Monitoring aquatic plants proliferation in Lake Victoria using satellite data

by

CHERUIYOT, Elijah K. (BSc.)

Jan, 2012

A thesis submitted in partial fulfilment of the requirements for the award of the degree of Master of Science (Physics) of the University of Nairobi.

Declaration

This thesis is my original work and has not been presented for examination in any other university.

Signature.............................

Date.........06/02/2012.........

CHERUIYOT, Elijah K., BSc. Hons

The undersigned supervisors certify that they have read and hereby recommend for acceptance by the University of Nairobi a thesis entitled ***Monitoring aquatic plants proliferation in Lake Victoria using satellite data***, in fulfilment of the requirements for the degree of Master of Science (Physics).

1. **Dr. Collins Omulo Mito**
   Department of Physics
   University of Nairobi

   Signature.....................
   Date.....06/02/2012

2. **Dr. Kenneth Amiga Kaduki**
   Department of Physics
   University of Nairobi

   Signature.....................
   Date............6/2/2012

3. **Prof. Giovanni Laneve**
   Centro di Ricerca Progetto San Marco (CRPSM)
   University of Rome

   for Signature.....................
   Date.......06/02/2012

To my dad Reuben and mum Rebecca.

Abstract

This work reports the monitoring of aquatic plants proliferation in Lake Victoria using satellite data over the period 2003 to 2010. The lake, which is the second largest freshwater body in the world and an important economic resource, is facing serious environmental challenges including growth of invasive plants. The role of some selected water quality parameters as well as meteorological conditions in aquatic plants proliferation is also investigated. Multispectral MERIS (Medium Resolution Imaging Spectrometer) imagery, were obtained from ESA in the framework of TIGER Initiative. Images were selected on the basis of spatial and temporal coverage, spatial and spectral resolution and the severity of cloud cover. Atmospheric correction was carried out as pre-processing procedure to improve on the image interpretability. The images were processed using *BEAM 4.8* and *ENVI 4.2* image processing and analysis software. Image derived endmembers were used to classify the images using linear spectral unmixing classification technique. Temporal variation of vegetation was obtained, and the spatial distribution was presented by cover maps. Mapping was done using *ArcGIS 9.3*. TSM and Chl-a values were retrieved from the images using *MERIS Eutrophic Lakes Processor 1.4.1* while rainfall data was obtained from Kenya Meteorological Department (KMD). Spectral unmixing technique performed well with a mean classification accuracy of 99.48%, based on RMSE. Most images showed a high concentration of Chl-a and TSM along the shores of the lake, especially the Winam Gulf, and most of the aquatic vegetation was observed in the same regions. Vegetation cover in the Winam Gulf which was kept below 100 km^2 during the years 2003 to 2006 increased to a peak of about 200 km^2 in 2007, before decreasing again to below 100 km^2 during the years 2008 to 2010. Regression results for Winam Gulf show that vegetation abundance has a weak linear correlation with rainfall, TSM and Chl-a of $R = 0.67$ ($R^2 = 0.44$), $R = 0.46$ ($R^2 = 0.21$) and $R = 0.57$ ($R^2 = 0.32$), respectively after a response period of two to three months. From these relations, vegetation abundance prediction models are proposed. At no time delay, however, vegetation abundance showed no significant relationship with these parameters, while TSM and Chl-a are significantly dependent on each other with $R = -0.77$ ($R^2 = 0.6$). While traditional methods of monitoring vegetation and water quality parameters is both expensive and time consuming, remotely sensed satellite data provides reliable, consistent and repeatable information that is suitable for this study.

Acknowledgements

I wish to acknowledge and appreciate my supervisors; Dr. Collins O. Mito, Dr. Kenneth A. Kaduki and Prof. Giovanni Laneve, for the supervision, constructive comments, encouragement and guidance they have provided me throughout my research. Their invaluable support and advice is unforgettable.

I am grateful to the University of Nairobi for the award of scholarship. I also acknowledge my colleagues and all the members of the Department of Physics for their ideas and constructive comments and moral support. Their immense contribution is highly appreciated.

Much appreciation goes to ESA through the TIGER Initiative, for the opportunity to attend their workshops and a short term experience in Delft University of Technology, The Netherlands. I'm also very grateful to Prof. Massimo Menenti and the entire Optical Laser and Remote Sensing (OLRS) team of Delft University of Technology, The Netherlands, for their support and friendliness during my stay with them.

I also appreciate the Regional centre for Mapping of Resources for Development (RCMRD) and all the staff for their invaluable support during my attachment.

Of course this work would not have been a success without the kind support, love and prayers of my beloved family and friends. I owe a lot to them.

Finally, and most importantly, I thank the Lord Almighty, for His love. He is the one who gave me the grace to carry on, good health and provided all I needed to carry out this research. He also brought my way all the above mentioned persons and organizations at the most appropriate time to accomplish His will for my life. I forever will worship Him!

Glossary

BEAM	Basic ERS and ENVISAT (A)ATSR and MERIS Toolbox
Chl-a	Chlorophyll-a
CRPSM	Centro di Ricerca Progetto San Marco
ECMWF	European Centre for Medium-Range Weather Forecasts
ENVI	ENvironment for Visualizing Images
ENVISAT	ENVIronment SATellite
ESA	European Space Agency
FR	Full Resolution
GIS	Geographic Information System
GPS	Global Positioning System
IOP	Inherent Optical Properties
KMD	Kenya Meteorological Department
LVEMP	Lake Victoria Environmental Management Project
MERIS	Medium Resolution Imaging Spectrometer
NIR	Near Infrared
RGB	Red-Green-Blue colour composite image
RMSE	Root Mean Squared Error
ROI	Region of Interest
SAR	Synthetic Aperture Radar
SMAC	Simplified Method for Atmospheric Corrections
TOA	Top of atmosphere
TSM	Total Suspended Matter
UTM	Universal Transverse Mercator
WGS	World Geodetic System

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

1.1 Background

Lake Victoria, the largest of all African Lakes, is also the second largest freshwater body in the world, with a surface area of 68 800 km^2 (Albright *et al.* 2004, Osumo 2001). Its extensive surface belongs to three countries; the northern half to Uganda, the southern half to Tanzania, and part of the north-eastern sector to Kenya; shared in the ratio 45%, 49% and 6% respectively (Osumo, 2001). The Lake Victoria basin in Kenya, Tanzania and Uganda has an estimated population of 23.7 million, representing about 30% of the total population of the three countries (LVFO, 2008). The lake supports one of the largest inland fisheries, producing around one million tonnes per year and providing a livelihood for around 200 000 fishermen and their families, as well as being an important source of export revenue for the riparian countries (Marshall *et al.* 2009). It has however, since late 1980s, been faced with environmental challenges and human impacts which have perturbed the ecological balance affecting its biodiversity (Gichuki, 2010). The most prevalent of them is the growth of aquatic plants, especially the water hyacinth (Mailu *et al.* 2000).

Efforts to control the aquatic plants have been made, attracting organizations like World Bank in conjunction with the three riparian countries through the LVEMP, and ESA through the TIGER Initiative. ESA, under its TIGER Initiative, launched a capacity building campaign, with the aim of training on conservation of water resources in Africa.

Monitoring aquatic vegetation in an extensive area, as in the case of Lake Victoria, can be quite challenging as it require constant collection of data, and mapping activities typically require the collection of extensive ground-truth data. Remotely sensed data, however, have the potential to provide much of the necessary detailed information, e.g. extent and distribution of vegetation in the lake (Jollineau and Howarth, 2002). Remotely sensed satellite data provides consistent and repeatable information (Albright *et al.* 2004). MERIS FR imagery used in this study has both the spatial (300 m) and spectral (400 – 900 nm) resolution, as well as adequate revisit period (three days) that is suitable for the study.

1.2 Study area

1.2.1 Lake Victoria: Geographic location

It stretches 412 km from north to south between 0° 30' N and 3° 12' S and 355 km from west to east between 31° 37' E and 34° 53' E. It lies across the equator at an altitude of 1135 m above sea level. Its extensive surface belongs to the three countries; the northern half to Uganda, the southern half to Tanzania, and part of the northeastern sector to Kenya; shared among the riparian countries in the ratio 45%, 49% and 6% respectively (Osumo, 2001). Figure 1-1 is a map showing the geographic location of the lake.

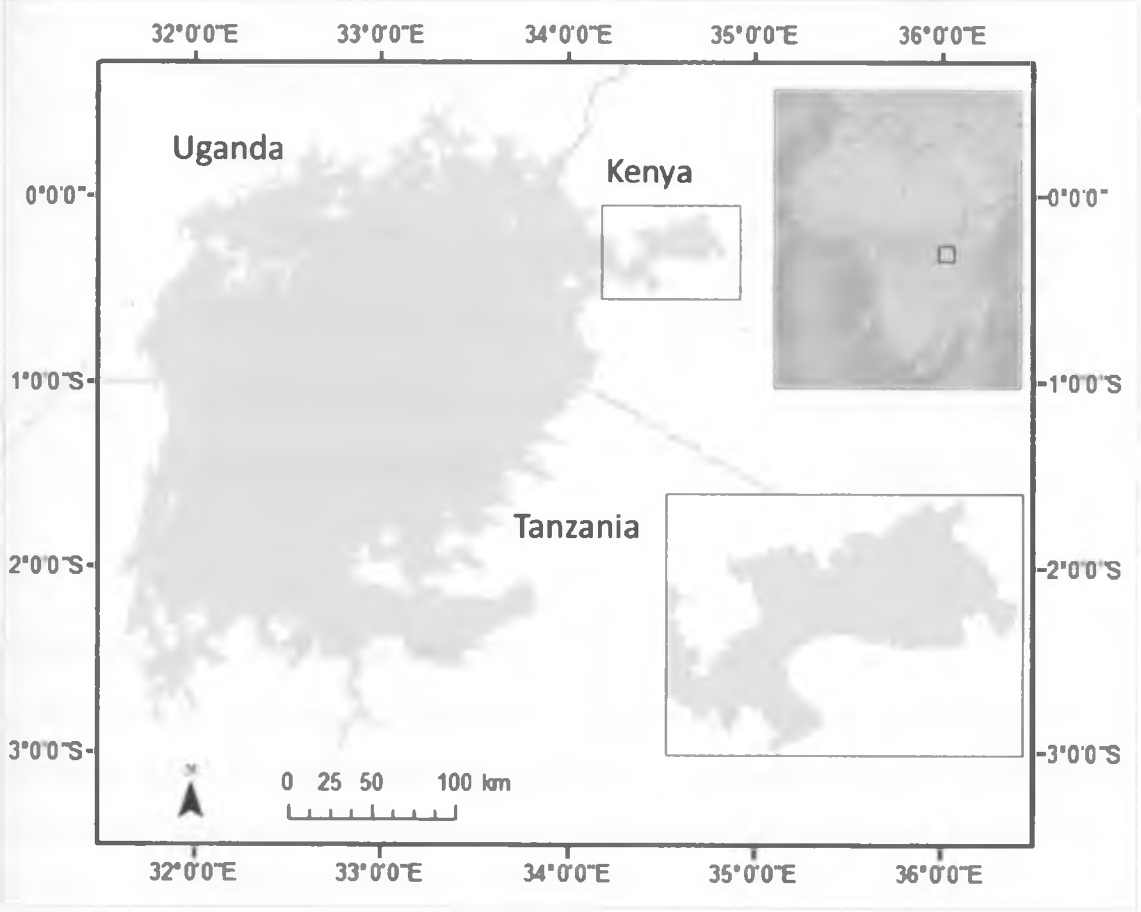

Figure 1-1: Geographic location of Lake Victoria. Inset is the location of Lake Victoria in Africa. Also inset (enclosed in red) is the zoomed-in Winam Gulf section of the lake. Spatial data obtained from DIVA-GIS (2011)

1.2.2 Lake Victoria: Physical dimensions

The lake is relatively shallow, with a recorded maximum depth of about 84 m and an average depth of 40 m. It has a water volume of about 2 750 km^3. It has a long indented shoreline (about 3 440 km), enclosing innumerable small, shallow bays and inlets. It contains numerous islands (Osumo, 2001). Table 1-1 gives the physical dimensions of the lake.

Table 1-1: Physical dimensions of Lake Victoria. Data source: ILEC (2012)

Surface area [km^2]	68 800
Volume [km^3]	2 750
Maximum depth [m]	84
Mean depth [m]	40
Water level	Regulated
Length of shoreline [km]	3 440
Residence time [yr]	23
Catchment area [km^2]	184 000

1.2.3 Lake Victoria: Drainage basin

The Lake Victoria basin in Kenya, Tanzania and Uganda has an estimated population of 23.7 million, representing about 30% of the total population of the three countries (LVFO, 2008). The catchment of Winam Gulf in the Kenyan side of the lake is the main water catchment for the whole lake, and lies between 1000 to 2000 m above sea level (Osumo, 2001). The main rivers flowing into the lake from the Tanzanian catchment are Mara, Kagera, Mirongo, Grumeti, Mbalageti, Simiyu and Mori. From the Kenyan catchment, the main rivers are Nzoia, Sio, Yala, Nyando, Kibos, Sondu-Miriu, Kuja, Migori, Riaria and Mawa, while from the Ugandan catchment the main rivers are Kagera, Bukora, Katonga and Sio.

1.3 Statement of the problem

Water hyacinth and other aquatic plants pose serious problems to Lake Victoria. These include obstruction to navigation, fishing and interference with other aquatic life. To effectively control the proliferation of these invasive plants, reliable information on the identification of the infested areas and the extent of the infestation, as well as the rate of proliferation is required. Control of the plants using mechanical and biological methods both require proper timing of activities. A time series monitoring of aquatic plants is needed to aid in decision making regarding the type of control action that is most suitable for the prevailing conditions, proper timing of the aquatic plants control activities, as well as assessing the efficacy of such activities.

1.4 Objectives / Goal

The aim of this research is to use satellite data to monitor the spatial distribution and temporal variation of the abundance of aquatic plants in Lake Victoria in the period 2003 – 2010. It is also aimed at finding out whether there exists a correlation between the time series variation of vegetation with some selected water quality parameters. In principle, such a correlation would be useful in developing algorithms for the prediction of the state of growth of the plants based on the conditions of the concentrations of the water quality parameters.

1.4.1 Specific objectives

The specific objectives of this project are;

1. To obtain spectral signatures of the dominant image constituents in the lake and develop an endmember spectral library
2. To use spectral unmixing technique to classify images and obtain cover-maps for the lake
3. To use the classified images to estimate the abundance of the aquatic plants in the lake
4. To obtain the temporal (time series) variation of aquatic plants abundance (vegetation phenology)
5. To find out if there exists a correlation between the aquatic plants distribution and Chl-a and TSM water quality parameters as well as rainfall.

1.5 Hypothesis

It is a known fact that aquatic plants exist in Lake Victoria, and they cause undesirable effects. But how do these plants vary spatially? Are some areas more prone to aquatic plants infestation? Do they vary temporally? If they do, could such variations be estimated using satellite data? Could this proliferation be accelerated by the conducive environment provided by some water quality parameters like Chl-a and TSM, introduced into the lake through nutrient enrichment caused by rain in its drainage basin? If it does, does a direct time relationship exist between these parameters and the aquatic plants abundance in the lake, so that an increase in aquatic plants abundance is preceded shortly by a rise in concentrations of Chl-a and TSM? By studying the temporal variation of these water quality parameters, is it possible to predict the future occurrence of 'abnormal growth' of the aquatic plants, and make appropriate plans to control them?

1.6 Report outline

This study is introduced in chapter one, where a brief description of the study area is given. Chapter two discusses literature review. Chapter three describes the theoretical background and discusses the concept of multispectral imaging and classification based on the spectral response of various class features using the spectral unmixing technique. K-means clustering as an unsupervised classification technique is also briefly discussed and a brief description of the two selected water quality parameters, Chl-a and TSM, is given in this chapter.

The methodology used to achieve the specific objectives is described in chapter four, where data requisition and selection criterion is described. Reprojection and atmospheric corrections as image preprocessing procedures as well as classification procedures are also described here. Classification results and accuracy assessment are presented in chapter five. Mapping and monitoring of the spatial and temporal variation of vegetation and water quality parameters as well as their correlations is presented in this chapter, with sample maps given. Analysis and discussions are also given here.

Conclusions and recommendations are presented in chapter six.

2 LITERATURE REVIEW

Over the years, since its infestation in the late 1980's, the rise and fall in the aquatic plants abundance in Lake Victoria has been reported. Researchers have suggested that the rise could be caused by nutrient enrichment in the lake. The fall on the other hand is mainly due to the weed control activities carried out in the lake. Debate arose, however, on whether the decline of water hyacinth abundance in 1998 and later years was due to the biological factors or due to the 1997-1998 El Niño weather pattern that caused stormy and wet weather in the region (Albright *et al.* 2004 and Wilson *et al.* 2007). With availability of periodic and frequent flow of satellite data, and accurate and reliable monitoring techniques, such uncertainties would not occur.

Several methods have been used to control the proliferation of the aquatic plants in Lake Victoria, which include mechanical (shredding of the weed using the 'Swamp Devils') and biological (introducing water hyacinth weevils) (Wilson *et al.* 2007). However, the choice of the control method to employ and the proper timing of such control activities has remained a challenge to all these methods. Biological method, for instance, requires proper timing on the release of the weevils (Gichuki, 2010). During its early stages of infestation, until early 2000s, the available information pertaining to the extent, distribution, and status of water hyacinth in Lake Victoria was largely based on anecdotal accounts, local field observations, and rough estimates (Albright *et al.* 2004). Schouten *et al.* (1999) demonstrated the potential of synthetic aperture radar (SAR) imagery for estimating water hyacinth distribution and extent by providing estimates on three dates in 1998 for selected bays in Uganda and Kenya. The need for reliable information to gauge the severity of the aquatic plants infestation through time, and to relate its abundance to environmental factors, identify areas requiring management action, and assess the efficacy of such actions was highlighted by Albright *et al.* (2004). Remotely sensed satellite data provides consistent and repeatable information. There is therefore a greater need to accurately map and monitor wetlands and their change (Jollineau and Howarth, 2002). Referring to Lake Cuitzeo in Mexico, Ramirez (2006) recommended timely generation of vegetation maps for the continued monitoring of the changes and relations in the lake.

Lake Victoria covers a very large spatial extent, and is therefore not very easy to accurately determine the extent of growth of aquatic plants in it. Remote sensing, however, affords a good estimation by exploiting the spectral features of the image constituents to characterize the remotely sensed satellite images. Spectral feature is regarded as one of the most important pieces of information for remote sensing image interpretation (Qian and Ping, 2007). This is especially true for Lake Victoria because the most dominant image constituents i.e. water and vegetation, have very unique spectral characteristics (spectral signatures), which satellite sensor can detect, and can easily be distinguished using any classifier (Kahlid and ConocoPhillips, 2005). An image is classified by comparing the spectral characteristics of its constituents with the spectral signatures of the known features (endmembers). Multispectral satellite imagery like MERIS allows identification of features by exploiting their spectral responsiveness. MERIS sensor has fifteen spectral bands in the visible and part of NIR regions, ranging from 412.5 nm to 900 nm (ESA, 2010).

In order to effectively use the satellite remote sensing data for land-use and land-cover applications, an appropriate image classification technique must be identified. Akgün *et al.* (2010) emphasized on the selection of the most proper satellite image, band combination, and the classifier for more effective use of the satellite remote sensing information. Spectral feature is a key tool in classifying images so as to generate cover maps. The most unique features of vegetation's spectral responsiveness which allow them to be discriminated especially from water are found in the visible and NIR regions, which fall within the range of acquisition of MERIS data (400 – 900 nm). It also has a swath width of 1150 km, which is wide enough to adequately cover the lake. Furthermore, the provision of MERIS imagery by ESA for this study and the freely available image processing software for MERIS (BEAM) prompted the selection of such data for the project. Incorporated into BEAM also are the bio-optical methods for retrieving water quality parameters from atmospherically corrected MERIS imagery.

Several classifiers exist, which include parallelepiped, minimum distance to mean and maximum likelihood. However, these conventional classification algorithms have a shortcoming of assuming that each pixel consists of only a single endmember (Foppa *et al.* 2002). When these algorithms are applied, for example, to estimate vegetation cover, they require some threshold value to discriminate vegetation from other image constituents and generate binary maps containing 'vegetation' or 'not vegetation'. In practice, this is usually not true since most pixels contain several classes of cover types, especially when the spatial resolution is relatively low. MERIS images, for example, have spatial resolution of about 300 m (ESA, 2010), so that one pixel extents over a large surface area which might cover more than one class features.

Spectral unmixing is a more sophisticated classification technique, which is based on the assumption that the spectrum of a pixel consists of a linear combination of the spectra of several individual land cover types at various proportions (abundances) (Foppa *et al.* 2002, Kumar *et al.* 2007, Matthias and Martin 2003 and Ramirez 2006). Linear mixture modeling considers that the signal received at the sensor is composed of a linear mixture of pure-element reflections (endmembers), where the weights are the percentage of the pixel area occupied by each element. Ideally, if all the endmembers were accurately identified spectrally, then the abundance values of all the endmembers in a pixel would sum to unity. A certain amount of error is however inevitable for different reasons (Foppa *et al.* 2002), so that accuracy of a linear mixture model is measured by the amount of error. For this model to work properly, two constraints are set; that the abundance values of all endmembers plus the residual (called the error) must sum up to unity, and that the abundance value must be positive.

Linear spectral unmixing has been used in several studies in various fields, which include snow cover estimation (Foppa *et al.* 2002), land cover mapping (Kumar *et al.* 2007), and imperviousness of surface distributions (Matthias and Martin, 2003). Before applying the linear spectral unmixing, endmembers for the given study area must be defined by a process called data training. The endmember spectra can be derived either from the image to be classified (image derived endmembers) or from field (*in situ*) measurements (field derived endmembers). Field derived endmembers are obtained by taking *in situ* measurements of the reflectance values of the targeted species using a field spectroradiometer. The maximum number of endmembers is, however, limited by the number of spectral bands of the satellite image (Foppa *et al.* 2002).

It has been established that vegetation is influenced greatly by its environmental factors, so that vegetation information can be derived from the knowledge of the relationship between the vegetation and its environment (Zhigang and Zhuang 2007). The possibility that the problem of macrophyte encroachment in Lake Victoria is greatly enhanced by nutrient enrichment was suggested by Muli (1996). Subsequent studies by Osumo (2001) highlighted the need for a study in nutrient fluxes into the lake, in order to control the problem.

In this work spectral unmixing is applied to detect vegetation in the lake by exploiting its spectral features, so as to map its spatial distribution and obtain a time series variation over the study period. Further, by establishing the influence of factors like TSM, Chl-a and rainfall on the growth of aquatic vegetation, a means to predict vegetation abundance is proposed.

3 THEORETICAL BACKGROUND

3.1 Concept of multispectral images

A multispectral (and hyperspectral) image is one obtained by detecting the spectral response of a scene in more than one spectral band. Pixel values in each band represent the reflectance of the image constituents at that particular wavelength region. Figure 3-1 is an illustration of the multispectral imagery.

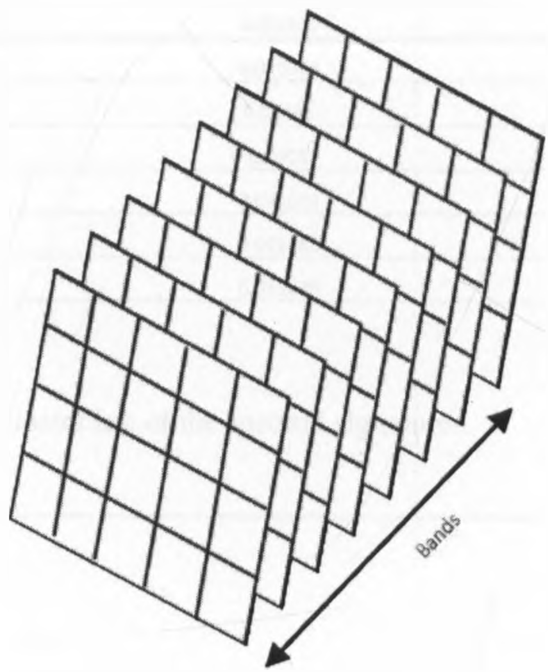

Figure 3-1: Graphical illustration of multispectral imagery

3.2 Spectral signatures

Objects respond differently to radiation of different electromagnetic range. Spectral characteristic of a material is its response to electromagnetic radiation at different wavelengths. Table 3-1 is an example of an endmember library.

Table 3-1: Example of an endmember spectral library for water and vegetation. Data generated from an image in BEAM

	Reflectance values	
Wavelength (nm)	Water	Vegetation
412.691	0.027391	0.015167
442.559	0.02766	0.020045
489.882	0.02758	0.025904
509.819	0.02695	0.031417
559.694	0.027664	0.059096
619.601	0.022619	0.051995
664.5731	0.022334	0.048802
680.821	0.022509	0.049704
708.329	0.022529	0.112117
753.371	0.020584	0.243117
761.5081	0.02485	0.240286
778.4091	0.0208	0.257854
864.876	0.021636	0.284601
884.944	0.021182	0.29052
900.0001	0.031679	0.351611

Figure 3-2 is a graphical illustration of the spectral signatures.

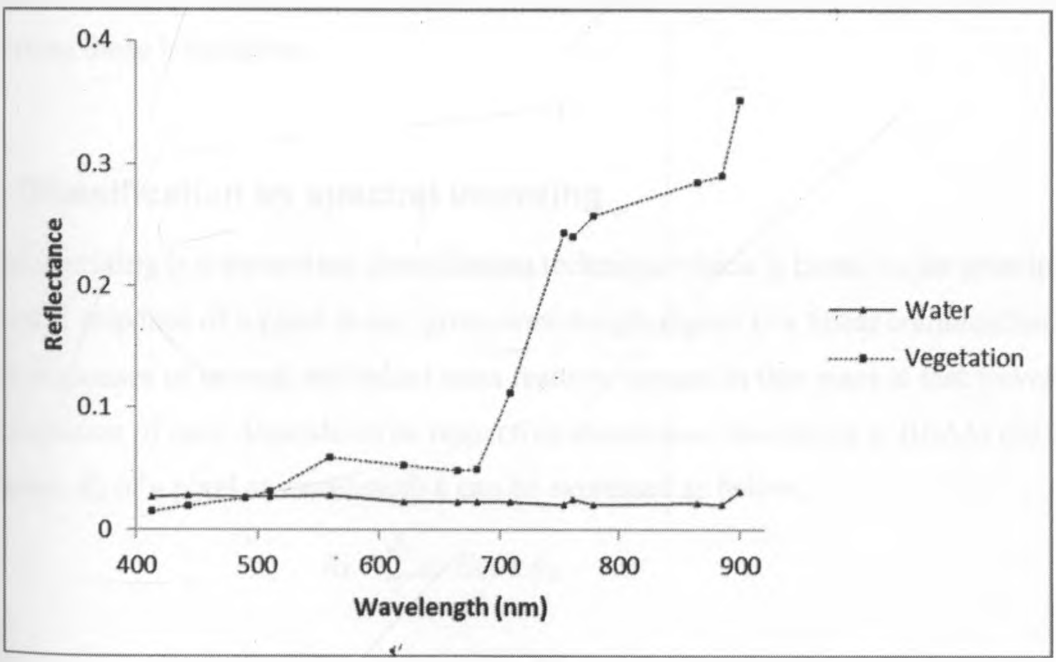

Figure 3-2: A graphical illustration of an endmember spectral library with signature spectra of vegetation and water. Data generated from an image in BEAM

3.3 *Principles of classification*

The spectral properties of a remote sensor can be used to classify images. Different objects (class features) have varying response to electromagnetic radiation at different spectral ranges, leading to unique spectral signatures for each class feature. Ideally, all pixels covering the same class feature would have exactly the same spectral signatures, so that any pixel in an image with that signature would be identified as that very class feature for which the signature represents. Classifying an image in this manner for several class features would end up in a map of classes. In reality, however, a pixel often covers more than one class features so that its resultant spectral response is not exactly the same as that of a pure class feature but rather produce a variety of spectral signatures.

To deal with variability, a pixel's reflectance is represented in an n-dimensional space, so that it occupies a point in that space. This effectively places pixels of each feature class at different points in the n-dimensional space. Here, 'n' is the number of spectral bands. 15-band MERIS images, for instance, have 15 spectral dimensions, and each pixel represents a point in a 15-dimensional space. With variability, the pixels of each feature class now occupy a region, not a point, of n-dimensional space. Vegetation pixels, for example, occupy a different region from that of water in the n-dimensional space. In principle, to classify an image is to delineate boundaries of classes in n-dimensional space and assign class names to pixels using those boundaries.

3.3.1 Classification by spectral unmixing

Spectral unmixing is a supervised classification technique which is based on the principle that the spectral response of a pixel at any given wavelength region is a linear combination of the spectral responses of several individual class features present in that pixel at that wavelength, the contribution of each depends on its respective abundance. According to BEAM (2010) the reflectance, R_k of a pixel at wavelength k can be expressed as below;

$$R_k = \sum_{i}^{n} a_i \cdot E_{i,k} + \varepsilon_k \qquad (1)$$

where

$$\sum_{i}^{n} a_i = 1 \qquad (2)$$

and

$$0 < a_i < 1 \tag{3}$$

$E_{i,k}$ is reflectance of endmember i at wavelength k, a_i is the abundance of endmember i, n is the number of endmembers, and ε_k is the error at wavelength k.

Equations (2) and (3) introduce the constraints that fractions (abundance) sum to one and are non-negative. The system of linear equation shown above can be solved by a least square solution which minimizes the sum of squares of errors. The accuracy of the unmixing is based on ε_k of equation (1), squared and summed over all m channels and is expressed as below

$$RMSE = \sqrt{\left[\sum_{k}^{m} \varepsilon_k^2\right]^{-m}} \tag{4}$$

where m is number of wavelengths in the discrete spectrum.

The spectrum of a pixel can thus be expressed in terms of the individual endmember spectra. With the knowledge of the spectral characteristics of each of the land cover types (called the endmembers), the spectra can thus be classified into its constituent spectra. This analysis results in abundance maps, as many as the defined endmembers.

3.3.2 K-Means clustering

K-means (MacQueen, 1967) is one of the simplest unsupervised classification algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters representing k feature classes) fixed a priori. The main idea is to define k centroids, one for each cluster. The first step is to randomly choose k pixels whose samples define the initial cluster centres. The next step is to assign each pixel to the nearest cluster centre as defined by the Euclidean distance, thus completing the first groupage. Next step is to recalculate the cluster centres as the arithmetic means of all samples from all pixels in a cluster, from which a new binding has to be done between the same data set points and the nearest new centroid. At this point, a loop (iteration) has been generated. As a result of this loop the k centroids change their location step by step. This loop is repeated until the convergence criterion is met. The convergence criterion is met when the specified maximum number of iterations is exceeded or when the cluster centres do not change between two iterations. Finally, this algorithm aims at minimizing an objective function, in this case a squared Euclidean distance function;

$$J = \sum_{j=1}^{K}\sum_{i=1}^{n}\left\|x_i^{(j)} - c_j\right\|^2 \quad (5)$$

which is an indicator of the distance of the n data points from their respective cluster centres, where $\left\|x_i^{(j)} - c_j\right\|^2$ is the distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j.

Figure 3-3 is an illustration of how class means m_1 and m_2 move into the centres of two clusters.

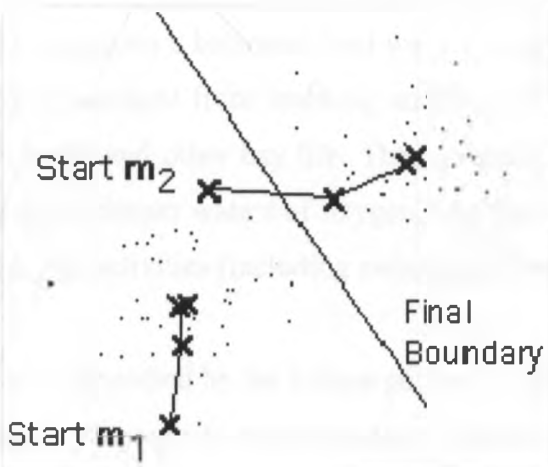

Figure 3-3: An illustration of K-Means clustering. Source: Matteucci (2010)

3.4 Water quality parameters

In this study, two water quality parameters were considered; the concentrations of chlorophyll-a (Chl-a) and the total suspended matter (TSM). The concentrations of these parameters were retrieved from satellite data using algorithms that derive data of their Inherent Optical Properties (IOPs) at 443 nm (MERIS band 2), from which the concentration values are computed. Lake water constituents comprise a large number of different substances, which include mineralic dissolved and particulate compounds, a large variety of organic macromolecules, living organisms such as phytoplankton, zooplankton and bacteria, and their debris and excrements. All of these water constituents have different optical properties concerning scattering and absorption and partly fluorescence (Doerffer and Schiller, 2008 (a)).

3.4.1 Chlorophyll-a (Chl-a)

Chlorophyll *a* is a type of chlorophyll that is most common and predominant in all oxygen-evolving photosynthetic organisms such as higher plants, red and green algae. It highly absorbs electromagnetic radiation in the 400 – 450 nm and 650 – 700 nm wavelength ranges. Its molecular formula is $C_{55}H_{72}O_5N_4Mg$. Chlorophyll is the green pigment that allows plants (including algae) to convert sunlight into organic compounds during photosynthesis. Of the several kinds of chlorophyll, chlorophyll *a* is the predominant type found in algae. High amounts of chlorophyll *a* in the bay's waters are an indicator of nutrient pollution because excess nutrients fuel the growth of algae.

Chlorophyll *a* is often used to measure the amount of algae present in the bay. The bay needs the right amount of algae to maintain a balanced food web. Too much algae can cause large-scale algae blooms that block sunlight from reaching underwater bay grasses, which are an important habitat for fish, crabs and other bay life. They eventually sink to the bottom and decay in a process that depletes deeper waters of oxygen, and they have negative impacts on both underwater life and human activities (including swimming, boating and fishing).

A dissolved substance may be identified by the unique pattern of wavelengths absorbed, since every substance has a unique response to electromagnetic radiation. Chlorophyll in plants absorb strongly in the blue wavelengths (about 450 nm) and red wavelengths (about 650 nm) but reflect in the green wavelengths (about 525 nm), explaining why leaves are green. A plot of absorbance versus visible wavelengths (400 to 700 nm) for a solution of chlorophyll (Figure 3-4) shows two major peaks, one at around 400 nm and one at around 700 nm, and a valley from 500 to 625 nm. This spectrum is characteristic for chlorophyll a for identification.

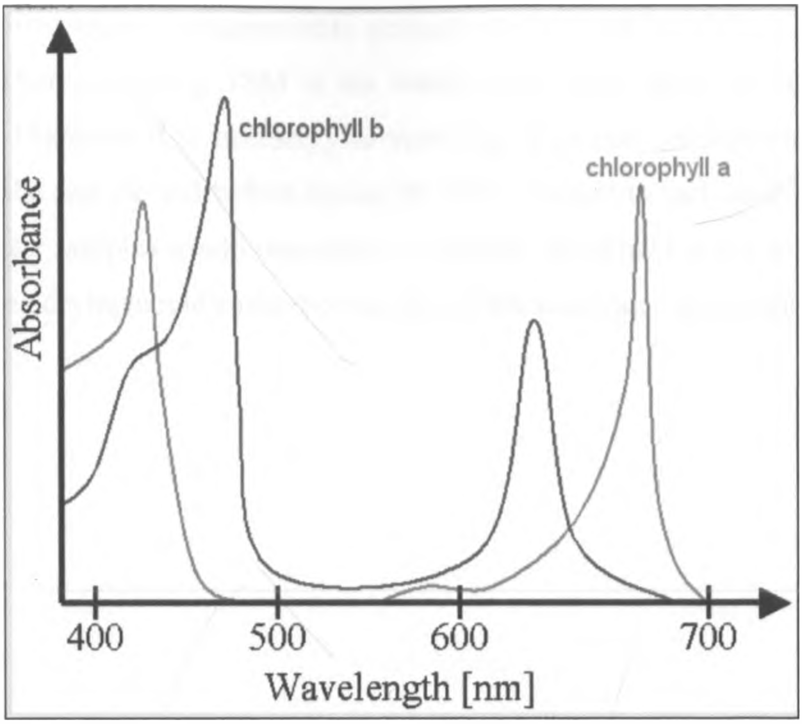

Figure 3-4: A diagram showing the absorbance characteristics of Chlorophyll. It shows two major peaks at around 400 nm and 700 nm. Source: Harrison (2010)

3.4.2 Total suspended matter (TSM)

Total Suspended Matter (TSM), also called Total Suspended Solids (TSS), is a water quality measurement (initially called non-filterable residue (NFR)), which refers to the dry-weight of particles trapped by a filter, typically of a specified pore size. TSM is defined, at the lower end by a cut-off established by properties of the filter being used (pore size) and at the upper end by the exclusion of all particulates too large to be suspended in water. Traditionally a pore size of 0.45 μm was used to define TSM, but nowadays 0.2 μm is used (Doerffer and Schiller, 2008 (a)).

TSM of a water sample is determined by pouring a carefully measured volume of water (typically one litre; but less if the particulate density is high, or as much as two or three litres for very clean water) through a pre-weighed filter of a specified pore size, then weighing the filter again after drying to remove all water. The gain in weight is a dry weight measure of the particulates present in the water sample expressed in units derived from the volume of water filtered (typically milligrams per litre).

If, however, the water contains an appreciable amount of dissolved substances (as certainly would be the case when measuring TSM in sea water), these will add to the weight of the filter as it is dried. Therefore it is necessary to wash the filter and sample with deionized water after filtering the sample and before drying the filter. Failure to include this step when working with sea water samples would completely invalidate the results as the weight of salts left on the filter during drying could easily exceed that of the suspended particulate matter.

4 METHODOLOGY

4.1 Image data requisition and selection criteria

Data used in this study are those taken by satellite borne sensor; Medium Resolution Imaging Spectrometer (MERIS), on board ESA's environmental research satellite, ENVISAT. MERIS data was obtained from ESA in the framework of TIGER Initiative. Table 4-1 below shows the MERIS product specifications.

Table 4-1: MERIS product specifications. Source: Sotis (2007)

Channel Number	Centre Wavelength (nm)	Bandwidth (nm)	Application
1	412.5	10	Yellow substance and detrital pigments
2	442.5	10	Chlorophyll absorption maximum
3	490	10	Chlorophyll and other pigments
4	510	10	Suspended sediment, red tides
5	560	10	Chlorophyll absorption minimum
6	620	10	Suspended sediment
7	665	10	Chlorophyll absorption and fluorescence reference
8	681.25	7.5	Chlorophyll fluorescence peak
9	708.75	10	Fluorescence reference, atmospheric corrections
10	753.75	7.5	Vegetation, cloud
11	760.625	3.75	Oxygen absorption R-branch
12	778.75	15	Atmosphere corrections
13	865	20	Vegetation, water vapour reference
14	885	10	Atmosphere corrections
15	900	10	Water vapour, land

The level one MERIS Full Resolution products (MER_FR_1P) are geocoded with calibrated TOA radiance and spatial resolution which varies in the across track direction, between 0.26 km at nadir and 0.39 km at swath extremities. Along-track sampling is close to 0.29 km (ESA, 2010). Its spectral resolution includes visible and NIR bands from 400 nm to 900 nm, and has a revisit time of three days. It also has the resampled ECMWF data: mean sea level pressure, total column ozone, total column water vapour and wind speed.

Images were selected on the basis of acquisition time, image area coverage, spatial and spectral resolution and the severity of cloud cover. Figure 4-1 is a flow chart which shows the criterion for selection of image data for the study.

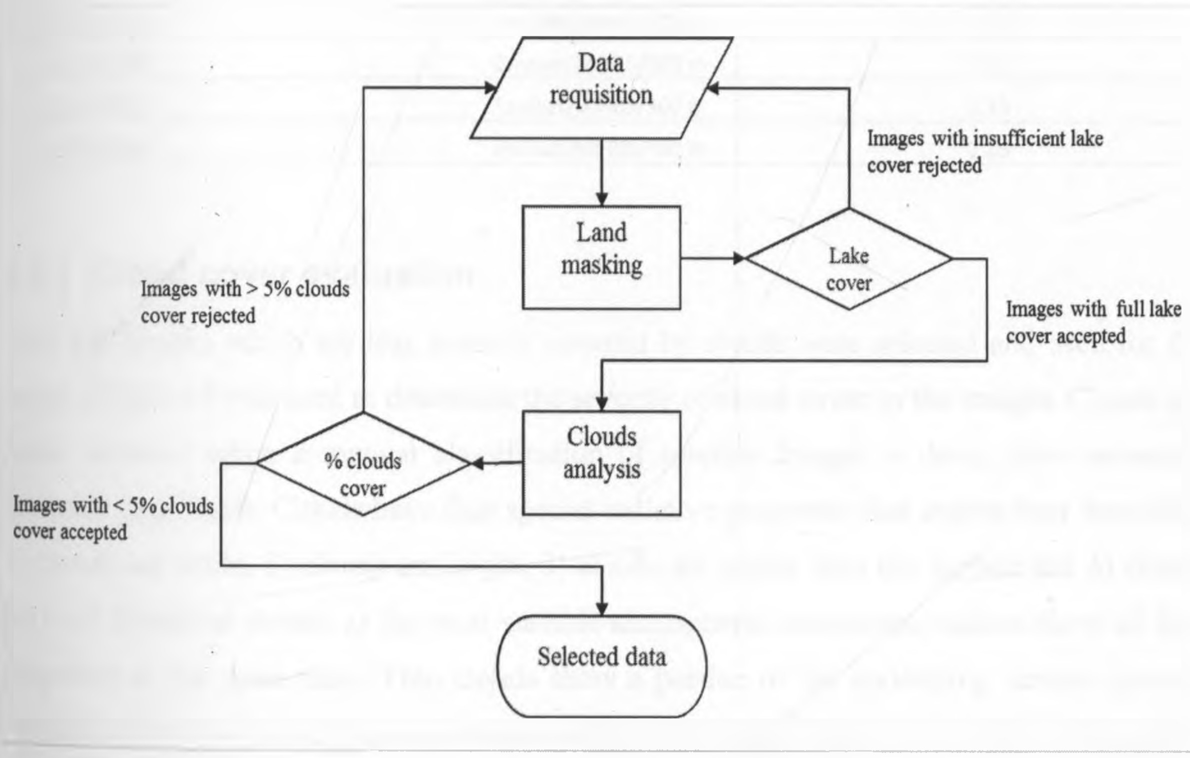

Figure 4-1: A flow chart showing the procedure for image data selection

Of the 174 images that were ordered and received from ESA, 31 were rejected due to insufficient lake coverage. Another 128 images were rejected since they fell beyond the set percentage cloud cover threshold of 5% of the image, leaving only 15 images for analysis. For water quality assessment, however, 93 images of Winam Gulf spread over the study period were used. Table 4-2 shows the image dates and product specifications for data used to study the whole lake.

Table 4-2: Image acquisition dates and product specifications for data used. Source: ESA (2010)

Acquisition date	Platform/Sensor/Resolution	Percentage cloud cover
26/12/2003	Envisat/MERIS/300 m	1.10
21/02/2005	Envisat/MERIS/300 m	0.15
17/07/2005	Envisat/MERIS/300 m	0.67
02/01/2006	Envisat/MERIS/300 m	2.61
06/02/2006	Envisat/MERIS/300 m	0.09
16/08/2006	Envisat/MERIS/300 m	1.51
12/10/2006	Envisat/MERIS/300 m	4.81
20/02/2007	Envisat/MERIS/300 m	0.03
27/09/2008	Envisat/MERIS/300 m	2.43
14/02/2009	Envisat/MERIS/300 m	0.92
21/03/2009	Envisat/MERIS/300 m	2.86
12/06/2009	Envisat/MERIS/300 m	1.27
02/08/2009	Envisat/MERIS/300 m	1.63
28/09/2009	Envisat/MERIS/300 m	2.57
12/02/2010	Envisat/MERIS/300 m	1.15
03/08/2010	Envisat/MERIS/300 m	2.85

4.1.1 Cloud cover evaluation

Only the images which are less severely covered by clouds were selected and used for the study. *BEAM 4.8* was used to determine the severity of cloud cover in the images. Clouds are easily detected when a manual classification of satellite images is done, their automatic detection is difficult. Clouds have four special radiative properties that enable their detection: 1) clouds are white, 2) clouds are bright, 3) clouds are higher than the surface and 4) clouds are cold. However clouds, as the most variable atmospheric constituent, seldom show all four properties at the same time. Thin clouds show a portion of the underlying surface spectral properties, and low clouds are sometimes quite warm. Additionally some surface types, like snow and ice have spectral properties that are very similar to some of the cloud properties (BEAM, 2010).

Considering Lake Victoria, where water and vegetation are the only dominant features (for there is no land or built-up area in the lake, except of course, the islands and boats), the 'brightness' attribute of clouds alone is sufficient to accurately identify clouds in the image. This attribute was used to estimate the severity of clouds cover in the image. A percentage threshold value of bright (cloud-covered) pixels over the whole image was set, to select the images to be used for the study. Only images with less than 5% cloud cover were selected.

4.2 *Image processing*

Image processing procedures involved image pre-processing, image classification and extraction of results by applying the ROI (Region of Interest) to limit the results to just within the lake. Analysis and presentation of results involved water quality analysis, mapping and regression analysis. These procedures are presented in a flow chart in Figure 4-2.

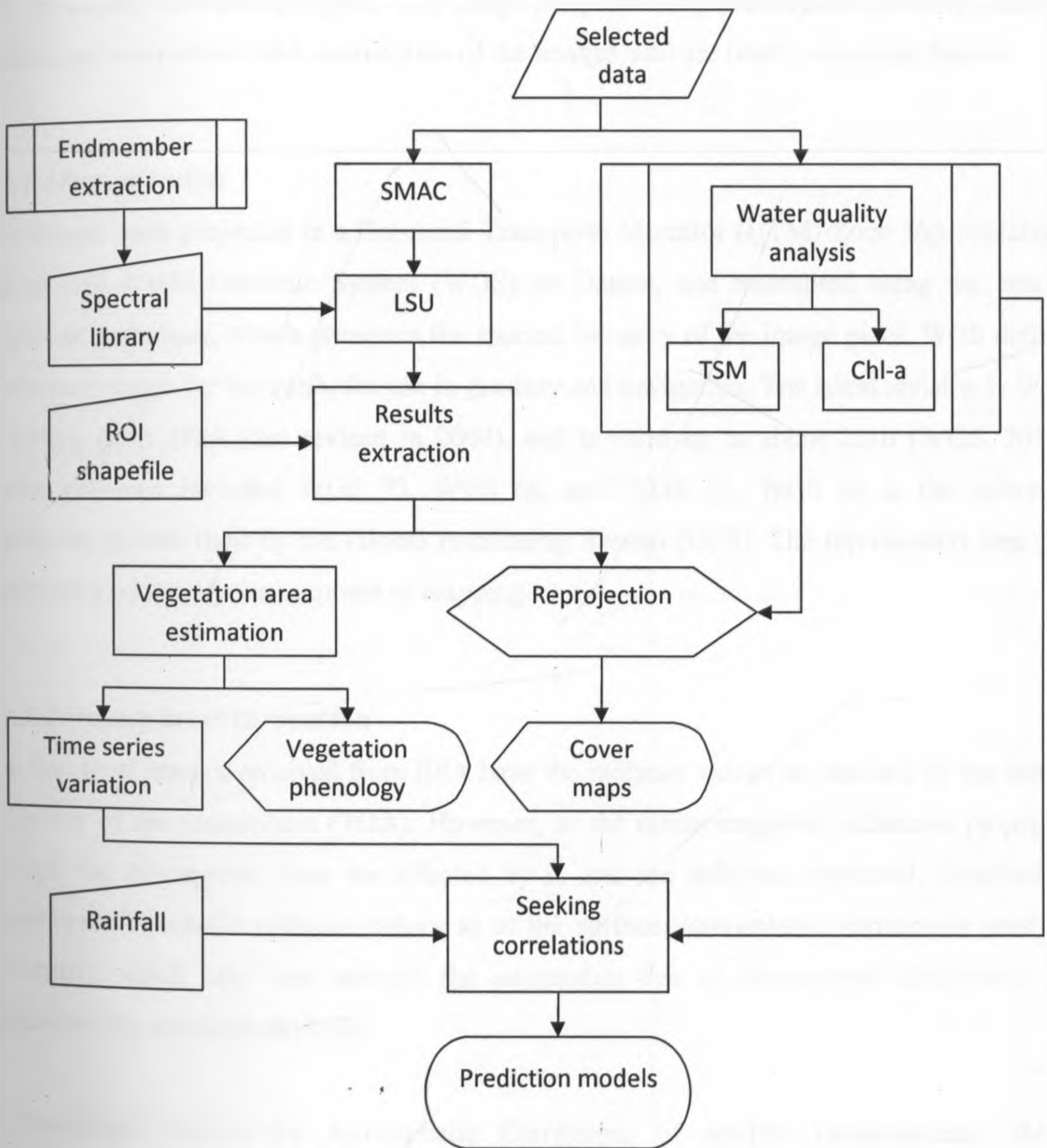

Figure 4-2: A flow chart showing image processing and analysis procedures

4.2.1 Image pre-processing

Due to the sensitive nature of spectral studies, pre-processing of satellite images prior to vegetation extraction is essential to remove atmospheric effects and increase the interpretability of image data (Idawo *et al.* 2004). The acquired images were processed using *BEAM 4.8* and *ENVI 4.2* image processing and analysis software. Before any processing was done, the images were first resized (creation of a spatial subset of the area of interest). The spatial subset is more convenient because it takes lesser processing time than the full scene and, of course, less storage space. The image pre-processing procedures performed include atmospheric corrections and reprojection of the images, and are briefly discussed below.

4.2.1.1Reprojection

The images were projected in a Universal Transverse Mercator (UTM) Zone 36S coordinate system and World Geodetic System (WGS) 84 Datum, and resampled using the nearest neighbour technique, which preserves the spectral integrity of the image pixel. WGS defines a reference frame for the earth, for use in geodesy and navigation. The latest revision is WGS 84 dating from 1984 (last revised in 2004), and is valid up to about 2010 (WGS, 2011). Earlier schemes included WGS 72, WGS 66, and WGS 60. WGS 84 is the reference coordinate system used by the Global Positioning System (GPS). The reprojection step was particularly necessary for purposes of mapping.

4.2.1.2Atmospheric correction

The first level images received from ESA have the radiance values as detected by the sensor at the top of the atmosphere (TOA). However, as the electromagnetic radiations propagate through the atmosphere, they are affected by it, and are reflected, refracted, absorbed or transmitted. To obtain radiance values as at the surface, atmospheric corrections must be performed, which take into account the attenuation due to atmospheric absorption and radiance of the scattered skylight.

A Simplified Method for Atmospheric Corrections of satellite measurements *SMAC Processor 1.5.203* (Rahman *et al.* 1994) incorporated in BEAM was used to perform atmospheric corrections on the images. It is a semi-empirical approximation of the radiative transfer in the atmosphere. The signal at the satellite is written as the sum of the following components, which are then expressed in simple analytical terms: Atmospheric spherical albedo, Total atmospheric transmission, Rayleigh scattering and Aerosol scattering.

Rayleigh scattering is the elastic scattering of the electromagnetic radiation by particles much smaller than the wavelength of the radiation, which may be individual atoms or molecules. Rayleigh scattering can be defined as scattering in the small size parameter regime;

$$\frac{2\pi r}{\lambda} << 1 \qquad (6)$$

where r is the characteristic dimension of the particle and λ the radiation wavelength.

With SMAC technique the radiative transfer in the atmosphere can be computed much faster than with a full model. A comparison has shown that the gain in computation time is several hundred times in comparison with the full model (BEAM, 2010). Because of its "speed", this method is best suited for application to large data volumes.

The SMAC requires as input, in addition to the measured top of atmosphere radiances, the surface pressure, the ozone content and the water vapour content, and, most importantly, the aerosols. Continental aerosol model was selected, with the default aerosol optical depth of 0.2 at 550 nm, while ECMWF meteorological data for pressure, ozone and humidity was used. ECMWF data files contain meteorological information for each pixel. To mask out clouds in the image, the code below that is in-built in BEAM was used:

(l1_flags.LAND_OCEAN or water) and not (l1_flags.INVALID or l1_flags.BRIGHT)

This masks out the invalid and bright (cloud) pixels, while retaining water masks in the image.

4.2.2 Classification

4.2.2.1 Creating a spectral library (an endmember file)

An endmember file is a compiled list of spectral signatures, with a signal file for every informational class containing spectral characteristics of the cover classes they describe. Endmembers can be derived from the image (image derived endmembers), or measured in the field using a field spectroradiometer (field derived endmembers). Image derived endmembers were used in this study, as opposed to field derived endmembers because they are in the same state as the image from which they are derived, so that the effects of atmospheric distortions and inaccuracies in the atmospheric corrections are minimized. Proper definition of endmember file is crucial when using spectral unmixing classification technique since the classification results are greatly determined by the input signatures. Much attention was given to the derivation of the endmember file that was used for classification.

A field study was conducted on 14 December 2010. During the field study, the geographic coordinates of the location in the lake that was covered by vegetation was taken using a GPS receiver. Water hyacinth, which is the predominant aquatic vegetation species in the lake, is free-floating and is therefore highly dynamic. It is easily carried away by tides and wind. On the date of the study, however, there was a huge piece of the hyacinth mat, with outgrowth of other aquatic vegetation, which had been trapped by the shore of the lake along the Rakwaro and Kisiege beaches for close to three months according to the local residents. This was an assurance that any image acquired within this period would contain the desired class feature static in the same location, so that a one day field work was sufficient to collect the geographic coordinates of the identified training site. It was about a kilometre into the lake, and at least a kilometre along the shore. This was large enough to be represented by several pixels of a MERIS FR image with spatial resolution of about 300 m.

A satellite image of 15 December 2010 was used to extract the vegetation signatures. Eight pixels were selected, at one pixel accuracy, at different points within the vegetation mat, and their respective signatures extracted, from which the mean signature of the vegetation was computed.

Four major 'water species' were visually identified from the resulting image of an unsupervised classification (see section 5.1). Eight pixels were selected at one pixel accuracy from each of these four regions, from which the spectral signatures were extracted, and an average spectrum computed. The signatures of the four water classes together with the extracted vegetation signature were then compiled into a spectral library, called the endmember file, which was then used to classify the image.

4.2.2.2 Classification of the Image

In this stage of classification, spectral unmixing technique was applied to derive the abundances of the feature classes specified in the endmember spectral library. The library defined earlier (section 4.2.2.1) was used as an input spectral data to *BEAM 4.8* spectral unmixing program which applied the unmixing algorithm described in equation (1) to derive the abundance values and equation (4) to derive the RMSE values of each pixel. These abundance values were then displayed in cover maps.

4.2.3 Shapefile (Shoreline)

Digitization of the lake was done using *ENVI 4.5*. The lake shapefile was necessary for extracting the classification results from just within the lake area. The accuracy of area estimation is subject to the accuracy of the shapefile (or Region of Interest, ROI). A shapefile was created by digitizing an RGB of a selected clouds free MERIS FR imagery of February 20, 2007. *Google maps* were used to identify the islands and to distinguish them from the floating mats of vegetation.

5 RESULTS, ANALYSIS AND DISCUSSION

5.1 *Classification results*

5.1.1 Endmember extraction

Unsupervised classification of a selected clouds-free image using K-Means clustering produced four major classes of water, two in the main lake and two in the Winam Gulf. These are the predominant water 'species' in the lake. The number of endmembers was determined visually from the classified image by considering only the major class features within the lake. Figure 5-1 below shows the results of this classification process. The reflectance spectra of each of these feature classes were then extracted.

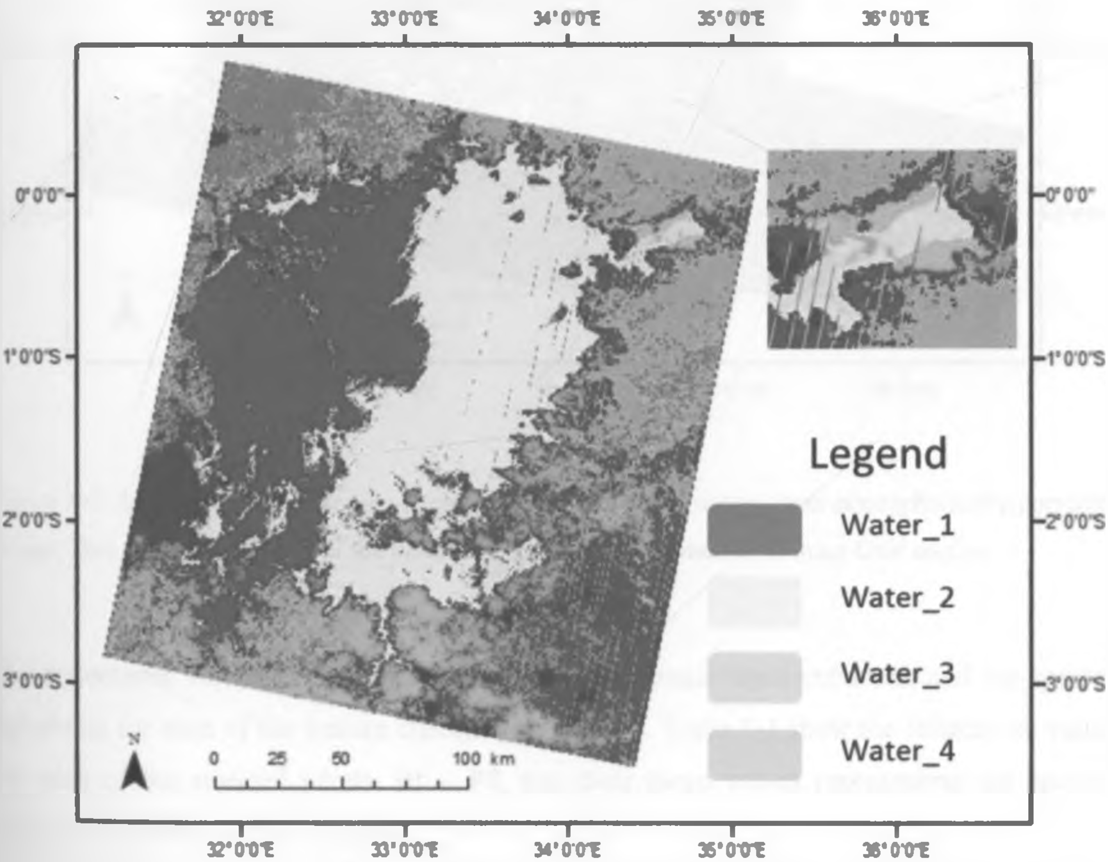

Figure 5-1: Results of unsupervised classification on a Lake Victoria image using K-Means clustering. There are four major classes of water, two in the main lake and two in Winam Gulf. Inset is the Winam Gulf section of the lake

Eight pixels were selected from each of the identified feature classes, as shown in Figure 5-2 below, with one pixel accuracy. The reflectance spectra of these pixels were extracted from which their average class spectra were computed.

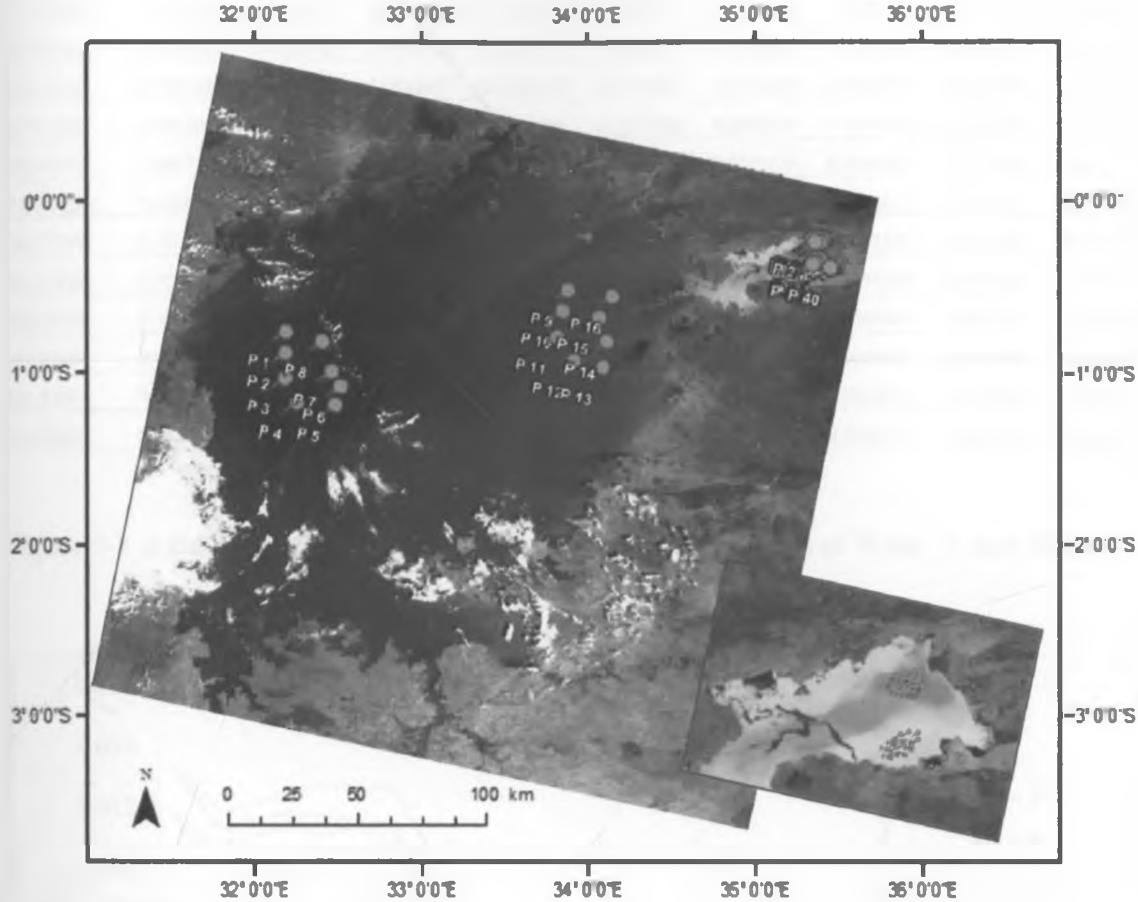

Figure 5-2: Identification of training sites and endmember extraction from atmospherically corrected image. This is an RGB image of the main lake. Inset is the zoomed-in Winam Gulf section

The reflectance values at each of the fifteen MERIS bands were extracted, and the spectral signatures for each of the feature classes were derived. Table 5-1 show the reflectance values for each of the selected pixels, P1 – P8, and their mean values representing the spectral response of Water_1 class feature.

Table 5-1: Water_1 individual pixel reflectance values

Wavelength	P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	Water_1
412.69101	0.0132147	0.0131077	0.0115945	0.0098727	0.0134366	0.0104973	0.0114912	0.0113837	0.0118248
442.55902	0.0112426	0.0112499	0.0095101	0.0079339	0.0118167	0.0098039	0.009737	0.0108213	0.0102644
489.88202	0.0118186	0.0117735	0.0098772	0.0082228	0.0124392	0.0096512	0.0097625	0.0112612	0.0106008
509.81903	0.0110075	0.0117452	0.0095283	0.0085097	0.0126787	0.0096517	0.0095362	0.0115427	0.010525
559.69403	0.0115682	0.0113118	0.0094022	0.0083787	0.0126597	0.0091801	0.0093778	0.0108761	0.0103443
619.60101	0.0061982	0.0060046	0.0043947	0.0029848	0.0078289	0.0042684	0.0052345	0.0051692	0.0052604
664.57306	0.0052098	0.005579	0.0040816	0.0025985	0.0073935	0.0042993	0.0044359	0.0049591	0.0048196
680.82104	0.0057271	0.0060234	0.004552	0.0033108	0.0079726	0.0045227	0.0049812	0.0051281	0.0052772
708.32904	0.0056269	0.00582	0.0044304	0.0029433	0.0075176	0.004107	0.0045327	0.005123	0.0050126
753.37103	0.0032821	0.0034904	0.0022164	5.21E-04	0.0051692	0.0018934	0.0022585	0.0030032	0.0027292
761.50806	0.0079266	0.0069151	0.0056907	0.0057392	0.0076806	0.0055601	0.0056548	0.0069868	0.0065193
778.40906	0.0034804	0.0039096	0.0024506	8.43E-04	0.0054024	0.0019897	0.0024324	0.0031251	0.0029541
864.87604	0.0037468	0.0042072	0.0029839	0.00116	0.0059631	0.0023072	0.0027567	0.0035905	0.0033394
884.94403	0.0035363	0.0045765	0.0027927	0.0011169	0.005691	0.0024673	0.0033025	0.0037889	0.003409
900.00006	0.0114656	0.0117771	0.010296	0.0072081	0.0139256	0.0085532	0.0100642	0.010676	0.0104957

Figure 5-3 is the graphical representation of the spectral signatures of Water_1 class feature.

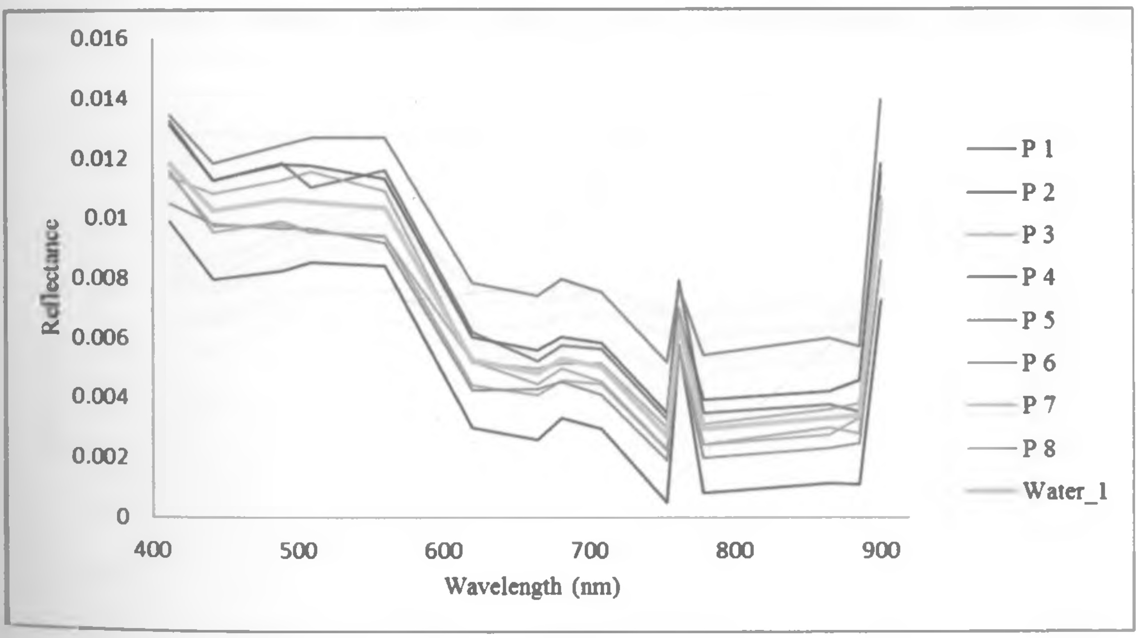

Figure 5-3: Individual pixel spectra and the resultant mean signature for Water_1 class feature

Table 5-2 shows the reflectance values for each of the selected pixels, P9 – P16, and their mean values representing the spectral response of Water_2 class feature. . Figure 5-4 is the graphical representation of these spectral signatures.

Table 5-2: Water_2 individual pixel reflectance values

Wavelength	P 9	P 10	P 11	P 12	P 13	P 14	P 15	P 16	Water_2
412.69101	0.0243727	0.0266212	0.0268793	0.0298173	0.0313346	0.0321359	0.0302128	0.0305414	0.0289894
442.55902	0.0240202	0.0273428	0.0265645	0.0296117	0.0321286	0.0311687	0.0302522	0.030369	0.0289322
489.88202	0.0265538	0.0288732	0.0283504	0.0314787	0.0338956	0.0343442	0.0330003	0.0332891	0.0312232
509.81903	0.0266608	0.0293837	0.0289332	0.0328336	0.0340626	0.0353743	0.0347019	0.0342653	0.0320269
559.69403	0.0267268	0.0289233	0.0291494	0.0326637	0.0346372	0.0361399	0.036856	0.035934	0.0326288
619.60101	0.0232687	0.0259238	0.0243093	0.0268864	0.0290271	0.0281522	0.0283777	0.0280572	0.0267503
664.57306	0.0236243	0.0261383	0.0244321	0.0267386	0.0290998	0.0281908	0.0278303	0.0281029	0.0267696
680.82104	0.0241981	0.0274235	0.0249535	0.0270515	0.0297058	0.0288658	0.0285973	0.0282144	0.0273762
708.32904	0.0243703	0.02728	0.0249482	0.0267082	0.0291695	0.0289158	0.028249	0.0291342	0.0273469
753.37103	0.0220989	0.0252065	0.0232547	0.0251448	0.0276041	0.0273315	0.0265532	0.0267768	0.0254963
761.50806	0.0205635	0.024231	0.0231232	0.0245203	0.0266459	0.0256253	0.0245948	0.0257685	0.0243841
778.40906	0.0224168	0.0251573	0.0233353	0.025145	0.0280015	0.0274706	0.0267794	0.0274191	0.0257156
864.87604	0.0231601	0.0260934	0.0238343	0.0257701	0.0285687	0.0278386	0.0272371	0.0280447	0.0263184
884.94403	0.023401	0.0258219	0.0240693	0.0255282	0.0286252	0.0282738	0.0278562	0.0280739	0.0264562
900.00006	0.033443	0.0358801	0.0328752	0.034818	0.037356	0.0372209	0.0365304	0.0373102	0.0356792

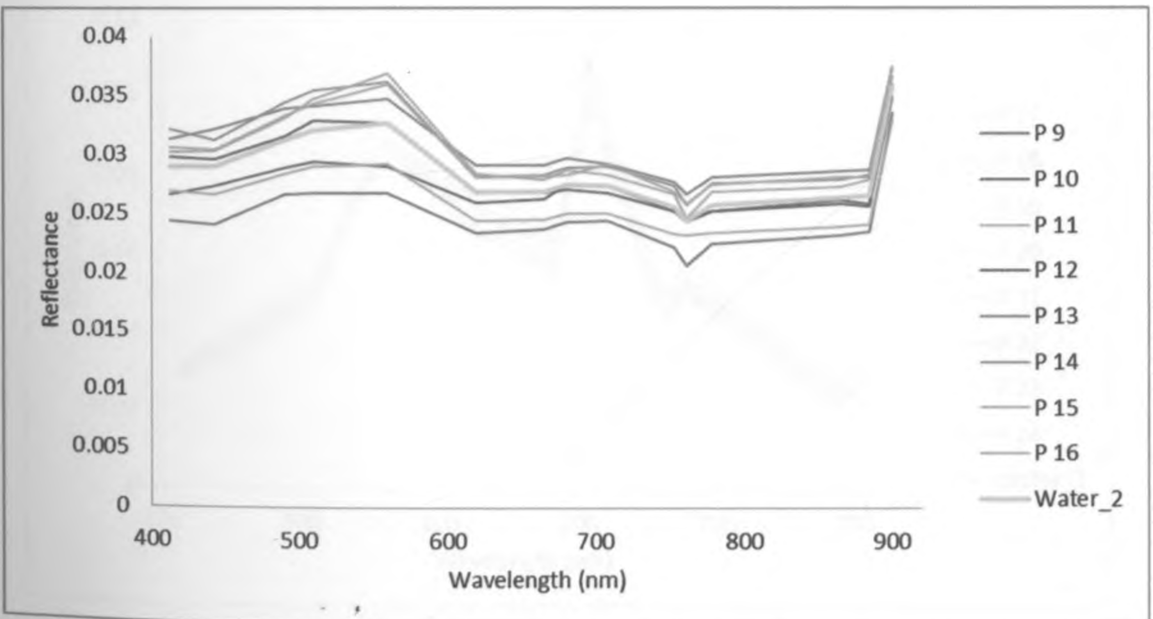

Figure 5-4: Individual pixel spectra and the resultant mean signature for Water_2 class feature

Table 5-3 shows the reflectance values for each of the selected pixels, P17 – P24, and their mean values representing the spectral response of Water_3 class feature. Figure 5-5 is the graphical representation of these spectral signatures.

Table 5-3: Water_3 individual pixel reflectance values

Wavelength	P 17	P 18	P 19	P 20	P 21	P 22	P 23	P 24	Water_3
412.69101	0.0300386	0.0312624	0.0315721	0.0300362	0.0298421	0.0287266	0.028151	0.0283671	0.0297495
442.55902	0.0357173	0.0379978	0.0384861	0.0370489	0.0362328	0.0332536	0.0328894	0.0325082	0.0355168
489.88202	0.0435126	0.0472364	0.047952	0.0463803	0.0446868	0.0406672	0.0420054	0.0403729	0.0441017
509.81903	0.0503115	0.0538788	0.0536548	0.0524878	0.0512895	0.047474	0.0481902	0.047108	0.0505493
559.69403	0.0893349	0.0896333	0.0895634	0.088659	0.087437	0.0872581	0.0877237	0.087807	0.0884271
619.60101	0.0702975	0.0754815	0.07615	0.0726866	0.0702949	0.0658416	0.0681607	0.0674581	0.0707964
664.57306	0.0642341	0.0700786	0.0703213	0.0674265	0.0645478	0.059499	0.0617974	0.0614302	0.0649169
680.82104	0.0575376	0.0637765	0.0639684	0.0607492	0.0579054	0.0523519	0.0548537	0.0539938	0.0581421
708.32904	0.1079724	0.1011698	0.0997985	0.1028432	0.1037564	0.1003607	0.0974122	0.1003639	0.1017096
753.37103	0.0565482	0.0492523	0.0479834	0.0510532	0.0523277	0.050427	0.0486443	0.0533943	0.0512038
761.50806	0.0497797	0.0437765	0.0436942	0.0450456	0.0477486	0.0453449	0.0428565	0.0478252	0.0457589
778.40906	0.0570671	0.0499543	0.0489064	0.0518126	0.0532388	0.0508252	0.0494702	0.0536565	0.0518664
864.87604	0.0318238	0.0265805	0.0259623	0.0280554	0.0289995	0.0280341	0.0271924	0.0307973	0.0284307
884.94403	0.0257824	0.0213901	0.0208523	0.0228359	0.023491	0.0230967	0.0224935	0.0257446	0.0232108
900.00006	0.0294838	0.0253881	0.0238331	0.0260374	0.026588	0.0260254	0.0256829	0.0305595	0.0266998

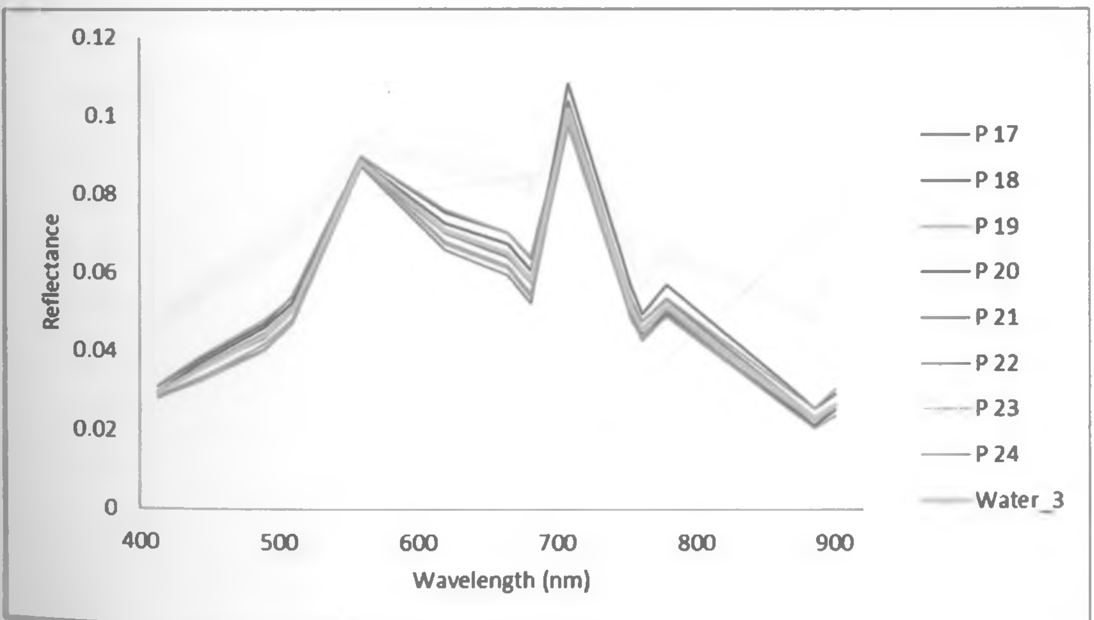

Figure 5-5: Individual pixel spectra and the resultant mean signature for Water_3 class feature

Table 5-4 shows the reflectance values for each of the selected pixels, P25 – P32, and their mean values representing the spectral response of Water_4 class feature. Figure 5-6 is the graphical representation of these spectral signatures.

Table 5-4: Water_4 individual pixel reflectance values

Wavelength	P 25	P 26	P 27	P 28	P 29	P 30	P 31	P 32	Water_4
412.69101	0.074211	0.0744292	0.0757986	0.076804	0.0725293	0.0721061	0.0695293	0.072502	0.0734887
442.55902	0.0884856	0.0874122	0.0891227	0.0923776	0.0856668	0.0850469	0.0831646	0.0860611	0.0871672
489.88202	0.1031536	0.1009493	0.1040275	0.107682	0.1015629	0.0995267	0.0969754	0.1006506	0.101816
509.81903	0.1099984	0.108487	0.1113109	0.1143201	0.1079683	0.1058544	0.1043681	0.1077389	0.1087557
559.69403	0.1424647	0.1410831	0.1429346	0.1451878	0.1400265	0.1374835	0.1385817	0.1396073	0.1409212
619.60101	0.1376035	0.1297226	0.1365103	0.145317	0.1326638	0.1315158	0.1319899	0.1356903	0.1351267
664.57306	0.1344816	0.1267449	0.1334981	0.1419264	0.1285543	0.1280461	0.1313267	0.1318098	0.1320485
680.82104	0.1294405	0.121694	0.1290906	0.1392061	0.1240235	0.1236933	0.1264054	0.1275167	0.1276338
708.32904	0.150462	0.1520966	0.1486496	0.146124	0.1485117	0.1486851	0.160218	0.1471298	0.1502346
753.37103	0.1004998	0.1041102	0.0996776	0.0973807	0.1010444	0.1017719	0.111315	0.0983738	0.1017717
761.50806	0.0866508	0.0902165	0.0885523	0.0860783	0.0887874	0.0891405	0.0972812	0.086621	0.089166
778.40906	0.1007301	0.1047428	0.1009508	0.0980732	0.1016844	0.1019873	0.1125882	0.0987516	0.1024385
864.87604	0.0806759	0.0830423	0.0805422	0.0791047	0.081217	0.0811328	0.0876322	0.0788719	0.0815274
884.94403	0.0762721	0.0777449	0.0755686	0.074841	0.0762881	0.0761093	0.0806349	0.0736224	0.0763852
900.00006	0.0910112	0.0921648	0.0896294	0.0886492	0.0911008	0.0907741	0.0951034	0.0874637	0.0907371

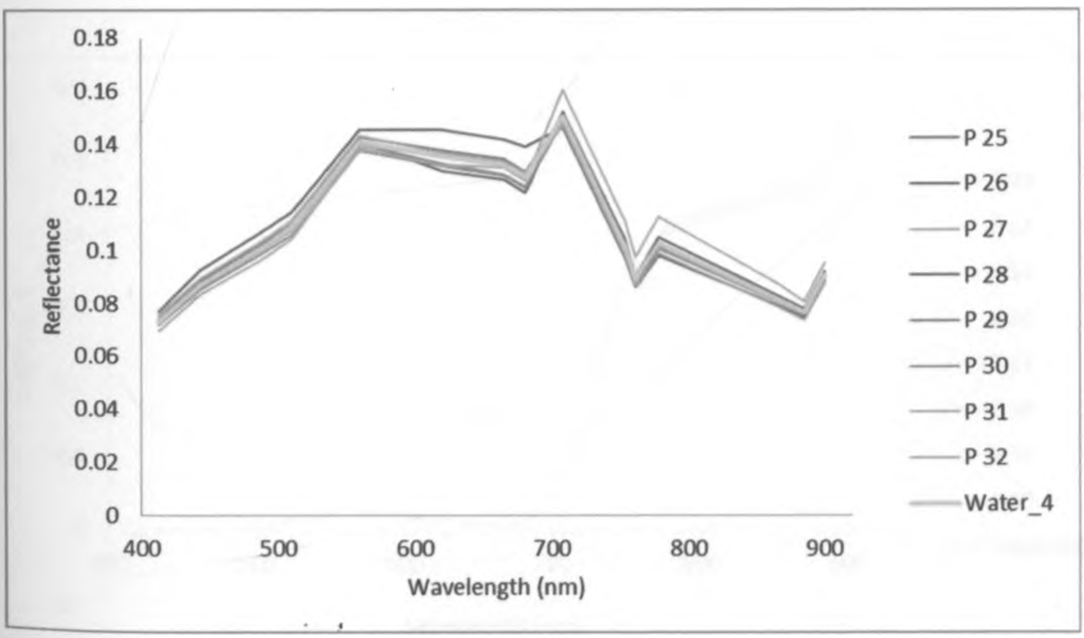

Figure 5-6: Individual pixel spectra and the resultant mean signature for Water_4 class feature

Similarly, the vegetation spectra were extracted, and are presented below. Table 5-5 shows the reflectance values for each of the selected pixels, P33 – P40, and their mean values representing the spectral response of vegetation class feature. Figure 5-7 shows the spectral signatures of each of these pixels and the mean spectra for vegetation.

Table 5-5: Vegetation individual pixel reflectance values

Wavelength	P 33	P 34	P 35	P 36	P 37	P 38	P 39	P 40	Vegetation
412.69101	-0.001959	-0.001945	-0.001392	-0.001347	-0.001595	-0.001767	-0.002162	-0.002302	-0.001809
442.55902	0.0021243	0.0017102	0.0013263	0.0012751	0.0010141	0.0014587	0.0014798	0.0014399	0.0014786
489.88202	0.0031707	0.003586	0.003591	0.0036546	0.0031446	0.0036964	0.0036329	0.0034209	0.0034871
509.81903	0.0080715	0.0079421	0.0072539	0.0074173	0.0074076	0.0078032	0.0079592	0.0080065	0.0077326
559.69403	0.0306687	0.0305968	0.031061	0.0304532	0.0306915	0.0307146	0.0308109	0.0308799	0.0307346
619.60101	0.0148636	0.0157221	0.0154964	0.0147963	0.0150411	0.0157639	0.0155378	0.0159636	0.0153981
664.57306	0.0082889	0.0089658	0.0086559	0.0081525	0.0083625	0.0088852	0.0088082	0.0090759	0.0086494
680.82104	0.0085411	0.008761	0.0090009	0.0082192	0.0082567	0.0088782	0.0087785	0.0090759	0.0086889
708.32904	0.0753904	0.0756991	0.0762407	0.0751903	0.0754019	0.0757727	0.0755418	0.0756348	0.075609
753.37103	0.3670259	0.3689279	0.3716893	0.3774043	0.370991	0.367342	0.3666497	0.3639067	0.3692421
761.50806	0.3327982	0.3361682	0.3349264	0.3444608	0.3423674	0.3316746	0.3350875	0.3261114	0.3354493
778.40906	0.4044282	0.407468	0.412389	0.4178461	0.4113943	0.4064367	0.4051486	0.4003839	0.4081868
864.87604	0.4583679	0.4611595	0.4660071	0.4708236	0.4641501	0.4591731	0.4574814	0.4498461	0.4608761
884.94403	0.4642478	0.4667602	0.4717672	0.4764155	0.469423	0.4647206	0.4636599	0.4560123	0.4666258
900.00006	0.5564654	0.5585285	0.5624438	0.5674535	0.5602568	0.5517254	0.5537859	0.5419457	0.5565756

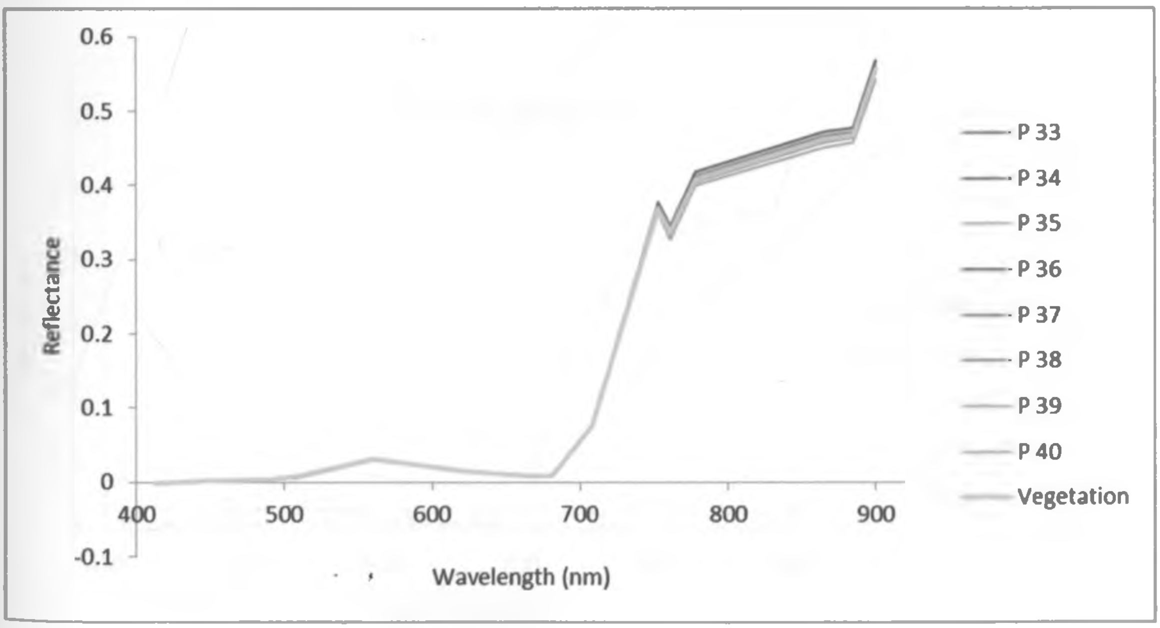

Figure 5-7: Individual pixel spectra and the resultant mean signature for Vegetation class feature

The reflectance values for all the four water classes together with that of vegetation at each of the fifteen MERIS bands were the combined to form an endmember spectral library shown in Table 5-6. Figure 5-8 shows the final endmember spectral signatures.

Table 5-6: Image derived endmember spectral library consisting of vegetation and various water classes

Wavelength (nm)	Vegetation	Water_1	Water_2	Water_3	Water_4
412.6910095	-0.00181	0.011824797	0.028989399	0.029749512	0.073488681
442.559021	0.001479	0.010264428	0.028932207	0.03551676	0.087167184
489.882019	0.003487	0.010600774	0.031223177	0.044101706	0.101815986
509.8190308	0.007733	0.010524974	0.032026923	0.050549326	0.108755748
559.6940308	0.030735	0.010344325	0.032628782	0.088427059	0.140921159
619.6010132	0.015398	0.005260422	0.026750303	0.070796357	0.13512666
664.5730591	0.008649	0.004819596	0.02676962	0.064916859	0.132048499
680.8210449	0.008689	0.005277234	0.027376227	0.058142072	0.127633767
708.3290405	0.075609	0.005012609	0.02734691	0.101709615	0.150234602
753.3710327	0.369242	0.002729237	0.02549631	0.051203812	0.101771683
761.5080566	0.335449	0.006519258	0.024384077	0.04575891	0.089166028
778.4090576	0.408187	0.002954085	0.025715631	0.0518664	0.10243854
864.8760376	0.460876	0.003339408	0.026318366	0.028430654	0.081527385
884.9440308	0.466626	0.003409013	0.026456181	0.023210826	0.076385161
900.000061	0.556576	0.01049573	0.035679223	0.026699776	0.090737084

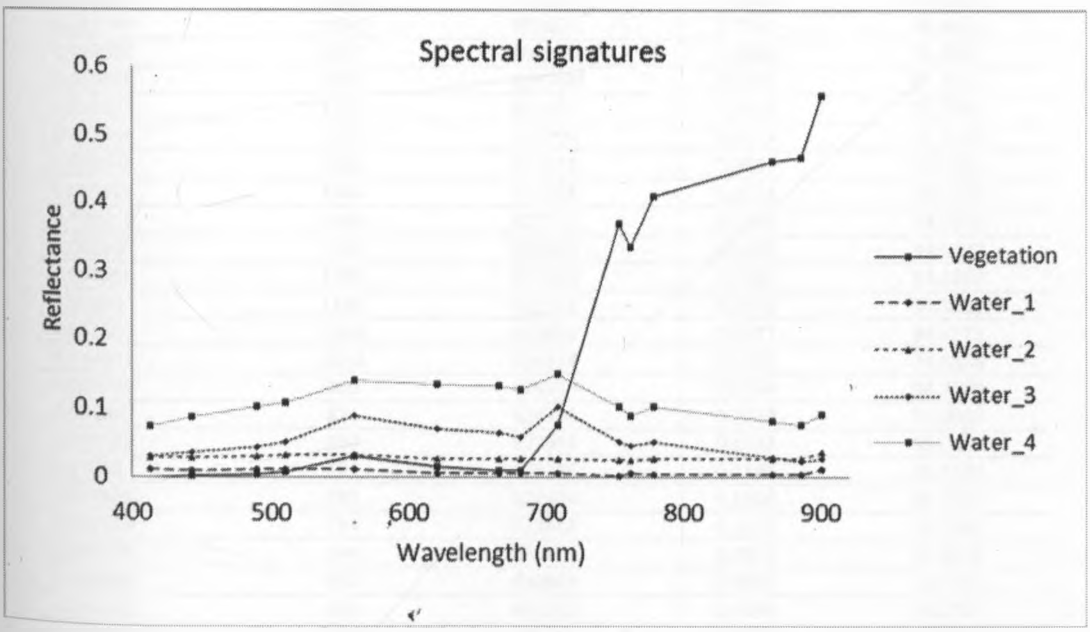

Figure 5-8: The image derived endmember spectral library consisting of five spectral signatures, four for various water classes and one for vegetation class feature

Spectral response of water in the lake was found to be varying spatially, possibly according to the concentrations of dissolved or suspended matter in it, indicating the extent of nutrient enrichment in the lake. Purer water in the main lake displayed low reflectance values, while that near the shores and especially at the almost enclosed Winam Gulf had generally higher reflectance values.

5.1.2 Linear spectral unmixing classification results

The resulting images were the abundance maps of each of the classes defined in the endmember file, an error map for each of the 15 MERIS bands, and a summary error band. The shapefile (shoreline) of the region of interest (ROI) described in *section 4.2.3* was overlaid on the images using ENVI 4.2 and a statistics file of the pixel values from these images were generated. Table 5-7 shows a statistics file of the vegetation abundance map for the 15 December 2010 image.

Table 5-7: A statistical summary of LSU classification results showing pixel values for 15-12-2010 image

Pixel value (Vegetation abundance)	Number of pixels	Total	Percentage	Accumulated Percentage
0	410823	410823	62.2354	62.2354
0.003922	123420	534243	18.6969	80.9323
0.007843	53775	588018	8.1464	89.0787
0.011765	11645	599663	1.7641	90.8428
0.015686	5557	605220	0.8418	91.6846
0.019608	3549	608769	0.5376	92.2222
0.023529	2732	611501	0.4139	92.6361
0.027451	2207	613708	0.3343	92.9704
0.031373	1906	615614	0.2887	93.2592
0.035294	1680	617294	0.2545	93.5137
0.039216	1449	618743	0.2195	93.7332
0.043137	1317	620060	0.1995	93.9327
0.047059	1294	621354	0.196	94.1287
0.05098	1198	622552	0.1815	94.3102
0.054902	1103	623655	0.1671	94.4773
0.058824	1018	624673	0.1542	94.6315
0.062745	960	625633	0.1454	94.7769
0.066667	871	626504	0.1319	94.9089
0.070588	890	627394	0.1348	95.0437
0.07451	782	628176	0.1185	95.1622
0.078431	732	628908	0.1109	95.2731
0.082353	734	629642	0.1112	95.3843
0.086275	644	630286	0.0976	95.4818
0.090196	657	630943	0.0995	95.5813
0.094118	590	631533	0.0894	95.6707
0.098039	596	632129	0.0903	95.761
0.101961	542	632671	0.0821	95.8431
0.105882	506	633177	0.0767	95.9198
0.109804	524	633701	0.0794	95.9992
0.113725	467	634168	0.0707	96.0699
0.117647	440	634608	0.0667	96.1366

0.121569	488	635096	0.0739	96.2105
0.12549	456	635552	0.0691	96.2796
0.129412	432	635984	0.0654	96.345
0.133333	401	636385	0.0607	96.4058
0.137255	374	636759	0.0567	96.4624
0.141176	419	637178	0.0635	96.5259
0.145098	432	637610	0.0654	96.5913
0.14902	393	638003	0.0595	96.6509
0.152941	330	638333	0.05	96.7009
0.156863	375	638708	0.0568	96.7577
0.160784	323	639031	0.0489	96.8066
0.164706	359	639390	0.0544	96.861
0.168627	297	639687	0.045	96.906
0.172549	318	640005	0.0482	96.9541
0.176471	289	640294	0.0438	96.9979
0.180392	265	640559	0.0401	97.0381
0.184314	282	640841	0.0427	97.0808
0.188235	258	641099	0.0391	97.1199
0.192157	231	641330	0.035	97.1549
0.196078	199	641529	0.0301	97.185
0.2	191	641720	0.0289	97.214
0.203922	185	641905	0.028	97.242
0.207843	178	642083	0.027	97.2689
0.211765	169	642252	0.0256	97.2945
0.215686	158	642410	0.0239	97.3185
0.219608	159	642569	0.0241	97.3426
0.223529	166	642735	0.0251	97.3677
0.227451	174	642909	0.0264	97.3941
0.231373	157	643066	0.0238	97.4179
0.235294	146	643212	0.0221	97.44
0.239216	151	643363	0.0229	97.4629
0.243137	134	643497	0.0203	97.4832
0.247059	132	643629	0.02	97.5031
0.25098	135	643764	0.0205	97.5236
0.254902	120	643884	0.0182	97.5418
0.258824	105	643989	0.0159	97.5577
0.262745	117	644106	0.0177	97.5754
0.266667	118	644224	0.0179	97.5933
0.270588	91	644315	0.0138	97.6071
0.27451	111	644426	0.0168	97.6239
0.278431	119	644545	0.018	97.6419
0.282353	115	644660	0.0174	97.6593
0.286275	104	644764	0.0158	97.6751
0.290196	122	644886	0.0185	97.6936
0.294118	113	644999	0.0171	97.7107
0.298039	109	645108	0.0165	97.7272
0.301961	105	645213	0.0159	97.7431
0.305882	107	645320	0.0162	97.7593
0.309804	108	645428	0.0164	97.7757
0.313725	101	645529	0.0153	97.791
0.317647	102	645631	0.0155	97.8064
0.321569	110	645741	0.0167	97.8231
0.32549	93	645834	0.0141	97.8372
0.329412	100	645934	0.0151	97.8523
0.333333	96	646030	0.0145	97.8669
0.337255	99	646129	0.015	97.8819
0.341176	93	646222	0.0141	97.896
0.345098	112	646334	0.017	97.9129
0.34902	128	646462	0.0194	97.9323
0.352941	97	646559	0.0147	97.947
0.356863	108	646667	0.0164	97.9634
0.360784	85	646752	0.0129	97.9762
0.364706	109	646861	0.0165	97.9928
0.368627	110	646971	0.0167	98.0094
0.372549	114	647085	0.0173	98.0267
0.376471	113	647198	0.0171	98.0438
0.380392	102	647300	0.0155	98.0593

0.384314	94	647394	0.0142	98.0735
0.388235	116	647510	0.0176	98.0911
0.392157	123	647633	0.0186	98.1097
0.396078	101	647734	0.0153	98.125
0.4	112	647846	0.017	98.142
0.403922	107	647953	0.0162	98.1582
0.407843	112	648065	0.017	98.1752
0.411765	138	648203	0.0209	98.1961
0.415686	116	648319	0.0176	98.2136
0.419608	123	648442	0.0186	98.2323
0.423529	145	648587	0.022	98.2542
0.427451	122	648709	0.0185	98.2727
0.431373	121	648830	0.0183	98.291
0.435294	131	648961	0.0198	98.3109
0.439216	140	649101	0.0212	98.3321
0.443137	127	649228	0.0192	98.3513
0.447059	134	649362	0.0203	98.3716
0.45098	146	649508	0.0221	98.3938
0.454902	150	649658	0.0227	98.4165
0.458824	128	649786	0.0194	98.4359
0.462745	132	649918	0.02	98.4559
0.466667	149	650067	0.0226	98.4784
0.470588	145	650212	0.022	98.5004
0.47451	141	650353	0.0214	98.5218
0.478431	146	650499	0.0221	98.5439
0.482353	139	650638	0.0211	98.5649
0.486275	185	650823	0.028	98.593
0.490196	166	650989	0.0251	98.6181
0.494118	138	651127	0.0209	98.639
0.498039	142	651269	0.0215	98.6605
0.501961	164	651433	0.0248	98.6854
0.505882	182	651615	0.0276	98.7129
0.509804	161	651776	0.0244	98.7373
0.513725	177	651953	0.0268	98.7641
0.517647	153	652106	0.0232	98.7873
0.521569	166	652272	0.0251	98.8125
0.52549	168	652440	0.0255	98.8379
0.529412	156	652596	0.0236	98.8616
0.533333	150	652746	0.0227	98.8843
0.537255	154	652900	0.0233	98.9076
0.541176	186	653086	0.0282	98.9358
0.545098	184	653270	0.0279	98.9637
0.54902	151	653421	0.0229	98.9865
0.552941	172	653593	0.0261	99.0126
0.556863	161	653754	0.0244	99.037
0.560784	162	653916	0.0245	99.0615
0.564706	161	654077	0.0244	99.0859
0.568627	173	654250	0.0262	99.1121
0.572549	185	654435	0.028	99.1401
0.576471	180	654615	0.0273	99.1674
0.580392	135	654750	0.0205	99.1879
0.584314	149	654899	0.0226	99.2104
0.588235	141	655040	0.0214	99.2318
0.592157	134	655174	0.0203	99.2521
0.596078	154	655328	0.0233	99.2754
0.6	134	655462	0.0203	99.2957
0.603922	129	655591	0.0195	99.3153
0.607843	149	655740	0.0226	99.3378
0.611765	105	655845	0.0159	99.3537
0.615686	121	655966	0.0183	99.3721
0.619608	121	656087	0.0183	99.3904
0.623529	142	656229	0.0215	99.4119
0.627451	125	656354	0.0189	99.4309
0.631373	103	656457	0.0156	99.4465
0.635294	122	656579	0.0185	99.4649
0.639216	114	656693	0.0173	99.4822
0.643137	124	656817	0.0188	99.501

0.647059	114	656931	0.0173	99.5183
0.65098	123	657054	0.0186	99.5369
0.654902	87	657141	0.0132	99.5501
0.658824	107	657248	0.0162	99.5663
0.662745	89	657337	0.0135	99.5798
0.666667	94	657431	0.0142	99.594
0.670588	110	657541	0.0167	99.6107
0.67451	109	657650	0.0165	99.6272
0.678431	86	657736	0.013	99.6402
0.682353	107	657843	0.0162	99.6564
0.686275	82	657925	0.0124	99.6688
0.690196	91	658016	0.0138	99.6826
0.694118	73	658089	0.0111	99.6937
0.698039	70	658159	0.0106	99.7043
0.701961	75	658234	0.0114	99.7157
0.705882	77	658311	0.0117	99.7273
0.709804	78	658389	0.0118	99.7391
0.713725	86	658475	0.013	99.7522
0.717647	79	658554	0.012	99.7641
0.721569	86	658640	0.013	99.7772
0.72549	63	658703	0.0095	99.7867
0.729412	71	658774	0.0108	99.7975
0.733333	54	658828	0.0082	99.8056
0.737255	76	658904	0.0115	99.8172
0.741176	58	658962	0.0088	99.8259
0.745098	55	659017	0.0083	99.8343
0.74902	68	659085	0.0103	99.8446
0.752941	67	659152	0.0101	99.8547
0.756863	64	659216	0.0097	99.8644
0.760784	52	659268	0.0079	99.8723
0.764706	54	659322	0.0082	99.8805
0.768627	42	659364	0.0064	99.8868
0.772549	38	659402	0.0058	99.8926
0.776471	37	659439	0.0056	99.8982
0.780392	45	659484	0.0068	99.905
0.784314	35	659519	0.0053	99.9103
0.788235	31	659550	0.0047	99.915
0.792157	34	659584	0.0052	99.9202
0.796078	32	659616	0.0048	99.925
0.8	32	659648	0.0048	99.9299
0.803922	29	659677	0.0044	99.9343
0.807843	21	659698	0.0032	99.9374
0.811765	21	659719	0.0032	99.9406
0.815686	19	659738	0.0029	99.9435
,0.819608	20	659758	0.003	99.9465
0.823529	22	659780	0.0033	99.9499
0.827451	15	659795	0.0023	99.9521
0.831373	14	659809	0.0021	99.9543
0.835294	11	659820	0.0017	99.9559
0.839216	20	659840	0.003	99.9589
0.843137	14	659854	0.0021	99.9611
0.847059	10	659864	0.0015	99.9626
0.85098	8	659872	0.0012	99.9638
0.854902	10	659882	0.0015	99.9653
0.858824	5	659887	0.0008	99.9661
0.862745	6	659893	0.0009	99.967
0.866667	10	659903	0.0015	99.9685
0.870588	9	659912	0.0014	99.9699
0.87451	8	659920	0.0012	99.9711
0.878431	8	659928	0.0012	99.9723
0.882353	3	659931	0.0005	99.9727
0.886275	4	659935	0.0006	99.9733
0.890196	7	659942	0.0011	99.9744
0.894118	5	659947	0.0008	99.9752
0.898039	5	659952	0.0008	99.9759
0.901961	4	659956	0.0006	99.9765
0.905882	4	659960	0.0006	99.9771

0.909804	12	659972	0.0018	99.9789
0.913725	4	659976	0.0006	99.9795
0.917647	3	659979	0.0005	99.98
0.921569	6	659985	0.0009	99.9809
0.92549	5	659990	0.0008	99.9817
0.929412	5	659995	0.0008	99.9824
0.933333	8	660003	0.0012	99.9836
0.937255	6	660009	0.0009	99.9845
0.941176	8	660017	0.0012	99.9858
0.945098	10	660027	0.0015	99.9873
0.94902	3	660030	0.0005	99.9877
0.952941	3	660033	0.0005	99.9882
0.956863	5	660038	0.0008	99.9889
0.960784	6	660044	0.0009	99.9899
0.964706	6	660050	0.0009	99.9908
0.968627	7	660057	0.0011	99.9918
0.972549	1	660058	0.0002	99.992
0.976471	3	660061	0.0005	99.9924
0.980392	2	660063	0.0003	99.9927
0.984314	6	660069	0.0009	99.9936
0.988235	2	660071	0.0003	99.9939
0.992157	5	660076	0.0008	99.9947
0.996078	2	660078	0.0003	99.995
1	33	660111	0.005	100

5.1.3 Classification accuracy

The accuracy of linear spectral unmixing classification was measured by the amount of mean RMSE of the image, obtained by considering the RMSE of each individual pixel, given by equation (4) (page 13), in the image. This is the residual error which occurs as a result of some inevitable inaccuracies in defining the endmember spectral library, so that some class features present in a pixel do not appropriately match any of the spectral signatures provided in the input endmember library. Using the RMSE values as an indicator of the classification accuracies, the results displayed very high accuracy levels. Table 5-8 shows the statistics of RMSE pixel values for the 15-12-2010 image.

Table 5-8: A statistical summary of RMSE pixel values for 15-12-2010 image

Pixel Value (RMSE)	Number of pixels	Total	Percentage	Accumulated Percentage
0.000068	70650	70650	10.7027	10.7027
0.000589	245595	316245	37.2051	47.9079
0.001111	118885	435130	18.0098	65.9177
0.001632	55094	490224	8.3462	74.2639
0.002153	31631	521855	4.7918	79.0556
0.002674	21039	542894	3.1872	82.2428
0.003195	14196	557090	2.1505	84.3934
0.003716	10746	567836	1.6279	86.0213
0.004237	8406	576242	1.2734	87.2947
0.004759	6821	583063	1.0333	88.328
0.00528	5780	588843	0.8756	89.2036
0.005801	4555	593398	0.69	89.8937
0.006322	3989	597387	0.6043	90.498
0.006843	3414	600801	0.5172	91.0151

0.007364	24864	625665	3.7666	94.7818
0.007886	2564	628229	0.3884	95.1702
0.008407	2233	630462	0.3383	95.5085
0.008928	1960	632422	0.2969	95.8054
0.009449	1763	634185	0.2671	96.0725
0.00997	1560	635745	0.2363	96.3088
0.010491	1389	637134	0.2104	96.5192
0.011012	1209	638343	0.1832	96.7024
0.011534	1142	639485	0.173	96.8754
0.012055	1056	640541	0.16	97.0353
0.012576	975	641516	0.1477	97.183
0.013097	899	642415	0.1362	97.3192
0.013618	824	643239	0.1248	97.4441
0.014139	789	644028	0.1195	97.5636
0.01466	756	644784	0.1145	97.6781
0.015182	682	645466	0.1033	97.7814
0.015703	626	646092	0.0948	97.8763
0.016224	494	646586	0.0748	97.9511
0.016745	537	647123	0.0813	98.0325
0.017266	493	647616	0.0747	98.1071
0.017787	478	648094	0.0724	98.1795
0.018308	451	648545	0.0683	98.2479
0.01883	438	648983	0.0664	98.3142
0.019351	403	649386	0.0611	98.3753
0.019872	364	649750	0.0551	98.4304
0.020393	373	650123	0.0565	98.4869
0.020914	342	650465	0.0518	98.5387
0.021435	302	650767	0.0457	98.5845
0.021957	317	651084	0.048	98.6325
0.022478	310	651394	0.047	98.6795
0.022999	256	651650	0.0388	98.7182
0.02352	265	651915	0.0401	98.7584
0.024041	258	652173	0.0391	98.7975
0.024562	244	652417	0.037	98.8344
0.025083	227	652644	0.0344	98.8688
0.025605	194	652838	0.0294	98.8982
0.026126	194	653032	0.0294	98.9276
0.026647	180	653212	0.0273	98.9549
0.027168	187	653399	0.0283	98.9832
0.027689	149	653548	0.0226	99.0058
0.02821	147	653695	0.0223	99.028
0.028731	142	653837	0.0215	99.0496
0.029253	127	653964	0.0192	99.0688
0.029774	149	654113	0.0226	99.0914
0.030295	140	654253	0.0212	99.1126
0.030816	133	654386	0.0201	99.1327
0.031337	145	654531	0.022	99.1547
0.031858	97	654628	0.0147	99.1694
0.03238	132	654760	0.02	99.1894
0.032901	109	654869	0.0165	99.2059
0.033422	114	654983	0.0173	99.2232
0.033943	114	655097	0.0173	99.2404
0.034464	120	655217	0.0182	99.2586
0.034985	112	655329	0.017	99.2756
0.035506	93	655422	0.0141	99.2897
0.036028	93	655515	0.0141	99.3038
0.036549	99	655614	0.015	99.3188
0.03707	88	655702	0.0133	99.3321
0.037591	92	655794	0.0139	99.346
0.038112	79	655873	0.012	99.358
0.038633	109	655982	0.0165	99.3745
0.039154	97	656079	0.0147	99.3892
0.039676	84	656163	0.0127	99.4019
0.040197	101	656264	0.0153	99.4172
0.040718	94	656358	0.0142	99.4315
0.041239	84	656442	0.0127	99.4442
0.04176	89	656531	0.0135	99.4577

0.042281	92	656623	0.0139	99.4716
0.042802	85	656708	0.0129	99.4845
0.043324	85	656793	0.0129	99.4974
0.043845	76	656869	0.0115	99.5089
0.044366	93	656962	0.0141	99.523
0.044887	73	657035	0.0111	99.534
0.045408	79	657114	0.012	99.546
0.045929	83	657197	0.0126	99.5586
0.046451	68	657265	0.0103	99.5689
0.046972	74	657339	0.0112	99.5801
0.047493	75	657414	0.0114	99.5914
0.048014	79	657493	0.012	99.6034
0.048535	67	657560	0.0101	99.6135
0.049056	72	657632	0.0109	99.6245
0.049577	72	657704	0.0109	99.6354
0.050099	63	657767	0.0095	99.6449
0.05062	61	657828	0.0092	99.6541
0.051141	91	657919	0.0138	99.6679
0.051662	74	657993	0.0112	99.6791
0.052183	70	658063	0.0106	99.6897
0.052704	56	658119	0.0085	99.6982
0.053225	61	658180	0.0092	99.7075
0.053747	62	658242	0.0094	99.7169
0.054268	64	658306	0.0097	99.7266
0.054789	68	658374	0.0103	99.7369
0.05531	69	658443	0.0105	99.7473
0.055831	57	658500	0.0086	99.756
0.056352	63	658563	0.0095	99.7655
0.056873	79	658642	0.012	99.7775
0.057395	52	658694	0.0079	99.7853
0.057916	57	658751	0.0086	99.794
0.058437	69	658820	0.0105	99.8044
0.058958	74	658894	0.0112	99.8156
0.059479	62	658956	0.0094	99.825
0.06	60	659016	0.0091	99.8341
0.060522	45	659061	0.0068	99.8409
0.061043	55	659116	0.0083	99.8493
0.061564	68	659184	0.0103	99.8596
0.062085	48	659232	0.0073	99.8668
0.062606	52	659284	0.0079	99.8747
0.063127	33	659317	0.005	99.8797
0.063648	40	659357	0.0061	99.8858
0.06417	36	659393	0.0055	99.8912
0.064691	43	659436	0.0065	99.8977
0.065212	47	659483	0.0071	99.9049
0.065733	33	659516	0.005	99.9099
0.066254	39	659555	0.0059	99.9158
0.066775	35	659590	0.0053	99.9211
0.067296	31	659621	0.0047	99.9258
0.067818	26	659647	0.0039	99.9297
0.068339	17	659664	0.0026	99.9323
0.06886	16	659680	0.0024	99.9347
0.069381	28	659708	0.0042	99.9389
0.069902	12	659720	0.0018	99.9408
0.070423	21	659741	0.0032	99.9439
0.070945	18	659759	0.0027	99.9467
0.071466	23	659782	0.0035	99.9502
0.071987	20	659802	0.003	99.9532
0.072508	19	659821	0.0029	99.9561
0.073029	12	659833	0.0018	99.9579
0.07355	10	659843	0.0015	99.9594
0.074071	12	659855	0.0018	99.9612
0.074593	15	659870	0.0023	99.9635
0.075114	10	659880	0.0015	99.965
0.075635	14	659894	0.0021	99.9671
0.076156	12	659906	0.0018	99.9689
0.076677	17	659923	0.0026	99.9715

0.077198	7	659930	0.0011	99.9726
0.077719	10	659940	0.0015	99.9741
0.078241	11	659951	0.0017	99.9758
0.078762	11	659962	0.0017	99.9774
0.079283	10	659972	0.0015	99.9789
0.079804	8	659980	0.0012	99.9802
0.080325	3	659983	0.0005	99.9806
0.080846	3	659986	0.0005	99.9811
0.081367	6	659992	0.0009	99.982
0.081889	9	660001	0.0014	99.9833
0.08241	4	660005	0.0006	99.9839
0.082931	1	660006	0.0002	99.9841
0.083452	6	660012	0.0009	99.985
0.083973	6	660018	0.0009	99.9859
0.084494	4	660022	0.0006	99.9865
0.085016	4	660026	0.0006	99.9871
0.085537	3	660029	0.0005	99.9876
0.086058	3	660032	0.0005	99.988
0.086579	3	660035	0.0005	99.9885
0.0871	3	660038	0.0005	99.9889
0.087621	3	660041	0.0005	99.9894
0.088142	3	660044	0.0005	99.9899
0.088664	4	660048	0.0006	99.9905
0.089185	7	660055	0.0011	99.9915
0.089706	1	660056	0.0002	99.9917
0.090227	3	660059	0.0005	99.9921
0.090748	1	660060	0.0002	99.9923
0.091269	0	660060	0	99.9923
0.09179	4	660064	0.0006	99.9929
0.092312	5	660069	0.0008	99.9936
0.092833	1	660070	0.0002	99.9938
0.093354	3	660073	0.0005	99.9942
0.093875	1	660074	0.0002	99.9944
0.094396	0	660074	0	99.9944
0.094917	1	660075	0.0002	99.9945
0.095438	3	660078	0.0005	99.995
0.09596	0	660078	0	99.995
0.096481	1	660079	0.0002	99.9952
0.097002	0	660079	0	99.9952
0.097523	0	660079	0	99.9952
0.098044	1	660080	0.0002	99.9953
0.098565	2	660082	0.0003	99.9956
0.099087	1	660083	0.0002	99.9958
0.099608	0	660083	0	99.9958
0.100129	2	660085	0.0003	99.9961
0.10065	2	660087	0.0003	99.9964
0.101171	1	660088	0.0002	99.9965
0.101692	0	660088	0	99.9965
0.102213	1	660089	0.0002	99.9967
0.102735	1	660090	0.0002	99.9968
0.103256	2	660092	0.0003	99.9971
0.103777	1	660093	0.0002	99.9973
0.104298	0	660093	0	99.9973
0.104819	2	660095	0.0003	99.9976
0.10534	0	660095	0	99.9976
0.105861	1	660096	0.0002	99.9977
0.106383	2	660098	0.0003	99.998
0.106904	1	660099	0.0002	99.9982
0.107425	1	660100	0.0002	99.9983
0.107946	0	660100	0	99.9983
0.108467	1	660101	0.0002	99.9985
0.108988	0	660101	0	99.9985
0.10951	0	660101	0	99.9985
0.110031	0	660101	0	99.9985
0.110552	0	660101	0	99.9985
0.111073	1	660102	0.0002	99.9986
0.111594	1	660103	0.0002	99.9988

0.112115	0	660103	0	99.9988
0.112636	0	660103	0	99.9988
0.113158	0	660103	0	99.9988
0.113679	0	660103	0	99.9988
0.1142	0	660103	0	99.9988
0.114721	0	660103	0	99.9988
0.115242	0	660103	0	99.9988
0.115763	2	660105	0.0003	99.9991
0.116284	1	660106	0.0002	99.9992
0.116806	1	660107	0.0002	99.9994
0.117327	1	660108	0.0002	99.9995
0.117848	0	660108	0	99.9995
0.118369	0	660108	0	99.9995
0.11889	0	660108	0	99.9995
0.119411	1	660109	0.0002	99.9997
0.119932	0	660109	0	99.9997
0.120454	0	660109	0	99.9997
0.120975	0	660109	0	99.9997
0.121496	0	660109	0	99.9997
0.122017	0	660109	0	99.9997
0.122538	0	660109	0	99.9997
0.123059	0	660109	0	99.9997
0.123581	0	660109	0	99.9997
0.124102	0	660109	0	99.9997
0.124623	0	660109	0	99.9997
0.125144	0	660109	0	99.9997
0.125665	0	660109	0	99.9997
0.126186	0	660109	0	99.9997
0.126707	0	660109	0	99.9997
0.127229	0	660109	0	99.9997
0.12775	1	660110	0.0002	99.9998
0.128271	0	660110	0	99.9998
0.128792	0	660110	0	99.9998
0.129313	0	660110	0	99.9998
0.129834	0	660110	0	99.9998
0.130355	0	660110	0	99.9998
0.130877	0	660110	0	99.9998
0.131398	0	660110	0	99.9998
0.131919	0	660110	0	99.9998
0.13244	0	660110	0	99.9998
0.132961	1	660111	0.0002	100

From table 5-8, it is noticed that more than 96.3% of the pixels were classified with less than 0.01 RMSE, or more than 99% accuracy. Figure 5-9 shows a histogram of the RMSE values of one of the images.

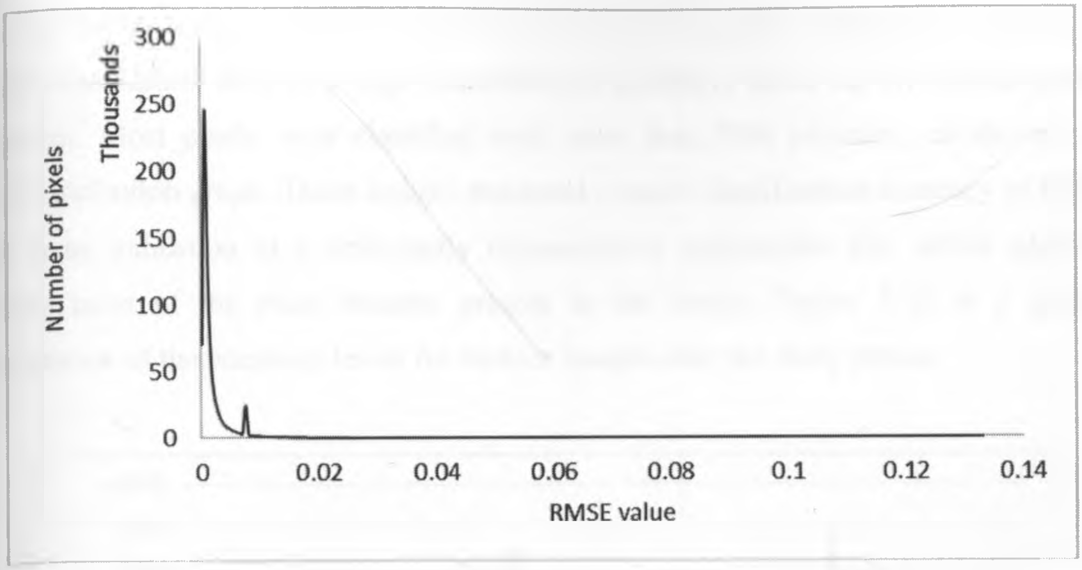

Figure 5-9: A histogram of RMSE values for one of the images. Majority of the pixels have RMSE values less than 0.02

Mean RMSE for an image was then computed by taking into account the number of pixels for each RMSE and the total number of pixels in the image. The mean RMSE value for this image is 0.002709 giving a percentage error of 0.27%, and thus the overall classification accuracy for the image is 99.73%. Table 5-9 shows the mean RMSE values and the corresponding percentage classification accuracies for various images.

Table 5-9: Classification accuracy assessment

Date	Mean RMSE	Percentage Error	Percentage Accuracy
26/12/2003	0.004331	0.4331	99.5669
21/02/2005	0.005201	0.5201	99.4799
17/07/2005	0.007342	0.7342	99.2658
02/01/2006	0.002851	0.2851	99.7149
06/02/2006	0.002936	0.2936	99.7064
16/08/2006	0.004417	0.4417	99.5583
12/10/2006	0.012835	1.2835	98.7165
20/02/2007	0.005826	0.5826	99.4174
27/09/2008	0.005384	0.5384	99.4616
14/02/2009	0.002114	0.2114	99.7886
21/03/2009	0.004325	0.4325	99.5675
12/06/2009	0.003222	0.3222	99.6778
02/08/2009	0.007701	0.7701	99.2299
28/09/2009	0.00418	0.418	99.582
12/02/2010	0.004333	0.4333	99.5667
03/08/2010	0.005996	0.5996	99.4004

From the Table 5-9, it is observed that the spectral unmixing technique produced good classification results with very high classification accuracy, based on the RMSE accuracy assessment. Most pixels were classified with more than 99% accuracy, as shown in the RMSE distribution graph. These images produced a mean classification accuracy of 99.48%, which is an indication of a sufficiently representative endmember file, which adequately describes most of the class features present in the image. Figure 5-10 is a graphical representation of the accuracy levels for various images over the study period.

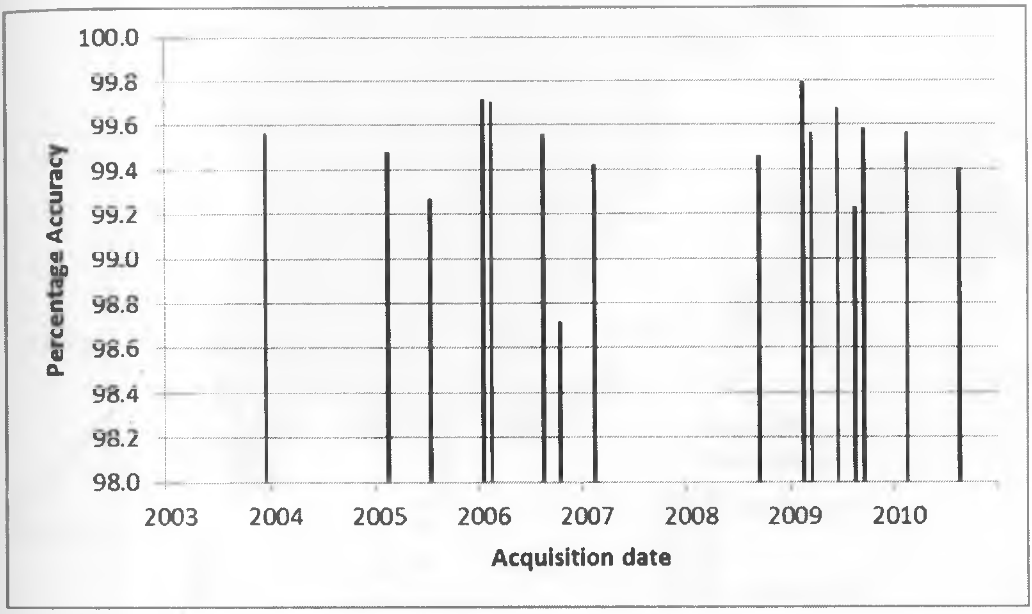

Figure 5-10: A bar graph showing percentage classification accuracy for various images. All images were classified with more than 98% accuracy

5.2 *Mapping and Monitoring*

5.2.1 Monitoring spatial distribution of the aquatic plants

When classification was complete, vegetation cover maps were then generated using *ArcGIS 9.3* software. The images used in this study were received when they were already geocoded, and were easily imported to *ArcGIS 9.3* software for mapping. The images were projected in a *UTM Zone 36S* coordinate system and *WGS·84* Datum, and resampled using the nearest neighbour technique, which preserves the spectral integrity of the image pixel. Vegetation maps were then generated. Figure 5-11 is a map showing the spatial distribution of aquatic plants in the lake on December 15th, 2010. It is observed from the map that there is a massive infestation of the aquatic plants in the Winam Gulf, the almost enclosed section in the Kenyan side of the lake.

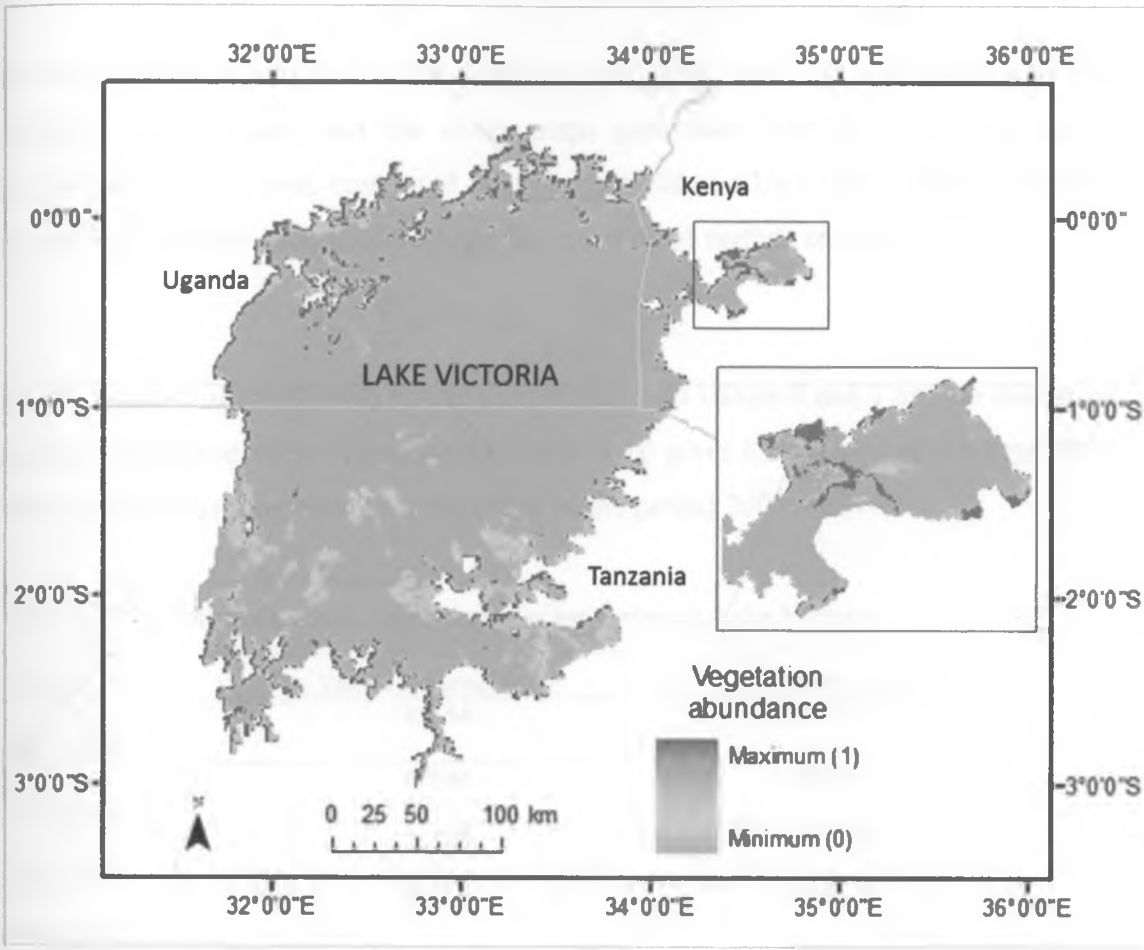

Figure 5-11: Map showing the spatial distribution of aquatic vegetation in Lake Victoria on 15-12-2010. The map displays the fractional abundance of vegetation per pixel, where minimum means the pixels displays open water and maximum means pixel is fully covered by vegetation. Inset (enclosed in red) is the Winam Gulf section of the lake

5.2.2 Monitoring temporal variation of the aquatic plants

5.2.2.1 Surface area estimation

Spatial extent of a particular image constituent is computed by determining the fractional abundance of that feature in all the pixels, as well as the mean pixel area. In BEAM, the mean pixel area is obtained by considering the spatial resolution of the image and putting into consideration the bow-tie effect due to the earth's curvature. For MERIS Full Resolution data, spatial resolution varies in the across track direction, between 0.26 km at nadir and 0.39 km at swath extremities. Along-track sampling is close to 0.29 km. With the earth's radius of 6370.997 km, the mean pixel area is estimated at 0.074 km^2 (ESA, 2010).

To monitor the spatial distribution of the aquatic plants in the lake, the area covered by the aquatic plants was computed and the cover maps generated. The total surface area of vegetation in the lake (A_v) was computed using an algorithm which takes into account the abundance of vegetation in each pixel (a_i) and the mean pixel surface area (A_p);

$$A_v = \sum_{i=0}^{1} a_i \cdot n \cdot A_p \tag{7}$$

where n is the number of pixels with abundance value a_i and values 0 and 1 are the minimum and maximum abundance values respectively. Table 5-10 gives a summary of the vegetation surface areas for various images of the entire lake in the period 2003 – 2010.

Table 5-10: Results of aquatic vegetation area estimation in Lake Victoria

Date	Vegetated area (km^2)	Vegetated area (ha)
26/12/2003	817.94	81794.01
21/02/2005	628.88	62888.43
17/07/2005	738.60	73859.81
02/01/2006	670.28	67027.71
06/02/2006	524.40	52440.13
16/08/2006	615.99	61598.77
20/02/2007	746.16	74616.38
27/09/2008	1070.25	107025.32
14/02/2009	1004.56	100456.41
21/03/2009	885.82	88581.72
12/06/2009	663.43	66342.79
02/08/2009	642.17	64217.26
28/09/2009	770.76	77075.54
12/02/2010	655.61	65560.99
03/08/2010	679.36	67936.34

Lake Victoria covers a very wide area spatially (about 68 800 km^2). Many images were therefore rejected for failure to cover the entire lake. Further, the area is prone to clouds, so that many more images were rejected for severe cloud cover, beyond the preset percentage cloud cover threshold of 5%. This greatly reduced the number of available data for use in the study, so that the time series trends of the vegetation coverage in the lake had several no-data gaps. When a smaller region, the Winam Gulf was considered, the number of usable data increased significantly from 15 to 93 images. Table 5-11 is a summary of the vegetation surface area estimations of the Winam Gulf for various acquisition dates within the period 2003 – 2010.

Table 5-11: Results of aquatic vegetation area estimation in Winam Gulf

Date	Vegetated area (km^2)	Vegetated area (ha)
11/05/2003	33.82171	3382.171
17/05/2003	12.56434	1256.434
02/06/2003	17.46435	1746.435
30/08/2003	15.98224	1598.224
16/12/2003	28.41325	2841.325
26/12/2003	20.94951	2094.951
15/02/2004	21.83479	2183.479
20/05/2004	42.55147	4255.147
01/07/2004	31.73996	3173.996
30/08/2004	15.19065	1519.065
18/09/2004	45.63649	4563.649
26/12/2004	29.86872	2986.872
05/02/2005	13.04411	1304.411
21/02/2005	7.730967	773.0967
19/04/2005	63.28403	6328.403
12/06/2005	33.30988	3330.988
12/08/2005	16.41349	1641.349
03/09/2005	28.67832	2867.832
29/09/2005	43.8736	4387.36
08/10/2005	86.36262	8636.262
03/11/2005	45.82429	4582.429
14/12/2005	11.63109	1163.109
30/12/2005	15.52643	1552.643
02/01/2006	60.83674	6083.674
24/01/2006	77.95841	7795.841
06/02/2006	12.91612	1291.612
25/02/2006	25.7966	2579.66

29/03/2006	22.52676	2252.676
17/04/2006	26.35305	2635.305
30/04/2006	38.69969	3869.969
06/05/2006	43.43046	4343.046
15/07/2006	76.40806	7640.806
16/08/2006	20.21652	2021.652
01/09/2006	43.10439	4310.439
03/10/2006	49.83714	4983.714
13/11/2006	45.86347	4586.347
12/12/2006	60.2305	6023.05
06/01/2007	52.63732	5263.732
20/02/2007	58.64196	5864.196
08/03/2007	119.4526	11945.26
20/05/2007	114.6957	11469.57
02/06/2007	200.0298	20002.98
22/07/2007	86.81768	8681.768
07/08/2007	74.02762	7402.762
23/08/2007	90.143	9014.3
08/09/2007	71.85243	7185.243
21/09/2007	69.50482	6950.482
29/10/2007	62.75957	6275.957
14/11/2007	62.59728	6259.728
30/11/2007	58.49004	5849.004
06/12/2007	57.05684	5705.684
29/12/2007	91.9234	9192.34
04/01/2008	41.76006	4176.006
02/02/2008	68.48107	6848.107
24/02/2008	32.4542	3245.42
11/03/2008	42.82396	4282.396
27/04/2008	39.18437	3918.437
16/05/2008	52.0122	5201.22
01/06/2008	40.66646	4066.646
27/06/2008	22.31888	2231.888
06/07/2008	40.38542	4038.542
10/08/2008	31.26372	3126.372
10/10/2008	33.62286	3362.286
26/10/2008	37.31923	3731.923
17/11/2008	58.89212	5889.212
06/12/2008	32.01996	3201.996
22/12/2008	32.27649	3227.649

27/02/2009	27.8922	2789.22
21/03/2009	57.7322	5773.22
03/04/2009	41.911	4191.1
05/05/2009	69.97878	6997.878
24/05/2009	44.77441	4477.441
02/06/2009	70.89753	7089.753
12/06/2009	54.81892	5481.892
18/06/2009	45.65801	4565.801
17/07/2009	22.15358	2215.358
27/08/2009	41.4565	4145.65
09/09/2009	21.10147	2110.147
28/09/2009	37.03736	3703.736
15/11/2009	31.2164	3121.64
20/12/2009	43.02216	4302.216
30/01/2010	28.33534	2833.534
12/02/2010	31.2164	3121.64
20/04/2010	47.01235	4701.235
21/05/2010	81.60867	8160.867
31/05/2010	56.22057	5622.057
10/06/2010	79.62372	7962.372
27/07/2010	57.00062	5700.062
03/08/2010	36.2957	3629.57
23/09/2010	56.09027	5609.027
15/11/2010	99.79671	9979.671
04/12/2010	95.78852	9578.852
15/12/2010	75.52587	7552.587

5.2.2.2 Time series variation (Vegetation phenology)

Temporal variation in the abundance of the aquatic plants was monitored by graphically analyzing the variation in its spatial extent (surface area coverage) with time, using images covering a wide temporal extent. Figure 5-12 shows the time series variation of vegetation abundance in the lake over the period 2003 – 2010.

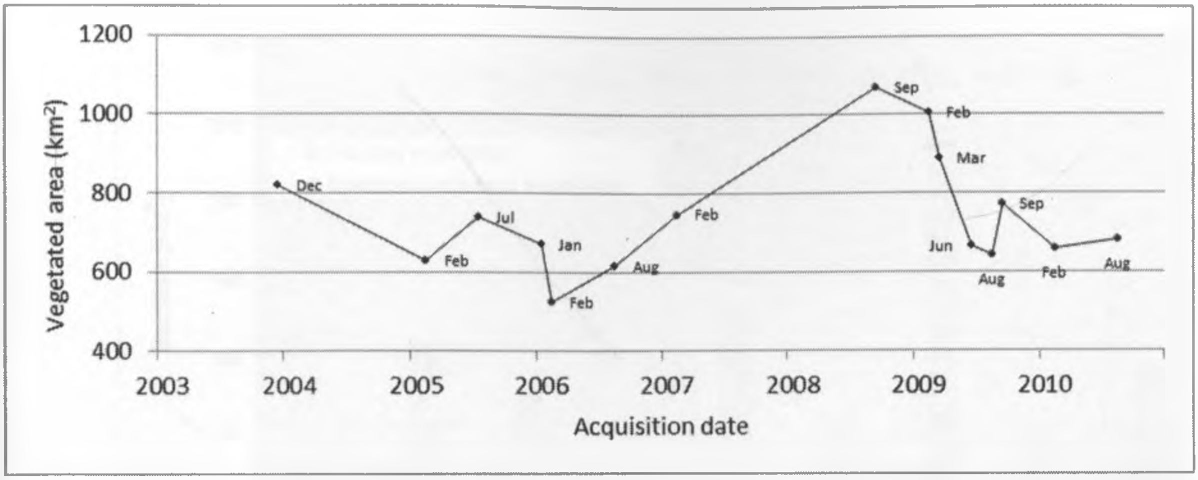

Figure 5-12: Time series variation of vegetation abundance (surface area) in Lake Victoria in the period 2003 – 2010

Figure 5-13 presents the time series variation of aquatic plants in the Winam Gulf over the same study period. See Table 5-11 for the source data.

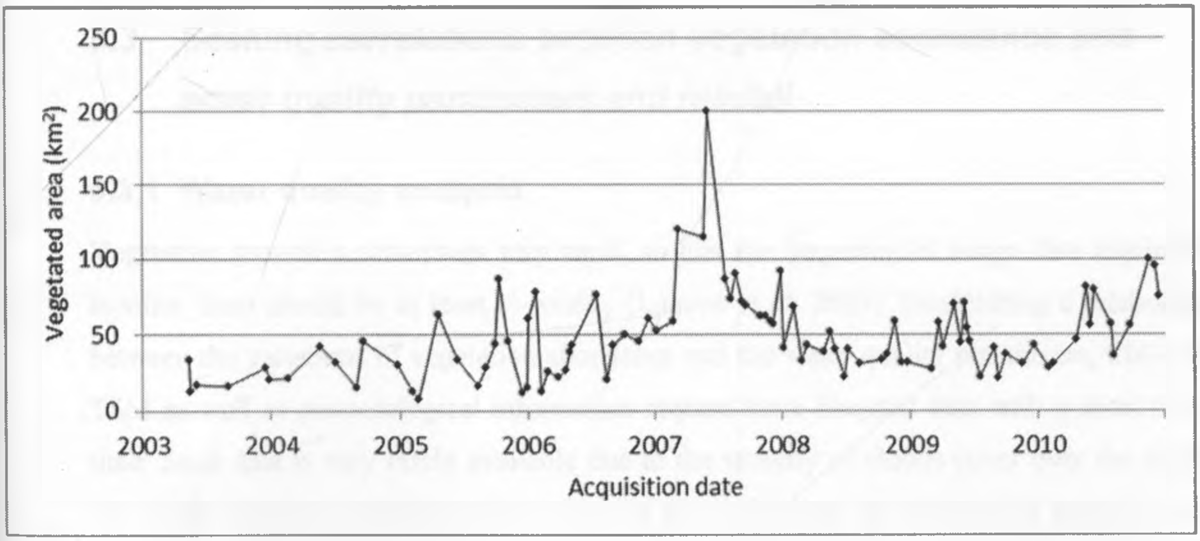

Figure 5-13: Time series variation of vegetation abundance in the Winam Gulf section of Lake Victoria in the period 2003 – 2010

These results show that vegetation cover in the Winam Gulf which was kept below 100 km^2 during the years 2003 to 2006 increased to a peak of about 200 km^2 in 2007, before decreasing again to below 100 km^2 during the years 2008 to 2010. This trend is similar to that of Laneve *et al.* (2010), shown in Figure 5-14, which were obtained using Landsat data.

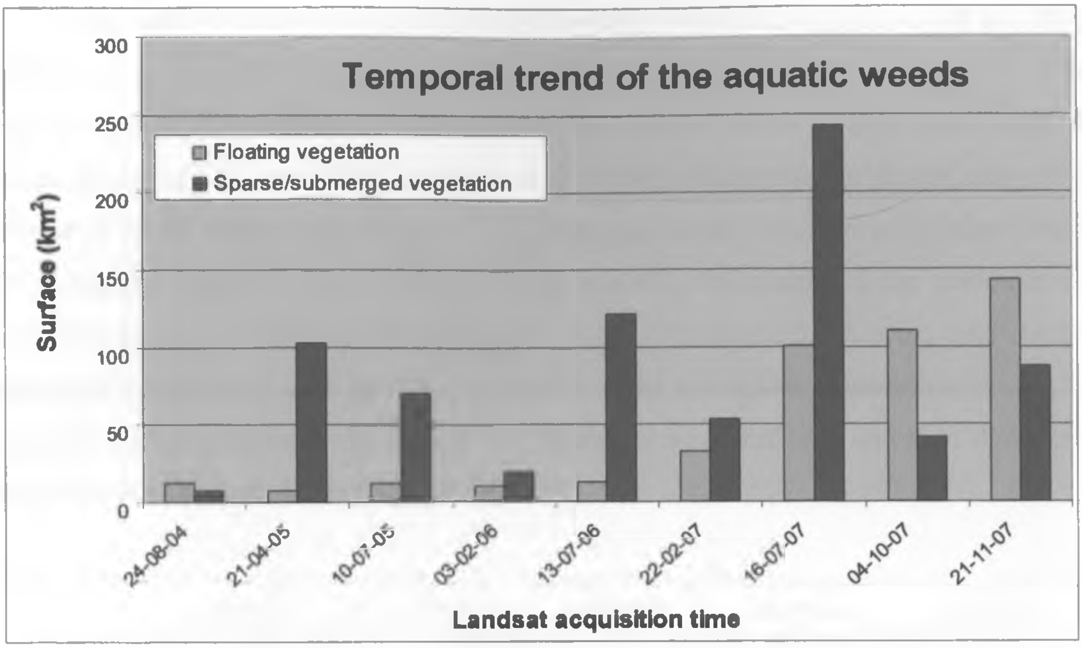

Figure 5-14: Histogram of floating and sparse/submerged vegetation computed from Landsat ETM+ temporal series classification of the Winam Gulf. Source: Laneve *et al.* (2010)

5.3 Seeking correlations between vegetation abundance and water quality parameters and rainfall

5.3.1 Water quality analysis

Vegetation growth is sometimes very rapid, so that the frequency of image data required to monitor them should be at least bi-weekly (Laneve *et al.* 2010). Establishing a relationship between the variations of vegetation abundance and the water quality parameters; Chl-a and TSM as well as meteorological information require more frequent data with a short revisit time. Such data is very rarely available due to the severity of clouds cover over the region. The mean values of concentrations of Chl-a and TSM over the whole lake were not very useful in deriving any correlation with the aquatic plants proliferation because much of the vegetation is along the shore and in the shallow waters, especially the Winam Gulf, and little is found in the deeper waters at the main lake so that the mean values would water down the information. Williams *et al.* (2007) commented that Lake Victoria is the second largest lake in the world and to condense the system into a single graph is an over simplification of the spatial complexity. For this reason therefore, only the Winam Gulf was considered in seeking correlations between vegetation variation and the water quality parameters.

Some water quality parameters have optical properties, which the satellite sensor can detect. *MERIS Eutrophic Lakes Processor 1.4.1* (Doerffer and Schiller, 2008 (b)) in *BEAM 4.8* was used to retrieve the abundance values for some selected water quality parameters; the concentrations of Chl-a and TSM. Koponen et al. (2008) validated the processor using *in situ* measurements for some eutrophic lakes in Europe and Africa which include Lake Victoria. The validation results of Lake Victoria showed a good performance of the processor with correlation between satellite derived data and *in situ* measurements showing coefficient of determinations of R^2 = 0.77 and R^2 = 0.92 for TSM and Chl-a respectively. Using this processor, the pixel concentrations of these parameters were retrieved and their distribution maps were then generated using *ArcGIS 9.3*.

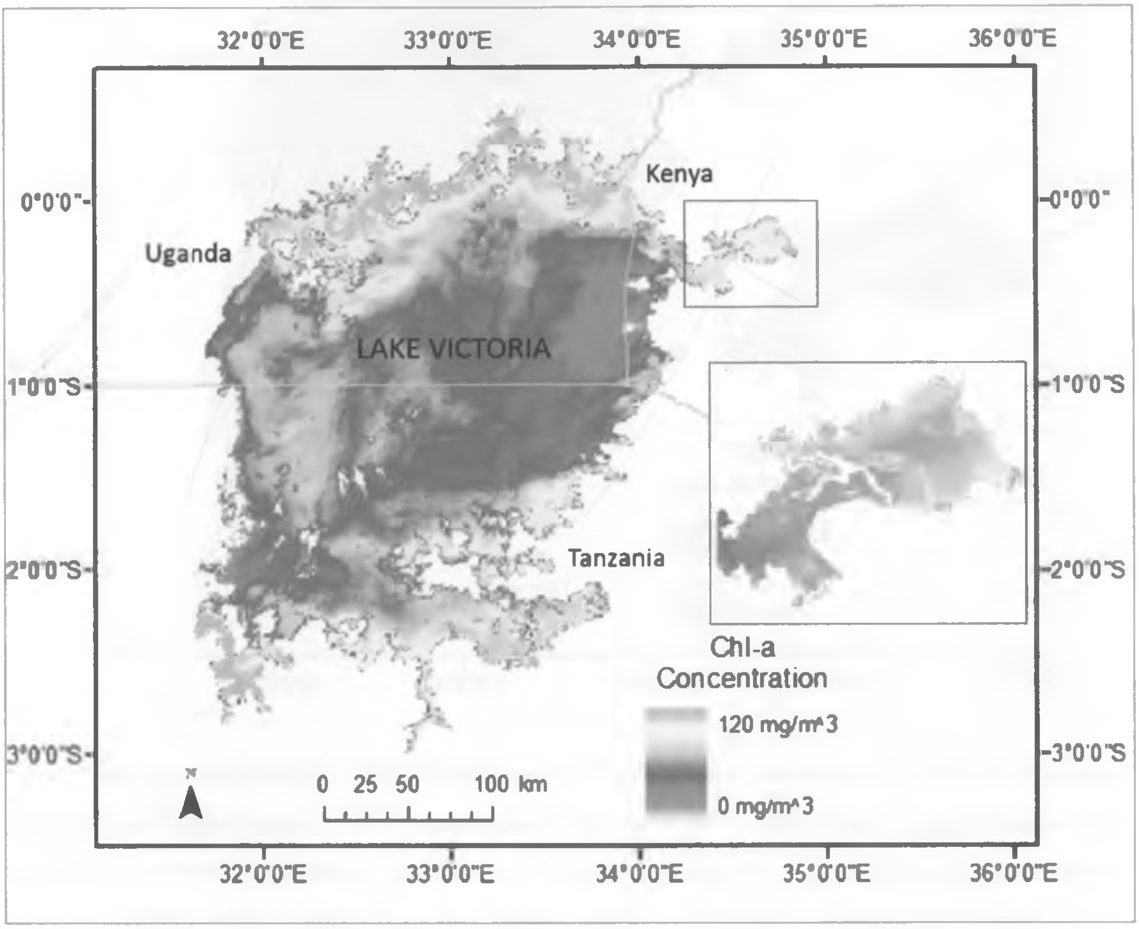

Figure 5-15: Map showing the spatial distribution of Chl-a in Lake Victoria on 15-12-2010. The map displays pixel concentration of Chl-a ranging from zero to 120 mg/m^3. In this image, higher concentration of Chl-a is observed along the shores especially on the Ugandan section of the lake. Inset (enclosed in red) is the Winam Gulf section of the lake

Figure 5-15 shows the distribution of the concentration of Chl-a in the lake on December 15th, 2010. From this figure, high concentrations of Chl-a is observed along the shores of the lake, which generally decreases towards the central part of the lake. Figure 5-16 shows the distribution of the concentration of TSM in the lake on December 15th, 2010. This figure shows a very high concentration of TSM in the Winam Gulf and in some other bays. The central part of the lake is however free from these sediments.

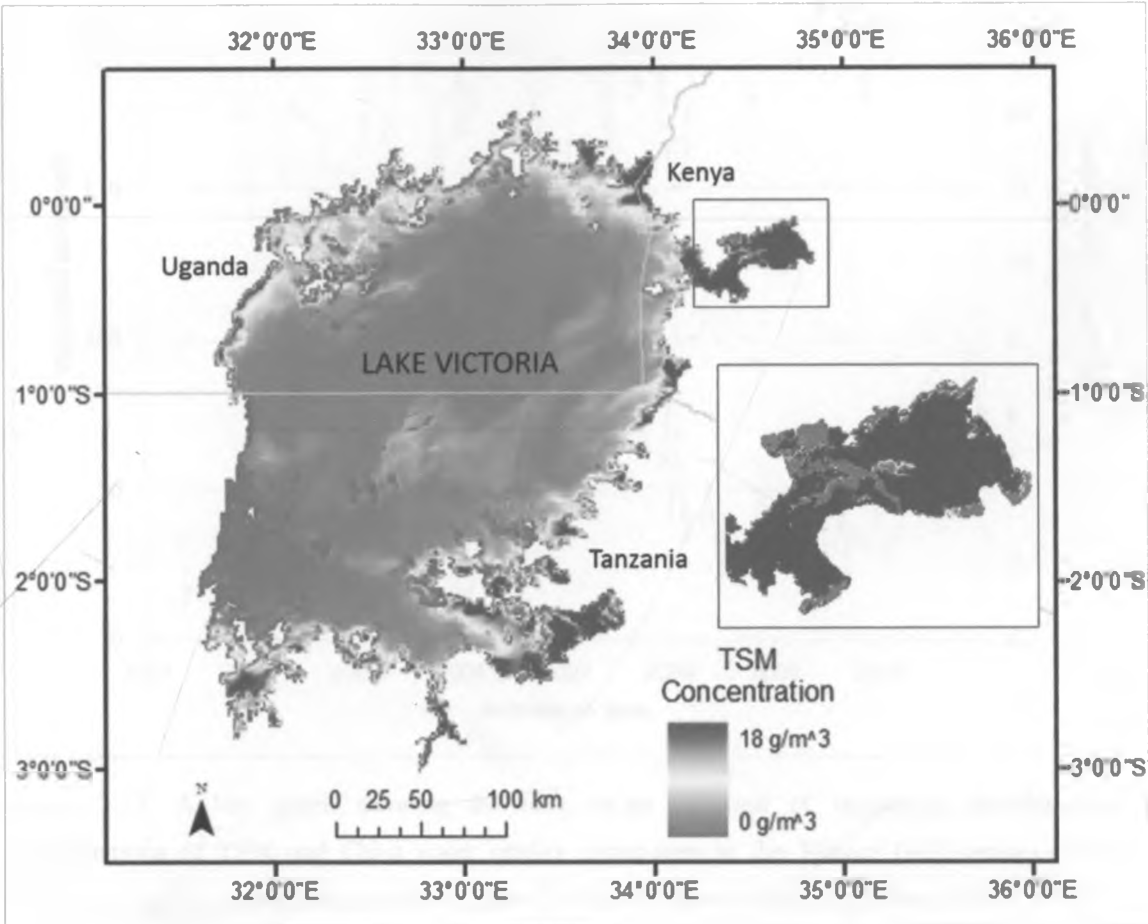

Figure 5-16: Map showing the spatial distribution of TSM in Lake Victoria on 15-12-2010. The map displays pixel concentration of TSM ranging from zero to 18 g/m^3. In this image, higher concentration of TSM is observed on the shallow sections of the lake especially the Winam Gulf and some sections on the Tanzanian section of lake. Inset (enclosed in red) is the Winam Gulf section of the lake

The temporal variation of these water quality parameters were compared with those of the aquatic plants, with a view of determining if any correlation exists. The mean values of concentrations of Chl-a and TSM over some selected open water regions of Winam Gulf were computed for every image. Figure 5-17 shows the time series variation of vegetation abundance with these water quality values.

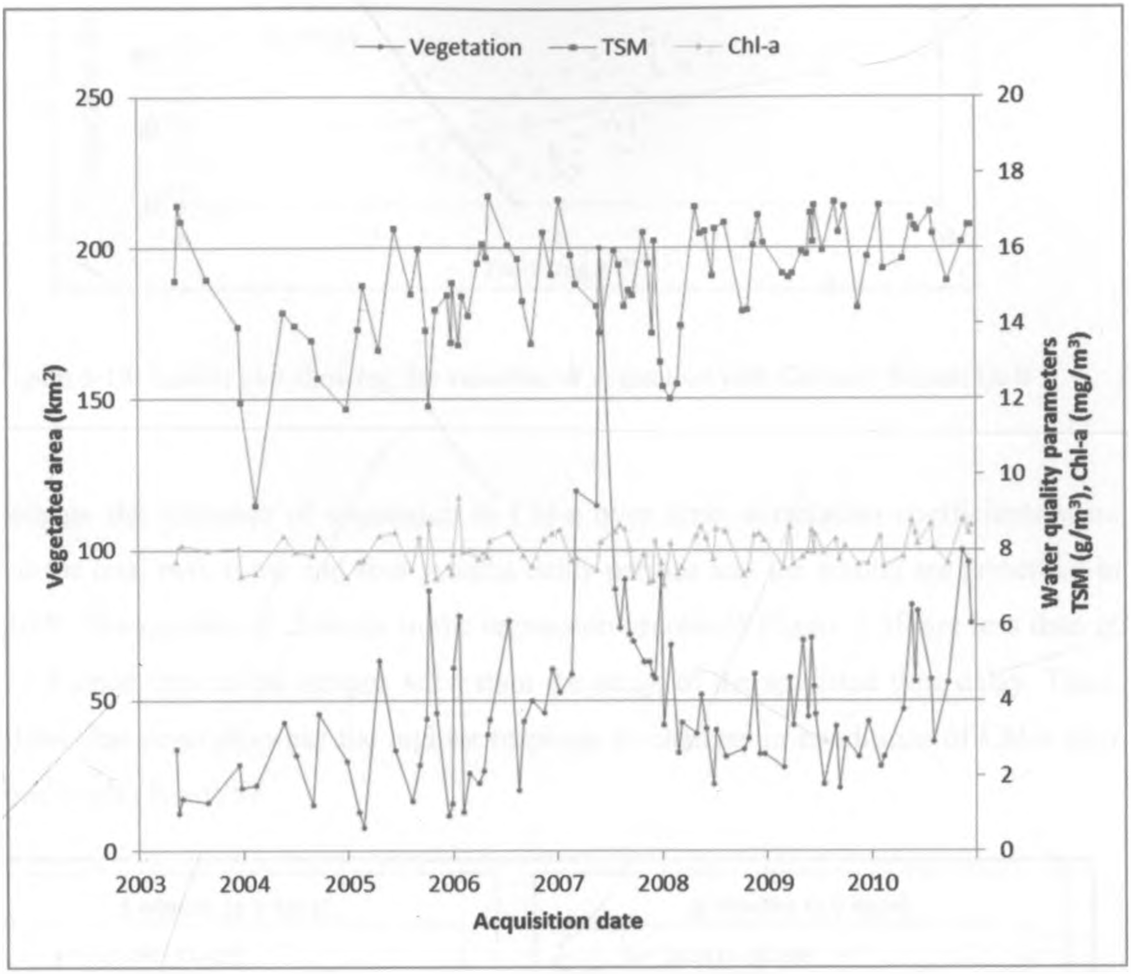

Figure 5-17: A line graph showing the time series variation of vegetation abundance and concentrations of TSM and Chl-a water quality parameters in the Winam Gulf section of Lake Victoria. A peak vegetation abundance of about 200 km^2 is observed at around June 2007

To further seek the relationship between the vegetation variation and the variation of the water quality parameters, regression analysis was conducted. Regression results show that for no time delay, vegetation abundance has no significant relationship with Chl-a over the period 2003 to 2010, with correlation coefficient of $R = 0.34$. Figure 5-18 shows the relationship between vegetation abundance and Chl-a values for no time delay. This is because vegetation would be expected to take some time to respond to the changes in the quality of water.

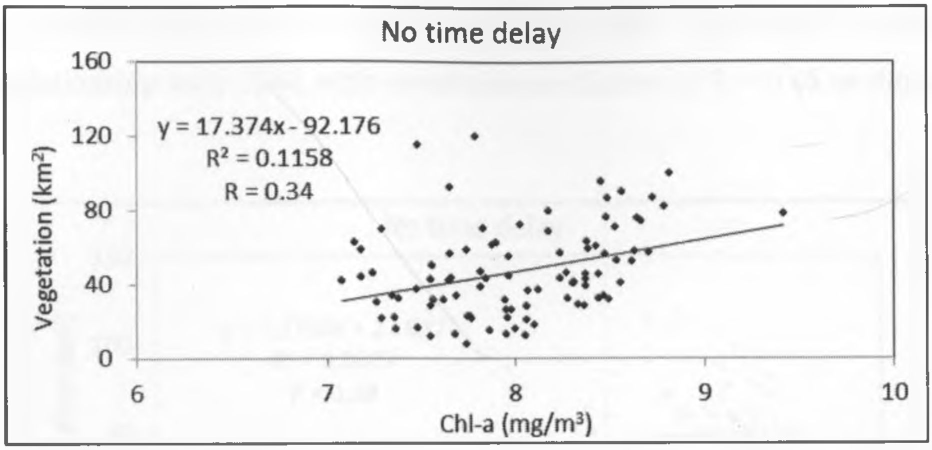

Figure 5-18: Scatter plot showing the variation of vegetation with Chl-a in Winam Gulf

To investigate the response of vegetation to Chl-a over time, correlation coefficients were computed for one, two, three and four months delay periods and the results are presented in Figure 5-19. The number of datasets in the regression graphs of Figure 5-19 are less than in Figure 5-18 since only a few images fell within the range of the specified time delay. These results show that vegetation has the highest response to changes in conditions of Chl-a after three months with $R = 0.57$.

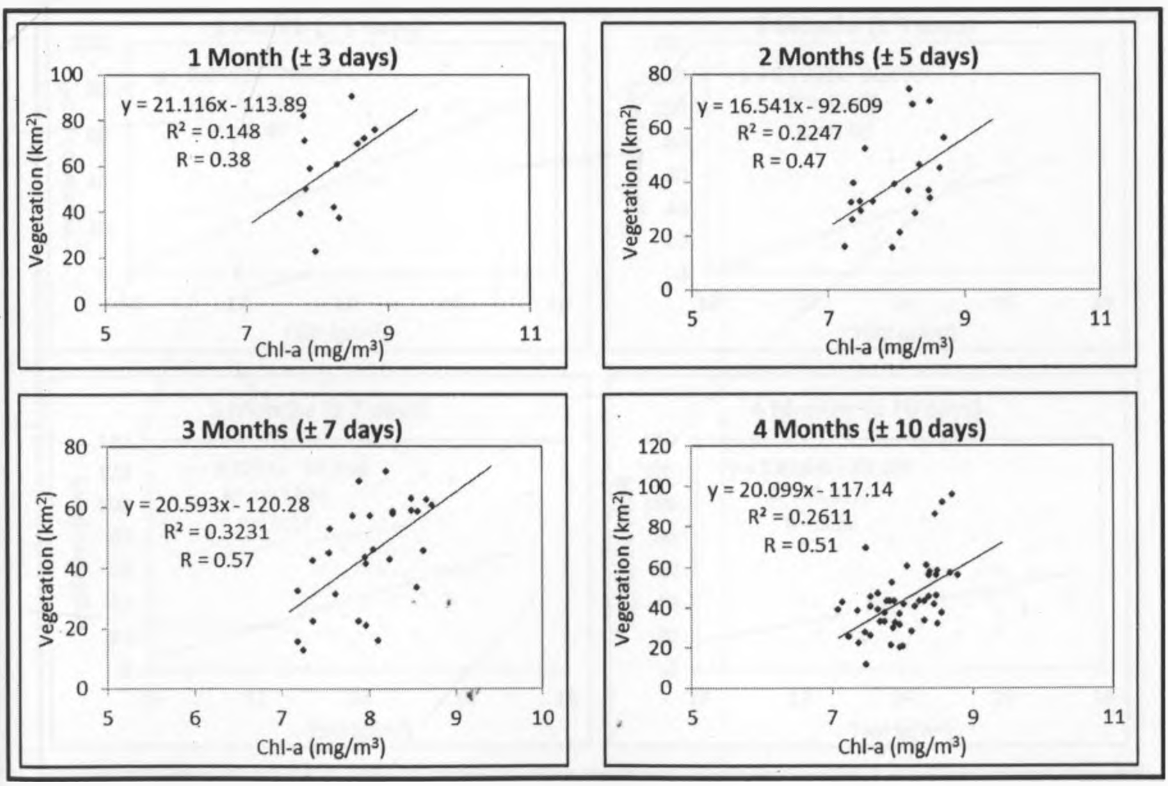

Figure 5-19: Variation of vegetation with Chl-a in Winam Gulf for various delay periods

Similarly, regression results showed that for no time delay, vegetation abundance has no significant relationship with TSM, with correlation coefficient of R = 0.08 as shown in Figure 5-20.

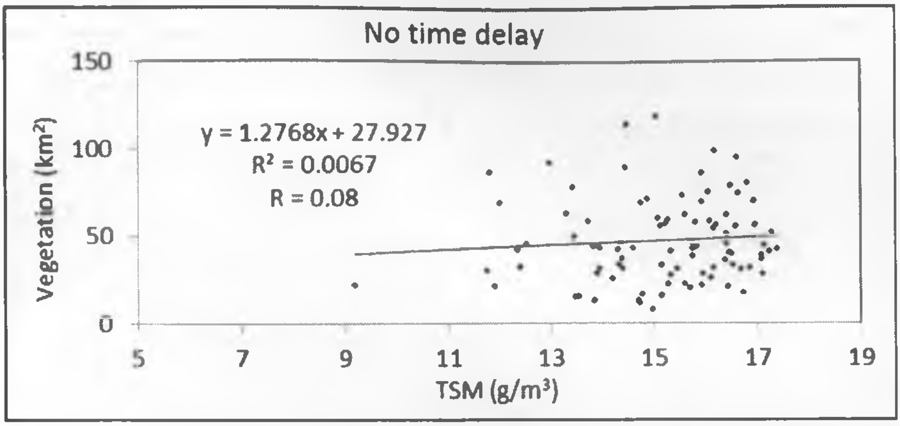

Figure 5-20: Scatter plot showing the variation of vegetation with TSM in Winam Gulf

Response of vegetation to TSM over time was investigated and the regression results are presented in Figure 5-21. These results show that vegetation has the highest response to changes in conditions of TSM after two months with R = 0.46.

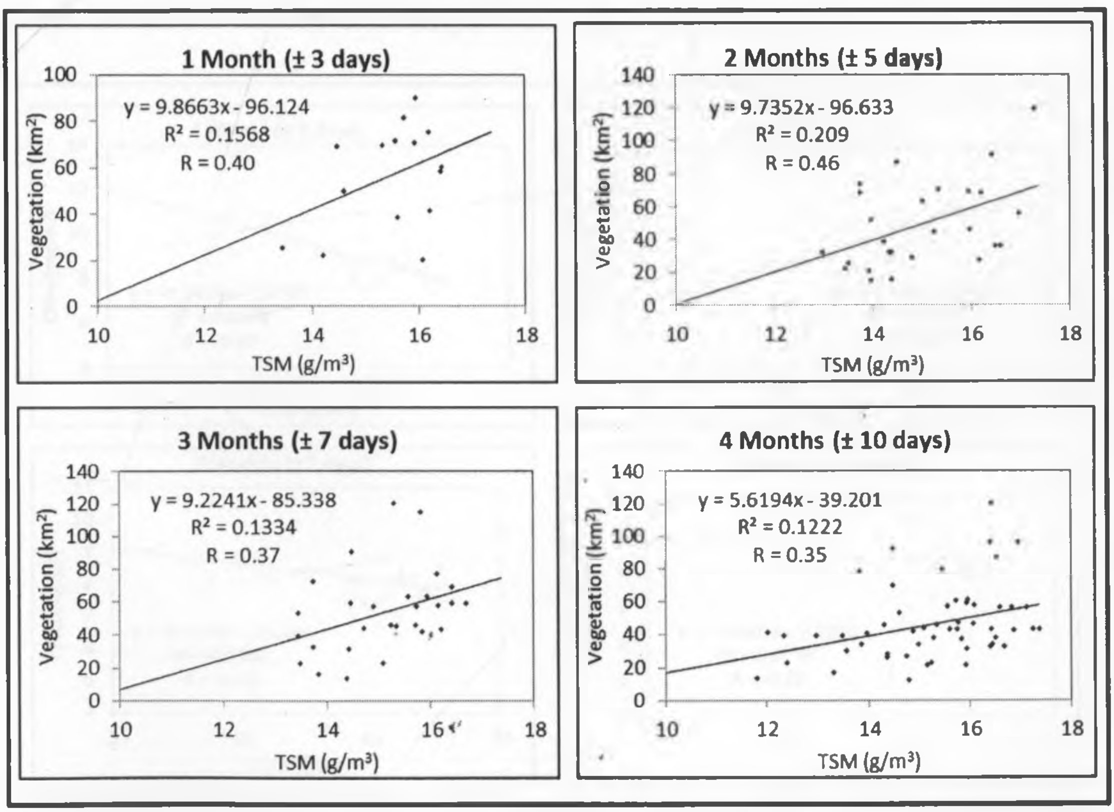

Figure 5-21: Variation of vegetation with TSM in Winam Gulf for various delay periods

With the view of determining the dependence of these water quality parameters on one another, regression analysis was carried out for Chl-a and TSM. Variation of Chl-a with TSM (Figure 5-22) was found to be linear, with correlation coefficient of R = -0.77 for no time delay.

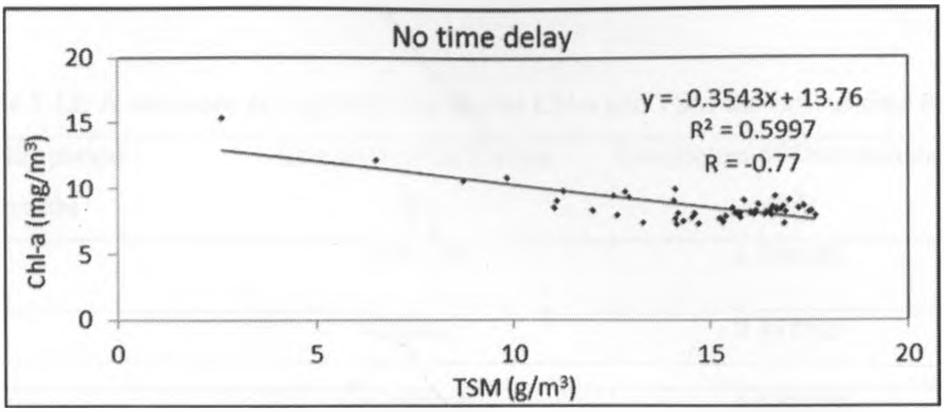

Figure 5-22: Scatter plot showing the relationship between Chl-a and TSM in Winam Gulf at no time delay

After a delay period of a few months, the dependence of Chl-a on TSM dropped gradually to R = -0.28 after four months. The linearity, however, was conserved. Figure 5-23 shows the regression results for various time delays.

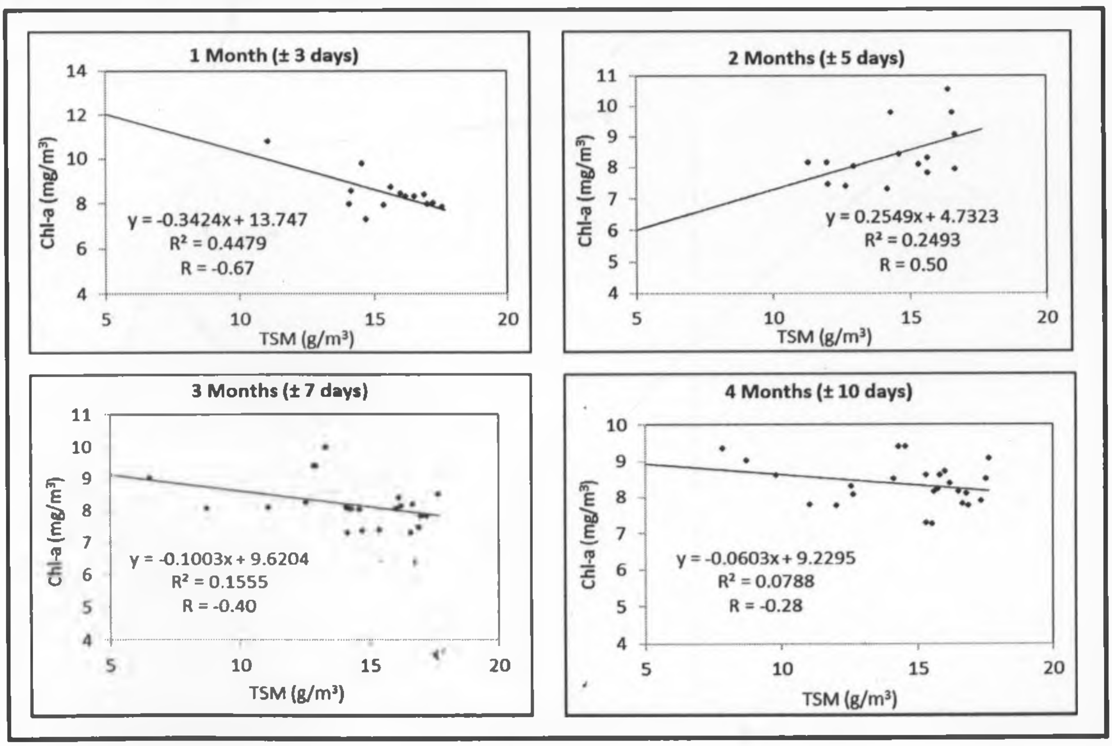

Figure 5-23: Variation of Chl-a with TSM in Winam Gulf for various delay periods

Regression results showed that for no delay period there is a fairly strong linear and inverse relationship between TSM and Chl-a, with R = -0.77. The relationship, however, gradually dropped in the subsequent delay periods to R = -0.28 by the fourth month. Table 5-13 is a summary of these regression results.

Table 5-12: A summary of regression results for Chl-a and TSM for various time delays

Delay period (months)	Correlation Coefficient (R)	Coefficient of Determination (R^2)
0	-0.77439	0.599686
1	-0.66927	0.447925
2	0.499259	0.249259
3	-0.3944	0.155549
4	-0.28071	0.078797

The possible explanation to this is that while rain water decreases the concentrations of already existing Chl-a in the lake by diluting it, run-off water sweeps sediments and nutrients into the lake, thus increasing that of TSM.

5.3.2 Rainfall

The presence of TSM in the lake is most likely to be a result of run-off water sweeping sediments into the lake during heavy rainy seasons. The role of rain in the proliferation of aquatic vegetation was investigated. Rainfall data for Kisumu rain station was obtained from the Kenya Meteorological Department (KMD) for the period 2003 to 2009. Figure 5-24 shows the variation of the vegetation with weekly average rainfall over the period.

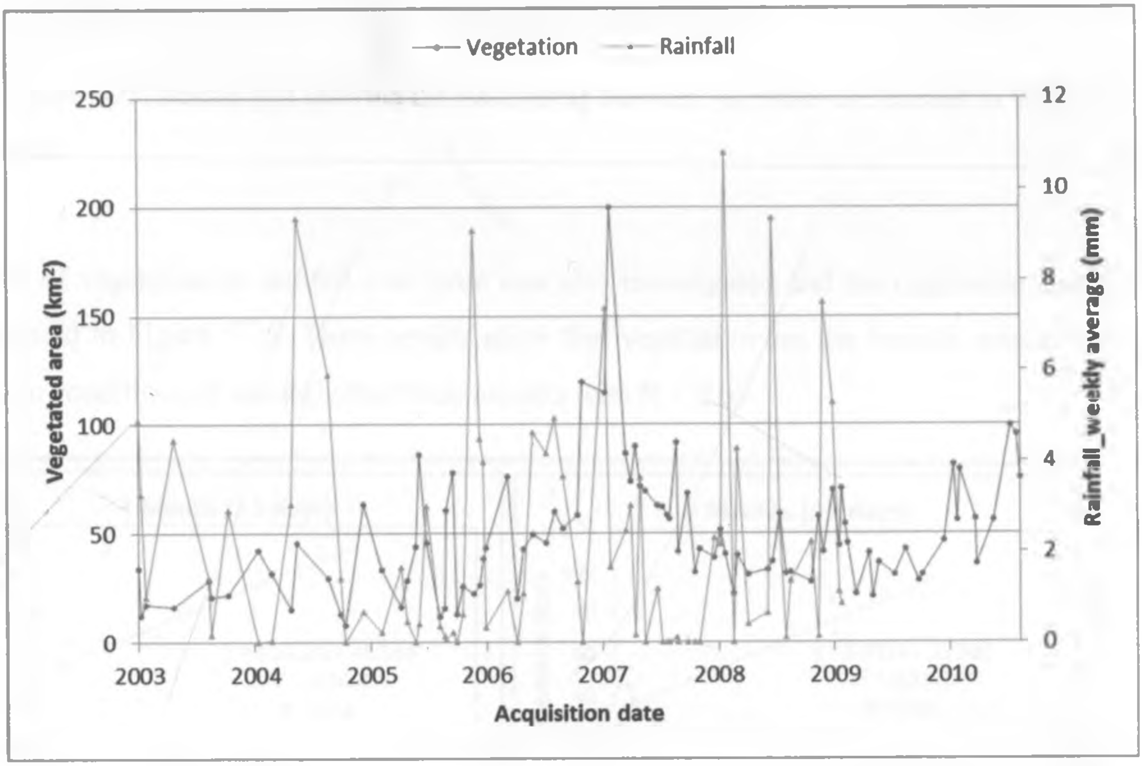

Figure 5-24: A line graph showing the time series variation of vegetation abundance with rainfall in the Winam Gulf section of Lake Victoria. Kisumu rainfall data source: Kenya Meteorological Department (KMD)

Regression analysis was conducted to further seek the relationship between vegetation variations with that of rainfall, and the results show that at no time delay, there is no significant correlation between vegetation and rainfall, with R = 0.08 as shown in Figure 5-25. This is as expected, because if vegetation proliferation is influenced by the rainfall pattern in its drainage basin, then the response would take place after some time, possibly a few months.

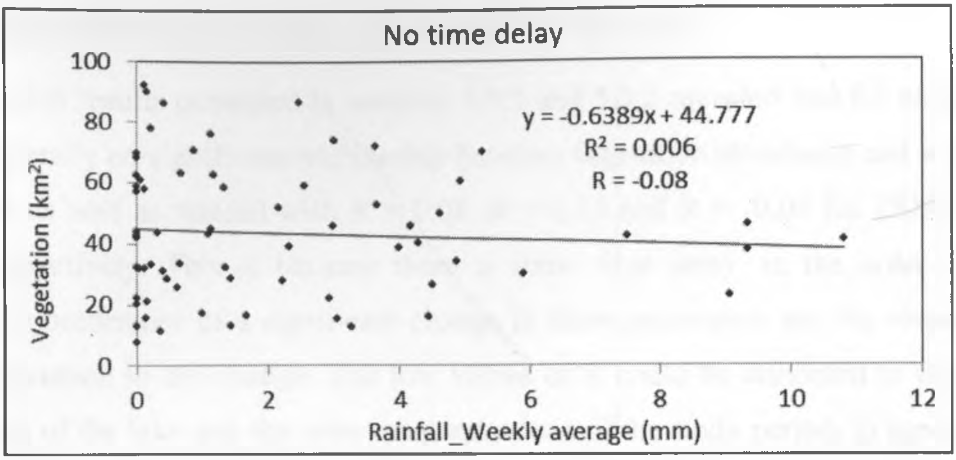

Figure 5-25: Scatter plot showing the relationship between vegetation and rainfall in Winam Gulf

Response of vegetation to rainfall over time was also investigated and the regression results are presented in Figure 5-26. These results show that vegetation has the highest response to changes in conditions of rainfall after three months with R = 0.67.

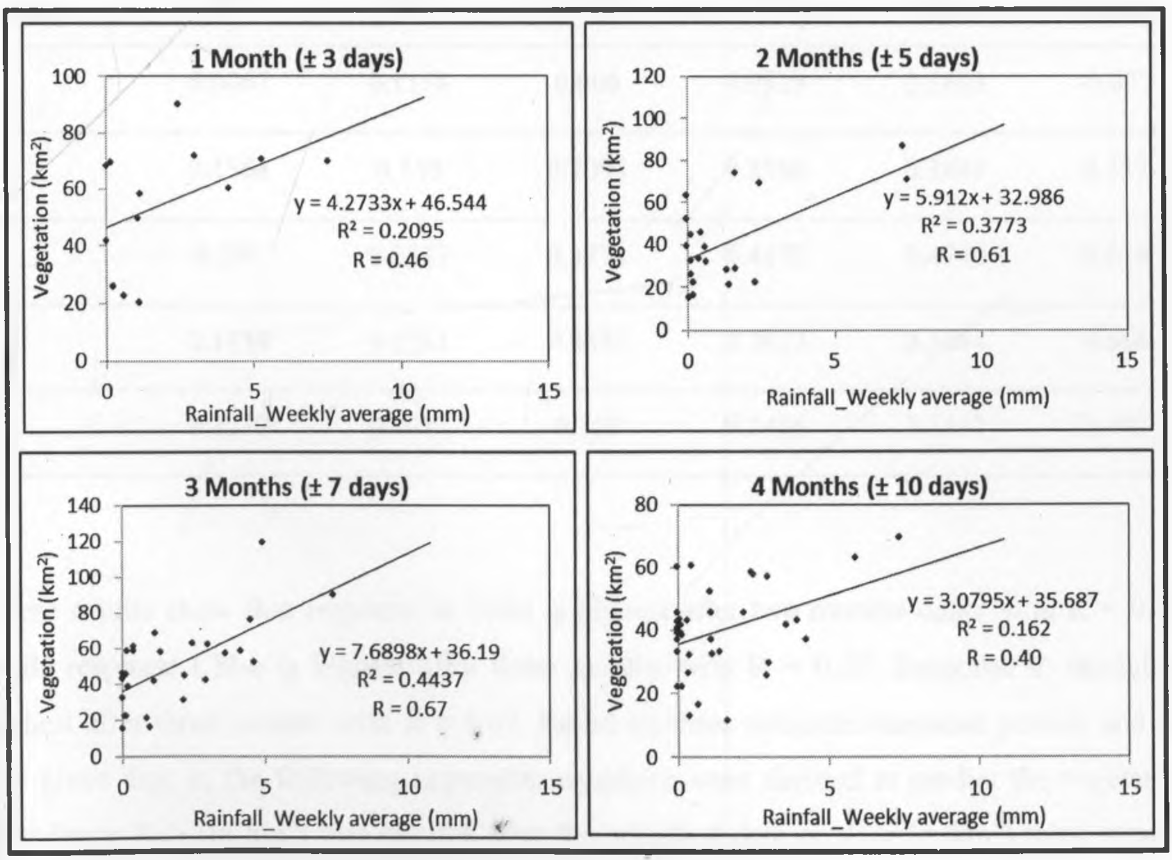

Figure 5-26: Variation of vegetation with rainfall in Winam Gulf for various delay periods

5.4 Vegetation abundance prediction models

The regression results presented in sections 5.3.1 and 5.3.2 revealed that for no time delay, there is generally no significant relationship between vegetation abundance and water quality parameters as well as rainfall with $R = 0.08$, $R = 0.34$ and $R = -0.08$ for TSM, Chl-a and rainfall respectively. This is because there is some time delay, in the order of months, between the occurrence of a significant change in these parameters and the response of the aquatic vegetation to the change. The low values of R could be attributed to the extensive surface area of the lake and the wide temporal extent of the study period, in agreement with Zhang *et al.* (2011). A summary of the regression results is presented in Table 5-12.

Table 5-13: A summary of regression results for the variation of vegetation abundance with TSM, Chl-a and rainfall for various time delays

Delay period (months)	Coefficient of Determination (R^2)			Correlation Coefficient (R)		
	TSM	Chl-a	Rainfall	TSM	Chl-a	Rainfall
0	0.0067	0.1158	0.006	0.0819	0.3403	-0.0775
1	0.1568	0.148	0.2095	0.3960	0.3847	0.4577
2	0.209	0.2247	0.3773	0.4572	0.4740	0.6142
3	0.1334	0.3231	0.4437	0.3652	0.5684	0.6661
4	0.1222	0.2611	0.162	0.3496	0.5110	0.4025

These results show that response to TSM is highest after two months delay with $R = 0.46$, while response Chl-a is highest after three months with $R = 0.57$. Response to rainfall is highest after three months with $R = 0.67$. Based on these optimum response periods and for any given day, *n*, the following regression equations were derived to predict the vegetation abundance A_{n+2} (in km^2) two months after the specified date or A_{n+3} (in km^2) three months after the specified date;

$$A_{n+2} = 9.7 \cdot TSM_n - 96.6 \quad (8)$$

$$A_{n+3} = 20.6 \cdot Chl_n - 120.3 \quad (9)$$

$$A_{n+3} = 7.7 \cdot Rain_n + 36.2 \quad (10)$$

where $TSM_{(n)}$ is the mean concentration of TSM (measured in g m^{-3}), $Chl_{(n)}$ is the mean concentration of Chl-a (measured in mg m^{-3}) while $Rain_{(n)}$ is the average weekly rainfall (measured in mm), all at the specified date, n.

The relationships between variation of aquatic vegetation with those of water quality parameters and rainfall were found to be generally weak. It is possible that vegetation growing in the Winam Gulf is carried away by winds and currents and exit through the narrow opening into the main lake. This effect lowers the coefficient values and reduces the ability to predict future occurrence of vegetation growth based on the information about the condition of the lake.

6 CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

An endmember spectral library of the predominant class features in the lake was developed by deriving their spectral response characteristics from a multispectral satellite imagery. It consists of five individual endmember files, one for vegetation and four for various water classes. The individual endmember files were computed as the mean of the spectral signatures of eight discrete pixels covering that endmember feature, each identified at one pixel accuracy. The predominant class features were identified following the results of unsupervised classification with K-Means clustering.

Spectral unmixing as a supervised classification technique was found to be very suitable for application with multispectral data with relatively low spatial resolution. This is because of the ability of the algorithm to decompose the large mixed pixels into various constituent class features. Together with the image derived endmembers, the algorithm performed very well, producing a mean classification accuracy of 99.48% based on RMSE accuracy assessment. These classification results were then presented in the spatial distribution cover maps, which revealed that the almost enclosed Winam Gulf was more severely affected by the aquatic plants infestation.

Using the classified data, the vegetation abundance (surface area coverage of aquatic plants) in the lake was estimated. Algorithm was used which took into account the fraction of vegetation in each pixel obtained from classified data and the estimated pixel size of MERIS FR imagery. The abundance results were presented in tables for the period 2003 – 2010, both for the entire lake and for Winam Gulf section.

The temporal variation of the abundance of aquatic plants in the lake (vegetation phenology) over the study period 2003 – 2010 was presented in graphs. These results showed that vegetation cover in the Winam Gulf which was kept below 100 km^2 during the years 2003 to

2006 increased to a peak of about 200 km^2 in 2007, before decreasing again to below 100 km^2 during the years 2008 to 2010.

Spatial distribution maps showed a high concentration of Chl-a and TSM at the Kenyan side of the lake, the Winam Gulf. It is also in this side of the lake that most of the aquatic vegetation was found. Regression results revealed that vegetation proliferation responds to the variations in the conditions of the water quality parameters and meteorological information after a delay period of about two to three months. The optimal response periods were found to be about two and three months for TSM and Chl-a with correlation coefficients $R = 0.46$ and $R = 0.57$ respectively, while that of rainfall was about three months with $R = 0.67$. An inverse linear relationship between Chl-a and TSM was observed, with $R = -0.77$. With these optimal response periods and their respective regression equations, vegetation abundance prediction models were developed.

6.2 Recommendations

A comparison could be made to find out that which produces better results between the image derived and field derived endmember spectral libraries. It could be also of importance to find out the level of eutrophication each of the water classes represents.

A comparison should be made between spectral unmixing and other classification techniques to ascertain the efficiency of each and determine the most appropriate method for detecting aquatic vegetation.

There is need to search for a more accurate time delay between the occurrence of a significant change in the quantities and condition of the water quality parameters and the meteorological information and, the response of the aquatic vegetation to these changes. In order to determine more accurate time delays, the effect of vegetation movement into and out of Winam Gulf (the region of interest considered in developing prediction models) as well as the human activities such as weed harvesting should be considered.

Availability of suitable data is one of the greatest challenges to the proper monitoring of vegetation proliferation in Lake Victoria, owing to the large extent of the lake and the severe clouds cover in the region. For proper monitoring of aquatic vegetation, at least a bi-weekly data frequency, and possibly acquired locally, is recommended for developing an automatic monitoring system.

7 REFERENCES

Akgün, A., Eronat, A. H. and Türk, N., 2010, Comparing Different Satellite Image Classification Methods: An Application in Ayvalik District, Western Turkey. Available online at: http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/505.pdf (accessed 7 October 2010).

Albright, T. P., Moorhouse, T.G. and McNabb, T. J., 2004, The Rise and Fall of Water Hyacinth in Lake Victoria and the Kagera River Basin, 1989-2001. *Journal of Aquatic Plant Management*, 42, pp. 73–84.

Basic ERS and ENVISAT (A)ATSR and MERIS Toolbox (BEAM), 2010, BEAM help. Available online at: http://www.brockmann-consult.de/BEAM/doc/help/index.html (accessed 22 November 2010).

DIVA-GIS website, *http://www.diva-gis.org/Data*, (accessed 3 May 2011).

Doerffer, R. and Schiller, H., 2008 (a), MERIS Lake Water Algorithm for BEAM, ATBD of bio-optical models, GKSS Research Centre 21502 Geesthacht Version 1.0, 10 June 2008. Available online at: http://www.brockmann-consult.de/beam-wiki/display/LAKES/Home (accessed 14 February 2011).

Doerffer, R. and Schiller, H., 2008 (b), MERIS Regional Coastal and Lake Case 2 Water Project - ATBD of the Atmospheric Correction, GKSS Research Centre 21502 Geesthacht Version 1.0, 18 May 2008. Available online at: http://www.brockmann-consult.de/beam-wiki/display/LAKES/Home (accessed 14 February 2011).

European Space Agency (ESA) website: http://envisat.esa.int/instruments/meris/ (accessed 12 August 2010)

Foppa, N., Wunderle, S. and Hauser, A., 2002, Spectral unmixing of NOAA-AVHRR Data for snow cover estimation. Proceedings of *EARSeL-LISSIG-Workshop, Observing our Cryosphere from Space*, 11–13 March 2002, Bern, pp. 155–162.

Gichuki, J. W., 2010; A native grass destroys Water hyacinth in Lake Victoria East Africa, Luvei Times, Kenya's Social Online Magazine. Available online at: http://www.luvei.com/?p=335 (accessed 27 July 2010).

Harrison, T. E., 2010, Atomic Structure and Processes, and the Nature of Light. Available online at: http://astronomy.nmsu.edu/tharriso/ast105/Ast105week07.html (accessed 9 September 2010).

Idawo, C. and Laneve, G., 2004, Hyperspectral Analysis of Multispectral ETM+ Data: SMA Using Spectral Field Measurements in Mapping of Emergent Macrophytes. IGARSS Conference, 20–24 September 2004, Alaska, *IEEE*, pp. 246–249.

International Lake Environment Committee (ILEC) website, *http://www.ilec.or.jp/database/afr/afr-05.html*, (accessed 23 January 2012).

Jollineau, M. and Howarth, P., 2002, Use of High-Resolution Imagery to Map Wetland Environments in South-Central Ontario, Canada, *IEEE*, 5, pp. 3089–3091.

Kahlid, S. and ConocoPhillips, 2005, Image Classification. Satellite Remote Sensing Lecture. Available online at: http://topex.ucsd.edu/rs/classification.pdf (accessed 7 October 2010).

Koponen, S., Ruiz-Verdu, A., Heege, T., Heblinski, J., Sorensen, K., Kallio, K., Pyhälahti, T., Doerffer, R., Brockmann, C. and Peters, M., 2008, Validation Report, Development of MERIS lake water algorithms. Available online at: http://www.brockmann-consult.de/beam-wiki/display/LAKES/Home (accessed 16 May 2009).

Kumar, U., Kerle, N. and Ramachandra, T. V., 2007, Constrained Linear Spectral Unmixing Technique for Regional Land Cover Mapping Using MODIS Data. *International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering (CIS²E 07)*, 3–12 December 2007, University of Bridgeport, IEEE, pp. 1–8.

Laneve, G., Cavalli, R. M., Fusilli, L., Palombo, A., Pignatti, S. and Santini, F., Lake Victoria aquatic weeds monitoring by high spatial and spectral resolution satellite imagery.

Proceedings of *Hyperspectral 2010 Workshop*, 17 – 19 March 2010, Frascati, Italy, ESA SP-683, May 2010.

LVFO, 2008. The Fisheries Management Plan for Lake Victoria 2009–2014, LVFO, Jinja.

MacQueen, J. B., 1967, Some Methods for classification and Analysis of Multivariate Observations. Proceedings of *5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1, pp. 281–297.

Mailu, A. M., Ochiel, G. R. S., Gitonga, W. and Njoka, S. W., 2000, Water hyacinth: an environmental disaster in the Winam Gulf of Lake Victoria and its control. Proceedings of the *1st IOBC water hyacinth working group*, pp. 101–10. Available online at: http://hdl.handle.net/1834/1281 (accessed 6 September 2010).

Marshall, B. E., Ezekiel, C. N., Gichuki, J., Mkumbo, O. C., Sitoki, L. and Wanda, F., 2009, Global warming is reducing thermal stability and mitigating the effects of eutrophication in Lake Victoria (East Africa). Available online at: hdl:10101/npre.2009.3726.1 (accessed 6 September 2010).

Matteucci, M., 2010, K-Means Clustering, A Tutorial on Clustering Algorithms. Available online at: http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/kmeans.html (accessed 16 September 2010).

Matthias, B. and Martin, H., 2003; Mapping imperviousness using NDVI and linear spectral unmixing of ASTER data in the Cologne-Bonn region (Germany), Proceedings of the *SPIE 10th International Symposium on Remote Sensing*, 8–12 September 2003, Barcelona, Spain.

Muli, R. J., 1996, Environmental problems in Lake Victoria (East Africa): What the international Community can do. Lakes and Reservoirs: Research and Management, Nairobi, Kenya, 2, pp. 47–53.

Osumo, W. M., 2001, Effects of water hyacinth on water quality of Winam Gulf, Lake Victoria. Kenya Marine Fisheries Research Institute, Kisumu, Kenya. Available online at: http://www.unuftp.is/static/fellows/document/osumoprf.pdf, (accessed 6 September 2010).

Qian, Y. and Ping, G., 2007, Multispectral Remote Sensing image classification with multiple features. Proceedings of *the Sixth International Conference on Machine Learning and Cybernetics,* 19–22 August 2007, Hong Kong, IEEE, pp. 360–365.

Ramirez, J. S., 2006, Monitoring of Wetland Vegetation in Lake Cuitzeo, Mexico. MSc. thesis, International Institute for Geo-information Science and Earth Observation, Enschede, The Netherlands.

Rahman, H. and Dedieu, G., 1994, SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum, International Journal of Remote Sensing, 15, pp. 123–143.

Schouten, L. S. M., van Leeuwen, H. J. C., Bakker, J. G. M and Twongo, T., 1999, Water Hyacinth Detection in Lake Victoria by Means of Satellite SAR, *USP–2 report*, Netherlands Remote Sensing Board, Delft, The Netherlands, USP–2, pp. 98–128.

Sotis, G., 2007, ENVISAT-1 products specifications, Volume 11: MERIS products specifications, Issue 5. European Space Agency (ESA) website. Available online at: *http://earth.esa.int/pub/ESA_DOC/ENVISAT/Vol11_Meris_5b.pdf*, (accessed 19 July 2011).

Williams, A. E., Hecky, R . E. and Duthie, H. C., 2007, Water hyacinth decline across Lake Victoria—Was it caused by climatic perturbation or biological control? A reply, *Science Direct, Aquatic Botany,* 87, pp. 94–96.

Wilson, J. R. U., Ajuonu, O., Center, T. D., Hill, M. P., Julien, M. H., Katagira, F. F., Neuenschwander, P., Njoka, S. W., Ogwang, J., Reeder, R. H. and Van, T., 2007, The decline of water hyacinth on Lake Victoria was due to biological control by Neochetina spp., *Science Direct, Aquatic Botany,* 87, pp. 90–93.

World Geodetic System (WGS) website of the National Geospatial-Intelligence Agency (NGA), *https://www1.nga.mil/ProductsServices/GeodesyGeophysics/WorldGeodeticSystem/Pages/default.aspx*, (accessed 27 April 2011).

Zhang, Y., Lin, H., Chen, C., Chen, L., Zhang, B. and Gitelson, A. A., 2011, Estimation of chlorophyll-a concentration in estuarine waters: case study of the Pearl River estuary, South China Sea, IOP Science, Environmental Research Letters, Vol. 6, No. 2.

Zhigang, X. and Zhuang, D., 2007; The methodology of detailed vegetation classification based on environmental knowledge and remote sensing images, IEEE, pp. 2074–2077.