Application of Geostatistical Models for Estimating Spatial Variability of Rock Depth

Pijush Samui¹, Thallak G. Sitharam²

¹Centre for Disaster Mitigation and Management, VIT University, Vellore, India
²Department of Civil Engineering, Indian Institute of Science, Bangalore, India

e-mail: pijush.phd@gmail.com, sitharam@civil.iisc.ernet.in

Received April 30, 2011; revised May 20, 2011; accepted June 10, 2011

Abstract

Rock depth information of a site is a significant factor for geotechnical engineering and earthquake ground response analysis. In this paper, reduced level of rock at Bangalore is arrived from the 652 boreholes in the area covering 220 km². Geostatistical modeling based on kriging (simple and ordinary) techniques has been applied for estimating reduced level of hard rock in Bangalore. The models are used to compute variance of estimated reduced level of the rock. A new type of cross-validation analysis proves the robustness of the developed models. The comparison between the simple and ordinary kriging model demonstrates that the ordinary kriging model is superior to simple kriging model in predicting reduced level of rock in the subsurface of Bangalore.

Keywords: Geostatistics, Semivariogram, Kriging

1. Introduction

Rock depth for a site is a very useful parameter to the geotechnical and earthquake engineers to find their basic requirement of hard strata and ground motion at rock level. In most geotechnical investigations, knowledge of the hard strata or hard rock depth is essential to design a suitable foundation. In ground response analysis, Peak Ground Acceleration (PGA) at rock level and response spectrum for the particular site is evaluated. This information is also essential to evaluate liquefaction hazards of a site and to estimate earthquake induced forces on structures.

With an objective of evaluating hard rock depth in Bangalore, an attempt has been made to develop a two dimensional map of reduced level of rock for Bangalore based on kriging (simple and ordinary) technique. Rock is identified by borelogs data available in the area and identified by visual observation of the cores taken at these locations. Hard disintegrated rock is also identified as rock and depth from the ground level has been used to evaluate reduced level of rock at any location.

The kriging method was developed during the 1960s and 1970s and has been acknowledged as a good spatial interpolator [1-3]. The most important features of this method are that the interpolator 1) is linear and unbiased 2) gives minimum estimation error, and 3) is exact and gives an evaluation of uncertainty for interpolated values. This technique is widely used in the field of earth sciences, including mining, geochemistry, remote sensing, etc. A major advantage of kriging is that it is more flexible than other interpolation methods such as inverse-distance weighting, deterministic splines and Thiessen polygons. The weights are not selected arbitrarily, but depend on how the variable of interest (in this case hard rock elevation) varies in space. In kriging, the variable weights have been used based on the scale of variability whereas in Thiessen polygon, one has to apply same weights, whether the function exhibits small-or large-scale variability. The main goal of kriging is to predict (in the interpolation sense) a variable where no measurements were made using the semivariogram as a model characterizing spatial variability. Semivariogram is the analytical tool used to evaluate and quantify the degree of spatial autocorrelation. It contains three elements of information. First of all the semivariogram is an appreciation of the dispersion (or scatter) of the parameters, which equates to the half variance. Secondly, it gives an autocorrelation distance that represents the radius of influence of a measurement made at a given point. Thirdly, it provides the type of variability that indicates how values fluctuate in space.
A valid estimation of the variance of estimated reduced level of rock depths is important for developed ordinary as well as simple kriging model. So, the developed models have been used to compute variance of estimated reduced level of the rock depths. In general, the variance depends both on the semivariogram and the location of the measurements. A number of publications can be found in the literature which presents the theory and application of kriging [4-12].

In this paper, a semivariogram model has been developed along with the kriging model for the reduced level of the rock in the subsurface of Bangalore. Semivariogram analysis is used to detect trends in structure or internal properties of deposits and estimated values can be obtained at points where no data is available. A new method for cross-validation analysis of developed models has been proposed. The cross-validation of the model has been done based on the examination of residuals. A comparative study has been also done between developed simple kriging and ordinary kriging.

2. Subsurface of Bangalore and GIS Model Development

Bangalore covers an area of over 220 square kilometres and ground Reduced levels (GRL) also vary a lot in the city. It varies from 810 m in north-east part to 940 m in south-western part of Bangalore. Ground reduced levels do not vary much in the central and north-western parts of the city. The population of greater Bangalore exceeds 6 million and it is the fifth biggest city in India. It is situated on latitude of 1208' North and longitude of 77037' East. From geology, subsurface of Bangalore region is a Gneiss complex, which was formed by several tectonic-thermal events with large influx of sialic material, occurring between 3 to 3.4 billion years ago giving rise to an extensive group of gray gneisses designated as the “older gneiss complex”. These gneisses act as the basement for a widespread belt of schist’s. The younger group of gneissic rocks mostly of granodioritic and granitia composition is found in the eastern part of the state, representing remobilized parts of an older crust with abundant additions of newer granite material, for which the name “younger gneiss complex” has been given [13]. The soil is mostly a residual soil from granite gneiss due to weathering action. In the most recent past, there were more than 400 lakes, and more than 340 lakes were dried up and have been encroached for residential and industrial development. In the old tank beds, silty sand/clay is also found as overburden.

A Geographic Information System (GIS) model (see Figure 1) of Bangalore with several layers on a scale of 1:20,000 has been developed with a purpose of carrying out microzonation of Bangalore. The Bangalore map forms the base layer for GIS. The map entities have been developed for locating boreholes to the utmost accuracy and at each location borelogs have been attached along with geotechnical data of each layer up to the hard rock. The digitized map has several layers of information. Some of the important layers considered are the boundaries (outer and Administrative), Highways, Major roads, Minor roads, Streets, Rail roads, Water bodies, Drains, Ground Contours and Borehole locations. The locations of boreholes are shown in Figure 1 along with ground reduced level with an interval of 10 m (see Figure 2). Distribution of collected boreholes in Bangalore is shown in Figure 3, indicating a very good distribution of the boreholes in each quadrant of Bangalore from the city center. Figure 1 also depicts grids of 1kmX1km along with the corporate boundary of Bangalore and outer boundary circumscribing the ring road. Figure 1 gives a clear view of the spatial distribution of boreholes in Bangalore region. An average of about three boreholes data is available within the grid of 1 km × 1 km.

![Figure 1. Borehole location in Bangalore Map (scale: 1:20000).](image1)

![Figure 2. GIS model of borehole locations with respect to contours.](image2)
3. Methodology

In this paper, ordinary kriging and simple kriging are adopted for evaluating reduced level of the rock in subsurface of the Bangalore. For both the methods, there is a need to introduce the terminology of the covariance function or semivariogram. The covariance function between two points is defined as:

\[ C(h) = E[(d(x) - m)(d(x') - m)] \tag{1} \]

where \( m \) is the mean of \( d(x) \) and \( C(h) \) is the covariance function with lag \( h \), with \( h \) being the distance between two samples \( x \) and \( x' \):

\[ h = \|x - x'\| = \sqrt{(x - x')^2 + (y - y')^2} \tag{2} \]

The semivariogram for an intrinsic random function [14,15] is defined as:

\[ \gamma(h) = 0.5 * E[(d(x) - d(x'))^2] \tag{3} \]

Figure 4 illustrates the different components of the semivariogram.

If the variance exists (the random function is second order stationary), the relation between the covariance function and the semivariogram is as follows:

\[ \gamma(h) = C(0) - C(h) \tag{4} \]

where, \( \gamma(h) \) is the semivariogram and \( C(0) \) is the variance.

For semivariogram, the model used in this analysis is Gaussian model. The general equation for this model looks like:

\[ \gamma(h) = c_0 + c \left( \frac{3h}{2a} - \frac{1h^3}{2a^3} \right) \text{ for } h \leq a \tag{5} \]

\[ \gamma(h) = c_0 + c \text{ for } h \leq a \tag{6} \]

![Figure 4. A typical semivariogram.](image)
For semivariogram model, geoanisotropic model has been used to reduce the anisotropy into isotropy by a linear transformation of coordinates [16]. Once the model of semivariogram is constructed, the weights are computed for kriging. The method for ordinary kriging, simple kriging and cross-validation of the models are given as below:

3.1. Ordinary Kriging

Ordinary kriging is a linear geostatistical method. It gives local estimation by interpolation. The basic equation used in ordinary kriging is as follows:

\[ d(x,y) = \sum_{i=1}^{n} w_i d_i \]  

(7)

where \( n \) is the number of scatter points in the subsurface of the Bangalore, \( d_i \) is reduced level of the rock at point \( i \) in the subsurface of the Bangalore and \( w_i \) is weights assigned for point \( i \). This equation is essentially the same as the equation used for inverse distance weighted interpolation except that rather than using weights based on an arbitrary function of distance. The weights used in kriging are based on the model semivariogram. For example, to interpolate reduced level of the rock (\( d_p \)), at a point \( 'P' \) in the subsurface of the Bangalore based on the surrounding points \( P_1, P_2 \), and \( P_3 \), the weights \( w_1, w_2, \) and \( w_3 \) must be found. The weights are found through the solution of the simultaneous equations:

\[ w_1 \gamma(h_{11}) + w_2 \gamma(h_{12}) + w_3 \gamma(h_{13}) = \gamma(h_{pp}) \]  

(8)

\[ w_1 \gamma(h_{21}) + w_2 \gamma(h_{22}) + w_3 \gamma(h_{23}) = \gamma(h_{pp}) \]  

(9)

\[ w_1 \gamma(h_{31}) + w_2 \gamma(h_{32}) + w_3 \gamma(h_{33}) = \gamma(h_{pp}) \]  

(10)

where \( \gamma(h_{ij}) \) is the model semivariogram evaluated at a distance equal to the distance between points \( i \) and \( j \) in the subsurface of the Bangalore. For example \( \gamma(h_{ip}) \) is the model semivariogram evaluated at a distance equal to the separation of points \( P_1 \) and \( P \). Since it is necessary that the weights sum to unity, a fourth equation is added.

\[ w_1 + w_2 + w_3 = 1.0 \]  

(11)

Since there are now four equations and three unknowns, a slack variable, is added to the equation set. The final set of equations is as follows:

\[ w_1 \gamma(h_{11}) + w_2 \gamma(h_{12}) + w_3 \gamma(h_{13}) + \gamma = \gamma(h_{pp}) \]  

(12)

\[ w_1 \gamma(h_{21}) + w_2 \gamma(h_{22}) + w_3 \gamma(h_{23}) + \gamma = \gamma(h_{pp}) \]  

(13)

\[ w_1 \gamma(h_{31}) + w_2 \gamma(h_{32}) + w_3 \gamma(h_{33}) + \gamma = \gamma(h_{pp}) \]  

(14)

\[ w_1 + w_2 + w_3 = 1.0 \]  

(15)

The equations are then solved for the weights \( w_1, w_2, \) and \( w_3 \). The \( d_p \) value of the point \( p \) is then calculated as:

\[ d_p = w_1 d_1 + w_2 d_2 + w_3 d_3 \]  

(16)

where, \( d_1, d_2 \) and \( d_3 \) are reduced level of the rock at point \( P_1, P_2 \) and \( P_3 \) respectively. The variance can be calculated at each interpolation point as:

\[ S^2_i = w_1 \gamma(h_{p1}) + w_2 \gamma(h_{p2}) + w_3 \gamma(h_{p3}) + \gamma \]  

(17)

when interpolating to an object using the kriging method, variance data set is always produced along with the interpolated data set. In some cases, specific spatial data distributions give rise to negative kriging weights, causing interpolated values to be negative or out of data limits and physically not compatible with data. For this reason negative and particular positive weights are set to 0 according to the rules proposed by Deutsh [17], ensuring the sum of weights is equal to one.

3.2. Simple Kriging

Simple kriging uses the average of the entire data set. The basic equation used in simple kriging is as follows:

\[ d_s (x,y) = m + \sum_{i=1}^{n} \lambda_i [d_i - m] \]  

(18)

where \( \lambda_i \) is weight assigned for point \( i \) and \( m \) is the average value.

The weights \( \lambda_i \) are found through the solution of the simultaneous equations:

\[ \gamma_1 [C(0) - \gamma(h_{11})] + \gamma_2 [C(0) - \gamma(h_{12})] \]

\[ \gamma_3 [C(0) - \gamma(h_{13})] = C(0) - \gamma(h_{pp}) \]  

(19)

\[ \gamma_1 [C(0) - \gamma(h_{21})] + \gamma_2 [C(0) - \gamma(h_{22})] + \gamma_3 [C(0) - \gamma(h_{23})] = C(0) - \gamma(h_{pp}) \]  

(20)

\[ \gamma_1 [C(0) - \gamma(h_{31})] + \gamma_2 [C(0) - \gamma(h_{32})] + \gamma_3 [C(0) - \gamma(h_{33})] = C(0) - \gamma(h_{pp}) \]  

(21)

The variance of estimation then become equal to

\[ S^2_{id} = \left(1 - \sum_{i=1}^{n} \lambda_i \right) C(0) + \sum_{i=1}^{n} \lambda_i \gamma(h_{ip}) \]  

(22)

3.3. Cross-Validation of the Models

A new type of cross-validation analysis for kriging has been presented in this study. In practice, cross-validation is based on statistical tests involving the residuals. Residuals are differences between observation and model predictions. The detailed description of residuals in the case of kriging is given by Kitanidis [16]. It has been
The probability density function (pdf) of \( x \) and this normalized value can be written as, respectively:

\[
\varepsilon_k = \frac{\delta_k}{\sigma_k}, \quad \text{for } k = 2, \ldots, n
\]

A cross-validation \( Q_1 \) and \( Q_2 \) are used to check the statistical distribution of the residuals between the observed data and kriged values at the original observation location by using the same kriging parameters and semivariogram model parameters. To perform \( Q_1 \) and \( Q_2 \) cross validation, a normalized residual array \( (\varepsilon_k) \) is constructed as suggested by Kitanidis (1997). \( Q_1 \) is the mean of the residual \( (\varepsilon_k) \) and it is written as:

\[
Q_1 = \frac{1}{n-1} \sum_{k=2}^{n} \varepsilon_k
\]

Under the null hypothesis, \( Q_1 \) is normally distributed with mean 0 and variance \( \frac{1}{n-1} \). The probability density function (pdf) of \( Q_1 \) is:

\[
f(Q_1) = \frac{1}{\sqrt{2\pi(n-1)}} \exp \left( -\frac{Q_1^2}{2(n-1)} \right)
\]

where, \( n \) is the number of data points. If the experimental value of \( Q_1 \) turns out to be acceptable close to zero then this test gives no reason to question the validity of the model. \( Q_2 \) is the variance of \( \varepsilon_k \) and it is written as:

\[
Q_2 = \frac{1}{n-1} \sum_{k=2}^{n} \varepsilon_k^2
\]

(\( Q_2 \))*\((n-1) \) approximately follows the chi-square distribution with parameter \((n-1)\). Where, \( n \) is the number of data points. The mean and variance of \( Q_2 \) are 1 and \( \frac{2}{n-1} \) respectively. The pdf of \( Q_2 \) is given by the following equation:

\[
f(Q_2) = \frac{(n-1)^{n-3}}{2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right)} Q_2^{\frac{n-3}{2}} \exp\left(-\frac{(n-1)Q_2}{2}\right)
\]

where, \( \Gamma \) is the gamma function. For robust model, the experimental value of \( Q_2 \) should be close to one. In this work, kriging model and cross-validation have been programmed using MATLAB software.

4. Results and Discussion

In case of ordinary kriging, the semivariogram of reduced level of rock obtained from the experimental values is shown in Figure 5. The Gaussian model has been plotted in Figure 5 and gives a reasonable fit to the values obtained. The range, sill and nugget of the semivariogram are 0.95, 1.202 and 0.097 respectively. In the semivariogram, on the x-axis “relative to the full length scale” means normalized lag distance. The estimation of reduced level of rock has been done by using developed model of semivariogram (shown in Figure 5). Figure 6 shows the kriging surface of the reduced level of rock for ordinary kriging. For simple kriging analysis, the experimental semivariogram has been calculated using the reduced level of rock data. A Gaussian model has been fitted with parameters: 0.95 for range, 1.178 for sill and 0.079 for nugget. By using the model of the semivariogram shown in Figure 7, reduced level of rock has been estimated in the subsurface of the Bangalore. The result is shown in Figure 8.

One of the most important finding of this study is that the semivariogram for both models is free from white noise or a pure nugget effect. The pure nugget effect corresponds to the total absence of auto-correlation. A weighted nonlinear least squares method has been used to fit semivariogram model. The points closer to the origin are given higher weights than points further away, because they are inherently more accurate, as they are calculated using more data pairs. Gaussian model, which has been used for semivariogram, ensures well-conditioned

Figure 5. Semivariogram model using ordinary kriging.
kriging matrix. So this study did not exhibit any numerical stability problem. For both models, the semivariogram stops increasing beyond a certain distance. This semivariogram is called “transition” models, and corresponds to a random function which is not only intrinsic but also stationary. A function is stationary if it consists of small scale fluctuations about some well-defined mean value. For a stationary function, the length scale at which the sill is obtained describes the scale at which two measurements of the variable become practically uncorrelated. The advantage of intrinsic model is that it has been used to summarize incomplete information and patterns in noisy data. It has been also used to interpolate unknown data from observation of data. The semivariogram has a sill, which indicates that having extreme values has a very low probability. The both kriging models gave a unique solution. The reason for unique solution is explained below.

Considering a case of two different measurements of reduced level of rock obtained at the same location, if the semivariogram for both models is continuous (with zero nugget) the function \( d(x) \) will be continuous. This means that one of the two measurements is redundant. Thus, one must be discarded; otherwise, a unique solution cannot be obtained because the determinant of the matrix coefficients of the kriging system vanishes. This problem is solved by adding a nugget term to the semivariogram. As a result, the Gaussian model adopted shows the nugget effect. In practical sense, nugget effect gives the kriging equations a stability and robustness. Without a nugget effect, inverting the kriging matrices may lead to computational round-off errors. Nugget effect also confirms that the contour map of estimate has a discontinuity at each observation point. As the sampling distance decreases, it is possible to obtain a better estimate of nugget effect. But the cost of the exploration program increases enormously. The model for the experimental semivariogram has been chosen based on the examination of residuals (differences between observation and model predictions). The predicted values from simple and ordinary kriging are different because of their different properties. As a result, the residuals from simple and ordinary kriging are different. For this reason, the semivariograms are different for simple and ordinary kriging. In this study, the residuals are always uncorrelated. The lack of correlation in the residuals has been explained below:

If the residuals are correlated, one can use this correlation to predict the value of \( \varepsilon_i \) from the value of \( \varepsilon_{i+1} \ldots \varepsilon_{i-1} \) using a linear estimator. So, one can reduce further the mean square error of estimation of \( d(x) \). But this is impossible because \( d(x) \) is already the minimum-variance estimate. Thus, the residuals must be uncorrelated.

Kriging maps (Figures 6 and 8) provide a qualitative difference between the ordinary kriging and simple kriging methods. The advantage of this study is that it provides the magnitude of the variance of the estimated reduced level of rock. Figure 9 is the variance map of reduced level of rock generated by ordinary kriging. Figure 10 is a variance map of reduced level of rock generated by simple kriging. Using kriging variance map (Figures 9 and 10), one can give an indication of the quality of estimate. From Figures 9 and 10, it has been seen that the variance increases with increasing distance between estimated points and the actual point. The overall pattern of Figures 9 and 10 give an indication of where in the field adequate or inadequate sampling occurred. In the Figures 9 and 10, it is clear that the variance of the estimated data from simple kriging analysis is...
always greater than the variance of the estimated data from ordinary kriging analysis. Prediction of variance also depends on the behavior of the semivariogram at the origin, and it is known that without a nugget effect the predicted variance is often underestimated.

In case of cross-validation of kriging model, the acceptable region is defined in the Figures 11 to 14 (between the two vertical lines). For a good model, the $Q_1$ as well as $Q_2$ must fall in this acceptable regions as shown in Figures 11-14. In case of ordinary kriging, the value of $Q_1$ and $Q_2$ is 0.002, 1.069 respectively. The $Q_1$ and $Q_2$ values are well within the acceptable region (shown in Figures 11 and 12). In case of simple kriging, the value of $Q_1$ and $Q_2$ is 0.01, 0.911 respectively. The $Q_1$ and $Q_2$ fall in the acceptable region (shown in Figures 13 and 14). For both the models, the value of $Q_1$ and $Q_2$ are close to 0 and 1 respectively. The cross-validation indicates that the developed ordinary kriging as well as simple kriging models are robust models for the estimation of the reduced level of rock in the subsurface of Bangalore. However it is clear from the results that the ordinary kriging seems to be predicting better than simple kriging.

In order to compare between the ordinary and simple kriging models, five points have been chosen randomly from known reduced level of the rock of 652 points in
the subsurface model of Bangalore. The predicted values of these points are shown in Table 1. It can be seen from the table that the ordinary kriging model has given better prediction than simple kriging model. For the data sets used in this paper, ordinary kriging has shown to be a better estimator than simple kriging in terms of reduced kriging variance and the comparison between an estimated and actual value. This result is expected, since the simple kriging uses the average of the entire data set while ordinary kriging uses a local average (the average of the scatter points in the kriging subset for a particular interpolation point). The reduced level of rock at a half-space point could be more accurately estimated from the reduced level of rock at neighboring half-space points than that at distinct location. As a result, simple kriging is less accurate than ordinary kriging.

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the file for “MSW US ltr format”.

Table 1. Comparison between ordinary kriging and simple kriging model.

<table>
<thead>
<tr>
<th>Bore hole No.</th>
<th>Longitude (degree)</th>
<th>Latitude (degree)</th>
<th>Actual reduced level of the rock (m)</th>
<th>Predicted value(m) by simple kriging</th>
<th>Predicted value(m) by ordinary kriging</th>
</tr>
</thead>
<tbody>
<tr>
<td>275-1</td>
<td>77.5765</td>
<td>12.9448</td>
<td>885.2</td>
<td>896.2</td>
<td>890.4</td>
</tr>
<tr>
<td>965-3</td>
<td>77.6237</td>
<td>12.9447</td>
<td>884.46</td>
<td>875.46</td>
<td>883.46</td>
</tr>
<tr>
<td>15-1</td>
<td>77.6641</td>
<td>12.9924</td>
<td>893.5</td>
<td>888.5</td>
<td>896.5</td>
</tr>
<tr>
<td>104-1</td>
<td>77.5874</td>
<td>12.9331</td>
<td>896.6</td>
<td>908.6</td>
<td>903.6</td>
</tr>
<tr>
<td>344-6</td>
<td>77.5368</td>
<td>13.0293</td>
<td>900</td>
<td>893</td>
<td>896</td>
</tr>
</tbody>
</table>

5. Conclusions

This study has demonstrated the usefulness of kriging as a tool to determine the reduced level of the rock in Bangalore considering a large data set (652 points) distributed over 220 sq·km area. Geostatistics has permitted the development of a semivariogram model for predicting reduced level of the rock in Bangalore.

The power of geostatistics has become even more apparent through the estimated reduced level of the rock in a way that is consistent with what we know from the field data. By the use of semivariogram, it is possible to make estimation of the reduced level of the rock at points of the site where reduced levels of the rock were not known. The models have been developed by using simple as well as ordinary kriging methods. Variance map for both kriging techniques have been generated and presented. A new type of cross-validation analysis (Q1 and Q2) which proves the robustness of the developed kriging models has been also presented in this study. Ordinary kriging yielded better results than simple kriging for estimation of reduced level of the rock in the subsurface of Bangalore.

6. References


