Detecting Musk Thistle (*Carduus nutans*) Infestation Using a Target Recognition Algorithm

Mustafa Mirik¹, Yves Emendack¹, Ahmed Attia¹, Sriroop Chaudhuri¹, Mimi Roy², Georges F. Backoulou³, Song Cui⁴

¹Texas A & M AgriLife Research and Extension Center, Vernon, USA
²Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, USA
³Oklahoma State University, 127 Noble Research Center, Stillwater, USA
⁴Middle Tennessee State University, Murfreesboro, USA

Email: song.cui@mtsu.edu

Received 14 June 2014; revised 10 July 2014; accepted 18 August 2014

Abstract

The outbreaks of invasive plant species can cause great ecological and agronomic problems through aggressively competing for environmental resources that could be otherwise utilized by other desirable species. Thus, it is crucial for detecting small infestations before they reach a significant extent that can cause ecological and economic damages over a large geological area. Remote sensing is a proven method for mapping invasion extent and pattern based on geospatial imagery and indicated great repeatability, large coverage area, and lower cost compared with traditional ground-based methods before. We investigated the feasibility and performances of adopting multispectral satellite imagery analyses for mapping infestation of musk thistle (*Carduus nutans*) on native grassland, crop field, and residential areas in early June using spectral angle mapper classifier. Our results showed an overall classification accuracy of 94.5%, indicating great potential of using moderate resolution multispectral satellite-based remote sensing techniques for musk thistle detection over a large spatial scale.

Keywords

Accuracy Assessment, Invasive Plant, Weed Management, Weed Infestation, Remote Sensing, Geospatial Data, Nodding Thistle

1. Introduction

Non-native invasive plant species are well known for their successful exploitation of environmental resources...
M. Mirik et al.

(e.g., water, space, light and nutrients), largely due to their aggressive and competitive ability, prolific seed production and seed longevity [1] [2]. Along with having a competitive advantage in resource utilization over native species, many non-native weeds grow in the absence of natural enemies [1] [2]. This allows non-native invasive weeds to easily establish and compete with native plants in the environment [1]. Due to these characteristics, non-native invasive weeds usually move from small, manageable infestations to larger areas, reaching levels where control is economically prohibitive [3]. Some non-native invasive species could dominate the vegetative canopy and eventually could form monotypic stands [3]. Invasion may alter nutrient accumulation and cycling, hydrology, carbon sequestration and grass production [4] [5]. Weed infestations have been recognized as a primary reason for loss in global biodiversity and even species extinction [6] [7]. Therefore, non-native invasive plant species have long been of interest to natural resource managers, ecologists and biological conservationists [8].

One non-native weed of major ecological and economic importance in the US, Canada, Australia, New Zealand and southern Africa is musk thistle or nodding thistle \((\text{Carduus nutans} \text{L.})\), an aggressive member of the sunflower (Asteraceae) family that is native to North Africa and Eurasia [9]-[14]. Musk thistle infests a considerable amount of land area in the US (found in 45 states) and Canada (found at least in five provinces) and has a long history in these countries as an invasive species [12] [13]. Musk thistle has been declared a noxious weed in many US states, Canadian provinces and Australia. It infests forest, pasture and rangelands, roadsides, railroad rights-of-way, ditch banks, stream beds, waste or disturbed areas, agricultural fields [9] [10], as well as residential areas. The presence of this weed can impede grazing by livestock and decreases the forage production due to competition with grasses [9] [10]. Since livestock do not graze musk thistle, it easily outcompetes grassland forage species and can dominate entire fields with up to 150,000 plants per hectare [14] [15]. It has been determined that a single musk thistle per 1.5-m\(^2\) reduces pasture yields by 23% [16]. It has been also reported that musk thistle reduced 72% of the amount of forage utilized in beef-cattle and sheep-grazed pastures [17].

Early determination of the distribution and severity of rapidly spreading invasive populations of musk thistle is needed to implement mitigation treatments, but accurate assessments are often difficult or impossible to obtain with ground surveys, because of the extensive land area involved, time and labor required and inaccessibility of many areas [18]. Therefore, remote sensing has received considerable attention as a rapid, inexpensive and non-destructive method for assessing non-native noxious species invasions [4] [18]. It is a technology that provides a complete assessment instead of an assumed level based on traditional field-based sub-sampling methods.

A wide range of sensor systems, including aerial photographs, airborne and satellite multispectral and hyperspectral images have been successfully utilized for mapping the distribution of certain plant species [19]-[22]. An important characteristic of remotely sensed imagery for mapping plant species is the spatial resolution or pixel size representing the smallest area identifiable in the imagery. Higher spatial resolution imagery often contains detailed intraspecies spectral variability than lower resolution imagery, when target species occur in small patches or cover smaller areas than the pixel size [19] [23] [24]. Timing, location, species composition, spectral and spatial resolutions are all important for vegetation mapping at the species level, because spectral reflectance from vegetation is determined by structural and biochemical properties, including leaf area, leaf angle distribution, architecture, biochemical and pigment (lignin, cellulose, chlorophylls, carotenes, nitrogen and water) contents [25] over time and space. Remote detection of invasive plant species using geospatial imagery may substantially improve monitoring, planning, and management practices by overcoming some of the shortcomings of ground-based surveys such as observer bias and inaccessibility to certain locations. Remote sensing techniques for accurate mapping of invasion offer a unique set of advantages including repeatability, large area coverage and cost-effectiveness over ground-based methods over time and space [26]-[28].

One way to extract information about the invasive species from optical imagery involves using classification methods, broadly grouped as unsupervised and supervised. Supervised classification can be per-field, per-pixel, and sub-pixel types [29]. Sub-pixel classification approaches are deemed more advantageous, than the per-pixel methods, as it provides more appropriate representation and accurate area estimation of land cover types, especially with coarse spatial resolution and high heterogeneity [29]. Some frequently used sub-pixel algorithms include fuzzy-set and spectral mixture analysis classifications [29]. The latter includes linear spectral unmixing (LSU), matched filtering (MF), mixture tuned matched filtering (MTMF), spectral angle mapper (SAM), adaptive coherence estimator (ACE), orthogonal subspace projection (OSP), and constrained energy minimization (CEM) techniques among others. The key to sub-pixel classification is to identification of spectral endmembers in the imagery [30], which may be derived from laboratory or field measurements using ground-based portable spectroradiometers. The endmembers can also be simulated using radiative transfer models or extracted directly.
Sub-pixel information extraction methods, such as SAM, are among the most useful supervised target detection algorithms. Six land cover classes were assessed with an overall accuracy of 70% using SAM algorithm [32]. Burn areas in forestland were mapped with an overall accuracy of 84% using SAM classifier [33]. SAM method was able to distinguish spotted knapweed (Centaurea maculosa) cover as little as 1% to 40% [34]. Hoary crest (Cardariadraba) was discriminated with an overall accuracy of 65% using SAM [35]. SAM was used to detect the distribution of five invasive plant species saltcedar (Tamarix spp.), Russian olive (Elaeagnus angustifolia), Canada thistle (Cirsium arvense), musk thistle, and reed canary grass (Phalaris arundinacea) with an overall accuracy of >74% [22].

Despite the use of remotely sensed data for identification and mapping of non-native invasive plants and the large distribution of musk thistle in the US and other countries, the application of moderate resolution (e.g., 30-m) publicly available multispectral imagery for mapping musk thistle on grassland, rangelands, cropland, and residential areas is not well studied. Utilizing moderate resolution publicly available multispectral imagery to discriminate and separate invasive musk thistle from highly variable vegetation comprised of many species of grasses, broadleaved forbs, and crops is needed in order to improve monitoring population dynamics and developing long-term adaptive management practices for musk thistle. Typically, in the southern Great Plains (in Texas and Oklahoma, USA), the diversity of plant species at any one site includes species with cool-season (C3) or warm-season (C4) photosynthetic pathways that usually have very different periods of physiological activity during the growing season and would have markedly different spectral properties at any one point in time.

Our objective was to evaluate the use of moderate resolution publicly available multispectral imagery and SAM sub-pixel classification method for detecting and discriminating musk thistle from the co-occurring vegetation in a rangeland, cropland, and residential setting. Since the multispectral remote detection of musk thistle has not been well documented, our study presents the first application of moderate resolution publicly available multispectral data for discriminating musk thistle using SAM. The reason why this classifier has potential is because it can distinguish and map single target objects rather than all objects present in the images.

2. Materials and Methods

2.1. Site Description

Musk thistle infestations were observed on areas located in southwest of City of Amarillo and in northwest of City of Canyon in Randall County, Texas High Plains, USA (Figure 1) during the mid-spring in 2007.

A ground survey was conducted to locate infested areas by musk thistle and record coordinates using a hand-held field computer equipped with the Farm Works software package (CTN Data Service, Hamilton, IN) and a GPS receiver. Heavy infestations of musk thistle at flowering growth stage were found at 36 locations over the study area (Figure 2). A cloud free 7-seven bands Landsat 5 Thematic Mapper (TM) scene with 30-m spatial resolution that were dated June 7, 2007 covered the northwestern Texas High Plains with path 30 and row 36 and scene center latitude 34.27˚ and longitude −102.19˚ were obtained. Images were provided by the United State Geological Service (USGS) via the Center for Space Research, University of Texas at Austin with geometric and radiometric corrections. Image digital numbers were converted to reflectance using Environment for Visualizing Images software package (ENVI: Exelis Visual Information Solutions, Boulder, CO). Image was subset for a 44,796 ha (21 km by 21 km) area that included all locations of musk thistle infestations identified on the ground (Figure 3(a)).

2.2. Imagery Classification

One location with a heavy musk thistle infestation was used to collect endmember spectra from the image for classification. This location was verified using some landmarks (e.g., roads, neighboring fields, field shapes etc.) and overlying the point coordinates over the image. Endmember spectra of musk thistle manually extracted from the imagery, consisted of a total of 25 pixels. Meanwhile, another total of 25 pixels of green vegetation at three locations were selected as non-target spectra to reject in classification.

After visually evaluating a few sub-pixel classification methods (SAM, MF, MTMF, LSU, ACE, OSP, and CEM) using “the target detection wizard” in Environment for Visualizing Images (ENVI; Exelis Visual Information Solutions, Boulder, CO). SAM classifier was selected for our objectives due to its superior ability to
detect musk thistle. SAM matches image spectra to reference target spectra in $n$ dimensions. SAM compares the angle between the target spectrum (considered an $n$-dimensional vector, where $n$ is the number of bands) and each pixel vector in $n$-dimensional space. Smaller angles represent closer matches to the reference spectrum. When used on calibrated data, this technique is relatively insensitive to illumination and albedo effects (ENVI Users Guide). SAM was used with 0.30 rule threshold and binary preview after applying minimum noise fraction transform during the classification (Figure 3(b)). Resulting classified musk thistle was converted to vector format (shape file) for ground truthing. Steps involved in image classification are given in Figure 4.
2.3. Field Validation

Field validation (accuracy assessment) of the classification performance was carried out using verification data (ground control points) over the study area (Figure 5). In addition to 36 musk thistle locations, another 164 field validation points were created using the “create random points” function in ArcGIS (ESRI Inc., Redland, CA). The verification points were exported to a field computer equipped with Farm Works software and a global positioning system receiver and navigated on the ground at each verification point within a 15-m radius area,
Figure 4. Workflow of steps involved in image classification for musk thistle detection using spectral angle mapper.

Figure 5. Detected musk thistle and ground verification locations displayed over Landsat 5 Thematic Mapper band 1 (a); and musk thistle endmember, a total of 25 pixels (b).
presence (at least more than one musk thistle plants found) and absence (no musk thistle plant found) of musk thistle were recorded.

2.4. Accuracy Assessment

Classification of remotely sensed imagery is the process of assigning variables into discrete categories of useful information. However, errors of classification occur because of mixed pixels, similar spectra of different materials, positional errors and human errors [36]. The most effective way to represent classification accuracy is via an error matrix [37]. Therefore, accuracy assessment for classification was made by constructing an error matrix for classified image. An error matrix is a square, with the number of columns and rows being equal to the numbers of categories whose classification accuracy is being evaluated [37]. Error matrix for classification map was generated by comparing the classified classes with the field validation points. Error matrices to evaluate the classification accuracy were calculated, including overall, user’s and producer’s accuracies.

There is no single set standard for selection of the image and ground areas for comparison [37] [38]. Because a pixel in an image has an arbitrary location on the ground and because positional errors of maps and global positioning system receivers become significant with small pixel sizes, areas based on geographic information system polygons are used frequently. However, using individual pixels is appropriate if a per-pixel classification is assessed for accuracy, which avoids problems caused by generating “homogeneous” polygons on a landscape [39]. It has been shown that pixel positional error results in a conservative bias of the accuracy assessment [40]; therefore, the unavoidable positional error introduced into this assessment would result in lower or conservative estimates of mapping accuracy [36].

3. Results and Discussion

The classification map for musk thistle is illustrated in Figure 3(c) and Figure 5. The error matrix that compared ground verification data with classified data is contained in Table 1. An overall mapping accuracy of 95% with a kappa coefficient of 0.82 was achieved for musk thistle detection. About 19% of the other land cover types (a total of eight locations) was included into musk thistle category, whereas about 8% of the musk thistle (a total of three locations) was misclassified as other classes by the classification method performed.

SAM method identified several additional musk thistle infestations that were not located during the ground survey due to the remote areas from the roads. It was noted during the ground-truthing that undetected musk thistle at three locations consisted of yards of two houses and a barnyard building with musk thistle infestation not larger than 30-m in diameter. This indicates that building materials suppressed musk thistle detection. Eight locations where musk thistle undetected were infested by mustard (Brassica spp.) and clover (Trifolium spp. and Melilotus spp.) with light to heavy densities. This indicates that these species had similar reflectance characteristics to musk thistle.

There is no set standard for classification accuracy as to what level of accuracy is adequate. The acceptable level of mapping accuracy depends on the project-based choice that differs based on the needs of the end users and utilization of the classification results [36] [37]. An accuracy target of 85% was recommended by [41], and an overall accuracy target of 85% with no individual class accuracy < 70% was suggested by [42]. Although there is

<table>
<thead>
<tr>
<th>Table 1. Confusion matrix generated from the ground verification and classified data for musk thistle.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual category</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Present</td>
</tr>
<tr>
<td>Absent</td>
</tr>
<tr>
<td>Column total</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
</tr>
<tr>
<td>Overall accuracy (%)</td>
</tr>
<tr>
<td>Kappa statistics</td>
</tr>
</tbody>
</table>
no information related to the acceptable classification accuracies specified for remote detection of musk thistle, a threshold of the user’s accuracy > 70% is required for mapping leafy spurge [43]. Based on these recommended accuracy guidelines, our classification accuracies (overall, user’s and producer’s accuracies) were above the recommended guidelines. One of the most likely utilizations of these accuracy maps would be to track changes in musk thistle population density over time and space. With consistent and repeatable classification methodology, the level of accuracies we accomplished in this study would present very accurate musk thistle distribution over the study area.

Image classification methods for weed identification and mapping have a tendency to overestimate the target species [44]. In our study, the producer’s accuracy of 88% was slightly lesser than the user’s accuracy of 90% for image classification. Producer’s accuracy is more important than user’s accuracy for weed detection, because undetected or omission of target weeds (false negative) is the factor that most likely hampers control efforts [3] [45]. There is greater potential for misclassification of musk thistle when it grows in close proximity to dense patches of background live vegetation, in locations where musk thistle with canopy coverage is <15% or when individual musk thistle plants are small in size. All of these factors affected classification accuracy in the present study. Our classification results were consistent with the findings of other studies. For example, leafy spurge (Euphorbia esula) with cover as low as 10% [43] and musk thistle with varying densities [3] were successfully detected using hyperspectral images with pixel sizes of 3.5 by 3.5-m and 1 by 1-m, respectively.

With the history of remote sensing, researchers have sought the optimal strategy to extract as much information as possible from remotely sensed data by using the best possible physiological, compositional and structural distinctions in target species, classification algorithm, data transformation and band selection. The accuracy of a classification task could be improved by focusing on most informative bands or spectral regions. Just a few species-specific spectral bands may be sufficient to detect target species. Since our image classification provided sufficient accuracies to be used for management decisions regarding musk thistle treatment, our study provides an application of moderate resolution multispectral imagery and classification method for discriminating and mapping musk thistle.

It is more beneficial for the resource manager to have information regarding weed infestations as early as possible. Detecting and mapping of musk thistle on an early spring image before it reaches peak flowering, seed production and dispersal stage could be more favorable to implement treatments when control measures are needed for the current growing season. When control measures are not needed or planned for the current growing season, mapping of musk thistle infestation at the flowering stage is the better option for more accurate mapping and planning for future control measures.

4. Conclusion

Remote sensing techniques offer important opportunities for detecting plant species and mapping the extent of invasions in a timely and spatial manner, especially when the target species have unique characteristics from the co-occurring vegetation mosaic. This study investigated the use of 30-m spatial resolution multispectral imagery and spectral angle mapper target detection method for detecting musk thistle infestation in a rangeland, cropland and residential setting. Given the economic and ecological outcomes of invasion by this weed, our result indicates that publicly available Landsat imagery is a useful tool for mapping musk thistle infestation. The accuracy assessment revealed that the overall mapping success was 94.5% with the user’s and producer’s accuracy >80% and >91%, respectively. Our results have implications for rangeland management and productivity. Maps generated at a 30-m scale provide the spatial extent and dynamics of musk thistle infestation over time. These maps can be used for monitoring, planning and control measures. Such maps have also been shown to provide animal distribution, grazing effects and the effects of natural and anthropogenic disturbances in rangeland, cropland, and residential areas. Using maps at a 30-m scale, the reduction in range productivity can be estimated, because a single musk thistle per 1.5-m² decreases pasture yields by 23%. The use of a 30-m spatial resolution image provided valuable estimates of musk thistle infestation. Therefore, we recommended that this methodology and technology should be considered when moderate scale maps are needed for rangeland research and management.

References


M. Mirik et al.


Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.