Determining the Suitability and Accuracy of Various Statistical Algorithms for Satellite Data Classification

Khalid Omar Murtaza, Shakil A. Romshoo
Department of Earth Sciences, University of Kashmir, Srinagar Kashmir (India)
komurtaza@gmail.com

ABSTRACT

Land use and land cover (LULC) data is very important for determining the nature and mechanism of different land surface, hydrological processes. The production of land use land cover map, using an image classification is one of the most common applications of remote sensing. However, image classification is a complex process that may be affected by many factors including spatial resolution, classifier used, training sets, etc. This paper briefly reviews the suitability of different methods of classification that are commonly used and their impact on classification accuracy. Three different supervised classification techniques (Maximum likelihood, Mahalanobis Distance, and Minimum Distance) were applied in Kashmir valley for the classification of the IRS LISS-III (2008) image in thirteen different LULC classes; agriculture, aquatic vegetation, barren land, built-up, exposed rock, forest, horticulture, pastures, plantation, riverbed, scrub land, snow, and water. The classified maps where then visually compared with each other and the accuracy of classified map was assessed using the reference data sets which consisted of a large number of ground samples collected in each land cover category. The overall accuracy for Maximum likelihood classifier was 89%, for Mahalanobis distance was 54% and for Minimum distance was 48%. It was observed that the Maximum likelihood method gave the best results and good agreement between classes extracted from the classified maps and field observations. Mahalanobis distance method has overestimated agriculture land, plantation and built-up. Minimum distance method overestimated water, built-up and horticulture and underestimated agriculture. The selection of a suitable classification method is significant for improving classification accuracy.

Keywords: Remote sensing, land use, land cover, Kashmir, Maximum likelihood, Mahalanobis Distance, and Minimum Distance.

1. Introduction

Land use Land cover (LULC) studies are important for understanding land surface processes, hydrology and the interactions of the human activities with the environment (Altaf et al., 2012; Meraj et al., 2013). It is beyond doubt that human activities have modified the natural environment considerably and nowadays the intensity and scale of these modifications has increased significantly (Goldewijk 2001). Anthropogenic activities have profound impacts upon the natural setting of global systems (Romshoo and Muslim 2011; Rashid and Romshoo 2012; Bhat et al., 2013, 2014) and the most striking human induced changes of the current era is because of the increase in population (Seiferling et al., 2012; Weinzettel et al., 2013). Remote Sensing (RS) technologies are playing an important role in studying dynamics of LULC. A number of these technologies can supply data to solve problems, and can often be accomplished at a lower relative cost than many other traditional methods (James and Daniel, 2002). Remote sensing data of the earth's surface could be made readily available in digital format (Richards and Jia, 1998). A principal application of remotely sensed data is to create a
The classified map of the identifiable or meaningful features or classes of land cover types (Jasinski, 1996). Therefore, the principal product is a thematic map with themes like land use land cover, geology and vegetation types. These advantages have attracted great interest in the scientific and engineering community (Lyon, 1995). The reasons of remote sensing priorities over traditional methods are because of several unique aspects including the capability to measure spatial, spectral, and temporal information as opposed to point data, ability to assess the state of the earth’s surface over large areas, and to assemble long-term data sets and the capability to measure inaccessible areas (Qi et al., 1994; Ritchie and Rango, 1996; and Rango and Shalaby, 1998). Researches on image classification based remote sensing have long attracted the interest of the remote sensing community since most environmental and socioeconomic applications are based on the classification results (Lu and Weng, 2007). The “landscape-scale” requires methods to gather spatially distributed information and this requires repeated sampling of the variables of interest to acquire information over large areas. The costs and logistics of these actions can be high, and work is usually constrained by available resources. However, remote sensing is considered the most efficient technology to handle these problems and to observe the spatially distributed variables (Lyon, 1995).

Image classification is an important part of the remote sensing, image analysis and pattern recognition. In some instances, the classification itself may be the object of the analysis. The image classification therefore forms an important tool for examination of the digital images. The term classifier refers loosely to a computer program that implements a specific procedure for image classification (Campbell, 2002). Proper classification of LULC is a very essential requirement for all modelling tasks in environmental problems (Romshoo and Muslim 2011; Rashid et al., 2011). At present, there are different image classification procedures used for different purposes by various researchers (Butera, 1983; Ernst and Hoffer, 1979; Lo and Watson, 1998; Ozesmi and Bauer, 2002; Dean and Smith, 2003; Pal and Mather, 2003; and Liu et al., 2002). Scientists and practitioners have made great efforts in developing advanced classification approaches and techniques for improving classification accuracy (Gong and Howarth, 1992; Kontoes et al., 1993; Foody, 1996; San Miguel-Ayanz and Biging, 1997; Aplin et al., 1999; Stuckens et al., 2000; Franklin et al., 2002; Pal and Mather, 2003; Gallego, 2004). Nevertheless, knowing the best classification method to perform this task is a very important aspect in order to utilize the right approach for classification. Although much previous research is specifically concerned with image classification (Tso and Mather, 2001; Landgrebe, 2003), a comprehensive up-to-date review of classification approaches for higher accuracy is not available. Continuous emergence of new classification algorithms and techniques in recent years necessitates such a review, which will be highly valuable for guiding or selecting a suitable classification procedure for a specific study. The focus of this paper was on providing a summarization of different classification methods and briefly reviews the status of land use land cover classification accuracy using, Maximum likelihood, Mahalanobis distance, and Minimum distances classifiers.

2. Study area and data sets used

The study is focused in Kashmir valley, having area of 15948 sqkm². Kashmir lies between the Himalayan range in the north-east and the Pir Panjal range in the south-west (Figure 1). Geographically, the valley lies between 33°55’-34°50’ North latitude and 74°30’-75°35’ East longitude. Valley is subdivided into 24 watersheds. Jhelum River is one of the main tributaries of the upper Indus basin. The Valley receives precipitation both in the form of rain and snow. The analysis of the average monthly rainfall shows that the area receives highest
rainfall in the month of March. The weather in the area remains pleasant from April to October, but in rest of the year, particularly in winters, the study area experiences extreme cold and heavy snowfall. Its relief is diverse, comprising of steep slopes, plateaus, plains, and alluvial fans. The plains of the valley are very fertile, hence ideal for agriculture, whereas the higher reaches comprise dense pine forests and lush green alpine pastures. Kashmir's economy is centered on agriculture and tourism. In geological literature, the Kashmir valley is referred as one of the most tectonically active areas in the world that is also known as “the Nappe zone of Kashmir,” which lies to the south of the higher Himalaya crystalline and Panjal thrust to the south.

In the present study Indian Remote Sensing (IRS) LISS-III data with 23.5m spatial resolution of 21st Oct 2008 and path/row of 92/46 (band 1, 2, 3 and 4) was used to generate land use land cover information of the study area using supervised classification. Besides using satellite data, extensive field survey was also done.

![Figure 1: Location map of the study area](image)

3. Methodology
3.1 Mapping approach

Supervised Image classification was used in the study for classification. Supervised classification of multispectral remote sensing imagery is commonly used for LULC determination (Duda and Canty, 2002). In supervised classification, the analyst supervises the pixel categorization process by specifying the computer algorithm of the statistical parameters, viz. mean vector and variance-covariance matrix of various thematic classes present in the scene. The quality of a supervised classification (Palaniswami et al., 2006) depends on the quality of the training sites. All the supervised classifications usually have a sequence of operations that must be followed. 1), Defining of the Training Sites. 2), Extraction of Signatures. 3), Classification of the Image. The training sites are done with digitized features. Usually two or three training sites are selected. The more training sites are selected, the better results can be gained. This procedure assures both the accuracy of classification and the true interpretation of the results. After the training site areas are
digitized then the statistical characterizations of the information are created. These are called signatures. Finally the classification methods are applied. Three supervised classification methods were used in this study; these are Maximum likelihood, Mahalanobis distance, and Minimum distance for classification of different LULC classes.

3.2 Maximum likelihood

Maximum likelihood is one of the most popular supervised classification method used with remote sensing image data. This classification method uses the training data as a means of estimating means and variances of the classes, which are then used to estimate probabilities. Maximum likelihood classification considers not only the mean or average values in assigning classification, but also the variability of brightness values in each class. It is the most powerful of the classification methods as long as accurate training data is provided (Asmala, 2012). Therefore this method requires excellent training data. An advantage of this method is that it provides an estimate of overlap areas based on statistics. This method is different from parallelepiped that uses only maximum and minimum pixel values. The Maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions. The Maximum likelihood algorithm assumes that the histograms of the bands of data have normal distributions. If this is not the case, you may have better results with the parallelepiped or Minimum distance decision rule, or by performing a first-pass parallelepiped classification. It is one of the commonly used supervised classifications (Benediktsson et al., 1990). Maximum likelihood classification calculates the following discriminant functions for each pixel in the image:

\[ g_i(x) = \frac{1}{np(\omega_i)} - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (x - m_i)^T \Sigma_i^{-1} (x - m_i) \]  

(1)

where: \( i \) = the \( i \)th class, \( x = n \)-dimensional data (where \( n \) is the number of bands), \( p(\omega_i) \) = probability that a class occurs in the image and is assumed the same for all classes, \( |\Sigma_i| \) = determinant of the covariance matrix of the data in a class, \( \Sigma_i^{-1} \) = the inverse of the covariance matrix of a class, \( m_i \) = mean vector of a class.

3.3 Mahalanobis Distance

Mahalanobis distance classification is similar to Minimum distance classification, except that the covariance matrix is used in the equation. This algorithm assumes that the histograms of the bands have normal distributions (Perumal and Bhaskaran, 2010). Variance and covariance are figured in so that clusters that are highly varied, lead to similarly varied classes, and vice versa. Mahalanobis distance between each pixel and the mean of each class. For example, when classifying urban areas typically a class whose pixels vary widely correctly classified pixels may be farther from the mean than those of a class for water, which is usually not a highly varied class. The Mahalanobis distance decision rule uses the covariance matrix in the equation. Variance and covariance are figured in so that clusters that are highly varied will lead to similarly varied classes, and vice versa. The Mahalanobis distance algorithm assumes that the histograms of the bands have normal distributions. If this is not the case, you may gain better results with the parallelepiped or Minimum distance decision rule, or by performing a first-pass parallelepiped classification. Mahalanobis distance classification calculates the Mahalanobis distance for each pixel in the image to each class:

\[ D_i(x) = \sqrt{(X - m_i)^T \Sigma_i^{-1} (x - m_i)} \]

(2)
where: $D =$ Mahalanobis distance, $i =$ the $i$th class, $x =$ $n$-dimensional data (where $n$ is the number of bands), $\Sigma_i^{-1} =$ the inverse of the covariance matrix of a class, $m_i =$ mean vector of a class.

### 3.4 Minimum Distance

The Minimum distance decision rule (also called spectral distance) calculates the spectral distance between the measurement vector for the candidate pixel and the mean vector for each signature. This classification method derives distance between any pair of pixels after defining training data. The Minimum distance classifier can be used as a supplement to the parallelepiped classification method, which can leave unanswered pixels. The classification technique takes pixels of known identity and then includes pixels closest to it as training pixels. Like the other methods of classification, this method uses two bands to evaluate the training data. Minimum Distance classification calculates the Euclidean distance for each pixel in the image to each class:

$$D_i(x) = \sqrt{(x - m_i)^T \Sigma_i^{-1} (x - m_i)}$$

where: $D =$ Euclidean distance, $i =$ the $i$th class, $x =$ $n$-dimensional data (where $n$ is the number of bands), $m_i =$ mean vector of a class.

### 3.5 Classification Accuracy Assessment

Accuracy assessment of classification can be defined as the process of comparing the classification with geographical data that are assumed to be true, in order to determine the accuracy of the classification process. Usually, the assumed-true data are derived from ground-truth data. Accuracy assessment measures the agreement between a standard (assumed to be correct) and a classified map. This represents the correctness of the classified map (Campbell, 1987, 96). After a classification is performed it is to be evaluated, so as to check the authenticity and reliability of the classified data. Reference pixels are points on the classified image for which actual data are (or will be) known. The reference points or pixels were selected from ground field points. Points from different classes were taken in order to check the accuracy of as many as possible classes. The number of reference pixels is an important factor in determining the accuracy of the classification. 388 reference pixels were used to estimate the overall accuracy. The error matrix approach is the one most widely used in accuracy assessment (Foody, 2002a). Error matrix is a table that displays statistics for assessing image classification accuracy by showing the degree of misclassification among different classes. It can be used to generate various statistics that characterize the accuracy of a classification technique. Overall classification accuracy is given by following formula (Veregin1995):

$$\rho = \left( \frac{n}{N} \right) \times 100$$

Where ‘$\rho$’ is classification accuracy, ‘$n$’ is number of points correctly classified on image, and ‘$N$’ is number of points checked in the field.

However, this statistics can be misleading since a certain number of correctly classified pixels are expected to occur by chance alone. The Kappa coefficient is a measure of overall statistical agreement of an error matrix, which takes non-diagonal elements into account. Kappa analysis is recognized as a powerful method for analysing a single error matrix and for comparing the differences between various error matrices (Congalton, 1991; Smits et al., 1999; Foody, 2004). The Cohen’s Kappa statistics allows for chance, and ranges from 0 in
Determining the Suitability and Accuracy of Various Statistical Algorithms for Satellite Data Classification

Khalid Omar

International Journal of Geomatics and Geosciences
Volume 4 Issue 4, 2014

the case of the most confused classification to 1 in the case of the most accurate classification (Romshoo and Rashid 2012; Showqi et al., 2013). The Cohen's Kappa (k) statistics allows accessing the accuracy that takes into account the chance of random agreement (Smits et al., 1999; Dymond and Jonson 2002; Selamat et al. 2002). The equation for k is:

\[ k = \frac{pr(a) - pr(e)}{1 - pr(e)} \]

Where Pr (a) is the relative observed agreement among raters, and Pr (e) is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly saying each category. If the raters are in complete agreement then \( k = 1 \). If there is no agreement among the raters other than what would be expected by chance (as defined by Pr (e)), \( k = 0 \).

4. Results

Supervised image classification of IRS LISS-III image of 2008, was classified in thirteen different LULC classes; agriculture, aquatic vegetation, barren land, built-up, exposed rock, forest, horticulture, pastures, plantation, riverbed, scrub land, snow, and water using Maximum likelihood, Mahalanobis distance and Minimum distance classifiers. Figure 2a,b,c shows the thematic maps from three classifiers. Table 1 shows the status of LULC classes using different classifiers. When compared with field data, it was observed that Mahalanobis distance method has overestimated agriculture land, plantation and built-up. It misclassified barren land as agriculture land, horticulture as plantation and exposed rock with built-up. Minimum distance method overestimated water, built-up and horticulture and underestimated agriculture which is one of the main land use in the study area. It misclassified forest as water, exposed rock and river bed as built-up, agriculture as horticulture and aquatic vegetation as scrub land. Maximum likelihood classifier classified all the classes correctly with minor misclassification of horticulture with plantation. A good agreement was seen between the classified map and field reality.

Accuracy assessment of all hard supervised classified images was carried out. Table 2, 3, 4 shows the Error matrix table of three classified images. The obtained overall accuracy for Maximum likelihood was 89%, 54% for Mahalanobis distance and for Minimum distance was 48%. Kappa coefficient robust indicator of the accuracy estimation (Jensen 1996; Foody 2002a) for classification products were 0.89, 0.54 and 0.47 for Maximum likelihood, Mahalanobis distance and Minimum distance respectively. Table 2, 3, 4 shows error matrix table for LULC maps using Maximum likelihood, Mahalanobis distance and Minimum distance respectively. The Maximum likelihood classifier has highest overall accuracy (Thakur et al., 2012; Akgün et al., 2004; Ahmadi and Hames, 2008) and highest Kappa coefficient in classification followed by Mahalanobis distance and Minimum Distance.

Table 1: Land use land cover status

<table>
<thead>
<tr>
<th>LULC Classes</th>
<th>Maximum Likelihood</th>
<th>Mahalanobis Distance</th>
<th>Minimum Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>2691.39</td>
<td>2960.84</td>
<td>1409.80</td>
</tr>
<tr>
<td>Aquatic vegetation</td>
<td>117.29</td>
<td>63.30</td>
<td>29.48</td>
</tr>
<tr>
<td>Built up</td>
<td>214.33</td>
<td>805.65</td>
<td>1241.62</td>
</tr>
<tr>
<td>Barren land</td>
<td>508.88</td>
<td>86.86</td>
<td>436.38</td>
</tr>
<tr>
<td>Exposed rock</td>
<td>1183.82</td>
<td>1217.27</td>
<td>1144.81</td>
</tr>
<tr>
<td>Forest</td>
<td>3913.69</td>
<td>2374.82</td>
<td>2148.60</td>
</tr>
</tbody>
</table>
Figure 2: LULC map using (a) Maximum likelihood classifier, (b) Mahalanobis distance classifier and (c) Minimum distance classifier.
Determining the Suitability and Accuracy of Various Statistical Algorithms for Satellite Data Classification

Khalid Omar

Table 2: Error matrix for LULC map using Maximum likelihood classifier

<table>
<thead>
<tr>
<th>CLASS</th>
<th>HT</th>
<th>PL</th>
<th>RB</th>
<th>FO</th>
<th>BU</th>
<th>AG</th>
<th>SL</th>
<th>PA</th>
<th>WA</th>
<th>AV</th>
<th>ER</th>
<th>SN</th>
<th>TO</th>
<th>UA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT</td>
<td>56</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>63</td>
<td>88</td>
</tr>
<tr>
<td>PL</td>
<td>4</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>85</td>
</tr>
<tr>
<td>RB</td>
<td>0</td>
<td>3</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>75</td>
</tr>
<tr>
<td>FO</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>100</td>
</tr>
<tr>
<td>BU</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>103</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>99</td>
</tr>
<tr>
<td>AG</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>43</td>
<td>88</td>
</tr>
<tr>
<td>SL</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>88</td>
</tr>
<tr>
<td>PS</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>90</td>
</tr>
<tr>
<td>WA</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>68</td>
</tr>
<tr>
<td>AV</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>ER</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>SN</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>11</td>
<td>82</td>
</tr>
<tr>
<td>TO</td>
<td>65</td>
<td>50</td>
<td>18</td>
<td>28</td>
<td>110</td>
<td>42</td>
<td>16</td>
<td>21</td>
<td>15</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>388</td>
<td></td>
</tr>
<tr>
<td>PA (%)</td>
<td>86</td>
<td>82</td>
<td>66</td>
<td>96</td>
<td>93</td>
<td>90</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Table 3: Error matrix for LULC map using Mahalanobis distance classifier

<table>
<thead>
<tr>
<th>CLASS</th>
<th>HT</th>
<th>PL</th>
<th>RB</th>
<th>FO</th>
<th>BU</th>
<th>AG</th>
<th>SL</th>
<th>PA</th>
<th>WA</th>
<th>AV</th>
<th>ER</th>
<th>SN</th>
<th>TO</th>
<th>UA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT</td>
<td>34</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>61</td>
<td>55</td>
</tr>
<tr>
<td>PL</td>
<td>4</td>
<td>25</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>45</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>RB</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>16</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>FO</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>18</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>27</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>BU</td>
<td>1</td>
<td>11</td>
<td>17</td>
<td>0</td>
<td>45</td>
<td>20</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>99</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>42</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>SL</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>18</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>22</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>22</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>AV</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>SN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>22</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>TO</td>
<td>47</td>
<td>58</td>
<td>44</td>
<td>22</td>
<td>54</td>
<td>84</td>
<td>3</td>
<td>16</td>
<td>10</td>
<td>5</td>
<td>29</td>
<td>16</td>
<td>388</td>
<td></td>
</tr>
<tr>
<td>PA (%)</td>
<td>72</td>
<td>43</td>
<td>22</td>
<td>81</td>
<td>83</td>
<td>38</td>
<td>100</td>
<td>81</td>
<td>90</td>
<td>60</td>
<td>6.9</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

International Journal of Geomatics and Geosciences
Volume 4 Issue 4, 2014
Table 4: Error matrix for LULC map using Minimum distance classifier

<table>
<thead>
<tr>
<th>CLASS</th>
<th>HT</th>
<th>PL</th>
<th>RB</th>
<th>FO</th>
<th>BU</th>
<th>AG</th>
<th>SL</th>
<th>PA</th>
<th>WA</th>
<th>AV</th>
<th>ER</th>
<th>SN</th>
<th>TO</th>
<th>UA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT</td>
<td>23</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>17</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>62</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>4</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>45</td>
<td>40</td>
</tr>
<tr>
<td>RB</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>FO</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>27</td>
<td>62</td>
</tr>
<tr>
<td>BU</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>56</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>99</td>
<td>56</td>
</tr>
<tr>
<td>AG</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>20</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>41</td>
</tr>
<tr>
<td>SL</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>PS</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>22</td>
<td>45</td>
</tr>
<tr>
<td>WA</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>AV</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>ER</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>SN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>22</td>
<td>72</td>
</tr>
<tr>
<td>TO</td>
<td>38</td>
<td>51</td>
<td>5</td>
<td>17</td>
<td>74</td>
<td>85</td>
<td>22</td>
<td>15</td>
<td>15</td>
<td>3</td>
<td>3</td>
<td>39</td>
<td>388</td>
<td></td>
</tr>
<tr>
<td>PA (%)</td>
<td>60</td>
<td>35</td>
<td>60</td>
<td>100</td>
<td>75</td>
<td>23</td>
<td>31</td>
<td>66</td>
<td>86</td>
<td>66</td>
<td>33</td>
<td>41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of k-coefficient were plotted against different classifiers. It was observed that classification accuracy was achieved highest by using Maximum likelihood classifier followed by Mahalanobis distance classifier. The Mahalanobis distance classifier filtered out shadows very accurately. A possible explanation for this is that the parametric Mahalanobis distance classifier is able to classify highly varied clusters, such as was evident with the various shaded area, into similarly varied classes. Lowest accuracy was achieved by using Minimum distance classifier (Campbell, 2001). The graph obtained after plotting these results is shown in figure 3.

Figure 3: Graph showing comparison of k statistics
5. Conclusions

For more effective use of the satellite remote sensing, land use managers should be aware of the limitations and advantages of satellite data and should choose from their available land use mapping options accordingly. Remote sensing is especially proper for initial reconnaissance mapping and continued monitoring of land use over large areas. In this study three different classifiers were used to classify the IRS LISS-III (2008) image. The focus of the study was to assess the classification accuracies from different classifiers. The overall accuracy and Kappa value obtained were compared against the three different classifiers. Of all the three methods--Maximum likelihood classification method, Mahalanobis distance and Minimum distance method; the Maximum likelihood classification method produced more accurate results than the Mahalanobis distance and Minimum distance. And thus was considered as best method for image classification.

Acknowledgement:

The research work was conducted as part of the DST, Govt. of India sponsored research project on Flood Vulnerability Assessment in Jhelum basin and the financial assistance received under the project to accomplish thisresearch is thankfully acknowledged. The authors express gratitude to the anonymous reviewers for their valuable comments and suggestions on the earlier version of the manuscript that greatly improved the content and structure of this manuscript.

6. References


Determining the Suitability and Accuracy of Various Statistical Algorithms for Satellite Data Classification

Khalid Omar


