

Comparison between Sub-pixel Classifications of MODIS images: Linear Mixture Model and Neural Network Model

HASAN ROOSTA^{a,*} and RAHMATOLAH FARHUDI^a, MOHAMAD EBRAHIM AFIFI^b

^a Faculty of Geography, University of Tehran
North Kargar Avenue, Enqleb Sq.
TEHRAN, IRAN

^b Islamic Azad University, Larestan Branch
LARESTAN, IRAN

* hroosta@gmail.com

Abstract: This study employs sub-pixel classification methods to estimate crop acreage using low resolution satellite images. Linear mixture and Neural Network Methods examined the relationship between the value of Normalized Difference Vegetation Index (NDVI) in the context of Iran and crop acreage that reported by administrative office. The Moderate Resolution Imaging Spectroradiometer (MODIS) offers a unique combination of spectral, temporal, and spatial resolution compared to previous global sensors, making it a good candidate for large-scale crop type mapping. However, because of sub-pixel heterogeneity, the application of traditional hard classification approach to MODIS may result in significant errors in crop area estimation.

Fars province in south of Iran was selected as test zone, because of the cover type of the large majority of agricultural fields. Neural network model and Linear Mixture model were investigated and result in area fraction images (AFIs). The AFIs contain for each 250 m pixel the estimated area proportions occupied by the different cover types (crops or other land use). The algorithm was trained with both of reference data and in situ data which collected by GPS in Marvdasht District. For the major classes (winter wheat, maize and other crops) the obtained acreage estimates showed good agreement with the true values (NNR²≈90% and LMMR²≈85%). These methods seem attractive for wide-scale, regional area estimation in the countries that appropriate data are not available.

Key-Words: - Mixed pixel; neural network; Linear Mixture Model; Area fraction images; land cover; MODIS time series.

1 Introduction

It is important to be able to estimate the area of an agricultural field to manage agricultural production system. In order to operate effectively, the crop area estimation has to be automated. Since Iran is very vast, this task is too large to be carried out by investigators on site. A further complication is that on many of parcels that are in use several crops are grown in succession each year, making a regular observation necessary. To cope with these problems, we need an automated system based on satellite remote sensing to support investigators in their task. Since each remotely sensed image covers a large region, many observations can be made in a relatively cost-effective way. Furthermore, the observation can be performed regularly because the satellite passes over the same area on a frequent basis. It is because of these properties that an automated system based on satellite remote sensing is ideally suited for crop area estimation.

On the other hand, crop monitoring at regional or national level requires synoptic methods. It is not

practical to use detailed methods, as they are field data demanding. Remote sensing, due to its high temporal resolution data (e.g. NOAA-AVHRR and Terra-MODIS), could provide synoptic observations. But this data has low spatial resolution and the recorded radiometry is a mixed signal. The mixed pixel problem is a strong limitation for the use of low-resolution images in crop monitoring. This is especially a problem when agricultural farms are scattered or small. In order to improve the utility of these images, mixed signals must be disaggregated to the level of individual crops (land covers).

1.1 The mixed pixel problem

In every remotely sensed image, a considerable number of mixed pixels are present. A mixed pixel is a picture element representing an area occupied by more than one ground cover type. Basically, there are two situations in which mixed pixels occur. The first case concerns the pixels that are located at the edges of large objects like agricultural fields, for

instance. The second case arises when objects are imaged that are relatively small compared to the spatial resolution of the scanner. This can be long linear features such as rivers or highways, but also objects that are small in both dimensions such as farms or ponds, or even bushes in the sparsely vegetated semi-arid rangelands.

For a given scanner, the number of mixed pixels greatly depends on the landscape that is imaged. Irons et al. [5] reported proportions of probable mixed pixels in TM-images ranging from 29.6% for the category water to 68.3% for grass patches, while Schoenmakers [4] claimed that in some Mediterranean countries (also same condition in Iran), where the average field size is small, the proportion of mixed pixels can easily be as high as 30%. These figures indicate that mixed pixels have a significant influence on the information that can be derived. Classification of mixed pixels leads to errors that make the subsequent area estimation inaccurate. These errors are caused by the premise of classification that all pixels are pure, i.e. consisting of a single ground cover type, while in fact they are not. The Spectral confusion caused by mixing of ground cover types is outlined in Fig. 1.

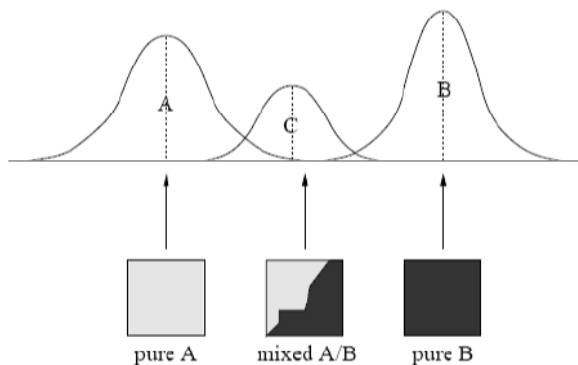


Fig 1: Spectral confusion caused by mixing of ground cover types.

The mixed pixel problem is not solved simply by increasing the spatial resolution. In general, the proportion of mixed pixels decreases as the spatial resolution becomes finer, for the smaller pixel size allows more pure pixels to be fit within the object boundaries. In some cases, however, the proportion of mixed pixels can actually increase because the finer detail resolves features not recorded before, thus introducing new spectral classes (Campbell [3]). For example, the image of a forested area, which seemed uniform at coarse resolution, may display individual trees of different species interspersed with open spaces at finer resolution (Woodcock and Strahler [6]). But even if the spectral classes remain the same and the proportion

of mixed pixels decreases, the classification results can still deteriorate (Markham and Townshend [8], Irons et al. [5]).

The main reason for this effect is that at finer resolutions the within-class variation increases as local differences in humidity, elevation, illumination, etc. become more apparent. Another reason is that the increase in spatial resolution usually is achieved at the expense of the spectral or radiometric resolution, because the reduction in received energy due to a smaller IFOV must be compensated for, e.g. by broadening the spectral band at which the reflectance is measured. A further disadvantage of fine spatial resolution is that the number of pixels can become very large, which adds to the costs of processing.

2 Study area and DATA

Because of its cover type of the large majority of agricultural fields, Marvdasht region located in Fars province in south of Iran was selected as test zone. Some random fields were selected and their positions were specified using GPS. To validate the use of linear mixture model with remotely sensed data, an ASTER image of Marvdasht was used to generate a land use map of the study area as reference data set (Figure 2).



Fig 2: The study area in south of Iran.

The following image data were used to carry out this research:

- One ASTER image (06 Jul 2005).
- Selected MODIS time series images, February-05 till July-05.
- A classified ASTER image (crop map).

- MODIS time series NDVI and simple ratio index for ASTER image.

The Normalized Difference Vegetation Index (NDVI) is a commonly used, space-observed measure for the amount of green vegetation. The current analysis started with the re-projection of images to the same co-ordinate system as used for the GPS data (UTM, Zone 39N-250 m (MODIS)). At the end, NDVI-composites were computed with the maximum NDVI-criterion. These images were considered completely free of missing values (cloud, snow, other noise). To have just the crop land area, ASTER image was masked by the ratio of Band 3 to Band 2 of the image. The mask defined by threshold which originates from in situ data. ASTER image was classified by two common classification methods: maximum likelihood and Support Vector Machine (SVM) for classification the ASTER image and comprise the results (Table 1).

Method	Overall Accuracy (%)	Overall Kappa (%)
Maximum likelihood	90.2	88
Support Vector Machine (SVM)	91.6	90.1

Table 1: Accuracy of the classification methods

Therefore, SVM classification was selected to produce crop map for this study. The reference crop map which produced by classifying ASTER image containing four main classes of wheat (green), barley (blue), maize (red) and orchards (cyan) and is shown in figure 3.

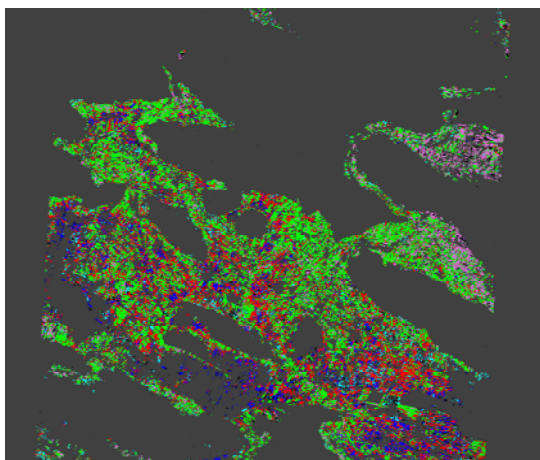


Fig 3: Area fraction image produced by crop map.

The reference crop map which produced by classifying ASTER image was not directly compatible with the low resolution NDVI-images. Therefore, it was transformed into a set of area fraction images (AFIs), one for each of the classes. The procedure for the AFI-creation is outlined in Fig. 4.

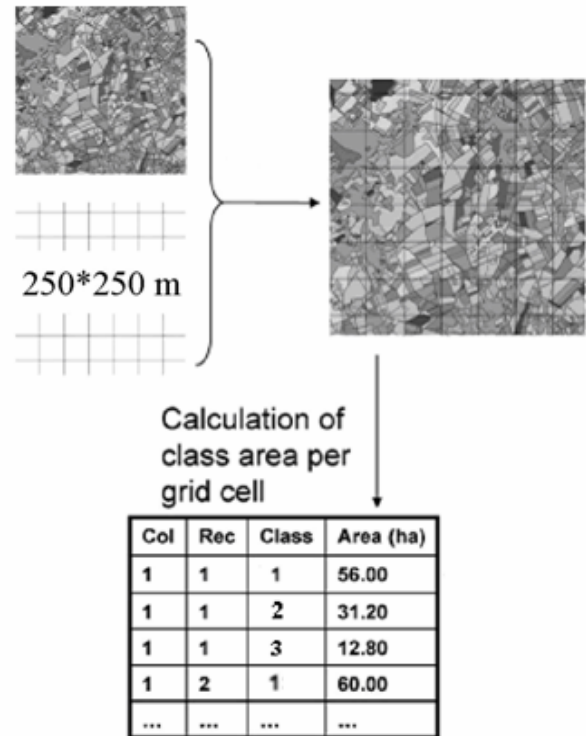


Fig 4: Creation of area fraction images by crop map.

These AFIs have the same 250 m resolution as the MODIS images and they give for each pixel the area fraction occupied by the considered classes (the fractions sum up to 1 per pixel). First, a 250*250 m² grid was created with the same spatial characteristics (projection, resolution, and framing) as the NDVI-images. This grid was superimposed over the crop map generated from ASTER image, and the area fractions of the 3 main classes in the test site (wheat, barley and maze) within each grid cell were computed.

The fraction images that are produced are shown in succeeding figures (Figure 5, 6, 7).

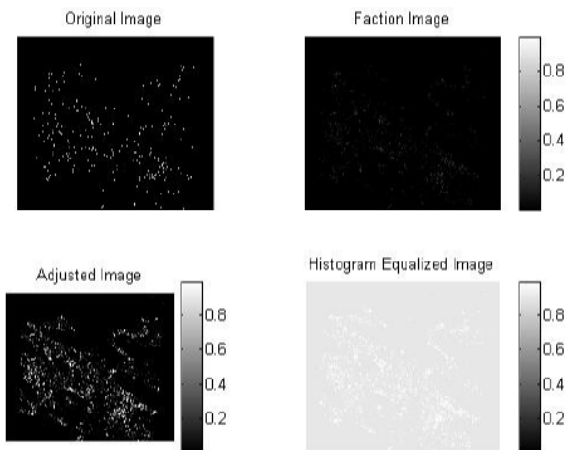


Fig 5: Area fraction images for maize that produced by crop map.

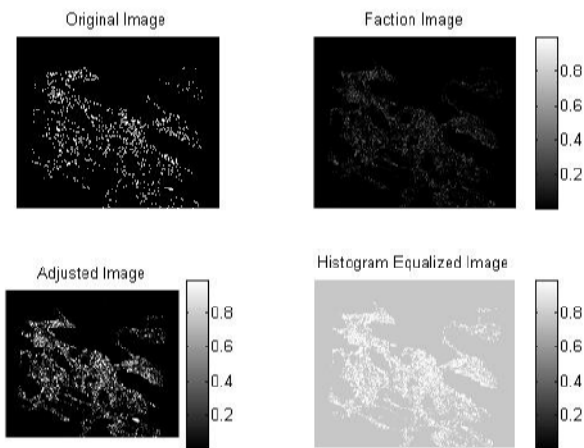


Fig 6: Area fraction images for wheat that produced by crop map.

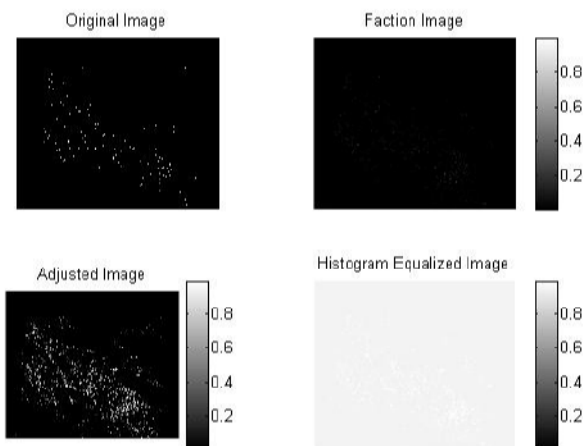


Fig 7: Area fraction images for barley that produced by crop map.

3 Area Fraction Methodology

The usual approach to carry out spectral mixture analysis is by modeling of spectral mixtures. Mixture modeling is the process of deriving mixed

signals from pure endmember spectra while spectral unmixing aims at doing the reverse, deriving the fractions of the pure endmembers from the mixed pixel [7]. Several models have been proposed to unmix pixels and determine proportions of their components. The more particular ones are linear, artificial neural network, probabilistic, geometric or geometric-optical and stochastic geometric models [9]. These models are comprised of known and unknown parameters. The known parameters are always the observed reflectance from the pixel and the pure spectra of the pixel components (or endmembers). The unknown parameters then will have to be determined by properly using known parameters.

Dennison and Roberts [10] describe linear mixture model as “essential tool for remote sensing vegetation analysis”. Therefore, numerous studies have tested and developed this approach and obtained strong correlation between the actual land covers and estimated land covers but In general, the neural network outperformed the linear mixture model compared with linear and/or parametric approaches. The results obtained have been used to describe land cover change, seasonal change in vegetation, fractional vegetation cover and regeneration after disturbance (Dennison and Roberts [10], Foody and Cox [11], Lobell and Asner [2], Sagardia [1]). Both of these methods consider the reflectance of a pixel as combination of two or more “pure” spectra called endmembers (classes or components) and report the fraction of each endmember in each pixel [10]. Hence, mathematically, the observed reflectance r_i for a pixel in band i will be;

$$r_i = f_1 a_{i,1} + f_2 a_{i,2} + \dots + f_c a_{i,c} + e_i$$

Where e is an error term, f is fraction of an endmember in a pixel, c is possible number of endmembers in the scene, a is the pure (or characteristic) spectra from the respective endmember. If we replace classes 1 to c with j , then the equation can be simplified as:

$$r_i = \sum_{j=1}^c f_j a_{ij} + e_i$$

Hence, for a multispectral image of n bands; $i = 1, \dots, n$, there will be n linear equations. In addition to these n equations, there will be another equation, which is called sum-to-unity constraint equation as:

$$f_1 + f_2 + \dots + f_c = 1$$

It states that the sum of component proportions for each pixel should sum to 1 (provided that none of the fraction is negative). Hence we have a system of linear equations, which can be solved in a number of

ways. With the matrix form we have the basic equation as:

$$Y = F * X$$

Matrix F contains the class area fractions for each pixel, matrix X is the pure class responses, and matrix Y is the mixed image observations. This formulation is commonly used for two different purposes: sub-pixel classification and "endmember" assessment.

Sub-pixel classification aims to assess the matrix F (class area fractions per pixel) from the satellite-registered Y and a priori known pure class responses X (in this context often termed 'endmembers'). An important practical limitation of this approach in linear mixture model is that the number of classes cannot exceed the number of available bands (plus 1 due to extra sum-to-unity constraint equation). This is achieved by solving the basic equation as:

$$F = Y * X' * (X * X')^{-1} = Y * A$$

The second application of these models oppositely tries to assess the matrix X with pure class responses (endmembers) from the image set Y and external knowledge on the area fractions F . As a solution, the basic equation is inverted as follows:

$$X = (F' * F)^{-1} * F' * Y$$

In this case the limitation is seldom problematic.

4 Experimental results

This study explored the potential of sub-pixel classification for regional crop area estimation. By area fraction images obtained through ASTER image, two methods which solve the problem for X are Neural Network model and Linear Mixture Method have been investigated.

Neural network model was the first method that has been investigated (Figure 8).

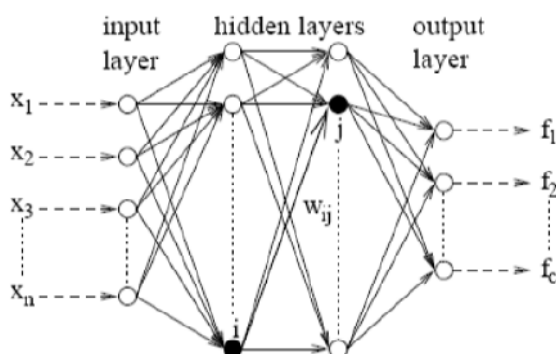


Fig 8: A multi-layer perceptron designed for the composition of mixed pixels.

The result will be pure class responses. In addition, the limitation of number of available bands is not a problem in neural network model. The sub-pixel classification was performed with a simple, three layer back-propagation neural network, with 4 nodes in the input layer (monthly NDVIs) and 3 nodes in the output layer (class area fractions). In the case of vegetation, the number of nodes in the intermediate (hidden) layer was set to twice the input nodes plus one (i.e. 9 in this case), Kavzoglu and Mather [12]. The hyperbolic tangent sigmoid as:

$$y(x) = \frac{2}{1 + \exp(-2x - 1)}$$

was used as transfer function for the hidden layer and the log-sigmoid function

$$y(x) = \frac{1}{1 + \exp(-x)}$$

for the output layer. As the log-sigmoid function scales its outputs between zero and one, all fraction estimates automatically remained between the physical bounds (0.0 to 1.0). The neural network was trained with a sample of 1000 pixels selected randomly from data set. After training (i.e. definition of the neural network weights), the neural network was applied on the entire image.

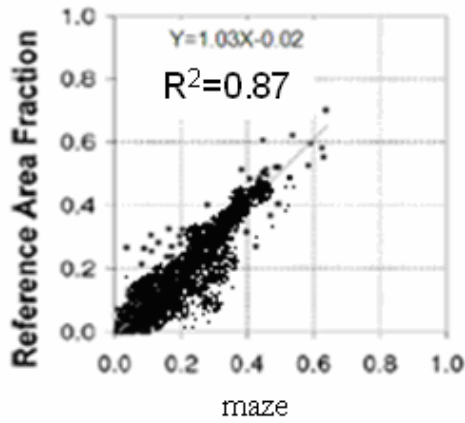
On the other hand we used Linear Mixture Model to discrimination the mixed pixels. The model that we used is:

$$F = Y * X' * (X * X')^{-1}$$

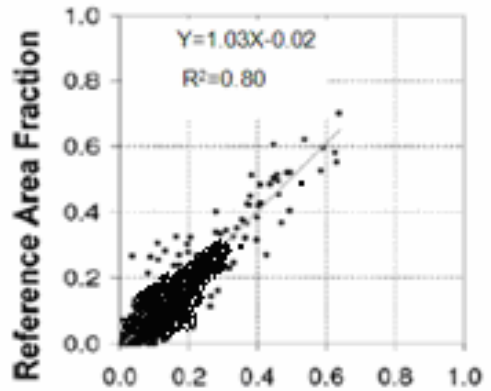
To compute the regressions and estimate the accuracy of the model, the result produced by neural network and reference data (obtained from ASTER image) were compared.

Fig. 9 and 10 display the scatter plots and regression for the three main crops produced with NN model and Linear Mixture Model.

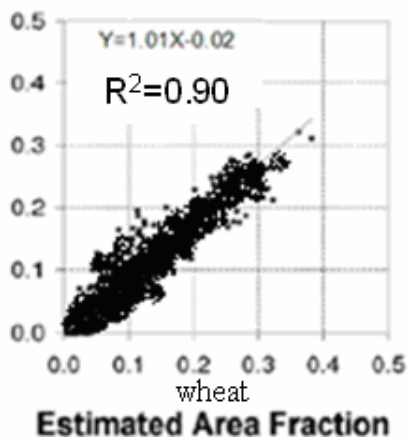
The R^2 value between neural network fractions and Reference data obtained from ASTER are 0.90, 0.87 and 0.75 for winter wheat, maize and barley respectively.



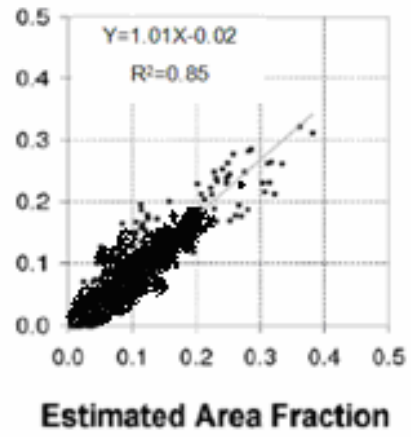
(a)



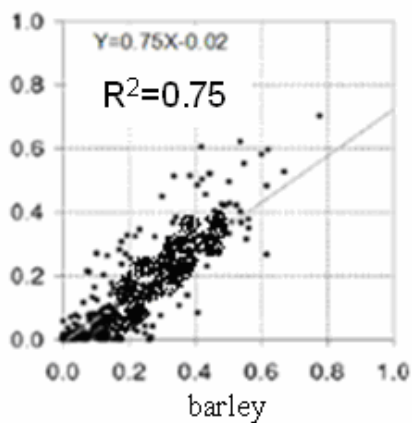
(a)



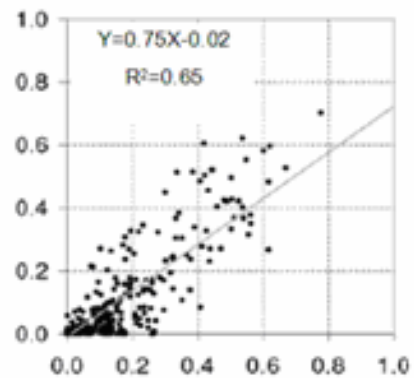
(b)



(b)



(c)



(c)

Fig 9: Scatter plots and linear regressions (REF produced by crop map = $a + b * EST$ by neural network) for the validation of the area fractions of maize, winter wheat and barley (a to c) estimated with the neural network model.

The R^2 value between linear mixture model fractions and Reference data obtained from ASTER are 0.85, 0.80 and 0.65 for winter wheat, maize and barley respectively.

Fig 10: Scatter plots and linear regressions (REF produced by crop map = $a + b * EST$ by linear mixture model) for the validation of the area fractions of maize, winter wheat and barley (left to right) estimated with the linear mixture model.

Scatter plots and linear regressions show that the accuracy of the neural network method for crop area estimation is high enough to have reliable data for making proper decisions. Due to the spectral

mixture of barley with other crops, the relevant R^2 value is unbiased estimation and the regression line does not approach to the 1:1 diagonal (*i.e.* $a \approx 0.0$ and $b \approx 1.0$).

5 Conclusions

We have evaluated the feasibility of MODIS time series images for estimating crop area proportion. The crop map was used as a reference to investigate the potential of sub-pixel classification of low spatial resolution satellite sensor imagery for regional area estimation. Area fraction images that were produced using ASTER image were used as input data for neural network model. The neural network approach does not experience the inherent limitation of the linear mixture model where the number of classes is restricted to the number of input variables. The results strongly suggest that neural network subpixel classification provide useful regional information about crop area estimation.

6 Acknowledgements

The MODIS time series images were kindly provided by Iranian Space Agency (ISA), and high resolution image (ASTER image) was obtained by Agricultural Jihad Organization of Fars Province in Iran.

References:

- [1] Sagardia, R., Use of subpixel classifier for wetland mapping: a case study of the Cuitzeo Lake, Mexico. Enschede, ITC, 65-78, 2005.
- [2] Lobell, D. B. and G. P. Asner, Moisture Effects on Soil Reflectance, *Soil Science Society of America Journal*, Vol. 66, 2002, pp. 722-727.
- [3] Campbell, J.B., *Introduction to remote sensing*. Taylor & Francis Ltd, London, 2nd edition, 1996.
- [4] Schoenmakers, R.P.H.M., Integrated methodology for segmentation of large optical satellite images in land applications of remote sensing. PhD thesis, Department of Informatics, University of Nijmegen, The Netherlands, 1995
- [5] Irons, J.R., B.L. Markham, R.F. Nelson, D.L. Toll, D.L. Williams, R.S. Latty, and M.L. Stauffer, The effects of spatial resolution on the classification of Thematic Mapper data. *International Journal of Remote Sensing*, VOL. 6(8), 1985, pp. 1385-1403.
- [6] Woodcock, C.E. and A.H. Strahler, The factor of scale in remote sensing, *Remote Sensing of Environment*, VOL. 21, 1987, pp. 311-332.
- [7] Van der Meer, F., *Spatial Statistics for Remote Sensing*, Kluwer Academic Publishers, first edition, 1999.
- [8] Townshend, Land cover classification accuracy as a function of sensor spatial resolution. In Proceedings of the 15th International Symposium on Remote Sensing of Environment, 1981, pp. 1075-1085.
- [9] Hartemink, R., World database, <http://www.ngw.nl/indexgb.htm>, 2002. Accessed 02-2007.
- [10] Dennison, P.E. and D.A. Roberts, The effects of vegetation phenology on endmember selection and species mapping in Southern California chaparral, *Remote Sensing of Environment*, VOL. 87, 2003, pp. 295-309.
- [11] Foody, G.M. and D.P. Cox, Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions, *International Journal of Remote Sensing*, VOL. 15, 1994, pp. 619-631.
- [12] Kavzoglu, T. and Mather, P.M., The use of back-propagating artificial neural networks in land cover classification, *International Journal of Remote Sensing*, VOL.24(23), 2003, pp. 4907-4938.