k^T cannot be set to zero. But one can expect that s_k, s_k^T will be "nearly zero" if x_k is rounded to the nearest integers:

$$x_k = \left[\hat{x}_k\right]^{\text{int}} \quad (12)$$

Assume that index *k* increases monotonically from *1* to *(m-1)* with a step size 1. By using Equations (10) to (12), the elements of s_k and s_k^T can become close to zero. After the recursive process up to *(m-1)* step, the (*m*m*) upper-left block of the last matrix in Equation (10) will have "decorrelated" elements all over the off-diagonal area.

If another form Z_h^T of transformation matrix is taken, Equations (10) to (12) will become Equation (13) to (15), respectively:

$$Z_h^T Q_{\hat{a}} Z_h = \begin{bmatrix} I^h & 0 & 0 \\ 0 & 1 & x_h^T \\ 0 & 0 & I^{n-h-1} \end{bmatrix}$$

$$* \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12}^T & q_{22} & q_{23} \\ q_{13}^T & q_{23}^T & q_{33} \end{bmatrix} * \begin{bmatrix} I^h & 0 & 0 \\ 0 & 1 & 0 \\ 0 & x_h & I^{n-h-1} \end{bmatrix}$$

$$= \begin{bmatrix} q_{11} & t_h & q_{13} \\ t_h^T & p_h & s_h^T \\ q_{13}^T & s_h & q_{33} \end{bmatrix} \quad (13)$$

$s_h = q_{33} x_h + q_{23}^T$, $t_h = q_{13} x_h + q_{12}$,

$p_h = (x_h^T q_{33} + q_{23}) x_h + x_h^T q_{23}^T + q_{22} = s_h^T x_h + x_h^T q_{23}^T + q_{22}$

$$s_h = q_{33} \hat{x}_h + q_{23}^T = 0 \text{ or } \hat{x}_h = -q_{33}^{-1} q_{23}^T \quad (14)$$

$$x_h = \left[\hat{x}_h\right]^{\text{int}} \quad (15)$$

Comparing with Equation (10), positions of *s* and *t* are exchanged and the role of q_{11} is now given to q_{33} in Equation (13). If the index *h* decreases from *(n-2)* to *m* with the step size -1, the block q_{33} will augment from (1*1) to (*n-m-1)*(n-m-1*) matrix in the recursive process of Equations (13), (14) and (15). In addition to that, there is another useful block matrix multiplication:

$$Z_g^T Q_{\hat{a}} Z_g = \begin{bmatrix} I^g & 0 \\ y_g^T & I^{n-g} \end{bmatrix} * \begin{bmatrix} q_{11} & q_{12} \\ q_{12}^T & q_{22} \end{bmatrix} * \begin{bmatrix} I^g & y_g \\ 0 & I^{n-g} \end{bmatrix}$$

$$= \begin{bmatrix} q_{11} & s_g \\ s_g^T & p_g \end{bmatrix} \quad (16)$$

$s_g = q_{11} y_g + q_{12}$,

$p_g = (y_g^T q_{11} + q_{12}^T) y_g + y_g^T q_{12} + q_{22} = s_g^T y_g + y_g^T q_{12} + q_{22}$

Again, the upper-left block q_{11} does not change after transformation. The off-diagonal blocks s_g, s_g^T will be zero if $\hat{y}_g$ is a solution of:

$$q_{11} \hat{y}_g + q_{12} = 0 \text{ or } \hat{y}_g = -q_{11}^{-1} q_{12} \quad (17)$$

Because of the conditions C1, C2 imposed to Z_g^T, it is only possible to set s_g, s_g^T to near zero by rounding the solution of (17) to nearest integers:

$$y_g = [\hat{y}]^{\text{int}} \tag{18}$$

Equations (10) to (18) provide an idea of decorrelating $Q_{\hat{a}}$. Dividing $Q_{\hat{a}}$ into four blocks of nearly equal size (see Figure 1), the upper-left block **A** can be decorrelated using Equations (10) to (12). Equations (16) to (18) can be used to make lower-left $\mathbf{B}^T$ and upper-right **B** near zero, and then the lower right block **C** will be decorrelated by Equations (13) to (15). Instead of full-size matrix $Q_{\hat{a}}$, this process deals with smaller (or half-size) blocks **A**, **B**, and **C.** Thus it will speed up the decorrelation process of $Q_{\hat{a}}$.

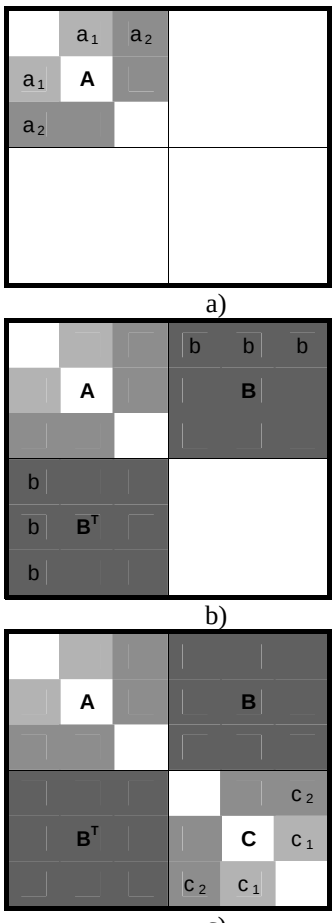


Figure 1. Graphical visualization of block decorrelation method for a 6*6 variance-covariance matrix: (a) - matrix Q_A; (b) - Q_{AB}; (c) - $Q_Z = Q_{ABC}$; a_1, a_2, c_1, c_2 show the order of substeps in the corresponding step; elements with gray color have near zero values.

Assume that there are *n* ambiguities in Equation (1), *m=n/2* if *n* is even, and *m=(n+1)/2* if *n* is odd number. BDM suggests decorrelating $Q_{\hat{a}}$ within a few numbers of iteration; each consists of the following steps:

Step 1: Permute matrix $Q_{\hat{a}}$ to obtain $\vec{Q}_{\hat{a}}$ so that its first *m* diagonal elements are minimal and stay in increasing order, i.e.,

$$[\vec{Q}_{\hat{a}}]_{11} \le [\vec{Q}_{\hat{a}}]_{22} \le \ldots \le [\vec{Q}_{\hat{a}}]_{mm} \le [\vec{Q}_{\hat{a}}]_{jj} \ (m < j \le n)$$

This step is necessary to achieve a better decorrelation of block **A** (Figure 1a). Assume that the minimal diagonal element of $Q_{\hat{a}}$ currently stays at row *r*. To make it the first diagonal element, one can use the permutation:

$$\vec{Q}_{\hat{a}} = H_{A1}^T Q_{\hat{a}} H_{A1} \tag{19}$$

where H_{A1}^T is a permutation matrix, obtained from the identity matrix by exchanging rows 1 and *r*. Repeating the procedure above for the second, third, … and m^{th}-diagonal elements, it yields to:

$$\begin{aligned} \vec{Q}_{\hat{a}} &= H_{Am}^T H_{A(m-1)}^T \ldots H_{A1}^T Q_{\hat{a}} H_{A1} \ldots H_{A(m-1)} \ldots H_{Am} \\ &= H_A^T Q_{\hat{a}} H_A \end{aligned} \tag{20}$$

Note that H_{Aj}^T and H_A^T satisfy conditions C1 and C2 and are admissible transformation matrices.

Step 2: Apply the decorrelation process to the upper-left block **A** of $\vec{Q}_{\hat{a}}$ using Equations (10), (11) and (12) with index *k* increasing from 1 to *(m-1)*. The upper-left diagonal block q_{11} in Equation (10) will be gradually augmented by s_k, s_k^T and p_k to fill up **A** as shown in Figure 1a. To speed up the process, the following formula can be used for inverting the augmented matrix q_{11}^k, based on the inverted matrix $[q_{11}^{k-1}]^{-1}$ in previous substep *k-1*:

$$\begin{aligned} &\left[q_{11}^k\right]^{-1} = \begin{bmatrix} q_{11}^{k-1} & s_{k-1} \\ s_{k-1}^T & p_{k-1} \end{bmatrix}^{-1} = \begin{bmatrix} G_{11} & G_{12} \\ G_{12}^T & G_{22} \end{bmatrix} \\ &R = s_{k-1}^T \left[q_{11}^{k-1}\right]^{-1} \\ &G_{22} = (p_{k-1} - R s_{k-1})^{-1}, \\ &G_{12} = -R^T G_{22}, \\ &G_{11} = \left[q_{11}^{k-1}\right]^{-1} - G_{12} R \end{aligned} \tag{21}$$

Thus the process of Step 2 is recursive. It produces a partially decorrelated matrix Q_A:

$$\begin{aligned} Q_A &= Z_A^T \vec{Q}_{\hat{a}} Z_A = Z_A^T H_A^T Q_{\hat{a}} H_A Z_A \\ &= [Z_{A(m-1)}^T \ldots Z_{A2}^T Z_{A1}^T H_A^T] Q_{\hat{a}} [H_A Z_{A1} Z_{A2} \ldots Z_{A(m-1)}] \end{aligned} \tag{22}$$

Step 3: Decorrelate blocks **B** and $\mathbf{B}^T$ of Q_A (Figure 1b) using Equations (16) to (18). The size of the block q_{11} in Equation (16) is *m*m*. To get its inverse, Equations (21) and $[q_{11}^{m-1}]^{-1}$ from Step 2 can be used. As a result, Q_{AB} is obtained by:

$$Q_{AB} = Z_B^T Q_A Z_B \tag{23}$$

Step 4: Permute Q_{AB} to yield $\overleftarrow{Q}_{AB}$ so that the diagonal elements of its lower-right block **C** (Figure 1c) stays in decreasing order, i.e.,

$$[\overleftarrow{Q}_{AB}]_{nn} \le [\overleftarrow{Q}_{AB}]_{(n-1)(n-1)} \le ... \le [\overleftarrow{Q}_{AB}]_{(m+1)(m+1)}.$$

Analogous to the Step 1, this procedure can be done by the matrix H_C^T:

$$\overleftarrow{Q}_{AB} = H_C^T Q_{AB} H_C \tag{24}$$

The upper-left block **A** remains unchanged in this step.

Step 5: Decorrelate the last block **C** in $\overleftarrow{Q}_{AB}$ using Equations (13) to (15) with index *h* decreasing from *(n-2)* to *m*. The lower-right block q_{33} in Equation (13) will be gradually augmented by s_h, s_h^T and p_h to fill up **C** as shown in Figure 1. The formula for inverting the augmented matrix is similar to Equation (21):

$$\left[q_{33}^{h}\right]^{-1} = \begin{bmatrix} p_{h-1} & s_{h-1}^T \\ s_{h-1} & q_{33}^{h-1} \end{bmatrix}^{-1} = \begin{bmatrix} G_{11} & G_{12} \\ G_{12}^T & G_{22} \end{bmatrix}$$

$$R = \left[q_{33}^{h-1}\right]^{-1} s_{h-1}$$

$$G_{11} = (p_{h-1} - s_{h-1}^T R)^{-1}, \tag{25}$$

$$G_{12} = -G_{11} R^T,$$

$$G_{22} = \left[q_{33}^{h-1}\right]^{-1} - R G_{12}$$

The first m elements of the vector q_{12} and the first m rows of q_{13} in Equation (13) are near zero because the block **B** is already decorrelated in Step 3. Thus the first m elements of the vector $t_h = q_{13} x_h + q_{12}$ in Equation (13) also become near zero. Hence, elements of $\mathbf{B}^T$ and **B** remain near-zero, even though they can change at Step 5. It provides a decorrelated matrix $Q_{\hat{z}i}$:

$$Q_{\hat{z}i} = Q_{ABC} = Z_C^T \overleftarrow{Q}_{AB} Z_C \tag{26}$$

Combining Equations (20), (22), (23), (24) and (26) yields:

$$\begin{aligned} Q_{\hat{z}i} &= Z_C^T H_C^T Z_B^T Z_A^T H_A^T Q_{\hat{a}} H_A Z_A Z_B H_C Z_C \\ &= [Z_{Cm}^T ... Z_{C(n-2)}^T H_C^T Z_B^T Z_{A(m-1)}^T ... Z_{A1}^T H_A^T] Q_{\hat{a}} \\ &\quad * [H_A Z_{A1} ... Z_{A(m-1)} Z_B H_C Z_{C(n-2)} ... Z_{Cm}] \end{aligned} \tag{27}$$

The transformation matrix Z_i^T in current i^{th}-iteration is obtained by Equation (27):

$$Z_i^T = Z_{Cm}^T ... Z_{C(n-2)}^T H_C^T Z_B^T Z_{A(m-1)}^T ... Z_{A1}^T H_A^T \tag{28}$$

Since Z_{Ch}^T, Z_B^T, Z_{Ak}^T $(h = n-2,...,m;\ k = 0,...,m-1)$ and H_A^T, H_C^T satisfy conditions C1 and C2, it is readily seen that Z_i^T satisfies these conditions too.

The procedure described in steps 1 to 5 can be repeated until Z_i^T becomes an identity matrix i.e., until no further decorrelation of $Q_{\hat{a}}$ can be obtained. The final transformation matrix Z^T is then computed by:

$$Z^T = Z_M^T Z_{M-1}^T ... Z_1^T = \prod_{i=M}^{1} Z_i^T \tag{29}$$

where *M* denotes the number of iteration. An estimation shows that each iteration without explicit calculation of Z_i^T requires about $7n^3/8$ multiplication. Therefore, the decorrelation process requires about $7Mn^3/8$ multiplication.

3 Numerical Example

In this section, BDM is used for decorrelating two sets of ambiguities. Ambiguities of these numerical examples are highly correlated. To quantify the decorrelation of $Q_{\hat{a}}$, two kinds of measures are used:

- The correlation coefficients;
- The condition number ***c*** which is the ratio of the largest and the smallest singular value of variance-covariance matrix.

If *e* denotes the elongation of the ellipsoid in Equation (2) then:

$$c = e^2 = R_{\max}^2 / R_{\min}^2 \tag{30}$$

where $R_{\max}$ and $R_{\min}$ are the largest and smallest axes of the ellipsoid, respectively.

Table 1 shows the main characteristics of the decorrelation process: the condition numbers ca and cz, the smallest $\rho_{\min}$ and the largest $\rho_{\max}$ correlation coefficients of original matrix $Q_{\hat{a}}$ and decorrelated

matrix $Q_{\hat{z}}$; the number of iteration circles M. To compare with the existing methods, Table 1 also shows the results obtained by United Decorrelation Method (Liu, 1999) and Gauss transformation.

NN	Description	$Q_{\hat{a}}$		$Q_{\hat{z}}$		# of iter.
		c	$\rho_{\min}$	c	$\rho_{\max}$	
1	6 ambiguities, BDM	$2.2*10^7$	0.8442	12.2	0.3990	6
2	6 ambiguities, United decorrelation method	$2.2*10^7$	0.8442	12.2	0.3990	6
3	6 ambiguities, Gauss transformation	$2.2*10^7$	0.8442	11.7	0.4275	6
4	12 ambiguities, BDM	$2.1*10^5$	0.9448	24.8	0.5056	9
5	12 ambiguities, Gauss transformation	$2.1*10^5$	0.9448	54.5	0.4778	10

Table 1. Parameters of the correlation processes: c, $\rho_{\min}$, $\rho_{\max}$ denotes condition numbers, minimal and maximal absolute values of correlation coefficients.

The test results show that highly correlated ambiguities are significantly decorrelated: the condition number and the corresponding elongation of the search ellipsoid drastically reduced from 10^5-10^7 to less than 100, and the average correlation coefficients diminished more than 2 times. In a relatively small number of ambiguities, all three methods give almost identical results, but for a larger number of ambiguities, BDM produces a better result. The MatLab implementation of the algorithm proves that BDM is 70-120% faster than Gauss transformation depending on the original matrix $Q_{\hat{a}}$ in the cases of 12 ambiguities.

4 Conclusions

The decorrelation process plays an important role in resolving integer ambiguities of GPS carrier phase measurements. In this paper, a new method for the decorrelation is presented. The method is based on dividing the variance-covariance matrix into 4 small blocks and decorrelating them simultaneously. The decorrelation of each block is processed recursively so that the result of the previous step is not affected by the next step. This algorithm reduces the dimension of the original variance-covariance matrix and therefore increases the speed of the decorrelation process. The proposed algorithm provides comparable or better result than that of the existing algorithm.

References

De Jonge, P.J. and Tiberius, C.C.J.M. (1996), ***The LAMBDA method for integer ambiguity estimation: implementation aspects,*** Delft Geodetic Computing Center LGR series, No. 12.

Grapharend, E.W. (2000), ***Mixed integer-real valued adjustment (IRA) problem: GPS initial cycle ambiguity resolution by means of the LLL algorithm,*** GPS Solution, No. 4, pp. 31-44.

Liu, L. T. et al. (1999), ***A new approach to GPS ambiguity decorrelation,*** Journal of Geodesy, No 73, pp. 478-490.

Strang, G. (1997), ***Linear algebra, geodesy and GPS,*** Wellesley-Cambridge Press.

Teusnissen, P.J.G. and Kleusberg, A. (1998), ***GPS for Geodesy,*** Springer.

Xu, P. (2000), ***Random simulation and GPS decorrelation,*** Journal of Geodesy, No 75, pp. 408-423.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 296-301

Journal of Global Positioning Systems
2004
© CPGPS

Development of SydNET Permanent Real-time GPS Network

C. Rizos, T.S. Yan

School of Surveying & Spatial Information Systems, University of New South Wales, Sydney, NSW 2052, Australia
e-mail: thomas.yan@unsw.edu.au Tel: + 61 2 9385 4189; Fax: +61 2 9313 7493

D.A. Kinlyside

Land and Property Information, Department of Lands, Bathurst, NSW 2795, Australia
e-mail: doug.kinlyside@lands.nsw.gov.au Tel: +61 2 6332 8372; Fax: +61 2 6332 8479

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. Over the past few years, there has been substantial growth in multiple-reference-station networks used to overcome the limitations of standard real-time kinematic (RTK) systems. SydNET is a project to establish a permanent real-time GPS network in the Sydney basin area providing Network-RTK support to users in the area. SydNET is being developed by NSW Department of Lands in partnership with the School of Surveying & SIS at the University of New South Wales. This paper presents recent developments of the SydNET network. Preliminary test results will be presented which will show the network's performance, achievable accuracy. It will outline the SydNET system, its operation, current status and vision of future development as a high precision positioning service infrastructure.

Key words: RTK, Network-RTK, GPS networks, CORS infrastructure

1 Introduction

1.1 Background

The standard mode of precise differential GPS positioning is for one reference receiver to be located at a base station whose 3D coordinates are known in a geocentric reference frame so that the second receiver's coordinates are determined relative to this reference receiver. This is the principle underlying pseudorange-based differential GPS (or DGPS for short) techniques. To achieve high accuracy, carrier phase data must be used but this comes at a cost of system complexity because the measurements are ambiguous. Therefore Ambiguity Resolution (AR) algorithms must be incorporated as an integral part of the data processing software. Such high accuracy techniques are the result of progressive R&D innovations which have been subsequently implemented by the GPS manufacturers in their top-of-the-line 'GPS Surveying' products e.g., Rizos (2002a).

Over the last decade several significant developments have resulted in this high accuracy performance also being available in real-time. That is, immediately following the making of measurements, and after the data from the reference receiver has been transmitted to the field receiver for processing via some data communication links (e.g., VHF or UHF radio, mobile phone, FM radio sub-carrier or satellite link), accurate positions are produced in the field. Real-time precise positioning is then possible when the GPS receiver is in motion. These systems are commonly referred to as real-time kinematic or RTK systems and make feasible the use of GPS for many time-critical applications such as engineering surveying, GPS-guided earthworks/excavations, machine control and other high precision navigation applications e.g., Lachapelle *et al.* (2002).

The limitation of single-base RTK is the distance between reference receiver and the user receiver due to distance-dependent biases occurring such as orbit error and ionospheric and tropospheric signal refraction. This has restricted the inter-receiver distance to 10km or less if very rapid Ambiguity Resolution (AR) is desired (ie less than a few seconds).

Wide Area Differential GPS (WADGPS) and the Wide Area Augmentation System (WAAS) on the other hand,

use a network of base stations separated by hundreds of kilometres over a wide geographic area. The measurement biases can be modelled and corrected at the users' receiver and therefore the positioning accuracy will be almost independent of the inter-receiver distance. However, these are predominately pseudorange-based systems intended to deliver accuracies at the metre to sub-metre level.

Continuously Operating Reference Stations (CORS) have been deployed to support very high accuracy geodetic applications since the 1980s (Evans *et al*., 2002). Geodetic techniques are by their very nature 'multi-station', taking advantage of the geometric strength, reference datum stability (and redundancy) afforded by network-based positioning. Such CORS networks have been deployed globally, as well as in geodynamic 'hot spots' like Japan and Southern California where there is significant tectonic motion (Ibid, 2002).

In Europe, as in many other countries, countrywide 'active control stations' have been established, consisting of CORS that collect data specifically for survey and mapping applications. In Australia, the state-wide Victorian CORS network, *GPSNet* serves the same purpose. Until recently however, such CORS networks have contributed to improving surveying productivity by obviating the need for GPS surveyors to operate a reference receiver. They have not been used in an optimal manner to address the distance constraint of single-base positioning (real-time or post-mission) in the same way that the WADGPS/WAAS techniques have done so far for pseudorange-based positioning.

1.2 Network-based positioning

Rapid static and kinematic GPS surveying techniques can become more productive by taking advantage of CORS networks in such network-based implementations as *Network-RTK* or more generally *Network-based Positioning* (Lachapelle *et al*., 2002).

CORS networks support multi-station (e.g. geodetic) processing of data from reference receivers simultaneously with the user receiver data. These days, this is typically achieved by a web-based service such as AUSPOS (http://www.auslig.gov.au/geodesy/sgc/wwwgps/), which requires the user to upload their data to the web-engine, subsequently returning the results of the processing to the user. AUSPOS is not a real-time service and currently only supports static positioning for occupations of several hours or more using dual-frequency user receivers. GPS surveyors use these online processing services to establish high-order geodetic control but the services are obviously unsuitable for high productivity engineering-type surveys. AUSPOS and similar services rely on data collected and archived by the International GPS Service (http://igscb.jpl.nasa.gov), hence inter-receiver distances are many hundreds (even thousands) of kilometres. The Australian Regional GPS Network (http://www.auslig.gov.au/geodesy/argn) is an example of a sparse network that augments the IGS global network and provides data to web-engines such as AUSPOS.

To address applications other than geodesy/geodynamics, many countries and states have established CORS networks that collect data for users to subsequently access and process themselves. This is an important distinction; as such networks only provide 'passive' services such as data downloads of RINEX-formatted measurement files. As with 'geodetic' CORS networks, user only needs to operate one GPS receiver. However, because the survey user must process data using software typically provided by the GPS manufacturer, and rapid GPS survey techniques are used (e.g. kinematic, rapid static, 'stop-and-go', etc. – Rizos, 2002a), the distance between the user receiver and the closest reference receiver must be less than the maximum recommended for GPS surveying applications. This is less than 10km for very rapid AR and typically 20-30km for rapid static techniques.

The Hong Kong GPS Network (Kwok, 2002) is an example of a CORS network with a density of base stations that a user is always within 10km of a reference receiver (and usually two, to permit checking). On the other hand, the state-wide GPSNet (http://www.land.vic.gov.au/GPSnet/) established in Victoria, is a typical example of a 'passive' CORS network, with base station spacing of between 50 and 100km. In order to upgrade such a CORS network to real-time operations would require the implementation of a Network-RTK system if no user were to be disadvantaged by being more than 10km from a base station.

Several European countries have upgraded their CORS networks to implement RTK. In some cases, such as in Denmark where the density of base stations is high, of the order of 10-20km station spacing, it is possible to use standard single-base RTK techniques (Leica, 2003, personal communication). In Germany, the Satellite Positioning Network (SAPOS) (Elsner, 1996) has been upgraded in recent years to offer a Network-RTK service across all German states. This is a model that is likely to be followed by other 'passive' CORS networks as they upgrade to real-time operations.

1.3 Network-RTK concept

Network-RTK is the logical outcome of the continuous search for a GPS positioning technique that challenges the current constraints of single-base RTK, namely the

need to be within 10km of the base station if the highest performance is to be achieved. It is a centimetre-accuracy, real-time, carrier phase-based positioning technique capable of operating over inter-receiver distances up to many tens of kilometres with equivalent performance to current single-base RTK systems. The most crucial characteristic of contemporary RTK techniques that must be preserved is very rapid time to ambiguity resolution (AR), measured in seconds. Hence the base stations must be deployed in a pattern dense enough to model distance-dependent errors to such accuracy that residual double-differenced carrier phase observable errors can be ignored in the context of such rapid AR (Rizos, 2002b).

Network-RTK requires a data processing 'engine' with the capability to resolve the integer ambiguities between the static reference receivers that make up the CORS network. The 'engine' must be capable of handling double-differenced data from receivers 50-100km apart, operate in real-time, instantaneously for all satellites at elevation cut-off angles down to a couple of degrees (even with high noise data that is vulnerable to a higher multi-path disturbance). The Network-RTK correction messages can then be generated.

The benefits of the Network-RTK messages (as opposed to standard RTK messages) are:

- Elimination of orbit bias and ionosphere delay.
- Reduction of troposphere delay, multi-path disturbance and observation noise.
- RTK can be extended to what might be considered 'medium-range' baselines (up 100km).
- Low-cost single-frequency receivers can be used for RTK and rapid static positioning.
- Very high accuracy applications using low-cost GPS receivers (e.g. deformation monitoring, geodetic control network, etc.) are possible.
- Improve the accuracy, reliability, integrity, productivity and capacity of GPS positioning.

In addition to the *data processing engine*, the Network-RTK system needs to have a *data management system* and a *data communication system*. It needs to manage corrections generated in real-time, the raw measurement data, multi-path template for each reference stations (for multi-path mitigation), ultra-rapid IGS orbits, etc. There are two aspects to the data communication system: (a) between the master control station (MCS - where the *data processing engine* and data archive are located) and the various reference stations, and (b) communication between the MCS and users. From the Network-RTK implementation point of view, there are three possible architectures (Rizos, 2002b): (1) generation of the Virtual Reference Station and its corrections, (2) generation and broadcast of an Area Correction Model, or (3) broadcast the raw data from all of the reference stations. These are briefly described below.

2.1.1 Virtual Reference Station (VRS)

- At the MCS server, the VRS can be generated and the RTCM 20/21 message created and transmitted once the server knows the position of the roving user. There is no further request from the roving user if the rover supports RTCM 20/21 format, except that the user needs to send their location to the server.
- Two-way communication is required, with the user informing the server where they are, and the server continuously sending data to the user for RTK applications.
- There are some limitations on the number of simultaneous users accessing the VRS service due to server capacity.
- This configuration has been used by Trimble/Terrasat in their commercial product, *The Trimble Virtual Reference Station* (Vollath *et al*., 2000).

2.1.2 Area correction model broadcasting

- At the MCS server, the corrections, e.g. dispersive and non-dispersive atmospheric correction terms or carrier phase measurement residuals for each satellite at each reference stations, will be generated using data from the CORS network.
- The corrections can be used to generate an interpolation model or the VRS at the user end. The correction generation algorithms can be different.
- One-way communication is sufficient and there is no limit on the number of users. This requires a new data format, and the volume of transmitted data is more than in the case of a single reference station.
- This configuration has been proposed as a Network-RTK RTCM format by Leica and Geo++, and will be implemented in RTCM version 3 (Han, 2003, *personal communication*).

2.1.3 Raw data broadcasting

- Broadcast raw measurements (CMR or RTCM 18/19 message format) from either the MCS server or from the multiple reference stations individually.
- Generate the VRS, or corrections, at the user site. The computation load is therefore shifted to the user.
- This requires a new data format.
- One-way communication is sufficient.
- There is no limit on the number of users.

A discussion of the pros and cons of each type of implementation is beyond the scope of this paper. Tests will need to be conducted to determine which of these is best suited for the type of applications that will be addressed by the network service.

1.4 The Singapore Integrated Multiple Reference Station Network (SIMRSN)

Due to the complexity and cost (typically between $30-$50,000 per station) involved in establishing CORS networks, the data links and the data processing/management servers at the control centre, there have been comparatively few university-based Network-RTK systems established to support research. During the last few years, to the best of the authors' knowledge, only the Singapore Integrated Multiple Reference Station Network (SIMRSN) has been operating *both* as a research facility and an operational Network-RTK service for the benefit of GPS surveyors. The SIMRSN is a joint research and development initiative between the Surveying and Mapping Laboratory, Nanyang Technological University (NTU), Singapore; the School of Surveying & Spatial Information Systems at the University of New South Wales (UNSW); and the Singapore Land Authority (currently the main user) (Chen *et al*., 2000).

The SIMRSN consists of five CORS, connected by dedicated ISDN data lines to the control centre at NTU. It is a high quality and multi-functional network designed to serve the various needs of real-time precise positioning, such as surveying, civil engineering, precise navigation, road pricing, etc. (Chen et al., 2000). The SIMRSN also services off-line users, who can access archived RINEX data files via the Internet. The inter-station distances are of the order of several tens of kilometres at most. However, tests conducted in 2001 have shown that even a network with such comparatively short baselines had difficulty in modelling the disturbed ionosphere in equatorial regions during the last solar maximum period of the 11 year sunspot cycle (Hu et al., 2002). Unique facilities such as SIMRSN can therefore act as a test bed for network-based positioning techniques. The SIMRSN model of a network that is both a research facility and an operational network service for users is being adopted for the SydNET network being established in the Sydney metropolitan area.

2 SydNET CORS Network

2.1 Introduction

The SydNET real-time CORS network is being established with network-based positioning capability from the very start, including Network-RTK. The project is funded by the NSW Department of Lands (Lands) as an initiative of State Government infrastructure. Lands has been active in using GPS for a variety of surveying and mapping applications for over a decade (Kinlyside, 1999, 1995, 1993). The development of a Network-RTK system for the state's largest capital city is a natural and logical extension of the organisation's previous and current involvement in GPS applications for surveying and geodesy.

SydNET is also an important research facility for the Cooperative Research Centre for Spatial Information (http://www.spatialinfocrc.org/programs.html). It will be available for testing various network-based positioning techniques, both commercial products and those developed by research organisations. Additionally, SydNET can also be utilised for experiments on non-positioning applications such as 'GPS meteorology'.

2.2 Project Management and Structure

As with most other current projects within Lands, SydNET Project is managed using the PRINCE2 project management methodology. Major authorisations for the project are made through a project board which currently comprises two executives from Lands, one external member from UNSW and one external member from the Roads & Traffic Authority (RTA).

The School of Surveying & SIS at the University of New South Wales is the main IP supplier and development contractor. The first phase of SydNET is only servicing the Sydney basin region - an area of approximately 100x100km - but it is planned for expansion over time to cover other areas in NSW. In this initial phase, SydNET is implemented using the SIMRSN Network-RTK algorithms which support VRS-style Network-RTK and provide an online service for RINEX data download.

2.3 Locations and Coverage

In 2003, Lands approached RailCorp with the idea of utilising their fibre-optic network in order to provide the communication links between receivers at the reference stations and servers at the Network Control Centre (NCC). RailCorp is a state-owned corporation which provides passenger rail network throughout NSW. RailCorp has fibre-optic network installed extensively throughout their electrified railway network.

As many as eighteen sites on RailCorp's network were inspected for feasibility of installing a GPS reference station with the intention of providing coverage over Sydney basin area based on 15 to 20km spacing. Seven sites are subsequently chosen because they provide the best possible sky view, least radio interference, safety and suitable infrastructure for installing a permanent GPS antenna and receiver. These sites are located in Chippendale, Villawood, Waterfall, Mulgrave, Springwood, Cowan and Menangle. At the time of writing, six stations have been installed at these locations. Menangle is still to be constructed.

Figure 1 Location of SydNET sites and their coverage

2.4 Reference station hardware

Each reference station consists of a dual-frequency geodetic grade GPS receiver with choke-rings antenna and a device converting serial data streams into TCP/IP packets over Ethernet. SydNET is using existing GPS receivers belonging to Lands and UNSW, with the intention of reducing the amount of initial capital investment required. Currently, three Trimble 4000SSe receivers from Lands are being deployed in SydNET stations with a number of Leica System 500 receivers from UNSW available for testing and ad-hoc stations. Four new Ashtech µZ-CGRS receivers have also been procured and deployed to augment the network. The receivers will be upgraded as new GPS signals and the Galileo GNSS signals become available to users.

2.5 SydNET communication link

Data from the reference stations is transmitted via RailCorp's network back to the NCC in TCP/IP packets over Ethernet. There is a 10 Mbps physical link to RailCorp's network at each station. This is more than sufficient for the current data stream which is in the range of 20 to 30 kbps. Currently there is one main data stream of raw GPS measurements at a rate of 1Hz. With significant remaining bandwidth available, it is possible to add more data streams or equipments for testing and research purposes. In some of the stations, multi-ports serial converters are being installed to enable different streams of data for operational, research and signalling purposes.

2.6. SydNET Network Control Centre

The Network Control Centre (NCC) for SydNET is hosted by the Australian Centre for Advanced Computing and Communications (ac3) located at the Australian Technology Park, Redfern NSW. Ac3 provides a professionally managed, premium facility that was purpose built for high availability, is highly secure and highly connected to the Internet backbones, including a connection to RailCorp's fibre-optic network.

Data streams from SydNET reference stations are aggregated at ac3. A media converter was installed by RailCorp in Lands equipments cabinet which connects RailCorp's fibre-optic network to Lands' network. Measurements data from all SydNET stations are processed by SydNET servers to provide real-time correction and also post-processing data.

Traditionally, correction data is transmitted to users via radio on a specifically allocated frequency. This requires specialised hardware both on the service provider's side and on the users' side and also specific radio spectrum license. A more advanced technique is to use GSM mobile network but this still requires service provider to invest in expensive modem banks and routers. A dial-in system such as GSM has limited number of lines and hence users at the same time. Another disadvantage is the cost associated with using mobile network.

2.7 Real-time data distribution

Real-time correction data from SydNET is distributed via the Internet. The choice is then given to users to arrange access to the data on the Internet. A suggested method is

by using General Packet Radio System (GPRS) service which is available in most part of Sydney. With a suitable GPRS-enabled device, users can access the data and connect it to their GPS receiver. For testing and internal use, Lands and UNSW use either a Compaq iPAQ PDA with a GPRS jacket or an O2 XDA with built-in GPRS phone. Both devices run Pocket PC operating system. Client software has been developed to access SydNET server and retrieve correction data. This data is then transmitted on the device's serial port which is connected to the GPS receiver in the field.

Other alternatives to GPRS network are CDMA1x and 3G which are both available in Sydney as well. Similar to GPRS, users can get either an add-on CDMA 1X card or a PDA with built-in CDMA 1X capability such as the i-mate PDA2K.

3 Conclusions

The proliferation of CORS networks at all scales, global, national, state and local, will be a challenge to organisations that seek the implementation of common standards of service, and those that wish to see a seamless network-based positioning capability across all networks. The integration of networks with different operators, and different functionalities, is an added challenge. In Australia, there is the opportunity to address these challenges at the national and state level through such initiatives as the 'network research' to be undertaken by the CRC-SI (http://www.spatialinfocrc.org).

The SydNET network may be considered a 'second generation' CORS network, as it will be established with network-based positioning capability from the very start. The physical infrastructure, the communication links and the database are all controlled by one agency, the NSW Department of Lands. By providing such a framework, new reference receivers will be able to be connected to this 'backbone' on an *ad hoc* basis. The SydNET CORS network is the first step in ultimately replacing NSW's primary geodetic network of trig stations with an extensive network of 'active control stations'.

References

Chen X.; Han S.; Rizos C.; Goh P.C. (2000): ***Improving real-time positioning efficiency using the Singapore Integrated Multiple Reference Station Network (SIMRSN)***, 13th Int. Tech. Meeting of the Satellite Div. of the U.S. Institute of Navigation, Salt Lake City, Utah, 19-22 September, 9-18.

Elsner C. (1996): ***Real-time differential GPS: The German approach***, GIM Int. Journal for Geomatics, 10(5), 76-79.

Evans A.G.; Swift E.R.; Cunningham J.P.; Hill R.W.; Blewitt G.; Yunck T.P.; Lichten SM, Hatch RR, Malys S, Bossler J (2002): ***The Global Positioning System Geodesy Odyssey***, Navigation, 49(1), 7-34.

Hu G.; Khoo H.S.; Goh P.C.; Law C.L. (2002): ***Performance of Singapore Integrated Multiple Reference Station Network (SIMRSN) for RTK positioning***, GPS Solutions, 6(1-2), 65-71.

Kwok S. (2002): ***The Hong Kong GPS network and reference stations***, Journal of Geopsatial Engineering, 2(2), 57-65.

Kinlyside D.; Jones G. (1999): ***DGPS Applications in the Land Information Centre***, Procs 4th Int. Symp. on Satellite Navigation Technology & Applications, Brisbane, QLD, 20-23 July 1999.

Kinlyside D.; Bannister M.; (1995): ***Using OmniSTAR and GPS in Outback NSW***, Procs Satellite Navigation Technology Conference, UNSW, Brisbane, QLD, 26-28 June 1995.

Kinlyside D.; Dickson G.; (1993): ***DGPS Activities of the Land Information Centre***, Procs Satellite Navigation Technology Conference, UNSW, Kensington, NSW, 19-21 July 1993.

Lachapelle G.; Ryan S.; Rizos C. (2002): ***Servicing the GPS user***, Chapter 14 in Manual of Geospatial Science and Technology, J. Bossler, J. Jenson, R. McMaster & C. Rizos (eds.), Taylor & Francis Inc., 201-215.

Rizos C. (2002a): ***Making sense of the GPS techniques***, Chapter 11 in Manual of Geospatial Science and Technology, J. Bossler, J. Jenson, R. McMaster & C. Rizos (eds.), Taylor & Francis Inc., 146-161.

Rizos C. (2002b): ***Network RTK research and implementation: A geodetic perspective***, Journal of Global Positioning Systems, 1(2), 144-150.

Vollath U.; Buecherl A.; Landau H.; Pagels C.; Wagner B. (2000): ***Multi-base RTK positioning using Virtual Reference Stations***, 13th Tech. Meeting of the Satellite Div. of the U.S. Institute of Navigation, Salt Lake City, Utah, 19-22 September 2000, 123-131.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 302-307

Journal of Global Positioning Systems
2004
© CPGPS

Kinematic GPS Precise Point Positioning for Sea Level Monitoring with GPS Buoy

Wu Chen, Congwei Hu, Zhihua Li, Yongqi Chen, Xiaoli Ding, Shan Gao and Shengyue Ji

Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong
e-mail: lswuchen@polyu.edu.hk Tel: + 852 27665969; Fax: 852 23302994

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. In this paper, the basic precise point positioning model has been reviewed. A recursive least square algorithm that separates the position coordinates and other parameters, such as ambiguities and tropospheric delays, is proposed for kinematic PPP applications. A test was carried out to test the method proposed in this paper, which made use of a GPS buoy equipped with a pressure and a tilt meter to monitor the sea level in Hong Kong. The initial results from kinematic PPP positioning compared with conventional kinematic positioning methods shows the accuracy of decimetre level positioning accuracy can be achieved by the PPP method.

Key words: kinematic GPS; precise point positioning; sea level; GPS buoy

1 Introduction

Buoys equipped with GPS receivers have been used to measure water levels, atmospheric parameter and other physical conditions in sea, river or lake for the purposes of navigation, tide correction, the altimeter range calibration, ocean environment and pollution monitoring, flood control, and fisheries (Rocken et al. 1990, Key et al. 1998, Moore et al. 2000). For high positioning accuracy applications, the relative positioning RTK technique is widely adopted in most of GPS buoy positioning. With the improvement of IGS orbit and clock error products, the precise point positioning (PPP) using IGS satellite orbit and clock products has been attracted more and more attention in resent years. It has been demonstrated that centimeter level positioning precision can be achieved for static points on land with the PPP techniques, which is comparable in quality with the conventional differential mode while the computational efficiency has been improved greatly for large scale GPS network (Zumberge et al., 1997; Huang et al., 2002). Decimeter level Kinematic PPP has also been developed in precise orbit determination or meteorological study in low earth orbit (LEO) spacecrafts such as TOPEX/POSEIDON and CHAMP mission (Bisnath et al. 2001; Bock et al., 2002). The main algorithms and correction models for the PPP have been discussed by Kouba et al (2001) and the most widely used data type is un-difference ionosphere-free combination with phase and code measurements. An alternative data type used by some studies is code-phase ionosphere-free combination that aims at accelerating convergence speed (Gao et al., 2001). In this paper, we will mainly concentrate on the kinematic PPP technique for GPS buoy to measure the sea level, tide and tropospheric parameter. The different combination of data types for the PPP, namely un-difference (UD), satellite difference(SD), time difference(TD) and time-satellite difference(TSD), are presented in part 2 of this paper. A recursive least square algorithm for kinematic precise point positioning is presented in part 3. A test was carried out to test the method proposed in this paper, which made use of a GPS buoy equipped with a pressure and a tilt meter to monitor the sea level in Hong Kong. The kinematic PPP positioning solution is promising by comparing with conventional kinematic positioning methods. An initial field work is performed on Hong Kong island and data processing results are presented in part 4 while the conclusion and suggestions are given in the part 5.

2 Basic PPP models

The most widely used observation model for the Precise Point Positioning (PPP) is ionosphere-free un-difference combination (UD) for dual frequency phase or code data. To meet different positioning requirements, some simpler

combinations from UD can be formed. For simplicity, only carrier phase data are used and the correlation matrix form will not be presented here. For centimetre precision positioning, some related error corrections have to be applied, such as the relativity, satellite phase centre offset, satellite wind up, earth body tide, ocean load correction etc. (Kouba,2000) which will not be discussed here.

The UD data type is formed by combination of dual frequency ionosphere-free GPS data (Holfman-Wellenhof el al 2002, Kouba J et al 2001,Han et al 2001). It takes the form of

$$UD^k(t) = \rho^k(t) + c(dt_r(t) - dt^k(t)) + N^k + Tr^k(t) + \varepsilon \quad (1)$$

where

UD - ionosphere-free carrier phase observation (m),

ρ - geocentric topocentric distance for satellite,

N - real valued ionosphere-free carrier phase ambiguity,

$dt_r(t)$ - clock error of receiver r,

$dt^k(t)$ - clock error of satellite k,

$Tr^k(t)$ - tropospheric delay,

t - epoch time,

c - speed of light,

r - subscript for receiver station identifier,

k - superscript for satellite identifier,

ε - measurement noise.

The tropospheric delay can be expressed as a function of zenith path delay and a mapping function

$$UD^k(t) = \rho^k(t) + c(dt_r(t) - dt^k(t)) + N^k + m^k(z)T_r(t) + \varepsilon \quad (2)$$

where $m^k(z)$ is the mapping function, e.g. Neil mapping function (Niell, 1996), z is the zenith of the satellite and $T_r(t)$ is the zenith tropospheric path delay.

Therefore, parameters to be estimated with the UD model for the PPP positioning are station position coordinates (X,Y,Z), receiver clock $dt_r(t)$, ambiguity N and tropospheric zenith path delay $T_r(t)$. Whether in static or kinematic positioning, the receiver clock error is an epoch by epoch unknown parameter. The ambiguity is a real valued constant over time. As the troposphere changes very slowly over time, the tropospheric zenith path delay can be treated as a constant parameter within short period time, such as 1 hour interval or 15-minute interval. On the other side, because the receiver clock parameters and the ambiguity parameters are correlated, to estimate the absolute epoch by epoch receiver clock errors, code data has to be added to overcome the deficit of the normal equation matrix with a small weight. Otherwise a reference satellite has to be assigned in data processing, and in this way, only a relative receiver clock error can be obtained.

Based on the UD model, some combinations can be formed as to eliminate some of common parameters in an epoch or between epochs. Namely the satellite difference (SD), time difference (TD) and time and satellite difference (TSD) can be formed between satellite k and l or between epochs i and j or both. The definitions of these combinations can be as:

$$SD^{kl}(t) = UD^l(t) - UD^k(t) \quad (3)$$

$$TD^k(t) = UD^k(t_j) - UD^k(t_i) \quad (4)$$

$$TSD^{kl}(t) = SD^{kl}(t_j) - SD^{kl}(t_i) \quad (5)$$

From Eq. (3), the epoch by epoch receiver clock parameters are canceled out in the SD model. Only position, ambiguity and tropospheric parameters are remained. Compared with the UD model, the total parameters to be estimated reduced dramatically. While in the TD model (Eq. (4)), there are no ambiguity parameters left as the ambiguity is a constant over time as long as there are no cycle slips or loss of lock. The receiver clock parameters are transformed to clock difference between subsequent epochs. It can also be expressed as the clock difference relative to first epoch. Even though the ambiguities are removed in the TD model, the parameters to be estimated are far more than those in the SD model because of the epoch by epoch receiver clock difference parameters. In the TSD model (Eq. (5)), both receiver and ambiguity parameters are eliminated, only position components and tropospheric zenith delay parameters remain in the observation model. However, the TD and TSD models only provide distance change information (similar to Doppler measurements) and it may need longer periods for the position solution to be convergent.

3 A recursive least square algorithm for kinematic PPP

In our approach, we will use the UD and SD data models for the kinematic PPP data processing. The Kalman filtering method has been widely used for kinematic GPS data processing. To apply the Kalman filtering algorithms, it required to establish a dynamic model to describe the relationship of the estimated parameters in different epochs and the constant velocity or acceleration models are commonly used. However, with some

applications, such as GPS buoys, the above models cannot be used to describe the motion of the GPS receiver. In this paper a recursive least square method is proposed to separate the ambiguity and position parameters. Assuming that there are two kinds of parameters to be estimated as X and Y, the observation equation can be written as

$$AX + BY = L + V, P \tag{6}$$

where X is a vector for coordinate, receiver clock parameters which need to be estimated at each epoch. Y is a vector for ambiguity, zenith tropospheric parameter which does change rapidly between epochs. P is the weight matrix. The algorithm used in this paper tries to simplify the observation equation by eliminating parameter X in the observation equation (Eq. (6)). The normal equation for Eq. (6) can be written as:

$$\begin{bmatrix} A^T PA & A^T PB \\ B^T PA & B^T PB \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} = \begin{bmatrix} A^T PL \\ B^T PL \end{bmatrix} \tag{7}$$

$$\text{with } \begin{bmatrix} I & 0 \\ -Z & I \end{bmatrix}, \text{ and } Z = N_{21}N_{11}^{-1} \tag{8}$$

$$\begin{bmatrix} I & 0 \\ -Z & I \end{bmatrix}\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ 0 & \overline{N}_{22} \end{bmatrix} = \begin{bmatrix} A^T PL \\ \overline{B}^T PL \end{bmatrix}$$

$$\overline{N}_{22} = B^T PB - B^T PAN_{11}^{-1}A^T PB \tag{9}$$

$$\text{Assuming } J = AN_{11}^{-1}A^T P \tag{10}$$

we have

$$\overline{N}_{22} = B^T (I - J)^T P(I - J)B \tag{11}$$

$$\text{with } \overline{B} = (I - J)B \tag{12}$$

Then we have new normal equation

$$\overline{B}^T P\overline{B}Y = \overline{B}^T PL$$

This is equivalent to that we have a new observation equation

$$\overline{B}Y = L + U, P \tag{13}$$

In Eq. (13) only the parameter Y, such as ambiguity and tropospheric parameters are included while coordinate parameter X is eliminated. On the other hand, the observable L and its weight matrix keep unchanged. Therefore, in our algorithms, we will estimate the parameter Y first, together with a simple dynamic model to describe the ambiguity and tropospheric delays, by applying a Kalman filter. After parameter Y is determined, the parameter X (i.e. the station coordinates and receiver clock errors) can be estimated using the following equation.

$$\overline{X} = N_{11}^{-1}(A^T PL - N_{12}Y) \tag{14}$$

4 Data processing results

In this paper, two sets of data are used to demonstrate the positioning results using the algorithm proposed in the section 3. The first data set is static GPS data observed in Hong Kong. The data were observed on March 4, 2001 in Hong Kong GPS active network, totally 24 hours with 5 seconds observation interval. The data is processing in both state mode (to estimate one solution with all data available) and the kinematic mode (to estimate position every epoch). Figure 1 position errors in 3D with the static mode. It can be seen that it requires about 2 hours for the coordinate errors to reduce to less than 0.1 m. With 24 hour data, the coordinate difference compared with the known coordinate are 0.008m, -0.028m and -0.001m in north, east and height components. In the kinematic mode, we solved for the ambiguities and tropospheric delay parameters first. Then the position coordinates are solved every epoch in a 5 second interval using Eq. (14). In data processing, the IGS precise orbit and 5 min satellite clock error data are used. As in the kinematic mode 5 second interval data is used, we use a linear interpolation of 5 minute satellite clock error to get the satellite clock errors at the epoch. Figure 2(a) shows the position errors with 5 second interval. The position RMS errors with 5 second interval data are 0.037m, 0.043 and 0.104m in North, East and Height component respectively. To assess the errors caused by the interpolation of satellite clock errors, we process the data again with 5 minute interval, with the epoch we have exact satellite clock errors as IGS provided. The positioning errors are shown in Figure 2(b). Comparing Figures 2(a) and (2(b), it can be seen that position errors with the epochs that have the exact IGS clock error (Figure 2(b)) are much less the solution with the interpolated satellite clock error (Figure 2(a)). The position RMS errors with the solution that is without using the clock interpolation are 0.021m,0.030m and 0.055m in North, East and Height component respectively. This clearly shows that the satellite clock accuracy plays an important role on the PPP method.

To test the overall performance of the kinematic PPP method in a kinematic environment, we use GPS buoy data with 1s data interval, from 22 to 24 October 2004, at Repulse Bay, Hong Kong island. Two Leica SR530 dual frequency GPS receivers were used, one is set on shore as a fixed station for differential positioning test and another receiver is installed on a buoy in the sea. (Figure 3).

Firstly we estimate the position of GPS buoy with the relative position mode. As the reference station is very close to the GPS buoy (a few hundred metres), most GPS measurement errors can be cancelled out by the relative

position mode. The position accuracy of centimetre level can be easily achieved with this method. In this paper, we will use the solution from relative positioning mode as the baseline to evaluate the accuracy of the kinematic PPP method. Figure 4 shows the height variation of the GPS buoy from the relative positioning mode. The dark black line is the moving average of 5 minute position

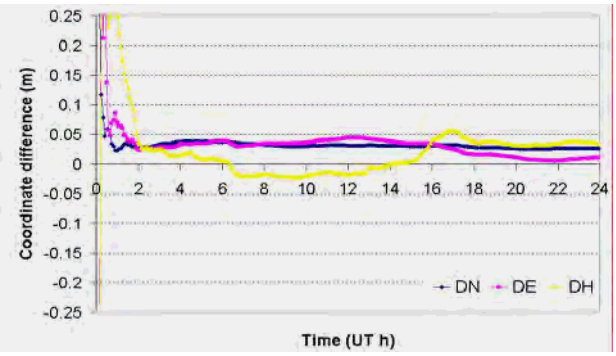


Fig. 1 Coordinate components convergence processing

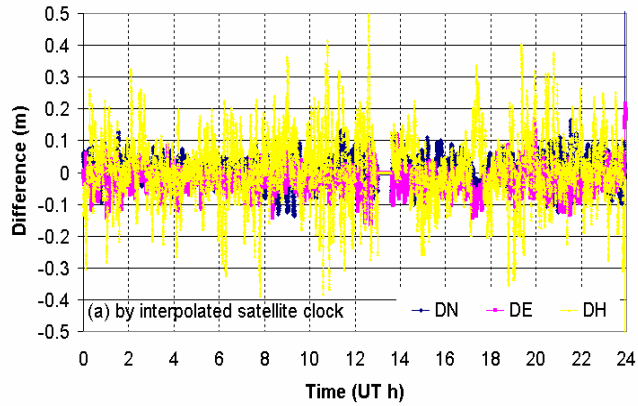


Figure 2(a) Kinematic PPP results with interpolated satellite clock errors

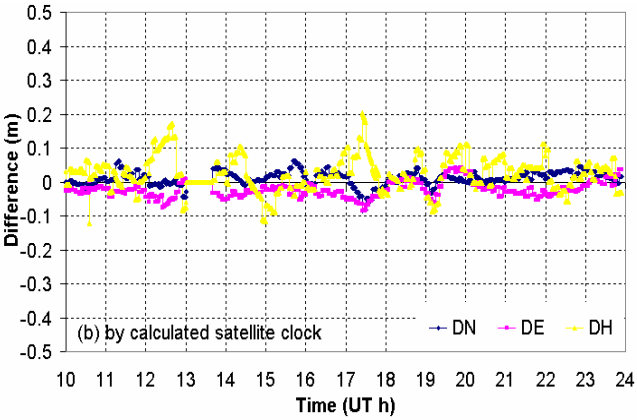


Fig. 2(b) Kinematic PPP results with the IGS satelite clock errors

Figure 3(a) The location of the reference station

Fig. 3(b) The location of the GPS buoy

results, which may represent the tide variation over the two days, while the light line represents the sea waves. In general, the mean sea surface observed by GPS is quite smooth, except two abrupt changes happened (see Figure 4) and that are due to the periods few satellites were observed with large DOP value.

Then we estimate the position of the buoy using the kinematic PPP method proposed in this paper. The IGS rapid products (satellite orbit and clock errors) with the update rate of 15 minutes for orbit and 5 minutes for clock errors are used in data processing, as the IGS final products will only be available after a week of the observation. Figure 5 shows the height of the buoy from the kinematic PPP method by using 1 second GPS data from buoy only. The satellite clock errors are interpolated to every second by the 5 minute interval clock errors. The overall tide trend can be seen in Figure 5. However the results are much noisier compared with the relative positioning mode (Figure 4).

Figure 6 compares the height difference between the relative mode and the PPP mode with 30 second interval. The errors of the PPP mode are quite large to over 0.5 m, with the RMS difference of 20 cm. Similarly, we process the GPS buoy data again with 5 minute update interval with the epoch exactly the same as the IGS clock error update. The height difference between the relative mode and the PPP mode is shown in Figure 7. The positioning differences are less than 20 cm, with an RMS difference

of 12 cm. This confirms again that accuracy satellite clock error estimation is important to the PPP data processing.

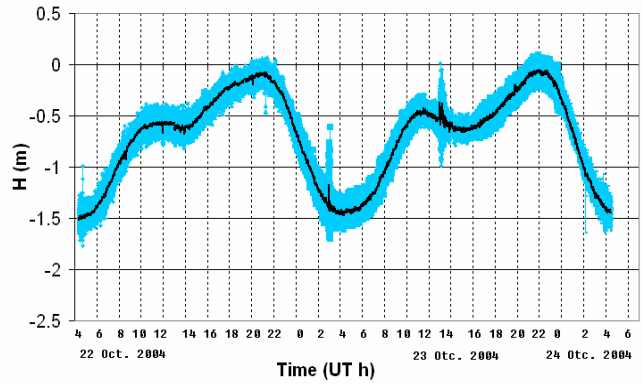


Fig. 4 Height result for GPS buoy by kinematic DGPS

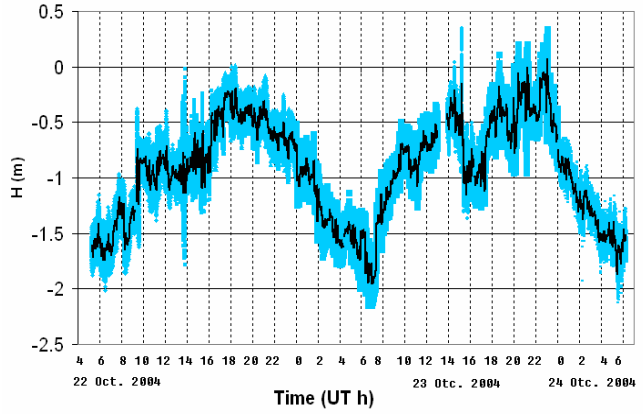


Fig. 5 Height result for GPS buoy by kinematic PPP with IGS rapid orbit

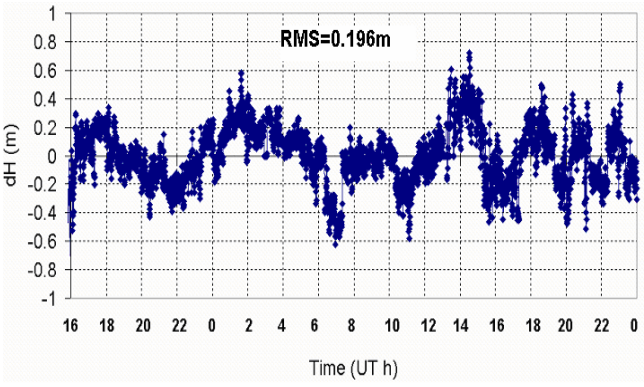


Fig. 6 Height difference between the PPP and relative mode with 30 second interval

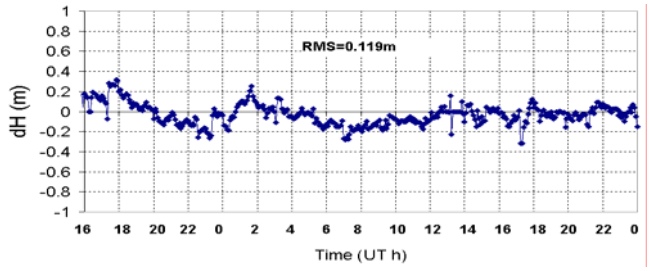


Fig. 7 Height difference between the PPP and relative mode with 5 minutes interval

5 Conclusion and suggestions

In this paper, the basic precise point positioning (PPP) models reviewed and a new algorithm for kinematic PPP is proposed. The key of this algorithm is to separate parameters with their property on the change with time. We will estimate the parameters that des not change much with the time first (such as ambiguity and tropospheric delays). And then the receiver coordinate swill be calculated by removing ambiguities and troposheric delays. The experiment results shows that this algorithm can effectively separate the position parameters (large changes with time) and those parameters with slow changes (tropospheric delays) or no changes (ambiguities) and the decimetre level of positioning accuracy can be achieved with this method. The experiment results also show that the quality of satellite clock errors is important for the kinematic PPP mode. In the data processing, with the interpolation of 5 minute satellite clock error data, the positioning errors are almost doubled for both the static and kinematic cases. Thus, for the satellite clock errors, we should avoid any interpolation from the IGS products. Currently, the IGS already started to provide 30 second interval data and experiments have been carried out to provide 1 second interval data. We expect the positioning accuracy can be improved by using these products.

This paper only shows the preliminary results from the experimental data. Further studies are required to improve the positioning accuracy. Firstly we will process the data again with final IGS orbit and high rate satellite clock production to examine the improvement on positioning accuracy. Secondly, compared with figures 4 and 5, there are a number of positioning jumps in figure 5 with the kinematic method. These are mainly due to the fact that the change of satellite observed (dropped or new satellites). We will also try to use the previous position and satellite ambiguities to constrain the solution to reduce the size of the discontinuity on position.

ACKNOWLEDGEMENTS

This study is under the support by the Hong Kong RGC projects: PolyU 5075/01E and PolyU 5062/02E.

References

Han SC, Kwon JH, Jekeli C (2001): ***Accurate Absolute GPS Positioning through Satellite Clock Error Estimation***, Journal of Geodesy, 77: 33 - 43.

Héroux P.; Kouba J (2001): ***GPS Precise Point Positioning Using IGS Orbit Products***. Physics and Chemistry of the Earth (A), Vol. 26 No. 6 – 8: 573 – 578.

Hofmann-Wellenhof B.; Lichtenegger H.; Collins J. (2003): ***GPS Theory and Practice (5th Ed.)***. Springer, Wien New York.

Moore T.; Zhang K.; Close G.; Moore R. (2000): ***Real-Time River Level Monitoring Using GPS Heighting***. GPS Solution, Vol. 4, Number 2: 63 - 67.

Key K.W.; Parke M.E.; Norn G.H. (1998): ***Mapping the Sea Surface Using a GPS Buoy***. Marine Geodesy, 21: 67 – 79.

Niell; A. E. (1996): ***Global Mapping Functions for the Atmosphere Delay at Radio Wavelengths***. Journal of Geophysical Research, 101(B2): 3227 - 3246.

Rocken C.; Kelecy T.M.; Norn G.H.; Young L.E.; Purcell G.H.; Jr.; Wolf S.K. (1990): ***Measuring Precise Sea Level From a Buoy Using the Global Positioning System***. Geophysical Research Letters, 17: 2145 – 2148.

Zumberge JF; Watkins MM; Web FH (1998): ***Characters and Applications of Precise GPS Clock Solutions of Every 30 Seconds***. Journal of the Institute of Navigation, 44(4): 449 – 456.

Zumberge JF; Heflin MB; Jefferson DC; Watkins MM; Webb FH (1997): ***Precise Point Processing for the Efficient and Robust Analysis of GPS Data from Large Networks***. Journal of Geophysical Research, 102 (B3): 5005 - 5017.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 308-321

Journal of Global Positioning Systems
2004

Would a GNSS need a backup?

Walter Blanchard

Royal Institute of Navigation, Trundle Tower Hill, Dorking, Surrey, UK
e-mail: blanch@pncl.co.uk Tel: ++44 1306 884539

Received: 15 Nov 2004 / Accepted: 3 Feb 2005

Abstract. No navigator likes to be totally dependent on only one navaid – it is an article of faith for many that there should always be a backup system. Several systems have been put forward as possible backups for a GNSS but they seem to have originated more in a generalised feeling that there ought to be one rather than a dispassionate examination of what is involved. GPS/Galileo are radical departures from any previous concepts of radio navigation aids and a full-blown GNSS is an even more radical proposal. There is a good deal more involved than simply engineering and technical matters. There are the questions of who controls them; what the customer interface is; who certifies them for use in safety-related situations; and what legal recourse there is. On the answers to these questions depends whether a backup is needed and if so what form it should take. It is found in this paper that for many non-critical users there is no need for a backup, and that others who may be involved in safety-critical situations already have a backup in the form of their current systems. It is also found that in fact it may be extremely difficult to compose a GNSS in the form it is generally given; that is, a combination of GPS, Galileo and perhaps Glonass. The problem lies not in on the engineering side, but in matters of legality and the sovereignty of individual nations. For these reasons it is concluded that the development or implementation of a new system purely to act as a backup for a GNSS is not necessary.

Key words: Satellites, Navigation, GNSS.

1 Introduction

Most people have a slightly uneasy feeling that if they rely a great deal on something there ought to be a backup for it, or at least some sort of fall-back provision. For example, cars; if our car breaks down unexpectedly we may have a major problem. No-one wants to be stuck several hundred miles from home. Although fortunately this is now a rare occurrence with modern cars if it happens a mobile telephone will bring a breakdown truck to the rescue. This combination of basic reliability and the availability of rescue services means that very few drivers are to be found towing a spare car behind them "just in case" (except perhaps for large RV owners!). What was really a similar consideration occurred in the aviation world some time ago when the wisdom of allowing twin-engine commercial airliners to transit oceans was much debated. It was the legacy of comparatively unreliable piston engines that caused the debate; the fact that the modern jet engine is far more reliable took some time to establish and now most trans-oceanic aircraft are twin-engined.

The same sort of thing is occurring in the navigation business now we have now come to use GPS so much. There is a popular feeling there ought to be a "backup" in case it is "switched off" or somebody jams it. This feeling is often reinforced by previous experience with radio-based systems when a transmitter failed or a receiver packed up, and of course there is always the wise old saw of never relying totally on one system. But, looking ahead, how much of this will apply to GNSS, and, if it does, what type of backup should there be? The rather vague feeling that "there ought to be a back-up" is not much of a guide. Many factors are involved ranging from the effect withdrawal of GNSS would have on a users' operations and how a back-up would enable him to continue them, to the cost of installing back-up infrastructure and, not least, the cost to individual users of installing back-up user equipment.

In any case, just what IS a back-up? How do you define the point at which "loss of GNSS services" becomes so severe that a backup is needed? Should it be able to completely take over every function of a GNSS; most of them, or only the most essential? If it can do the same job as a GNSS, why hold it in reserve? It would be expensive and if not used except on very rare occasions would be a

financial disaster. Does every user actually need a back-up? Do different users need different back-ups? Who would pay for it, run it, control it?

At the moment we have no direct experience with a GNSS; all we have had so far is GPS, with a few people using Glonass as well. Perhaps our experience with GPS might provide a guide, but we must be clear about the distinction. GPS is owned, operated and controlled entirely by one country, the United States. A GNSS, by definition, will consist of a combination of several distinct systems possibly in multiple ownership but essentially under consensus-control authority.

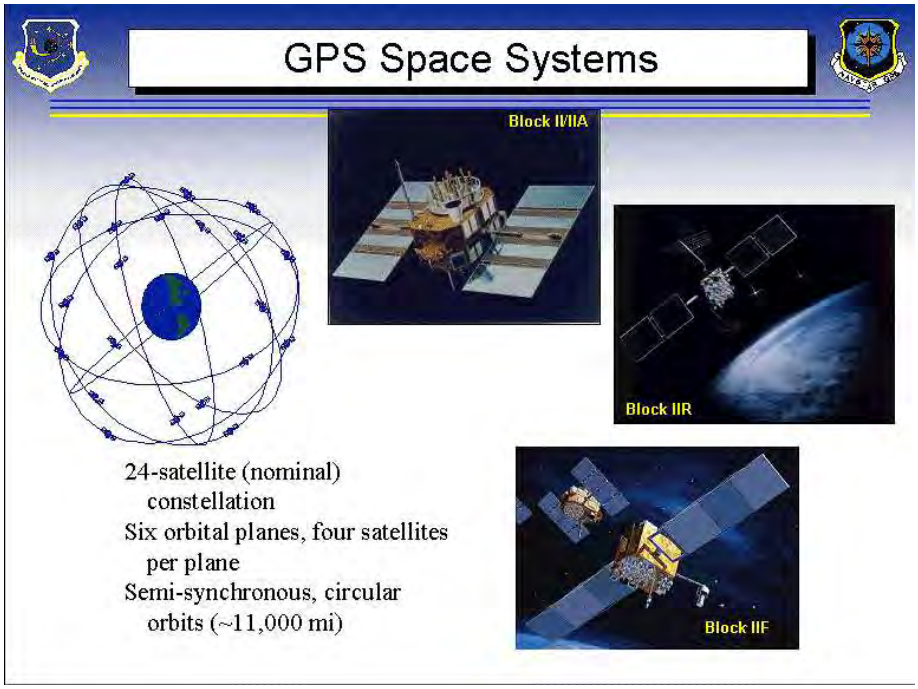


Fig. 1 GPS

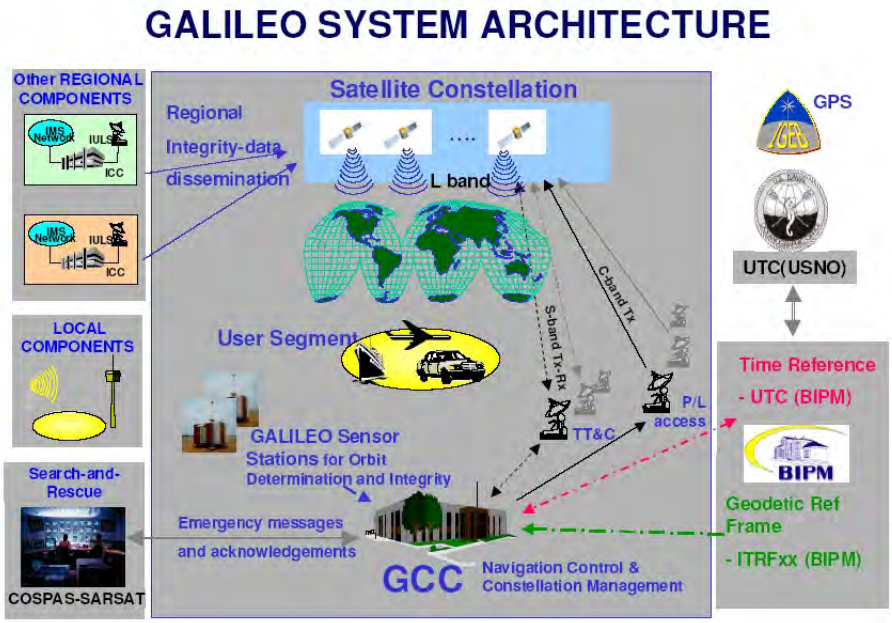


Fig. 2 Galileo .

2. THE GPS SITUATION; NON-MANDATORY USERS.

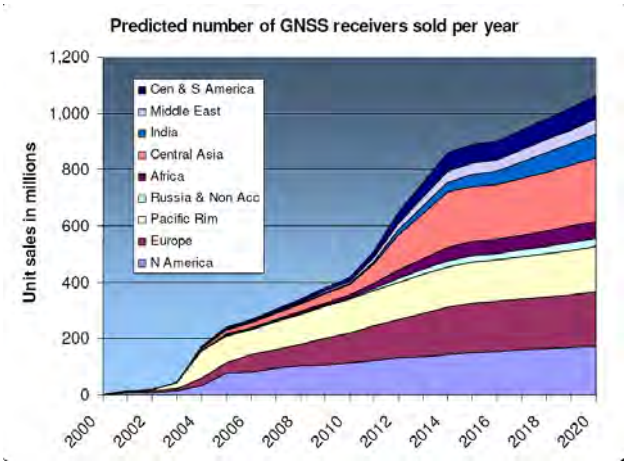


Fig. 3 GNSS Market (from a diagram by EC GJU).

This diagram shows that by far the great majority of users of a GNSS will in the category of "non-mandatory" users – those who are not compelled by law to carry a satnav system. They might be termed "voluntary" users.

Fig. 4 "Voluntary" users.

So let's start by looking at how GPS is used by them. The category includes the millions of "pleasure" sailors and aviators; the even more numerous car-drivers who found a GPS in their new car; and all those others using it on what might be called a "hobby" basis.

For most of them, GPS is a luxury and many do not even bother to learn how to use it properly. Its loss would be inconsequential and they would just fish out the old map every car has under the seat. Commercial vehicles use it rather more seriously to find addresses and alternative routes but here again no commercial driver is without his A-to-Z and a good map. Discretional vehicle-tracking systems (those not required by law) would undoubtedly miss it but not all use GPS. If any of these users really felt it was essential to have it at all times they would either install a system not based on GPS, or make arrangements for their own back-up. Datatrak and other non-satellite systems are well-established and available for those who do not want to trust GPS.

As regards light aviation, the standard GA aircraft radio-navigation fit is still VOR, maybe a DME, an ADF and a transponder, which will be there whether or not GPS is fitted. There is no proposal to withdraw any of these in the foreseeable future. Although GPS is considered an excellent navaid they will always be there to fall back on should GPS fail.

Small-boat sailors invariably use GPS now, but before GPS they seemed to manage quite well using traditional systems – DR, perhaps a sextant, radio bearings, etc. The only tricky situation that might happen is if they set out with GPS and found halfway through the trip it disappeared. But exactly the same thing would happen if their GPS batteries went flat and presumably they always have this possibility in mind.

In the ultimate, if any of this category of user cannot make a trip without GPS, they are in the happy position of being able to cancel it, unlike commercial users. The expense of installing another system purely as a GPS backup is simply not justified for these non-mandatory users.

3 THE GPS SITUATION; COMMERCIAL USERS.

But what about those for whom GPS is important in their commercial activities? What about professional aviators and sailors; land surveyors; offshore surveyors; those using GPS timing for keeping cellphones working, etc?

Fig. 5 Commercial Users.

We can dismiss commercial aviation immediately, for much the same reasons as light aviation. VOR/DME, ILS and INS are mandatory as they have been for many years, and no aircraft can be certificated for IFR flight or allowed to land at major airports without them in some combination or other. Fig. 6 illustrates the profusion of such aids available in a major terminal area.

Fig. 6 London FIR Aeronautical Chart.

Although GPS is often also fitted it is only as a supplementary aid which means prima facie the owner must demonstrate he can navigate without it, in other words, that he has an adequate fit of the primary systems. There is therefore no need for the commercial aviator to have a back-up for GPS, he does not rely on it. Nor is he likely to do so for many years – there is no proposal to allow the use of unaided GPS as a sole system, and ICAO has stated no such intention. One very good reason is that it has been adequately demonstrated that unaided GPS does not have sufficient availability for sole use and it will certainly need the addition of another system such as Glonass or Galileo, thus turning it into a GNSS, before it can be considered. GPS augmentation systems such as WAAS/LAAS and EGNOS have their own problems and it may be some time before they are worked out. The professional aviator will need some very hard convincing before he will be happy to ditch his well-established VOR/DME in favour of a satnav alternative and it is only then that talk of a backup will become relevant. Even then, the more likely course is that having already amortised their conventional VOR/DME, etc

installations, airlines (and civil aviation administrations) will insist on them remaining as the major reversionary system for GNSS. Talk of a different backup, requiring completely new equipment and installation costs, will be dismissed instantly. It is often forgotten that the real costs of putting a new piece of kit into an aircraft lie in its installation and integration costs, the costs of setting up workshop support, retraining engineers, and so on, not just the bare cost of the boxes. Quite apart from these technical arguments, the problem of legislation in countries having no control over GPS seems insoluble at present. For these reasons, arguments using commercial aviation as justification for a GPS back-up are poorly based.

Turning to commercial shipping, it is here without doubt that GPS has made the biggest inroads. No doubt the reason has been that there is an IMO requirement for ships to carry either a "GNSS" or, since one is not yet available, an equivalent which since the demise of Omega has been universally interpreted to mean GPS, although IMO does not mention GPS by name. Practically every ship now has GPS on board for deep-water navigation, and not merely one; most have two and sometimes as many as six. GPS is also used in more demanding situations where considerable accuracy is required, such as berthing, although it is then used in differential mode rather than raw. The question of a back-up thus might seem rather more important for commercial shipping than for any other class of user, but let us examine the problem in more detail.

Before GPS nearly every European vessel had Decca Navigator and swore by it as the most accurate and reliable system of its day.

Fig. 7 Decca Navigator Mk 21 Marine Receiver.

That didn't stop them deserting it in droves once they discovered GPS, making Decca unviable commercially and eventually leading to its switch-off. One of the main users of Decca were fishermen but they similarly went over to GPS once they found they could buy a GPS for less than a Decca. They didn't seem to have much concern about keeping it as a back-up for GPS even though they had already installed and paid for it. This lack of protest over the demise of Decca demonstrates that there is really not much concern amongst these users about the availability of a back-up for GPS. No doubt they consider that if they were out at sea and lost their GPS they would simply revert to long-standing traditional methods of navigation.

The loss of GPS while on the oceans would not actually be a matter of deep concern to the average Master; standard navigation could still be carried on quite successfully by traditional methods and the only loss would be perhaps a certain amount of extra fuel used. In close waters a Masters' main concern these days is avoiding other shipping and for this purpose his radars are his main tool. They also serve admirably as a means of avoiding land and other obstacles and do a job no GPS or GNSS could do.

More critically, some vessels, for instance high-speed ferries, now use differential GPS for berthing and its sudden loss could be serious. However, they still station lookout men at each critical point on the vessel and berthing without GPS is regularly practised. Its loss would not stop operations although it might slow them down. It should also be remembered that if it were desired to continue operations uninterruptedly should GPS disappear it would be essential that any backup had the same precision as the local special-purpose DGPS systems they use, claimed to be accurate to the region of half a metre. There is no current marine radio aid that can achieve this, differentially operated or not, so the selection of a backup radio aid would be extremely difficult.

Some of the most critical users of GPS are the marine and land surveyors. Both have adopted it because of its uniquely high accuracy, particularly in differential mode, and its height-measuring capability.

There are no other systems that would enable them to continue operations uninterruptedly; no land-based system can measure height or get anywhere near DGPS accuracy. For this reason its loss would be a very serious matter for them that could not be solved by any current standard system. However, neither operate in "loss-of-life" situations and, although costly, could suspend operations temporarily. The argument for a backup here is not technical – it is economic.

Then we have the proposed use of GPS for road-tolling; tracking offenders, and similar "Government-inspired"

ideas. There is actually a much bigger problem here than whether the signals might be lost - it is whether this sort of use is legal. If a government says to its citizens "We intend to make a law that you must fit GPS so that we can track you and charge you for using roads/going too fast/parking in the wrong places, etc." then it is legitimate for the citizen to ask what measures the government has taken to ensure GPS will always be there providing the necessary accuracy to do these things. Has it obtained a guarantee from the Americans that they will never alter it in any way without first consulting them about what effect it might have on their legislation? Of course not, and unless the USA changes its attitude towards foreign users very radically it never will. Accordingly any such use can only be on an experimental basis and that being so there is no need at all for a backup.

Lastly we come to the timing business. Most data systems rely on accurate timing, including much radio and television broadcasting as well as cellphones and broad-band web data. Their basic timing is often derived from standards incorporating a GPS-disciplined reference oscillator but it is incorrect to claim that were GPS to disappear these standards would become unusable. Their specification usually includes the ability to operate sufficiently accurately for a week or more without GPS correction and although this does not cover complete and permanent loss of GPS it certainly covers short periods of accidental or deliberate jamming. In really critical applications timing is derived from an atomic standard that does not require GPS correction.

We see therefore that although individual users of GPS might want a backup they really cannot have it both ways – GPS is a totally free service provided by the generosity of a foreign government, and they cannot blame their own governments for not using taxpayers money to backup a system which they did not install or authorise, and do not control.

Actually, many Governments have recognised the reality of the situation; that is, that a large number of navigators will be using GPS regardless, and have installed Government-run differential and augmentation systems. Thus, IALA (the International Association of Lighthouse Authorities) has encouraged its members to use their old MF radio-beacon sites to provide differential services, and general-purpose systems like WAAS and EGNOS, based on satellites, have appeared.

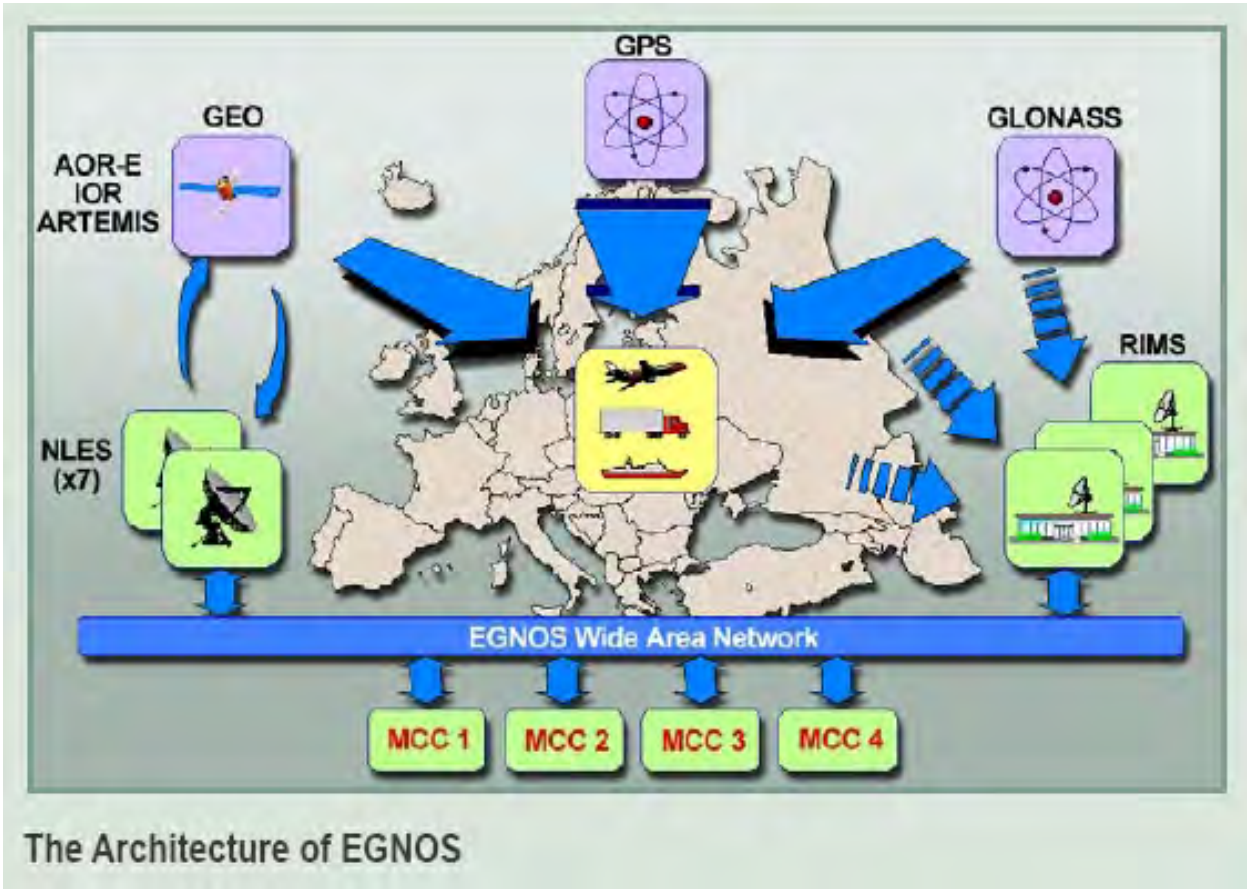


Fig. 8 EGNOS system (from an EC diagram) .

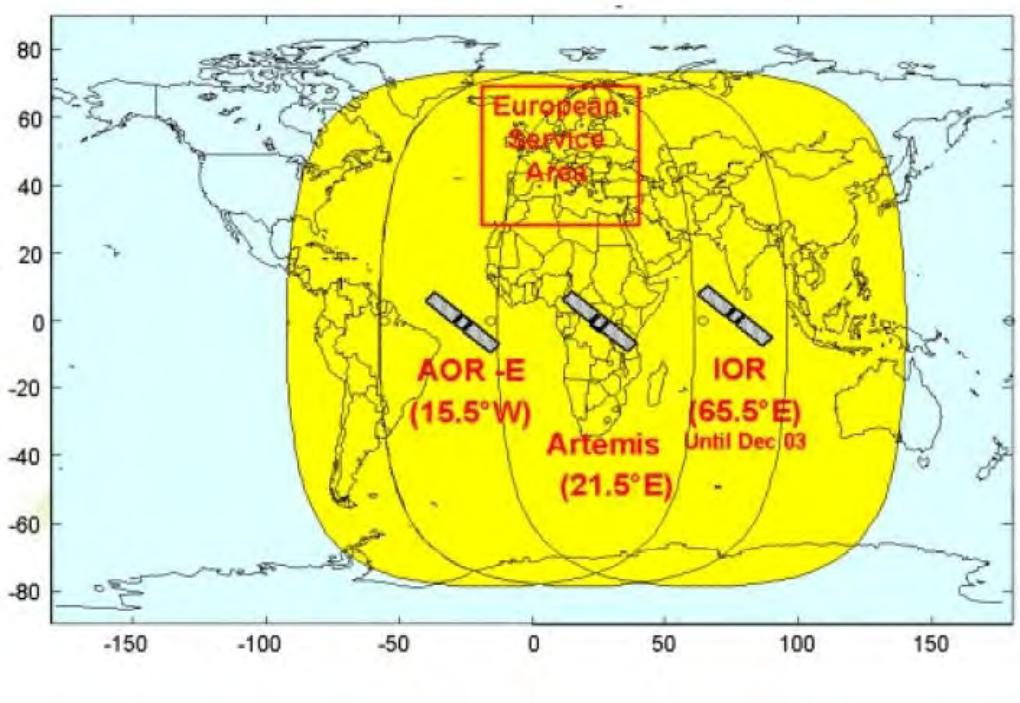


Fig. 8 EGNOS system (from an EC diagram) .

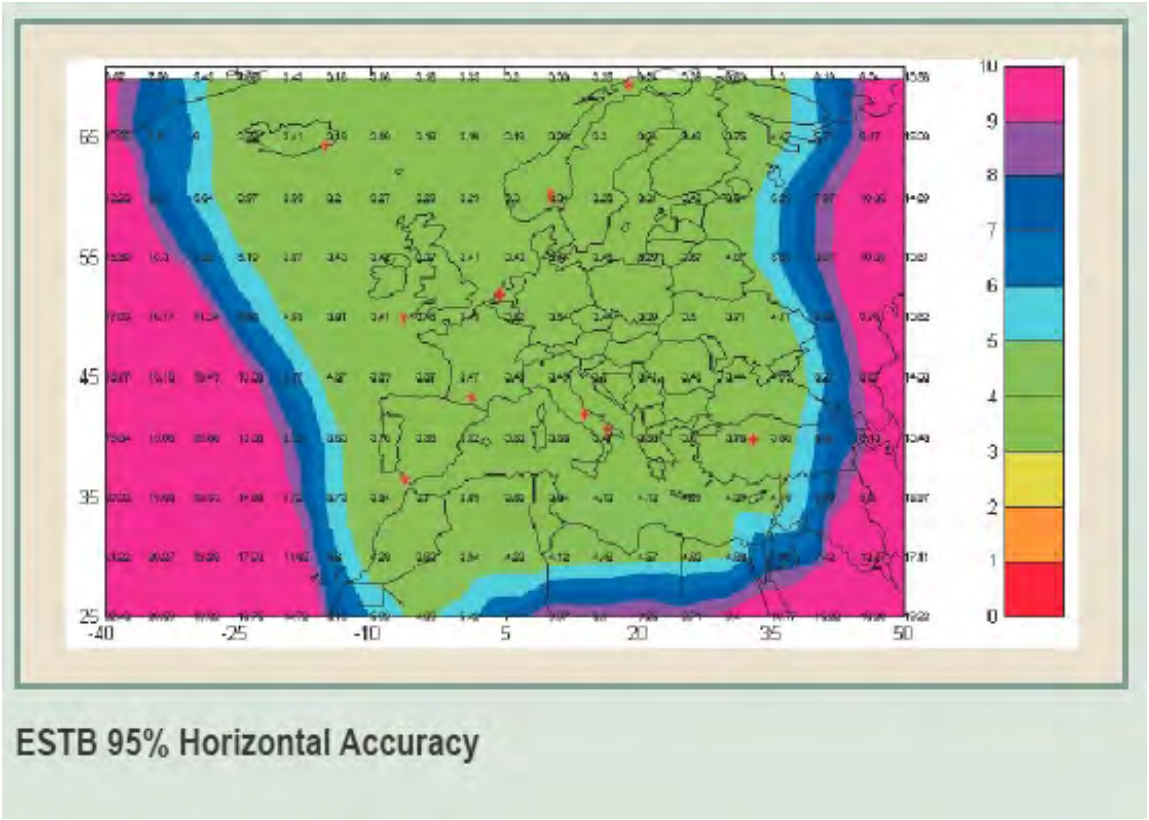


Fig. 9 EGNOS cover (from an EC diagram).

These systems monitor GPS continuously and transmit correction information so that if a GPS satellite fails it is instantly apparent. They also transmit other information enabling users to obtain improved accuracy. These systems are not fall-back systems, in that if GPS fails totally they could be used to continue navigation, but at least they are an independent check on its performance.

4. THE GNSS SITUATION – DOES IT DIFFER FROM GPS?

Does any of this change when we start thinking about a GNSS instead of just GPS? Yes, it does, in quite major ways, but perhaps not those we might expect. A GNSS will be an integrated satnav system made up of several components owned perhaps by different bodies but under the control of one central authority. The one most often talked about is a combination of GPS with Galileo, the European system, so let's have a look at that.

GPS, it cannot be said too often, is a military system with civil users allowed on sufferance.

The main driver behind Galileo is civil control with the military allowed in on sufferance. To combine the two under a common set of rules will require the very considerable reconciliation of US military with European civil views which, as we have seen repeatedly quite recently, is extremely difficult to attain and may never happen. The likelihood is that the US will continue to operate GPS under its own military control and Galileo will be operated mainly under some form of international civil control. "Civil control", in this context, means amongst other things that there is full disclosure of basic operating engineering parameters and standards in order that specifications can be drawn up for safety standards and minimum operating performance specifications. That has not occurred for GPS and is not likely to. It is why ICAO has never certified it for sole use. That being so, any civil use for safety-critical purposes of a GNSS that combines the two could only be done on the basis of what Galileo alone offers, thus negating the whole purpose of integration. It might as well be done completely ignoring GPS from the outset. This is so irrespective of whether they are technically integrated; that is not particularly difficult and in fact has already been agreed formally. The proponents of Galileo have said from its inception that it must be "inter-operable" with GPS and technical co-operation over such things as modulation, coding and frequencies has been excellent.

Quote:

"The Council agrees :

.....

That Galileo should be interoperable with existing satellite navigation systems ; it should in particular be interoperable with GPS and its successor systems through an EU-US agreement that should be negotiated as soon as possible"

(Council of European Transport Ministers Statement, 26th March 2002)

Since then, a formal agreement has been concluded. They can now co-exist on the same frequencies and there is no chance that one will jam the other. It will not be difficult to make a receiver that will receive both simultaneously. Both will provide a free service and for those users who do not have to use satnav for mandatory purposes it will be marvellous to have upwards of 60 satellites in the sky and performance in cities and other obstructed situations will be much improved. But the problem of legal recourse will still remain for the operating authorities who must consider how to reconcile GNSS with their existing regulatory structures which have been built up over many years and are so embedded they are extremely difficult, if not impossible, to change.

It has already been emphasized that there is a great difference between non-mandatory and mandatory users but there is no harm repeating it. We must be very careful to draw a distinction between non-mandatory systems and those fitted to ensure compliance with regulations of one type or another. Once the fitting of a system is required by law a whole set of new parameters appear in which a prime consideration is who owns, operates and controls the basic system. Who is it that guarantees its performance? You cannot make a law requiring the fitting of a system you cannot guarantee. There are no guarantees for GPS; to be sure there is a set of basic operating goals set out in documents like the "SPS Performance Standards" document but they are not guarantees. This is one of the main reasons for Galileo, which, being fundamentally under civil control will have the background legal structure to enable guarantees to be given.

It may be thought too much emphasis is being laid on this aspect. After all, a prime example of a radio navaid that was never under direct legislative control but nevertheless widely used was the Decca Navigator system. At its peak one of the most widely-used marine radio navaids in the world it never came under Government or international regulation and was operated solely on the basis of common commercial prudence. But because of this lack of regulation it was never adopted as a mandatory navaid in either aircraft or ships and therefore whether one was carried or not was entirely at the discretion of the Master. Over 30,000 vessels and 10,000 aircraft decided it was worth the risk and fitted it. Decca was so successful because it filled a gap in the armoury of marine navaids and was operated by a commercial company that had to be responsive to its customer's wishes. There were free alternatives to Decca in the form of Loran and Transit but although they were sometimes fitted the same ships

almost always had Decca as well. Decca became a byword for reliability and performance and those are attributes valued by a navigator above almost all else. The key was that it was a commercially responsive company that ran it and spent a great deal of effort ensuring its customers were happy with it. Not so with GPS. It is not a commercial offering and the aims of its owners are not primarily to satisfy civil users. Consider a little recent history. Most years, there are during the Spring months significant periods – half-an-hour or more - on some days when GPS satellite availability is quite marginal over the UK - only four satellites at elevations above 20° and in poor positions for good fix accuracy.

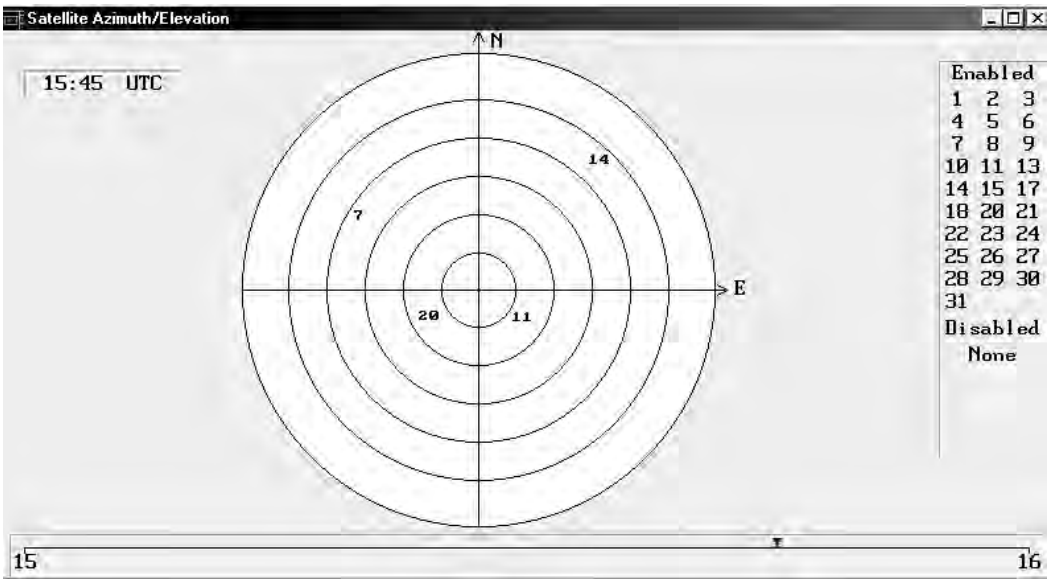


Fig. 10 GPS availability, UK.

This might seem a high cutoff elevation but it is one many people are using and is reasonable for cars travelling in built-up areas where even higher cut-offs might be necessary - 40° has been quoted. These four satellites often only provide an HDOP of 20 or so (35 has been seen on occasion) and the result is very poor accuracy - 200 metres or more.

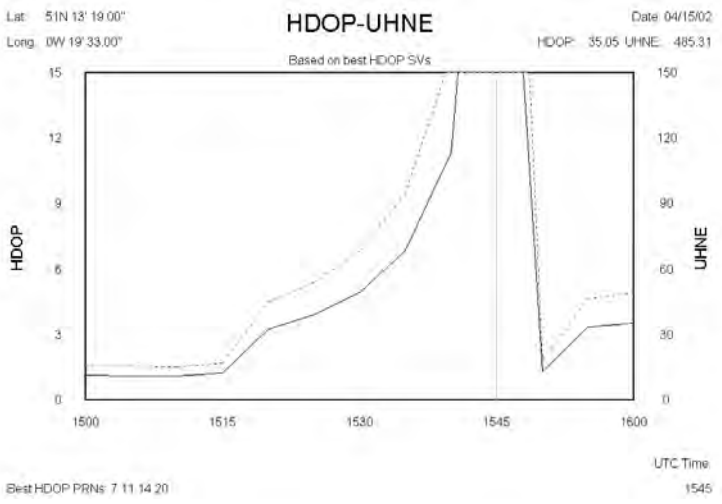


Fig. 11 Poor GDOP

On these occasions there was nothing wrong with the GPS system; there were no satellites off-air; no maintenance was being done; and there was nothing in NANU's about it. These periods of poor HDOP were within the published performance limits for the GPS Standard Positioning Service and therefore there was no cause for complaint. The problem was just a fortuitous combination of satellite positions caused by some satellites not being quite on station and others being drifted around their orbits to new positions. Only Northern Europe was affected and then only for periods less than the outage periods permitted by the statistical availability parameters. It is entirely possible that this could recur in future - *and there is absolutely nothing to be done about it!* Would an EU Government be prepared to have its expensively-implemented road pricing scheme collapse even for half-an-hour at unpredictable times?

While this is true of GPS, one of the major drivers, if not the main driver, behind Galileo was to remedy this lack of a GPS performance guarantee by initiating a civil-based system operated with civil users as its main concern and totally responsible to them.

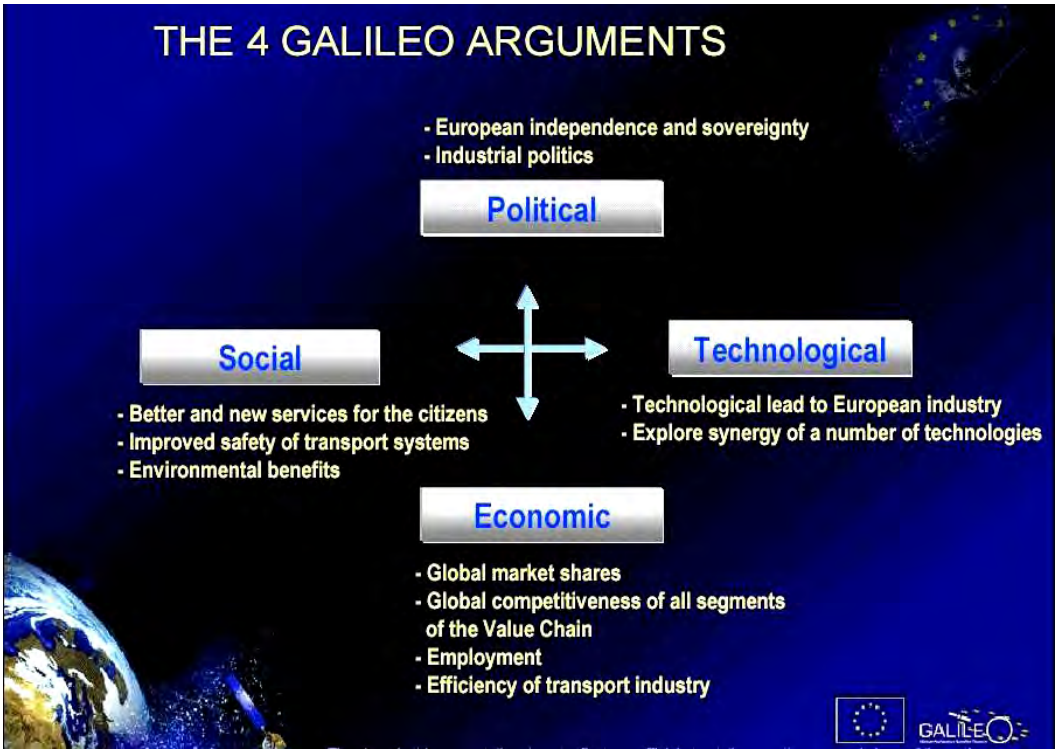


Fig. 12 The Four Galileo Drivers.

If it were to be used for road-pricing then no doubt extreme care would be taken to ensure that this sort of situation did not occur and legal guarantees would be much easier to enter into. Although the military will also use Galileo, (under the guise of a "Government Service"), they will not be in the same powerful position as in GPS since they will not be providing the major funding, probably much to their relief! The PPP concept is designed in large part to provide some degree of Governmental control over a privately-funded enterprise and is being actively discussed for Galileo. That is not to say there are no problems at all using Galileo but only that they will be present to a lesser degree.

5. COULD A LEGALLY-ENFORCEABLE GUARANTEE REGARDING GPS EVER BE GIVEN?

The US DOD, like any other national military organisation, is charged with the defence of its national citizens, not the provision of civil services even to its own citizens. That is what other Government departments are for. So, although there might be an understanding between civil and military within the US itself, possible because they both operate under the same fundamental legal system, there is no chance that that understanding could be extended to foreigners who considering it at the most venal level contribute nothing in financial terms. The DOD is not a commercial organisation committed to providing value for money, and foreign part-control of

one of its vital systems would, rightly, be anathema to it. There have been attempts in the USA to persuade us that GPS is not really under military control and is fully responsive to civil needs. It is pointed out that it is managed by IGEB, an organisation on which 7 US civil GPS-user agencies are represented as against only 2 military, and is operated by the US Air Force only on behalf of the Government as a whole. This is very commendable but one might ask where is the *foreign* civil representation? Of course, there isn't any, and one would not expect there to be - it is a US National system, not an international one. And that is the problem.

Fig. 13 GPS control

Sovereignty is jealously guarded by all countries. After all, without it a country is not in control of its own affairs and subordination of sovereignty to a foreign power, which is what legislation for the use of a foreign-owned system amounts to, has no historical precedent in any major country. For instance, the UK Civil Aviation Act of 1921 was passed with the express object of ensuring the UK maintained complete sovereignty over its airspace as well as land, a principle maintained rigorously by every country ever since. We need only look at the problems still being encountered with attempts to establish a common air traffic control system for Europe to see the complications even a slight derogation of this principle can cause.

6. WILL GALILEO ALONE FILL THE BILL?

We are therefore in a situation where it is impossible to certify GPS for use in any regulated civil situation and if this situation persists we will have to rely totally on Galileo. That means that Galileo could NOT be reinforced by GPS and we would still have only the 30 Galileo satellites in the sky capable of being used for these purposes. What then of the much-vaunted reinforcement of GPS by Galileo in city canyon areas on which so much depends? We would simply have the same situation as today, not enough satellites.

Galileo at least starts off with the premise it is a civil system setting out to satisfy civil needs.

That being so, all the usual panoply of civil and commercial procedures and safeguards can apply. If a proper Public Private Partnership is forged, in which a not-for-profit, probably quasi-Governmental, organisation oversees the safety and regulatory aspects while private enterprises seek profit in the system where they may, it should be possible to ensure service guarantees sufficient to enable legislation for its use to be feasible. The difficulty Galileo will have is in the formation and working of the overseeing authority. Since aviation will be only a minority partner it cannot be the sole or even the major determinant of policy and an organisation such as Eurocontrol will be inappropriate. However, there is plenty of precedent for a new organisation to be formed along the lines of the European meteorological and telecommunications satellite consortia EUMETSAT and EUTELSAT; EUNAVSAT perhaps?

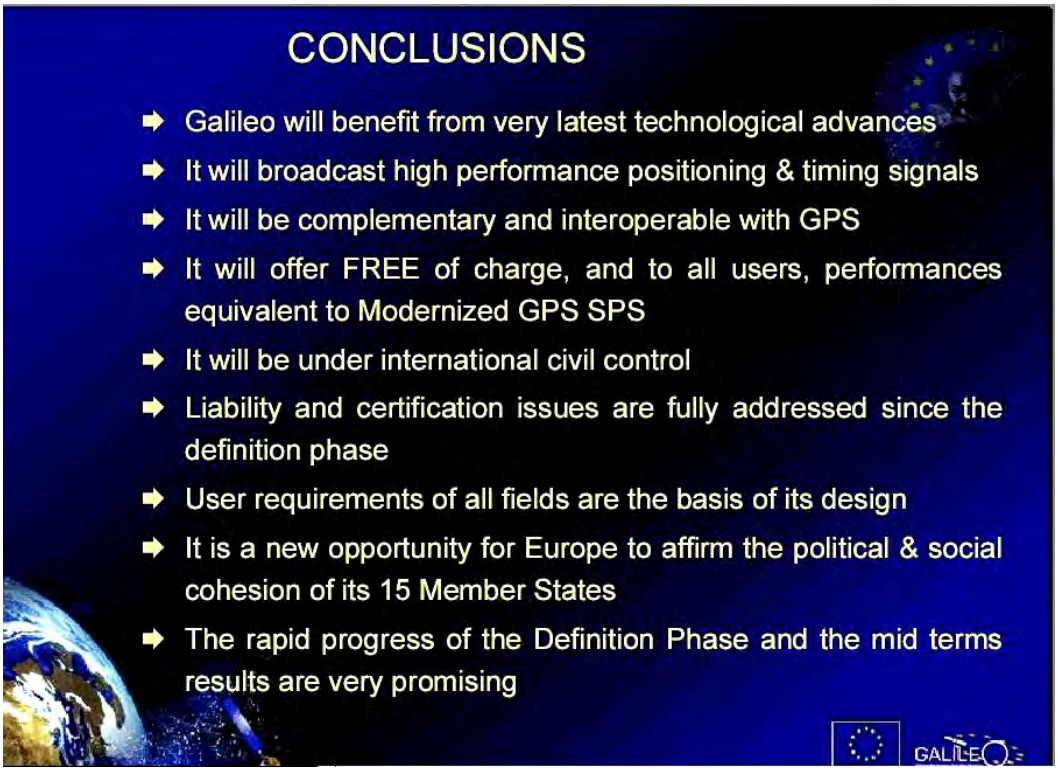


Fig. 14 Galileo aims.

The EU need not prevent non-EU countries joining in, as already demonstrated by non-EU participation in these and other organisations. Although no doubt the usual tedious and lengthy consultation would be involved there appears to be no really fundamental problem preventing such an organisation being formed. Some political and sovereignty issues would still remain, of course, but in a less severe form considering the existence of the European Union. The big difference between these older organisations and EUNAVSAT would be that are no legal compulsions surrounding them. It is not yet a jailing matter if you do not use EUTELSAT or EUMETSAT, but it might well become one if you refuse to have EUNAVSAT in your car!

But a word of warning. Little is being said publicly about what type of organisation might eventually run Galileo, for all sorts of very good reasons. One is that it will take a very long time to thrash out and it would be premature to go on the record now. So the likelihood is that the initial operation of Galileo will be under the auspices of ESA as a test and trial system but for how long will that last? 20 years? While it is in the status of a T and T system no legislation at all could be passed so it might not be until 2030 or so we will have a fully civil-responsible Galileo run by a legally-responsible organisation and all its putative advantages secured.

7. 2010, GALILEO, GPS III, AND ALL THAT.

Let us ignore such gloomy forebodings and assume that in 2010 we have a civilly-responsible Galileo system around which legislation could be framed but with which GPS is unusable in any legislated situation. Unfortunately it will be a Catch-22 situation because the technical problems we already have with GPS will still exist - a restricted number of satellites causing non-performance in city areas - so no performance guarantee to adequate standards will be possible and that would prevent legislation anyway. And what of the millions of drivers who would have GPS-only receivers - are they to be told to get rid of them and fit new, certified, Galileo-only models when they would not see any significant difference in performance? At, possibly, A$2000 a time (it would be claimed certification costs are not negligible) when they can get GPS sets for a few hundreds?

Aircraft do not experience the same restrictions on satellite visibility as cars and provided the legal situation is resolved the use of Galileo for approach and landing could, in principle, be authorised. So it might - for European and other Galileo-subscribing non-US commercial air traffic, but what about intercontinental traffic? The US would no doubt not want to authorise Galileo, not having any control over it over their own

territory, in the same way Europeans could not authorise GPS. But if Europe insisted on legislating for Galileo the result would be that international traffic would have to carry dual capability *and separate* GPS/Galileo equipment instead of a combined GNSS box. This would come about because mixing a certified with a non-certified system is strictly a no-no. This might not matter as far as cockpit operation was concerned but the important point is that they would be carried *not to reinforce one another* but so that the appropriate system could be switched in when over different countries. If an aircraft had an incident in the USA while using a dual GPS/Galileo equipment that integrated the two instead of treating them separately the US authorities might wash their hands of the problem on the grounds that a non US-approved system was in use, and the reverse might happen in Europe. This is the exact situation airlines now fear and is why there will almost inevitably be a strong reaction if it is ever proposed that GPS/Galileo or any other satnav system should be legalised as a sole system. In the face of this, satnav would stay where it is now; a useful support system for INS but not much else. Since it would itself then be in the position of a back-up for other aids there would not be much point installing a back-up to back-up the back-up.

8. A BACK-UP FOR A GNSS?

If the above is correct then the obvious inference is that we will never have a true GNSS. There will be GPS, Galileo, perhaps Glonass, possibly all integrated together *at the engineering level* and receivable on one single receiver, marvellous for the non-mandatory user but useless when legislation is involved. It may well eventually be possible to legislate around Galileo, since that is one of its prime purposes, but hardly around either GPS or Glonass while they maintain their military stance. So the question is moot, there will not be a GNSS to back-up.

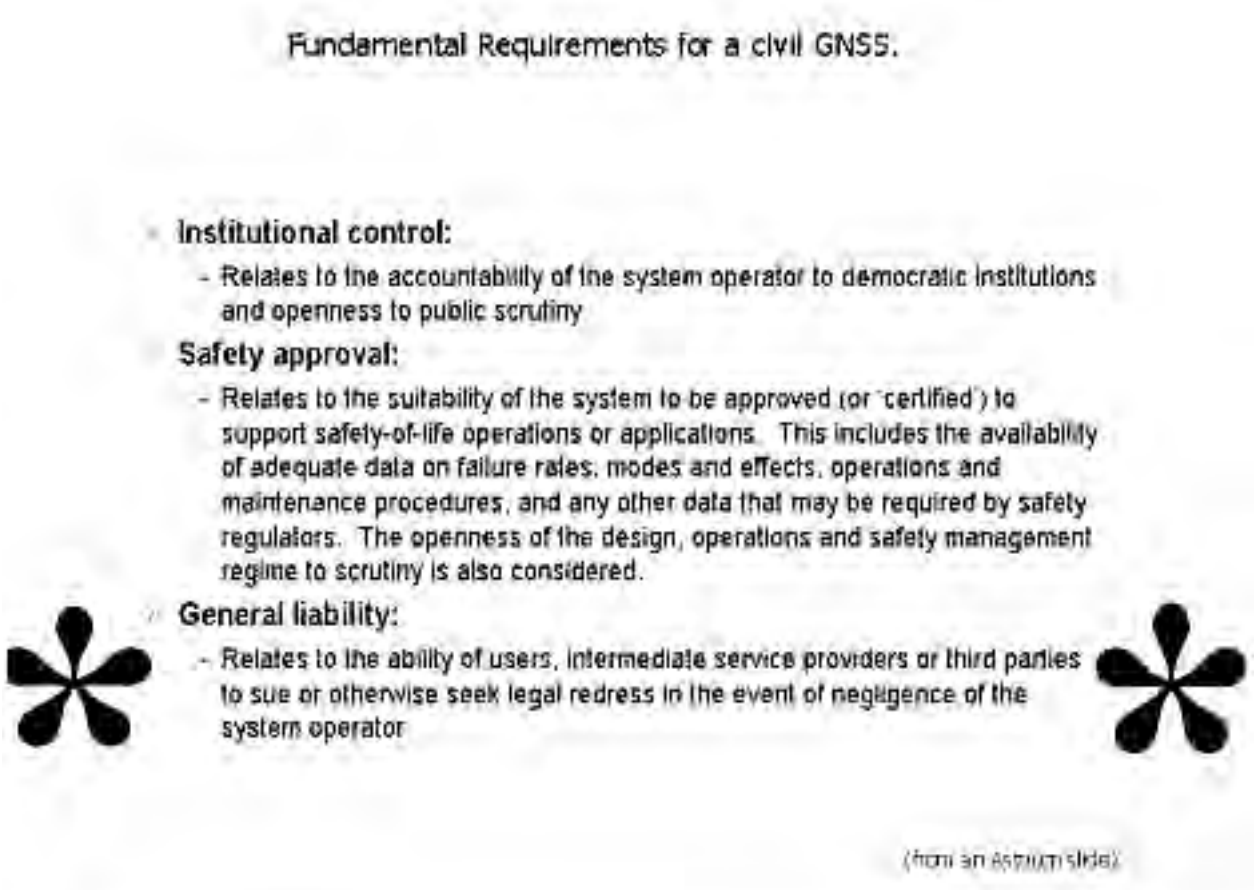


Fig. 15 Requirements for a GNSS.

9. A BACK-UP FOR GPS OR GALILEO, THEN?

We have seen above that a back-up is required only for those for whom satnav is an essential tool, for the others it is only a matter of convenience. But it turns out that neither of the major professional users, aviators and mariners, require a back-up because they already have well-proven systems in place; VOR/DME/ILS for one

and radar for the other. For "voluntary" users, no back-up is required either because it is not generally a safety matter. The cases of amateur sailors and aviators who might get lost if they suddenly lose their satnav are the "hard cases" that make bad law. What is it that might cause them to lose their satnav anyway? Ignoring sillinesses like failing to carry spare batteries there is actually very little. The hard case we think about is where satnav has been working perfectly well on a trip and it suddenly disappears because all the satellites have been switched off. This is actually an impossibility because they could not be made to all disappear instantly; they would disappear one by one as they came within range of command stations that could command a switch-off. And it would be done only in a "grave emergency", we are told, requiring an international situation so bad that it would have been general public knowledge well in advance and the possibility of a GPS switch-off well-known. As regards a total failure of the GPS command system, that is about as likely as all four engines failing simultaneously on a 747. Everything is duplicated, if not triplicated, and even if the command centre in Colorado were to fail completely, there is a reserve one in California. Failures in the satellites themselves are certainly possible and have occurred. The guard against these is proper monitoring and Galileo will have much improved built-in monitoring compared with GPS, the key being instant notification to users of a failure through the satellite signals themselves transmitted from other satellites.

What about "natural" causes? There are no significant propagation effects that could affect all visible satellites simultaneously. Scintillation can occasionally cause temporary fades of individual satellites but it is a very localised effect and there are always sufficient satellites in other parts of the sky. There are no precipitation static effects or skywaves as with low-frequency systems. The only significant impediments to propagation are blockage by buildings, trees, and other physical objects and the remedy is obvious.

The ease of jamming the relatively weak signals emitted by satellites is often quoted as if no other radio signals are that weak. To take what might be called "official" jamming first, it is well-known that the preferred "official" method of denial of GPS service to an enemy is by local jamming to wipe out the SPS while leaving the PPS unaffected. If things are bad enough to require this then perhaps the non-military user should not be in the area in the first place, but in any case the possibility will have been well-publicised. Secondly, it has been claimed that a terrorist, or, perhaps worse, an electronic hobbyist would find it easy to jam GPS for his own purposes. The first thing to be said about this is that it would not affect professional users in the slightest - no aircraft or ship would be totally dependent on satnav as we have seen. The amateur seamen and aviators who might be relying on it rather more would not necessarily be affected very greatly either, since unless it were a major and very professional attack, they would either be out of range of a land-based jammer or would fly through the area of jamming relatively quickly. It is only local land-based users who would be really affected and for them there is no safety issue, only inconvenience. "Hobbyist" jammers only do it for some easily-visible effect, but unlike jamming communications, jamming a navigation system does not show any such effects. It would be quite easy to jam an ILS for instance, with possibly more visible and instant effects even if it is only the sight of a jumbo having to go round again, but there have not been any recorded cases so far. As regards a radio-based backup for satnav to guard against jamming, that too could very likely be easily jammed. Loran-C has been quoted in this context but in fact over much of its coverage area its signals are not at all strong and they are just as liable to malicious jamming.

10. CONCLUSIONS.

The possibility of total and instantaneous failure of a satnav system like GPS or Galileo is so remote as to be discounted. No other radio navigation system has ever had the advantage of multiply-redundant independent transmitters unaffected by power supply outages, storm damage and insurrection. Natural causes of loss of signal are rare and are guarded against by the multiple satellite sources. Such possibilities of signal loss as there are come mainly from human activities that are limited in their effect. The reliability that results from these factors is such that no ground-based radio aid could possibly approach it and it would be ridiculous to have a back-up radio system that was less reliable than the system it was supporting.

On examination, it is seen that in fact there are really very few users who actually need a new type of backup for satnav anyway, whether the satnav is GPS, Galileo or some combination that might be called a GNSS. For those for whom the loss of satnav would be important there are alternatives already available; for many others it is not a safety-of-life matter. It is hardly an effective use of resources to install a completely different system that would almost never be used. The important matter is not a back-up but putting into place the correct international arrangements to enable the genuine co-use of GPS, Galileo, and Glonass as one truly integrated GNSS system.

Journal of Global Positioning Systems (2004)
Vol. 3, No. 1-2: 322-330

Journal of Global Positioning Systems
2004
© CPGPS

Mitigating Residual Tropospheric Delay to Improve User's Network-Based Positioning

Tajul A. Musa, Jinling Wang, Chris Rizos, Young-Jin Lee

School of Surveying and Spatial Information Systems, the University of New South Wales, NSW, Australia
e-mail: Tajul.Musa@student.unsw.edu.au; Tel: +61-2-9385 4208; Fax: +61-2-9313 7493

Azhari Mohamed

Department of Survey and Mapping, Malaysia, Jalan Semarak, 50578 Kuala Lumpur, Malaysia
e-mail: azhari@jupem.gov.my; Tel: +60-3-26170971; Fax: 03-26912757

Received: 15 November 2004 / Accepted: 3 February 2005

Abstract. Existing apriori tropospheric models are not sufficiently accurate to remove tropospheric delay from GPS observations. Remaining effects of residual tropospheric delay need to be estimated to ensure high accuracy and reliability of GPS positioning. Other researchers have shown that implementations of network-based positioning techniques can adequately model the residual tropospheric delay as well as ionospheric delay and orbit biases. However, the effectiveness in removing residual tropospheric delay is highly dependent on the degree to which the wet component from the troposphere can be estimated or mitigated, an effect which shows strong variation with time and space. The aim of this paper is to illustrate the performance of an existing apriori tropospheric model and to discuss some issues concerning the estimation of the (total) tropospheric delay in the equatorial area. Finally, the network approach is applied to mitigate the effect of residual tropospheric delay. Some preliminary results from test experiments using GPS network data from an equatorial region, a location with the highest effect of tropospheric delay, are presented.

Key words: Residual tropospheric delay, zenith path delay, network-based GPS positioning

1 Introduction

The two propagation mediums which contribute to signal delay of satellite observations are the ionosphere and the neutral atmosphere. The ionosphere is a dispersive medium for microwave, i.e the refractivity depends on the frequency of the propagation signal. The ionosphere delay can be determined and eliminated (at least to first order) by making observations on both GPS frequencies. Meanwhile the neutral atmosphere delay is mainly attributed to the earth's troposphere layer. The troposphere consists of dry gases and water vapour, and is a non-dispersive medium to radio frequency. Therefore the delay effect cannot be estimated in the same way as that of the ionosphere.

The neutral atmospheric delay can be estimated by integrating the tropospheric refractivity along the GPS signal path through the atmosphere. This is referred to as the tropospheric path delay. It is possible to separate tropospheric refractivity into a hydrostatic component (or simply known as "dry") and a wet component, where the former is due to the dry atmosphere and the latter due to the presence of water vapour in the atmosphere. The (total) troposphere path delay needs to be mapped along a path of arbitrary orientation, which can be represented as the product of zenith delay and a specified mapping function. The simplest mapping function is approximated by cosec of the elevation angle. There is a difference in mapping of wet and dry components, but they differ very slightly and in practice usually they are lumped into a single mapping function. The (total) Zenith Path Delay (ZPD) can be written as:

$$ZPD = Z_{dry}\, m(\theta) + Z_{wet}\, m(\theta) \qquad (1)$$

where Z_{dry} and Z_{wet} are the zenith dry delay and zenith wet delay respectively, $m(\theta)$ is the mapping function with θ as the satellite elevation angle (for $m(\theta) \approx cosec(\theta)$). There are many troposphere models that have been developed, e.g, Saastamoinen, Hopfield, Davis, Lanyi and Chao. Most of these models effectively model the zenith dry

delay, which contributes about 80%-90% of the total delay (Hoffman-Wellenholf *et al.*, 1994). However, all the models have difficulty in modelling the wet delay due to the high spatial and temporal variability of the water vapour . As a result, a residual tropospheric delay remains in the measurements after application of the model.

Over the past few years network-based GPS positioning has been widely discussed in the literature (e.g. Wanninger, 2002; Chen *et al.*, 2000; Landau *et al.*, 2002; Rizos and Han, 2003). External information about the GPS measurement biases provided by the network technique has enabled the performance of conventional single reference station, carrier phase-based techniques to be extended over longer baselines. This is possible because the network technique attempts to model distance-dependent errors (i.e, atmospheric and orbit effects) in the local network (Han, 1997; Chen, 2001).

First section on this paper discusses the performance of two apriori tropospheric delay models. Secondly, problems of residual tropospheric delay are discussed and some issues concerning the estimation of the (total) tropospheric delay are mentioned. A review of the GPS network-based positioning approach is given in the third section. Section four describes how the network approach can be used in order to mitigate the residual tropospheric delay.

2 Testing on apriori tropospheric delay modelling

To test the performance of the apriori tropospheric delay model, tests were conducted using GPS datasets in a near-equatorial region of the earth. The data was collected by stations of the Malaysia Active GPS System (MASS) (Figure 1). The GPS double-differenced (DD) measurement model based on the ionosphere-free (IF) carrier phase combination is used to eliminate the ionospheric delay effect. The data processing methodology resolves the "wide-lane ambiguity" first and then fixes the "narrow-lane ambiguity" during subsequent processing (Rothacher and Mervart, 1996; Sun *et al.*, 1999). Therefore for long baselines, the tropospheric delay will dominate the DD IF residual errors, assuming that other errors (geometric errors and multipath) are minimised (for example by using the precise GPS orbit data, multipath-free location and precise receiver coordinates). GPS data of Day of Year (DoY) 29/03 for a 24 hours span was processed by the method described above. A Satellite elevation cut-off angle of 15º was used for the analyses. Station IPOH is excluded in the test due to bad observations.

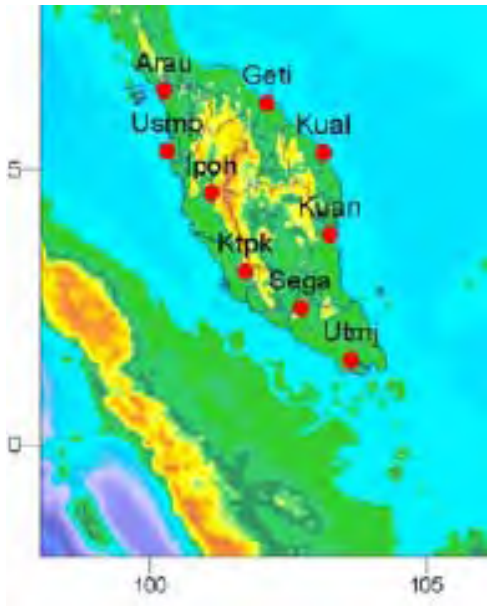


Fig. 1 Part of the MASS (Peninsular Malaysia)

Two apriori tropospheric delay models were chosen for the test: the Saastamoinen model and the Modified Hopfield model. Both models used values that are derived from a standard atmosphere model. The test methodology is as follows; Test 1: no apriori model is applied; Test 2: applying only the dry model; and Test 3: applying both the dry and wet troposphere models. Time series of the above tests are shown in Figure 2, Figure 3(a) and 3(b), and Figure 4(a) and 4(b) for a selected baseline KTPK-ARAU. Table 1 and Table 2 give details of the results.

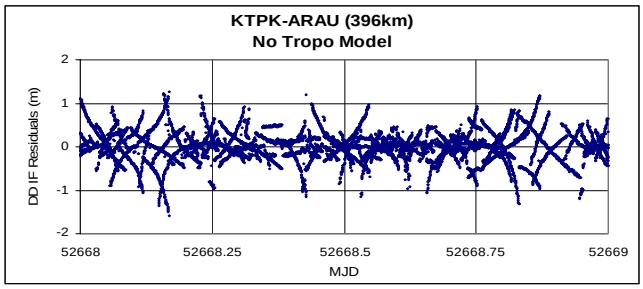


Fig. 2 Test 1 (no apriori troposphere model)

From Figure 2 and Table 1, the differential tropospheric delay can be observed as being as large as 1.5m and a RMS of up to 0.3m if no apriori troposphere model is applied. Comparing Figure 1 to Figure 3 and Figures 4, it is clear that both apriori models can mitigate the tropospheric delay, as the maximum value decreases to 0.2m and the RMS of DD IF residuals is 0.05m. This is also true for the other baselines in the tests. The test statistic in Table 1 also shows that the error increases with baseline length, which confirms that the residual (DD) tropospheric delay can be categorised as a distance-dependent error. Results in Table 2 show that a 73%-87% improvement is achieved after applying the dry model. Only a small improvement (1%-2%) is observed by

applying both the dry and wet models. In general, the DD IF residuals after applying the apriori model are between 0.03m to 0.05m, mostly due to the wet component. There is no significant difference in the test results between the two apriori troposphere models.

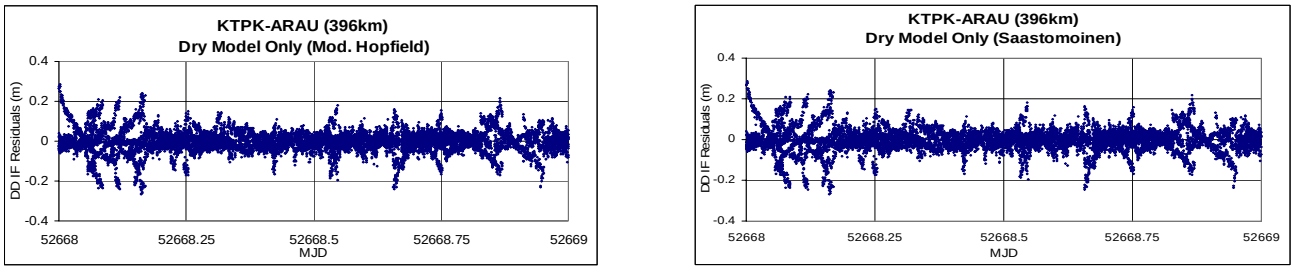


Fig. 3 Test 3(a) left, dry Modified Hopfiled model and Test 3(b) right, dry Saastamoinen model

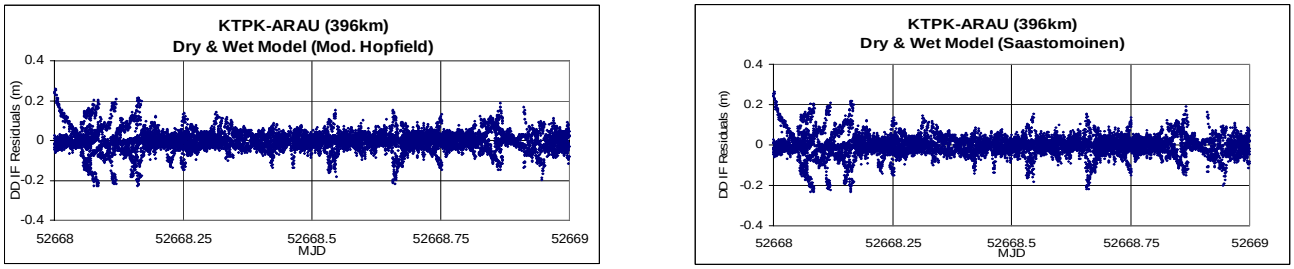


Fig. 4 Test 4(a) left, dry and wet Modified Hopfiled model. Test 4(b) right, dry and wet Saastamoinen model

Tab. 1 Statistic of Test 1, Test 2 and Test 3 of DD IF measurements.Station KTPK as reference and station height is 99m

Stn KTPK to:	Length (km)	Stn HEIGHT (m)	DD IF RMS NO MODEL (m)	DD IF RMS DRY SAAS (m)	DD IF RMS DRY M. HOPFIELD (m)	DD IF RMS DRY&WET SAAS (m)	DD IF RMS DRY&WET M.HOPFILED (m)
ARAU	396	82	0.310	0.054	0.053	0.049	0.048
GETI	341	100	0.373	0.057	0.056	0.049	0.048
USMP	288	80	0.267	0.048	0.048	0.045	0.044
KUAL	285	45	0.254	0.040	0.040	0.034	0.034
UTMJ	278	19	0.221	0.053	0.052	0.047	0.047
KUAN	196	74	0.127	0.027	0.027	0.025	0.025
SEGA	136	75	0.104	0.028	0.028	0.025	0.025

Tab. 2 Percentage improvement after applying dry model only and dry & wet model for DD IF measurements

Stn KTPK to:	DRY SAAS (%)	DRY M.HOPFILED (%)	DRY&WET SAAS (%)	DRY&WET M.HOPFIELD (%)
ARAU	82.6	82.9	84.2	84.5
GETI	84.7	85.0	86.9	87.1
USMP	82.0	82.0	83.1	83.5
KUAL	84.3	84.3	86.6	86.6
UTMJ	76.0	76.5	78.7	78.7
KUAN	78.7	78.7	80.3	80.3
SEGA	73.1	73.1	76.0	76.0

3 Issues on residuals tropospheric delay

At this stage, it is clear that the apriori troposphere model cannot effectively handle the residual tropospheric delay. High accuracy GPS positioning requires the residuals to be reduced through appropriate modelling. The approach usually is to introduce additional unknown parameters in the least square estimation process, and to, for example, solve for one scale factor for every station per session. The estimation of the scale factor tends to average the residual tropospheric delay, thus improving the results. However, the scale factor is only a constant offset to the apriori model and does not reflect the time varying nature of the atmosphere. Alternatively, a time-varying polynomial scale factor can be introduced to estimate several troposphere parameters per session. Another viable approach is to use stochastic estimation to model using a first-order Gauss-Markov or random walk process (Dodson *et al*., 1996).

To this extent, it is convenient to discuss the residual tropospheric delay in the context of the total ZPD. The estimated troposphere parameter together with the apriori model value and associated mapping function gives the GPS derived (total) ZPD. Typically the process of GPS ZPD estimation requires a large network of GPS reference stations to achieve a stable value of absolute ZPD (discussion in next section). A good example is the global network of the International GPS Service (IGS) which already is in use, publishing 2 hour absolute ZPD values. This IGS estimate should be included in the processing of regional/local GPS network data to benchmark the ZPD value derived from regional/local solution.

3.1 Absolute vs relative tropospheric delay

Relative delay is more important than absolute delay for GPS positioning. Beutler *et al.* (1988) gave a rule of thumb that relative delay causes height errors which are amplified by the factor of cosec(θ_{min}) (2.9 for θ_{min} =20°). Meanwhile an absolute delay of 10cm will cause scale biases of 0.05ppm in the estimated baseline lengths (Rothacher and Mervart, 1996). However, an accurate and absolute ZPD value is crucial for GPS meteorology applications. Equation 1 indicates that one of the important factors in total ZPD estimation is the satellite elevation angle. Duan *et al.* (1996) have shown that for small sized GPS networks, the total ZPD is sensitive to relative ZPD but not to absolute ZPD. This is due to the small elevation angle difference observed between two GPS receivers in the network. On the other hand, a large network is needed to have large elevation angle variations in order to get a better estimation of the absolute ZPD.

To analyse the relationship between absolute and relative delay and the network size, a few regional IGS stations around the local MASS network were used (Figure 5). IGS station NTUS however is treated as a local station because of the small distance to the MASS network (KTPK-NTUS is only 297km). This will give an advantage to the MASS network analysis in order to benchmark the absolute ZPD value to the IGS estimate.

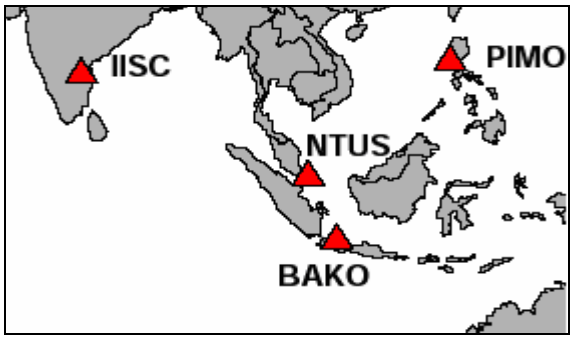


Fig. 5 Regional GPS network

Two weeks data were selected for the test, DoY204-210/03 (Jul23-Jul29 03), i.e. during a dry month, and DoY323-329/09 (Nov19-Nov25 03), i.e. during a wet month. For this analysis, the precise IGS orbits were used; satellite elevation cut-off angle was set at 10°, 15° and 20°; a simple cosec mapping function was used and the precise coordinates of all the reference stations were supplied by the network operator. Tropospheric parameters were estimated as piecewise linear functions at two hour intervals for all the stations, using the BERNESE software (Rothacher and Mervart, 1996). Only results for the case of 15° cut-off elevation angle for station NTUS is shown in Figures 6 (a) & 6(b), 7(a) & 7(b) and 8(a) & 8(b), for both weeks. Table 3 and Table 4 give the statistics of all the test results for station NTUS.

Inspecting Figures 6(a) and 6(b) and Table 3, it can be found that the absolute ZPD value (compared to the IGS value) derived from the regional network is accurate to about 3mm (in the dry season) and 5mm (in the wet season), in terms of RMS values when compared to the local network. Figures 7(a) and 7(b) show the extracted values of absolute ZPD for both networks. All tests (different 20°, 15°, 10° cut-off elevation angles) show that the differences between the regional and local absolute ZPD are within 1-3mm (in the dry season) and 5-8mm (in the wet season). The higher elevation angle observed from regional network can provide a better estimation of absolute ZPD. Both local and regional absolute ZPD estimates differ by about 18mm-32mm in their RMS to the IGS values, where the maximum difference occurs during the wet season for the 10° cut-off elevation angle.

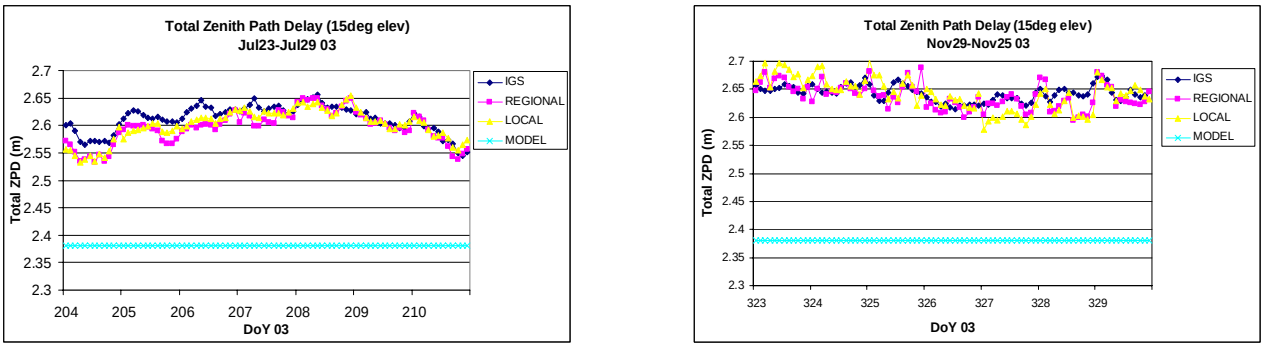


Fig. 6 6(a) left, dry season and 6(b) right, wet season of total ZPD for station NTUS derived from different network size. IGS value is obtained from combined ZPD solution published by IGS. Saastamoinen apriori model ZPD value is used (derived from standard atmosphere value)

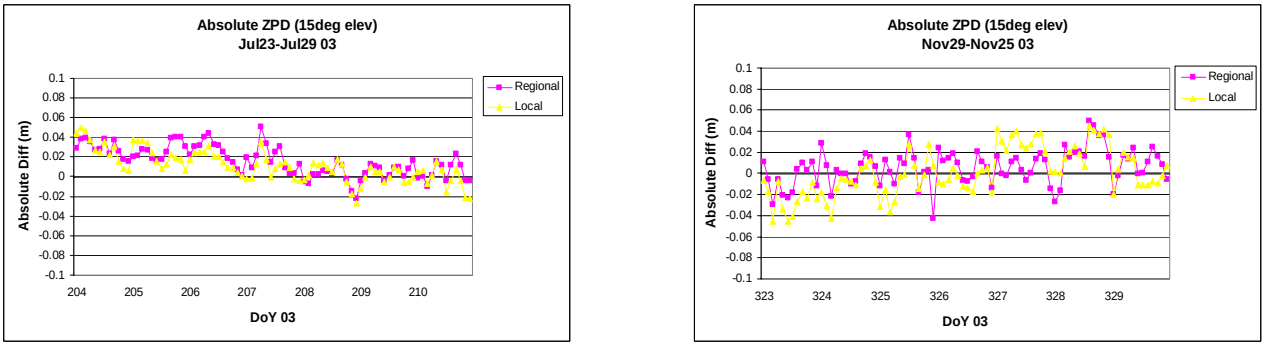


Fig. 7 7(a) left, dry season and 7(b) right, wet season of absolute total ZPD difference for station NTUS to absolute IGS value using different network size

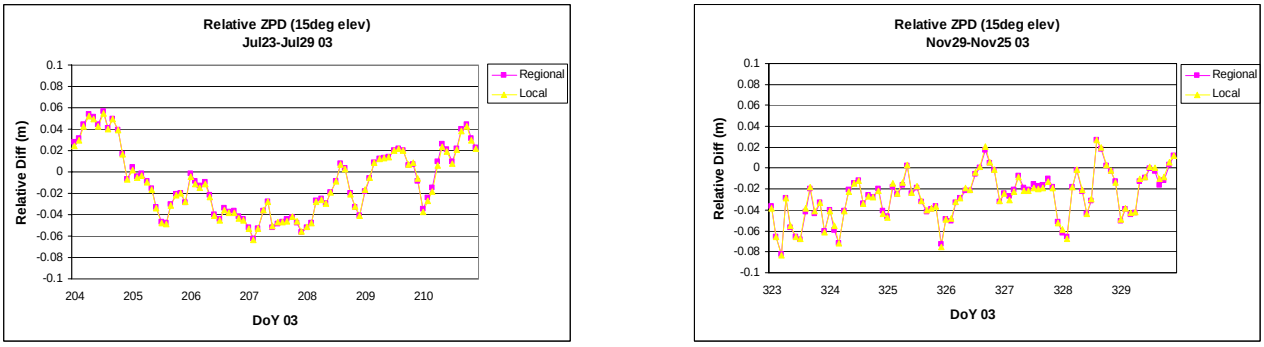


Fig. 8 8(a) left, dry season and 8(b) right, wet season of relative total ZPD difference for station NTUS. Station KTPK taken as reference

Tab. 3 Statistic of absolute ZPD difference (to value published by IGS) for station NTUS

		Dry Season (Jul23-Jul29 03)			Wet Season (Nov19-Nov25 03)		
Elevation	Network	Mean (m)	Stdv (m)	RMS (m)	Mean (m)	Stdv (m)	RMS (m)
20º	REGIONAL	0.005	0.024	0.024	-0.013	0.020	0.023
	LOCAL	-0.001	0.024	0.024	-0.016	0.023	0.028
15º	REGIONAL	0.016	0.015	0.019	0.006	0.017	0.018
	LOCAL	0.011	0.016	0.022	0.001	0.023	0.023
10º	REGIONAL	0.014	0.016	0.021	0.019	0.014	0.024
	LOCAL	0.017	0.015	0.022	0.027	0.017	0.032

Tab. 4 Statistic of relative ZPD difference for station NTUS, KTPK as reference station

		Dry Season (Jul23-Jul29 03)			Wet Season (Nov19-Nov25 03)		
Elevation	Network	Mean (m)	Stdv (m)	RMS (m)	Mean (m)	Stdv (m)	RMS (m)
20º	REGIONAL	-0.004	0.041	0.041	-0.024	0.023	0.033
	LOCAL	-0.006	0.041	0.041	-0.026	0.024	0.035
15º	REGIONAL	-0.009	0.032	0.033	-0.027	0.023	0.035
	LOCAL	-0.010	0.032	0.033	-0.027	0.023	0.035
10º	REGIONAL	-0.007	0.028	0.029	-0.021	0.019	0.029
	LOCAL	-0.007	0.027	0.028	-0.021	0.019	0.028

Comparing Figures 8(a) and 8(b), there is almost no difference seen for the relative ZPD value estimation between both networks (also true in the dry and wet seasons). Further confirmation is found by inspecting Table 4 for the rest of the tests. Results in Table 4 also show that the RMS of the relative delay after applying the apriori model is between 0.03m to 0.05m, which agrees with the result in Table 1. Thus, the delay needs to be estimated and removed from the measurements to ensure high accuracy positioning, especially in the context of ambiguity resolution.

4 Review of network-based positioning technique

Based on the Linear Combination Method (LCM) (Han and Rizos, 1996), the single-differenced functional model for the virtual measurements with *n* reference stations can be written as:

$$\sum_{i=1}^{n} \alpha_i \Delta\phi_i = \Delta\phi_{u,m} - [\alpha_1 \Delta\phi_{1,m} + ... + \alpha_{n-1} \Delta\phi_{n-1,m}] \qquad (2)$$

where ϕ is the carrier phase observation, α_i is the weight for the i reference station determined to be inversely proportional to the distance from i reference stations to the user station *u*, *m* is the master reference station and Δ is the single-differenced operator. The second term on the right hand side of Equation (2) is the network correction for the single-difference.

The DD functional model for the virtual measurements can be derived from Equation (3) as:

$$\Delta\nabla\phi_{u,m} - [\alpha_1 V_{1,m} + ... + \alpha_{n-1} V_{n-1,m}] = \Delta\nabla p_{u,m} + \lambda\Delta\nabla N_{u,m} + \varepsilon_{\sum_{i=1}^{n} \alpha_i \Delta\nabla\phi_i} \qquad (3)$$

where *V* is defined as the DD residual vectors from the master station (*m*) to the other reference stations after the ambiguities (*N*) have been resolved, $[\alpha_1 V_{1,m} + ... + \alpha_{n-1} V_{n-1,m}]$ is the DD network corrections term, $\varepsilon_{\sum_{i=1}^{n} \alpha_i \Delta\nabla\phi_i}$ is the DD linear combination carrier phase observation noise, λ is the wavelength of the carrier wave, *p* is the satellite position vector minus the station position vector, and $\Delta\nabla$ is the DD operator. The virtual measurement ambiguity then should be fixed to its integer value.

In general, the network processing can be summarised in four major steps (Tajul *et al.*, 2003):

⇒ Processing Master-Reference Stations – to get fixed residuals of master station to other reference stations after fixing the network ambiguities.

⇒ Calculation of Network Corrections – the network corrections were calculated through Linear Combination Method (LCM), i.e by applying linear interpolation techniques to the fixed residual vectors.

⇒ Generating the so-called "Virtual Measurements" – the network corrections were applied to master-user measurements, epoch-by-epoch and satellite-by-satellite basis to form a new set of measurement (the "virtual measurements").

⇒ Fixing the ambiguities – from master to user station

This processing can be implemented in either the post-processing or real-time modes.

5 Network approach to mitigate residuals tropospheric delay

One of the reasonable assumptions used in the network technique described above is that the residual tropospheric delay (i.e after applying the apriori model) should be mitigated to some extent through the application of the LCM (Han, 1997). To this point, it is not clear how good the network technique will mitigate the residual tropospheric error. The reason is because the

network corrections provided by the network are lumped together with other distance-dependence errors, mostly dominated by the ionosphere. The performance of the network technique to account for the residual tropospheric delay can be studied using the DD IF measurements explained in section 2, to replace Equation (3). This technique was successfully applied by Zhang (1999) in his study using the NetAdjust method. The purpose is to generate only residual tropospheric delay corrections from the network stations, and it should be applied to the user's station in order to asses the performance of this technique.

For this study part of the MASS network, stations ARAU, KUAL, KUAN, KTPK, SEGA and NTUS (IGS station), were selected (Figure 1). The reason for selecting only these stations is to avoid the computational burden in generating the network corrections, however the design still gives good coverage over the study area. For this selected network design there are five (n = 5) reference stations - KTPK is selected as a master station and SEGA as the user station because of it location inside the network (KTPK-SEGA is 136km). All the measurements are handled in post-processing (static) mode, the precise IGS orbits are used and the satellite elevation cut-off angle was set to 10°. The procedure used for this network processing strategy was:

- Generate n-1 DD IF residual vectors from the network measurements using the methodology described in section 2.
- Calculate the residual tropospheric delay corrections from the network based on the LCM.
- Apply the corrections to the DD IF measurements at the user site.

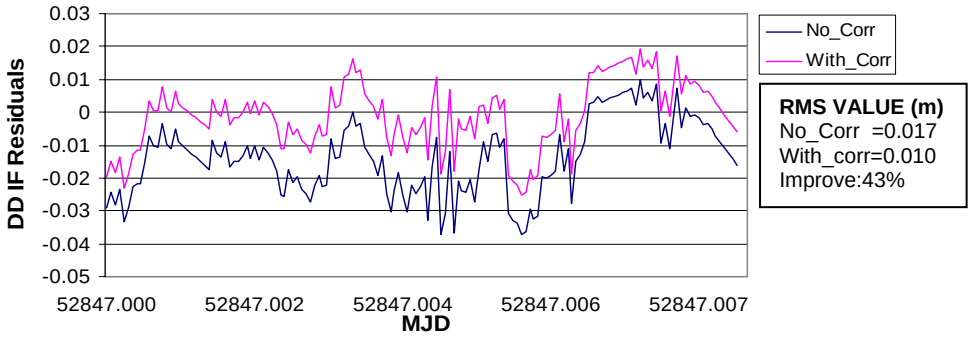


Fig. 10 DD IF residuals (m) of Satellite pair PRN26-05

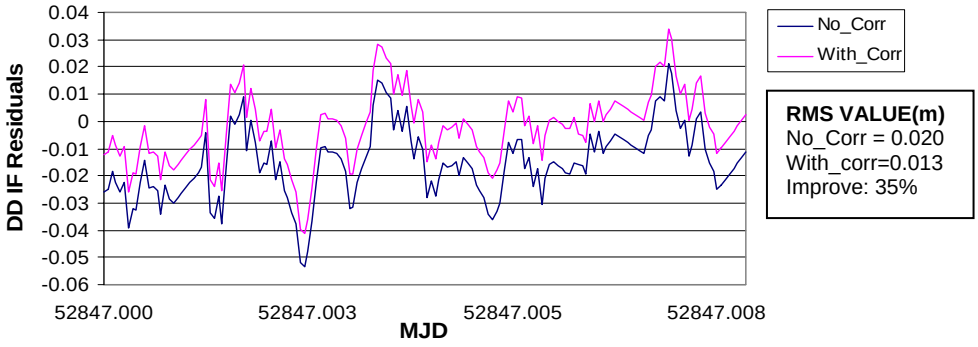


Fig. 11 DD IF residuals (m) of Satellite pair PRN26-30

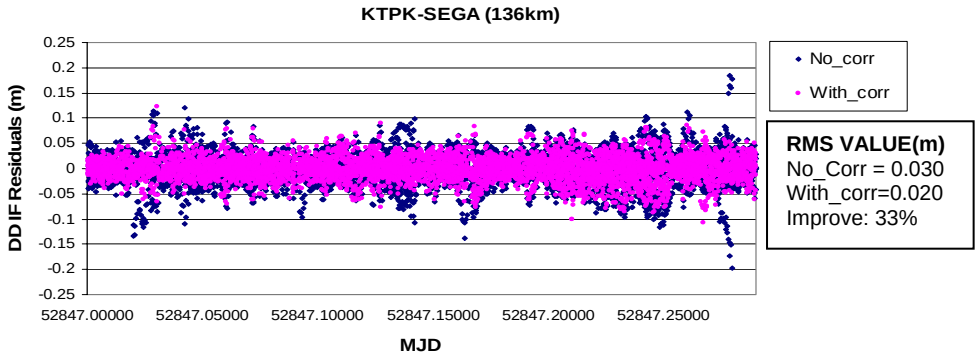


Fig. 12 DD IF residual of all satellites pairs

Figure 10 and Figure 11 show the DD IF residuals and statistics for two satellite pairs PRN26-05 and PRN26-30. Figure 12 shows all combinations before and after applying the correction. An improvement of about 33%-43% in the RMS value is found for both pairs, 33% in the case of all combinations after applying the corrections, confirming the effectiveness of the network technique in mitigating the residual tropospheric delay.

6 Concluding remarks

Results from this study using data from the MASS network and the regional IGS stations show that:

Apriori tropospheric models effectively removed the dry delay of the tropospheric delay by up to 73%-87%. Small improvement (1%-2%) is achieved after applying the wet model, indicating the difficulty in modelling the wet component (mostly due to high variations of water vapour in this region).

Accuracy of absolute ZPD value estimation (compared to the IGS values) using the regional network is found to be better than 3mm (dry season) and 8mm (wet season) compared to local network estimation. Meanwhile, almost no difference is found for the relative ZPD value estimation for both networks.

Residual tropospheric delay can be mitigated in user's location using the network approach, where improvements of up to 33% have been achieved.

Acknowledgements: We gratefully thank Department of Survey and Mapping Malaysia (DSMM) for providing us with the data used in this study.

References

Beutler, G, Bauersima I, Gurtner W, Rotacher M, Schildknecht T, and Geiger A (1988): ***Atmospheric refraction and other important biases in GPS carrier phase observation.*** In: Atmospheric Effects on Geodetic Space Measurements, Monograph 12, School of Surveying, UNSW, Australia, 15-43.

Chen X, Han S, Rizos C, and Goh PC (2000): ***Improving real-time positioning efficiency using the Singapore Integrated Multiple Reference Station Network (SIMRSN).*** Proceedings of 13th Int. Tech. Meeting of the Satellite Division of the Insitute of Navigation, Salt Lake City, Utah, USA 19-22 September, 9-18.

Chen HY (2001): ***A study on real-time medium-range carrier phase-based GPS multiple reference stations.*** Ph.D. dissertation. School of Surveying and Spatial Information Systems, The University of New South Wales, Sydney, Australia, 43-77.

Dodson AH, Shardlow PJ, Hubbard LCM, Elegered G, and Jarlemark POJ (1996) ***Wet tropospheric effects on precise relative GPS height determination.*** Journal of Geodesy, 70(4): 188-202.

Duan J, Bevis M, Fang P, Bock Y, Chiswell S, Businger S, Rocken C, Solheim F, Hove TV, Ware R, McClusky, Herring TA, and King RW (1996): ***GPS meteorology: direct estimation of the absolute value of precipitbale water.*** Journal of Applied Meterology, 35(6), 830-838.

Hoffmann-Wellenhof B, Licthenegger H, and Collins J (1994): ***Global Positioning System: Theory and Practice***. 4th edition, Springer, Berlin, Germany, 98-106.

Han S (1997): ***Carrier phase-based long-range GPS kinematic positioning.*** Ph.D. dissertation. School of Surveying and Spatial Information Systems, The University of New South Wales, Sydney, Australia, 72-114.

Han S, and Rizos C (1996): ***GPS network design and error mitigation for real-time continuous array monitoring systems.*** Proceedings of 9th Int. Tech. Meeting of the Satellite Division of the Insitute of Navigation, Kansas City, Missouri, USA, 17-20 September, 1827-1836.

Landau H, Vollath U, and Chen X (2002): ***Virtual Reference Station Systems.*** Journal of Global Positioning. 1(2):137-143.

Musa T, Wang J, and Rizos C (2004): ***A stochastic modelling method for network-based GPS positioning.*** Proceeding

of European GNSS Conference 2004, Rotterdam, The Netherlands, 16-19 May, CD-Rom proc.

Rothacher M, and Mervart L (1996): ***Manual of Bernese GPS Software.*** Version 4.0, Astronomical Institute, University of Berne, 159-188.

Rizos C, and Han S (2003): ***Reference station network based RTK systems - Concepts & progress.*** Wuhan University Journal of Nature Sciences, 8(2B): 566-574.

Sun H, Melgard T, and Cannon ME (1999): ***Real-time GPS Reference Network Carrier Phase Ambiguity Resolution.*** Proceedings of the ION National Technical Meeting, San Diego, California, 25-27 January, 193-199.

Wanninger L (2002): ***Virtual reference stations for centimeter-level kinematic positioning.*** Proceedings of 15th Int. Tech. Meeting Satellite Division of the Institute of Navigation, Portland, Oregon, USA, 24-27 September, 1400-1407.

Zhang J (1999): ***Investigation into the estimation of residuals tropospheric delay in a GPS network.*** Ph.D. dissertation. Department of Gematics Engineering, University of Calgary, Alberta, Canada, 53-125.