COMPARING DIFFERENT SATELLITE IMAGE CLASSIFICATION METHODS: AN APPLICATION IN AYVALIK DISTRICT,WESTERN TURKEY.

Aykut AKGÜN[a,*], A.Hüsnü ERONAT[b] and Necdet TÜRK[a]- (aykut.akgun@deu.edu.tr)
[a] Dokuz Eylul University, Department of Geological Engineering, Building-A, İZMİR
[b] Dokuz Eylul University, Institute of Marine Sciences and Technologies, İZMİR

KEYWORDS: Classification, GIS, Image, Satellite, Landsat

ABSTRACT

The different satellite image classification methods were compared using the satellite images of the Ayvalık district located on the western coast of Turkey covering approximately 560 km^2. For this purpose, landuse classification of the investigation area was made by different supervised image classification procedures and the results were compared with one another. Landsat 7 ETM+ satellite image, IDRISI Klimanjaro image processing and the GIS package were used in this study. Of the classified images, the maximum likelihood method is found to be more applicable and reliable for the satellite image classification purposes. While the minimum distance method has given more reliable results than the linear discriminant procedures, the parellelpiped method is found to give the least reliable results compared to the other methods.

1. INTRODUCTION

Image classification is an important part of the remote sensing, image analysis and patern recognition. In some instances, the classification itself may be the object of the analysis. For example, classification of landuse from remotely sensed data produces a map like image as the final product of the analysis (Campbell 2002). 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). The analyst must select a classification method that will best accomplish a specific task. At present, it is not possible to state which classifier is best for all situation as the characteristic of each image and the circumstances for each study vary so greatly. Therefore, it is essential that each analyst understand the alternative strategies for image classification so that he or she may be prepared to select the most appropriate classifier for the task in hand.

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&Bauer 2002, Dean&Smith 2003, Pal&Mather 2003, Liu et al 2002) . These techniques are distinguished in two main ways as supervised and unsupervised classifications. Additionally, supervised classification has different sub-classification methods which are named as parellelpiped , maximum likelihood, minimum distances and Fisher classifier methods.These methods are named as Hard Classifier.

In this study, the Ayvalık district located on the western coast of Turkey (Figure 1) was selected as a study area covering approximately 560 km^2 for comparing the

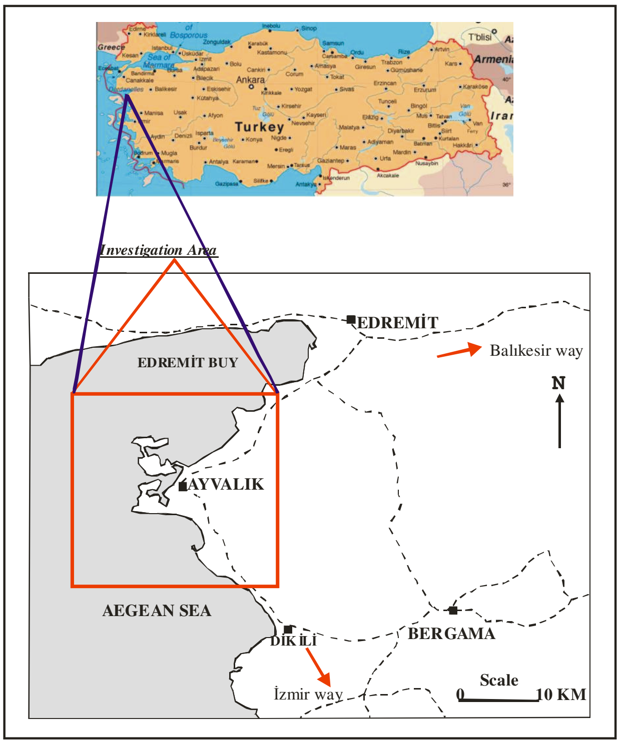

Figure 1. Location map of the study area

different satellite image classification methods. For this purpose, landuse classification of the study area was conducted by different supervised image classification procedures and the results were compared with one another. © Landsat 7 ETM+ satellite image acquised in 11.16.2001 and the IDRISI Klimanjaro image processing and the GIS package were used in this study. There are different image processing and GIS softwares using all around the world. A lot of them have the similar properties and capabilities for use remote sensing purposes. The IDRISI Klimanjaro image processing and the GIS package is one of the most useful and economic software of these image processing packages.

In this study, the Idrisi Klimanjaro was used for the different image classification chosen; Parellelpiped, Maximum Likelihood, Minimum Distance to Means and Fisher Classifier (Linear Discriminant Analysis) classifiers were used to determine which classifier is more effective and useful for this study purpose. To test these classifiers, a land use application was made in the study area. In this context, CORINE method was used for land use classification. 7 land classes were selected. Artifical Surfaces (Urban Areas), Agricultural Areas, Forests and Olive Trees, Wetlands and Water bodies (sea, lake) are the selected land classes according to CORINE land use method (CORINE, 1995). Bare Land class was added to the selected classes. For this purpose, PCA (Principal Component Analysis) composite image which was composed by PC2, PC4 and PC5 band combination was constituted because the each PC images reflects the most principle components on that band. Training sites have been digitized on screen and so a signature file to clasify the image have been made. After then, four image classifiers, Parellelpiped, Minimum Distance, Maximum Likelihood and Fisher, were applied to clasify the composite image respectively.

2. METHODOLOGY.

There is a consistent logic to all of the supervised classification routines in almost all image processing softwares, especially in IDRISI Klimanjaro, regardless of whether they are hard or soft classifiers (IDRISI Klimanjaro Guide, 2004). In addition, there is a basic sequence of operations that must be followed no matter which of the supervised classifiers is used. In this study the following sequence of operations were used.

1. Defining of the Training Sites.
2. Extraction of Signatures
3. Classification of the Image.

2.1 Defining of the Training Sites

The first step in undertaking a supervised classification is to define the areas that will be used as training sites for each land cover class. This is usually done by using the on-screen digitized features. For this purpose, band is chosen with strong contrast (such as a near infrared band) or a color composite image for use in digitizing. In this study, a color composite image which was made with PC2, PC4 and PC5 images was used. Generally, one should aim to digitize enough pixels so that there are 10 times as many pixel for each training class as there are bands in the image to classify. This should be made with at least two or three training sites but, the more training site is selected, the better results can be gained. However, this procedure assures both the accuracy of classification and the true interpretation of the results. In this context, each land class have been represented with two and three training sites.

2.2 Extracting of Signatures

After the training site areas have been digitized, the next step is to create statistical characterizations of each information. These are called signatures in Idrisi (Idrisi Klimanjaro Guide, 2004). With this module, categorization of infomation which of each pixels is possible. In this step, the goal is to create a signal (SIG) file for every informational class.The SIG files contain a variety of information about the land cover classes they describe. Each SIG file also has a corresponding SPF file that contains the actual pixel values used to create the SIG file. It is used only by HISTO histogram (HISTO) in displaying histograms of signatures). These include the names of the image bands from which the statistical characterization was taken, the minimum and mean values on each band, and the full variance /covariance matrix associated with that multispectral image band set for that class (IDRISI Klimanjaro Guide,2003).

2.3 Classification of the Image

The classification of the image is the third and the final step. This can be done with any of the hard or soft classifiers described below.

The Parellelpiped procedure (PIPED) is used for special pedalogic reasons only. Generally this procedure is not used for landuse mapping . When training sites are known to be strong, the MAXLIKE procedure is used (Richards 1995). However, if there are concerns about the quality of the training sites, the MINDIST procedure with standardized distances should be used (Richards 1995). The MINDIST module with the standardized distances option is a very strong classifier and one that is less susceptible training site problems than MAXLIKE. The FISHER Classifier can perform exceptionally well when there are not substantial areas of unknown classes and when the training sites are strongly representative of their informational classes (IDRISI Klimanjaro Guide 2004).

2.4 Genereal Properties of Classifiers

In this study, supervised classification classifiers have been used to classify the image of the study area for land cover classification. The parellelpiped, maximum likelihood, minimum distance and fisher (lineer discrimination) classifiers are used for this purposes.

The parellelpiped classifier is a very simple supervised classifier that is, in principle, trained by inspecting histograms of the individual spectral components of the available training data (Richards, 1995).

Whilst the parellepiped method is, in principle, a particularly simple classifier to train and use, it has several drawbacks. One is that there can be considerable gaps between the parellelpipeds, and the pixels in those regions will not be classfied. By comparision the minumum distance and maximum likelihood classifiers will label all pixels in an image, unless thresholding methods are used. Another limitation is that prior probabilities of class membership are not taken into account of;nor are they for minimum distance classification. Finally, for the correlated data there can be overlap of the parellelpipeds since their sides are parallel to the spectral axes (Richards 1995).

The Minimum distance classifier is based on training site data. This classifier characterizes each class by its mean position on each band (IDRISI Klimanjaro Guide 2004).

Minimum distance classifier is highly recommended in all image classification applications (Richards 1995). The classification is performed by placing a pixel in the class of the nearest mean. The minimum distance algorithm is also more attractive since it is a faster technique than the maximum likelihood classification.

The maximum likelihood classification is the most common supervised classification method used with remote sensing image data (Richards 1995). This classifier is based on Bayesian probability theory (IDRISI Klimanjaro Guide 2004).

The Fisher classifier conducts a linear discriminant analysis of the training site data to form a set of linear functions that express the degree of support for each class. It is more difficult to describe graphically (IDRISI Klimanjaro Guide 2004).

3. IMAGE CLASSIFICATION AND RESULTS

Landsat 7 ETM+ images of the Ayvalık, were classified to obtain the landuse map of the area using above mentioned four classifiers. Of these hard supervised classifiers used in this study, the maximum likelihood and Fisher are clearly the most powerful as they make more reliable classification. But these realiabilities can change according to purpose of the study. In order to make an image classification for landuse mapping, selection of the most proper image is the first step. For this, Landsat 7 ETM + images processed with IDRISI Klimanjaro GIS and image processing package. Firstly, all visible and infared bands were corrected atmospherically and geometrically. These images can be used for different interpretations such as geomorphological, geologcial, landuse and land cover mapping. We have seen that using only normal composite and false color composite images to interprate may be missleading in view of discrimination of objects on the image. To eliminate this discrepancy, visible and infrared bands have been processed by principal component analysis. After then, the composite images were made by different PCA bands. 20 PCA composite images were formed to chose the most approppriate images to classify them for landuse mapping. The composite image composed with PC2, PC4 and PC5 was used to map landuse features. In this image, land properties such as agricultural sites, vegetation cover, settlement areas, bare lands, wetlands and others was more clear than the same compositon of false color composite image. After digitizing of the training sites, the signature file from defined training sites was constituted. And then, this signature file were used for four classifier.

CORINE method was used (CORINE, 1995) to make landuse map. In application of this method, 7 landuse classes were determined. Water bodies (sea and lake), wetlands, forest, urban areas, agricultural areas, bare lands were selected firstly. As the olive trees covers the study area spreadingly, a seventh class was added to the above mentioned six classes.

The Parellelpiped, Minimum distances, Maximum likelihood and Linear discriminant classifiers were applied respectively. Results obtained from the classified images were compared and each of these images were controlled by field verification.

The parellelpiped classification results (Figure 6) was simple and have not reflected the real features on the land. For example, the urban sites on the map could not be identified, bare lands and olive trees could not be distinguished from each other. At some locations the vegetation cover has been seen as black color. Because of these anomalies, this classifier was not found proper enough for landuse mapping purpose.

The map derived using the minimum distances classifier (Figure 7) seemed more reliable than the map produced by parellelpiped method. In this map, settlement sites were selectable, borders of vegetation cover, agricultural areas and olive trees were more clear than the parellelpiped classifier map. The Maximum likelihood classification result was much better than the previous two maps. In the maximum likelihood map, barelands-olive trees discrimination could be seen clearly, boundries of agricultural areas and forest were more apparent than both minimum distances and parellelpiped maps (Figure 8). In this map, some wetlands areas were indicated by orange color scale. This was a cause of sedimantation on that locations.

The map that has been produced by application of the linear discriminant classifier (Figure 9) was more suitable than the minimum distances map and was less proper than maximum likelihood map. In this map, urban sites were more clear than both maximum likelihood, minimum distances and parellelpiped maps. Olive trees and barelands borders were also identifiable. However, forest and olive trees boundries were not clear according to ground truth studies.

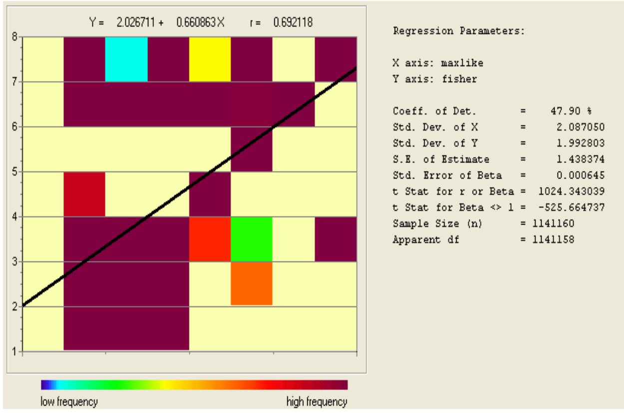

Figure 2. Regression of Maximum Likelihood and Linear discriminant classifier.

Classification results have compared with one another and regression analysis were made. The best corelation was obtained between the Maximum likelihood map and the minimum distance map (r= 0.79). Other regression results are r= 0.69, r= 0.52 and r= 0.76 for maximum likelihood-linear discriminant classifier, minimum distances-linear discriminant classifiers and maximum likelihood-parellepiped classifiers respectively. An interesting point related with these results is that the corelation between maximum likelihood and parellelpiped maps results have high corelation coefficient than maximum likelihood and linear discriminant classifiers corelation. This might be thingking usefullnes of parellelpiped map. But field studies have showed that parallelpiped map results do not reflect real properties on the land surface. It is thought that it may be due to the classification algorithm differences.

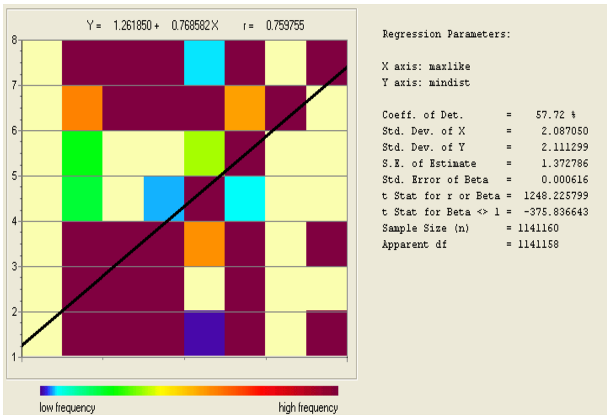

Figure 3. Regression of Maximum likelihood and Minimum distances classifiers.

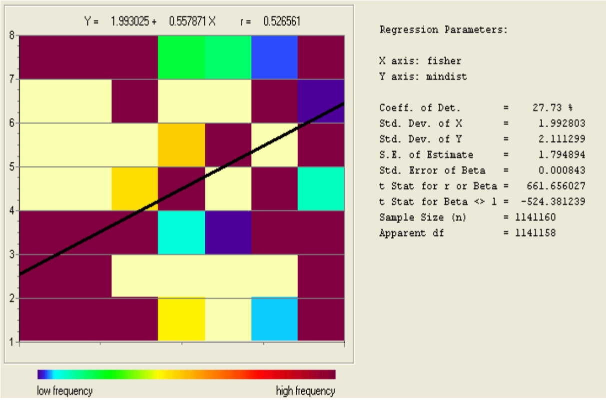

Figure 4. Regression of Linear discriminant and Minimum distances classifiers.

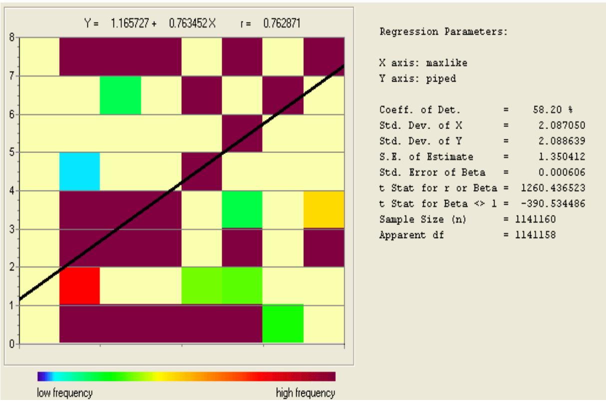

Figure 5. Regression of Maximum likelihood and Parellelpiped classifiers.

4. CONCLUSIONS

For more effective use of the satellite remote sensing, landuse managers should be aware of the limitations and advantages of satellite data and should chose from their avaible landuse mapping options accordingly. Remote sensing is especially proper for initial reconnaissance mapping and continued monitoring of landuse over large areas. In this context, techniques for improving the classification of landuse with satellite remote sensing data include the use of appropriate digital data. In order to achieve this task, selection of the most proper satellite image, band combination, and the classifier are very important. Additionaly, the image processing is important and different stages of it such as filtering of bands and principal component analysis should be applied before evaluation. All these points were applied to this study and it has been seen that maximum likelihood classifier was the most suitable classification method for landuse mapping purpose. Minimum distances classifeir was also determined as suitable as the maximum likelihood classifier.

4.1 References

Butera,M.K., 1983, Remote sensing of wetlands, *IEEE Transactions on Geoscience and Remote Sensing* GE-21,pp. 383-392

Campbell, 200?, *Introduction to Remote Sensing*,

CORINE Land Cover Technical Guide, European Commission, Luxemburg, pp. 21-53.

Dean, A.M and Smith, G.M., 2003, An evaluation of per – parcel land cover mapping using maximum likelihood class probabilities, *International Journal of Remote Sensing*, 24 (14), pp. 2905-2920

Ernst, C.L. and Hoffer, R.M., 1979, Digital processing of remotely sensed data for mapping wetland communities. LARS Technical Report 122079. Laboratory for Applications of Remote Sensing. Purdue University, West Lafayette, 119 pp.

Liu, X.H., Skidmore,A.K. and Oosten,V.H., 2002, Integration of classification methods for improvement of land-cover map accuracy, ISPRS Journal of Photogrammetry&Remote Sensing, 56, pp 257-268.

Lo, C.P. and Watson, L.J., 1998, The influence of geographic sampling methods on vegetation map accuracy evaluation in a swampy environment. *Photogrammetric Engineering and Remote Sensing* 64:1189-1200.

Ozesmi, S.L and Bauer, M., 2002, Satellite Remote Sensing of Wetlands, *Wetlands Ecology and Management*, 10, pp. 381-402

Richards, J.A., 1995, *Remote Sensing Digital Image Analysis:An Introduction*, Springer-Verlag, pp. 265-290.

IDRISI Klimanjaro, *2004, Guide to GIS and Image Processing Volume 2*, Idrisi Production, Clark Labs,USA, pp.57-82

4.2 Acknowledgments

The authors would like to thank to Dokuz Eylul University Scientific Research Fund for providing financal support for this project (Project no: 02.KB.FEN.052).

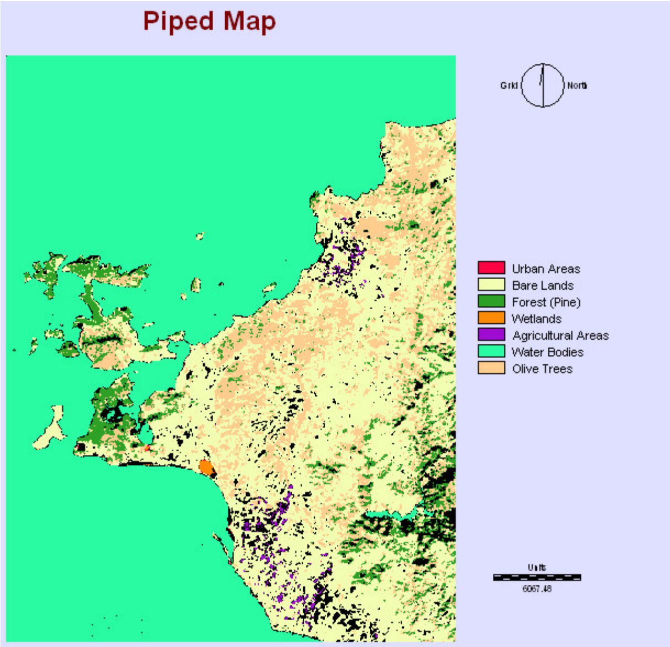

Figure 6. Parellelpiped classifier map

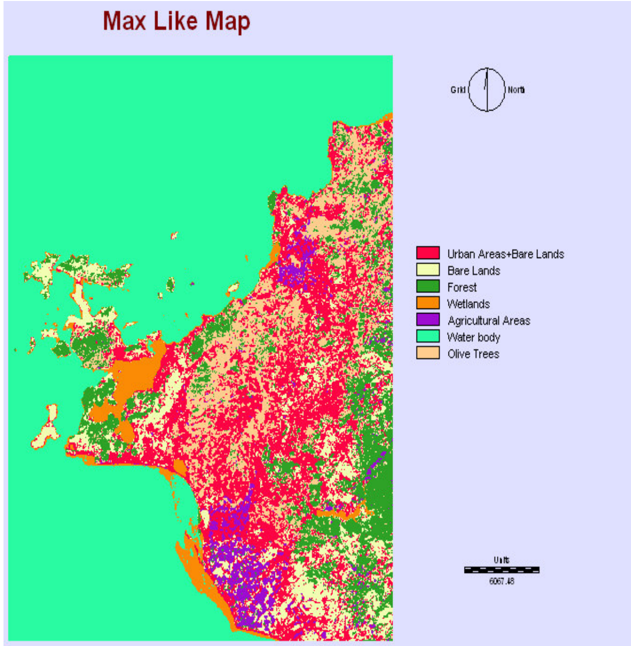

Figure 8. Mamimum Likelihood classfier map

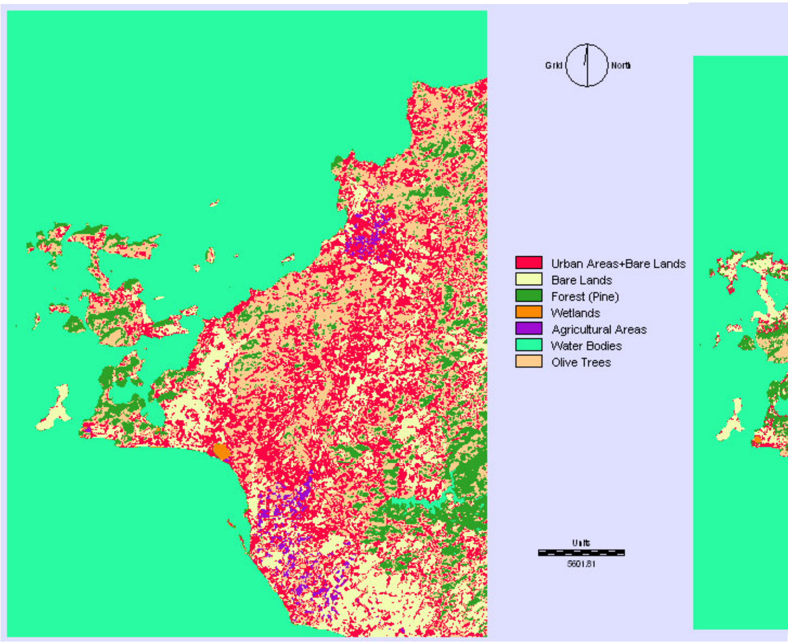

Figure 7. Minimum distances classifier map

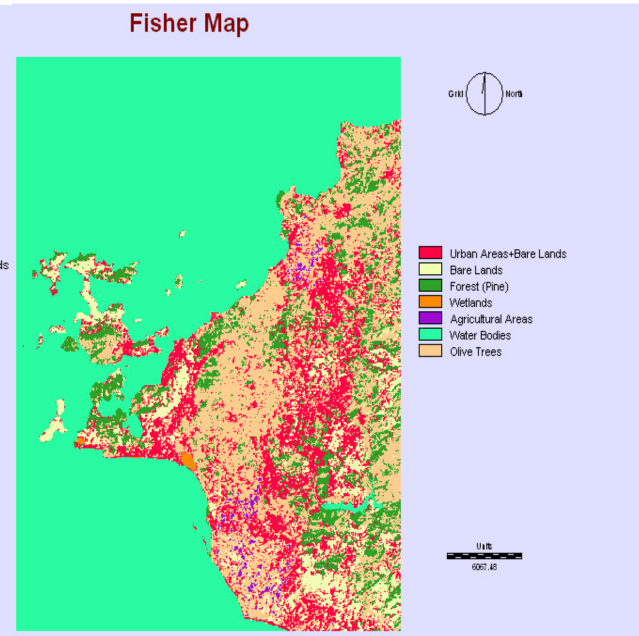

Figure 9. Fisher (Linear discriminant) classifier map