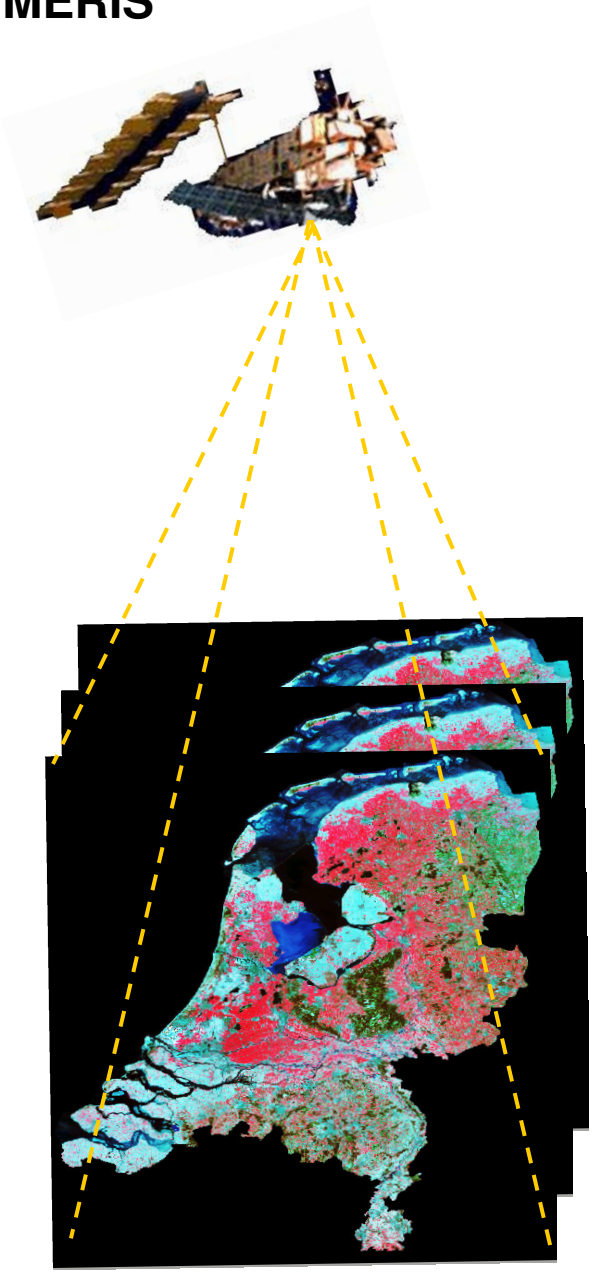
Centre for Geo-Information

Thesis Report GIRS-2005-01

Multitemporal Land Cover Classification in The Netherlands with MERIS

Teshome Eshete Bedlu



March 2005

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Multitemporal Land Cover Classification in The Netherlands with MERIS

Teshome Eshete Bedlu

Registration number 740824-229-080

Supervisors:

Dr. Ir. J.G.P.W. (Jan) Clevers
Mr. R. (Raul) Zurita Milla

A thesis submitted in partial fulfillment of the degree of Master of Science at Wageningen University and Research Centre, The Netherlands.

March 2005
Wageningen, The Netherlands

Thesis code number: GRS- 80337
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing
Thesis Report: GIRS-2005-01

.....to my beloved mother, Zewudye.....

Acknowledgments

In the name of God, most Gracious most Merciful. Praise be to the Almighty God sustainer of the world. Many thanks to my God for giving opportunity to me to broaden my academic horizons.

I would like to express my gratitude to all those people who have spent their time directly or indirectly to assist me in this work.

First and for most, my greatest thanks goes to my supervisors Dr. Jan Clevers and Mr. Raul Zurita for their close guidance, valuable criticism, advice, and support in completion of the research and writing the thesis report. Their professional guidance and supervision helped me to complete the work with its present form.

I would like to extend my thanks to my employer, the Ethiopian Agricultural Research Organization (EARO) for giving me the study leave. My thanks also will go to the Netherlands Minister for Development Co-operation via the Academic Program of the Netherlands Fellowship program (NFP-AP) for covering the full cost of my study.

I will never forget the people I have met here in the Netherlands. In particular my MSc room mates: Taye, Worku, Angela, Jie, Elisa, Tine, Lucie, and Hailu. I spent with them great and enjoyable time full of humors. They have given not only a good class environment with their love and jokes, but also they have been first consultants when having technical problems. I will always miss them.

Off course, I am very grateful for their support to my brothers Tilahun, Girma, Terefe, and Getachew. Their voice on telephone functioned as a source of inspiration to complete this research on time.

Last but not least, I want to convey my deepest and heartfelt thanks to my special friends Fekrete Guda, Yonas Alemu, and Abebech fikre for making my stay in the Netherlands enjoyable.

Abstract

The MEdium Resolution Imaging Spectrometer (MERIS) satellite data of 16th April, 14th July, 2003 and the combined image created by layer stacking the images of these two dates were used to produce land cover maps of the major land cover types present in the Netherlands. Classification accuracy for the single dates and the multitemporal MERIS image were compared using testing dataset collected from the aggregated Dutch land use database (LGN4) as a reference. To do the land cover classifications different classification methods were used: maximum likelihood (MLH), minimum distance to mean (MDM) and decision tree (DT). The latter classification method was performed using CART (Classification and Regression Tree) software. The classification accuracy achieved with these classifiers are presented and compared in this study. Finally, the effects of the training dataset size, the splitting criterion, and the data dimensionality on the decision tree classifier were also examined.

The overall classification accuracy achieved by the MLH and DT classifiers for the April image (69.76% and 68.56%, respectively) and for the combined image (70.96% and 70.84%, respectively) were almost the same. The accuracy obtained by the MLH classifier for the July image (71.11%) was higher than for the DT classifier (67.13%). The classification accuracies of the MDM for all the images were smaller than for both MLH and DT classifiers. The multitemporal classification by the DT and MDM classifiers were better than monotemporal classification with the same classifiers. The results indicate that classification accuracy of the DT increases with the training dataset increase up to a certain point. The splitting rule, Gini, offered the best overall classification accuracies in all the training dataset sizes. In this study, the data dimensionality has no effect on DT classifier. The study also indicates the ability of the CART model of the DT classifier to weight the importance of the discriminating variables in classification, thereby ignored redundant variables.

Keywords: MERIS, CART, MLH, DT, MDM, Land cover classification, Multitemporal, Remote sensing

Abbreviations

MERIS	MEdium Resolution Imaging Spectrometer
MLH	Maximum Likelihood
MDM	Minimum distance to mean
DT	Decision tree
CART	Classification and regression tree
LGN	Dutch land use database
TOA	Top of atmosphere
NIR	Near infrared
ENVISAT	ENVIronment SATellite
FR	Full resolution
RR	Reduced resolution
RD	Dutch national coordinate system
MODIS	Moderate Resolution Imaging Spectroradiometer
CORINE	Coordination of information on the environment
TM	Thematic mapper
NOAA	National Oceanic and Atmospheric Administration
AVHRR	Advanced Very High Resolution Radiometer

Table of contents

List of figures

List of tables

1. Introduction

1.1 Background

Land cover classification and monitoring using remotely sensed images have been widely applied for its effectiveness in time, labor and cost; particularly for studies at regional scales (Defries and Chan, 2000; Mucher et al., 2000). Various remotely sensed data ranging from high to low spatial resolution are used for land cover/change studies. Medium spatial resolution sensors, which have a pixel size about 250 – 300 meter, are appropriate for mapping of large areas since not too many images are needed and the images are not too coarse (Verstraete et al., 1999).

With increasing frequency, remotely sensed data sets have been used to classify land cover for monitoring large areas such as Europe (Clevers et al., 2003; Hansen et al., 2000). Earlier land cover mapping was based on the visual interpretation of Landsat-TM and SPOT-XS hard copies at a landscape level in the CORINE land cover project by producing an ecological legend (Bossard et al., 2000). Besides this, digital classification of the same kind of images is also used to map agricultural land cover at national level (Thunnissen et al., 1992; Fuller et al., 2002). Even though both approaches are well known, they use high spatial resolution images that make them too expensive and very time consuming for application at the European scale (Clevers et al., 2003).

Other approaches of land cover mapping are the use of coarse spatial resolution sensors such as the NOAA_AVHRR (Friedl et al., 1999; Mucher et al., 2000). However, according to Clevers et al. (2003) the use of coarse scale AVHRR imagery is limited for monitoring purposes due to the fine scale at which most land cover changes take place in Europe.

Medium spatial resolution images like the one from MERIS and MODIS can bridge the gap between high and low spatial resolution sensors for multiscale land cover assessment (Addink, 2001). MERIS pixel area is more than ten times smaller than an AVHRR pixel. That means more detailed information can be obtained from MERIS imagery.

As shown by Verstraete et al. (1999), the MERIS spatial resolution of 300m (full resolution) is sufficient to monitor heterogeneous land cover at continental and global scales.

Therefore, MERIS could be an interesting sensor for land cover mapping at regional scale. This study will illustrate the use of MERIS level 1b full resolution images for multitemporal land cover classification in the Netherlands. Land cover mapping using remotely sensed images is based on the classification of individual pixels or groups of pixels with similar spectral responses or spectral signatures (Zhan, 2003). Ideally, pixels are expected to be more or less grouped in the multispectral space in clusters corresponding to different land cover types (Keuchel et al., 2003). In remote sensing literatures there are different classification algorithms being used to group pixels into similar categories. One of these algorithms is the maximum likelihood (MLH) classifier, which is the most common technique presented in the literature (Erbek et al., 2004; Foody et al., 1992). The MLH was selected as a conventional classifier to perform multitemporal land cover mapping. The MLH approach has the limitation of assuming that the members of each class follow a Gaussian frequency distribution in the feature space and this is sometimes not fully true with remotely sensed data. The minimum distance to mean (MDM) classifier was also selected as another conventional classifier. The decision tree (DT) classifier, which is not yet tested with MERIS data, was chosen as an alternative to the traditional land cover classification methods. Unlike the MLH and MDM, decision trees are strictly non-parametric and do not require assumptions regarding the normal distribution of the input data. In addition, they have the ability to handle non-linear relationships between features and classes. They can also accept a wide variety of input data including non-remotely sensed ancillary data in the form of continuous and /or categorical variables. In contrast to neural networks, decision trees can be trained quickly, and are rapid in their execution. The analyst can easily interpret a decision tree. It is not a 'black box', like the neural networks, the hidden working groups of which are concealed from view (Pal and Mather, 2003; Friedl and Brodley, 1997).

The study also involves a comparison among (MLH), (MDM) (both are parametric methods) and (DT) (non-parametric methods) classifiers. The last classifier can be performed with the CART (Classification and Regression Tree) analysis software. Overviews of all the classifiers are given in chapter 2. A more detailed description of the objectives and purpose of this thesis can be found in the following section.

1.2 Problem definition

Preliminary classification of land cover at national scale using a single MERIS image gave promising results. However, some classes could not be well distinguished (Clevers et al., 2003). The spectral reflectance of vegetation changes with seasons, and it is also different for each vegetation type (Mucher et al., 2000). Different vegetation types often show similar spectral responses within a single image at one moment. Because of this spectral similarity, it is difficult to separate them. In general there is a significant temporal variation in the reflectance of the same object by plant growth (Yang et al., 2003). As a result the use of more than one image becomes crucial to get a proper land cover classification.

1.3 Research objectives

The general objective of this research is to study the advantage of multitemporal land cover classification by using full resolution MERIS images**.**

1.4 Detailed objectives

- To classify full resolution MERIS images using "Classical" classifiers like MLH and MDM
- To classify full resolution MERIS images using Decision Tree (DT) classifiers
- To study the effect of training size, data dimensionality, and splitting rules on the performance of a decision tree classifier algorithm
- To compare the classification accuracies between monotemporal and multitemporal full resolution MERIS images
- To compare the classification accuracies among MLH, MDM and DT classifiers
- To study which classes have high separablity

1.5 Thesis outline

The organization of this report is as follow:

Chapter 1 gives the introduction to the main subject, the problem definition, the main objective and the detail objectives of the thesis.

Chapter 2 reviews the most relevant literature related to the classifiers used in this study.

Chapter 3 explains the methodological aspects of this study.

Chapter 4 illustrates the results obtained from the analysis of data and discussion of the results.

Chapter 5 includes conclusions obtained from the results and recommendations.

2 Literature review

2.1 Land cover classification and remote sensing

The effective management and use of land resources requires knowledge of the properties and spatial distribution of these resources. The rapid evolution and increasing number of applications of remote sensing methods in the last 20 years have shown that such methods are becoming more widely accepted for resource survey, especially for the observation of land cover (Pal and Mather, 2003).

Land cover, i.e. the composition and characteristics of the land surface element, (Cihlar, 2000) is a fundamental environmental information. It is an important determinant of land use and thus of the value of that land to the society. Land cover mapping and classification is a product of the development of remote sensing. This is because 'viewing' large areas repeatedly is necessary for acquiring information about land cover. For the same reason, land cover mapping has been perhaps the most widely studied problem employing satellite data (Cihlar, 2000).

Multispectral and multitemporal properties of remotely sensed data are of great importance for land cover classification. Accurate classification of land cover from remotely sensed data is essential information for agricultural and forest monitoring, ecological monitoring of vegetation communities, land cover change detection, resource management and planning, and policy purposes (Brown et al., 2003).

A considerable body of work related to classification of land use/land cover categories using remotely sensed data is reported in the literature (e.g., Cihlar, 2000; Loveland and Belward, 1997; Mucher et al., 2000).

2.2 Multitemporal image classifications

A monotemporal image analysis, which relies on one image obtained at a single point in time, is usually used for classification of different vegetation types. However, sometimes it does not work well for classification of vegetation species types which show similar spectral signatures at that date (Paxlenney and Woodcock, 1997). Therefore, the use of multitemporal images can help to improve the classification accuracy of very similar vegetation types. Classification of landscapes characterized by a seasonal variation of vegetation features such as deciduous forests and

agricultural fields show improved accuracy when obtained from multidate imagery compared to single date imagery (Mucher et al., 2000).

Oetter et al. (2000) found that multitemporal data sets consisting of different dates from a single year provide significant improvements in accuracy of vegetation classification. In addition to this, classification accuracy is significantly improved by using multiseasonal imagery (Jeon and Landgrebe, 1999).

2.3 Classification methods

Different types of classifiers have been used for land cover mapping over large areas. Techniques range from unsupervised clustering and parametric supervised algorithms (like maximum likelihood) to non parametric and machine learning algorithms such as decision tree classifier and neural network (e.g. Loveland and Belward, 1997; Defries and Townshend, 1994; Hansen et al., 2000; Brown et al., 2003; Erbek et al., 2004) can be mentioned.

Generally classification procedures can be divided into two main broad categories based on the method used: Supervised and unsupervised classification.

The decision tree, maximum likelihood, and minimum distance to mean classifiers are supervised classification methods that were used in this study for classification. A brief summary of the properties of each classifier is given in this section.

2.3.1 Decision tree classifiers

Decision tree classifiers have not been as widely used within the remote sensing community as either the statistical or the neural methods (Pal and Mather, 2003) since it is a relatively new technique, which was developed about 20 years ago (Breiman et al., 1984). Recently, decision trees are increasingly being used for the analysis and classification of remotely sensed digital imagery particularly with regard to land cover classification at continental to global scales. DT has been used successfully for classification of multispectral (Friedl and Brodley, 1997; Hansen et al., 1996) and hyperspectral imagery (Lawrence and Labus, 2003). In addition, ancillary data can be added to the DT classification procedure to increase the final accuracy (Lawrence and Wright, 2001), and DT can also be used for change detection analysis (Rogan et al., 2003).

A decision tree is defined as a classification procedure that recursively partitions a data set into smaller subdivisions on the basis of a set of tests defined at each branch in the tree (Friedl and Brodley, 1997). Unlike conventional statistical and neural/connectionist classifiers which use all available features simultaneously and make a single membership decision for each pixel, the decision tree uses a multi-stage or sequential approach to the problem of label assignment. The labeling process is considered to be a chain of simple decisions based on the results of sequential tests rather than a single, complex decision. Sets of decision sequences form the branches of the decision tree, with tests being applied at the nodes (Pal and Mather, 2003). The node containing all of the data before any splitting is called the root node. A node from which the data can be split again is called a mother node. In all other cases, subsets are known as terminal (leave) nodes, and the subsequent split nodes are termed child nodes (Figure 1). Each node in a decision tree has only one mother node and two or more descendent nodes.

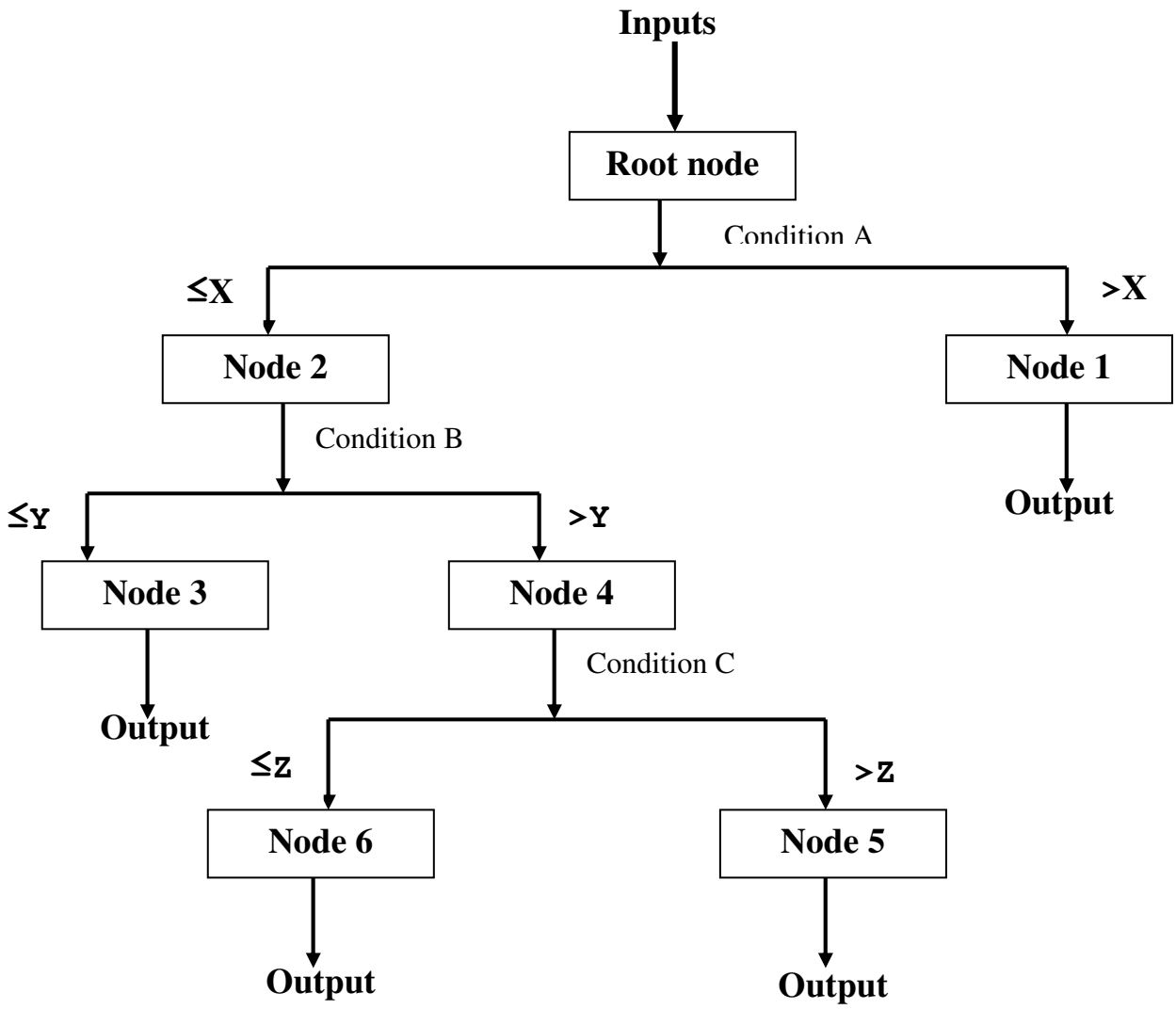


Figure 1. General Structure of a data mining decision tree

Figure 1 shows the general architecture of a DT model. The decision tree starts from the root node and applies the condition (A) to split the input data. If the input data, in the condition (A), is larger than the threshold (X), it is classified to node 1; otherwise, the input data is classified to node 2. For example, if the reflectance of waveband (A) in the remote sensing data is larger than the split point (X), the data is assigned to 1; otherwise to node 2. Then, if the reflectance of waveband (B) in the data at node 2 is less than or equal to the split point (Y), the data is assigned to node 3; otherwise to node 4. This procedure is repeated until all input data have been assigned to one of the target categories (output) (Yang et al., 2003).

More commonly, the splits defined at each internal node of a decision tree are estimated from training data. The techniques used for this work are called learning

algorithms. They require high-quality training data from which relations among the features and classes present within the data are "learned". Therefore, a set of training samples representative of the population to be classified must be available to construct an accurate decision tree (Friedl and Brodley, 1997). The decision tree classification algorithms can be distinguished according to whether a uniform or a heterogeneous set of algorithms is used to estimate the splits at internal nodes. Such algorithms are described as having homogeneous or heterogeneous hypothesis spaces, respectively. Univariate and multivariate decision trees are considered as homogeneous decision trees while a hybrid decision tree is a heterogeneous decision tree.

2.3.1.1 Univariate decision tree

In a univariate decision tree, the decision boundaries at each node of the tree are defined by the result of a test applied to a single feature that is evaluated at each internal node (Swain and Hauska, 1977). On the basis of the test outcome, the data are split into two or more subsets. Each test is required to have a discrete number of outcomes. A univariate decision tree classification proceeds by recursively partitioning the input data until a leaf node is reached, and the class label agreed with the leaf node is assigned to the observation. The value of the decision boundaries in a univariate decision tree are estimated empirically from the training data. In the case of continuous data, a test of the form $Y_1 > C$ is performed at each internal node of the decision tree, where Y_1 is the feature in data space and C is a threshold estimated from the observed feature Y_1. The value of C is estimated by using some objective measure that maximizes the dissimilarity or minimizes the similarity of the descendant nodes, using one feature at a time. As each test in a univariate decision tree is based on a single feature, it is restricted to a split through the feature space that is orthogonal to the axis representing the selected feature (Pal and Mather, 2003; Friedl and Brodley, 1997). For this study the univariate decision tree was used because of the availability of the software.

2.3.1.2 Multivariate decision tree

Multivariate decision trees are similar to univariate decision trees, but the splitting test at each node may be used on more than one feature of the input data. Specifically, a set of linear discriminant functions are estimated at each interior node of a

multivariate decision tree, where the coefficients for the linear discriminant function at each node are estimated from training data (Friedl and Brodley, 1997). The splitting test at each has the form $\sum_{i=1}^{n} a_i x_i \leq c$, where $\boldsymbol{x}_i$ represents a vector of measurements on the n selected features, $\boldsymbol{a}$ is a vector of coefficients of a linear discriminant function, and $\boldsymbol{c}$ is a threshold value.

2.3.1.3 Hybrid decision tree

A hybrid decision tree is a decision tree where different classification algorithms may be used in different subtrees of a large tree. The learning algorithms used to estimate a hybrid tree allow different splitting methods to be applied within different subtrees of larger decision trees (Friedl and Brodley, 1997).

2.3.2 Maximum likelihood classifier

The maximum likelihood (MLH) classifier is based on the assumption that the members of each class are normally distributed in the feature space. The MLH classifier is a pixel based method and can be defined as follows:

A pixel with an associated observed feature vector $\mathbf{X}$ is assigned to class $\boldsymbol{C}_j$ of $\boldsymbol{N}$ classes if

$$\mathbf{X} \in \boldsymbol{C}_j \text{ if } \mathbf{h}_j(\mathbf{X}) > \mathbf{h}_k(\mathbf{X})$$

For all $j \neq k$, with $j, k = 1,\ldots,\boldsymbol{N}$.

For a multivariate Gaussian distributions $\mathbf{h}_k(\mathbf{X})$ is given by:

$$\mathbf{h}_k(\mathbf{X}) = \ln(p(\boldsymbol{C}_j)) - {}^1/_2 \ln \left|\boldsymbol{\Sigma}_k\right| - {}^1/_2 (\mathbf{X} - \mathbf{M}_k)^t \boldsymbol{\Sigma}_k^{-1} (\mathbf{X} - \mathbf{M}_k),$$

Where $\mathbf{M}_k$ and $\boldsymbol{\Sigma}_k$ are the sample mean vector and sample covariance matrix for class k, and $\mathbf{h}_k$, the discriminating function.

Implementing the MLH classifier involves the estimation of class mean vectors ($\mathbf{M}_k$) and covariance matrices ($\boldsymbol{\Sigma}_k$) from a training dataset chosen from known examples of each particular class. MLH evaluates the variance and the covariance of the trained spectral response patterns when classifying an unknown pixel. Based on this value, the MLH classifier evaluates the membership probability of an unknown pixel for class j. The pixel is assigned to the class for which it has the highest membership probability value (Lillesand and Kiefer, 2000).

2.3.3 Minimum Distance to Mean classifier

The minimum distance to mean (MDM) classifier calculates the mean spectral value in each band and for each category, and relates each mean value by a vector function. MDM classifier computes the distance between the value of a pixel of unknown identity and each category mean in the feature space to assign the unknown pixel to the closest class. It is mathematically simple and computationally efficient, but it has certain limitations. Most importantly, it is insensitive to a different variance in the spectral response data. Because of such problems, it is not widely used in applications where spectral classes are close to one another in the measurement space and have high variance (Lillesand and Kiefer, 2000).

In a minimum distance to mean classifier, suppose we have nc known class centers $\mathbf{C} = \{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{nc}\}$, $\mathbf{C}i$, i = 1, 2, …, nc is the grey-level vector for class i.

$$\mathbf{C}_i = \begin{cases} \{DNi1, DNi2, \ldots, DNinb\}^T & \text{in digital number form} \\ \{ri1, ri2, \ldots, rinb\}^T & \text{in spectral reflectance form} \end{cases}$$

In a general form, an arbitrary pixel with grey-level vector $g = (g_1, g_2, \ldots, g_{nc})$, is classified as $\mathbf{C}_i$ if

$d(\mathbf{C}_i, g) = \min (d(C_{i1},g_1), d(C_{i2},g_2),\ldots, d(C_{inc},g_{nc}))$ where d is distance.

As an example, we show a special case in Figure 2 where we have 3 classes (nc = 3) and two spectral bands (nb = 2)

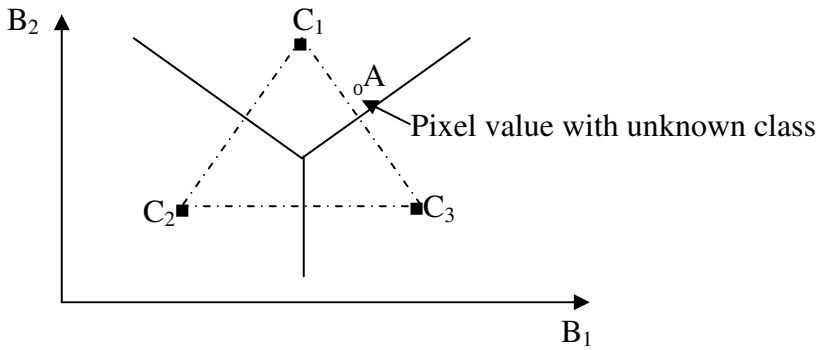


Figure 2.The minimum distance to means classification strategy

If we have a pixel with a grey-level vector located in the B_1-B_2 space shown as A (an empty dot) it is possible to determine to which class it should belong by calculating the distance between A and each of the centers. A is assigned to the class whose center has the shortest distance to A.

3 Material and Methodology

3.1 Study area

The study area covers the whole of the Netherlands. The Netherlands is located in Western Europe bordering the North Sea, between Belgium and Germany. It is geographically located at about 52^0 30'N latitude and 5^0 45'E longitude. The area coverage of the study area is about 41,526 km^2, in which water and land occupy about 7643 km^2 and 33,883 km^2, respectively (URL 1)

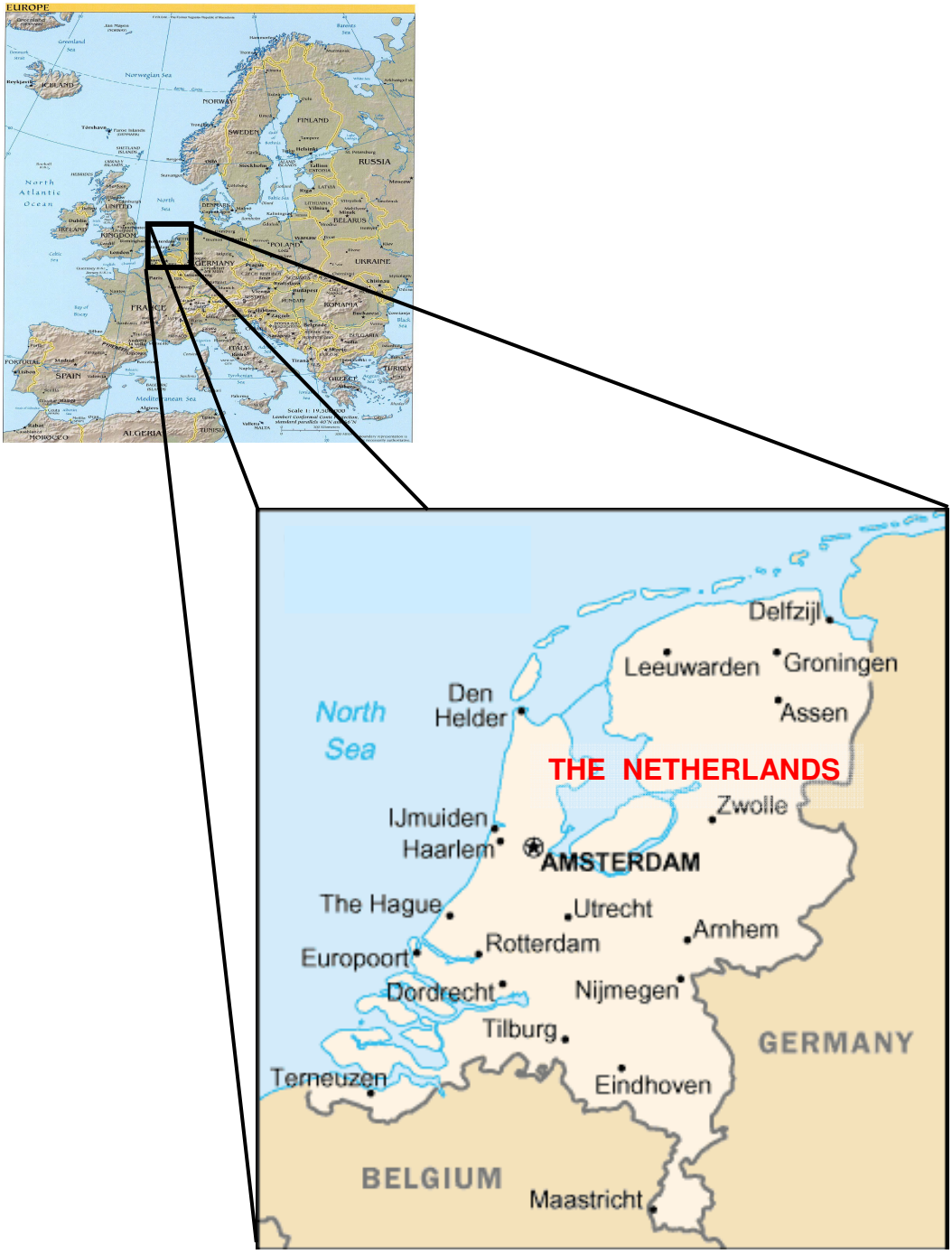


Figure 3. The study area (The Netherlands)

3.2 Data description

3.2.1 Main input data: MERIS data

The **ME**dium **R**esolution **I**maging **S**pectrometer (MERIS) is one of the main payload components of the European polar platform ENVISAT-1 system that was launched in March 2002. MERIS is a 68.5^0 field-of-view push-broom imaging spectrometer that measures the solar radiation reflected by the Earth, at ground spatial resolution of 300m (Full Resolution) and 1200m (Reduced Resolution). MERIS is a 15 band programmable imaging spectrometer, which allows for change in band position and bandwidths throughout its lifetime. It is designed to acquire data at variable bandwidth of 3.7 to 20nm over the spectral range of 390 – 1040nm. MERIS data will be provided at 3 different levels of processing, namely Level 0, Level 1, and Level 2 (URL 2). Table 1 shows the general characteristics of MERIS.

MERIS products will be available at two spatial resolutions: Full Resolution (FR) with a resolution at subsatellite point of 300 meter and Reduced Resolution (RR) with a resolution at subsatellite point of 1200 meter. In this study a Full Resolution (FR) MERIS level 1b data set for The Netherlands was used. Two cloud free MERIS images which were acquired on April 16, and July 14, 2003 were used for this study. The data included geocoded top-of-atmosphere (TOA) radiances ($Wsr^{-1}m^{-2}\ \mu m^{-1}$). The specification of the MERIS spectral channels of the two images are given in table 2.

Scenes dimension	Altitude	Field-of-view	Global coverage	Scanning	Spatial Resolution	Spectral band range(nm)	Swath width
575 km(FR) 1150 km(RR)	800 km	68.5^0	3 days	CCD arrays	300m (RR) 1200m (FR)	390 -1040	1150 km

Table 1. General characteristics of MERIS

Band No.	Band center(nm)	Band width(nm)
1	412.5	9.9
2	442.4	9.9
3	489.7	10
4	509.7	10
5	559.6	10
6	619.6	10
7	664.6	10
8	680.9	7.5
9	708.4	10
10	753.5	7.5
11	761.6	3.7
12	778.5	15
13	864.8	20
14	884.8	10
15	899.8	10

Table 2. The spectral bands of the MERIS images of April 16, and July 14, 2003

3.2.2 Reference data: LGN4 Database

The Dutch land use database (LGN) is a geographical database that describes the land use in the Netherlands. It has a grid structure with 25 meters cell size; the scale is about 1:50,000. The nomenclature of the LGN4 database contains 39 classes covering urban areas, water, forest, various agricultural crops and ecological classes. LGN is created for an important part on the base of satellite imagery, but also other data is integrated into the database. Currently 4 versions exist (LGN1 - LGN4) which span a time period of 1986 to 2000. For this study LGN4 was used as reference data for validation. The overall classification accuracy of LGN4 is 85-90% (URL 3). The 39 classes of LGN4 were recoded to nine main land cover classes (Appendix B-table I) and the grids were aggregated to 300 meters by assigning the most frequently occurring class as label. Thus, the aggregated LGN4 was used as reference data (Table 3).

Land cover	Total area (%)	Class value
Grassland	37.7	1
Arable land	24.3	2
Water	18.2	6
Built up area	9.3	7
Coniferous forest	4.8	5
Natural vegetation	1.7	9
Deciduous forest	3.0	4
Greenhouse	0.3	3
Bare soil	0.6	8

Table 3. Main land cover types in the Netherlands derived from LGN4.

3.3 Research methodology

The methodology used in this research involves four main steps.

The first step comprises pre-processing of the data set and preparation of training and testing areas.

The second step deals with supervised classification by maximum likelihood and minimum distance to mean classification for mono and multitemporal classification

The third step deals with preparation of input data for the implementation of CART analysis to study the effect of training size, data dimensionality, and splitting rule on classification accuracy of the decision tree classifier.

The fourth step deals with accuracy assessment and validation of classified images to draw conclusions for all the classifiers used.

The overall scheme of the whole methodology is illustrated in figure 4.

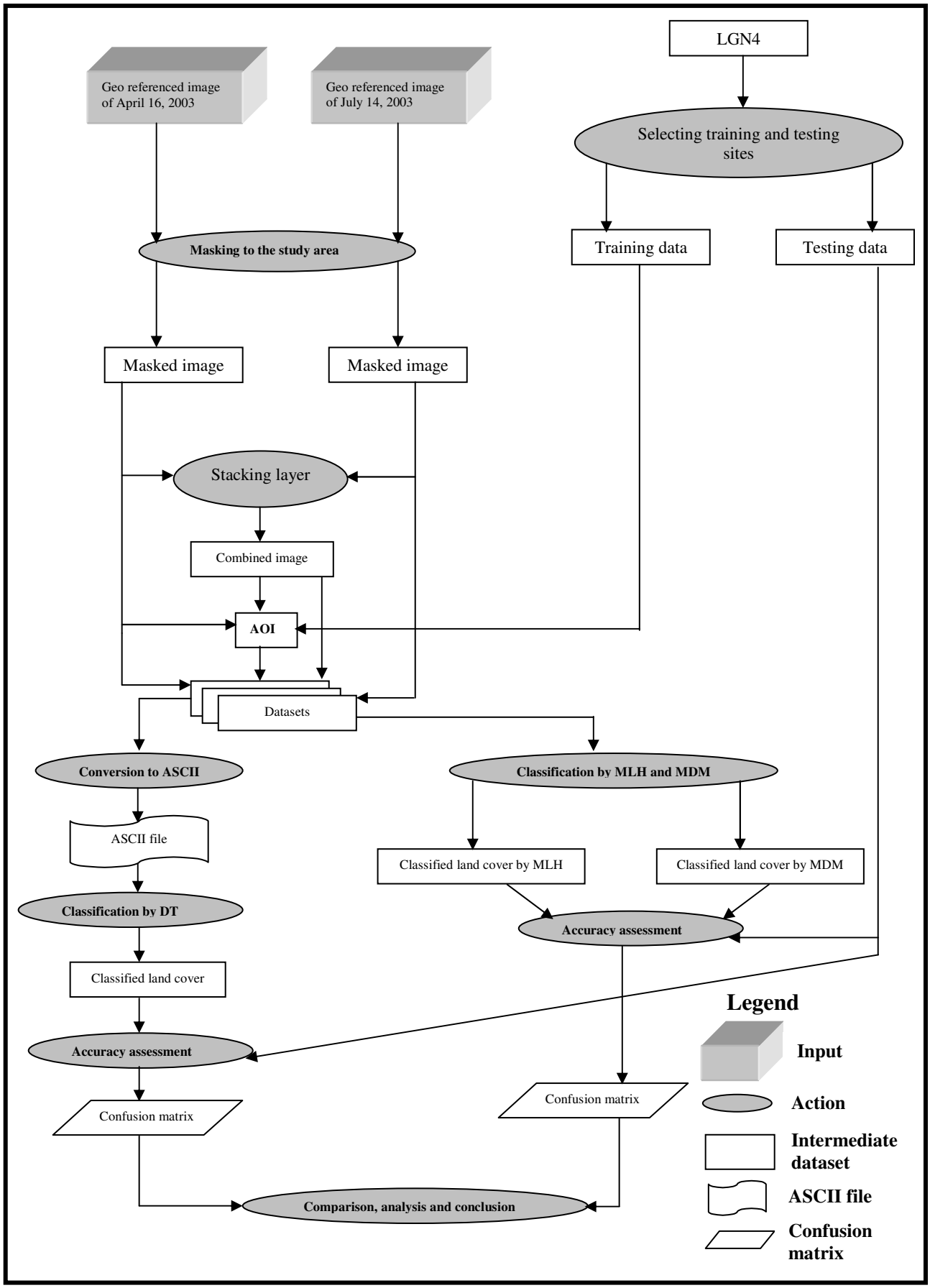


Figure 4. Overview of the main procedure involved in the methodological process.

3.3.1 Data pre-processing

After the two MERIS full resolution images (in ENVI 4.0 format) of the Netherlands were obtained, they were imported to Erdas Imagine 8.7 using the import function for further analysis. The images were already geo-referenced to the Dutch national coordinate system (RD). Masking was applied to have a data set covering the study area. Finally, the geocoded TOA radiances were multiplied by 10 to allow the inversion during classification by the MLH classifier.

3.3.2 Training and testing areas

Training and testing samples were delineated on the image by using the LGN4. They are used as learning (i.e. training site) paradigm for the classification of images and for the accuracy assessment (i.e. testing site) of classification (Lillesand and Kiefer, 2000). Both sites were selected in a random way, and at the same time keeping the spatial distribution of the classes over the whole study area.

For parametric classifiers that involve the estimation of statistical parameters, an important requirement is that the number of pixels included in the training data set for each class should be at least 10–30 times the number of features (Mather, 1999). Accordingly, four different training sets were tested and the best one, 9272 training pixels, was delineated on the image with reference to the aggregated LGN4 data (Appendix F-figure I).

Congalton and Green (1999) presented a method for estimating the sample size of the testing set for the accuracy assessment of classification results. The sample size n is derived from the relationship:

$$\mathbf{n = B\ II_i\ (1\text{-}II_i)/b_i^2}$$

Where, n = Minimum number of pixels in testing data

B = the upper (α/k) x 100^{th} percentile of the chi-square distribution with one degree of freedom

k = number of classes

α = required confidence level

b_i = the required precision

II_i = proportion of the area covered by class i

The approach can be made much clearer with a numerical example. In this study the number of classes (k) used was 8 (section 3.3.3). The proportion of the area covered

by grassland (II), the largest one, is 37.8%. The required precision is 0.05. The required confidence level is 95% i.e. α is 0.05. The value of chi-square for the probability level (0.05/8) = 0.00625 with one degree of freedom is 7.568.

$\mathbf{n = B\ II_i\ (1\text{-}II_i)/b_i^2}$

$n = 7.568\ (0.378)\ (1\text{-}0.378)/\ (0.05)^2$

n = 712

Thus, the minimum number of pixels for the testing set should be 712.

3.3.3 Land cover classes used for the study

Initially, the land cover types were recoded into 9 main land cover classes (Table 3). Greenhouses cover relatively only a small area. In the analysis presented in this thesis greenhouses were added to the built-up area. Therefore, the remaining 8 main land cover types were used for the analysis.

3.3.4 Multitemporal analysis of the MERIS image

To perform a multitemporal analysis, the two dates of MERIS images were registered into one master data set. The 12 bands (band 1, 2, and 11 were excluded (see section 4.1)) of the MERIS image from one date were combined with the same 12 bands (band 1, 2, and 11 were excluded) for an image acquired on another date, resulting into a 24 bands data set to be used in the classification. So, the two images of April 16 and July 14, 2003 were combined using the stack layer function in Erdas Imagine 8.7.

3.3.5 Classification

Supervised classifications were preformed using image signatures collected by training sets. But before performing classification, the signatures of the training sites were studied to see if signatures of different classes were well separated. Based on the spectral signatures analysis (see section 4.1), the first two bands of the visible region (band 1 and 2) and NIR band 11 of both images were not utilized for any analysis in this study. Three classification algorithms, MLH classifier, MDM classifier, and DT classifier, were tested over all input data. The 9272 training pixels and a test dataset of 2510 pixels were employed for the classification performed by the MLH and MDM classifiers. DT classifier was performed by CART software.

3.3.6 Overview of CART 5 model

CART (Classification and Regression Trees) can be used to analyse either categorical (classification) or continuous data (regression). The CART model is a well-known example of a univariate decision tree that is described by Breiman et al. (1984). A full CART analysis consists of three main steps.

The first step consists of tree building by splitting nodes. The building of a classification tree begins with a root node and then through a process of yes/no questions, it generates descendent nodes. CART finds the best possible variable to split the root node into two child nodes. In order to find the best variable, the software checks all possible splitting variables (called splitters), as well as all possible values of the variable to be used to split the node (Speybroeck et al., 2004). The extension to a categorical response variable (j categories) is obvious. CART then tries to obtain nodes, which contain as many subjects as possible belonging to only one of j categories. Finally, a saturated tree is obtained, which has one case or only "pure" cases (of only one category of the response) in each terminal node. Based on the distribution of classes in the training data set which would occur in that node and the decision cost matrix, each node is assigned to a predicted class. The assignment of a predicted class to each node occurs whether or not that node is subsequently split into child nodes (Lewis, 2000). This is necessary, as there is no way to know, during the tree building process, which nodes will end up being terminal nodes. Here the misclassification is not different for different categories and therefore the criteria for assigning classes to nodes can be based on the following rule:

Node is class I, if

$$\frac{N_i(t)}{N_j(t)} > \frac{N_i}{N_j}$$

With N_i is the number of class I in the data set

$N_i(t)$ is the number of class I in the node

N_j is total number of class in the data set

$N_j(t)$ is total number of class in the node

At the end of this process the result is often a very large and complex tree. In most cases, fitting a DT until all leaves contain data for a single class may overfit to the noise in the training data as some training samples may not be members of the class

that they purport to represent. If the training data contain any errors, then overfitting the tree to the data in this manner can lead to poor performance on unseen cases. To reduce the impact of this problem, the original tree can be pruned which is the second step of CART analysis. The tree pruning results in the formation of a sequence of simpler and simpler trees, through the cutting off of increasingly important nodes. The pruning relies on a complexity parameter $\boldsymbol{\delta}$, which is a measure of how much additional accuracy a split must add to the entire tree to warrant the additional complexity (Speybroeck et al., 2003). This parameter can be calculated for all internal nodes and the first pruning is done on the smallest one. By repeating this process a number of nested trees are obtained, going from the saturated tree to the root node only. As more nodes are pruned away, a simpler and simpler tree results.

The third step consists of selecting an optimal tree from the sequence of pruned trees. Large trees are highly accurate but they provide poor results when applied to new data sets (Friedl et al., 1999) and they are also difficult to interpret. The goal is thus to select the optimal tree with respect to the expected performance on an independent set of data, so that the information in the training dataset is fit but not overfit (Speybroeck et al., 2003). This depends on comparing the misclassifications for all the nested trees. The quality of a tree is a result reflecting the purity of its terminal node. If sufficient data is available, the simplest method is to divide this data into learning and test subsamples. The learning sample is then used to split nodes, continuing until the largest tree is grown. The test subsample is used to select the optimal tree among the different pruned trees.

Figure 5 shows the relationship between tree complexity, reflected by the number of terminal nodes, and a measure for misclassification cost for an independent test data set and for the original training dataset. As the number of nodes increases the decision cost decreases monotonically for the learning data. This corresponds to the fact that the maximal tree will always give the best fit to the training dataset. In contrast, the expected cost for an independent dataset (testing dataset) reaches a minimum, and then increases as the complexity increases. This reflects the fact that an overfitted and complex tree will not perform well on a new data set (Lewis, 2000).

For small data sets it is impossible to split the data into learning and testing data sets of reasonable sizes. A cross-validation method can be used, which consists of dividing the entire sample randomly into N (usually 10) subsamples, stratified by the response

variable. One sub-sample is then used as the test sample and the other N-1 are used to construct a large tree. The process is repeated N times, with a different subset of the data as the test dataset each time.

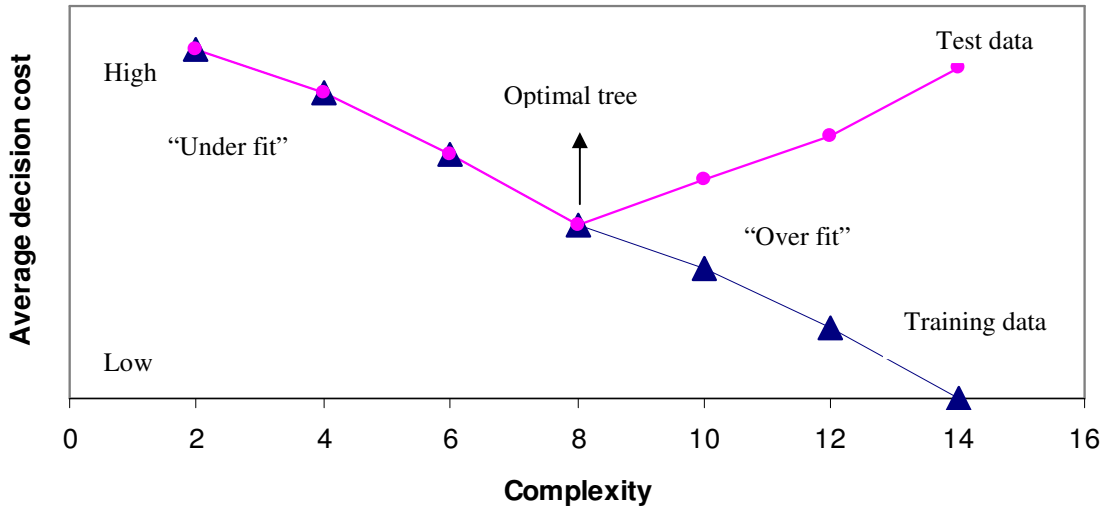


Figure 5. Building a classification tree: Finding a balance between overfitting the training data set and underfitting the test data set (Lewis, 2000)

3.3.7 Application of CART for DT classifier

The decision tree classification was carried out using CART software version 5.0. The training dataset with an identification number for each land cover type and the spectral bands of the MERIS image are the fundamental input for the CART analysis to make the DT classification. The 12 bands of the single image and 24 bands of the combined image were used as the predictor variables which are independent variables. The identification numbers for each land cover type (1, 2, 4, 5, 6, 7, 8, and 9 for grassland, arable land, deciduous forest, coniferous forest, water, built up, bare soil, and natural vegetation, respectively) represented the target variable which is a dependent variable. The land cover class itself was also considered as the categorical variable that is to be categorized according to the predictor variable.

Prior functions, one of the important options used in CART which help to shape the classification analysis, were used. By means of definition of class priors the CART software has the ability to compensate for over- or under-representation of classes (McIver and Friedel, 2002). By default the assumed priors are equal, that means all categories have equal probability which means that the classes are treated as if they were uniformly distributed in the sample. But in this study the probability was given

to each class based on the proportion of the area covered by each class (Table 3). For minimum node sizes, parent node minimum cases and terminal node minimum cases were set to 12 and 5, respectively. For testing the classification accuracy, a separate set of testing data was used.

To evaluate the effect of the size of training dataset on the overall classification accuracy of the DT classifier, seven subsets of training data for the April 16, 2003, MERIS image were formed by randomly sampling from the large training dataset (9272 pixels) used for the parametric classifiers. A total of 810 testing pixels was used to perform this experiment.

The numbers of pixels used in each of these training exercises were 800, 1330, 1860, 2390, 2920, 3450, and 3980 pixels, 100 pixels for each class at the starting point. Subsequently the training sizes were increased by 100 for four classes (due to their large area occupation and heterogeneity), 50 for one class, 30 for two classes and 20 for one class (Table 4). Finally, the large training data set (9272 pixels) was also used for comparison.

To study the effect of the splitting rule on the DT classifier, all the training data subsets (Table 4) used for studying the effect of training size were employed. Five splitting rules of CART software were analyzed, namely Gini, Symmetric Gini, Entropy, Twoing, and Order Twoing.

The Gini splitting rule looks for the largest class in the data set and isolates it from all other classes. Once the first split is made, Gini continues to split the data that require further segmentation based on the same class dominant criteria.

The twoing method operates by recursive segmentation trying to keep the same classes proportion (50%-50%) in each “child” node.

Entropy is another well-known splitting rule related to the likelihood function that, with multilevel targets, tends to look for splits where as many levels as possible are divided perfectly or near perfectly. As a result Entropy puts more emphasis on getting rare levels right relative to common levels than either Gini or Twoing.

Order Twoing is a modification of Twoing designed to handle ordered targets. This splitting rule only considers grouping together target classes adjacent to each other. It works best with targets with numerical levels.

To study the effect of the data dimensionality of the feature space over the performance of the DT classifier, the 9272 training and 810 testing pixels, and Gini splitting rule were used for the combined (April 16, and July 14, 2003) MERIS image.

A principal component analysis was performed for the combined image in order to know the relative importance of each band in the principal component analysis. Accordingly, the first four principal components contain about 99.7% of information of the combined image (Appendix D-table I). After that the information contributed by different bands to the first four principal components was ranked (Table 5). The number of features was initially set to 1 (the first bands that have high rank information in the combined image) and then increased by 1 at each iteration. Thus, the first experiment is based on one band of high rank information to the combined image 1, the second on bands 1-2, the third on 1-3, and so on.

Land cover type	Training data 1	Training data 2	Training data 3	Training data 4	Training data 5	Training data 6	Training data 7	Training dataset
Grassland	100	200	300	400	500	600	700	3619
Arable land	100	200	300	400	500	600	700	1999
Water	100	200	300	400	500	600	700	1372
Built up	100	200	300	400	500	600	700	909
Coniferous forest	100	150	200	250	300	350	400	499
Natural vegetation	100	130	160	190	220	250	280	313
Deciduous forest	100	130	160	190	220	250	280	326
Bare soil	100	120	140	160	180	200	220	235
Total	800	1330	1860	2390	2920	3450	3980	9272

Table 4. Different training set sizes used to study their effect on classification accuracy of the DT classifier.

Band contribution (%)	Real band number	Band number in the combined image
12.11	10*	10
12.03	12*	12
9.50	13*	13
9.00	14*	14
6.13	10	25
5.89	12	27
4.70	13	28
4.57	14	29
4.29	15*	15
3.34	3	18
2.98	4	19
2.87	3*	3
2.72	15	30
2.60	4*	4
2.42	5	20
2.41	5*	5
2.26	9*	9
2.14	9	24
1.66	6	21
1.42	7	22
1.36	6*	6
1.35	8	23
1.01	7*	7
0.93	8*	8

Table 5. Information contributed to combined MERIS image by different bands based on the first four principal components.
(*) = Band number from July image. The rest are the bands from April image.

3.3.8 Classification accuracy assessment

To compare the different classifiers, the CART results (based on 9272 training pixels) were converted to an image and validation was made using 2510 testing pixels with the same software, i.e. Erdas Imagine 8.7, as was used for MLH and MDM classification. The CART ASCII results were transformed to an image using the decision tree function of ENVI 4.0 software. The classified images were imported to Erdas imagine and masking was applied since the background (no data area) is also classified when ENVI 4.0 software is used for classification purpose. Therefore, the same test and training data sets were used for each classifier. Thus, any difference resulting from sampling variations is avoided.

No classification is complete until its accuracy has been assessed. In this context the term accuracy means the level of agreement between labels assigned by the classifier and class allocations based on ground data collected by the user, known as test data. Ground data do not necessarily represent reality, due to observation and recording errors, mislocation of test data sets, differences caused by change in land cover between the time of observation and the date of imaging, etc.

The accuracy of the classification process was evaluated using testing sets different from the training sets. For each classification the confusion (error) matrix was generated. The overall accuracy, the producer's accuracy, and the user's accuracy were calculated from the confusion matrix. The overall accuracy was used to compare the performance of classifiers. It is obtained by dividing the number of pixels correctly classified (i.e. the sum of the main diagonal entries of the confusion matrix) by the total number of pixels included in the validation process.

For individual classes, the producer's accuracy and user's accuracy were evaluated. The producer's accuracy measures the proportion of pixels in the test dataset that are correctly recognized by the classifier. The user's accuracy tells us the proportion of pixels identified by the classifier as belonging to class i that agrees with the test data.

4 Result and discussion

4.1 Spectral signatures analysis

To study the spectral reflectance of land cover types, the spectral signatures of main land covers were derived from training samples of different MERIS images. Figures 6 and 7 show the spectral signatures of the main land covers of April 16, 2003, and July 14, 2003, respectively.

In general, vegetation classes showed a steep slope between red (681nm) and NIR (778.5nm). The first two MERIS bands (412.5 and 442.4nm) of the visible spectrum have shown a relatively high radiance value when they are compared to other visible parts. This is due to atmospheric scattering in the blue part of the spectrum. On the other hand, in the NIR region band 11 (762nm) showed an absorption dip for all classes due to the oxygen absorption in the atmosphere. Clevers et al. (2003) have also obtained similar results on the spectral signature analysis of the MERIS image of June 16, 2003.

The spectral signatures of April 16, 2003, image (figure 6) showed a clear spectral signature for grassland, bare soil and arable land. Water is also clearly observed in the NIR part. A clear vegetation spectrum for grassland can be easily seen. The grassland has the highest radiance value in the NIR portion of the spectrum. On the contrary, water showed the lowest radiance value in the same region. Arable land showed a flat spectral reflectance like bare soil. This is what was expected for arable land since in early spring most arable land is bare. The spectral signatures of built up area, coniferous forest, deciduous forest and natural vegetation revealed high overlap and have also quite similar radiance values. So, these classes will be the most difficult to distinguish.

The spectral signatures for the July 14, 2003, image (figure 7) are different from the April image. All vegetation classes showed a nice vegetation spectrum due to increase in their biomass. The spectral signature of grassland is more or less similar to the April one even though it increased in biomass (radiance value showed further increment particularly in the NIR region). In contrast, the spectral signature of arable land completely changed to a clear vegetation spectrum. It has also more biomass than grassland. The built up area signature also has a bit a vegetation spectrum. This indicates that built up has some contamination of vegetation in its signature. The radiance value of bare soil is larger than the one in April. This is because the soil

becomes drier in July than in April. Water has almost the same spectrum as that of April. In general, in the visible part of the spectrum the radiance value of all land covers are quite close to each other which make them difficult to separate.

Generally, when comparing the spectral signatures of the two dates, there is a decrease in radiance value for vegetation classes in the visible part (band 7 and 8) for the July image. It can be observed that the most pronounced decrease is for arable land in this region of the spectrum. In the NIR part of the spectrum, temporal differences are more pronounced; the July image showed quite an increase of radiance values for all land covers except water. This was more pronounced for arable land and deciduous forest.

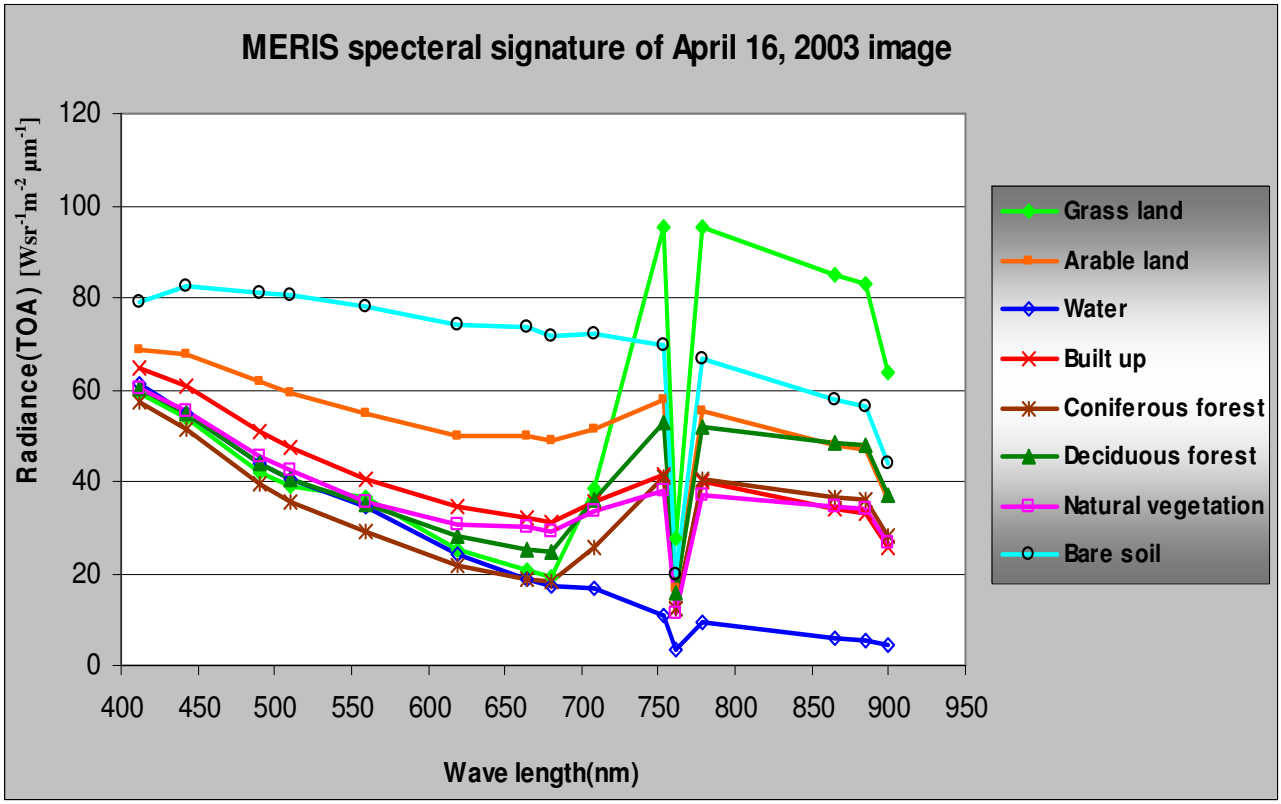


Figure 6. Spectral signature for the main land covers of April 16, 2003 MERIS image

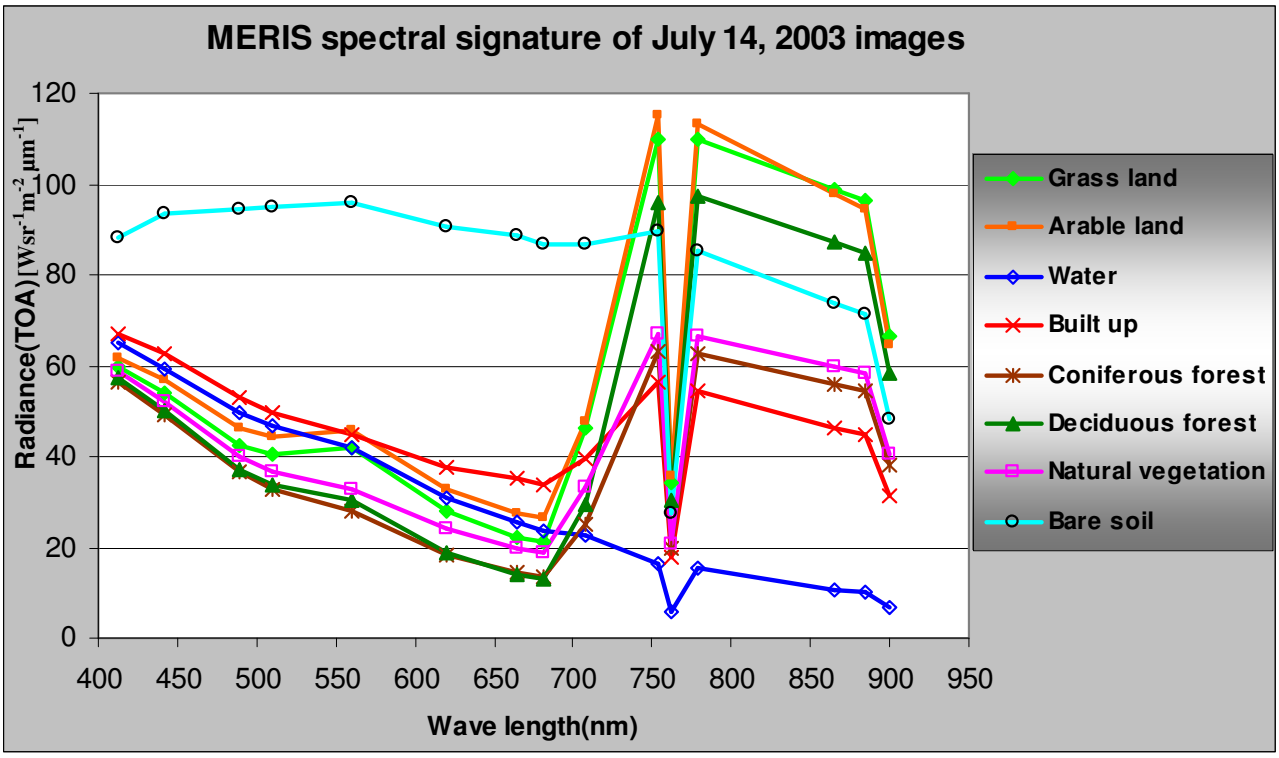


Figure 7. Spectral signature for the main land covers of July 14, 2003 MERIS image

4.2 Classification results of MLH classifier

The Maximum likelihood (MLH) classification was applied to all the MERIS images. Table 6, 7, and 8 show the results of classification by MLH classifier for April 16, July 14 and the combined images, respectively.

Table 6. Accuracy assessment of April 16, 2003, MERIS image by MLH classifier

Land cover types	**Producer's accuracy**	**User's accuracy**
Grassland	60.88%	81.39%
Arable land	83.72%	55.52%
Deciduous forest	53.76%	31.25%
Coniferous forest	51.61%	79.21%
Water	86.44%	94.95%
Built up	64.65%	77.19%
Bare soil	57.89%	50.00%
Natural vegetation	59.62%	45.59%

Overall Classification Accuracy = 69.76% *kappa statistics = 0.6147*

Table 7. Accuracy assessment of July 14, 2003, MERIS image by MLH classifier

Land cover types	**Producer's accuracy**	**User's accuracy**
Grassland	64.53%	80.21%
Arable land	81.75%	58.74%
Deciduous forest	53.76%	38.76%
Coniferous forest	47.74%	78.72%
Water	86.92%	95.99%
Built up	67.20%	79.03%
Bare soil	63.16%	32.43%
Natural vegetation	73.08%	36.89%

Overall Classification Accuracy = 71.11% *kappa statistics = 0.6304*

Table 8. Accuracy assessment of combined MERIS images by MLH classifier

Land cover types	Producer's accuracy	User's accuracy
Grassland	62.43%	82.00%
Arable land	86.58%	53.54%
Deciduous forest	58.06%	40.91%
Coniferous forest	43.23%	83.75%
Water	86.68%	95.72%
Built up	66.24%	81.89%
Bare soil	63.16%	48.00%
Natural vegetation	63.46%	63.46%

Overall Classification Accuracy = 70.96% *kappa statistics = 0.6268*

When the April image was analyzed, the overall accuracy was 69.76% (Table 6). The results for each class showed a producer's and user's accuracy ranging from 51.61% to 86.44% and 31.25% to 94.95%, respectively. Producer's accuracy (86.44%) and user's accuracy (94.95%) of water were quite satisfactory, which means that most of the water area can be correctly recognized. Although it was observed from the spectral signatures (figure 6) that grassland could be well differentiated from other classes, the producer's accuracy (60.88%) of grassland was not as good as expected. This may be due to variation in biomass for the various grassland pixels that results in larger variation than what was observed in the spectral signatures. But grassland has a good user's accuracy (81.39%). Arable land has a good producer's accuracy (83.72%) and a moderate user's accuracy (55.52%). Coniferous forest has the lowest producer's accuracy (51.61%) and good user's accuracy (79.21%). This indicates that although 51.61% of the pixels were correctly classified as coniferous forest, 79.21% of the areas labeled as coniferous forest actually belong to that class on the ground. Built-up area has 64.65% producer's accuracy and 77.19% user's accuracy which were nice for such highly variable class. Deciduous forest, bare soil, and natural vegetation have low user's accuracy (31.25%, 50%, and 45.59%, respectively) as compared to other classes.

The main confusion occurred between grassland and arable land. This may be due to the area covered with winter wheat, which will be confused with grassland at this season (Mucher et al., 2000). Small amounts of grassland were confused with built-up area, deciduous forest, and natural vegetation. Coniferous forest was confused with

grassland and arable land but it was mainly confused with deciduous forest. Deciduous forest was also confused with grassland and arable land. Built-up areas were particularly confused with arable land (Appendix C-table I).

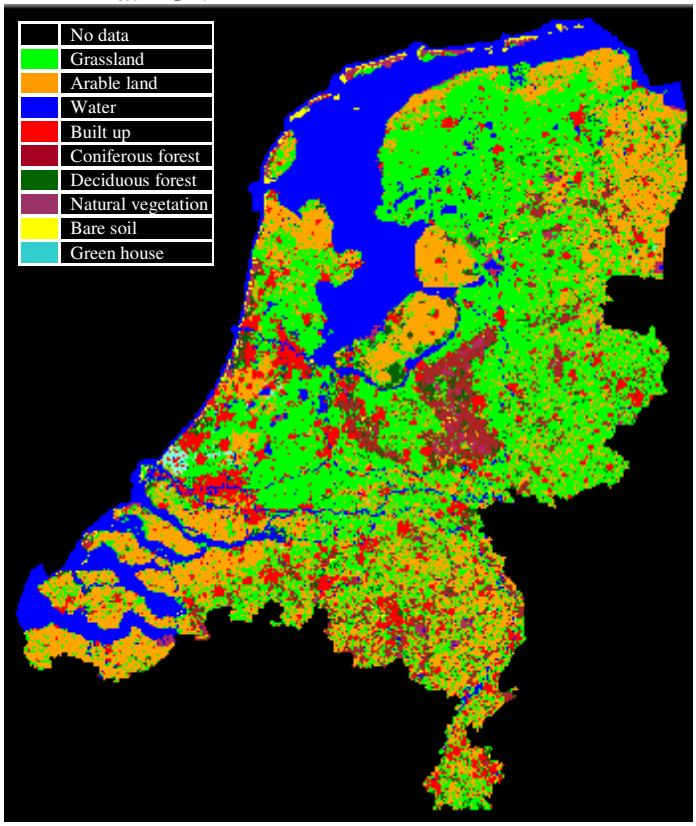
Using only the July image, the overall accuracy was 71.11% (Table 7). Producer's and user's accuracy of each class ranges from 47.74% to 86.92% and 32.43% to 95.99%, respectively. In this image also water has the highest producer's and user's accuracy. It has almost the same accuracies as for the April image. Similarly as in the April image, coniferous forest has the smallest producer's accuracy (47.74%) but bare soil has the lowest user's accuracy (32.43%). Grassland has good producer's accuracy (64.53%) that is larger than in the April image. Deciduous forest, bare soil and natural vegetation have a low user's accuracy like in the April image. These small area classes are found to have higher producer's accuracies than user's accuracies. The reason may be that a small area class often can not occupy the absolute majority in a 300m * 300m MERIS pixel.

Here also significant mixing was seen between grassland and arable land. It can be concluded that there is high spectral overlap between these two classes for both images. Coniferous forest was mainly confused with deciduous forest. It was also confused with natural vegetation. Built-up area was mainly confused with arable land. Some pixels of deciduous forest were also misclassified as grassland (Appendix C-table II).

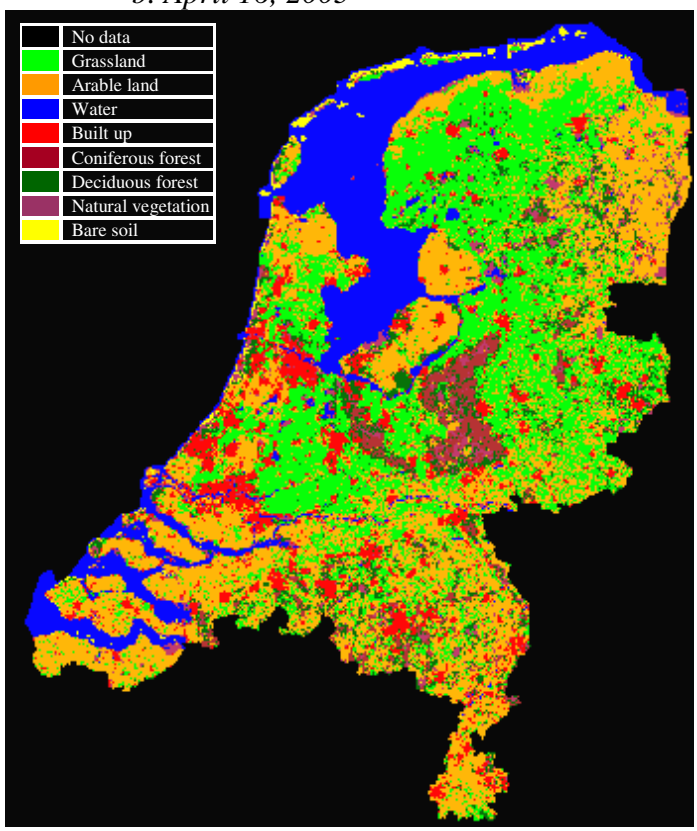
The overall accuracy of the combined MERIS image was 70.96% (Table 8) which is almost the same as for the April (69.76%) and July (71.11%) images. These results are contrary to Oetter et al. (2000) who found that multiseasonal images provide significant improvements in accuracy of vegetation classifications. This may be due to the number of images used in one growing season for the classification. The results of the combined image showed producer's and user's accuracy ranging from 43.23% to 86.68% and 40.91% to 95.72%, respectively. The only land cover type that showed improvement in both accuracies was deciduous forest, about 5% and 3% increment in producer's and user's accuracy, respectively (Table 6 and 7). Natural vegetation has shown about 27% significant increment in user's accuracy although 10% decrease has been observed in producer's accuracy (Table 6 and 7). About 8% decrease in producer's accuracy and 4% increment in user's accuracy have been observed in the class coniferous forest (Table 5 and 7). This class has still the lowest producer's

accuracy. Water has almost the same accuracies in all three images. There was no significant change in both accuracies for grassland. On the other hand, arable land has shown about 3% increment in producer's accuracy and it decreases about 2% in user's accuracy (Table 5 and 7). Figure 8 shows the classified land covers using the MLH classifier for monotemporal and multitemporal MERIS image with reference to LGN4.

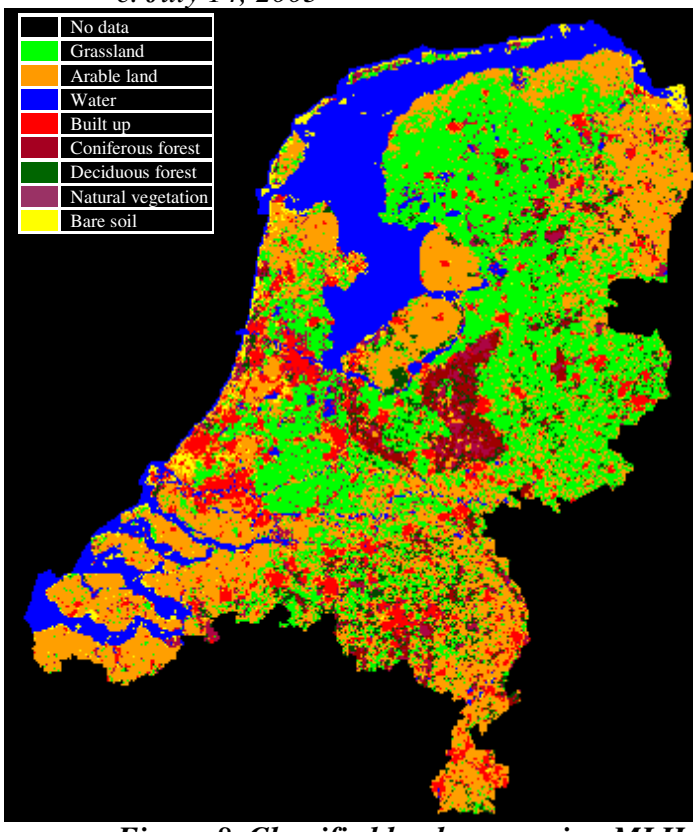
a. LGN4



b. April 16, 2003



c. July 14, 2003



d. Combined MERIS image

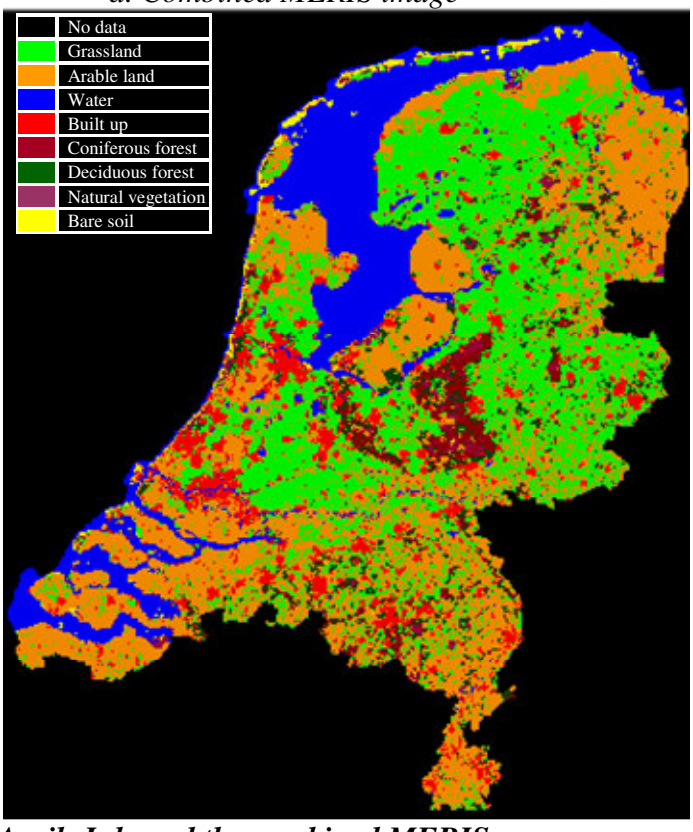


Figure 8. Classified land cover using MLH: April, July and the combined MERIS images with reference to LGN4

4.3 Classification results of MDM classifier

The minimum distance to mean (MDM) classification was also employed in order to compare with other classifiers. Table 9, 10 and 11 show the classification results of the MDM classifier for April, July and the combined MERIS image, respectively.

Table 9. Accuracy assessment of April 16, 2003, MERIS image by MDM classifier

Land cover types	**Producer's accuracy**	**User's accuracy**
Grassland	65.86%	79.26%
Arable land	52.95%	58.50%
Deciduous forest	47.31%	13.50%
Coniferous forest	56.77%	67.69%
Water	85.47%	94.64%
Built up	55.10%	57.28%
Bare soil	31.58%	75.00%
Natural vegetation	51.92%	24.32%

Overall Classification Accuracy = 63.07% *Kappa Statistics = 0.5376*

Table 10. Accuracy assessment of July 14, 2003, MERIS image by MDM classifier

Land cover types	**Producer's accuracy**	**User's accuracy**
Grassland	73.48%	62.80%
Arable land	35.78%	73.26%
Deciduous forest	60.22%	38.10%
Coniferous forest	44.52%	74.19%
Water	86.92%	94.72%
Built up	63.38%	67.69%
Bare soil	42.11%	44.44%
Natural vegetation	63.46%	13.64%

Overall Classification Accuracy = 63.31% *Kappa Statistics = 0.5281*

Table 11. Accuracy assessment of combined MERIS images by MDM classifier

Land cover types	Producer's accuracy	User's accuracy
Grassland	64.20%	80.25%
Arable land	64.04%	65.45%
Deciduous forest	67.74%	20.66%
Coniferous forest	46.45%	71.29%
Water	86.68%	93.72%
Built up	60.51%	70.90%
Bare soil	36.84%	36.84%
Natural vegetation	61.54%	19.88%

Overall Classification Accuracy = 66.18% *Kappa Statistics = 0.5776*

The classification result of the April image was 63.07% (Table 9). The producer's and user's accuracy ranged from 31.58% to 85.47% and 13.50% to 94.64%, respectively. Water has quite satisfactory producer's (85.47%) and user's (94.64%) accuracies. Grassland was well separated having 65.86% producer's and 79.26% user's accuracy. Arable land has a moderate producer's (52.95%) and user's (58.50%) accuracy which was not as expected because the signature of this class was well separated from other classes (Figure 6). The class deciduous forest was the most disappointing class that has 47.31% producer's and 13.50% user's accuracy. This is because most of its pixels were misclassified as coniferous forest, grassland and natural vegetation. Moreover, grassland, arable land, coniferous forest and built up area were confused with it (Appendix C-table IV). Substantial mixing was observed among grassland, arable land and deciduous forest. Arable land was also mixed with built up area since most of the arable land was still bare. The smallest area classes have low accuracies when compared to other large area classes. The reason may be that small area classes often can not occupy the absolute majority in MERIS pixels.

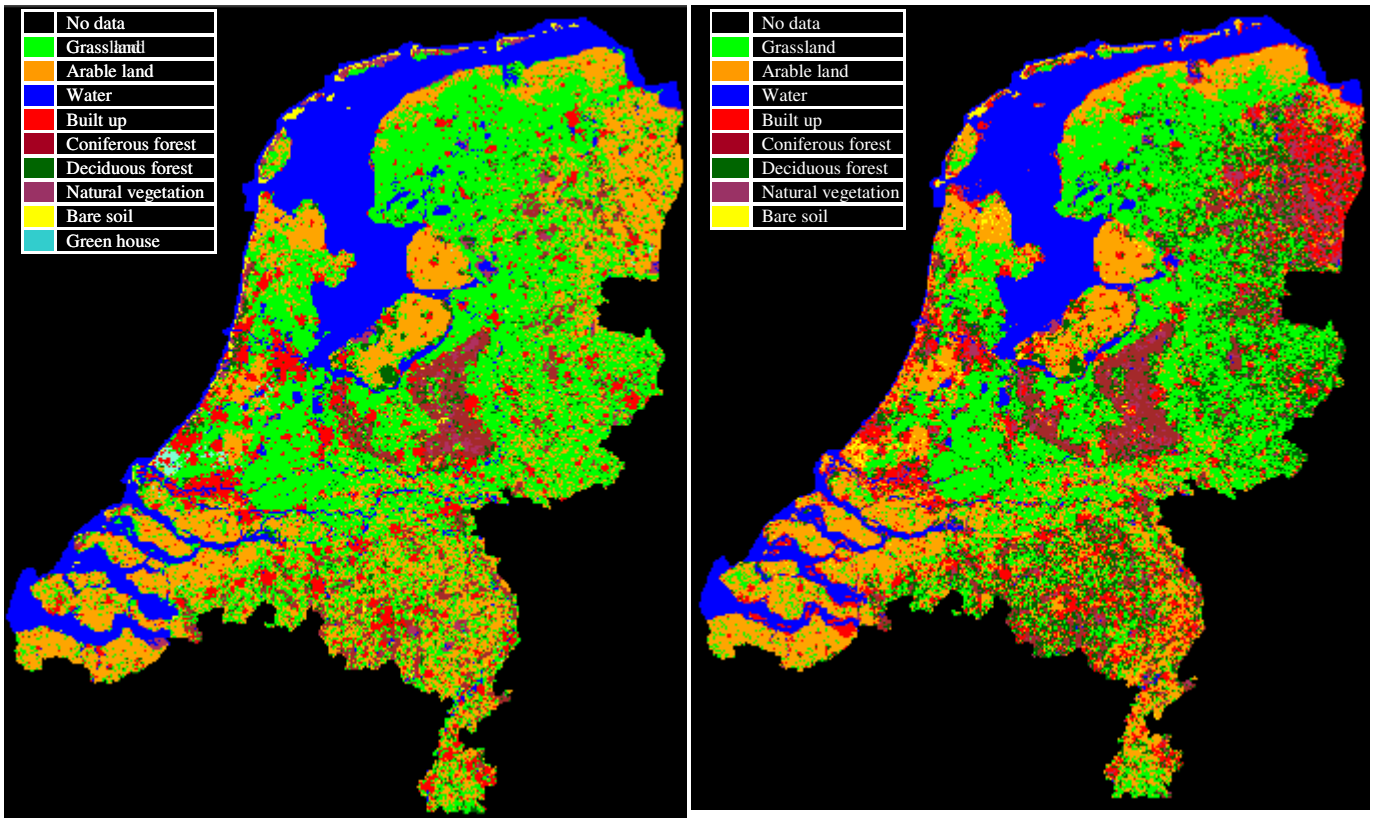
The result from the July image has shown 63.31% overall accuracy which was quite similar to the April image (Table 10). The producer's and user's accuracy ranged from 35.78% to 86.92% and 13.64% to 94.72%, respectively. Water has almost the same accuracies as for the April image. The producer's accuracy of grassland was 73.48% which is higher than for the April image, but its user's accuracy (62.80%) was

reduced by 16%. This is because more than half of the arable land reference pixels were classified as grassland (Appendix C-table V). For the same reason arable land has the smallest producer's accuracy (35.78%). It can be concluded that there was high spectral overlap between these two classes which was also observed from their spectral signatures (Figure 7). Coniferous forest and built up area were mainly confused with natural vegetation which has the smallest user's accuracy (13.64%). Both accuracies of built up area have shown an increment of almost 10% as compared to the April image since arable land is no more bare at this time.

The result from the combined MERIS image showed that the overall accuracy was 66.18% (Table 11) which was higher than for both single images. The producer's and user's accuracy ranged from 36.84% (bare soil) to 86.68% (water) and 19.88% (natural vegetation) to 93.72% (water), respectively. Arable land was the class that showed improvement in producer's and user's accuracy by 12% and 7%, respectively, when compared to the April image. Figure 9 shows the classified land covers of the April, the July and the combined MERIS image using the MDM classifier with reference to LGN4.

a. LGN4

b. April 16, 2003



c. July 14, 2003

d. Combined image

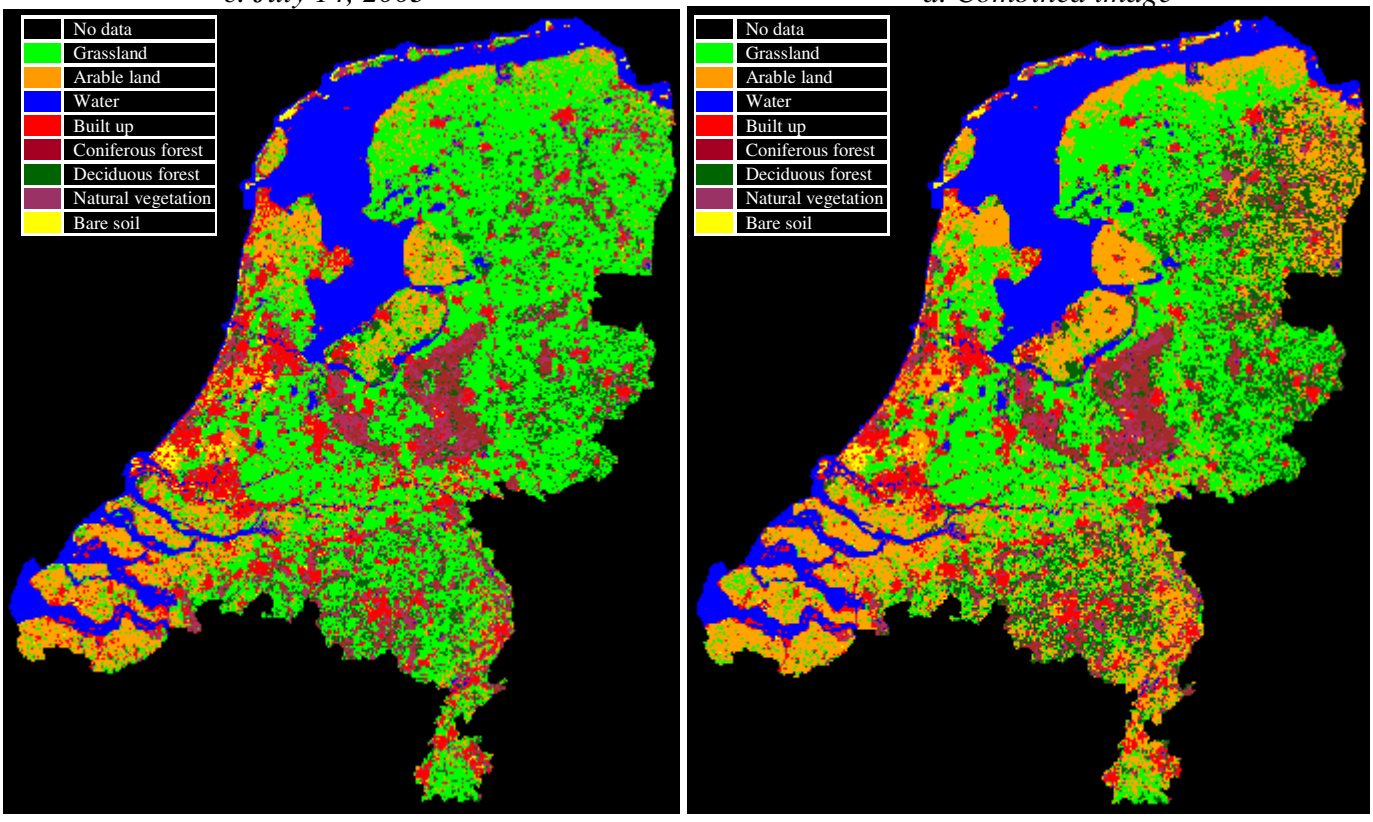


Figure 9. Classified land cover using MDM: April, July and the combined MERIS images with reference to LGN4

4.4 Classification results of DT classifier

The decision tree (DT) classification using the CART model was tested versus MLH and MDM. The performance of the decision tree was optimized by investigating the effect of the training dataset size, splitting rule, and data dimensionality on the classification accuracy of the DT.

The decision tree classifier starts by producing a tree from the root node that contains all input data. The thresholds for each class were generated by deriving information from different bands that were used as input data. After that, it recursively partitions the data set into simpler forms (See chapter 2 for a more detailed description).

Figure 10 shows the decision tree obtained from the CART analysis for the combined image. In this tree each node (root, internal and terminal nodes) contains a decision rule. All the terminal (leaf) nodes represent land cover types that are the target variables.

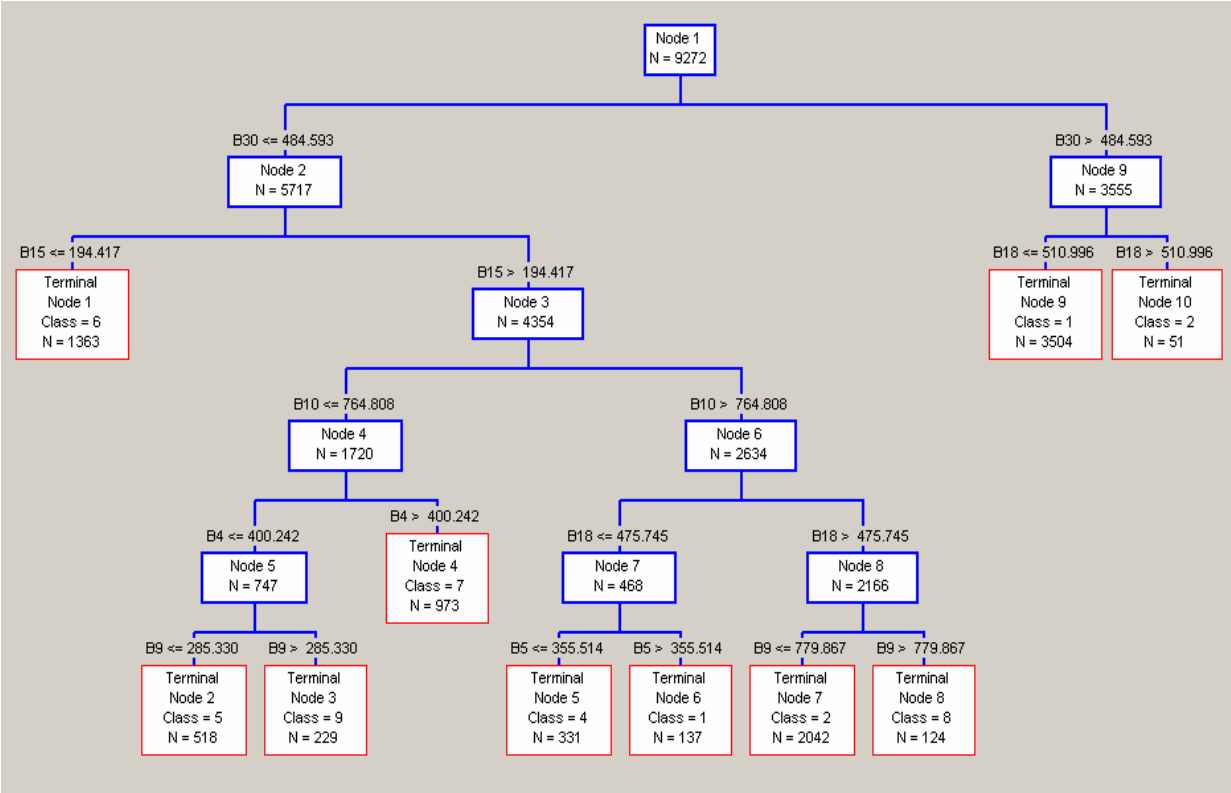


Figure 10. Decision tree generated by the CART model for the combined MERIS image

Numbers in shaded part indicate TOA radiances multiplied by 10

4.4.1 Effect of training dataset size and splitting rule

The characteristics of the data used to train a supervised classification have a considerable influence on the accuracy of the resulting classification (Thomas et al., 1987). It is essential that the training data provide a representative description of each class. For statistical classification the number of pixels incorporated in training set for each class must be at least 10-30 times the number of features. The required training set size must therefore be large when the classification involves a large number of classes or utilizing data acquired in many bands. This number could further rise due to the effect of other variables (for example, if variations in soil background have an effect on the spectral response of a crop) (Foody and Arora, 1997). Clearly acquiring such a large training datasets is difficult in such conditions. On the other hand, non parametric classifiers can perform successfully using training data sets that are smaller than those required to train statistical classifiers (Foody et al., 1995; Pal and Mather, 2003).

Figure 11 shows the relationship between classification accuracy and the training dataset size with the different splitting rules implemented in CART. The results indicate that the level of accuracy increases with the size of training set which is almost linear up to the third training data set over all splitting rules. However, the fourth training set (2390 pixels) showed a decrease in accuracy for all splitting rules. Afterwards, the level of accuracy started to increase until the sixth training set (3450 pixels) and decreased for the seventh data set (3980 pixels) for all splitting rules. The largest training set (9272) also showed almost the same level of accuracy as the seventh training set for all splitting rules except Gini. The highest accuracy was obtained at a 3450 pixels training set for all splitting rules (Figure 11). These results indicate that the accuracy of the decision tree classifier increased as the size of the training set increased but only up to a certain point. In addition to this, the decision tree classifier does not need very large training sets to be effective. These results match with the findings of Pal and Mather (2003) although the increment was not absolutely linear as the finding of Pal and Mather.

The CART program also allows to choose the splitting rule considering the levels of target variables (the number of land cover type). Accordingly, the CART software manual recommends to test the result with different splitting criteria such as Gini, Symmetric Gini, Entropy, Twoing and Ordered Twoing. The results of figure 11

indicate that the Gini splitting rule gave the best overall accuracy in all training set sizes. Particularly for training set size 3450 and 9272 pixels, Gini generated about 67.2% and 67.6% accuracy, respectively. This result contradicted with Salford systems (2002) which recommend the Twoing splitting rule that fits for a target variable analysis with 4 up to 9 levels. The results indicate that choosing splitting rules does not only depend on the level of target variables but also on another factor such as type of input data used for analysis.

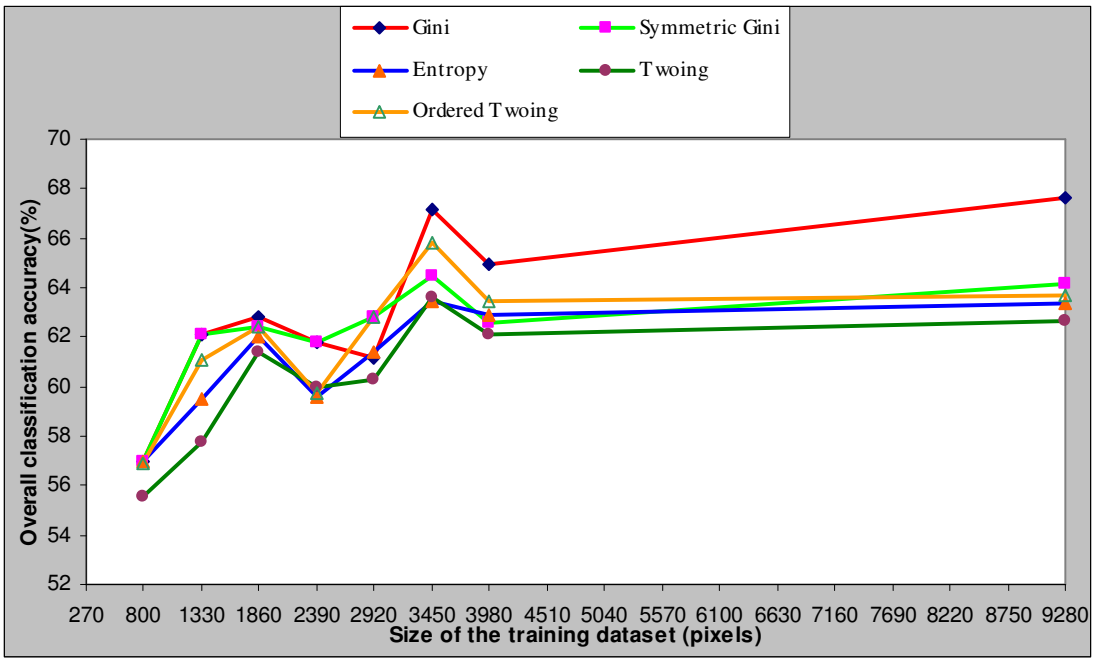


Figure 11. Variation of classification accuracy with increasing number of training size and different splitting rules using a decision tree classifier on April 16, 2003, MERIS image.

4.4.2 Dimensionality of the feature space

The ability to discriminate spectrally between classes is generally a function of the dimensionality of the data used. Typically, class separability is positively related to the number of discriminating variables used. These discriminating variables are generally data acquired in specific spectral wavebands, but could also be data acquired at different time periods. Shahshahani and Landgrebe (1994) have shown that the class separablity and ultimately classification accuracy rise initially with an increase in the number of wavebands used up to a point beyond which the addition of

data acquired in other spectral wavebands has either no significant effect or results in a decrease in classification accuracy.

Figure 12 shows the levels of overall classification accuracy obtained using the DT classifier on the combined MERIS image. These results indicate that the accuracy of DT classification increased for the first four bands of the July image. A slight decrease was observed for band 10 of the April image and then the accuracy started to increase on band 12 and 13 of the same image. By adding band 14 of the April and 15 of the July image, a change of level of accuracy was not observed. A possible reason for this may be due to the high correlation of NIR bands of MERIS images of the two dates (Appendix E –table I). At band 3 of both dates, the level of accuracy has shown high increment. This is due to the fact that band 3 of the MERIS image does not correlate to the NIR bands (Appendix E). Another increment of accuracy was observed at band 9 of the July image which is not correlated to visible and NIR bands of the MERIS image (Appendix E-table I). The rest of the bands have not any effect on the accuracy of DT classification since their information was already utilized by the former bands that correlated to them. From these results it is possible to conclude that the CART model of DT classifier can weight the importance of the discriminating variables in the classification, thereby ignoring redundant variables.

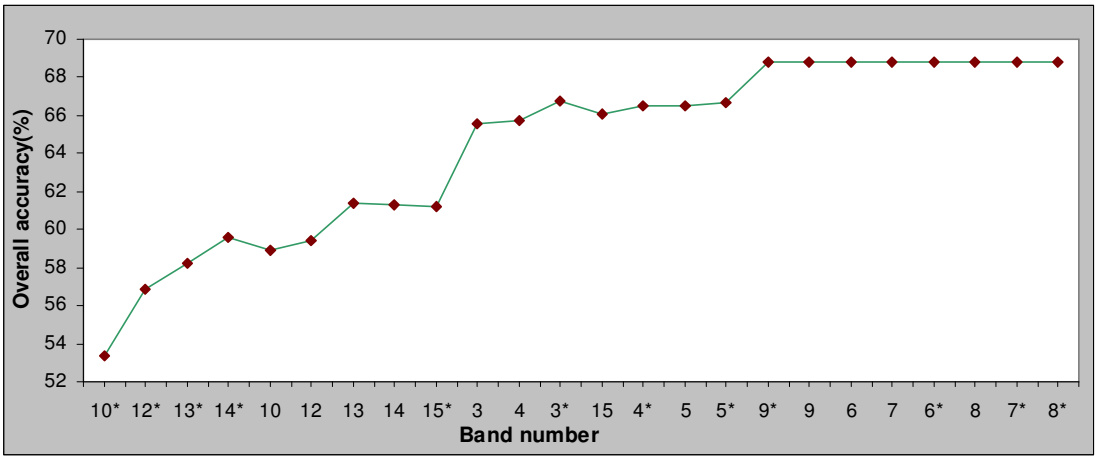


Figure 12. Classification accuracies using the combined MERIS image with 9272 training pixels and increasing number of features.
(*) = Band from July image. The rest is band from April image

4.4.3 Final Decision Tree classification

The final DT classification was done in order to compare its overall accuracy with the other classifiers: MLH and MDM classifiers. Table 12, 13 and 14 show the results of the DT classifier for April, July and the combined MERIS images, respectively.

Table 12. Accuracy assessment of April 16, 2003, MERIS image by DT classifier

Land cover types	**Producer's accuracy**	**User's accuracy**
Grassland	75.47%	72.27%
Arable land	57.42%	64.07%
Deciduous forest	49.46%	22.12%
Coniferous forest	52.26%	66.94%
Water	85.47%	94.64%
Built up	64.33%	68.01%
Bare soil	21.05%	80.00%
Natural vegetation	59.62%	51.67%

Overall Classification Accuracy = 68.56% *Kappa Statistics = 0.5925*

Table 13. Accuracy assessment of July 14, 2003, MERIS image by DT classifier

Land cover types	**Producer's accuracy**	**User's accuracy**
Grassland	71.27%	64.37%
Arable land	52.59%	56.11%
Deciduous forest	59.14%	40.44%
Coniferous forest	56.13%	72.50%
Water	87.89%	94.78%
Built up	69.11%	70.45%
Bare soil	5.26%	100.00%
Natural vegetation	44.23%	63.89%

Overall Classification Accuracy = 67.13% *Kappa Statistics = 0.5683*

Table 14. Accuracy assessment of combined MERIS images by DT classifier

Land cover types	Producer's accuracy	User's accuracy
Grassland	75.24%	71.08%
Arable land	64.04%	63.03%
Deciduous forest	54.84%	38.35%
Coniferous forest	55.48%	69.35%
Water	87.17%	94.74%
Built up	69.11%	70.45%
Bare soil	10.53%	100.00%
Natural vegetation	44.23%	62.16%

Overall Classification Accuracy = 70.84% *Kappa Statistics = 0.6212*

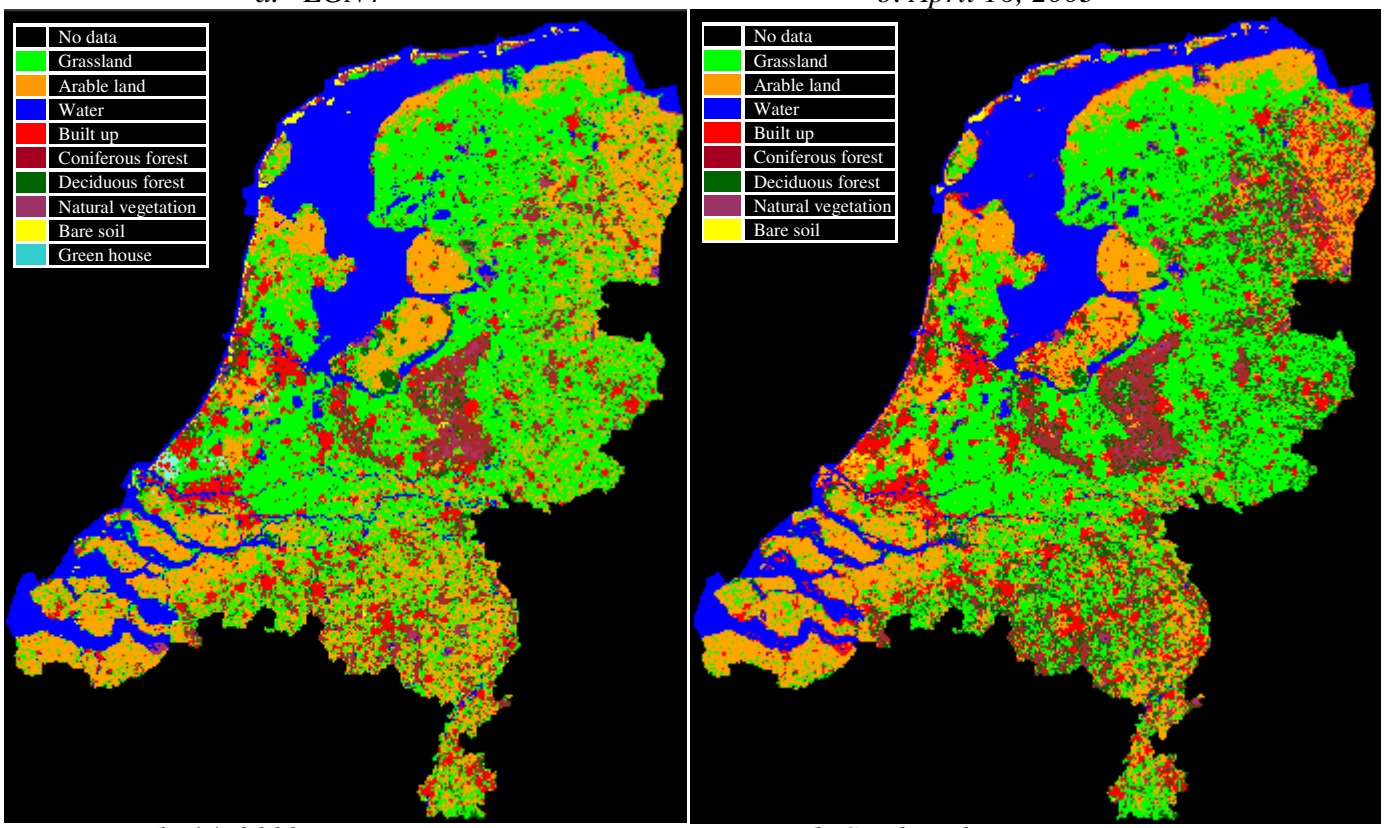
An overall accuracy of 68.56% was obtained for the April image (Table 12). The result of producer's and user's accuracy ranged from 21.05% to 85.47% and 22.12% to 94.64%, respectively. Water has the highest accuracies. Bare soil has the lowest producer's accuracy (21.05%) but it has a high user's accuracy (80%). That means 80% of the area labeled as bare soil was actually bare soil on the ground. Grassland has quite a good producer's (75.47%) and user's (72.27%) accuracy. The producer's accuracy (57.42%) of arable land was not as expected even though its user's accuracy (64.07%) was good. Some substantial mixing between grassland and arable land was also seen with the DT classifier (Appendix C-table VII). Deciduous forest was the class that has low producer's (49.46%) and user's (22.12%) accuracy. It was typically confused with grassland and coniferous forest. Coniferous forest was also confused with deciduous forest and grassland. The small area classes have also shown low accuracy as compared to large area classes.

The result from the July image has shown that the overall accuracy was 67.13% (Table 13) which was lower than for the April image. The results indicated that producer's and user's accuracy range from 5.26% to 87.895 and 40.44% to 100%, respectively. Grassland and arable land were the classes that have lower producer's and user's accuracy than for the April image. This was because confusion occurred between grassland and arable land which was already seen from their spectral signatures (figure 7). Bare soil was the class that showed the lowest producer's accuracy (5.26%) and highest user's accuracy (100%). Deciduous forest, coniferous forest and built up area have shown better accuracies than for the April image. Within

this image there was also confusion between coniferous and deciduous forest. Built up area was typically confused with grassland and arable land (Appendix C-table VIII).

The results from the combined MERIS image indicated that the overall accuracy was 70.84%, which was higher than for the April and July images by about 3%. The producer's and user's accuracy of this image ranged from 10.53% to 87.17% and 38.35% to 100%, respectively (Table 14). Arable land was the only class that has shown accuracy improvement in this combined image. Grassland and arable land were also confused as they were on other images. Deciduous forest was confused with coniferous forest and grassland. Coniferous forest was also confused with deciduous forest and grassland (Appendix C-table IX). Figure 13 shows the classified land covers of the April, the July and the combined MERIS image using the DT classifier with reference to LGN4.

a. LGN4 *b. April 16, 2003*



c. July 14, 2003 *d. Combined image*

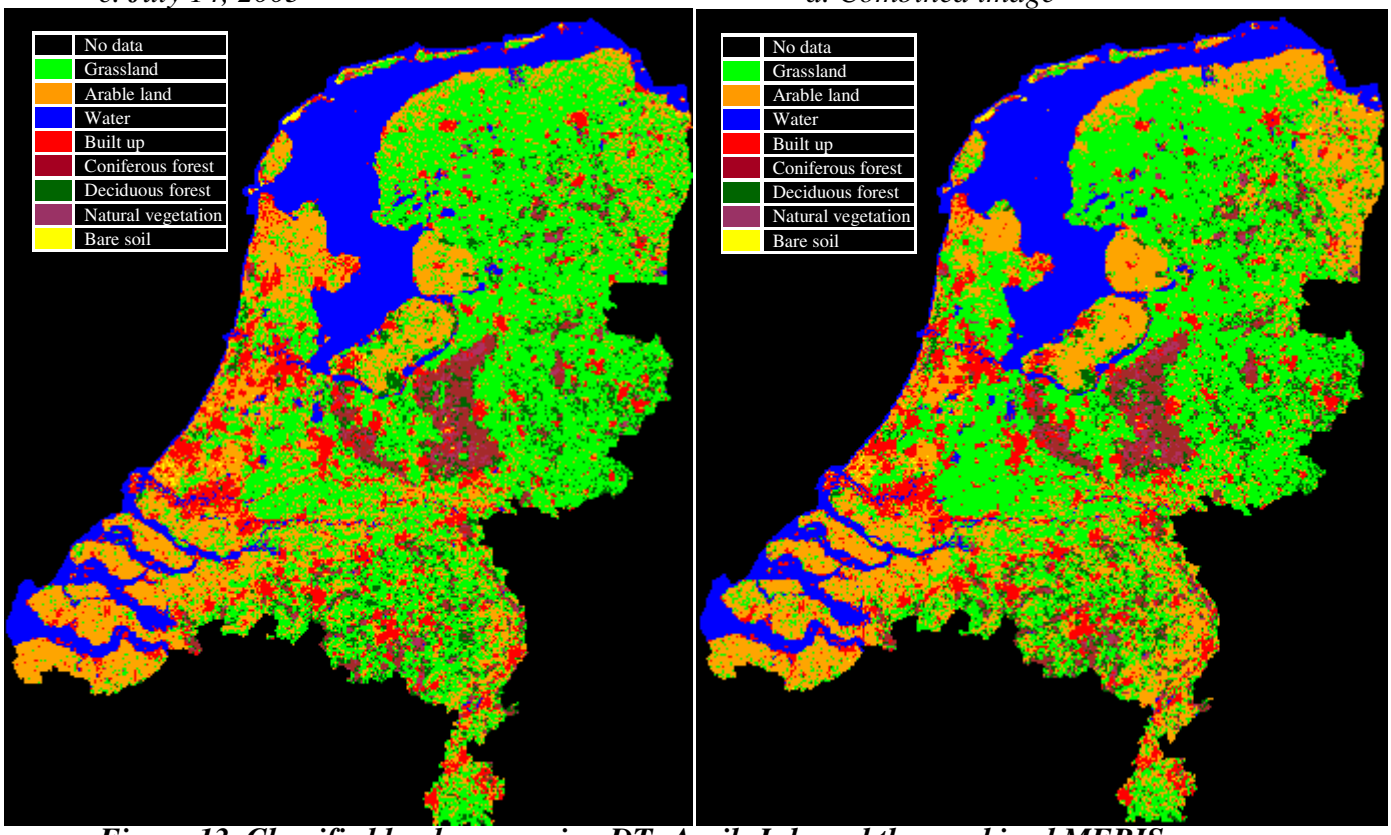


Figure 13. Classified land cover using DT: April, July and the combined MERIS images with reference to LGN4

4.5 Comparison among different classifiers

MLH, MDM, and DT classifiers were used for classification to see the performance of each classifier and to make comparison among them.

Type of classifier	April image	July image	Combined image
Maximum likelihood (MLH)	69.76%	71.11%	70.96%
Decision tree (DT)	68.56%	67.13%	70.84%
Minimum distance to mean (MDM)	63.07%	63.31%	66.18%

Table 15. Comparison of MLH, MDM, and DT classifiers using overall classification accuracy for the validation dataset *(see appendix C)*

From table 15 we can observe that the overall classification accuracy of MLH and DT classifiers were almost the same except for the July image. The performance of the MDM classifier was less than the two other classifiers for all images. The possible reason for this may be due to insensitivity of the MDM classifier for different variance in the spectral response data. The accuracy of the MLH classifier for the July image was higher than for the DT classifier. This may be due to the high spectral overlap between grassland and arable land that makes it difficult to separate them by the DT classifier which utilizes one spectral feature at a time for decision. The accuracy obtained by the DT and MDM classifiers for the combined image (multitemporal image) was better than the monotemporal images (the April and July image). This shows that, using only two dates, seasonality plays an important role in classification of different land covers.

The overall classification accuracy of all the images using the three classifiers was reduced by about 4-5% when using the whole land cover database (LGN4) as a reference (Appendix G-table I). This may be because the samples used for validation was not 100% representative of the total population dataset (LGN4).

5. Conclusion and Recommendation

5.1 Conclusion

The main objective of the study described here is to assess the advantage of multitemporal land cover classification using full resolution MERIS images and to compare different types of classification methods used in remote sensing. Accordingly, three classifiers (MLH, MDM and DT) were applied on two MERIS images from different dates (April 16, and July 14, 2003) and the combined MERIS image.

The result presented in this study provided some analysis on the use of different classification algorithms and the merit of a combined image for land cover classification in the Netherlands using full resolution MERIS images.

First, the results of all the three classifiers were satisfactory. The MLH and DT classifiers were performing better than the MDM classifier. The MLH classifier seems better than the DT classifier on the single date images. The classification with the combined imagery (multitemporal) has shown improved accuracies for the DT and MDM classifiers as compared to the MLH one. The size of the training dataset and splitting rule have great influence on the performance of the DT classifier. The training dataset at 3450 pixels with Gini splitting rule gave better result. The study also concludes that the data dimensionality has no effect on DT classifier. The advantage of the DT classifier was that it weights the importance of discriminating variables and thereby ignored redundant variables.

Water was quite well classified by all the classifiers in all the images. Considering different factors that affect classification accuracy, particularly within this study, grassland, arable land and built-up area were also well classified by the MLH and DT classifiers in all the images and by the MDM classifier in the combined image. Basically, coniferous forest, deciduous forest, natural vegetation and bare soil were not well classified in the images by all the classifiers. Coniferous forest and deciduous forest were the classes that confused with each other in all the images. Natural vegetation was mainly confused with grassland and arable land. This concludes that these classes have similar spectral behavior so they could not be classified well. Moreover, the main confusion was also occurring between grassland and arable land in all the images.

Generally, what can be concluded is that multitemporal analysis is quite useful to classify land cover types but the improvement of classification accuracy of land cover classification using multitemporal data is not only depending on the attributes of data but also on the algorithms used to classify land cover.

5.2 Recommendation

- From a more general perspective, it is clear that the classification results produced by supervised algorithms are highly reliant on the quality and representativeness of the training data used. Thus, care must be taken during selection of spectrally representative training data.
- There is no single classification algorithm that is expected to provide maximum accuracies with all data since the statistical method that can best distinguish between classes is likely dependent upon the specific attributes of the data such as resolution (spectral, spatial, radiometric, and temporal) and quality of training data. Accordingly,
    - MLH algorithm is highly recommended for land cover classification if no ancillary datasets are available.
    - If ancillary datasets are available, DT classifier will be highly preferable.
- Based on the results obtained, grassland and arable land could be merged to one class agriculture and coniferous and deciduous forest could also be merged to one class forest since they show significant spectral overlap in all images.
- Traditional methods of land cover classification have typically relied on image-derived variables, but evidence from multitemporal land cover mapping studies indicates that including non-spectral variables may help to improve discrimination between land cover classes (Rogan et al., 2003). Therefore, to improve the accuracy and to extend the work to regional scale use of more than two MERIS images at different dates in the season and use of ancillary datasets has to be given attention.

6. References

Addink, E.A. (2001). "Change detection with remote sensing; relating NOAA_AVHRR to environmental impact of agriculture in Europe." PhD thesis, Wageningen University, The Netherlands.

Bossard, M., Feranec, J., and Otahel, J., (2000). "CORINE land cover technical-Addendum2000." Technical report No 40. (visited on June 1, 2004, http://reports.eea.eu.int/tech40add/en/tech40add.pdf)

Breiman, L., Friedmen, J.H., Olshen, R.A., and Stone, C.J. (1984)."Classification and Regression Trees." Wadsworth, Belmont, CA.

Brown de Colstoun, E. C., Story, M. H., Thompson, C., Commisso, K., Smith, T. G., and Irons, J. R. (2003). "National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier." Remote Sensing of Environment **85**: 316-327.

Cihlar, J. (2000). "Land cover mapping of large areas from satellites: status and research priorities." International Journal of Remote Sensing **21**: 1093-1114.

Clevers, J.G.P.W., Bartholomeus, H.M., Mucher, C.A., and Wit, A.J.W. (2003) "Use of MERIS data for land cover mapping in The Netherlands." Proceedings of MERIS User Workshop, Frascati, Italy, November 10-13, 2003, Report SP-549, May 2004.

Congalton, R.G., and Green, K. (1999). "Assessing the accuracy of remotely sensed data: principles and practice." New York: Lewis publishers

Defries, R.S., and Chan, J.C.W. (2000). "Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data." Remote Sensing of Environment **74**: 503-515.

Defries, R.S., and Townshend, J.R.G. (1994). "NDVI-derived land cover classification at global scales." International Journal of Remote Sensing **15**: 3567-3586.

Erbek, F.S., Ozkan, C., and Taberner, M. (2004). "Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities." International Journal of Remote Sensing **25**: 1733-1748.

Foody, G.M., and Arora, M.K. (1997). "An evaluation of some factors affecting the accuracy of classification by an artificial neural network." International Journal of Remote Sensing **18**: 700 -810.

Foody, G.M., McCulloch, M.B., and Yates, W.B. (1995). "The effect of training set size and composition on artificial neural network classification." International Journal of Remote Sensing **16**: 1707 -1723.

Foody, G.M., Campbell, N.A., Trood, N.M., and Wood, T.F. (1992). "Derivation and application of probabilistic measures of class membership from the maximum likelihood classification." Photogrammetric Engineering & Remote Sensing **59**: 1335-1341.

Friedl, M. A. and Brodley, C. E. (1997). "Decision tree classification of land cover from remotely sensed data." Remote Sensing of Environment **61**: 399-409.

Friedl, M.A., Brodley, C.E., and Strahler, H. (1999). "Maximizing land Cover Classification Accuracy Produced by Decision Trees at Continental to Global Scales." IEEE Transactions on Geoscience and remote sensing **37**: 969-977.

Fuller, R.M., Smith, G.M., Sanderson, J.M., Hill, R.A., and Thomson, A.G. (2002). "The UK land cover map2000: construction of a parcel-based vector map from satellite image." The Cartographic Journal **39**: 15-25.

Hansen, M.C., Defries, R.S., Townshend, J.R.G., and Sohlberg, R. (2000). "Global land cover classification at 1km spatial resolution using a classification tree approach." International Journal of Remote Sensing **21**: 1331-1364.

Hansen, M.C., Dubayah, R., and Defries, R. (1996). "Classification trees: an alternative to traditional land cover classifiers." International Journal of Remote Sensing **17**: 1075-1081.

Jeon, B., and Landgrebe, D.A. (1999). "Decision fusion approach for multitemporal classification." IEEE Transactions on Geoscience and remote sensing **37**: 1227-1233.

Keuchel, J., Naumann, S., and Heiler, M. (2003). "Automatic land cover analysis for Tenerife by supervised classification using remotely sensing data." Contribution in magazine into remote Sensing of Environment.

Lawrence, R., and Labus, M. (2003). "Early detection of Douglas-fir beetle infestation with sub canopy resolution hyperspectral imagery." Western Journal of Applied forestry **18**: 202-206.

Lawrence, R.L., and Wright, A. (2001). "Rule –based Classification systems using classification and regression tree (CART) analysis." Photogrammetric Engineering & Remote Sensing **67**: 1137-1142.

Lewis, R.J. (2000). "An introduction to classification and regression tree (CART) analysis." Department of Emergency Medicine, Harbor-UCLA Medical center, Torrance, California

Lillesand, T.M., and Kiefer, R.W. (2000). "Remote sensing and Image interpretation." New York: Wiley and Sons.

Loveland, T.R., and Belward, A.S. (1997). "The IGBP-DIS global 1 km land cover data set, DIScovere: first results." International Journal of Remote Sensing **18**: 3289-3295.

Mather, P.M. (1999). "Computer processing of remotely-sensed images": An introduction (2nd ed.) Chichester, Wiley.

McIver, D.K., and Friedl, M. A. (2002). "Using prior probabilities in decision tree classification of remotely sensed data." Remote Sensing of Environment **81**: 253-261.

Mucher, C.A., Steinnocher, K.T., Kresler, F.P., and Heunks, C. (2000). "Land cover characterization and change detection for environmental monitoring of Pan-Europe." International Journal of Remote Sensing **21**:1159-1181.

Oetter, D.R., Cohen, W.B., Berterretche, M., Maiersperger, T.K., & Kennedy, R.E. (2000). "Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data." Remote Sensing of Environment **76**: 139-155.

Pal, M. and Mather, P.M. (2003). "An assessment of the effectiveness of decision tree methods for land cover classification." Remote Sensing of Environment **86**: 554-565.

Pax-Lenney, M. and Woodcock, C.E. (1997). "Monitoring agricultural lands in Egypt with multitemporal Landsat TM imagery: How many images are needed?" Remote Sensing of Environment **59**: 522-529.

Rogan, J., Millar, J., Stow, D., Franklin, J., Levien, L., and Fischer, C. (2003). "Land cover change monitoring with classification trees using Landsat TM and ancillary data." Photogrammetric Engineering & Remote Sensing **69**: 793-804.

Salford systems. (2002). "Classification and Regression Tree (Tree-Structured Non-parametric data analysis." San Diego, California.

Shahshahani, B.M., and Landgrebe, D.A. (1994). “The effect of unlabelled samples in reducing the small sample problem and mitigating the Hughes phenomenon.” ILEE. Transaction on Geoscience and Remote sensing **32**: 1087-1095

Speybroeck, N., Berkvens, D., Mfoukou-Ntsakala, A., Aerts, M., Hens, N., Van Huylenbroeck, G., & Thys, E. (2004). "Classification trees versus multinomial models in the analysis of urban farming systems in Central Africa." Agricultural Systems **80:**133-149.

Swain, P.H., and Hauska, H. (1977). “The decision tree classifier: Design and potential.” IEEE Transactions on Geo Science Electronics **3**: 142-147.

Thomas, I.L., Benning, V.M., and Ching, N.P. (1987). “Classification of remotely sensed image.” Bristol: Adam Hilger.

Thunnissen, H.A.M., Jaarsma, M.N., and Schouwmans, O.F. (1992). “Land cover inventory in the Netherlands using remote sensing; application in soil and ground water vulnerability assessment system.” International Journal of Remote Sensing **13**: 1693-1708

Verstraete, M.M., Pinty, B., and Curran, P.J. (1999). “MERIS potential for land application.” International Journal of Remote Sensing **20**: 1747-1756.

Yang, C.-C., Prasher, S. O., Enright, P., Madramootoo, C., Burgess, M., Goel, P. K., & Callum, I. (2003). "Application of decision tree technology for image classification using remote sensing data." Agricultural Systems **76**: 1101-1117.

Zhan, Q. (2003). “A hierarchical object-based approach for urban land use classification from remote sensing data.” Enschede; Wageningen: ITC; Wageningen University, PhD thesis, Summaries in Dutch and English (ITC Dissertation; 103)

Web site

URL 1 (http://www.travelblog.org/world/nl-geog.html accecced 26/11/2004)

URL 2 (http://envisat.esa.int/dataproducts/ accessed on 26/11/2004

URL 3 (http://www.lgn.nl accessed 27/11/2004)

7. Appendices

Appendix A. Projection information

Table I. Projection information

Projection type	Stereographic
Spheroid name	Clarke 1866
Datum name	Clarke 1866
Longitude of center of projection	5:23:15.000000E
Latitude of center of projection	52:09:22.006800N
False easting	155000.000000 meters
False northing	463000.000000 meters

Appendix B. Descriptions of the aggregated 9 main land covers from 39 LGN4 classes

Table I. Descriptions of the aggregated 9 main land covers from 39 LGN4 classes

Main land covers	Klasse naam(Dutch)	Class name(English)	Class value
Grass land	Gras	Pastures	1
	Gras in bebouwd gebied	Grass in urban area	1
	Veenweidegebied	Swampy pastures in peat areas	1
	Overig open begroeid natuurgebied	Other open nature areas	1
Arable land	Mais	Maize	2
	Aardappelen	Potato	2
	Bieten	Sugar beet	2
	Granen	Cereals	2
	Overige landbouwgewassen	Other agricultural crops	2
	Boomgaarden	Orchards	2
	Bloembollen	Flower bulb cultivation	2
Green house	Glastuinbouw	Green houses	3
Deciduous forest	Loofbos	Deciduous forest	4
	Loofbos in bebouwd gebied	Deciduous forest in urban area	4
	Bos in moerasgebied	Forest in swamp area	4
Coniferous forest	Naaldbos	Coniferous forest	5
	Naaldbos in bebouwd gebied	Coniferous forest in urban area	5
	Bos met dichte bebouwing	Forest with dense built up	5
	Bos in hoogveengebied	Forest in raised bogs	5
Water	Zoet water	Fresh water	6
	Zout water	Salt water	6
Built-up area	Bebouwing in agrarisch gebied	Building in agricultural area	7
	Stedelijk gebied	Urban area	7
	Bebouwing in buitengebied	Building in rural area	7
	Hoofdwegen en spoorwegen	Main roads and railroads	7
Bare soil	Kale grond in bebouwd gebied	Bare soil in rural area	8
	Open zand in kustgebied	Bare soil in coastal areas	8
	Open duinvegetatie	Sparsely vegetated dune	8
	Open stuifzand	Shifting sands	8
	Kale grond in natuurgebied	Bare soil in nature area	8
Natural vegetation	Kwelders	Salt marshes	9
	Gesloten duinvegetatie	Vegetated dune	9
	Duinheide	Heath in coastal area	9
	Heide	Heath lands(1)	9
	Matig vergraste heide	Heath lands(2)	9
	Sterk vergraste heide	Heath lands(3)	9
	Hoogveen	Raised peat bogs	9
	Overige moerasvegetatie	Other swamp vegetation	9
	Rietvegetatie	Reed swamp	9

Appendix C. Accuracy assessment of mono and multtemporal MERIS images

Table I. Confusion Matrix MLH classifier of April 16, 2003 MERIS image

	Reference data									
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Grass land	**551**	61	12	16	14	21	0	2	677	**81.39%**
Arable land	248	**468**	11	21	19	70	2	4	843	**55.52%**
Deciduous forest	38	19	**50**	33	2	8	3	7	160	**31.25%**
Coniferous forest	3	1	11	**80**	0	3	0	3	101	**79.21%**
Water	9	0	3	0	**357**	4	2	1	376	**94.95%**
Built up	38	8	2	3	8	**203**	1	0	263	**77.19%**
Bare soil	0	0	0	0	7	0	**11**	4	22	**50.00%**
Natural vegetation	18	2	4	2	6	5	0	**31**	68	**45.59%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**60.88%**	**83.72%**	**53.76%**	**51.61%**	**86.44%**	**64.65%**	**57.89%**	**59.62%**		
Overall accuracy = 69.76%								**Kappa Statistics = 0.6147**		

Table II. Confusion Matrix MLH classifier of July 14, 2003 MERIS image

	Reference data									
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Grass land	**584**	71	19	12	14	22	2	4	728	**80.21%**
Arable land	211	**457**	9	16	15	63	2	5	778	**58.74%**
Deciduous forest	28	12	**50**	34	2	0	0	3	129	**38.76%**
Coniferous forest	6	3	7	**74**	0	2	0	2	94	**78.72%**
Water	9	0	0	0	**359**	5	1	0	374	**95.99%**
Built up	39	8	2	2	4	**211**	1	0	267	**79.03%**
Bare soil	9	1	0	0	13	2	**12**	0	37	**32.43%**
Natural vegetation	19	7	6	17	6	9	1	**38**	103	**36.89%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**64.53%**	**81.75%**	**53.76%**	**47.74%**	**86.92%**	**67.20%**	**63.16%**	**73.08%**		
Overall accuracy = 71.11%								**Kappa Statistics = 0.6304**		

Table III. Confusion Matrix MLH classifier of combined (April 16 and July 14, 2003) MERIS image

Reference Data										
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Grass land	**565**	53	17	19	10	23	0	2	689	**82.00%**
Arable land	265	**484**	13	29	26	74	6	7	904	**53.54%**
Deciduous forest	24	14	**54**	34	1	2	0	3	132	**40.91%**
Coniferous forest	3	1	6	**67**	0	1	0	2	80	**83.75%**
Water	8	0	1	0	**358**	5	1	1	374	**95.72%**
Built up	29	7	1	2	7	**208**	0	0	254	**81.89%**
Bare soil	0	0	0	0	9	0	**12**	4	25	**48.00%**
Natural vegetation	11	0	1	4	2	1	0	33	52	**63.46%**
Column Total	905	559	93	155	413	314	19	**52**	2510	
Producer's accuracy	**62.43%**	**86.58%**	**58.06%**	**43.23%**	**86.68%**	**66.24%**	**63.16%**	**63.46%**		
Overall accuracy = 70.96%						**Kappa Statistics = 0.6268**				

Table IV. Confusion Matrix MDM classifier of April 16, 2003 MERIS image

Reference data										
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Unclassified	2	0	0	0	0	0	0	0	2	
Grass land	**596**	88	14	15	13	23	0	3	752	**79.26%**
Arable land	127	**296**	6	5	22	39	5	6	506	**58.50%**
Deciduous forest	114	70	**44**	37	5	52	2	2	326	**13.50%**
Coniferous forest	8	7	19	**88**	0	5	0	3	130	**67.69%**
Water	8	1	2	0	**353**	3	2	4	373	**94.64%**
Built up	39	64	1	1	13	**173**	4	7	302	**57.28%**
Bare soil	0	0	0	0	1	1	**6**	0	8	**75.00%**
Natural vegetation	11	33	7	9	6	18	0	**27**	111	**24.32%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**65.86%**	**52.95%**	**47.31%**	**56.77%**	**85.47%**	**55.10%**	**31.58%**	**51.92%**		
Overall accuracy = 63.07%						**Kappa Statistics = 0.5376**				

Table V. Confusion Matrix MDM classifier of July 14, 2003 MERIS image

Reference data										
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Unclassified	5	0	0	0	0	0	0	0	5	
Grass land	**665**	286	16	18	19	43	3	9	1059	**62.80%**
Arable land	58	**200**	2	1	4	8	0	0	273	**73.26%**
Deciduous forest	43	26	**56**	16	1	0	0	5	147	**38.10%**
Coniferous forest	6	2	10	**69**	0	2	0	4	93	**74.19%**
Water	10	0	0	0	**359**	6	3	1	379	**94.72%**
Built up	52	24	2	0	13	**199**	4	0	294	**67.69%**
Bare soil	1	1	0	0	6	2	**8**	0	18	**44.44%**
Natural vegetation	65	20	7	51	11	54	1	**33**	242	**13.64%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**73.48%**	**35.78%**	**60.22%**	**44.52%**	**86.92%**	**63.38%**	**42.11%**	**63.46%**		
Overall accuracy = 63.31%						**Kappa Statistics = 0.5281**				

Table VI. Confusion Matrix MDM classifier of combined (April 16 and July 14, 2003) MERIS image

Reference Data										
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Unclassified	3	0	0	0	0	0	0	0	3	
Grass land	**581**	82	6	13	12	28	0	2	724	**80.25%**
Arable land	124	**358**	7	2	10	34	4	8	547	**65.45%**
Deciduous forest	102	88	**63**	32	5	11	0	4	305	**20.66%**
Coniferous forest	8	2	11	**72**	0	3	0	5	101	**71.29%**
Water	10	0	3	0	**358**	7	3	1	382	**93.72%**
Built up	44	13	1	0	17	**190**	3	0	268	**70.90%**
Bare soil	2	0	0	0	6	4	**7**	0	19	**36.84%**
Natural vegetation	31	16	2	36	5	37	2	**32**	161	**19.88%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**64.20%**	**64.04%**	**67.74%**	**46.45%**	**86.68%**	**60.51%**	**36.84%**	**61.54%**		
Overall accuracy = 66.18%						**Kappa Statistics = 0.5776**				

Table VII. Confusion Matrix DT classifier of April 16, 2003 MERIS image

	Reference Data									
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Grass land	**683**	149	20	25	21	44	0	3	945	**72.27%**
Arable land	101	**321**	4	3	19	40	8	5	501	**64.07%**
Deciduous forest	53	34	**46**	40	9	18	1	7	208	**22.12%**
Coniferous forest	10	5	17	**81**	0	5	0	3	121	**66.94%**
Water	9	1	2	0	**353**	4	3	1	373	**94.64%**
Built up	44	32	3	2	10	**202**	2	2	297	**68.01%**
Bare soil	0	0	0	0	1	0	**4**	0	5	**80.00%**
Natural vegetation	5	17	1	4	0	1	1	**31**	60	**51.67%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**75.47%**	**57.42%**	**49.46%**	**52.26%**	**85.47%**	**64.33%**	**21.05%**	**59.62%**		
Overall accuracy = 68.56%						**Kappa Statistics = 0.5925**				

Table VIII. Confusion Matrix DT classifier of July 14, 2003 MERIS image

	Reference Data									
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Grass land	**645**	233	20	32	20	40	2	10	1002	**64.37%**
Arable land	154	**294**	5	4	9	44	8	6	524	**56.11%**
Deciduous forest	29	18	**55**	24	2	2	0	6	136	**40.44%**
Coniferous forest	11	3	10	**87**	1	3	0	5	120	**72.50%**
Water	9	0	0	0	**363**	6	5	0	383	**94.78%**
Built up	53	11	3	1	18	**217**	3	2	308	**70.45%**
Bare soil	0	0	0	0	0	0	**1**	0	1	**100.00%**
Natural vegetation	4	0	0	7	0	2	0	**23**	36	**63.89%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**71.27%**	**52.59%**	**59.14%**	**56.13%**	**87.89%**	**69.11%**	**5.26%**	**44.23%**		
Overall accuracy = 67.13%						**Kappa Statistics = 0.5683**				

Table IX. Confusion Matrix DT classifier of combined MERIS image

Reference Data										
Classified Data	Grass land	Arable land	Deciduous forest	Coniferous forest	Water	Built up	Bare soil	Natural vegetation	Row Total	**User's accuracy**
Grass land	**681**	159	20	31	21	40	2	4	958	**71.08%**
Arable land	121	**358**	7	7	12	43	7	13	568	**63.03%**
Deciduous forest	24	25	**51**	23	2	3	0	5	133	**38.35%**
Coniferous forest	13	4	12	**86**	1	3	0	5	124	**69.35%**
Water	9	0	0	0	**360**	6	5	0	380	**94.74%**
Built up	53	12	3	1	17	**217**	3	2	308	**70.45%**
Bare soil	0	0	0	0	0	0	**2**	0	2	**100.00%**
Natural vegetation	4	1	0	7	0	2	0	**23**	37	**62.16%**
Column Total	905	559	93	155	413	314	19	52	2510	
Producer's accuracy	**75.24%**	**64.04%**	**54.84%**	**55.48%**	**87.17%**	**69.11%**	**10.53%**	**44.23%**		
Overall accuracy = 70.84%						**Kappa Statistics = 0.6212**				

Appendix D. Principal component analysis of Combined MERIS image

Table I. Principal component analysis of Combined MERIS image

		Eigenvalue	% variance	Cumulative sum	Cumulative sum (%)
1	PC1	1678993.732	0.9142958	0.914296	91.43
2	PC2	100254.9768	0.0545938	0.96889	96.89
3	PC3	35138.59904	0.0191347	0.988024	98.80
4	PC4	16338.3813	0.0088971	0.996921	99.69
5	PC5	3817.400889	0.0020788	0.999	99.90
6	PC6	721.68027	0.000393	0.999393	99.94
7	PC7	353.370768	0.0001924	0.999586	99.96
8	PC8	238.34064	0.0001298	0.999715	99.97
9	PC9	155.451067	8.465E-05	0.9998	99.98
10	PC10	112.689531	6.137E-05	0.999861	99.99
11	PC11	96.222471	5.24E-05	0.999914	99.99
12	PC12	57.341948	3.123E-05	0.999945	99.99
13	PC13	25.684587	1.399E-05	0.999959	100.00
14	PC14	24.615182	1.34E-05	0.999972	100.00
15	PC15	14.218925	7.743E-06	0.99998	100.00
16	PC16	11.858936	6.458E-06	0.999987	100.00
17	PC17	7.843401	4.271E-06	0.999991	100.00
18	PC18	5.425079	2.954E-06	0.999994	100.00
19	PC19	4.931253	2.685E-06	0.999997	100.00
20	PC20	2.230814	1.215E-06	0.999998	100.00
21	PC21	1.385309	7.544E-07	0.999999	100.00
22	PC22	1.155124	6.29E-07	0.999999	100.00
23	PC23	0.987748	5.379E-07	1	100.00
24	PC24	0.536002	2.919E-07	1	100.00
	Total	1836379.059			

Appendix E. Correlation matrix for a subset of MERIS images

Table I. Correlation matrix for a subset of July 14, 2003 MERIS image

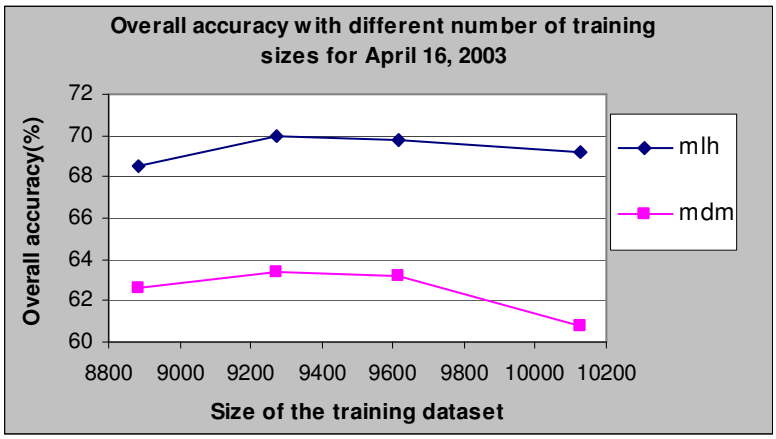
	3	4	5	6	7	8	9	10	12	13	14	15
3	1											
4	0.996	1										
5	0.911	0.939	1									
6	0.939	0.957	0.960	1								
7	0.926	0.940	0.924	0.992	1							
8	0.916	0.930	0.917	0.988	0.999	1						
9	0.355	0.407	0.645	0.602	0.590	0.603	1					
10	-0.315	-0.269	0.016	-0.096	-0.111	-0.093	0.696	1				
12	-0.337	-0.291	-0.007	-0.118	-0.133	-0.115	0.678	0.999	1			
13	-0.352	-0.306	-0.023	-0.128	-0.141	-0.122	0.678	0.997	0.998	1		
14	-0.353	-0.308	-0.024	-0.129	-0.141	-0.123	0.679	0.997	0.998	0.999	1	
15	-0.361	-0.315	-0.033	-0.137	-0.149	-0.130	0.673	0.994	0.996	0.999	0.999	1

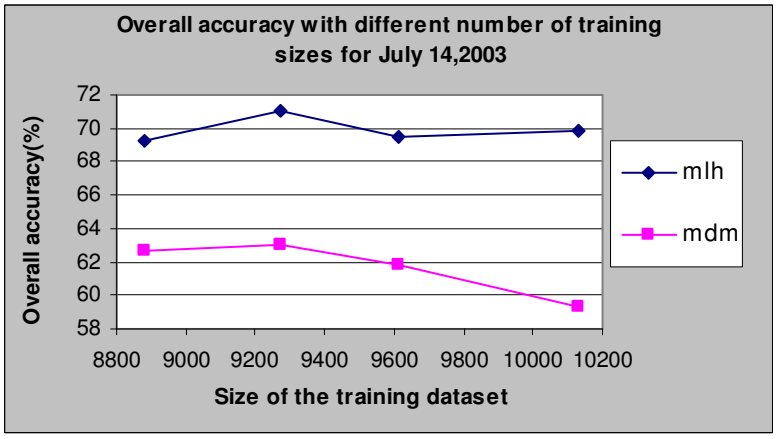
Table II. Correlation matrix for a subset of April 16, 2003 MERIS image

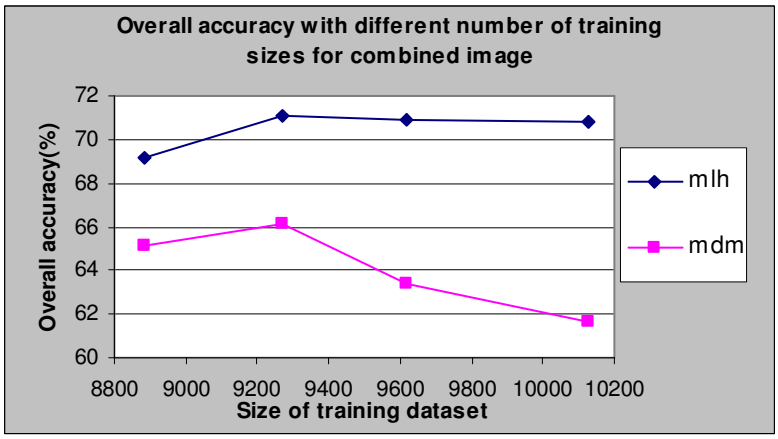
	3	4	5	6	7	8	9	10	12	13	14	15
3	1											
4	0.997	1										
5	0.972	0.985	1									
6	0.947	0.965	0.978	1								
7	0.891	0.915	0.931	0.985	1							
8	0.888	0.912	0.926	0.982	0.999	1						
9	0.727	0.754	0.823	0.846	0.863	0.865	1					
10	0.274	0.287	0.386	0.326	0.316	0.320	0.728	1				
12	0.128	0.146	0.257	0.207	0.214	0.219	0.657	0.985	1			
13	0.128	0.144	0.252	0.205	0.215	0.221	0.661	0.984	0.998	1		
14	0.205	0.218	0.318	0.264	0.264	0.270	0.695	0.993	0.990	0.994	1	
15	0.230	0.242	0.338	0.284	0.281	0.287	0.705	0.993	0.984	0.989	0.999	1

Appendix F. Overall classification accuracy of different sizes of training datasets of the MERIS images.

Figure I. Overall classification accuracy of April 16 and July 14, 2003, and the combined image based on different sizes of training datasets using MLH and MDM classifiers to select the best training datasets for final land cover classification.







Appendix G. Overall classification accuracy of April 16, and July 14, 2003, and the combined images by the MLH, MDM, and DT classifiers using the whole land database (LGN4) as reference.

Table I. Overall classification accuracy of all the images by the MLH, MDM, and DT classifiers using the whole land database (LGN4) as reference.

Type of classifier	April image	July image	Combined image
Maximum likelihood (MLH)	64.78%	66.11%	67.29%
Decision tree (DT)	62.88%	62.06%	66.26%
Minimum distance to mean (MDM)	57.48%	58.25%	61.83%