Comparing different classifications of satellite imagery in forest mapping (Case study: Zagros forests in Iran)

Maryam Niknejad¹, Vahid Mirzaei zadeh², Mehdi Heydari ³
1. MSc Student in Forestry, University of Ilam.
2. MSc Student in Forestry, University of Ilam.
3. Assistant professor, Department of Forest Science, University of Ilam.

Corresponding Author email: maryam612niknejad@yahoo.com

ABSTRACT: Forest mapping is essential to manage natural resources and environment, land use plans and also to determine land potential and it is defined as one of the main resource for adjusting development programs. Forest mapping is a difficult process that in all projects has financial and time constraints. Satellite data are one of the quickest and low cost methods for forest mapping. In recent years, researchers have produced different types of forest maps using these data. There are several methods for classification of satellite images. The present study was aimed to determine the most appropriate image classification of ETM” in Bivareh forests, a part of Zagros forests. Results showed that support vector machine (SVM) with Kappa coefficient 0.7069 and overall accuracy 88.65% is more accurate than other methods. The methods followed the order of accuracy form SVM, the maximum likelihood mahalanobis distance, the minimum distance, spectral information divergence, binary codes, parallelepiped to spectral angle mapping, respectively. The results of this study can be used as an effective tool for extracting forest maps with high accuracy.

Key words: Image classification, Forest mapping, Remote sensing, Kappa coefficient, Ilam province

INTRODUCTION

Land cover mapping in natural resources and environmental management and providing land use plans as well as determining land ability are necessary to create developing plans (Shataei et al, 2008). Excessive utilization causes adverse effects on forest and other parts of land. Therefore, the knowledge of natural resource for managers and experts is the main factor that can solve this problem (Mahini et al., 2012).

Forest mapping, especially in that part of the catchment basins with difficult availability is difficult and in most projects has financial and time constraints (Dontree, 2003). Remote sensing as a science of spatial information and GIS due to analytical capabilities can have a significant role on exploring and assessing forest cover. Various, cheaper, more qualitative and repetitive data can be obtained by images of remote sensing (Langford, 1997). Therefore, analyzing these data make an appropriate insight between human being and environment, especially using multiple images can help people to identify land cover (Brian and Michael, 2011). There are many applications of remote sensing data including land use and cover maps (Richards and Jia, 2006). Digital images have two-dimensional arrays of a pixel. Each array of this pixel demonstrates reflected energy from the ground up and it is has been stored in the corresponding area (sadegi-nad and Sahami-noshabadi, 1998). These numerical values describe various ground tasks on digital images of remote sensing. Hence, it is possible to classify these events and extract the ground information in related classes. Many classification methods typically use spectral information in the image bands. In these methods, it is expected that pixels with more or less light rate in multi spectral area classify in appropriate groups (Borri et al, 2005). In fact, the classification process of image is the conversion of data to understandable information (Mountrakis et al., 2011).

In the past, techniques to classify images were based on visual interpretation that often made systematic errors due to the experience of the interpreter’s ability to interpret the images (Lillesand et al., 2004). Digital categorization of images will help to classify process get closer to reality (Oommen, 2008). Classifying digital data is based on both supervised and unsupervised classification that both of them use quantitative and automatic
decision. In unsupervised classification, automatically and based on different effect of spectral and statistical calculations, each pixel belongs to a particular class. The downside of this method is that distinguishing the phenomenon is severely difficult when corresponding phenomenon has a little different spectral rather than others (Hord, 1982). Supervised classification overcomes this problem using training samples as a key tool that indicates the spectral effects of the special phenomenon (Lillesand et al., 2004). Supervised classification methods are divided into two groups: parametric and non-parametric. The main problem of parametric classification (e.g., maximum likelihood, minimum distance, etc.), is its dependence on the statistical distribution of the data (e.g. Gaussian normal distribution). So, non-parametric methods, such as Support Vector Machine (SVM) are those that have been used for classification of satellite images. Various studies have been conducted on classification methods. Rahdari et al. (2009) studied the comparison of different methods of IRS satellite images in Isfahan province, Iran. The results showed that the method of maximum likelihood classification based on kappa coefficient 0.90 was the most accurate method. Al-Ahmadi and Hames (2009) in a study in an arid region of Saudi Arabia, compared four unsupervised and supervised classifications (maximum likelihood, minimum distance and mahalanobis distance) using ETM* images. They concluded the maximum likelihood method with kappa coefficient 0.68 was the most accurate method. Su et al. (2009) concluded that data of MISR satellite indicated using SVM methods and maximum likelihood could get complete accuracy with 2/3 data. Konrn et al. (2009) used SVM method for forest classification in order to evaluate classification accuracy in a set of successive images. The study showed that this method also would be effective when the ground is covered with several plant covers. Lardeux et al. (2009) used SVM for classifying the density of tropical vegetation. The results indicated that the SVM classification had 20 percent accuracy rate more than Wishart method. They resulted SVM method could be much more appropriate when the radar data do not follow the Wishart distribution. Knudby et al. (2010) concluded that the method of support vector machine had the highest ranking among other used algorithms for plant diversity. Wang et al. (2011) used the combination method of SVM and ACA on the Landsat TM and ETM * and SVM-ACA was introduced as an efficient method in covered regions. Mountrakis et al. (2011) studied on SVM method and classified studied conducted on this method based on clarifying static and dynamic purposes, spatial and spectral resolution of images and the type of images.

Literature review has shown that despite many studies have been conducted in different parts of the world, but most of these studies have considered only few classification methods. In addition, in Zagros forests located in west of Iran due to their highly ecological values, there has not been applied an appropriate method. Therefore, this study was aimed to introduce the best method among various algorithms in order to creating the map of Zagros forests.

MATERIALS AND METHODS

Site description

The study area is a part of the forests in malekshahi city, Ilam province located in west of Iran, called Bivareh with area of 22577 ha. The geographical position is ranging from east longitude 46° 30’ 38” to 46 ° 40’ 23” and from north latitude 33˚ 19' 49 " to 33 ° 26 ' 52". The meteorological data shows that this area is mostly arid and semiarid and Quercus branti is the main and dominant tree species of these forests.

Figure 1. Location of study area
**Data used for this research**
In the study, Landsat ETM+ Imagery of 2007 was used and also the topographic maps 1:50,000 were selected for getting the goals.

**Image Preprocessing**
The data preprocessing stage is the most important stage in image processing, because all the calculations are performed based on the output images at this stage. Type and mode of operation are diverse depending on various factors such as the type of data used and the research purpose. Image map was applied for the geometric correction of the images by the number of points with the proper distribution on the image. After using polynomial parametric method and remove inappropriate points, geometric correction was done with residual points and error rate of pixel 0.21.

**Preparing the training samples**
 Basically classification of remote sensing images is a difficult process, because most supervised classification methods require training data been sufficiently large (Chi et al, 2008). On the other hand, how to select training samples is also one of the most important factors in classifying the pixels (Paola and Schowengerdt, 1995). To achieve better accuracy, this factor is even more important than classification algorithm (Hixon et al., 1980).

Global Positioning System was used for forest mapping and training samples were prepared for two classes of forest and non- forest areas. Training samples were divided into two classes that one of them was used for classification and another for supervising classification accuracy (Table 1). It should be noted that fixed training samples were applied and there was not any change in training samples during research process. The forest map was prepared after image correction according to studied algorithms using ENVI4.5 software.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Training samples</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classification</td>
<td>Evaluation</td>
<td>Classification</td>
<td>Evaluation</td>
</tr>
<tr>
<td>Forest</td>
<td>626939</td>
<td>363567</td>
<td>384</td>
<td>221</td>
</tr>
<tr>
<td>Non forest</td>
<td>897634</td>
<td>534523</td>
<td>549</td>
<td>328</td>
</tr>
</tbody>
</table>

**Classifier algorithms in satellite imagery**
For forest mapping, support vector machine was used based on supervised classification, maximum likelihood, minimum distance, parallelepiped, mahalanobis distance, spectral angle mapping, spectral information divergence and binary codes.

**The maximum likelihood classification**
The maximum likelihood method is a statistical- supervised approach for recognizing the patterns. Probability of a pixel belonging to each class is calculated and then the pixel will be allocated the class with the highest probability (Munyati, 2004). This method is more accurate than other existing methods for classification. The quantities of variance and correlation values for various bands can be calculated in this method. In other words, it used to study the distribution of pixel values for a range of statistical probability associated with one of the groups in their sample variance and the mean vector and the variance of the spectral values of the correlation. In this method, the user must carefully follow the Gaussian normal distribution and it is more suitable for the classification of multi-spectral (Qiu et al., 2004).

**Minimum distance classifier**
This approach calculates the variety within each pixel to mean spectra of each class and assigns the pixel to a group that has the least distance to mean. In this method, after determining spectral values, the distance of unclassified pixel is compared with mean pixels and corresponding pixels is belonged to the class that is nearest to mean. Therefore, all pixels of each image are belonged to related classes and different classes will be separated completely. This type of classifier is mathematically simple and computationally efficient, but according to the theoretical basis is weaker than maximum likelihood (Tso and Mather, 2009).

**Parallelepiped classifier**
This method is executed by defining the parallelepiped-shaped boxes for each class. Parallelepiped boundaries for each section are determined by the minimum and maximum of pixels in given classes. This indicates that the point represents a pixel in the interior of each parallelepiped is placed or not. This method is fast...
and easy to run, but there are errors, especially when more than one pixel is inside or outside a parallelepiped (Tso and Mather, 2009). Not enabling to classifying some pixels is the main reason that this method is always not used.

**Mahalanobis distance**

Mahalanobis Distance is a method of image classification. This method is very similar to the minimum distance method but covariance matrix is used in this method, it is assumed that the band histograms are normal (Richards, 1999).

**Spectral Angle Mapper**

Spectral angle mapping is based on spectral classification method. This method is used when the data are calibrated respect to reflection and fairly will be insensitive to light (Lu and Weng, 2007).

**Spectral Information Divergence**

Spectral Information Divergence is spectral classification that uses measuring the divergence for matching the pixels. In this method, less divergence brings more likely of pixels. In addition, pixels with more divergence than threshold maximum are not classified (Du et al., 2004).

**Binary Encoding**

Classification method for binary codes is a very simple method for classification of satellite images. This method is based on the band been above or below the average spectral coding that will be located between zero and one. (Mazer et al., 1988).

**Support Vector Machine (SVM)**

SVM is a non-parametric statistical method and is a new method for classification of satellite imagery that can be used to extract land-use map (Jensen, 2005). This method works on the assumption that there is no information on how to distribute the overall data. Ability to optimize the use of training data is the brilliant advantage that enables this method to categorize data better than common methods (Mountrakis et al., 2011). This advantage reduces costs and increases speed (Guatieri et al., 1998).

In fact, a binary support vector machine is a binary classifier that separates two classes using a linear boundary and linear classifications are generalized depending on the family (Srivastava and Bhambhu, 2009). SVM classifies data were analyzed with passing a page (boundary line) and using all bands as well as applying an optimization algorithm. So, first the samples constituted class boundary will be found. In other words, some training points that have the minimum distance to the decision boundary can be considered as a support vector. Large dimension of data can make an appropriate situation for this method that they can easily be distinguished. The main goal of this algorithm is to find the maximum distance between two classes and thus enhance the classification accuracy while generalizing will be reduced as possible (Zhang et al., 2008). Essential element that distinguishes the support vector machine is following the rule processing that is known as Structural Risk Modeling (SRM). SVM minimizes classification errors on unsupervised data without prior assumptions while statistical techniques such as maximum likelihood consider similarity of data as known phenomenon (Mountrakis et al., 2011).

**Classification accuracy**

Knowledge of data accuracy is essential to use those data correctly. After classification of satellite images using the training samples that are not involved in the classification process, the accuracy of classification was conducted. In the present study, overall accuracy, Kappa coefficient, the producer accuracy, user accuracy, commission error (rows of each class in the error matrix) and Committee error (columns of each class in the error matrix) were used to verify the classification. The total accuracy is the ratio of the number of correctly classified pixels in a class to the total number of correctly classified pixels in all classes. It is achieved according to the following equation:

\[
OA = \frac{1}{N} \sum P_{ii}
\]

Equation 1;

In this equation; \( OA \) = overall accuracy, \( N \) = the number of experimental pixels, \( \sum P_{ii} \) = the sum of principal diagonal elements of the error matrix. Kappa coefficient can be used because of some downsides of total accuracy. Kappa
Kappa coefficient describes the similarity between what on land (land truth) and what on paper. (Richards, 1999). So, Kappa coefficient is calculated from Equation 2:

Equation 2: \[ Kappa = \frac{p_0 - p_c}{1 - p_c} \]

In the above equation; \( p_0 \) observed correctly, \( p_c \) expected agreement.
Producer accuracy is the probability that a class of image classification in the same category get placed on ground and user accuracy is the probability that a given class of land should be placed in the same class as the image classification that is calculated from the following equation:

Equation 3: \[ PA = \frac{ta}{ga} \times 100 \]
Equation 4: \[ UA = \frac{ta}{n} \times 100 \]

In equation 4; \( PA \) =the percentage of class a for producer accuracy, \( ta \)= number of pixels correctly classified as class a, \( ga \)= number of pixels in class a of ground truth, \( UA \) percent accuracy of class a for user accuracy, \( n \)= the number of pixels of class a.

Error matrix shows the similarity of each classified class with ground truth and it is defined as following equation:

Equation 5: \[ C_e = 1 - UA \]
Equation 6: \[ O_e = 1 - PA \]

\( C_e \)=Pixels that in fact are not belong to the corresponding class, but classifier considers them as a part of that class.
\( O_e \)= Pixels that are belong to the corresponding class, but they are classified as a part of other classes (Arkhi and Niazi, 2010)
ed by SME and Liserl software.

RESULT

The forest cover map of Biuret prepared by support vector machine, maximum likelihood, minimum distance, parallelepiped, Mahalanobis distance, spectral angle mapping, spectral information divergence and binary codes is given in Figures 2 to 9.
Figure 4. maximum likelihood
Figure 5. minimum distance
Figure 6. spectral angle mapping
Figure 7. parallelepiped
In present study, eight classifications based on supervised methods in Zagros forests had been analyzed using overall accuracy, Kappa coefficient, the producer accuracy, user accuracy, commission and committee error that the results are presented in table 2.

Table 2. Accuracy coefficients of eight classifications based on supervised methods in Bivareh forests

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Land use</th>
<th>producer accuracy</th>
<th>user accuracy</th>
<th>commission error</th>
<th>committee error</th>
<th>Kappa coefficient</th>
<th>overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>binary codes</td>
<td>Forest</td>
<td>90/13</td>
<td>47/79</td>
<td>51/21</td>
<td>9/87</td>
<td>0/4547</td>
<td>74/19%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>69/21</td>
<td>95/73</td>
<td>4/27</td>
<td>30/79</td>
<td>0/6059</td>
<td>83/29%</td>
</tr>
<tr>
<td>mahalanobis</td>
<td>Forest</td>
<td>90/18</td>
<td>59/59</td>
<td>40/41</td>
<td>9/82</td>
<td>18/82</td>
<td>84/08%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>81/18</td>
<td>96/41</td>
<td>3/59</td>
<td>18/54</td>
<td>0/6262</td>
<td>84/08%</td>
</tr>
<tr>
<td>distance</td>
<td>Forest</td>
<td>92/63</td>
<td>60/58</td>
<td>39/42</td>
<td>7/37</td>
<td>18/54</td>
<td>84/08%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>81/46</td>
<td>97/29</td>
<td>7/21</td>
<td>18/54</td>
<td>0/6262</td>
<td>84/08%</td>
</tr>
<tr>
<td>maximum likelihood</td>
<td>Forest</td>
<td>87/05</td>
<td>54/85</td>
<td>45/15</td>
<td>12/95</td>
<td>0/5403</td>
<td>80/09%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>77/95</td>
<td>95/14</td>
<td>4/86</td>
<td>22/05</td>
<td>0/1909</td>
<td>49/01%</td>
</tr>
<tr>
<td>minimum distance</td>
<td>Forest</td>
<td>98/21</td>
<td>32/12</td>
<td>67/88</td>
<td>1/75</td>
<td>0/1909</td>
<td>49/01%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>33/86</td>
<td>98/80</td>
<td>1/20</td>
<td>66/14</td>
<td>0/379</td>
<td>7/51%</td>
</tr>
<tr>
<td>parallelepiped</td>
<td>Forest</td>
<td>15/40</td>
<td>100</td>
<td>0</td>
<td>84/60</td>
<td>0/379</td>
<td>7/51%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>5/08</td>
<td>98/67</td>
<td>1/33</td>
<td>94/92</td>
<td>0/379</td>
<td>7/51%</td>
</tr>
<tr>
<td>spectral angle</td>
<td>Forest</td>
<td>85/71</td>
<td>51/41</td>
<td>48/59</td>
<td>14/29</td>
<td>0/1948</td>
<td>38/39%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>23/83</td>
<td>97/75</td>
<td>2/25</td>
<td>76/17</td>
<td>0/7069</td>
<td>88/65%</td>
</tr>
<tr>
<td>mapping</td>
<td>Forest</td>
<td>86/83</td>
<td>71/25</td>
<td>28/75</td>
<td>13/17</td>
<td>0/7069</td>
<td>88/65%</td>
</tr>
<tr>
<td></td>
<td>Non forest</td>
<td>89/22</td>
<td>95/66</td>
<td>4/34</td>
<td>10/78</td>
<td>0/7069</td>
<td>88/65%</td>
</tr>
</tbody>
</table>

DISCUSSION

Generally, Landsat 7 (ETM’), having seven spectral bands with a panchromatic band can record data in the wavelength range of the various electromagnetic spectrum. So, the sensor detects and provides detection phenomena well. Since the phenomena of surface have different reflections in different bands of the electromagnetic spectrum, ETM’ with these abilities enables to identify and separate bands from each other perfectly (Mahini et al., 2012). As satellite images can be produced in different time series, they can be used in order to assess forest cover and land use. This result is similar to those obtained by Katris (1990), Dontri (2003) and Tipanitat and netion (2003).

In this study, after the necessary corrections and original image preprocessing, data were categorized into eight different methods, i.e. data were evaluated base on geometric and radiometric errors. Eight supervised
classifications including SVM, maximum likelihood, minimum distance, parallelepiped, Mahalanobis Distance, spectral angle mapping, spectral information divergence and binary codes were compared and analyzed in this research. Results of statistical analysis showed that SVM is the most accurate method with Kappa coefficients 0.7069 and total accuracy 88.68 %. Recent researches have shown that SVM is more accurate than other methods of classification. The main feature of this method is the high ability to use fewer training samples to reach higher accuracy compared to other methods (Mantero et al., 2005 and Mountrakis et al., 2011). SVM has some advantages such as high speed of computation and appropriate values for optimization. So it can be suitable method instead of other methods. The order of accuracy for all methods is SVM, maximum likelihood, Mahalanobis Distance, minimum distance, spectral information divergence, binary codes, parallelepiped and spectral angle mapping, respectively. These results are similar to those obtained by Knudby et al (2010). They concluded that SVM was the best method among six applied methods.

Su et al, (2009) in a study using MISR data showed that using SVM and maximum likelihood methods even with two-thirds of the data can achieve the accuracy that it is be obtained by all data. In addition, the present results are generally confirming previous studies in which the SVM method for classification of satellite images is more accurate than other methods. One of the advantages of SVM is the classification maps with higher accuracy than other methods using few training samples. The results of this study can be used for forest cover mapping using support vector machine and evaluating the forestry and natural resources projects in areas with the same requirements.

REFERENCES


Karteris MA. 1990. The utility of digital thematic mapper data for natural resources


