Uncertainty in lake extent trends related to time and frequency of observation

Julian Ijumulana
February, 2010

Uncertainty in lake extent trends related to time and frequency of observation

by

Julian Ijumulana

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: Geoinformatics

Thesis Assessment Board

Chair	: Prof.Dr. A. Stein
External examiner	: Ir. C.J. van der Sande
Supervisor	: Ms Dr.Ir. W. Bijker
Second supervisor	: Dr. N. Hamm

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Abstract

Increasing availability of large mixed archives of remotely sensed data have motivated their use in time series analysis in geographic phenomena. However, this increase requires appropriate image mining method as well as algorithms in which uncertainty inherent in these datasets is explicitly stated. In this study object oriented image analysis with subsequent fuzzy-rule based classification have been used to extract meaningful information about uncertain lake extent from a series of images acquired at irregular time interval by ASTER and ETM+ sensors between 1999 and 2009. The method involves transforming a discrete image into homogeneous regions that correspond to (part of) real world phenomenon through the process of segmentation. The multiresolution segmentation algorithm was used to generate a network of segments technically known as image objects. The algorithm applies optimization procedure that locally minimizes average heterogeneity of resulting image objects for a given resolution.

Fuzzy object-based classification was used in this study in which membership function values were determined based on various features of image objects. Each classified object was attributed with a membership function value expressing its uncertainty. Since the lake extent was determined by the total number individual objects, the weighted average membership was determined to express uncertainty in lake extent at a moment in time. Accuracy assessment of the resulting classes was evaluated by determining the class stability in which the least and best classified image object of that class was retrieved.

Lake extent time series was performed using linear regression modelling. Least squares method was used to estimate the best fit line representing the trend in lake extent. A combination of observations from the two sensors was perfomed after evaluating how strongly they are linearly related. The correlation coefficient of 0.71 was obtained revealing that the two sensors can be combined to obtain required number of observations for time series analysis in uncertain lake extent.

While analysing time series in lake extent, 97% of variation in the observations was well modelled by the prediction line during rainy season, 90% during dry season and 91% in all seasons. However, trend analysis in lake extent is complex because of wetlands with dominant vegetation whose size also varies with seasons. It is also difficulty due to natural and unpredictable water level fluctuations. With these preliminary results of time series analysis using remotely sensed data from mixed archives at irregular time interval, it is possible to study various geographic phenomena whose speed of change requires large number of observations for their detection.

Key words: Lake extent, uncertainty, object-oriented image analysis, multiresolution segmentation, fuzzy object-based classification, time series analysis

Acknowledgements

I take this opportune time to thank the following for their remarkable support during my MSc study in the Netherlands in general and the research period in particular.

First and foremost thanks go to the Netherlands Fellowship Programme (NFP) for granting me fully-funded scholarship that enabled me to pursue my studies comfortably.

Sincere thanks go to ITC administration for admitting me for MSc study, my lecturers and practical supervisors who enlighten me with new technologies related to earth sciences. May the Almighty God bless you all excessively!

In a special way I thank my supervisors Ms.Dr.Ir.W. Wietske Bijker and Dr. Nicholas Hamm for their continuous assessment and constructive ideas during research period. Their comments are really appreciated as they made me think deeply and come out with new ideas which made my research successful. Thank you very much and God bless you

I thank my wife, Elizabeth, for her consent to let me go away from her for a long period even without possibility of break. Her continuous prayers, courage and advice are all appreciated. May the Almighty God continue blessing her!

I also thank my parents, relatives and friends in- and outside Tanzania for their prayers during my stay in the Netherlands.

Lastly, I thank my fellow students GFM 2, 2008 for the support and company they have given me throughout my studies at ITC. Thank you very much for your company.

Table of contents

1. Introduction 1
   1.1. Introduction 1
   1.2. Background and problem statement 1
   1.3. Research objectives 3
      1.3.1. Research questions 4
   1.4. Inovation of this research 4
   1.5. Thesis structure 4
   1.6. Research approach 5
2. Literature review 6
   2.1. Introduction 6
   2.2. Uncertainty in geo-spatial objects 6
   2.3. Uncertainty in lake extent 6
      2.3.1. Wetlands 7
      2.3.2. Lake extent concept 8
   2.4. Modelling activities using remotely sensed data 9
   2.5. Sources of uncertainty in geographic objects 9
      2.5.1. Uncertainty during image acquisition 9
      2.5.2. Uncertainty during image pre-processing 10
      2.5.3. Uncertainty during modelling and monitoring 10
   2.6. Object-oriented image analysis 11
      2.6.1. Image segmentation 11
      2.6.2. Multiresolution image segmentation 11
      2.6.3. Image objects information 13
      2.6.4. Fuzzy logic classification 13
3. Study area, data and data pre-processing 15
   3.1. Introduction 15
   3.2. Study area description 15
   3.3. Study area selection criteria 16
   3.4. Characteristics of data used 16
      3.4.1. ASTER sensor 16
      3.4.2. ETM+ sensor 17
   3.5. Data availability and their quality 18
   3.6. Satellite image pre-processing 19
4. Methods 20
   4.1. Introduction 20
   4.2. Methodology flow chart 20
   4.3. Multiresolution image segmentation 21
      4.3.1. Homogeneity criteria 21
         4.3.1.1. Layer weighting parameters 21
         4.3.1.2. Scale parameter 24
         4.3.1.3. Colour and shape parameters 24

4.3.1.4. Compactness and smoothing parameters 24
4.4. Hierarchical image objects network 24
4.4.1. Segmentation parameters used 25
4.5. Fuzzy rule-based image objects classification 26
4.5.1. Estimation of fuzzy parameters 26
4.5.2. Layer mean intensity 26
4.5.3. Image object brightness 27
4.5.4. Individual layer contribution in each image object 27
4.5.5. Normalized Difference Vegetation Index (NDVI) 28
4.5.6. Implementation of NDVI model in Definiens software 29
4.5.7. Relative border to a defined class 29
4.5.8. Relative area of a defined class 30
4.5.9. Fuzzy object classification process 31
4.5.9.1. Definition of class hierarchy 31
4.5.9.2. Identification of image objects with class water membership 32
4.5.9.3. Identification of image objects that belong to class 'not water' 32
4.5.9.4. Parameters used to classify images 33
5. Results of object oriented image analysis 35
5.1. Introduction 35
5.2. Fuzzy objects classified as water 35
5.2.1. Estimated lake extent and its uncertainty from ASTER images 36
5.2.2. Estimated lake extent and its uncertainty from ETM+ images 36
5.3. Accuracy assessment of classified image objects 37
6. Time series analysis in lake extent 39
6.1. Introduction 39
6.2. Time series analysis in lake extent using ASTER images 39
6.2.1. Linear regression modelling: ASTER sensor 41
6.3. Time series analysis in lake extent using ETM+ images 43
6.3.1. Linear regression modelling: ETM+ sensor 44
6.4. Time series analysis using two sensors: ASTER and ETM+ sensor 46
6.4.1. Linear regression modelling of lake extent trend from two sensors 47
6.5. Comparing results from time series analysis 49
7. Discussion, conclusions and recommendations 50
7.1 Discussion 50
7.2 Conclusions 52
7.3 Recommendations 52
URL's: 55
References 55
Appendices 59

List of figures

Figure 1.1 Research methodology flow chart 5
Figure 2.1 Wetland occupy transition zone between water and land (Source : [12]) 8
Figure 2.2 Lake extent conceptual model 9
Figure 2.3 Composition of homogeneity criterion 12
Figure 2.4 Four-level hierarchical network of image objects in abstract illustration (Source: [14]) ... 13
Figure 2.5 Image objects information structure 13
Figure 3.1 Location of study area and corresponding Landsat7 ETM+ (A) and Terra ASTER (B) images acquired on the same date 15th October 2002 16
Figure 4.1 Methodology flow diagram of this research 20
Figure 4.2 Layer weighting procedure before segmentation: Landsat 7 ETM+) 23
Figure 4.3 Sample histogram for NIR band 23
Figure 4.4 Different image objects from different parameter sets 25
Figure 4.5 NDVI mathematical model implemented in Definiens 29
Figure 4.6 Relative border between neighbour objects 30
Figure 4.8 Fuzzy parameter estimation using green band and NDVI 31
Figure 4.9 Class hierarchy definition 32
Figure 4.10 Water objects identification 32
Figure 4.11 Classification map 33
Figure 5.1 Assessment of image objects classified and identification of objects that constitute to lake extent 38
Figure 6.1 Lake extent and its uncertainty during rainy season, ASTER images 40
Figure 6.2 Lake extent and its uncertainty during dry season, ASTER images 40
Figure 6.3 Lake extent and its uncertainty over time, all seasons, ASTER images 40
Figure 6.4 Lake extent trend from 2001 to 2009 from ASTER images 42
Figure 6.5 Lake extent and its uncertainty: Rainy season, ETM+ images 43
Figure 6.6 Lake extent and its uncertainty: Dry season, ETM+ images 44
Figure 6.7 Lake extent and its uncertainty over time, all seasons, ETM+ images 44
Figure 6.8 Lake extent trend from 1999 to 2002 from ETM+ images, all seasons 45
Figure 6.9 Lake extent and its uncertainty during rainy season : ASTER and ETM+ combined 46
Figure 6.10 Lake extent and its uncertainty during dry season: ASTER and ETM+ combined 46
Figure 6.11 Lake extent and its uncertainty from 1999 to 2009 :Observed from two sensors 47
Figure 6.12 Lake extent trend during rainy and dry season: Two sensors combined 48
Figure 6.13 Lake extent trend from 1999 to 2009: ASTER and ETM+ sensor combined 48

List of tables

Table 1.1 Research objectives and questions 4
Table 2.1 Various definitions of a lake and wetland 7
Table 3.1 3 spectral bands of ASTER (VNIR subsystem) 17
Table 3.2 Sensor/platform characteristics: ASTER 17
Table 3.3 7 spectral bands of ETM+ (Reflective) 17
Table 3.4 Sensor/platform characteristics: ETM+ 18
Table 3.5 Selected images and their corresponding acquisition dates 19
Table 4.1 Segmentation parameters applied in this study 26
Table 4.2 Rule set for classifying ASTER images 34
Table 4.3 Rule sets for classifying ETM+ images 34
Table 5.1 Image objects classified as water during dry season for ASTER images 36
Table 5.2 Image objects classified as water during rainy season for ASTER images 36
Table 5.3 Image objects classified as water during dry season for ETM+ images 37
Table 5.4 Image objects classified as water during rainy season for ETM+ images 37
Table 6.1 Weighted average lake extent information: Rainy season, ASTER images 39
Table 6.2 Weighted average lake extent information: Dry season, ASTER images 39
Table 6.3 Quality of the fitted trend line in different seasons: ASTER sensor 42
Table 6.4 Weighted estimated lake extent information: Rainy season, ETM+ images 43
Table 6.5 Weighted average lake extent information: Dry season, ETM+ images 43
Table 6.6 Quality of the fitted trend line in different seasons: ETM+ sensor 45
Table 6.7 Correlation between datasets during rainy season, dry season and all seasons in total 47
Table 6.8 Quality of the fitted trend line using measurements from two sensors 48
Table 6.9 Overview of results for ASTER, ETM+ and combined sensors 49

1. Introduction

1.1. Introduction

Sustainable use of available natural or artificial resources in a country is an important aspect towards restoring these resources. However, the sustainability of the resources requires detailed geo-information for proper planning and management. For centuries, ground surveys and aerial photographs have been used to collect detailed geo-information to characterise selected resources. Nevertheless, these methods are time- consuming and expensive since they require some investment and resources. In addition, these traditional survey methods were limited both in space and time thus limiting the monitoring activities. In space, the methods are limited by inaccessibility of some areas which results into improper sampling strategy when taking measurements while in time domain they are constrained by lack of enough capital (mostly less developed countries) to collect data that can be used to analyse trend of changing of resources of interest. Despite the limitations in these methods of surveying, they are considered crucial in providing geo-information that supports local resource management [11].

For the past few decades, earth observation (EO) technologies provide synoptic data at low or no cost, at refined spatial, spectral and temporal scales [6]. This advance in surveying technology has greatly mitigated constraints that hinder the performance of traditional surveys in a number of ways: provision of synoptic data at a point in time and over time regardless of intervisibility and accessibility, use of various visual characteristics of geographic objects such as colour, texture, shape and contextual information. Thus, this technology has been widely used in landcover/use inventorying, assessing and managing at national, regional and global levels [28]. However, the problem is how to extract meaningful information from these increasingly expanding mixed archives of remotely sensed data while quantifying inherent uncertainty in geographic phenomena. In addition to this problem, some of geographic phenomena require large number of observations for their detection due to the high speed of change in space and time which again limits the use of single sensor/platform combination observations. This again may require combining different sensor/platforms as a means of increasing the number of required observations necessary for change detection.

1.2. Background and problem statement

Increasing availability of remotely sensed data at low or no cost, at finer spatial, spectral and temporal resolutions has greatly raised many questions about applications of these data such as monitoring forest fires [49], flooding lakes [42] and dynamics in coastal landscape units [44]. However, many environmental processes occur at high speed which again requires an increased number of observations so as to track temporal patterns within geographic phenomena. In this context we may think of combining different measurements from different sources in the same time span. This will reduce uncertainty when modelling geographic phenomena of interest over time.

The aim of this study is to model the change in lake extent and its uncertainty from multiple sensors. The approach is to start from the definition of lake extent to identify what causes uncertainty in lake extent. Based on the lake definition as a large inland waterbody occupying a basin with no continuity to the sea or ocean [35], it is evident that the spatial extent of the lake is defined by the boundary points at which water interacts with land. Therefore, the lake extent is determined by the presence of water at the boundary of the lake. In most cases the lake boundary points are uncertain due to various environmental factors such as wind and water currents causing horizontal in- and outflow of water and other factors such as climate change or human influence. Therefore, what is observed at one time cannot be observed at another time thus causing uncertainty in identifying these points.

Depending on the history of formation of the lakes which determines the geology of the lake boundary, some of the areas fringing lakes turn into bogs, marshes and swamps related growth of vegetation. With time these areas no longer appear as part of lake, as vegetation become dominant and changes their characteristics. Since 1971 some of these areas have been recognised as wetlands of international importance by the international treaty for wise use of wetlands of international importance (RAMSAR) [29]. These areas again change with time particularly during rainy season when they are submerged. Thus, when observed at that time, they can be classified as part of lake therefore complicating the definition of the lake extent.

It is possible to argue that lake extent is larger in the rainy season than in the dry season. This is obvious since during rainy season fringing shoreline vegetation is submerged although it will depend on the type of vegetation and amount of rainfall. In contrast, the size of the lake will be small during dry season as water retreats towards the centre of the lake. Therefore, seasonality in mapping and monitoring of these important landscape features is of great importance for identifying underlying trend over time.

Remote sensing technology has been a major source of geo-information for many applications. The full usage of information contained in remotely sensed data requires appropriate tools that handle spatial relationships of patterns discernable on an image [30]. However, many applications rely on classification algorithms that were developed in the 1970's where a single pixel is class labelled in a multi-dimensional feature space [7]. These algorithms were developed based on signal processing concepts which cannot model the complex nature of real world objects. The development trend in classification algorithms shows that soft classifiers are developed to account for uncertainty inherent in remotely sensed data by incorporating fuzzy logic concepts [5, 51] developed in 1980's [52]. Despite this development real world objects relations are not modelled as these algorithms apply on pixel basis.

Recently, data mining methods for extracting meaningful information from large sets of observations with emphasis on uncertain objects are proposed [6, 41-42]. However, as a rule of thumb, proper selection of classification algorithm depends on understanding of the process to be modelled [41]. The knowledge about the phenomena to be modelled and monitored will lead to proper selection of a suitable dataset for that application as each dataset is collected for a particular application and has its own inherent uncertainty. In this study remotely sensed data from two sensors were used to analyze

trend in lake extent between 1999 and 2009. The datasets were selected on the basis of their similar characteristics.

In remote sensing, at a single point in time of observation, uncertainty depends on spatial resolution, spectral resolution and the definition of the geographic phenomenon under study [18, 41, 43]. Looking at the spatial resolution, the pixel itself contains uncertainty in location. The same object can appear different in size and shape, in consecutive images, if the pixels of both images are not recorded over the same areas of the Earth's surface. This problem has been addressed by Openshaw as a Modifiable Area Unit Problem (MAUP) [36]. In addition, the digital number recorded in a pixel contains uncertainty originating from ambiguity due to point spread function (PSF) of a sensor [17] and from distortions between object and the sensor or between source of illumination, object and sensor [6, 43]. However, there is also uncertainty in the geographic phenomenon to be modelled. Because of spatial resolution problem, objects may have crisp boundaries in reality but in images they are represented as vague objects and vice versa; this depends on the relative position of the object with respect to the sensor. Furthermore, there exists uncertainty due to sampling scheme when classifying images, since the grouping of pixels belonging to the same object involves subjective decision due to lack of spatial support to describe real world objects.

Since large sets of images are analysed together in image mining studies, uncertainty may arise from co-registration of the images, from uncontrollable differences in atmospheric parameters and illumination and from the temporal resolution of the observations, compared to the speed of the process [6]. Thus, conceptual understanding of different types and possible sources of uncertainty in image mining methods enables proper selection of remotely sensed data and classification algorithms incorporating models of uncertainty.

1.3. Research objectives

The overall objective of this research is to model the change and uncertainty in lake extent on series of remotely sensed images from multiple sensors. To achieve this broader objective, the following specific objectives should be achieved:

1. To estimate lake extent and its uncertainty at a point in time
2. To estimate lake extent and its uncertainty over time
3. To explore the impact of three sources of uncertainty in lake extent (object definition, data and model)

1.3.1. Research questions

Table 1.1 shows questions to be answered in order to achieve the aforementioned specific objectives

Table 1.1 Research objectives and questions

Research objective	Research questions
To estimate lake extent and its uncertainty at a point in time	1. How is the lake defined? 2. How should "lake extent" be defined 3. What is uncertainty in lake extent?
To estimate lake extent and its uncertainty over time	4. How does the lake extent change over time? 5. What is uncertainty in lake extent over time? 6. Are there time series patterns in lake extent?
To explore the impact of three sources of uncertainty in lake extent (object definition, data and classification)	7. Which temporal resolution is appropriate for studying temporal patterns in lake extent? 8. What are the key data issues (MAUP, temporal resolution and spectral resolution) influencing the monitoring of change in lake extent?

1.4. Inovation of this research

The novelty of this research is to model changes in lake extent and its uncertainty from a series of remotely sensed data from multiple sensors.

1.5. Thesis structure

This report comprises seven chapters. Chapter 1 introduces the problem to be solved through answering the posed research questions. Chapter 2 reviews the definition of the lake and the concept of lake extent is introduced. In addition, different sources of uncertainty in lake extent are identified in this chapter. Different approaches of handling uncertainty in geographic phenomena are also reviewed. Chapter 3 describes the case study area and datasets used in this study. It further describes the methods used to perform geometric correction of images. Chapter 4 explains the procedures followed in image analysis with subsequent classification. Chapter 5 presents the results of object-oriented image analysis and accuracy assessment methods used to test the reliability of image objects classified as water summing up to a total size of the lake. Chapter 6 explains procedures followed to perform lake extent time series analysis. Chapter 7 presents the discussion and conclusion based on the results. It ends up with some recommendations for further research based on the methods used in this research.

1.6. Research approach

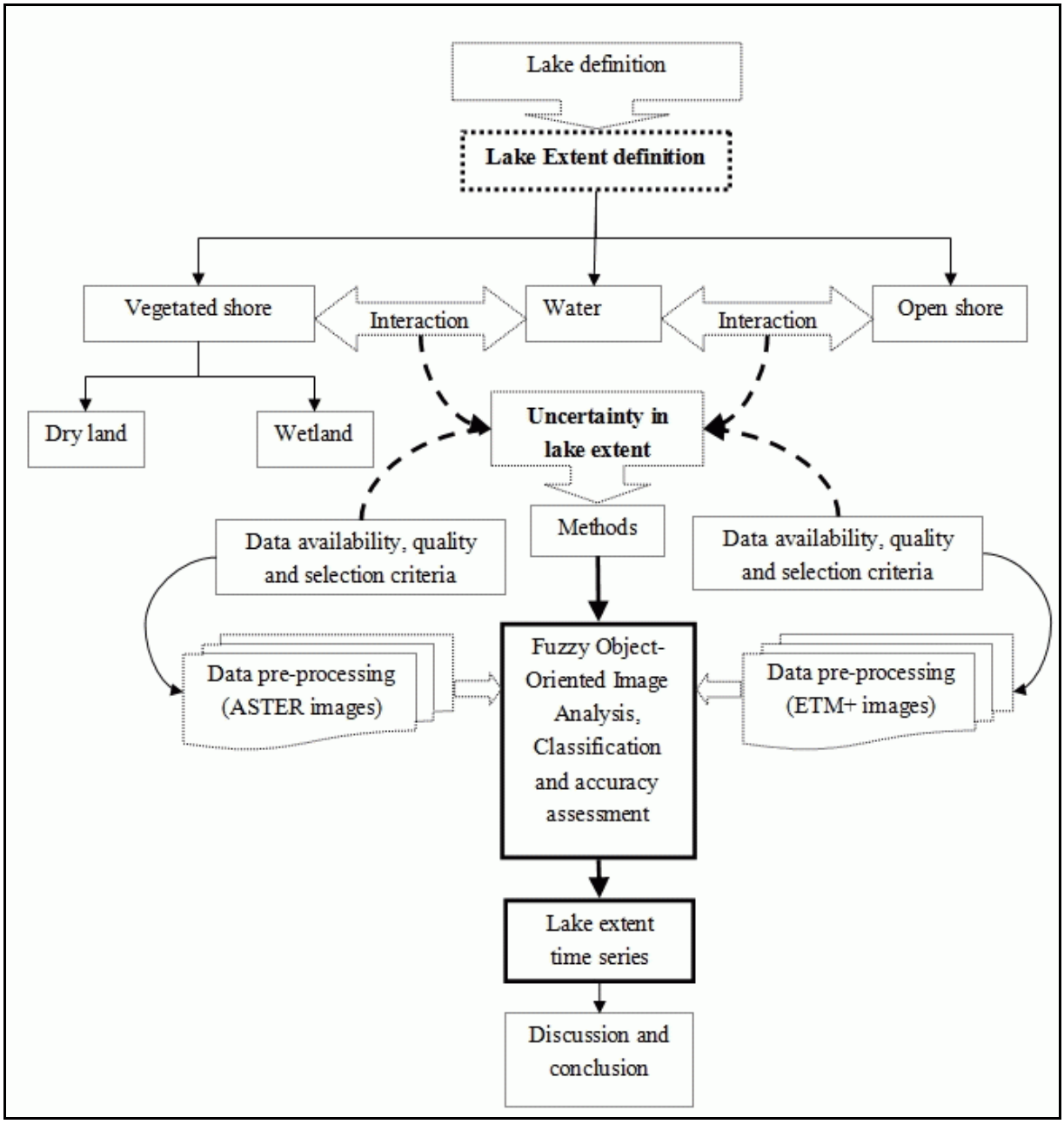


Figure 1.1 Research methodology flow chart

Figure 1.1 summarises procedures followed in this research. The first consideration was to review the definition of the lake in order to understand the lake extent while identifying sources of uncertainty in lake extent. The methods part briefly shows datasets and methods adopted to analyze images and quantification of uncertainty inherent in datasets. Lake extent time series part comprises methods used to estimate trend in lake extent using observations from two sensors.

2. Literature review

2.1. Introduction

This chapter reviews important concepts related to sources of uncertainty in geographical phenomena in general and sources of uncertainty in lake extent in particular.

2.2. Uncertainty in geo-spatial objects

Many monitoring activities of the environmental processes deal with phenomena that are uncertain. Their spatial extents are difficult to identify and their delineation depends on their thematic certainty [32]. For clearly defined objects, monitoring is not difficult as their geometric boundary points can easily be tracked over time. An object is considered vague and fuzzy if it cannot be precisely defined thematically and has transition zones at the boundary [42]. Mapping these objects requires a communicative capability that reflects the imprecise nature of a given geographical phenomenon. Traditionally, the mapping has been done using land surveying techniques and photogrammetry. However, these methods are limited in space and time to allow monitoring of changes that take place in these objects. This research proposes a new approach for handling uncertainty within the lake extent trends.

2.3. Uncertainty in lake extent

Definition of lake extent depends largely on the definition of the lake itself. However, I can define the lake extent as boundary points to which lake water can occupy. Table 2.1 summarises some of the concepts about the lake based on criteria as indicated in column 3 of the table. In this study the wetland understanding was of great importance as it occupies position between water and land [12]. At times this area can be classified as part of the lake while at another time as land thus complicating the definition of the lake extent especially when its monitoring for change detection is required. The classification as lake or water will again depend on the seasons when the observation was done. Apart from seasons, there are other explanatory variables of the extent of the lake such as wind that causes both vertical and horizontal water movement and other geomorphologic processes.

Many lakes are fringed by wetlands colonised by sub merged, deep-rooted and floating vegetation. Thus, the availability of these wetlands with variety of ecological units leads to vague definition of spatial extent of water (lake) at one moment in time. In addition, the problem becomes worse when monitoring activities to track changes in lake extent are required.

Table 2.1 Various definitions of a lake and wetland

Definition	Author	Criteria
A lake is a body of standing water occupying a basin and lacking continuity with the sea.	Forel [20] in O'Sullivan and Reynolds [35]	▪ Standing water in a basin ▪ No water outlet
A typical lake is deep enough for most of the bottom to be free of the rooted vegetation and is permanent	Bayly and Williams [3]	▪ Deep and permanent water ▪ Free of deep rooted vegetation
A wetland is an area of marsh, fen, peatland or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish or salt, including areas of marine water the depth of which at low tide does not exceed six metres."	Matthews [29]	▪ Permanent or temporary waterlogged areas ▪ Water depth < 6 meters
The term 'wetlands' groups together a wide range of inland, coastal and marine habitats which share a number of common features.	Dugan [15]	▪ Inland ecological habitats ▪ Coastal and marine habitats
The term 'wetland' is defined as temporarily or permanently wet ecosystems dominated by emergent vegetation	Harper D. M., *et al* [23]	▪ Temporary or permanent wet ecosystems ▪ Dorminant emergent vegetation

2.3.1. Wetlands

Wetlands are areas where there is either permanent or seasonal water associated with life of animals and plants and they occupy transition zones between land and water [12]. These areas exhibit both acquatic and terrestrial characteristics depending on seasons and other external influences such as wind and water currents. During rainy season, high water mark can be reached while water level is also possible to fall even below the low water mark. Therefore, water level observed at one moment in time can be different if observed at another time depending on the speed of water level fluctuation. Figure 2.1 shows the planimetric view of the wetland. During rainy season, the spatial extent of the lake will include (parts of) wetlands while during dry season the lake extent will exclude the wetland areas. This again will depend on the type of vegetation colonizing the wetland as well as the slope of the lake bed.

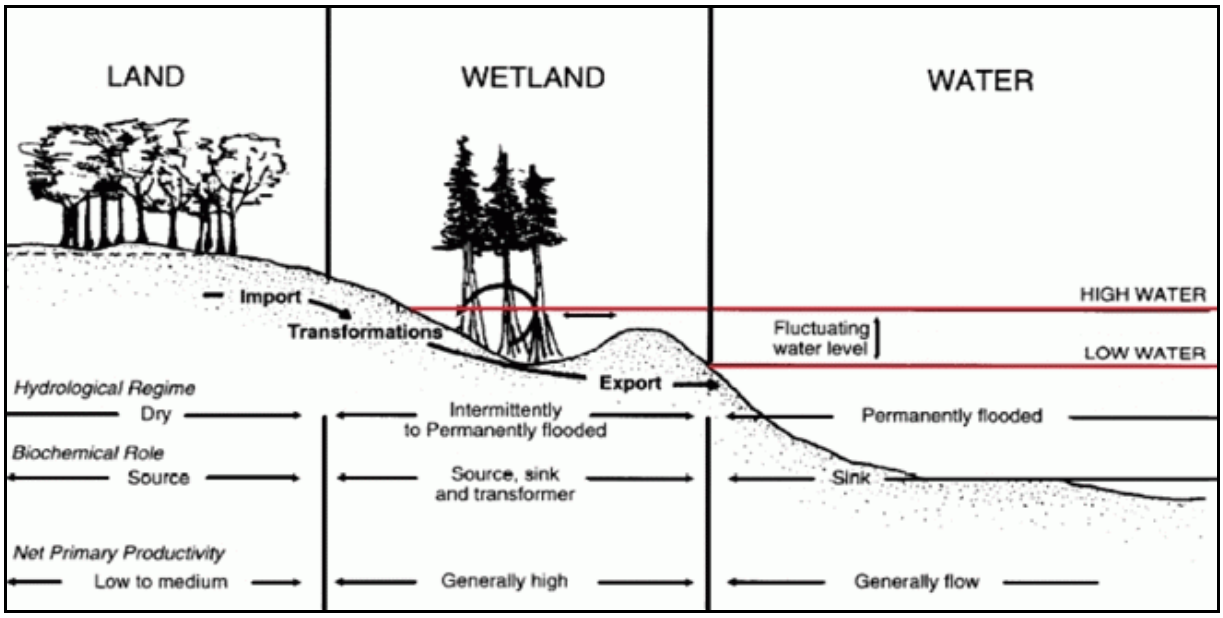


Figure 2.1 Wetland occupy transition zone between water and land (Source : [12])

Detailed inventory of wetlands of international importance has been a global agenda [16, 21, 29]. Although a detailed inventory of these wetlands has been prepared, less effort has been made to classify them as they depend on geographic latitude. The classification problem of wetland areas is due to their indeterminate nature especially where they are formed around large water bodies such as lakes, oceans, seas and rivers.

2.3.2. Lake extent concept

The lake can be defined as a large body of standing water localised to the basin, permanent and lacking continuity to the sea or ocean, deep enough for most of the bottom to be free of the deep rooted vegetation [3, 20, 35]. In the view of the wetland definition, majority of the lakes are likely to be surrounded by wetlands since the boundary points of the lake are interacting with land. However, the prevalence of the wetlands will depend on the nature of rocks and general slope of the lake basin. Figure 2.2 represents the situation that exists at the boundary of the lake. The spatial extent of the lake (defined by water-horizontal extension) at one moment in time depends on the state of shoreline. Some shorelines are ‘open’ while others are vegetated. Open shore in this context implies areas where at any moment in time water can clearly be discriminated from its background. For vegetated shore, it is difficult to discriminate water from vegetation especially where vegetation is not uniform throughout the lake boundary. Therefore, an appropriate strategy to map these areas is required so as to minimize the existing uncertainty.

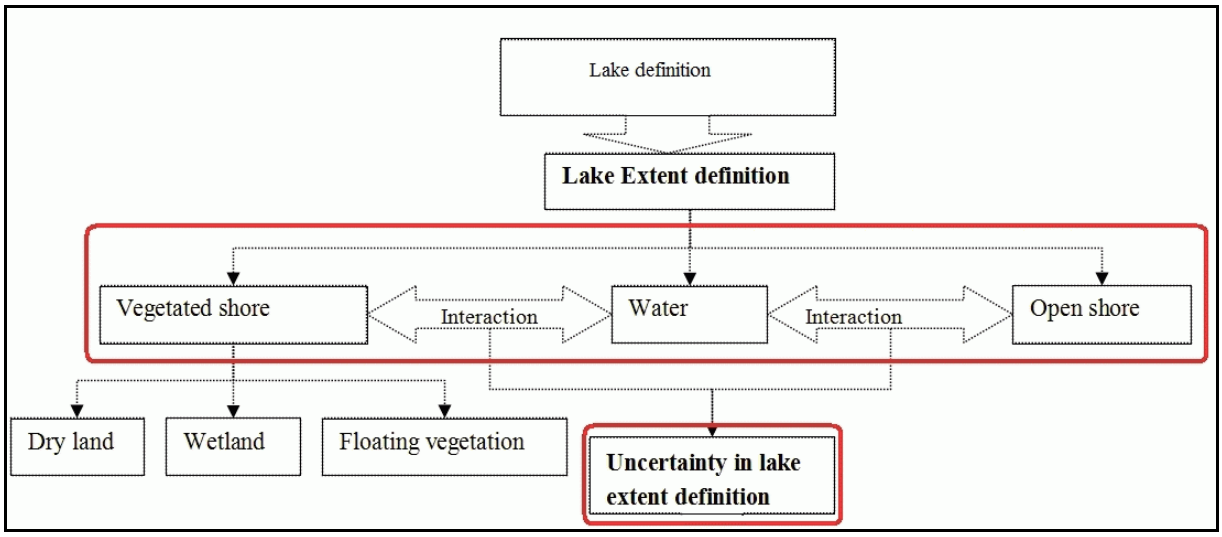


Figure 2.2 Lake extent conceptual model

2.4. Modelling activities using remotely sensed data

Over time, remote sensing technology has been used in monitoring activities of complex environmental processes such as forest fires [49] seasonal expanding lakes [42] and dynamics of coastal landscape features [44]. Full usage of information contained in remotely sensed data requires appropriate algorithms that incorporate models of uncertainty. Various models of uncertainty are well documented [18]. Although different techniques are well developed with sophisticated variations such as soft classifiers, sub-pixel classifiers and spectral un-mixing techniques, it is believed that they do not make use of real world object semantics such as shapes, size, orientation and texture. Recently, advanced image analysis algorithms have been developed. Markov random field (MRF) based classification is currently used in super resolution mapping of land cover[45-46]. Its suitability has been evaluated and it seems to provide promising results as it incorporates contextual information through the use of neighbourhoods [22]. However, the method is pixel based approach and does not include real world object semantics such as shape, size and spatial relationship of objects. Object-oriented image analysis is another advanced image analysis with subsequent classification approach in which various real world object features are considered during information extraction. With increasing availability of high-resolution satellite imagery, the use of spaceborne digital data the need for context-based algorithms and for object -oriented image processing are increasing as well. Recently available commercial products reflect this demand [14].

2.5. Sources of uncertainty in geographic objects

This section provides an overview of possible sources of uncertainty in geospatial objects. It further highlights possible approaches of handling them during image processing and subsequent modelling activities.

2.5.1. Uncertainty during image acquisition

Despite increasing the spatial resolution of satellite sensors, the uncertainty in locating geographic objects remains because of mixed land cover classes in an image. Due to pixel problem in location, some objects may appear vague on an image while in reality they are sharply defined. The problem

becomes serious when an object has transition zones at the boundary. On a fine resolution image, various objects with different meanings can be identified although it is the same object thus requiring appropriate functionalities for image analysis such as those found in Geographical Information Systems (GIS) [7].Looking at the spectral characteristics of satellite sensors, fine resolution sensors are limited in spectral range. Because of limitation in spectral coverage on a continuous electromagnetic spectrum, a single object may contain different patterns that can be classified as a different object. Therefore, appropriate sampling strategy to identify patterns and testing their correlation is required before land cover class labelling is performed. In addition, some of channels within a particular spectral range are highly affected by the atmospheric influence as the electromagnetic energy travels from the sun to the target on the earth's surface and reflected to the sensor [6, 41]. Thus, it is of value to examine the spectral dimension before image analysis in detail since the choice of spectral bands for a particular sensor significantly determines the information that can be extracted from the data for a particular application especially when multispectral data is used [38].

2.5.2. Uncertainty during image pre-processing

The first step in image analysis is to relate image coordinates to ground coordinate system. Various ground coordinate systems ranging from national to global level exist. The importance of relating these coordinate systems is to be able to compare patterns discernable on an image with existing GIS datasets so as to minimize uncertainty during classification. For better results, accurate GIS datasets are required such as topographical maps that were obtained using very accurate methods such as land surveying and photogrammetry. The problem arises when multiple sensors are integrated to optimize information for better modelling. Up-scaling and down-scaling of images from different sensors is a usual approach for fusing them with subsequent co-registration. However, this approach can result in loss of information as each sensor has its own resolving power and the scaling process is subjective based on the intended application [38]. For is instance, one may think of fussing Landsat7 Enhanced Thematic Mapper Plus(Landsat7 ETM+) data with Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (Terra ASTER) so as to optimize derived information. The decision depends on operator's interest. This can also affect information especially when mapping geographic phenomenona that are indeterminate in nature such as lake extent. In this case the idea could be to handle each data set separately and quantify inherent uncertainty and test the correlation of the results through statistical methods before fusion.

2.5.3. Uncertainty during modelling and monitoring

While the concept of data fusion is not new, the emergence of sensors with finer spatial and spectral resolution providing data routinely on regular basis in time, advanced image analysis techniques and improved processing hardware make real time fusion of data possible [26]. However, depending on the speed of the process to be modelled, a single sensor can be limited in time and thus identification of changes that happen when observation is impossible. To bridge the gap, different sensors that operate at irregular basis and in the same orbit can be combined for optimal information extraction. However, special attention in combining these datasets is required to ensure the quality of information extracted.

2.6. Object-oriented image analysis

Since 1970's image processing with subsequent geo-information extraction has been performed on pixel basis. Despite development of powerful classification algorithms that incorporate models of uncertainty such as the famous and commonly used fuzzy c-means clustering algorithm (FCM) [5], the possibilistic fuzzy c-means clustering algorithm (PFCM) [37], fuzzy supervised classification algorithm [51] and neural network classifiers borrowed from artificial intelligence systems, they do not consider the spatial context of real world objects. One of limitations in this development is lacking of image processing software that include functionalities found in GIS software. Recently commercial software are developed in which these functionalities are incorporated enabling access of spatial contextual information of an individual pattern on an image. Definiens Developer software is one the current commercial software in the market that provides GIS spatial analysis functions thus enabling image understanding before classification.

2.6.1. Image segmentation

Object-oriented image analysis comprises two important steps: Iterative image segmentation followed by classification. Image segmentation is defined as dividing an image into homogeneous regions that correspond to (part of) real world objects [13, 34]. Successful image segmentation implies that all pixels in a resulting image segment have similar grey values and form a connected region. Several segmentation algorithms exist [25]. In this study, a multiresolution segmentation algorithm has been adopted. The algorithm applies an optimization procedure which locally minimizes average heterogeneity of image objects for a given resolution [7, 14].

2.6.2. Multiresolution image segmentation

Definiens Developer software offers artificial object-oriented programming language upon which advanced image analysis algorithms can be developed. Basically, algorithms are developed based on object-oriented analysis principles and local adaptive processing procedures. Multiresolution segmentation is a bottom-up merging algorithm. It begins by considering a single pixel as a separate object and subsequently merging adjacent objects that fulfil user defined criterion [8, 13]. In this procedure, the merging decision is based on local homogeneity criterion that describes the similarity between adjacent image objects. Adjacent image objects having smallest increase in the defined criterion are merged. Nevertheless, this process stops when the smallest increase of homogeneity exceeds the defined scale parameter. Thus, the smaller the scale parameter, the smaller the resulting objects and vice versa although this depends on the nature of image data in the process.

During the multiresolution segmentation process, the homogeneity criterion is a combination of colour (spectral values) and shape properties. The shape criterion is further split up in smoothness and compactness parameters. An application of different scale parameters (levels) and colour/shape combinations results into a hierarchical network of image objects [8]. Figure 2.3 shows homogeneity combination criterion during image multiresolution segmentation routine.

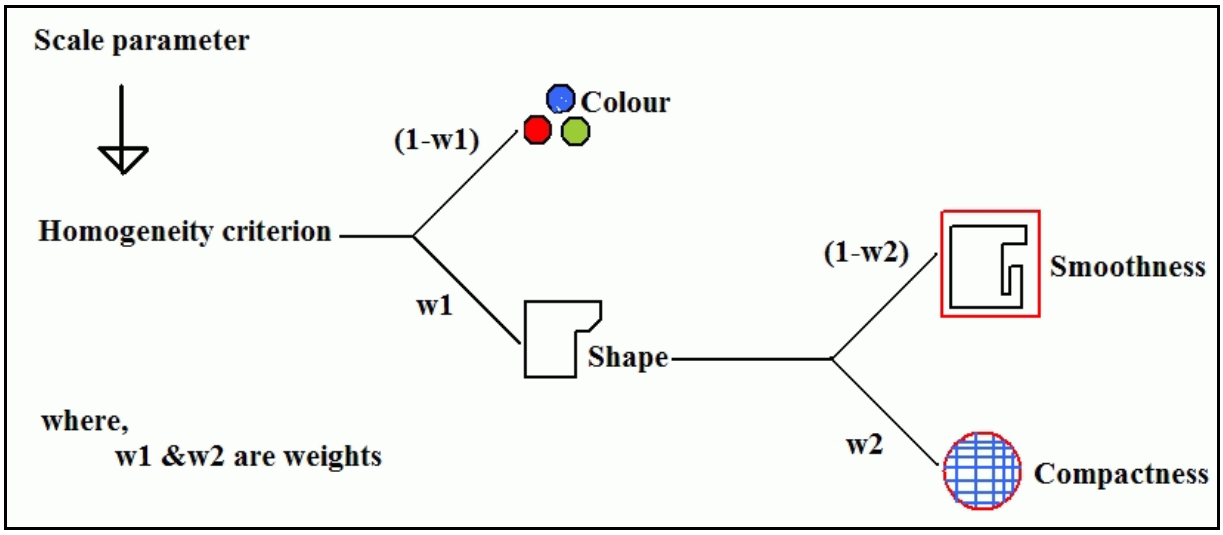


Figure 2.3 Composition of homogeneity criterion

The maximum admissible within-object heterogeneity in terms of spectral and spatial properties is given as follows [8]:

$$h_{scale} \geq w_{spectral} \cdot h_{spectral} + (1 - w_{spectral}) \cdot h_{spatial} \quad (1)$$

where;

$h_{spectral}$ measures the spectral variability of the object;

$h_{spatial}$ characterizes the object shape;

$w_{spectral}$ inversely weights $h_{spectral}$ and $h_{spatial}$.

The merging procedure continues until the within-object heterogeneity exceeds the user defined threshold h_{scale} [2].

Each composition of homogeneity criterion results into a network of objects. Figure 2.4 gives an abstract structure of the resulting hierarchical network of image objects. Each image object 'knows' its neighbour in the same level as well as its sub- and super-objects in the lower and upper level respectively within the network. This enables access to contextual information for each object in the subsequent analysis to identify meaningful objects that correspond to real world objects.

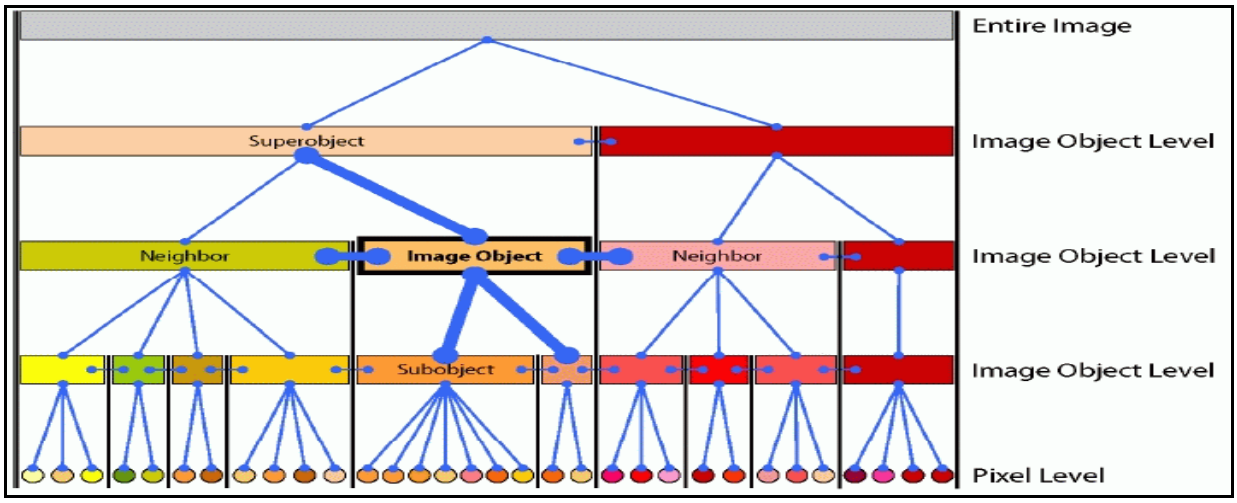


Figure 2.4 Four-level hierarchical network of image objects in abstract illustration (Source: [14])

2.6.3. Image objects information

Multiresolution image segmentation results into a network of three-dimensional topologically related image objects. Each image object has attributes such as layer mean intensity values, shape (area and position) and texture. Figure 2.5 shows a structure of image objects attributes after the segmentation procedure. Not only in-built functions are used to determine these features, but also an operator can program known mathematical functions such as layer indices. This is possible through the use of customizable algorithms within the customized category in Figure 2.5.

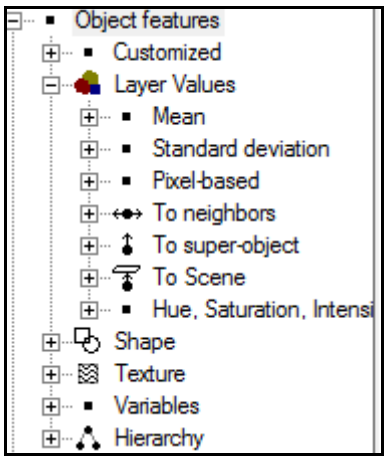


Figure 2.5 Image objects information structure

Image object features are used to develop a set of rules that is used to assign classes to meaningful image objects. Meaningful image objects are those objects of interest in a given application. In this study, meaningful image objects correspond to all objects that constitute a lake (water).

2.6.4. Fuzzy logic classification

Since 1970's pixel-based image classification algorithms have been developed and applied to remotely sensed data to optimize their value in various environmental modelling applications. However, these algorithms do not take into account spatial relationships among objects represented within an image. Recently, algorithms that incorporate spatial structure of each pixel in an image are developed [45-46].However, these algorithms do not consider object semantics and explicitly quantify

uncertainties that are inherent within real world objects. Object-oriented analysis incorporates real world object semantics by adopting the visual variables such as shape, size, colour and texture. As explained in section 4.2, object semantics are captured during segmentation process before classification.

Based on the objectives of this research, fuzzy rule-based classification was applied to derive meaningful image objects from the multiresolution segmentation process. The knowledge base of the lake extent was incorporated in this classification. As the lake extent is characterized by the expanse of water that interacts with shoreline fringing vegetation and submerged as well as floating patches of vegetation, two broad classes were defined, namely; Water and Not water.

In many representations of reality, classical set theory is used. In this approach a pixel has a full membership or zero membership to a land cover class. Clearly the system of representation is precise while the real system is uncertain. Therefore, the proper representation of a fuzzy system is to use fuzzy logic approach with statistical data analysis procedures while developing rules that reflect the object's context. In this scenario, a single attribute cannot be used to precisely discriminate objects from their surroundings. It is therefore important to make use of available real world object features discernable on an image such as shape, colour, orientation, texture and size. Many classification algorithms that incorporate these models of uncertainty apply on pixel basis within multi-dimensional feature space whereby a pixel belongs to a class to a certain degree expressed as percentage. This is a general representation of uncertain phenomena using fuzzy set theory.

Fuzzy sets as sets or classes that for various reasons cannot, or do not have sharply defined boundaries and can be described as follows [10, 51] :
If Z denotes a space of objects, then the fuzzy set A in Z is a set of ordered pairs

$$\mathrm{A} = \{z, \mathrm{M}F_{\mathrm{A}}(z)\} z \in \mathrm{Z} \qquad (2.1)$$

Where the membership function $\mathrm{M}F_{\mathrm{A}}(z)$ represents the grade of membership of z in A and $z \in \mathrm{Z}$ means that z is contained in Z. Usually $\mathrm{M}F_{\mathrm{A}}(z)$ is the number in the range $[0,1]$, with 0 representing non membership and 1 representing full membership of the set. The membership function values help in minimizing uncertainty especially where water interacts with land. Using water properties and vegetation, the range of thresholds could be defined prior to classification. Selection of thresholds depends on the objectives of intended study.

3. Study area, data and data pre-processing

3.1. Introduction

This chapter explores the study area, data used and their pre-processing procedure. Section 3.2 describes the study area; section 3.3 describes the remotely sensed data used and their quality while section 3.4 provides the processing procedure applied to selected remotely sensed data.

3.2. Study area description

Lake Naivasha in Kenya has been used in this study. The lake is located at approximately $00^0$45'00''S, $36^0$21'00'' E. It is situated in the West of Naivasha town in Kakuru district within Rift Valley Province. It is a shallow (mean depth of 6 m [URL 1]), endorheic, freshwater lake in warm and semi-arid conditions in the eastern Rift Valley of Kenya, lying within an enclosed basin at an altitude of 1886 m above mean sea level with surface area fluctuating between 100 and 150 km^2 [1]. It is world famous for its high biodiversity, especially for birds (more than 350 bird species) [31]. In the year 1995, lake Naivasha was declared as Ramsar site (Wetland of International Importance) because of its diverse acquatic and terrestrial ecosystems [URL 2].

The climate of this wetland area is hot and dry with a high potential evaporation exceeding the rainfall by around three times [23]. The area receives rainfall between April to June and October to December each year. The rest of the year is dry season.

The lake system has fringing swamps dominated by papyrus and submerged vegetation and an attendant riverine floodplain with a delta into the lake. These swamps vary in size especially during the rainy season resulting into uncertainty in lake/water boundary. However, there are acquatic plants (water hyacinth) that live and reproduce freely on the surface of fresh water or can be anchored in mud. There are also other submerged vegetation species. The presence of these types of vegetation makes the delineation of water/land boundary difficult especially when they are submerged due to water level rise.

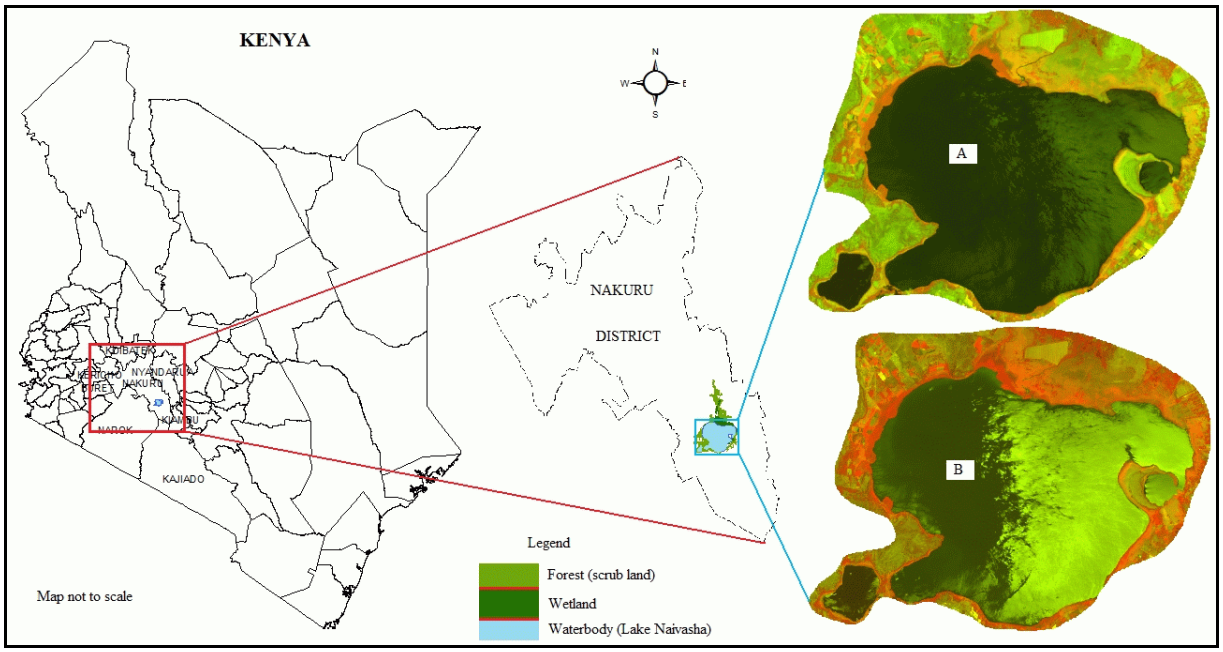


Figure 3.1 Location of study area and corresponding Landsat7 ETM+ (A) and Terra ASTER (B) images acquired on the same date 15th October 2002

Figure 3.1 shows the location of lake Naivasha, the study area in this research, and its corresponding images from ETM+ and ASTER sensors. The lake boundary is fringed by vegetation shown in red. In the north, there are two rivers, namely Malewa and Gilgil which discharge water to the lake. These two major rivers play a great role in balancing the lake water level. However, the two rivers form a delta which influences growth of various vegetation species that may impact the spatial extent of the lake with time.

3.3. Study area selection criteria

In this study, time series images from two sensors ASTER and ETM+ were used. The selection was based on free availability of these images and prior knowledge of the study area via literature. In addition, these sensors are on-board satellites in the same orbit. The study area is within equatorial zone, where cloud cover is a big problem for remote sensing. Images were selected with low cloud cover. In most of these images, the lake was cloud free.

3.4. Characteristics of data used

ASTER and ETM+ are imaging instruments that are flying on NASA's Terra and Landsat7 satellites respectively launched in December 1999. Although the two sensors are on different platforms and have different spatial and spectral characteristics, they are in the same orbit and have same revisit period of 16 days (temporal resolution). However, the two sensors have different spatial coverage indicated by their swath widths.

3.4.1. ASTER sensor

ASTER acquires 14 spectral bands and can be used to obtain detailed maps of land surface temperature, emissivity, reflectance and elevation. The sensor has three subsystems, namely visible near-infrared (VNIR), shortwave infrared (SWIR) and thermal infrared (TIR) subsystem with specific spatial resolution [39]. In this research images from VNIR subsystem were used to estimate the lake

extent and quantify its uncertainty. Table 3.1 gives the summary of spectral, spatial and radiometric characteristics of the sensor in VNIR subsystem while Table 3.2 shows the sensor/platform characteristics.

Table 3.1 3 spectral bands of ASTER (VNIR subsystem)

Spectral channel	Spectral range (μm)	Spatial resolution (m)	Dynamic range
1 (Green)	0.520 to 0.600	15	8 bit
2 (Red)	0.630 to 0.690		
3 (NIR)	0.760 to 0.860		

Table 3.2 Sensor/platform characteristics: ASTER

Launch Date	18 December 1999 at Vandenberg Air Force Base, California, USA
Equator Crossing	10:30 AM (north to south)
Orbit	705 km altitude, sun synchronous
Orbit Inclination	98.3 degrees from the equator
Orbit Period	98.88 minutes
Grounding Track Repeat Cycle	16 days
Resolution	15 to 90 meters
Swath	60 km

Source: [URL 3]

3.4.2. ETM+ sensor

While ASTER acquires information from 14 spectral bands, ETM+ corrects information in eight bands of the electromagnetic spectrum. Seven of these bands are reflective while band 6 is emissive. The sensor has large spectral coverage within the reflective and infrared portion of the spectrum though it is limited in spatial resolution compared to ASTER. Table 3.3 gives a summary of spectral, spatial and radiometric characteristics of ETM+ sensor while Table 3.4 describes the general characteristics of the sensor/platform system.

Table 3.3 7 spectral bands of ETM+ (Reflective)

Spectral channel	Spectral range (μm)	Spatial resolution (m)	Dynamic range
1(Blue)	0.45-0.52	30	8 bit
2(Green)	0.53-0.61		
3 (Red)	0.63-0.69		
4 (NIR)	0.78-0.90		
5(MidIR1)	1.55-1.75		
7(MidIR2)	2.09-2.35		
8(PAN)	0.52-0.90	15	

Table 3.4 Sensor/platform characteristics: ETM+

Launch Date	18 December 1999 at Vandenberg Air Force Base, California, USA
Equator Crossing	Between 10:00 and 10:15AM (north to south)
Orbit	705 km altitude, sun synchronous
Orbit Inclination	98.3 degrees from the equator
Orbit Period	98.88 minutes
Grounding Track Repeat Cycle	16 days
Resolution	15 to 60 meters
Swath	185km

Source: [URL 4]

Tables 3.1 and 3.3 describe the difference between ASTER and ETM+ sensor. ASTER has spatial resolution of 15m but limited in spectral coverage within the reflective portion of the electromagnetic spectrum. However, ETM+ has large spectral coverage though limited is spatial resolution. Tables 3.2 and 3.4 provide platform similarities and differences. The two platforms are in the same orbit with same inclination angle from the equator. They slightly differ in equatorial crossing time though the assumption is that within this short time interval the atmospheric conditions do not change. Therefore, the two sensors can be combined to derive useful geo-information as they complement each other both spectrally and spatially. However, the two sensors differ in over all spatial coverage by a single scene (swath)

3.5. Data availability and their quality

Considering the seasons of the year in the study area, table 3.5 shows selected images and their corresponding dates of acquisition from two sensors. However, it was difficult to obtain images in consecutive years due to high percentage of cloud cover. In addition, since 2003 images from ETM+ contain line dropouts due to scan line corrector (SLC) failure. Although these images were downloaded, they were not used in modelling change in lake extent and its uncertainty as an extra task to correct for gaps was required. Therefore, Level 1B images from ASTER sensor in combination with images from ETM+ (1999 to 2002) were used in this study.

Table 3.5 Selected images and their corresponding acquisition dates

Platform	Landsat7			Terra		
	Year	Season		Season		Year
		Rainy	Dry	Dry	Rainy	
	1999	23/10/1999	Cc	Cc	Cc	1999
	2000	25/10/2000	27/01/2000	Cc	Cc	2000
		Cc	12/02/2000			
			15/05/2000			
			22/08/ 2000			
	2001	12/10/2001	14/02/ 2001	22/01/2001	Cc	2001
		Cc	25/08/2001	Cc		
	2002	02/02/2002	13/09/2002	01/02/2002	15/10/2002	2002
		15/10/2002	Cc	17/02/2002	29/06/2002	
	2003	Data gaps due to SLC failure since 2003		15/08/2003	Cc	2003
				28/01/2003		
	2004			Cc	Cc	2004
	2005			Cc	01/11/2005	2005
					10/12/2005	
	2006			Cc	01/04/2006	2006
					29/06/2006	
	2007			11/09/2007	23/11/2007	2007
				24/02/ 2007	Cc	
	2008			01/01/2008	27/12/2008	2008
	2009			19/01/2009	Cc	2009
				30/07/2009		

Cc stands for cloud cover.

3.6. Satellite image pre-processing

Landsat ETM+ images were downloaded from the United States Geological Survey (USGS) archive [URL 5] in GeoTIFF format whereby each band was stored independently while ASTER images were provided in colour composite. The first step was to combine ETM+ image bands to generate colour composite images which enabled the co-registration process. A combination of band 5, 4 and 2 gave clear visualization of the lake and its surroundings within ERDAS Imagine 9.3. Since the interest was on lake extent, all images were subset in which only area of interest was retained. The image subsets were geometrically corrected respect to mapping system within the study area. In this procedure two topographic map sheets no. 133/4(Longonot) and 133/2(Naivasha) both at scale 1:50,000 were used. Permanent features like road junctions and bridges were selected on both images and topo sheets for ground control points (GCPs). First order polynomial transformation was adopted and applied. A total number of four GCPs were selected for geometric registration of a single image in UTM projection with Clarke 1880 (Modified) Spheroid, Arc 1960 datum and zone 37 S. In order to minimize errors, the selected GCPs were used during georeferencing procedure. The over all root mean square error of the georeferencing process was ± 0.016 m. The geometrically corrected image subsets were now ready for subsequent processing.

4. Methods

4.1. Introduction

A remotely sensed image is a raster grid with underlying scene comprising of interacting objects, existing at different scales which belong to different classes [14]. Traditionally, image analysis before classification is performed based on spectral information contained in an image. Powerful classification algorithms are developed based on signal processing concepts. However, these algorithms operate on pixel basis neglecting topological relationships of real world objects represented by the image. In advanced image analysis classification algorithms are trained to emulate human cognitive process in which the spatial context of real world objects and their interactions are considered [27, 40, 45, 47]. Object-oriented image analysis is one of the advanced remotely sensed data classification approaches and has been proved to provide accurate and valuable results for decision making [4, 30, 50]. This approach comprises two major parts, namely, image segmentation and classification of image segments. Section 4.2 describes image segmentation algorithm used while section 4.3 provides detailed description of fuzzy rule-based image classification applied in information extraction from the images.

4.2. Methodology flow chart

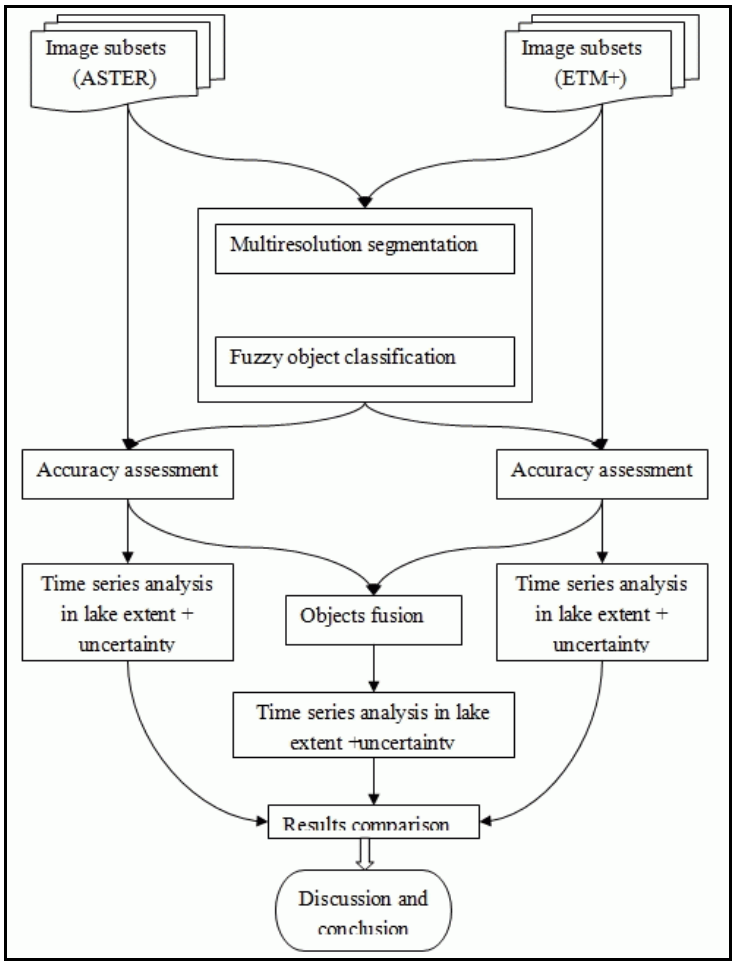


Figure 4.1 Methodology flow diagram of this research

Figure 4.1 summarises the methods used in this research. The first approach was segmentation of georeferenced image subsets from the two sensors followed by object-based classification using fuzzy sets theory. The next step was accuracy assessment of the classified image objects and finally time series analysis in lake extent.

4.3. Multiresolution image segmentation

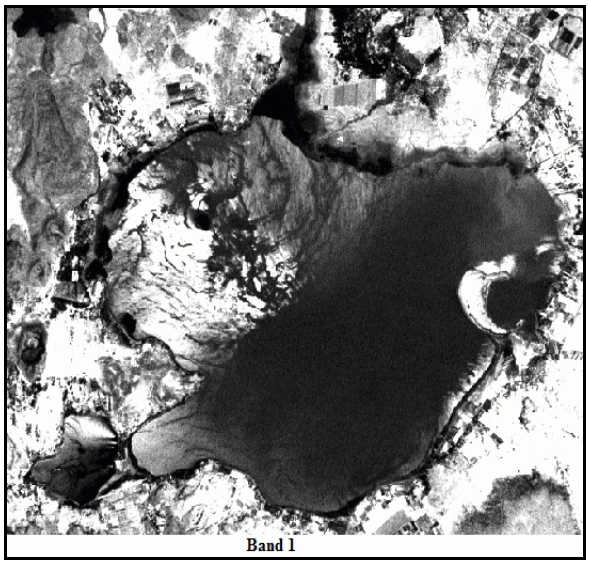
Multiresolution segmentation is a bottom-up merging algorithm. It begins by considering a single pixel as a separate object and subsequently merging adjacent objects that fulfil user defined criterion [8, 13]. The merging decision is based on local homogeneity criterion that describes the similarity between adjacent image objects. Adjacent image objects having smallest increase in the defined criterion are merged. Nevertheless, this process stops when the smallest increase of homogeneity exceeds the defined scale parameter. Thus, the smaller the scale parameter, the smaller the resulting objects and vice versa although this depends on the nature of image data used in the process [14]. In this study, different sets of homogeneity criteria were tested for images from ASTER and ETM+ sensor. The true set was obtained through 'trial and error' approach while observing the resulting image objects and associating them to the reality in the study area. The great emphasis was on the boundary of the lake which depends on the fringing shoreline dominated by deep-rooted, submerged and freely floating vegetation (water hyacinth).

4.3.1. Homogeneity criteria

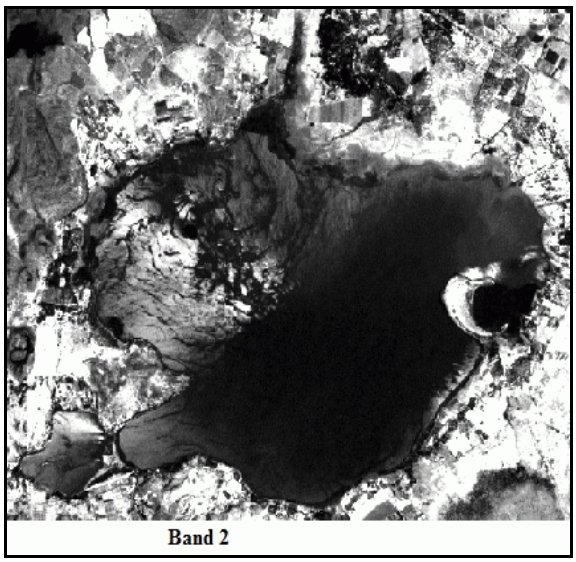
During the multiresolution segmentation process, the homogeneity criterion is a combination of colour (spectral values) and shape properties. The shape criterion is further split up in smoothness and compactness parameters. An application of different scale parameters (levels) and colour/shape combinations results into a hierarchical network of image objects (Figures 2.3 and 2.4).

4.3.1.1. Layer weighting parameters

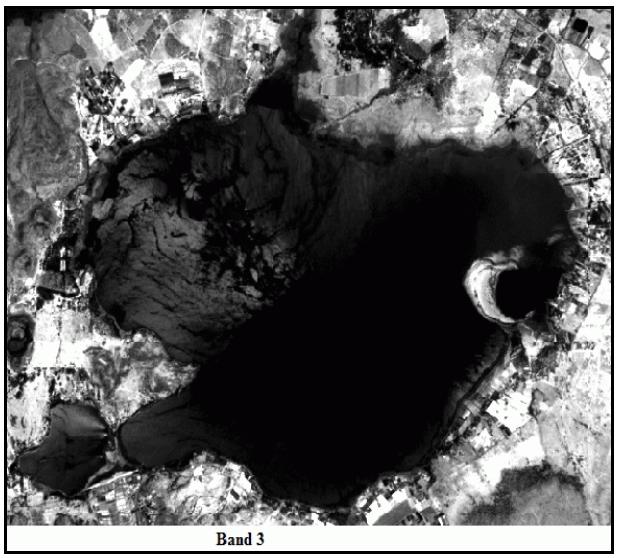
The layer weighting parameters depend on the spatial and spectral characteristics of a geographic phenomenon of interest and available image bands. Since two sensors were used to study the lake extent, I considered first the definition of the lake and its surrounding ecologic units namely, wetlands dominated by deep-rooted, submerged and freely floating vegetation whose size depend on water horizontal extension. Therefore, weight of 1 was assigned to NIR (band 4) while the rest of bands were given weight of 0.5 since water has low reflectance in NIR and vegetation has high reflectance in NIR.



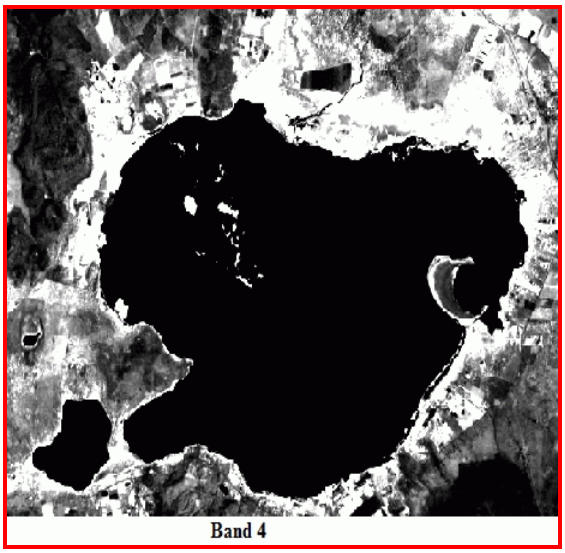
Band 1



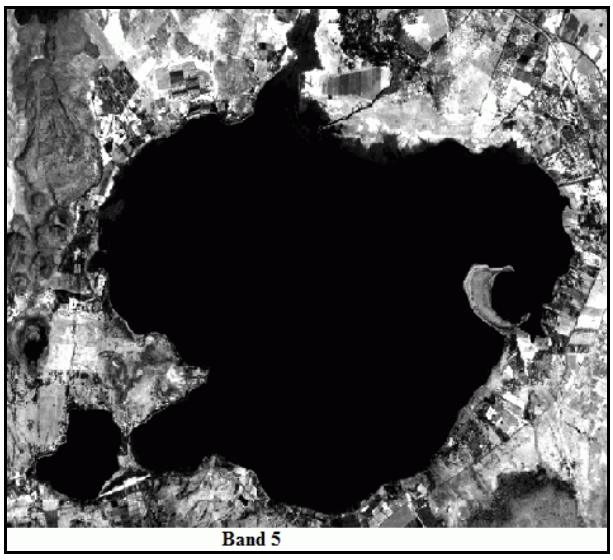
Band 2



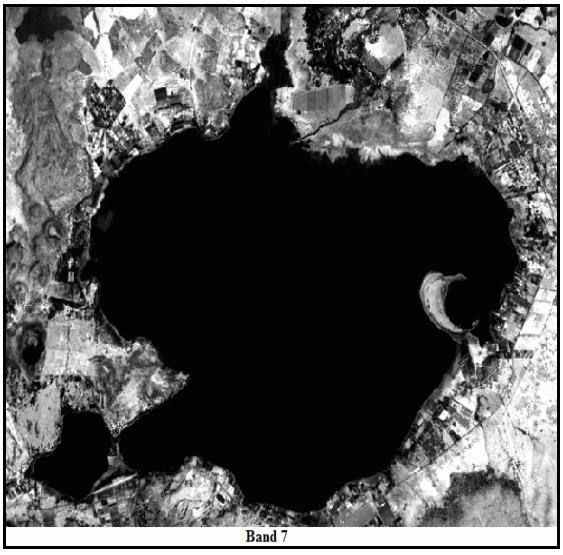
Band 3



Band 4



Band 5



Band 7

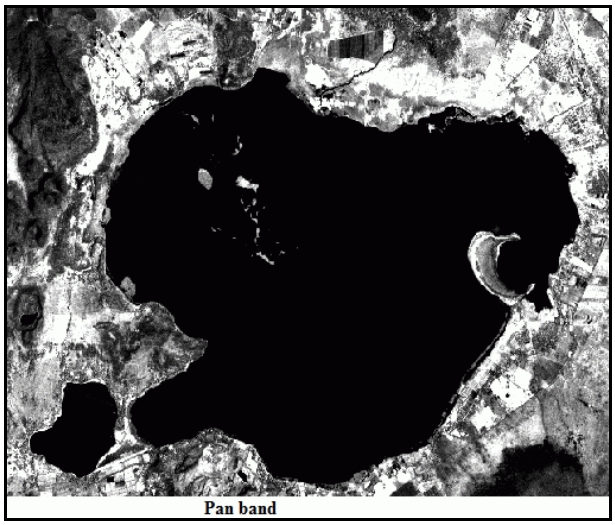


Figure 4.2 Layer weighting procedure before segmentation: Landsat 7 ETM+)

In Figure 4.2, band 4 (NIR) provides clear distinction between water and its surroundings. The shoreline is brighter than water indicating that there is vegetation. Therefore, there was a necessity to give more weight to band 4 since the resulting objects are more influenced by image bands that are input in the segmentation process.

Similar procedure was done for ASTER images. In addition to visual inspection of image bands, I inspected the histograms for all bands from each sensor. The NIR band had clear peaks indicating the possibility of discriminating water from surrounding vegetation. Figure 4.2 shows original and its corresponding equally stretched histogram of a selected image from ASTER sensor. Within the figure, there are clear peaks implying minimal land cover class mixture. Based on these visual aids, I was able to allocate weights according to the objective of this study as meaningful image segments depend on the prior knowledge about the area under study and available data for analysis.

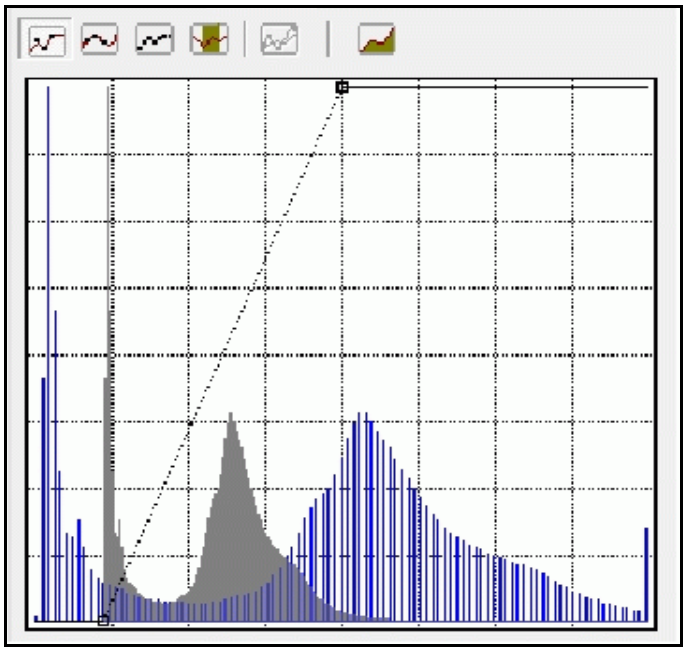


Figure 4.3 Sample histogram for NIR band

4.3.1.2. Scale parameter

The scale parameter is the most important parameter in multiresolution segmentation process. It determines the size of the resulting image objects [13]. Therefore, the larger scale parameter corresponds to fewer image objects constituted by a large number of pixels. However, the scale parameter allows multi-scale classification. This is because each scale parameter results in a network of geographically linked image objects. When applied sequentially, a hierarchy of image objects is constructed in which a 3-D topological relationship among objects is developed. Each object in the network 'knows' its neighbour at the same level, parent in the lower level and child in the upper level (Figure 2.4). This set helps in accessing the spatial context of the resulting image objects during subsequent analysis such as feature extraction for GIS applications or classification for monitoring purpose.

4.3.1.3. Colour and shape parameters

Basically the segmentation process is based on spectral information from the input bands. Multiresolution segmentation algorithm has an interface which allows an operator to choose how much colour or shape has to weight in the segmentation process. The shape parameter controls the shapes of the resulting image objects. The optimization of image objects shapes is possible within this algorithm by introducing another two parameters, namely compactness and smoothing parameters.

4.3.1.4. Compactness and smoothing parameters

The compactness parameter determines how compact the shapes of resulting image objects will be. The larger the value the more compact image objects will be and vice versa. The smoothing parameter operates on the borders of the resulting image shapes. The selection of these parameters should consider the data involved in the analysis. The two parameters complement each other is the sense that allocating weight to one implies the remaining weight is automatically allocated to another parameter (Definiens reference book).

4.4. Hierarchical image objects network

During object oriented image analysis, an image is converted into larger objects that correspond to real world objects by applying a set of parameters [13]. Each set of parameters determines the level of objects with different scale. A combination of these levels in hierarchical order results in a network of objects that are topologically linked thus allowing access of individual image object spatial context. Figure 4.4 illustrates image objects resulting from different segmentation parameter sets.

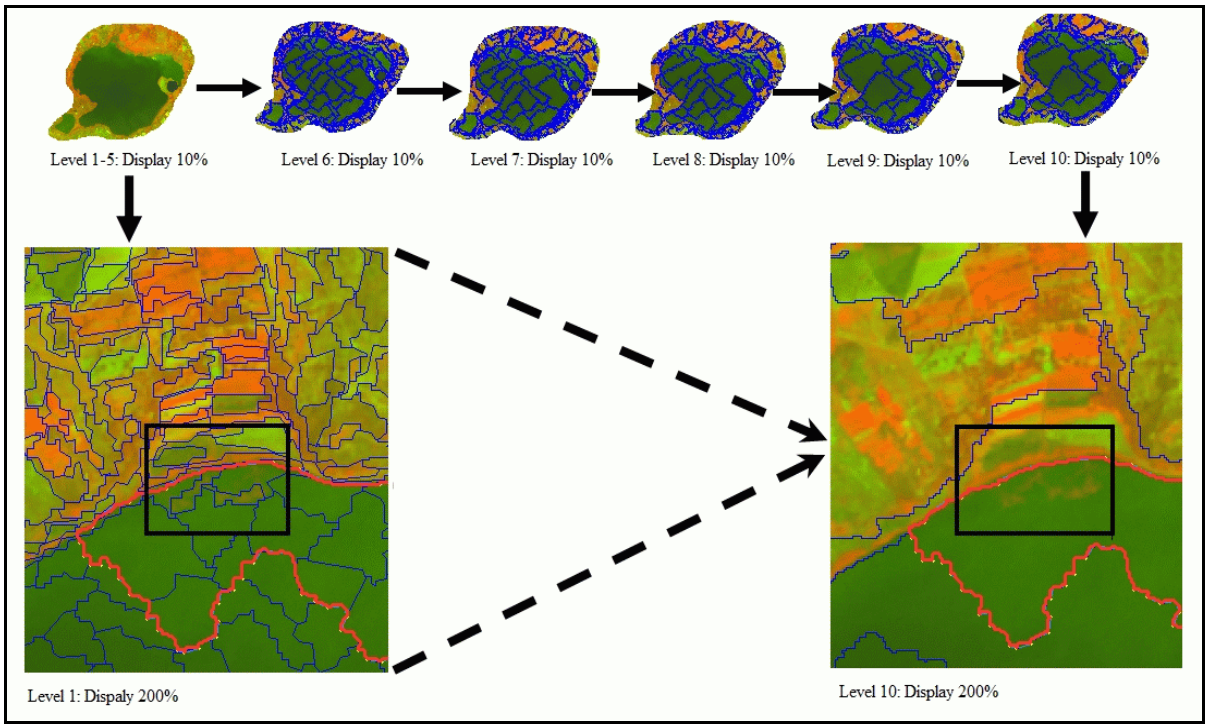


Figure 4.4 Different image objects from different parameter sets

4.4.1. Segmentation parameters used

Based on the study by Darwish *et al.*[13], different colour and shape parameters were tested. Table 4.1 shows a set of parameters that gave meaningful image objects that were subjected to fuzzy rule-based classification. As a rule of thumb, in order to be able to produce adequate classification results, the image objects have to represent objects of the classes that are to be discriminated in the subsequent classification. Objects which are to be assigned to different classes should not be merged. Furthermore, it is not possible to produce an image object level in which all image objects explicitly represent the classes to be extracted. Therefore, at each level of segmentation, classification should be performed to optimize information extracted from an image. This is not possible in many pixel-based image classification algorithms.

Segmentation parameter selection, however, depends on the imagery input in analysis. Highly textured images like Radar imagery would require more weight on shape parameters with little weight to colour criterion. In Definiens user guide, two main principles of segmentation are explicitly stated:

- Always produce image objects in the largest possible scale which in turn still distinguish different image regions (as large as possible and as fine as necessary)
- Use as much colour criterion as possible and as much shape criterion as necessary to come up with image objects of the least smooth and compact shape respectively.

Generally speaking, colour criterion requires more weight as in remote sensing spectral characteristics of real world objects is a source of information. Using too much shape criterion, therefore, can reduce the quality of segmentation results as it was observed in this study.

Table 4.1 Segmentation parameters applied in this study

Segmentation level		Scale parameter	Colour	Shape	
ETM+images				Compactness	Smoothness
	Level 1	30	0.7	0.2	0.8
	Level 2	60	0.7	0.2	0.8
	Level 3	90	0.7	0.2	0.8
	Level 4	150	0.7	0.2	0.8
ASTER images					
	Level 1	15	0.7	0.2	0.8
	Level 2	30	0.7	0.2	0.8
	Level 3	90	0.7	0.2	0.8
	Level 4	150	0.7	0.2	0.8

4.5. Fuzzy rule-based image objects classification

The first consideration in this study was the definition of lake extent. The definition depends on the spatial extent of water at a moment in time. Therefore, the geometric accuracy of the boundary points for the lake is determined by the thematic attributes (water) at that time. In addition, at the boundary of the lake there is vegetation which depend on the availability of water over time. However, these vegetation types are not uniform along the shoreline of the lake boundary. Therefore, I had to specify the range in spatial extent of the lake that will accommodate the uncertainty while estimating the size of the lake. Different image object features were considered in determining parameters to incorporate in the classification algorithm.

4.5.1. Estimation of fuzzy parameters

Since the lake extent is imprecisely defined due to the existing ecologic units at the boundary, I found that a single parameter could not be enough to discriminate water from its background. Therefore, I decided to estimate a set of fuzzy parameters that could minimize uncertainty in lake extent at a moment in time.

4.5.2. Layer mean intensity

For each meaningful image object, the layer mean value is calculated from the layer values of all pixels forming an image object. Mathematically (Definiens reference book),

$$\bar{C}_k(v) = \frac{1}{\#P_v} \sum_{(x,y)\in P_v} C_k(x,y) \qquad (4.1)$$

Where;

P_v: Set of pixels of an object v

P_v: $\{(x,y):(x,y)\in v\}$

$\#P_v$: total number of pixels contained in P_v

$C_k(x,y)$: image layer value at pixel (x,y)

$C_k^{\min}$: darkest possible intensity value of layer k

$C_k^{\max}$: brightest possible intensity value of layer k

$\bar{C}_k$: mean intensity of layer k

The mean intensity of each layer in a given object will take any value between the darkest and brightest possible intensity values of that layer.

4.5.3. Image object brightness

Image object brightness is determined as the mean value of the mean intensity values of an image object mathematically computed as,

$$\bar{C}(v) = \frac{1}{w^B} \sum_{k=1}^{K} w_k^B \bar{C}(v) \qquad (4.2)$$

Where;

$\bar{C}_k(v)$: mean intensity of layer k of an image object v

w_k^B : Brightness weight of layer k

$C_k^{\min}$: Darkest possible intensity value of layer k

$C_k^{\max}$: Brightest possible intensity value of layer k

$w_k^B := \begin{cases} 0 \\ 1 \end{cases}$, this quantity will be 0 if no reflected intensity in all layers

$w^B := \sum_{k=1}^{K} w_k^B$, for k=1, 2,…, K

The brightness of an individual object will take any value between the darkest and brightest mean intensity value of layer a given layer.

4.5.4. Individual layer contribution in each image object

The individual layer contribution to an image object of interest is determined as the ratio of mean intensity of that layer and total brightness of that image object. However, this operation is applicable to multi-spectral data which is believed to contain useful geo-information. The parameter is mathematically determined as,

$$Layer\ contribution = \frac{\bar{C}_k(v)}{\bar{C}(\mathrm{v})} \qquad (4.3)$$

Where;

$\bar{C}_k(v)$: mean intensity of layer k of an image object v

$\bar{C}(v)$: Brightness of image object v

From equation 3 above, it is clearly seen that individual layer contribution to any meaningful image object will take any value between [0, 1]. It will be zero if and only if there was no reflectance in that band during data acquisition while 1 if and only if the mean intensity of that layer equals the image object total brightness

4.5.5. Normalized Difference Vegetation Index (NDVI)

The NDVI is used to study vegetation phenology. However, the studies are based on processed MODIS data with course resolution of 1km [19, 53]. For small areas like Lake Naivasha and its surrounding wetlands, it can be difficult to study using these data as the interest is to estimate the spatial extent of the lake. However, the availability of medium spatial resolution with spectral coverage in the visible and near infrared portion of the electromagnetic spectrum and software with in-built GIS functions allow the use the data to derive physical indices such as NDVI.

Definiens software provides an interactive interface which enables determination of appropriate image object features (attributes). Thus, all in-built high level algorithms can be oriented to suit any application which depends on user needs. The interactive environment enables an operator to integrate the prior knowledge of the process under study. In this research, I decided to determine the NDVI for all meaningful image objects as the entire lake is surrounded by wetlands and dry lands. These landscapes are dominated by vegetation. The wetland is dominated by papyrus swamps while the dry land is dominated by scrubs. The following are assumptions made in this study:

- Water tends to have a low reflectance across all optical bands in the spectrum unless there are suspended sediments near its surface and/or has shallow depth.
- Regardless of the impurities in water and shallow depth, water has very low reflectance in the near infrared band of the spectrum
- Water will always have NDVI values in range of -1 and 0 based on mathematical formula (equation 4.4). However, this will depend on water constituents.

Based on the above assumptions and literature on Lake Naivasha, the near infrared band, green band and NDVI were considered to be of great importance in providing spectral information of the lake extent. Therefore, I determined the NDVI for all image objects using the following mathematical function:

$$NDVI_v = \frac{\left[\bar{C}_{NIR}(v) - \bar{C}_{Red}(v)\right]}{\left[\bar{C}_{NIR}(v) + \bar{C}_{Red}(v)\right]} \qquad (4.4)$$

Where;

$\bar{C}_{NIR}(v)$: Mean intensity value for NIR layer in object v

$\bar{C}_{Red}(v)$: Mean intensity value for red layer in object v

$$NDVI_v = \begin{cases} 0 & for\ \bar{C}_{NIR}(v) = \bar{C}_{Red}(v) \\ 1 & for\ \bar{C}_{Red}(v) = 0 \\ between\ [0,1] & for\ \bar{C}_{NIR}(v) > \bar{C}_{Red}(v) \\ -1 & for\ \bar{C}_{NIR}(v) = 0 \\ between[-1,0] & for\ \bar{C}_{Red}(v) > \bar{C}_{NIR}(v) \end{cases}$$

4.5.6. Implementation of NDVI model in Definiens software

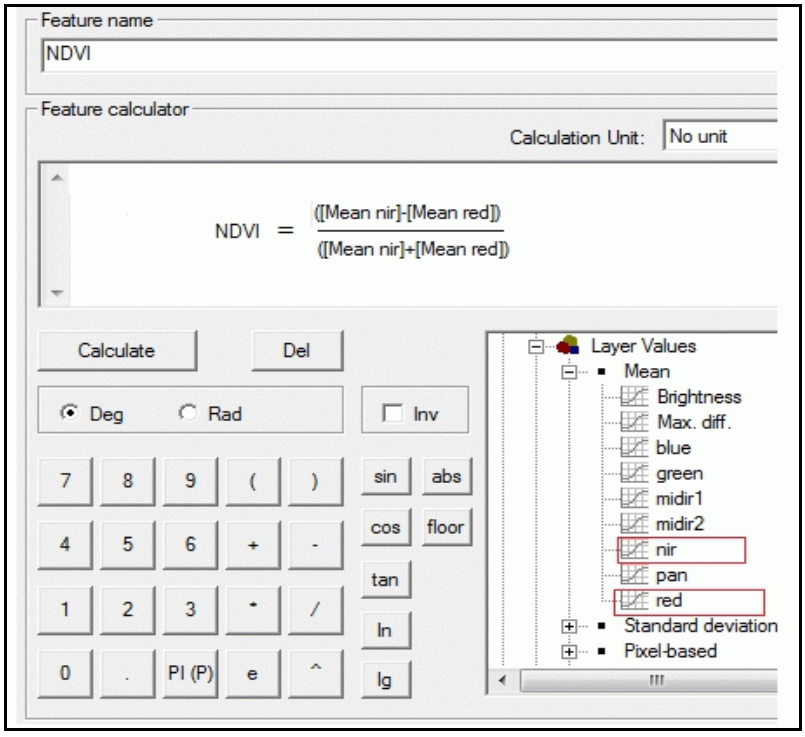


Figure 4.5 NDVI mathematical model implemented in Definiens

Figure 4.5 shows the interface through which various mathematical models can be programmed and optimize parameters for information extraction from images. The mean near infrared and red bands used to determine NDVI index for each image object are highlighted in red colour.

4.5.7. Relative border to a defined class

To identify all objects that share common boundary with water class, I used the feature "Relative border to" in Definiens (Definiens reference book). It refers to the length of the shared border of neighbouring image objects. The feature describes the ratio of the shared border length of an image object with a neighbouring image object assigned to a defined class to the total border length. If the relative border of an image object to image objects of a certain class is 1, the image object is totally embedded in these image objects. If the relative border is 0.5 then the image object is surrounded by half of its border.

$$\frac{\sum_{u \in N_v(d,m)} b(v,u)}{b_v} \tag{4.5}$$

Where;

$b(v,u)$: Topological relation border length

$N_v(d,m)$: Neighbours of an image object v in class m at a distance d

b_v : Image object border length

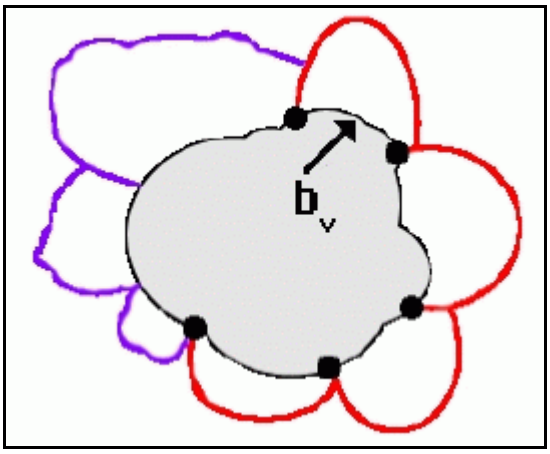


Figure 4.6 Relative border between neighbour objects

Figure 4.6 shows sample image object with relative border to image objects whose classes are known. The object with borders in black can be classified based on class related features such as "Relative border to" using the definition of the function in equation 4.5.

4.5.8. Relative area of a defined class

Using this parameter I was able to identify all objects adjacent to water class and vegetation class. The parameter is determined as area covered by image objects assigned to a defined class in a certain perimeter (in pixels) around the image object concerned divided by the total area of image objects inside this perimeter [14].Mathematically it is determined as follows,

$$\frac{\sum_{u \in N_v(d,m)} \# P_u}{\sum_{u \in N_v(d)} \# P_u} \tag{4.6}$$

Where;

$N_v(d)$: Neighbours to an image object v at distance d

$\# P_u$: Total number of pixels contained in P_u

For image objects where this parameter is 0 implies that the predefined class does not exist whereas 1 implies that objects belong to a predefined class. However, there are objects where this parameter is between 0 and 1 exclusively. In this case spectral information for these objects is required to ensure correct assignment to a desired class.

4.5.9. Fuzzy object classification process

In fuzzy object-based classification, the first step is to identify appropriate fuzzy parameters for classification algorithm. However, in analysing remotely sensed data, the problem is where to start in delineating boundary of a given class. The solution is to look for homogeneous regions and use these regions to train the classifier before actual classification. In addition we need to incorporate prior knowledge about the phenomenon we want to model. Corresponding to this I considered NDVI, NIR and green bands to be appropriate to estimate the range in which fuzzy lake extent can be estimated.

From literature, Lake Naivasha has shallow water which necessitated the use of green band to estimate the range in which its fuzzy boundary lies. In figure 4.4, the green colour represents objects in which green band has higher contribution than the other three bands while the blue colour represents objects in which green has low contribution. In addition, I determined the NDVI for all objects. Since water has low reflectance in NIR which might be lower than red reflectance depending on water constituents at the time of observation, all image objects that belong to water class are expected to have negative NDVI values (equation 4.4). Figure 4.6 shows the set up of getting the range in which the boundary of the lake can be estimated. The black colour represents objects with pure water elements with gradual change in brightness an indication of sediments within water. In the same figure, the green colour represents objects in which NIR band has more contribution than other layers. However, there are objects with low NIR but do not belong to water.

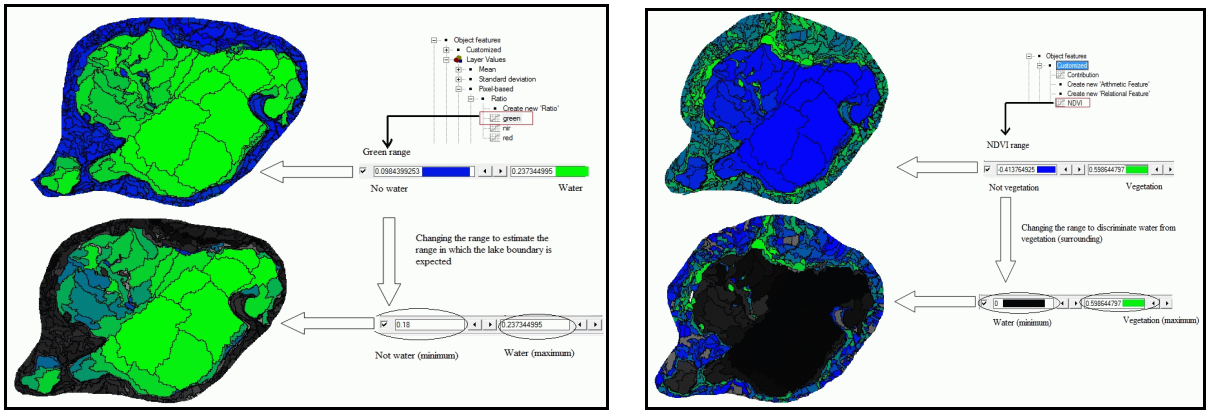


Figure 4.7 Fuzzy parameter estimation using green band and NDVI

Combining these parameters, I was able to identify approximate boundary of the lake and this was the starting point of developing rules used to identify all objects expected to belong to water class.

4.5.9.1. Definition of class hierarchy

From the knowledge acquired during literature review and image objects analysis, I first developed two broad classes namely, ‘water’ and ‘not water’. For not water class I further introduced other two classes namely, ‘vegetation’ and ‘adjacent to water’. The adjacent to water class was intended to identify all objects which are more uncertainty causing difficult in estimating the lake extent while the vegetation class was intended to identify those objects where NIR has more contribution than other bands. Figure 4.7 shows the class hierarchy developed and implemented during lake extent delineation from images.

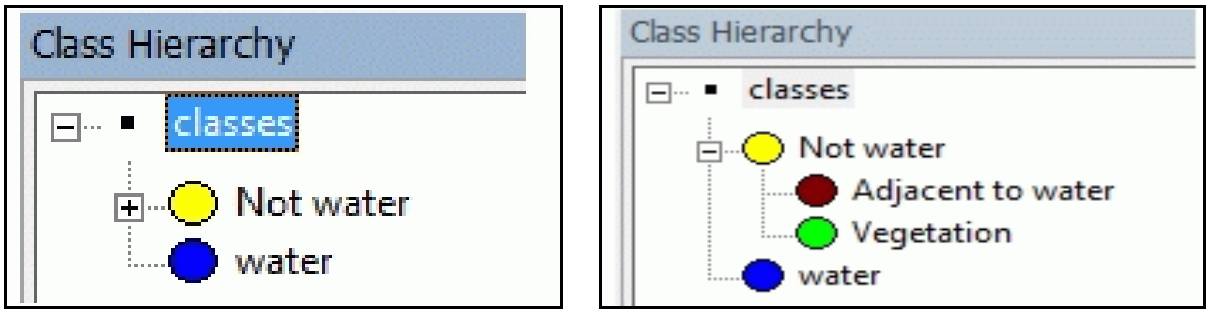


Figure 4.8 Class hierarchy definition

4.5.9.2. Identification of image objects with class water membership

In this procedure water class was identified using the membership function within the class description as shown in figure 4.8

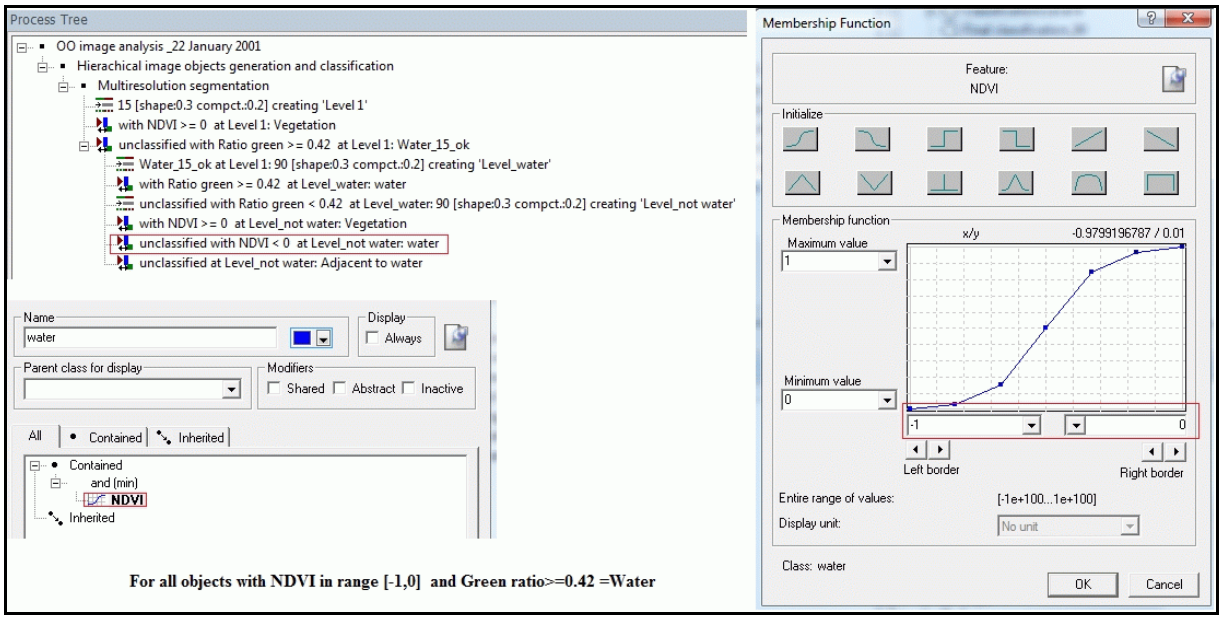


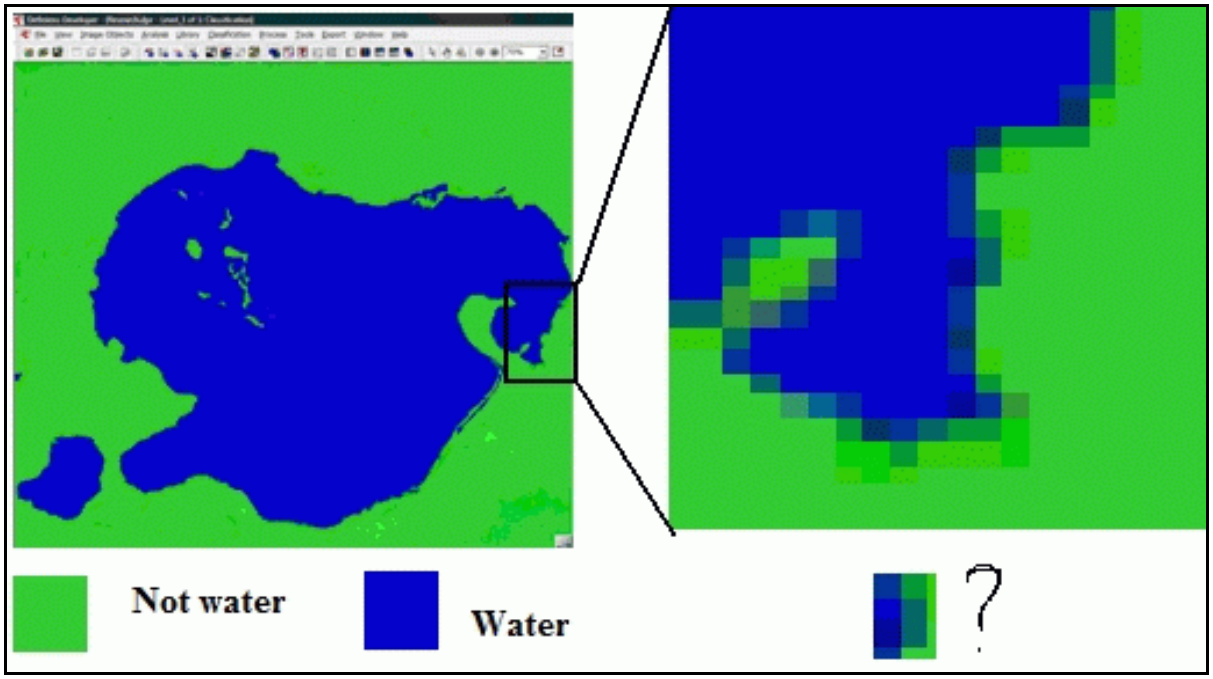
Figure 4.9 Water objects identification

Based on the condition as shown in the process tree in figure 4.8, the classification algorithm evaluates the membership function range for which we expect all objects that belong to water to lie. If the condition is fulfilled, all objects are assigned to class water with truth degrees between 0 and 1. Zero value implies that object does not belong to water and it is not assigned to it while 1 implies full membership. However, there are objects that have partial membership to the class and these are explicitly determined by the membership values within [0, 1] range.

4.5.9.3. Identification of image objects that belong to class 'not water'

Since all objects that belong to water were determined as described in section 4.5.2.2, I classified the remaining objects as 'not water' by negating the conditions used in classifying 'water' objects. However, there is a problem of mixed classes at the boundary of the lake because of the uncertainty of the pixel in location. The enlarged portion of the classification map in figure 4.9 shows the effect of this problem. Therefore, special consideration of all boundary pixels was required. In this case I used

the following function to identify all objects sharing border with class water. The class related features within Definiens software are available.



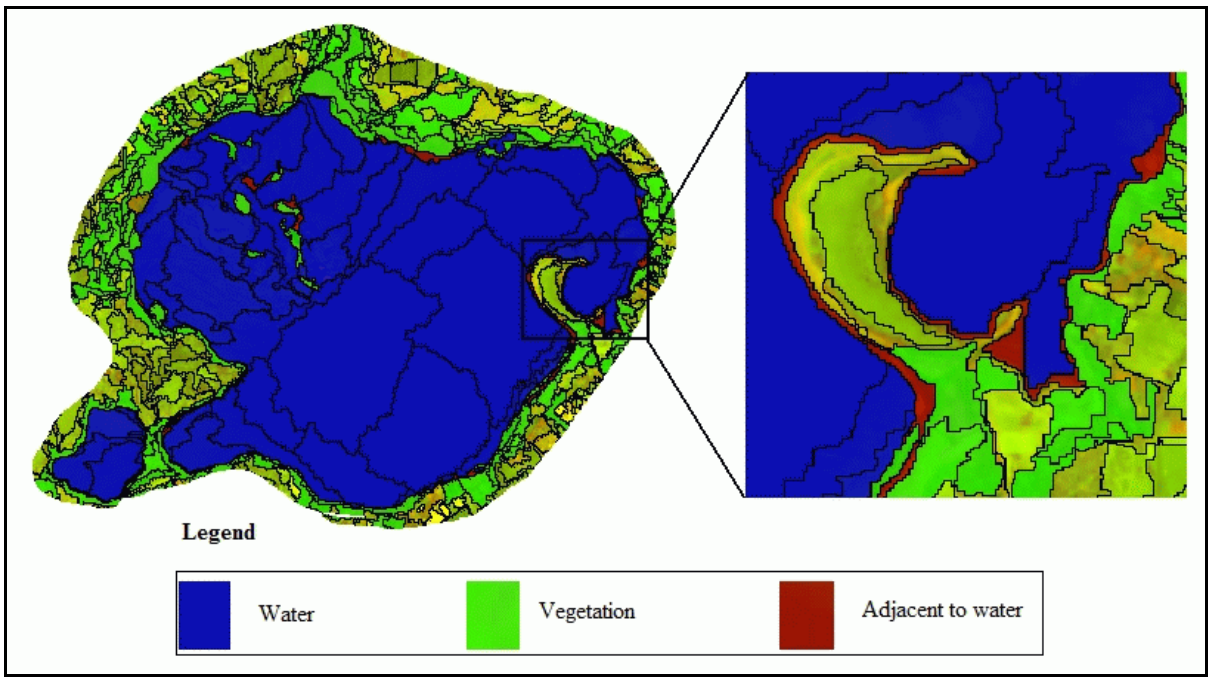


Figure 4.10 Classification map

4.5.9.4. Parameters used to classify images

Table 4.2 and 4.3 show a set of rules developed and applied to extract lake extent information and its uncertainty from ASTER and ETM+ images respectively.

Table 4.2 Rule set for classifying ASTER images

Scale	Class name	Condition	Membership range	Level name
15	Vegetation	NDV>=0	NDV[0, 0.56]	Level 1
	Water	Green>=0.42	NDVI[-1, 0]	
90	Water	Green>=0.42	NDVI[-1, 0]	Level_water
90	Vegetation	NDVI>=0	NDV[0, 0.56]	Level_not water
	Water	NDVI<0	NDVI[-1, 0]	
	Adjacent to water	-	Relative Border to water[0, 1]	

Table 4.3 Rule sets for classifying ETM+ images

Scale	Class name	Condition	Membership range	Level name
60	Vegetation	NDV>=0	NDV[0, 0.7]	Level 2
	Water	NDV<0	Ratio green[0.18, 0.237]	
	Adjacent to water	-	Relative Area to water[0, 1]	

5. Results of object oriented image analysis

5.1. Introduction

In this chapter the results of fuzzy-object-oriented image analysis are presented. The chapter comprises two sections. Section 5.2 presents the estimated lake extent and its uncertainty from ASTER and ETM+ sensor while section 5.3 describes the results of the method adopted to assess the reliability of the classified objects. In addition the concept of weighted average for membership values is introduced.

5.2. Fuzzy objects classified as water

During segmentation process the lake extent was identified as a network of uncertain image objects. Appendix 1 shows the number of image objects classified as water while tables 5.1-5.4 indicate the number of image objects that constitute the total area of the lake. Within the appendix, the number of objects classified as water is larger than the number of objects that constitute the spatial extent of the lake. The reason to this is that there are some objects which satisfied the applied classification rules and assigned to class water. Therefore, in order to identify these objects, I exported the classified objects to ArcGIS software. Within ArcGIS I considered image objects spatial distribution and their corresponding NDVI values. Having determined all objects that constitute the lake extent, I estimated the total area of the lake using the following formula,

$$\bar{A}_{total}(t) = \sum_{i=1}^{N} a_i(t) \qquad (5.1)$$

Where, $a_i(t)$ is an area of individual object *i* at time *t* and $\bar{A}_{total}(t)$ is an estimated total area of the lake at time t. Each identified image object was characterised by a membership function value specifying the degree of truth to constitute the lake extent as well as other spectral information. However, the membership function values differed considerably. Some objects had smaller membership values than others in the range of 0 and 1. In this regard, objects with low membership values are more uncertain than those with higher values. For the membership of the entire lake, weighted differently the image objects based on their size and estimated the weighted average membership values using the following mathematical model,

$$\bar{x}_w = \frac{\sum_{i=1}^{N} w_i x_i}{\sum w} \qquad (5.2)$$

Where; $\bar{x}_w$ is the weighted average for a given lake extent information, w_i is an object individual weight determined by its size (number of pixels), x_i is measured values for each classified object i,

N is a total number of objects that constitute the lake extent and $\sum w$ is total size of lake (number of pixels classified as water (lake)).

Note that the number of pixels w_i for each object was determined by dividing its size to area coverage by an individual pixel. For ASTER images the area coverage by one pixel equals 225 m^2 whereas it is 900 m^2 for ETM+ images.

5.2.1. Estimated lake extent and its uncertainty from ASTER images

Tables 5.1 and 5.2 shows the number of objects that constitute to the spatial extent of the lake as classified from ASTER images during dry and rainy season respectively. Column 5 indicates the uncertainty in lake extent quantified as weighted average membership values.

Table 5.1 Image objects classified as water during dry season for ASTER images

Date	#Objects	Area/m^2	#Pixels	Weighted average membership
22-Jan-01	47	125117775	556079	0.9718
02-Feb-02	43	122559750	544710	0.9355
17-Feb-02	44	124650450	554002	0.9365
15-Aug-03	38	123095700	547092	0.9127
24-Feb-07	42	116614350	518286	0.8930
11-Sep-07	69	117167175	520743	0.9098
01-Jan-08	61	119514825	531177	0.9503
19-Jan-09	36	116465400	517624	0.9301
30-Jul-09	33	116096400	515984	0.9615

Table 5.2 Image objects classified as water during rainy season for ASTER images

Date	#Objects	Area/m^2	No. of pixels	Weighted average membership
29-Jun-02	99	124176150	551894	0.9248
15-Oct-02	97	123729300	549908	0.9248
01-Nov-05	41	121145175	538423	0.9254
01-Apr-06	45	119251800	530008	0.9610
27-Dec-08	29	117136800	520608	0.9608

From tables 5.1 and 5.2, it is clear that the lake extent is uncertain in all seasons of the year. However, the degree of uncertainty varies from year to year and within a year. The variation depends on many factors such as water turbidity, non- uniform depth of lake bed, presence of vegetation at the boundary of the lake and other patches of floating vegetation which due to limitation of a pixel in space are classified as water or vegetation.

5.2.2. Estimated lake extent and its uncertainty from ETM+ images

Tables 5.3 and 5.4 present the number of objects classified as water constituting to spatial extent of the lake during dry and rainy season respectively. Their corresponding weighted average membership values representing the uncertainty in lake extent are presented in column 5 of each table.

Table 5.3 Image objects classified as water during dry season for ETM+ images

Date	#Objects	Area/m2	#Pixels	Weighted average membership
27-Jan-00	29	128388825	142654	0.9929
12-Feb-00	21	126134100	140149	0.9197
15-Mar-00	35	126387450	140431	0.7031
22-Aug-00	32	124987050	138875	0.8528
14-Feb-01	26	123016725	136685	0.9296
25-Aug-01	34	123837750	137598	0.8588
13-Sep-02	38	122213250	135793	0.9467

Table 5.4 Image objects classified as water during rainy season for ETM+ images

Date	#Objects	Area/m2	#Pixels	Weighted average membership
23-Oct-99	46	127221300	141357	0.8953
12-Oct-01	47	124294275	138105	0.9465
15-Oct-02	49	123316650	137019	0.9992
02-Dec-02	28	123318675	137021	0.9996

In tables 5.3 and 5.4, the lake extent is uncertain within all seasons of the year. However, there are seasons where it is more uncertain than others. For instance, on the 15th March 2000 the lake extent was more uncertain than the rest of the months in the same year. Looking at the original image, the lake extent is clearly identified visually. After classification, some of the classified image objects as water had low membership values signifying more uncertain although their spectral information revealed that they belong to water.

5.3. Accuracy assessment of classified image objects

For quality assurance of the classification results, I investigated the class stability of each object in each class. The algorithm is available within Definiens software. It determines the least and best classified image object in each class (Definiens reference book) and returns the number of objects classified as the predefined classes, mean, standard deviation, minimum and maximum membership function value. The results of this assessment are shown in appendix 1. Based on these statistical values I was able to ascertain whether the class is stable or not. For class water, these values were higher than other classes which again was not true that all objects classified as water had equal membership in that class. There were some objects which were more uncertain than others. In addition, some of these objects were spatially away from the lake boundary which required another step to identify only objects that constitute to total area covered by the lake. Therefore, another procedure was to track those objects that constitute to the total size of the lake. In this case, all classified image objects were exported to ERDAS Imagine as shapefiles. I overlaid the classified image objects over original image (left of figure 5.1) for visual inspection of the spatial distribution of these objects.

In identifying objects that constitute to total size of the lake, I exported water and adjacent to water classified objects to Arc Map. Using the NDVI condition that all objects with NDVI<0 are water, I

was able to identify only objects that sum up to the total size of the lake. The right map of figure 5.1 illustrates the results after applying that condition and again exporting the results to ERDAS for overlay on original image.

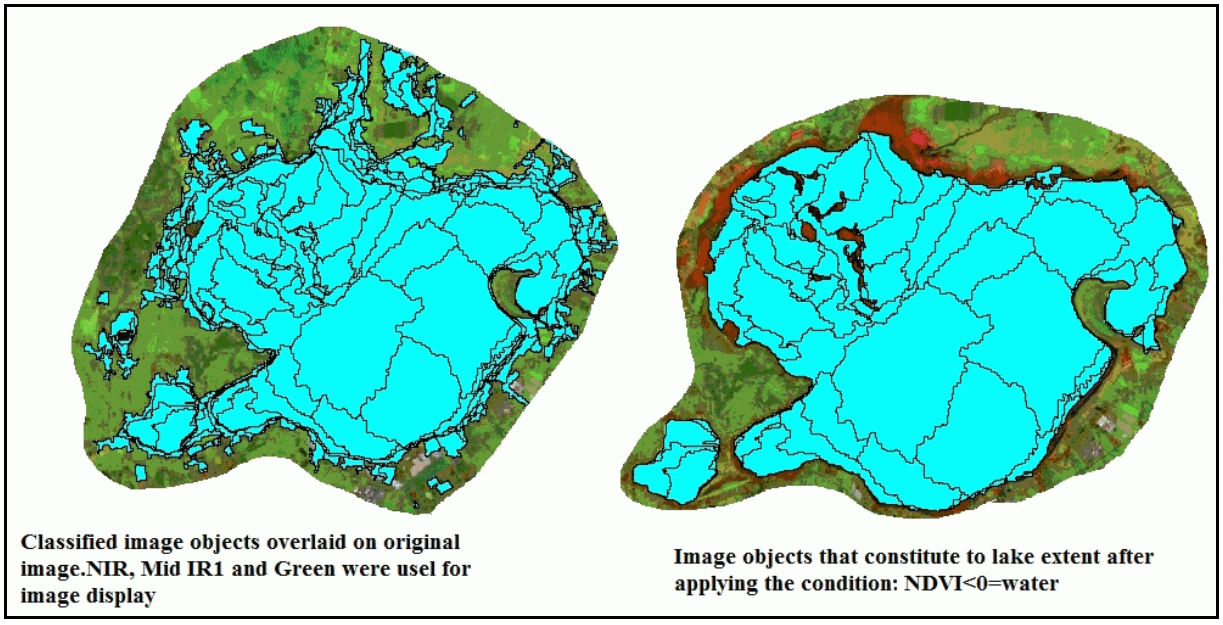


Figure 5.1 Assessment of image objects classified and identification of objects that constitute to lake extent

6. Time series analysis in lake extent

6.1. Introduction

This chapter describes time series analysis approach in lake extent from two sensors. It starts with analyzing time series in lake extent and its uncertainty from ASTER sensor in section 6.2 whereas section 6.3 presents time series analysis from ETM+ sensor. Section 6.4 describes time series analysis in lake extent by combining two sensors.

6.2. Time series analysis in lake extent using ASTER images

The aim of this analysis was to investigate how the estimated lake extent and its uncertainty vary over time as observed by ASTER sensor. In this case the seasons of the year for the study area were identified, namely, rainy and dry season. Tables 5.5 and 5.6 show the over all estimated lake extent information for two seasons respectively. Columns 3 to 5 are weighted average lake extent information computed using equation 5.2 of chapter 5.

Table 6.1 Weighted average lake extent information: Rainy season, ASTER images

Date	Area/m^2	NDVI	Green	Nir	Red	Membership
29-Jun-02	124176150	-0.2293	0.4822	0.1994	0.3184	0.9248
15-Oct-02	123729300	-0.2294	0.4824	0.1993	0.3183	0.9248
01-Nov-05	121145175	-0.2282	0.5081	0.1898	0.3021	0.9254
01-Apr-06	119251800	-0.1753	0.5064	0.2036	0.2900	0.9610
27-Dec-08	117136800	-0.1770	0.5197	0.1976	0.2827	0.9608

Table 6.2 Weighted average lake extent information: Dry season, ASTER images

Date	Area/m^2	NDVI	Green	Nir	Red	Membership
22-Jan-01	125117775	-0.1399	0.4801	0.2237	0.2962	0.9718
02-Feb-02	122559750	-0.2132	0.4644	0.2108	0.3249	0.9355
17-Feb-02	124650450	-0.2111	0.4633	0.2118	0.3249	0.9365
15-Aug-03	123095700	-0.2455	0.5029	0.1876	0.3095	0.9127
24-Feb-07	116614350	-0.2721	0.5045	0.1804	0.3150	0.8930
11-Sep-07	117167175	-0.2480	0.5103	0.1843	0.3055	0.9098
01-Jan-08	119514825	-0.1905	0.5108	0.1981	0.2911	0.9503
19-Jan-09	116465400	-0.2211	0.5145	0.1891	0.2965	0.9301
30-Jul-09	116096400	-0.1732	0.4904	0.2109	0.2987	0.9615

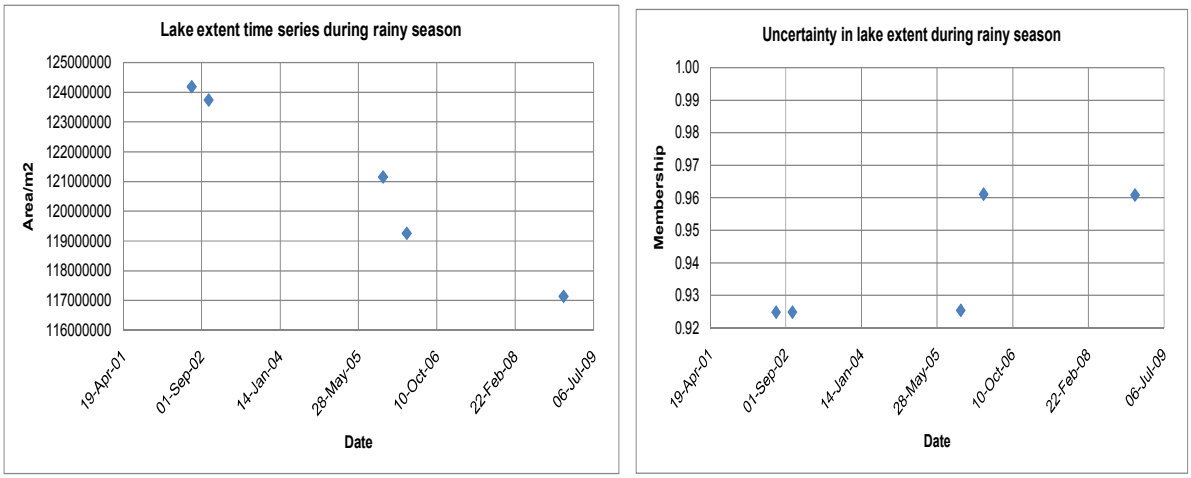


Figure 6.1 Lake extent and its uncertainty during rainy season, ASTER images

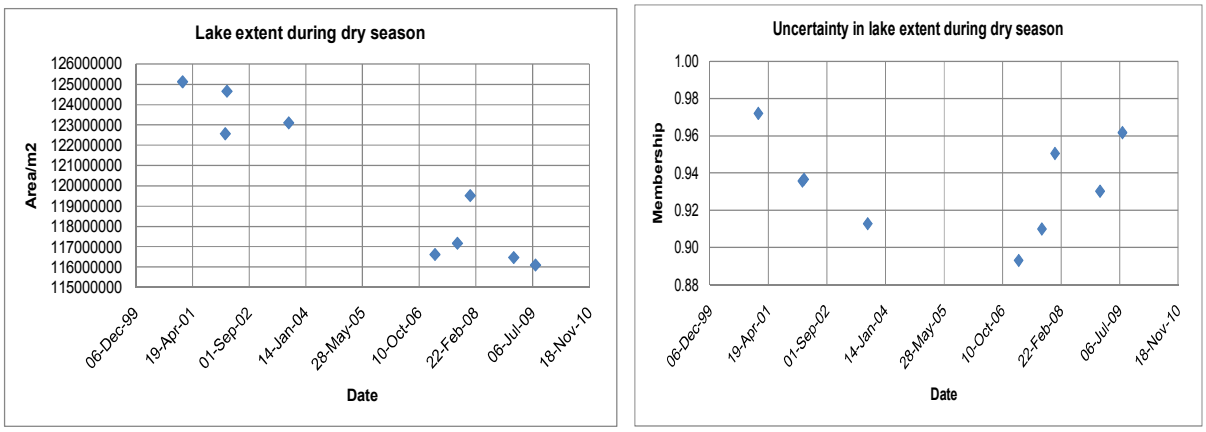


Figure 6.2 Lake extent and its uncertainty during dry season, ASTER images

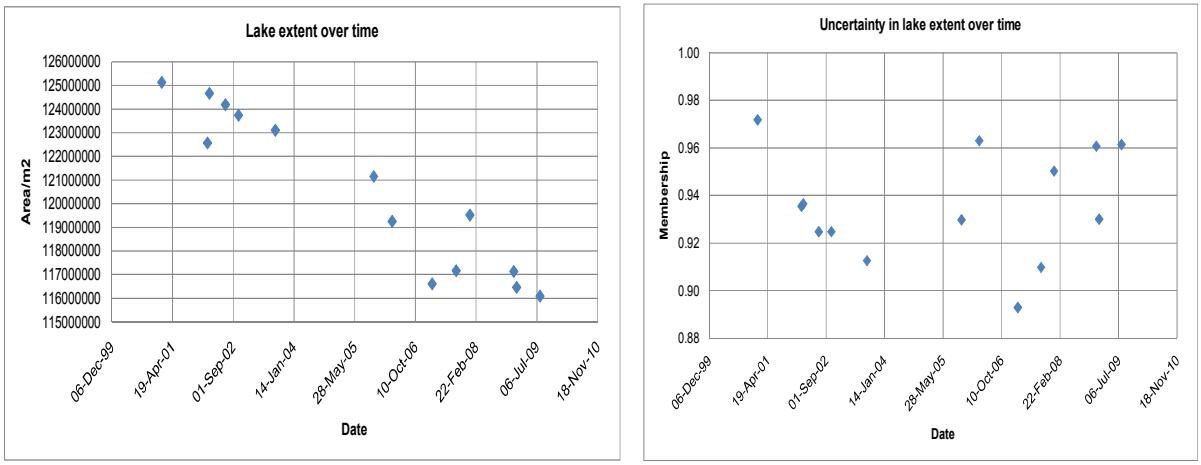


Figure 6.3 Lake extent and its uncertainty over time, all seasons, ASTER images

Figure 6.1 and 6.2 show the variations in spatial extent of the lake during rainy and dry season respectively between 2001 and 2009. The right graphs in the same figures represent the uncertainty in lake extent due to its vague definition and uncertainty in measurements during time of observation. Figure 6.3 shows the variation in lake extent in all seasons of the year. The scatterplots show that there is a falling trend in spatial extent of the lake from 2001 to 2009.

Having identified that there is a falling trend, the next procedure was to estimate the best fit line representing the falling trend from the measurements performed. In this case I performed linear regression modelling in which least squares method was used to estimate best fit parameters.

6.2.1. Linear regression modelling: ASTER sensor

Regression analysis is the statistical technique of mathematically finding relationships between variables and often used to predict the future. In this analysis, it is clearly seen from figure 6.3 that the lake extent varies with time which mathematically can be expressed as,

$$Y_i = b_0 + b_1 \mathrm{X}_i + \varepsilon_i \qquad (6.1)$$

Where; b_0 and b_1 are parameters to be estimated, ε_i are statistical errors assumed to be independent and normally distributed with mean 0 and standard deviation σ, Y_i is regression line used to predict future values from measurements X_i. Assuming $\hat{b}_0$ and $\hat{b}_1$ to be the estimates of b_0 and b_1 respectively and the residual $\hat{Y}_i = \hat{b}_0 + \hat{b}_1 \mathrm{X}_i$, the best fit line is the one which minimizes the sum of the squares of the residuals.

In this study the Y is a trend or prediction line which should be estimated in such a way that the sum of the squares of the residuals is minimal and this can be expressed mathematically as

$\sum (Y - \hat{Y}_i)^2$ = minimal. The idea is to estimate a set of parameters that characterize the fitted linear model and test the significance of the model, that is, whether it perfectly models the variability in the predicted values given sample measurements. The regression analysis was performed in Excel using available regression tool. In this case the measurements were arranged according to seasons and performing analysis while observing the parameters of the fitted model. Table 6.3 illustrates some of the quality assurance parameters of the fitted model with an emphasis on the P-value and the adjusted R-square. These values are only for the estimated coefficient as this is the most important parameter that influences the trend line.

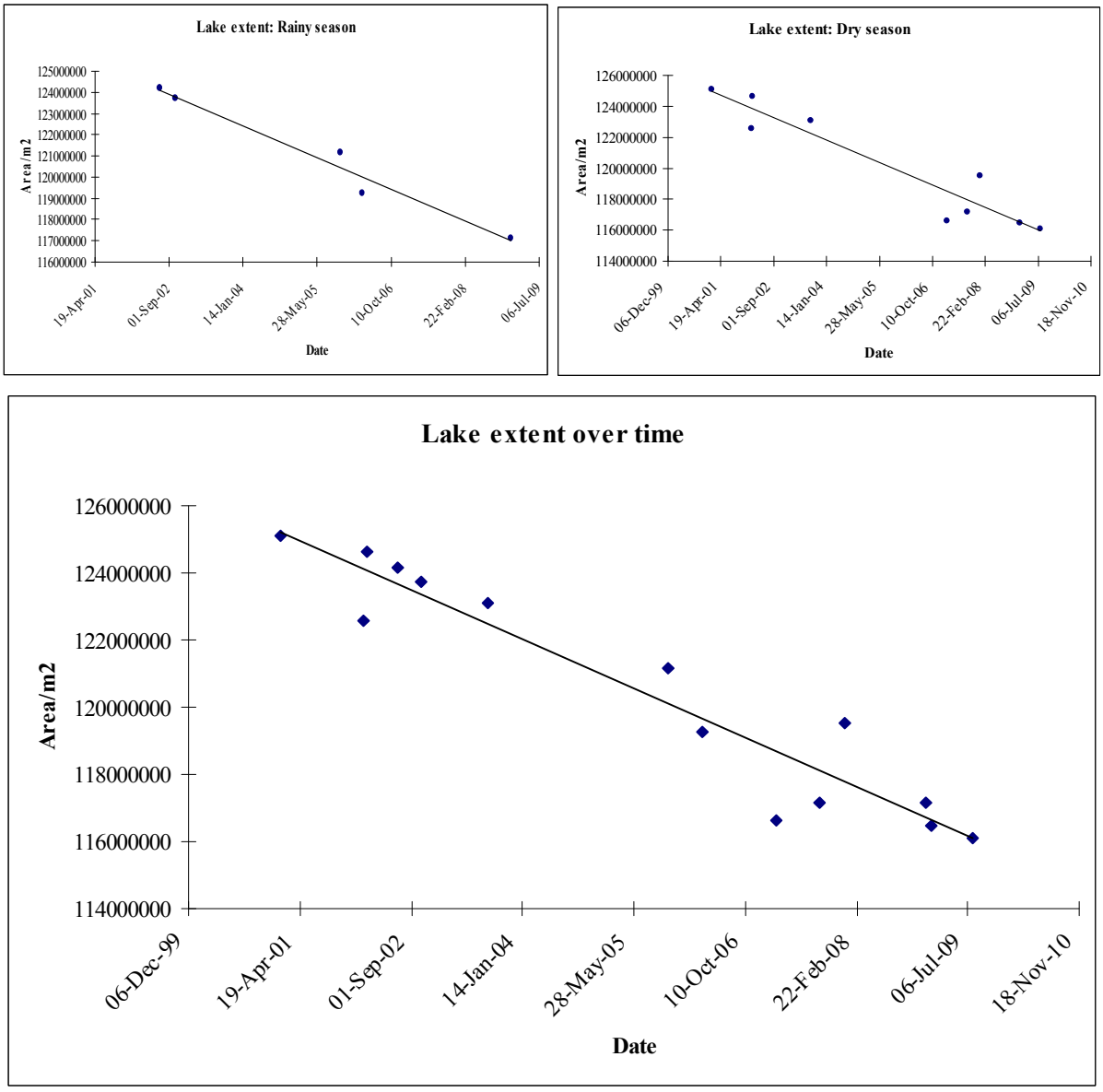


Figure 6.4 Lake extent trend from 2001 to 2009 from ASTER images

Table 6.3 Quality of the fitted trend line in different seasons: ASTER sensor

Date	R^2	Adjusted R^2	t-statistic	F-statistic	p-value
Rainy season	0.97	0.96	-9.85	97.12	$2.22\text{x}10^{-3}$
Dry season	0.90	0.89	-7.99	63.92	$9.15\text{x}10^{-5}$
Over time	0.91	0.90	-11.13	123.97	$1.00\text{x}10^{-7}$

Table 6.3 shows 96%, 89% and 90% of the variability in the data used to estimate the trend in lake extent during rainy season, dry season and over time respectively is well represented by the model. However, the P-values used to test the null hypothesis, in this case, no trend in lake extent decreases with an increase in the number of measurements. Therefore, the trend in lake extent can be studied in all seasons of the year.

6.3. Time series analysis in lake extent using ETM+ images

Table 6.4 and 6.5 present information about the lake extent as extracted from ETM+ images. Table 6.4 shows information during rainy season whereas table 6.5 indicates information during dry season. Looking at the column containing the area of the lake at a point in time, spatial extent changes over time in both seasons. The next step was to find out the mathematical model of this variation. The first step was to plot the scatterplots of the measurements to identify their distribution over time.

Table 6.4 Weighted estimated lake extent information: Rainy season, ETM+ images

Date	NDVI	Green	NIR	Red	Area/m2	#Pixels	Membership
23-Oct-99	-0.3618	0.2274	0.0873	0.1858	127221300	141357	0.8953
12-Oct-01	-0.3595	0.1901	0.0894	0.1897	124294275	138105	0.9465
15-Oct-02	-0.4589	0.2296	0.0721	0.1943	123316650	137019	0.9992
02-Dec-02	-0.4512	0.2452	0.0716	0.1890	123318675	137021	0.9996

Table 6.5 Weighted average lake extent information: Dry season, ETM+ images

Date	NDVI	Green	NIR	Red	Area/m2	#Pixels	Membership
27-Jan-00	-0.1051	0.2990	0.1891	0.2333	128388825	142654	0.9929
12-Feb-00	-0.5028	0.2255	0.0631	0.1909	126134100	140149	0.9197
15-Mar-00	-0.4695	0.2144	0.0690	0.1912	126387450	140431	0.7031
22-Aug-00	-0.4989	0.2222	0.0660	0.1980	124987050	138875	0.8528
14-Feb-01	-0.4746	0.2249	0.0673	0.1893	123016725	136685	0.9296
25-Aug-01	-0.4935	0.2218	0.0693	0.2041	123837750	137598	0.8588
13-Sep-02	-0.4818	0.2322	0.0693	0.1980	122213250	135793	0.9467

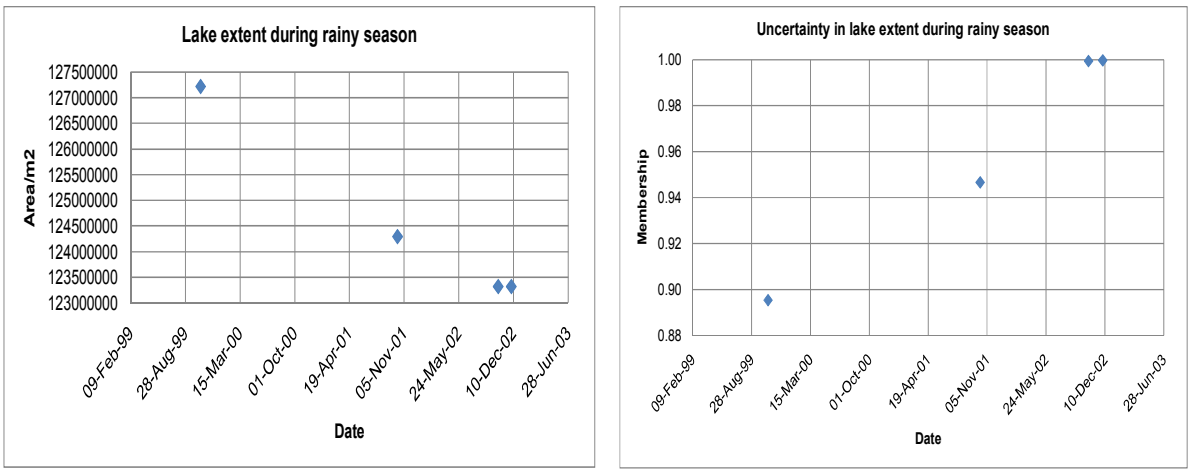


Figure 6.5 Lake extent and its uncertainty: Rainy season, ETM+ images

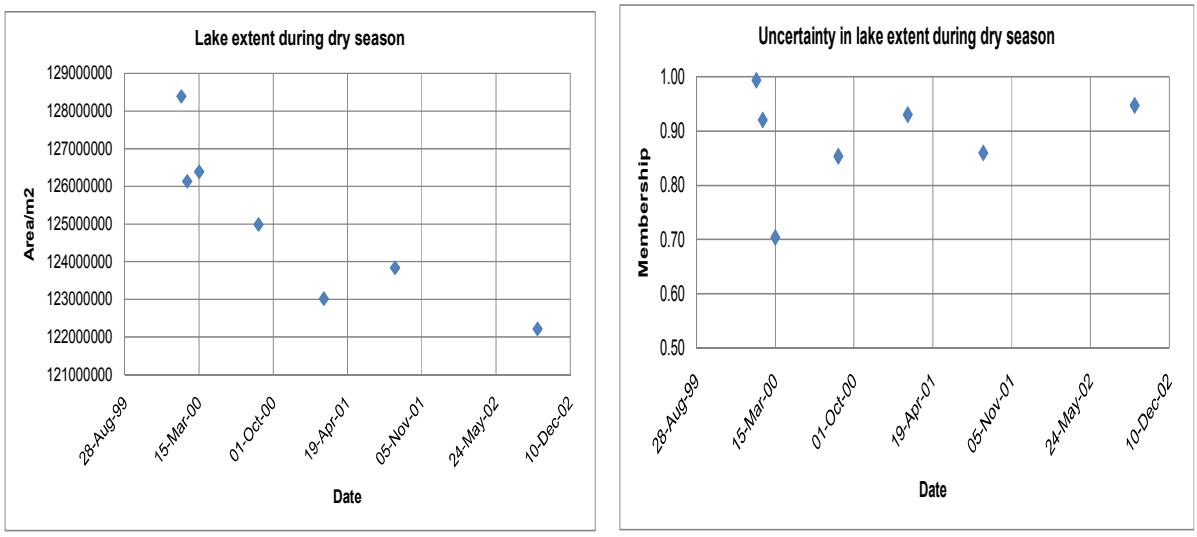


Figure 6.6 Lake extent and its uncertainty: Dry season, ETM+ images

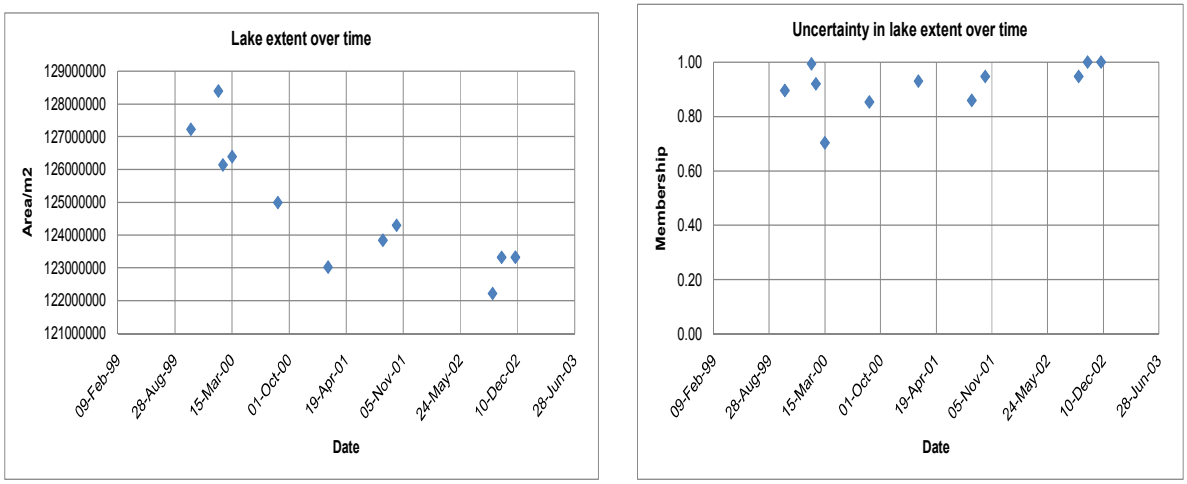
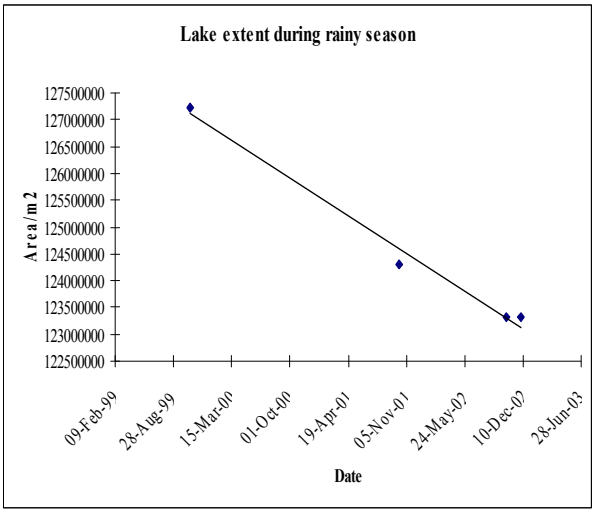


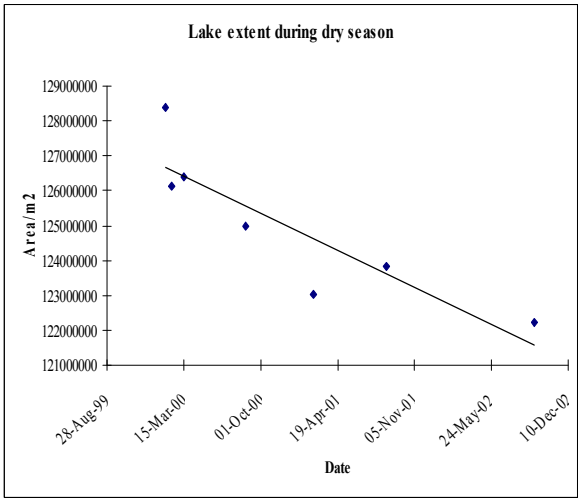
Figure 6.7 Lake extent and its uncertainty over time, all seasons, ETM+ images

The scatterplots in figures 6.4 to 6.6 indicate that there is a falling trend in lake extent between 1999 and 2002 as observed using ETM+ sensor. However, lake extent is more uncertain during dry season compared to rainy season. The next step was to estimate the linear model of the falling trend in lake extent. Therefore, linear regression modelling was performed in which least squares method of parameter estimation was again applied.

6.3.1. Linear regression modelling: ETM+ sensor

In this section, equation 6.1 was adopted. The following figures show the fitted regression line for rainy season, dry season and over time (combination of measurements made between 1999 and 2002).





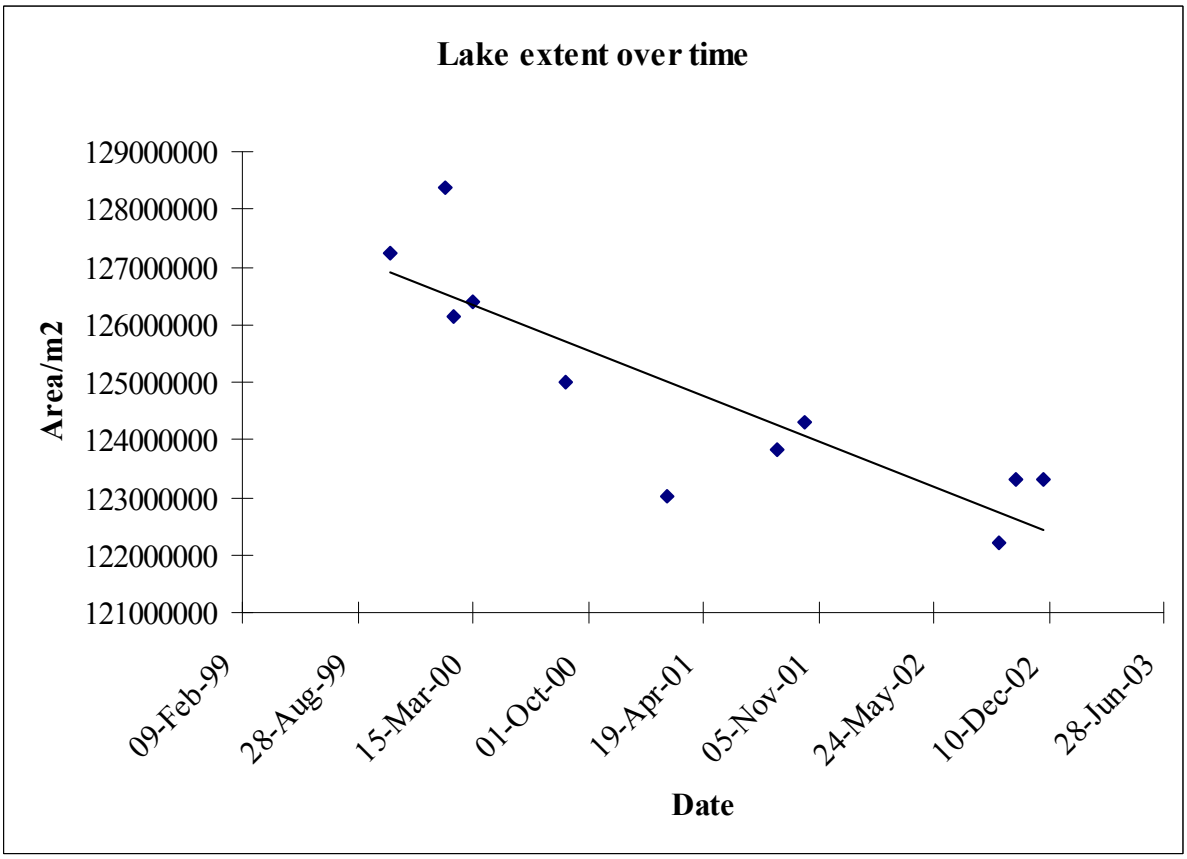


Figure 6.8 Lake extent trend from 1999 to 2002 from ETM+ images, all seasons

Table 6.6 Quality of the fitted trend line in different seasons: ETM+ sensor

Date	R^2	Adjusted R^2	t-statistic	F-statistic	p-value
Rainy season	0.99	0.98	-12.28	150.70	0.01
Dry season	0.76	0.72	-4.02	16.15	0.01
Over time	0.74	0.71	-5.08	25.76	6.67×10^{-4}

From table 6.6 it is clearly seen that 98%, 72% and 71% of the variability in the data used to estimate trend in lake extent during rainy season, dry season and combined seasons respectively is well represented by the fitted linear model. For the case of rainy season, the R-squared value is higher than

that of dry season and over time. Actually this is not the best fit because few measurements were used to estimate this trend as other images were covered by clouds during this season. However, the P-value was larger than that over time indicating again that the trend in uncertain lake extent can be studied by increasing the number of observations. Looking at the dry season, the P-value is larger than for the rainy season and for all seasons together. Again referring to membership column of table 6.5, the lake extent was more uncertain (0.70) compared to other dates. This might have influenced the estimated trend in lake extent though the number of observations increased.

6.4. Time series analysis using two sensors: ASTER and ETM+ sensor

While section 6.2 and 6.3 discussed the analysis of time series in lake extent from individual sensor, this section analyzes the trend identified in both sections by combining the two sensors by considering seasons and over time. In all cases the scatterplots for measurements are plotted to visualize the nature of the data before regression analysis.

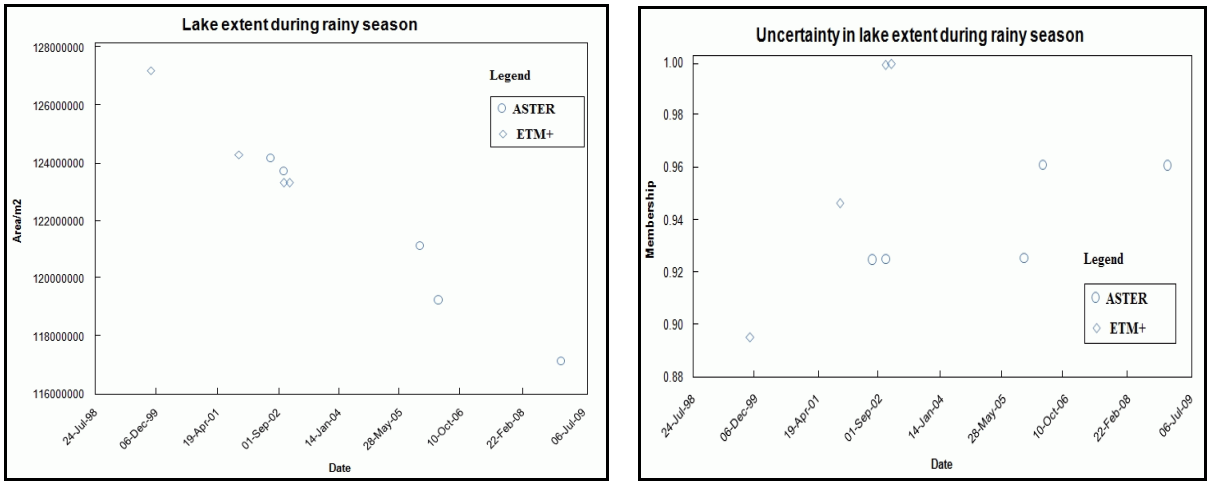


Figure 6.9 Lake extent and its uncertainty during rainy season : ASTER and ETM+ combined

By visual inspection of figure 6.9, there is an offset between the two datasets. The dataset from ETM+ sensor is lower than that from ASTER sensor. However, it is difficult to conclude that the offset is systematic since there is a little overlap in time of acquisition.

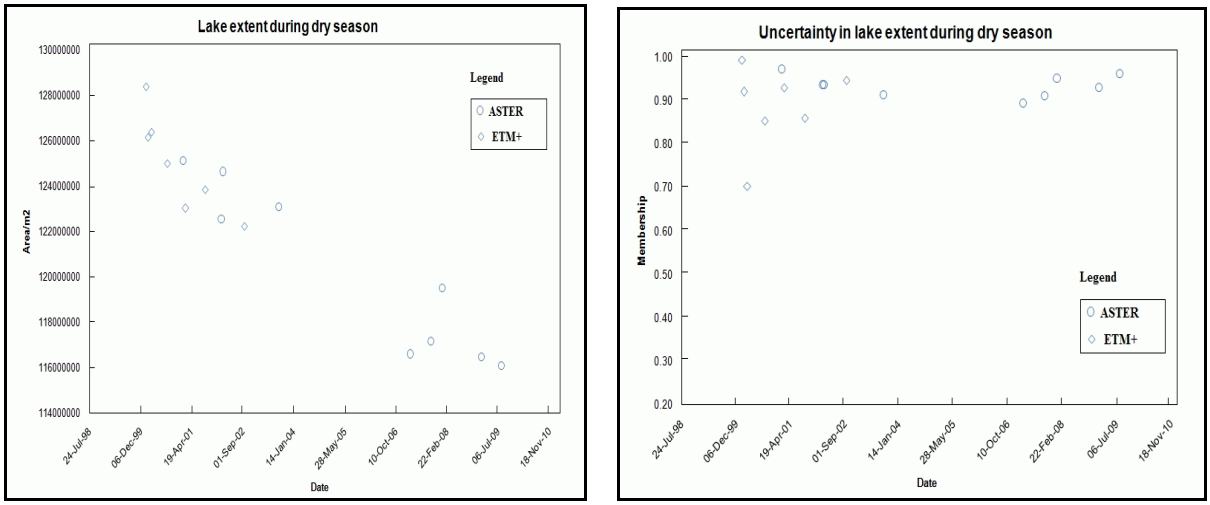


Figure 6.10 Lake extent and its uncertainty during dry season: ASTER and ETM+ combined

Figure 6.10 also shows that there is an offset between the two datasets. The dataset from ETM+ sensor is again lower than dataset from ASTER sensor.

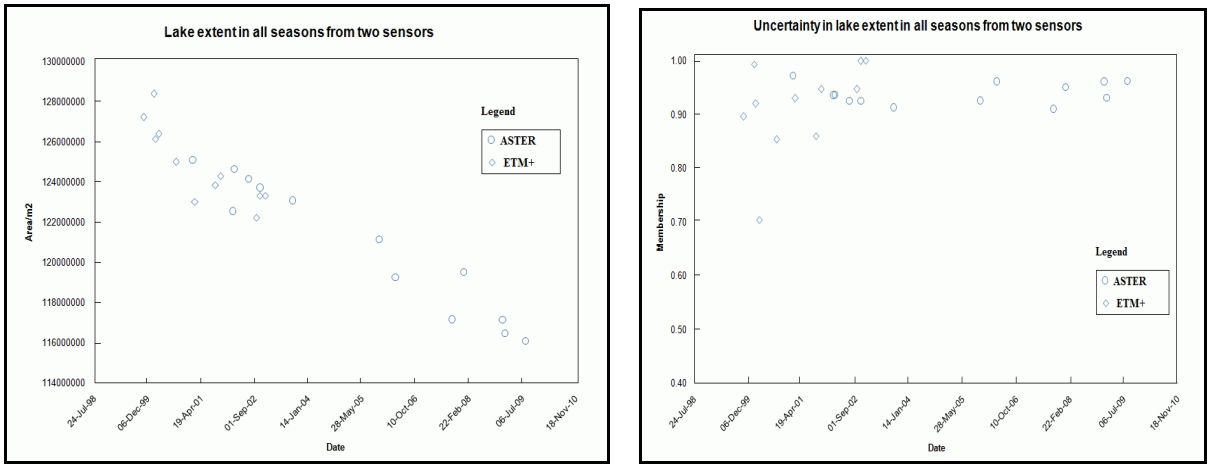


Figure 6.11 Lake extent and its uncertainty from 1999 to 2009 :Observed from two sensors

Figure 6.11 shows that there is an offset between the two datasets in all seasons of observation. However, the two datasets are linearly correlated as an indicator of representing the same phenomena.

Considering the two datasets as different random samples, I tested their correlation by determining the correlation coefficient. Table 6.7 shows the correlation coefficients in two seasons and all seasons combined.

Table 6.7 Correlation between datasets during rainy season, dry season and all seasons in total

Season	Correlation coefficient	
	Area	Membership
Rainy	0.77	0.54
Dry	0.85	0.28
All	0.71	0.04

Table 6.7 shows that the two datasets are positively correlated in all seasons. However, there is very weak positive correlation in the uncertainty within lake extent as indicated by the correlation coefficient when datasets acquired in dry and rainy seasons are combined. This implies that what is observed by one sensor at one time is different from that observed by the sensor at the same time. But this cannot be the reliable conclusion as there is no enough overlap between these datasets. Based on the evidence that the two datasets are positively correlated, I perfomed linear regression analysis to estimate the trend in lake extent by combining observations from the two sensors.

6.4.1. Linear regression modelling of lake extent trend from two sensors

In this section similar procedures for linear regression analysis as described in sections 6.3 and 6.4 were followed. Figures 6.12 and 6.13 illustrate the data points used to estimate the best fit trend line using measurements from two sensors. In figure 6.12 the trend in lake extent is estimated by combining measurements obtained during dry and rainy season. The left plot of the figure shows the estimated trend in lake extent by combining observations from ASTER and ETM+ sensor during rainy season while the right plot of the same figure shows the trend in lake extent during dry season. The plot in figure 6.13 shows the trend in lake extent from combined observations as acquired by the two sensors between 1999 and 2002 for ETM+ sensor and 2001 and 2009 for ASTER. By visual

inspection it is seen that by combining the observations from both sensors the slope of the line changes.

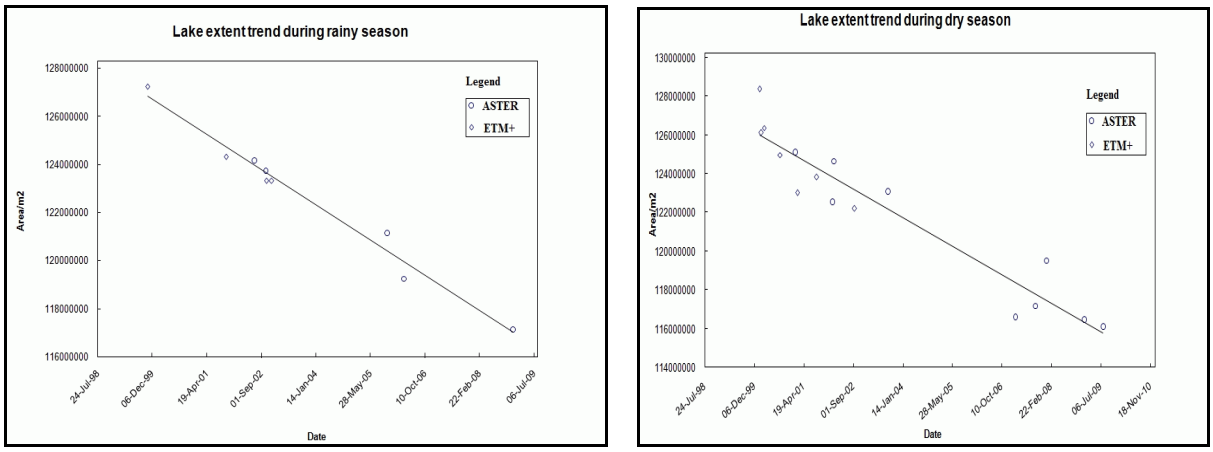


Figure 6.12 Lake extent trend during rainy and dry season: Two sensors combined

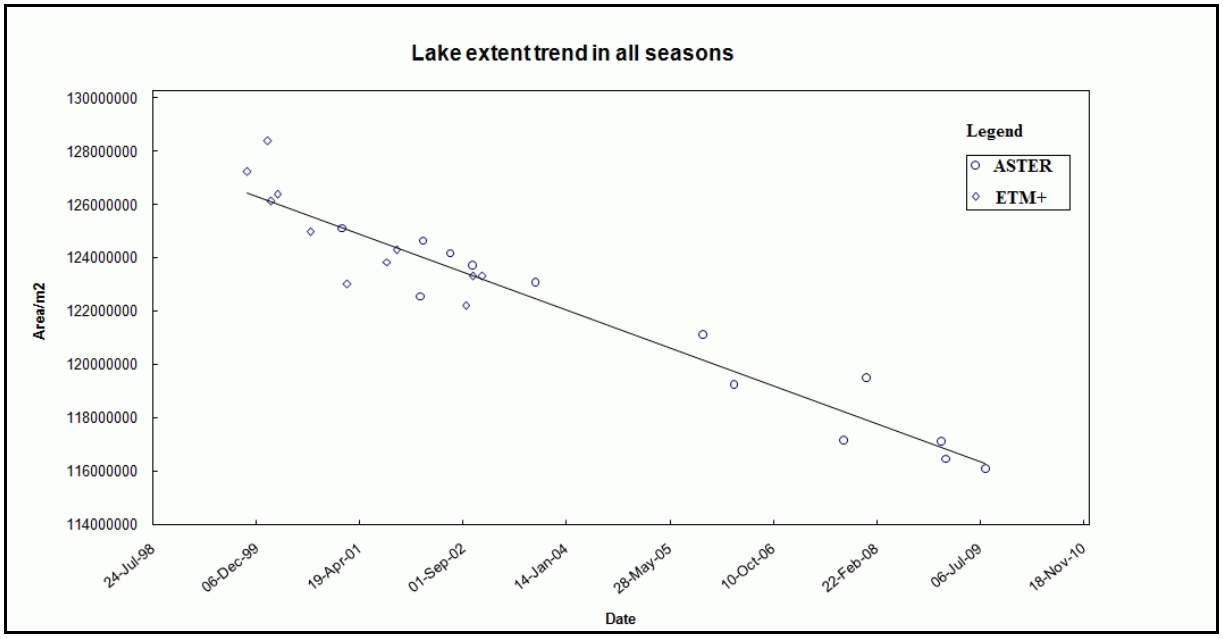


Figure 6.13 Lake extent trend from 1999 to 2009: ASTER and ETM+ sensor combined

Table 6.8 Quality of the fitted trend line using measurements from two sensors

Date	R^2	Adjusted R^2	t-statistic	F-statistic	p-value
Rainy season	0.98	0.97	-17.56	308.48	4.78×10^{-7}
Dry season	0.91	0.90	-11.61	134.72	1.43×10^{-8}
All seasons	0.92	0.91	-16.00	256.01	5.87×10^{-14}

In table 6.8, while 97% of variability in the datasets from two sensors during rainy season has been well represented by the model, 90% of the variability during dry season is represented by the model. Combining the two datasets in all seasons, 91% of the variability in the dataset is well represented by the fitted model. In addition, the p-value decreases giving evidence of goodness of the fitted model.

6.5. Comparing results from time series analysis

Table 6.9 gives an overview of the results from lake extent time series analysis. During rainy season the lake extent trend is more definite than dry season as indicated by adjusted R-square values from observations by the two sensors. During dry season the lake extent is more uncertain as indicated by the membership function values. This has also impact on estimating trend for future prediction as unpredictable seasonality at the boundary of the lake results into difficulty to identify boundary points at that time. Because of this seasonality, what is observed at one time is different from what will be observed at another time.

Table 6.9 Overview of results for ASTER, ETM+ and combined sensors

Date	Adjusted R^2,p-value ASTER	Adjusted R^2, p-value ETM+	Adjusted R^2,p-value two sensors
Rainy season	0.96, $2.22\text{x}10^{-3}$	0.98, $1.00\text{x}10^{-2}$	0.97, $4.78\text{x}10^{-7}$
Dry season	0.89, $9.15\text{x}10^{-5}$	0.72, $1.00\text{x}10^{-2}$	0.90, $1.43\text{x}10^{-8}$
All seasons	0.90, $1.00\text{x}10^{-7}$	0.71, $6.67\text{x}10^{-4}$	0.91, $5.87\text{x}10^{-14}$

7. Discussion, conclusions and recommendations

7.1 Discussion

Increasingly, available remotely sensed data at irregular time interval can be used to study changes taking place within our planet Earth which may be difficult to detect using a single sensor/platform combination. However, the use of these datasets requires appropriate image mining methods in which uncertainty inherent in each dataset can be explicitly quantified and stated. In addition special attention is required in observing impact of uncertainty in datasets and object definition on studying underlying trend in these objects. The results of this study indicate that natural and unpredictable seasonal variations in water level have an impact on the trend in lake extent. It was observed when combining measurements from different seasons there was decrease in slope of the fitted trend line.

In this study ASTER and ETM+ sensors onboard Terra and Landsat7 satellites respectively were used to analyse changes in lake extent and quantify its uncertainty over time. The first approach in understanding the definition of lake extent was to review the definition of the lake.

From the definition of the lake and wetlands, uncertainty in lake extent definition at one point in time is caused by the interaction of water and ecological units existing at the boundary of the lake. With a focus to the study area and available data, these ecological units were identified as floating vegetation, papyrus swamps with some emergent and submerged vegetation while the papyrus swamps share common boundary with grassland dominated by scrub [33].

Object-oriented image analysis was used to extract meaningful information corresponding to the lake extent. The procedure starts with transformation of raster grids into homogeneous regions that correspond to (part of) real world objects through the process of segmentation. In this study, multiresolution segmentation (bottom-up region growing) algorithm available in Definiens Developer software was used. Different image objects optimization parameters such as colour and shape parameters were used to obtain meaningful image objects that suit into subsequent analysis in lake extent trend and quantification of its uncertainty. Although there is great flexibility in this step, the problem is on obtaining the right set of parameters which give meaningful image objects of interest.

The scale parameter during image segmentation has greater impact on the resulting image objects. Large value of this parameter results into segments with large size which sometimes are not meaningful for a given application. However, the parameter is useful as it allows hierarchical image classification thus optimizing extracted information.

The shape parameter has also an impact on the resulting image objects. The compactness parameter results into compact shapes. The larger value of this parameter results into image objects with large

borders which sometimes may not be meaningful to a given application. The smoothing parameter results into image objects with smoothed borders. The magnitude of this parameter will depend on the shape of the objects to be extracted from the image.

Fuzzy rule-based classifiers were developed in Definiens software and applied to images from two sensors. The rules were developed based on image objects attributes resulting from segmentation routine. With these attributes, there is a great possibility of optimizing class description rather than depending on single attributes and minimizing uncertainty inherent within pixels. In addition, fuzzy logics were used to quantify this uncertainty.

The problem in this study was to validate the results of classification due to lack of sufficient reference dataset and field visit. Three procedures were adopted after classifying the image, namely looking at the class stability, visual inspection in ERDAS Imagine software after overlaying classified image over original image and consideration of NDVI values in Arc Map based on the assumption that all image objects with NDVI< 0 constitute to water.

From time series analysis, the lake extent has been diminishing from1999 to 2002 in all seasons of the year as observed by ETM+ sensor. The same falling trend in lake extent was observed by ASTER sensor from 2001 to 2009 in all seasons of the year. On the other hand, it was also clear that the lake extent was more uncertain during dry season than the rainy season as indicated by the membership function values. This is possible because during dry season water level decreases and some reflections from the bed of the lake are recorded by the sensor. In addition during dry season it might be that some submerged vegetation emerges due to water level fall and hence what is observed is not pure water but a mixture of water and vegetation resulting into low values of membership. However, it is difficult to make a conclusion about the uncertainty of lake extent during rainy season due to existing wetlands that also change with time. This is because during rainy season, water level rises and the areal extent of the lake increases since some parts of the wetlands are submerged causing uncertainty in lake extent at that time. Further consideration for reliable conclusion may be that we collect information about the history of that lake and its surrounding wetlands such as abundance and distribution of vegetation species surrounding the lake.

Combining observations from two sensors indicated some improvement in the fitted model as explained by an increasing index of correlation (R squared) from 90% to 91 and decreasing P-value. However, we cannot rely on this result to conclude that the two sensors represent the same phenomena since few observations from ETM+ were available for this study. In addition to this there was little overlap between the ETM+ and ASTER datasets to study the correlation between the two datasets. For further investigation on the correlation of these two sensors, there is a need to find out the means of correcting existing data gaps due to SLC failure by ETM+ since 2003.

Now the question to be answered based on the results is: why does the size of the lake change over time? According to literature, over centuries, there have been variations in climate and enormous shifts in rainfall patterns as evidenced on the meteorological data at Naivasha [48]. In addition, there is evidence that period of drought induced famine in the past were associated with low lake levels at Naivasha [48]. This implies that there might be migration of people, who depend on agriculture, from

drier areas to wetlands within Lake Naivasha basin in search arable land thus disturbing equilibrium in water cycle system. It is also believed that there is natural and unpredictable fluctuation of water levels which the lake has been experiencing, resulting in a drawn down zone of several vertical meters [24]. It is again believed that there is high rate of water abstraction from the lake due to various agricultural activities surrounding the lake [9].

7.2 Conclusions

The over all objective of this research was to model change in lake extent and its uncertainty given observations from multiple sensors. The determination of change in lake extent was achieved by estimating the size of the lake at one moment in time and tracking the same size on a series of images from two sensors using object oriented image analysis with fuzzy logics. Images from ETM+ sensor onboard Landsat7 and ASTER sensor onboard Terra were used to study the underlying trend in uncertain lake extent. The two sensor platforms are in the orbit with same temporal resolution of 16 days and dynamic range of 8 bits though they differ in spatial and spectral resolution.

Object oriented image analysis when combined with fuzzy logics can give reliable results appropriate for a particular application. However, the problem is validation of results especially when a series of images is mined together to derive meaningful information about environmental phenomena.

From the results of image analysis and classification, the lake extent varied from 124176150 to 117136800 m^2 during rainy season as observed by ASTER sensor from 2002 to 2009. The uncertainty in lake extent varied between 0.92 and 0.96. During dry season, the lake extent varied from 125117775 to 116096400 m2 while uncertainty varied between 0.89 and 0.97 as observed by ASTER sensor from 2001 to 2009.

The results from ETM+ images acquired during rainy season showed that the lake extent varied from 127221300 to 123316650 m2 between October 1999 and October 2002. There was an increase in lake extent from 123316650 to 123318675 m2 between October 2002 and December 2002. Uncertainty in all time intervals varied between 0.90 and 1. During dry season, the lake extent varied from 128388825 to 123016725 m2 between January 2000 and February 2001 while it increased to 123837750 m2 on August 2001. There was decrease in size to 122213250 m2 as observed on September 2002. In all cases the uncertainty in lake extent varied between 0.70 and 0.99.

Comparing results of linear regression analysis by individual sensor and when combined, the falling trend improved from R-squared 90% to 91% while the P-value decreased enormously from $1.00\text{x}10^{-7}$ and $6.67\text{x}10^{-4}$ for ASTER and ETM+ respectively to $5.87\text{x}10^{-14}$ giving evidence that the trend in uncertain lake extent can be best studied by combining these sensors thus increasing the number of required observations.

7.3 Recommendations

For improved results by adopting the methodology of this study, I recommend the following:

- Although the results of this study are promising, few remotely sensed data from ETM+ sensor were involved. This is because of data gaps since 2003 to date due to scan line corrector (SLC) failure. It could be a good idea to find out how to correct for dropout lines and combine

the two sensors within the same time span and compare the results. Therefore, if there are measurements from both sensors in the same time span, then we can test the correlation of the two trend lines from both sensors for strong evidence that the two sensors represent the same phenomenon

- If the correlation is good, then we can think of combining another sensor such as MODIS onboard Terra satellite and observe whether there is significant improvement in the trend within the uncertain objects such a the changing lake extent and forest boundary. However, it is important to note that MODIS sensor has course resolution compared to ASTER and ETM+.
- Furthermore, I recommend that for improved classification results at one moment in time, the following datasets are important
    - The digital elevation model (DEM) can be incorporated after classification especially when tracking those image objects that constitute to total size of the lake. This is because the lake is shallow with depth of 4 m on average where some deep-rooted vegetation can grow.
    - In addition to the DEM of this area, I recommend to have bathymetric data portraying the depth of the lake bed at one moment in time. If these models of lake depth and DEM are combined together, the certainty in boundary points for the lake can certainty be determined at that time. Thus, the monitoring activities can follow to track changes over time in the lake extent
    - In addition, it is important to have a complete history about the origin of the lake as well as the origin of the surrounding wetlands since the definition of the latter by the Ramsar convention may cause some ambiguity especially for shallow lakes as with time they turn into wetlands
- For concise understanding of underlying trend within geographical phenomena that are uncertain in nature such as lake extents, it is of vital importance to have dataset such as
    - Climate information about an area under study as this will improve the certainty in the derived information from remotely sensed data
    - Land use information which again can improve the confidence in information extracted from images.
    - Hydrology information about all streams that discharge water to the lake is of great importance in the study of trend in lake extent.
- While analysing trend in lake extent, all observations were given equal weight. However, each observation had its own uncertainty. For further investigation one may consider propagation of the uncertainty in modelling trend in lake extent by applying membership values to all residuals and observe the impact of this uncertainty in the resulting trend.
- Lastly, object oriented image analysis software like Definiens Developer is very expensive to purchase which could not be a good idea for research purpose. However, there are free software such as SPRING 5.0 developed by the National Institute for Space Research in Brazil (http://www.dpi.inpe.br/spring/english/index.html). Although I am not sure whether it can involve fuzzy set concepts in classification, one could try and see if it works.

URL's:

1. The Lake Naivasha Riparian Association (LNRA): http://web.ncf.ca/es202/naivasha/who_are_we.html, accessed: 2010-02-08
2. The Lake Naivasha Riparian Association (LNRA): http://web.ncf.ca/es202/naivasha/links.html, accessed: 2010-02-08
3. Satellite imaging corporation: http://www.satimagingcorp.com/satellite-sensors/aster.html, accessed: 2009-09-20
4. Satellite imaging corporation: http://www.satimagingcorp.com/satellite-sensors/landsat.html, accessed: 2009-09-20
5. USGS: http://glovis.usgs.gov, accessed: 2009-10-07

References

1. Adams, C.S.,Boar R. R.,Hubble D. S.,Gikungu M.,Harper D. M.,Hickley P. and Tarras-Wahlberg N., *The dynamics and ecology of exotic tropical species in floating plant mats: Lake Naivasha, Kenya.* Hydrobiologia, 2002. **488**(1): p. 115-122.
2. Baatz M. and Schäpe A. *Multiresolution Segmentation–an optimization approach for high quality multi-scale image segmentation*. 2000.
3. Bayly I.A.E. and Williams W.D., *Inland waters and their ecology*. 1973: Longmans, Melbourne, 316pp.
4. Benz Ursula C., Hofmann Peter, Willhauck Gregor, Lingenfelder Iris, and Heynen Markus, *Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information.* ISPRS Journal of Photogrammetry and Remote Sensing, 2004. **58**(3-4): p. 239-258.
5. Bezdek James C., Ehrlich Robert, and Full William, *FCM: The fuzzy c-means clustering algorithm.* Computers & Geosciences, 1984. **10**(2-3): p. 191-203.
6. Bijker W., *Needs and applications for data mining in large series of remotely sensed images* In: IGARSS 2009 : Proceedings of the IEEE International Geoscience & Remote Sensing Symposium : Earth observation, origins to applications, July 12-17, 2009, Cape Town, South Africa. IEEE, 2009. 4 p., 2009.
7. Blaschke T., Lang S., Lorup E., Strobl J., and Zeil P., *Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications.* Environmental information for planning, politics and the public, 2000. **2**: p. 555–570.
8. Bontemps Sophie, Bogaert Patrick, Titeux Nicolas, and Defourny Pierre., *An object-based change detection method accounting for temporal dependences in time series with medium to coarse spatial resolution.* Remote Sensing of Environment, 2008. **112**(6): p. 3181-3191.
9. Catherine Riungu. *Lake Naivasha is Dying - Government, Users Wake Up to Nightmare Reality*. 2009 [cited 2010 5 February]; Available from: http://allafrica.com/stories/200907201252.html.
10. Cheng, T., Molenaar M., and A. Stein, *Fuzzy Approach for Integrated Coastal Zone Management*, in *Remote Sensing and Geospatial Technologies for Coastal Ecosystem Assessment and Management*. 2009. p. 67-90.

11. Costa L.T., Farinha J.C., Tomàs Vives P., Hecker N., and Silva E.P., *Regional wetland inventory approaches: The Mediterranean example.* Wetland inventory, assessment and monitoring: Practical techniques and identification of major issues, 2001.
12. Coughanowr, C. (1998) *Wetland of the humid tropics, Water related issues and problems of the humid Tropics and other warm humid regions:* . IHP Humid tropics Programme series no. 12, UNESCO Paris: http://unesdoc.unesco.org/images/0011/001160/116084eo.pdf.
13. Darwish A., Leukert K., and Reinhardt W. *Image segmentation for the purpose of object-based classification*. 2003.
14. Definiens Imaging. *Definiens user guide [Online],accessed:17th November 2009*. Available from: http://www.definiens.com/.
15. Dugan P.J., *Wetland conservation: A review of current issues and required action*. 1990: The World Conservation Union (IUCN).
16. Finlayson C. M. and Valk A. G., *Wetland classification and inventory: A summary.* Plant Ecology, 1995. **118**(1): p. 185-192.
17. Fisher P., *The pixel: a snare and a delusion.* International journal of remote sensing, 1997. **18**(3): p. 679-685.
18. Fisher P.F., *Models of Uncertainty in Spatial Data.* In Geographical Information Systems:Principles, Techniques, Management and applications/ ed. by M.G. P.Longley, D.Maguire, and D.Rhind, 1999: Wiley and Sons, New York., 1999.
19. Foley Jonathan A., Levis Samuel, Costa Marcos Heil, Cramer Wolfgang, and Pollard David. *Incorpoarting Dynamic Vegetation Cover Within Global Climate Models.* Ecological Applications, 2000. **10**(6): p. 1620-1632.
20. Forel F.A., *Handbuch der Seenkunde: Allgemeine Limnologie*. 1901: J. Engelhorn, Stuttgart.
21. Gopal B. and Sah M., *Inventory and classification of wetlands in India.* Plant Ecology, 1995. **118**(1): p. 39-48.
22. Hailu Kassaye R., *Suitability of Markov random field based method for super resolution land cover mapping*. 2006, ITC: Enschede. p. 77.
23. Harper D. M., Mavuti K. M., and Muchiri S. M., *Ecology and Management of Lake Naivasha, Kenya, in Relation to Climatic-Change, Alien Species Introductions, and Agricultural-Development.* Environmental Conservation, 1990. **17**(4): p. 328-336.
24. Harper David, *The ecological relationships of aquatic plants at Lake Naivasha, Kenya.* Hydrobiologia, 1992. **232**(1): p. 65-71.
25. Hoover A.,Jean-Baptiste G.,Jiang X.,Flynn P.J.,Bunke H.,Goldgof D. and Bowyer K., *A Comparison of Range Image Segmentation Algorithms.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996. **18**(7): p. 673–689.
26. Klein L.A., *Sensor and data fusion concepts and applications*. 1999: Society of Photo-Optical Instrumentation Engineers (SPIE) Bellingham, WA, USA.
27. Lucieer A., Kraak M.J., and Stein A., *Uncertainties in segmentation and their visualisation*, in *ITC Dissertation;113*. 2004, ITC: Enschede. p. 174.
28. MacKay H., Finlayson C. M., Fernández-Prieto D., Davidson N., Pritchard D., and Rebelo L. M., *The role of Earth Observation (EO) technologies in supporting implementation of the Ramsar Convention on Wetlands.* Journal of Environmental Management, 2009. **90**(7): p. 2234-2242.
29. Matthews G.V.T., *The Ramsar Convention on Wetlands: its history and development*. 1993: Le Brassus, Switzerland.
30. Mavrantza O. D. and Karadimas N. V. *Fuzzy Representation of Land Cover Classes Using Digital Geo-Data*. in *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*. 2007.
31. Mena Lopez, S., *Papyrus conservation around Lake Naivasha : development of alternative management schemes in Kenya*. 2002, ITC: Enschede. p. 87.
32. Molenaar Martien and Cheng Tao, *Fuzzy spatial objects and their dynamics.* ISPRS Journal of Photogrammetry and Remote Sensing, 2000. **55**(3): p. 164-175.

33. NgariA. N., Kinyamario J. I., Ntiba M. J., and Mavuti K. M., *Factors affecting abundance and distribution of submerged and floating macrophytes in Lake Naivasha, Kenya.* African Journal of Ecology, 2009. **47**(1): p. 32-39.
34. Nock Richard and Nielsen Frank, *Semi-supervised statistical region refinement for color image segmentation.* Pattern Recognition, 2005. **38**(6): p. 835-846.
35. O'Sullivan P.E. and Reynolds C.S., *The lakes handbook. Limnology and limnetic ecology.* Limnology, 2004. **1**: p. 8.
36. Openshaw S., *The Modifiable Areal Unit Problem: Concepts and Techniques in Modern Geography 38.* Norwich: Geo books, 1984.
37. Pal N. R., Pal K., Keller J. M., and Bezdek J. C., *A Possibilistic Fuzzy c-Means Clustering Algorithm.* Fuzzy Systems, IEEE Transactions on, 2005. **13**(4): p. 517-530.
38. Richards J.A. and Jia X., *Remote sensing digital image analysis: an introduction*. 2006: Springer Verlag.
39. Rowan L.C. and Mars J.C., *Lithologic mapping in the Mountain Pass, California area using advanced spaceborne thermal emission and reflection radiometer (ASTER) data.* Remote Sensing of Environment, 2003. **84**(3): p. 350-366.
40. Sikazwe J., *Optimizing Markov random field for super - resolution land cover mapping*. 2007, ITC: Enschede,. p. 43.
41. Stein A., *Handling uncertainties in image mining for remote sensing studies.* In: Proceedings of the 8th international symposium on spatial accuracy assessment in natural resources and environmental sciences, Shanghai, P.R. China, June 25-27, 2008 Liverpool : M.F. World Academic Press, 2008. ISBN 978-1-84626-170-1 pp. 362-368, 2008.
42. Stein A., *Modern developments in image mining.* In: Science in China series E : technological sciences, 51(2008) supp. 1. pp. 13-25, 2008.
43. Stein A., Hamm N.A.S., and Qinghua Y., *Handling uncertainties in image mining for remote sensing studies.* International journal of remote sensing, 2009. **30**(20).
44. Tao Cheng, Molenaar M., and Stein A., *Fuzzy approach for integrated coastal zone management.* In: Remote sensing and geospatial technologies for coastal ecosystem assessment and management. / ed. by Xiaojung Yang. Berlin : Springer, 2009. ISBN: 978-3-540-88182-7. pp. 67-90, 2009.
45. Tolpekin V.A., Kassaye R.H., and Stein A., *Super resolution mapping from satellite images with Markov random field and simulated annealing* In: Abstracts of the workshop on modern aspects of remote sensing and geospatial data processing, NRSA - ITC, Hyderabad, India, 24-25 November 2006 / ed by. K.M.M. Rao, A. Stein, P.S. Roy and A.S. Manjunath. Hyderabad : National Remote Sensing Agency (NRSA), 2006. 4 p., 2006.
46. Tolpekin V.A. and Stein A., *Quantification of the effects of land - cover - class spectral separability on the accuracy of Markov - random - field - based superresolution mapping.* IEEE Transactions on geoscience and remote sensing, 2009. **47**(9).
47. Tso B. and Olsen R.C., *A contextual classification scheme based on MRF model with improved parameter estimation and multiscale fuzzy line process.* Remote Sensing of Environment, 2005. **97**(1): p. 127-136.
48. Tyson P. D., Lee-Thorp J., Holmgren K., and Thackeray J. F., *Changing Gradients of Climate Change in Southern Africa during the Past Millennium: Implications for Population Movements.* Climatic Change, 2002. **52**(1): p. 129-135.
49. Umamaheshwaran R., Bijker W., and Stein A., *Image mining for modeling of forest fires from meteosat images.* IEEE Transactions on geoscience and remote sensing, 2007. **45**(1).
50. Villa B., Emmolo D., and Orlando P., *Evaluation of capabilities of fuzzy logic classification of different kind of data.* 2008.
51. Wang F., *Fuzzy supervised classification of remote sensing images.* Geoscience and Remote Sensing, IEEE Transactions on, 1990. **28**(2): p. 194-201.
52. Zadeh L. A., *Fuzzy sets as a basis for a theory of possibility.* Fuzzy sets and systems, 1999. **100**(Supplement 1): p. 9-34.

53. Zhang Xiaoyang, Friedl Mark A., and Schaaf Crystal B., *Sensitivity of vegetation phenology detection to the temporal resolution of satellite data.* International journal of remote sensing, 2009. **30**(8): p. 2061 - 2074.

Appendices

Accuracy assessment ASTER images during dry season

22-Jan-01					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	120	0.5592	0.2806	0.1054	0.9801
Adjacent to water	6	0.5016	0.3301	0.1369	0.9575
water	84	0.9778	0.0139	0.9313	0.9993
01-Feb-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	53	0.3864	0.2473	0.1022	0.8968
Adjacent to water	47	0.4699	0.3320	0.1034	1.0000
water	120	0.9656	0.0260	0.8883	0.9997
17-Feb-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	57	0.4204	0.2513	0.1009	0.9727
Adjacent to water	56	0.4428	0.2959	0.1008	1.0000
water	114	0.9649	0.0271	0.8937	1.0000
15-Aug-03					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	46	0.4541	0.2635	0.1019	0.9903
Adjacent to water	56	0.3652	0.2594	0.1038	1.0000
water	116	0.9635	0.0337	0.8973	0.9999
24-Feb-07					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	118	0.4779	0.2734	0.1000	0.9511
Adjacent to water	16	0.4303	0.2494	0.1224	0.8832
water	119	0.9600	0.0412	0.8688	0.9999
11-Sep-07					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	116	0.4414	0.2571	0.1009	0.9746
Adjacent to water	14	0.4306	0.2752	0.1070	1.0000
water	126	0.9491	0.0389	0.8923	0.9989
01-Jan-08					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	78	0.3571	0.2522	0.1026	0.9709
Adjacent to water	44	0.4809	0.3213	0.1023	1.0000
water	185	0.9753	0.0210	0.9028	0.9998

19-Jan-09					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	53	0.3098	0.2346	0.1002	1.0000
Adjacent to water	66	0.4690	0.3021	0.1010	1.0000
water	188	0.9747	0.0244	0.9022	1.0000
30-Jul-09					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	33	0.3407	0.2325	0.1024	0.9953
Adjacent to water	47	0.4664	0.2701	0.1032	1.0000
water	131	0.9741	0.0211	0.9041	0.9995

Accuracy assessment: ASTER images during rainy season

29-Jun-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	35	0.4039	0.2529	0.1012	0.8989
Adjacent to water	21	0.4078	0.3090	0.1108	1.0000
water	166	0.9479	0.0313	0.8996	0.9987
15-Oct-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	42	0.4011	0.2373	0.1109	0.9518
Adjacent to water	22	0.3207	0.1871	0.1057	0.8058
water	148	0.9451	0.0321	0.9024	0.9992
01-Nov-05					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	48	0.4453	0.2570	0.1061	0.9686
Adjacent to water	33	0.3765	0.2569	0.1019	0.9775
water	117	0.9644	0.0303	0.8971	0.9998
01-Apr-06					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	71	0.4605	0.2488	0.1009	0.9781
Adjacent to water	5	0.7015	0.2488	0.3953	0.9286
water	78	0.9695	0.0185	0.9097	0.9989
23-Nov-07					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	91	0.3816	0.2452	0.1049	0.9723
Adjacent to water	39	0.5021	0.3017	0.1001	1.0000
water	142	0.9740	0.0259	0.9096	0.9999
27-Dec-08					
Class	Objects	Mean	StdDev	Minimum	Maximum
Vegetation	89	0.3670	0.2643	0.1066	0.9790
Adjacent to water	57	0.3605	0.2515	0.1030	1.0000
water	152	0.9816	0.0152	0.9156	1.0000

ETM+ images

Accuracy assessment: ETM+ images during dry season

27-Jan-00					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	33	0.8957	0.2731	0.1088	1.0000
Not water	0	0.0000	0.0000	0.0000	0.0000
Adjacent to water	32	0.6910	0.3001	0.1073	1.0000
12-Feb-00					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	23	0.8769	0.1494	0.3710	0.9968
Not water	37	0.2750	0.1997	0.1093	0.7942
Adjacent to water	22	0.7781	0.2108	0.3281	1.0000
15-Mar-00					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	38	0.6296	0.2210	0.2264	0.9258
Not water	16	0.2707	0.1635	0.1032	0.7127
Adjacent to water	27	0.7088	0.2297	0.2258	0.9891
22-Aug-00					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	34	0.7918	0.2006	0.3029	1.0000
Not water	20	0.2714	0.1575	0.1059	0.6055
Adjacent to water	28	0.8189	0.2194	0.2296	0.9882
14-Feb-01					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	27	0.9207	0.0688	0.7272	0.9934
Not water	456	0.3165	0.2106	0.1002	0.9521
Adjacent to water	25	0.7406	0.2541	0.1157	1.0000
25-Aug-01					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	35	0.7902	0.2185	0.1034	0.9811
Not water	41	0.2287	0.1630	0.1026	0.7707
Adjacent to water	28	0.7525	0.2615	0.1669	1.0000
13-Sep-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	40	0.8893	0.1945	0.1643	1.0000
Not water	31	0.2648	0.1656	0.1012	0.6742
Adjacent to water	31	0.7575	0.2607	0.1331	1.0000

Accuracy assessment: ETM+ images during rainy season

23-Oct-99					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	51	0.7350	0.2778	0.1156	1.0000
Not water	153	0.4123	0.2818	0.1010	0.9739
Adjacent to water	16	0.8257	0.2269	0.1449	0.9961
12-Oct-01					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	50	0.8981	0.1882	0.1287	0.9991
Not water	68	0.5932	0.3058	0.1249	1.0000
Adjacent to water	5	0.7097	0.2361	0.4193	1.0000
15-Oct-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	54	0.9596	0.1314	0.3681	1.0000
Not water	119	0.9403	0.1374	0.2782	1.0000
Adjacent to water	12	0.5411	0.3146	0.1341	0.9817
02-Dec-02					
Class	Objects	Mean	StdDev	Minimum	Maximum
Water	34	0.8850	0.2637	0.1239	1.0000
Not water	190	0.9060	0.2120	0.1172	1.0000
Adjacent to water	0	0.0000	0.0000	0.0000	0.0000