REMOTE SENSING AND PANEL DATA MODELS FOR ASSESSING WATER RESOURCES

JANUARY, 2014

MAGUMBA DAVID

Graduate School of Horticulture

CHIBA UNIVERSITY, JAPAN

Remote Sensing and Panel Data Models for Assessing Water Resources

January, 2014

A thesis Submitted to the Graduate School of Horticulture, Chiba University in partial fulfillment for the award of the Degree of

Doctor of Philosophy (Ph.D) Food and Resource Economics.

Magumba David

Graduate School of Horticulture,

CHIBA UNIVERSITY

Thesis defended on 31st January, 2014

at Chiba University, Japan

PREFACE

This PhD thesis is submitted to Chiba University in partial fulfillment for the award of a degree of Doctor of Philosophy (Ph.D) Food and Resource Economics. The study presents advanced techniques for assessment of water resources. Remote sensing is applied to study the time-series change caused by development in Kampala and Entebbe from Landsat (1995-2010). RapidEye satellite image of 2011 and field data were used to estimate the extent of *Cyperus papyrus* for Entebbe, Uganda at the Northern shore of Lake Victoria the second largest freshwater lake in the world. In water resource assessments, models are used in forecasting, this thesis presents new panel data models, the random effect and dynamic panel models were developed from a Cross-Country data-set of European lakes in 18 countries (1965-2009). Previous parameter structures are radically changed by the results of this thesis. The models were tested using 8 Japanese lakes (2000-2009). The thesis is divided into five chapters. Chapter one introduces the problem and the techniques for water resource assessments. Chapter two highlights the environmental pressures that are faced by lakes with reference to Lake Victoria and Lake Naivasha. A review of remote sensing with its application in deriving lake models and monitoring of spatio-temporal changes in wetlands is presented in chapter three. Chapter four presents panel data models which will change the estimations of Chlorophyll-a after publication of thesis, performance of the models are tested in this section. Finally chapter five states the conclusions and recommendations in light of the results, suggesting possibilities of applying the coefficients from this study in lake management strategies and the policy issues regarding phosphate and nitrate fertilizers.

ACKNOWLEDGEMENTS

I am very grateful to my family for encouraging me to pursue this advanced degree and for their support during the program. Great thanks to my academic advisors Prof. Michiko Takagaki, Assoc. Prof. Atsushi Maruyama, Dr. Akira Katou and Prof. Masao Kikuchi (Professor Emeritus) for their time to discuss the details of this research, review the articles and thesis sometimes past midnight.

Studies and research were only successful through MONBUKAGAKUSHO and JASSO scholarships for 4.5 years and KAKENHI research fund provided by the Japanese Government, we are grateful for the financial support, part of which enabled me to conduct one of the studies in my home country, Uganda. Thanks to Chiba University International Support Division and the Administration of The Graduate School of Horticulture that made our stay at Chiba University comfortable and peaceful.

Landsat images are used in the third chapter of this thesis, we appreciate the work done at USGS. Part of this research uses a huge online database provided by the European Environment Agency (EEA), we thank the Agency and Japanese Ministry of Environment, FloridaLAKEWATCH USA, Department of the Environment-Australia, for the data used in the simulations and Ministry of Water and Environment-Uganda for providing data on Lake Victoria.

Finally, I thank my friends in Japan and Uganda who made my stay interesting and enjoyable.

Magumba David

Chiba University, Japan

January, 2014

DECLARATION

I declare that this PhD thesis is original and was not presented in any other University for any degree. The program was conducted at Chiba University, Japan.

__________________________ Date _____________________

MAGUMBA DAVID

APPROVAL

ACADEMIC ADVISORS

Date

PROF. MICHIKO TAKAGAKI
(SUPERVISOR: LIFE CYCLE ASSESSMENT)

PROF. EMERITUS MASAO KIKUCHI
(CO-SUPERVISOR: AGRICULTURAL ECONOMICS)

ASSOC. PROF. ATSUSHI MARUYAMA
(CO-SUPERVISOR: ECONOMETRICS)

DR. KATOU AKIRA
(CO-SUPERVISOR: REMOTE SENSING)

THESIS COMMITTEE

______________________________ Date ______________________

PROF. MARUO TORU
(CHAIRPERSON)

PROF. MICHIKO TAKAGAKI

ASSOC. PROF. ATSUSHI MARUYAMA

ASSOC. PROF. KURIHARA SHINICHI

DR. KATOU AKIRA

ABSTRACT

The World's lakes and reservoirs have experienced a lot of environmental pressure among which is the effect of phosphate and nitrogenous fertilizers a global water resource problem. The thesis presents two methodologies for assessment of water resources; remote sensing and panel data analysis. Lake Victoria the second largest freshwater lake in the World has undergone diverse changes in the past 50 years due to introduction of Nile perch and flower production which started in the 1990s. An ecologically important plant species *Cyperus papyrus* (papyrus) surrounds Lake Victoria. Remote sensing is applied to estimate and identify, for Entebbe and Kampala areas on the northern shore of Lake Victoria, the distribution of papyrus wetland and its temporal change, using RapidEye and Landsat satellite images for the last 15 years. Results of RapidEye reveal that in 2011, 30% of Entebbe area was occupied by wetland, of which 70% was papyrus-covered. Urban land use increased in Kampala from 17% to 64%, and from 9% to 23% for Entebbe, with relatively low encroachment on wetlands until the mid-2000's. However, urban expansion in recent years has reached a stage to encroach wetlands. In assessment of water resources, it is valuable to grasp the global functional relationship between phytoplankton biomass (Chlorophyll-a; Chl-a), total phosphorous (TP) and total nitrogen (TN) in lake ecosystems. A comprehensive model was developed that explains the relationship between Chl-a, TP and TN in lakes under a wide range of environments. The conventional Ordinary Least Squares (OLS) model, random effect panel model and dynamic panel model are compared. Estimation based on water quality data for 396 lakes in 18 European countries from 1965-2009 show that TP and TN are significant determinants of Chl-a in the OLS model. Application of the non-conventional estimation method alters this parameter structure radically. The inclusion of auto-regressive effects makes TN insignificant. These models were tested by simulating the relationship for 9 Japanese lakes to show the superior performance of the non-conventional dynamic model.

TABLE OF CONTENTS

ACRONYMS

BBC	British Broadcasting Corporation
Coef	Coefficient
DT	Decision Trees
EEA	European Environment Agency
GIS	Geographic Information System
GLS	Generalized Least Squares
JICA	Japan International Cooperation Agency
KFC	Kenya Flower Council
ln Chl-a	Natural logarithm of Chlorophyll-a
ln TN	Natural logarithm of Total Nitrogen
ln TP	Natural logarithm of Total Phosphorus
LUD	Land Use Dummy
LVEMP	Lake Victoria Environmental Management Project
MLC	Maximum Likelihood Classification
MODIS	Moderate Resolution Imaging Spectroradiometer
MWE	Ministry of Water and Environment, Uganda
NaFIRRI	National Fisheries Resources Research Institute
NARO	National Agricultural Research Organisation
NEMA	National Environment Authority
NFA	National Forestry Authority
OLS	Ordinary Least Squares
SDT	Secchi Disk Transparency
SE	Standard error
SMD	Survey and Mapping Department
SVM	Support Vector Machines
TM	Thematic Mapper
UBOS	Uganda Bureau of Statistics
UEPB	Uganda Export Promotion Board
UFEA	Uganda Flower Exporter’s Association
USGS	United States Geological Survey
VicRes	Lake Victoria Research Initiative

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

1.1 Water and it's Assessment

The estimated volume of World's Water is 1.4 billion km^3, only 2.5 % is freshwater and most is salty (UNESCO, 2013). Lakes and reservoirs as major sources of water are under Agricultural related environmental pressure. Agriculture accounts for about 85 % of world's water use (Pfister et al., 2011) mostly for irrigation. Plant nutrition plays a central role in Crop production to provide food for the increasing Global Population. Phosphate and Nitrogenous fertilizers will continue to be used as sources of Phosphorus and Nitrogen.

Phosphate and Nitrogenous fertilizers are reported to have caused changes in the status of water resources by increasing phytoplankton biomass (e.g., Sebilo et al., 2013; Lundy et al., 2012; Nangia et al., 2010). Chlorophyll-a is the most popular indicator of phytoplankton biomass (Søndergaard et al., 2011; Carvalho et al., 2009), the other being Secchi-disk transparency (SDT) mostly applied in remote sensing of water clarity (Olmanson et al., 2008). This thesis presents an application of remote sensing for monitoring spatio-temporal changes in land cover and advanced statistical approach for assessment of water resources. Advanced statistical modeling specifically panel data analysis is done by developing random and dynamic panel models to estimate parameters for total nitrogen and total phosphorus in influencing lake Chlorophyll-a.

Statistical models for estimating and predicting Chlorophyll-a from nutrient data have been presented for several years (Bachmann et al., 2012; Reckhow, 1993), with the first preposition of Chlorophyll-nutrient relationship made by Sakamoto in 1966. Previous authors have presented coefficients in chlorophyll-nutrient relationships but the common estimation method is Ordinary Least Squares (OLS) (e.g., Bachmann et al., 2012; Huszar et al., 2006). Predictive models that apply OLS ignore the station-specific effects (sampling site) and previous concentrations of the Chlorophyll-a. Panel data analysis estimates a chlorophyll-nutrient relationship where these issues are considered, random-effect model for the former and dynamic panel model for the latter.

It is established in literature that phosphorus and nitrogen are the major nutrients that influence Chlorophyll-a in lakes (Abell et al., 2012; Lv et al., 2011; Carvalho et al., 2009; Brown et al., 2000). Phosphorus has higher coefficients of over 0.6 and nitrogen with an average of 0.4. Though fewer authors have reported higher coefficients for nitrogen than phosphorus, such results are notable of a single predictor variable e.g., Trevisan and Forsberg, (2007) and most of these relationships are modeled using OLS estimation.

Panel data (longitudinal or cross-sectional time-series data) enables study of different lake stations (sampling sites) over time. Panel data analysis controls for station heterogeneity which changes at specific stations but not across stations. Panel data models are commonly developed as fixed-effects and random effects models. Since we are interested in general inferences, it is justifiable to use the random effect's model. The detailed methodology for panel data analysis is explained by Hsiao (2007) and Oscar (N.d). Dynamic panel models are based on auto-regressive effects of Chlorophyll-a as a variable and other predictors to estimate its current concentration. The current concentration of Chlorophyll-a depends on the concentrations of

Chlorophyll-a in the earlier years (lags). Panel data analysis has been applied mostly in financial and economic studies. The major advantage of panel data analysis – random effect model and dynamic panel model is the ability to control station-specific effects and to incorporate previous lags in the regression equation.

Table 1.1 Applications of panel data analysis

Theme	Scope	Highlights	Author(s)
Capital structure of firms	7 Central and Eastern European Countries (CEE)	Cash flow is a significant determinant of firm leverage	Mateev et al., 2013
Fuel demand	Brazil	The market for ethanol is very dynamic and its demand is elastic compared to gasoline.	Santoso, 2013
CO_2 emissions	12 Countries in the Middle East	Energy consumption, FDI, GDP and trade determined CO_2 emissions.	Al-mulali, 2012
Renewable energy	24 European Countries	Traditional energy sources mainly control the rate of change to renewable energy sources.	Marques and Fuinhas, 2011
Exchange rate and Foreign Direct Investment (FDI)	9 Asian Economies	FDI increased with both higher value of the yen and exchange rate	Takagi and Shi, 2011
Financial development	G-7, Europe, East Asia and Latin America	Real income per capita and institutional quality are the determinants of Banking sector and capital market developments	Law and Habibullah, 2009
Financial factors for FDI	US, UK, Japan and Germany	Cointegrating relationship exist between FDI and real exchange rates	Choi and Jeon, 2007

1.2 Hypothesis

The recent developments in Kampala and Entebbe, Uganda have altered land use and land cover with implications for management of Lake Victoria, determination of the changes are feasible using satellite images.

Panel and dynamic panel models are statistically and geographically better performing models than the pooled Ordinary Least Squares model.

1.3 Objectives of the Study

The objectives of the thesis are to determine the spatio-temporal changes in wetlands for Entebbe and Kampala areas surrounding Lake Victoria and, estimate the parameters for total phosphorus and total nitrogen as determinants of Chlorophyll-a in the Chl-a = f(TP,TN) relationship from panel data analysis and test the robustness of the models by predicting the level of Chlorophyll-a in other lakes.

CONCEPTUAL FRAMEWORK

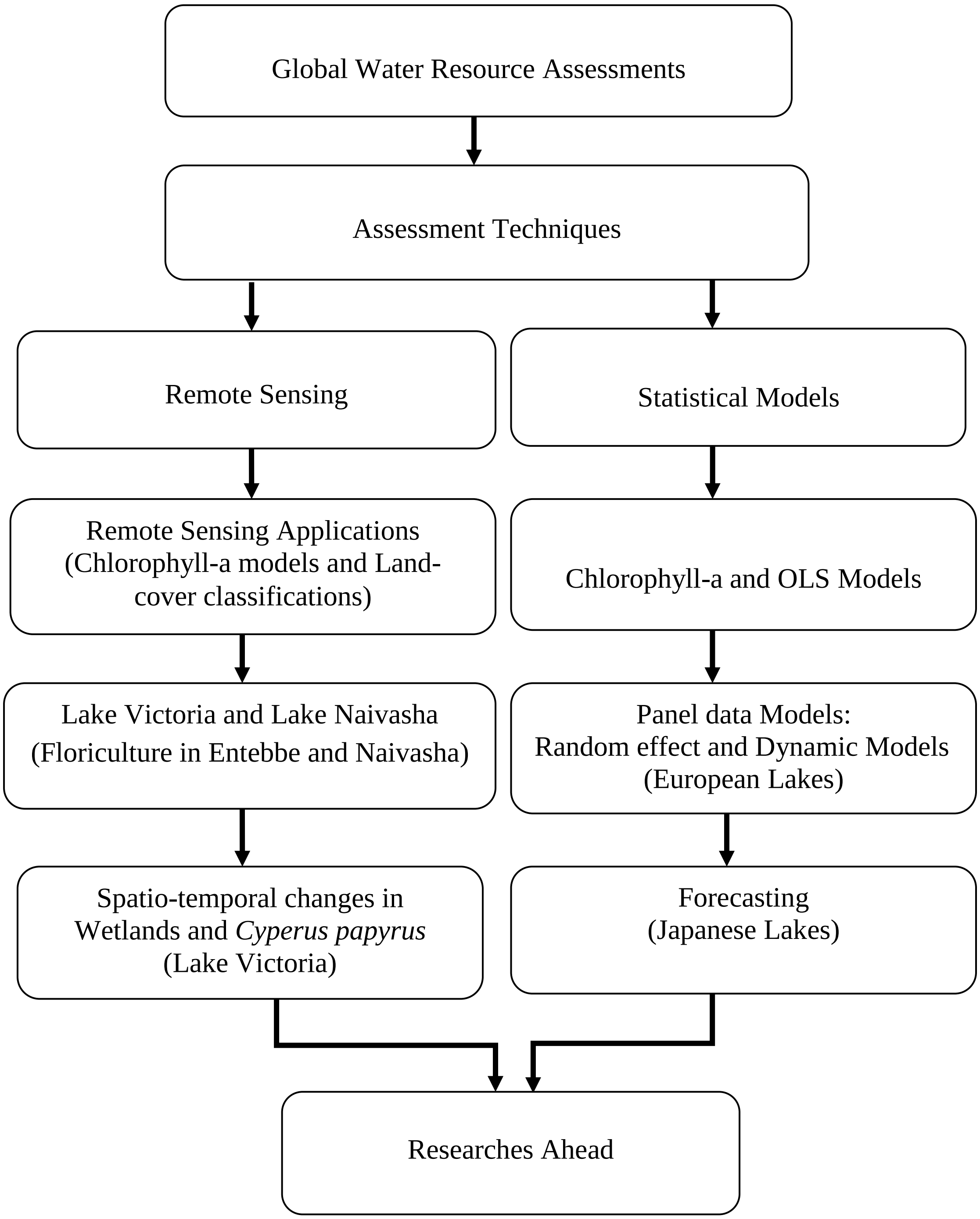


Fig. 1.1 Thesis conceptual framework

CHAPTER II

ENVIRONMENTAL PRESSURE ON LAKES: N AND P FERTILIZERS

2.1 THE CASE OF LAKE VICTORIA

2.1.1 Overview of Lake Victoria

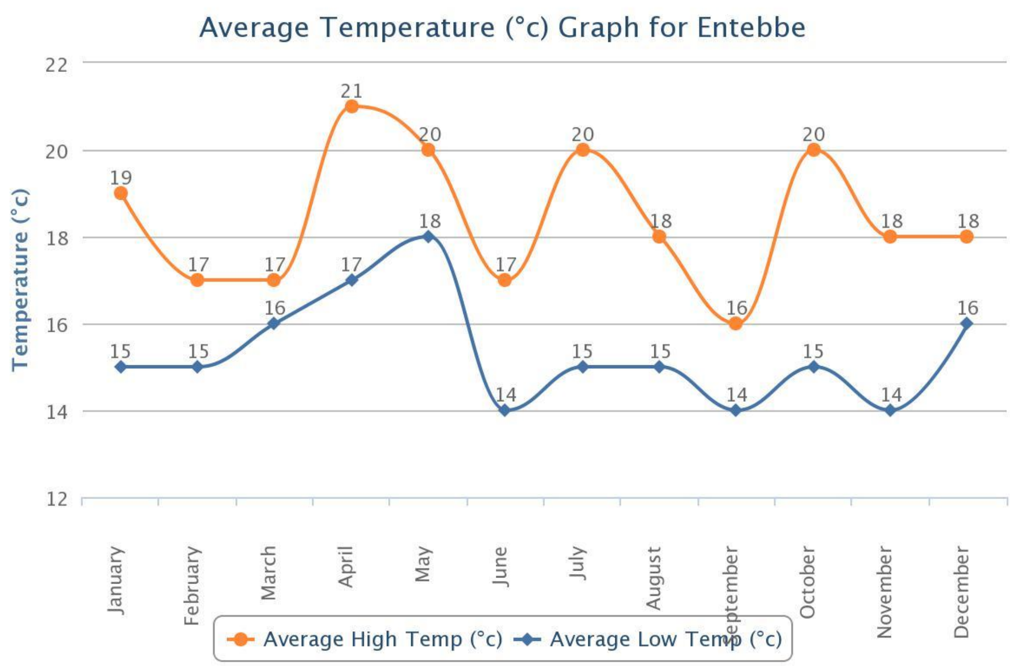
Lake Victoria (Nyanza, Ukerewe, Nalubaale) the second largest freshwater lake in the world with a surface area of 68,800 km^2, 400 km long and 320 km wide and its maximum and avearge depth of 83 meters and of 40 meters is a transboundary lake shared by three countries, Uganda (45%), Kenya (6%) and Tanzania (49%) (Muyodi et al., 2010). The Equator crosses the Northern tip of the lake near Entebbe International Airport at an altitude of 1133 meters. Lake Victoria has experienced three major environmental challenges, first was the introduction of Nile perch in 1950s, a predator which drastically reduced the population of Cichlids, endemic to Lake Victoria, Tanganyika and Malawi (Turner et al., 2001), though the reason was economically justifiable. Second, water hyacinth (*Eichhornia crassipes*) problem in the late 1980`s which lasted for over a decade (Opande et al., 2004; Kateregga and Sterner, 2007).Third, the flower farms established in the 1990s within Entebbe area which are fertilizer intensive systems and expansion of some farms resulted into uitilizing *Cyperus papyrus* area. In this thesis, the distribution of papyrus in 2011 and spatio-temporal changes in wetlands around Lake

Victoria are estimated. A brief background to fish exports, floriculture and management of Lake Victoria are presented in this section.

Fig. 2.1 Topographic map of East Africa showing Lake Victoria at the center of the countries, a section from ArcGIS online basemap

2.1.2 Temperature and Rainfall for Entebbe

A



B

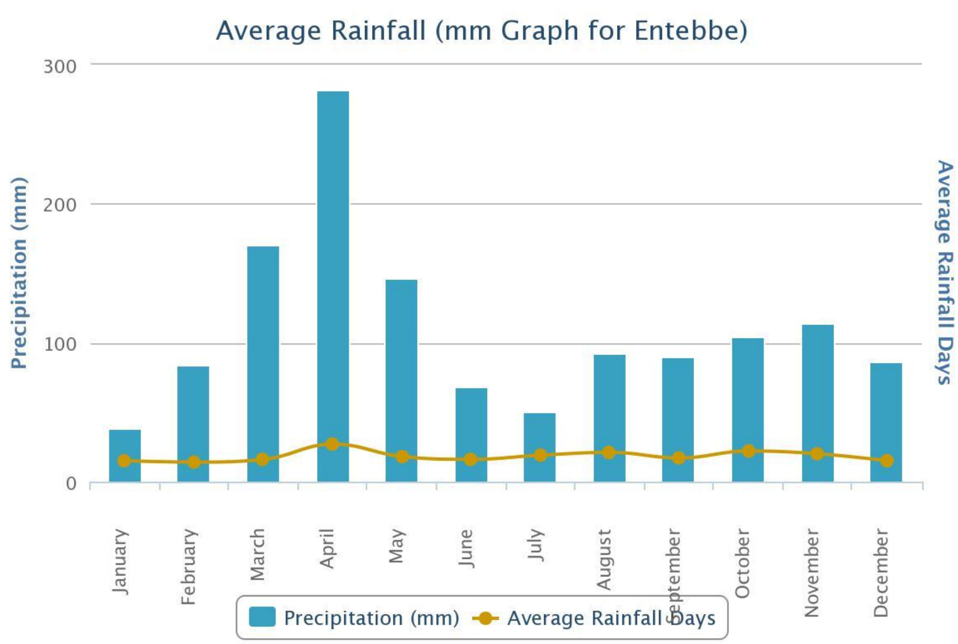


Fig. 2.2 Temperature and rainfall for Entebbe, Uganda
Source: Worldweatheronline
http://www.worldweatheronline.com/v2/weather-averages.aspx?q=EBB

2.1.3 Uganda Fish Exports

It is estimated that over 30 million people derive their livelihood from Lake Victoria. The most quantifiable economic value are exports of fish and fish products which was at 128 million US $ in 2010 (UEPB, 2013). In the biodiversity, Lake Victoria is home to over 700 endemic Cichlids (Turner et al., 2001), the introduction of Nile perch (the predator) in 1950s significantly altered their population structure.

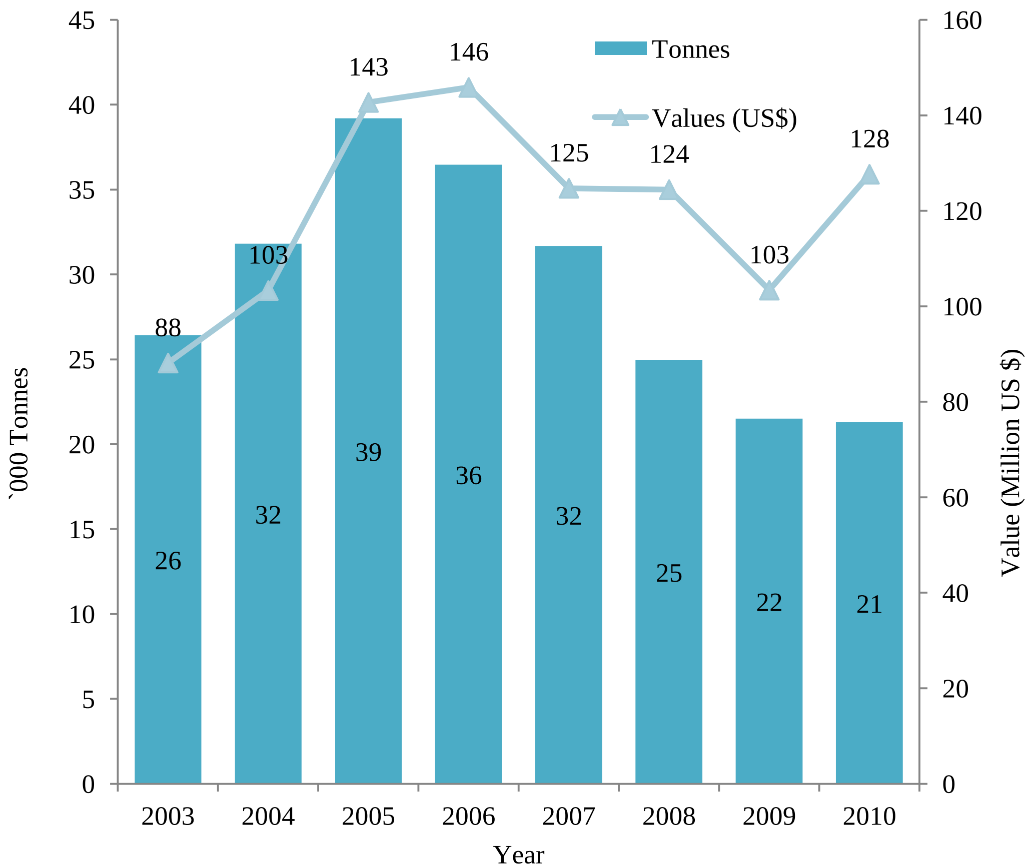


Fig. 2.3 Exports of Fish and Fish Products
Source: UEPB

Fig. 2.4 Fish from Lake Victoria at Kasenyi Landing Site.
Photograph taken during Geo-data collection in 2012.

2.1.4 Uganda Floriculture Industry

Commercial flower production in Uganda is the leading horticultural export cluster from mainly roses and chrysanthemum cuttings, it was valued in 2010 to have exceeded 22 million US $ (UEPB, 2013). The roses are mainly exported to the Flower Auction in the Netherlands. Having started in 1993, peak volume of exports was achieved in 2005. The sector employs 85,000 people and in 2010 there were 20 commercial flower farms though the number has reduced in the recent years. The global economic crisis of 2008 had tremedous effects on the sector.

TANK A
Calcium Nitrate
EDDHA
EDTA
Microfeed

TANK B
Potassiumnitrate
Potassiumsulphate
Magnesiumsulphate
Magnesiumnitrate
Ammoniumnitrate
Monopot.phosphate
Nitric Acid
Fosforic Acid 75%

Fig 2.5. Fertilizers for Flower production. Ethylenediamine bis(2-hydroxyphenyl)acetic acid (EDDHA) and Ethylenediaminetetraacetic acid (EDTA) are chelating agents used to supply and correct iron deficiency.
Source: Adapted from a Photograph taken during a flower farm visit in November, 2010.

Fig. 2.6 Roses produced in Uganda
Source: UFEA

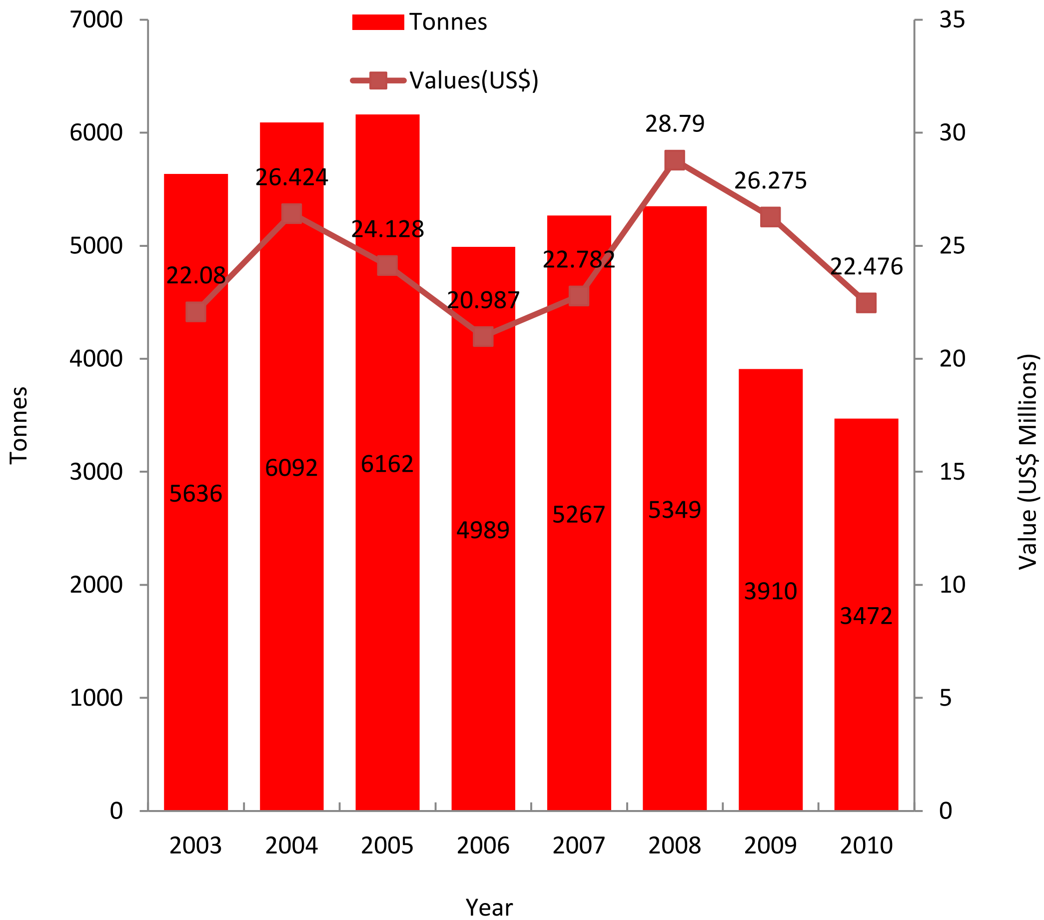


Fig. 2.7 Exports of Flowers and Cuttings, 2003-2010
Source: UEPB

2.1.5 Export Share of Horticultural Crops

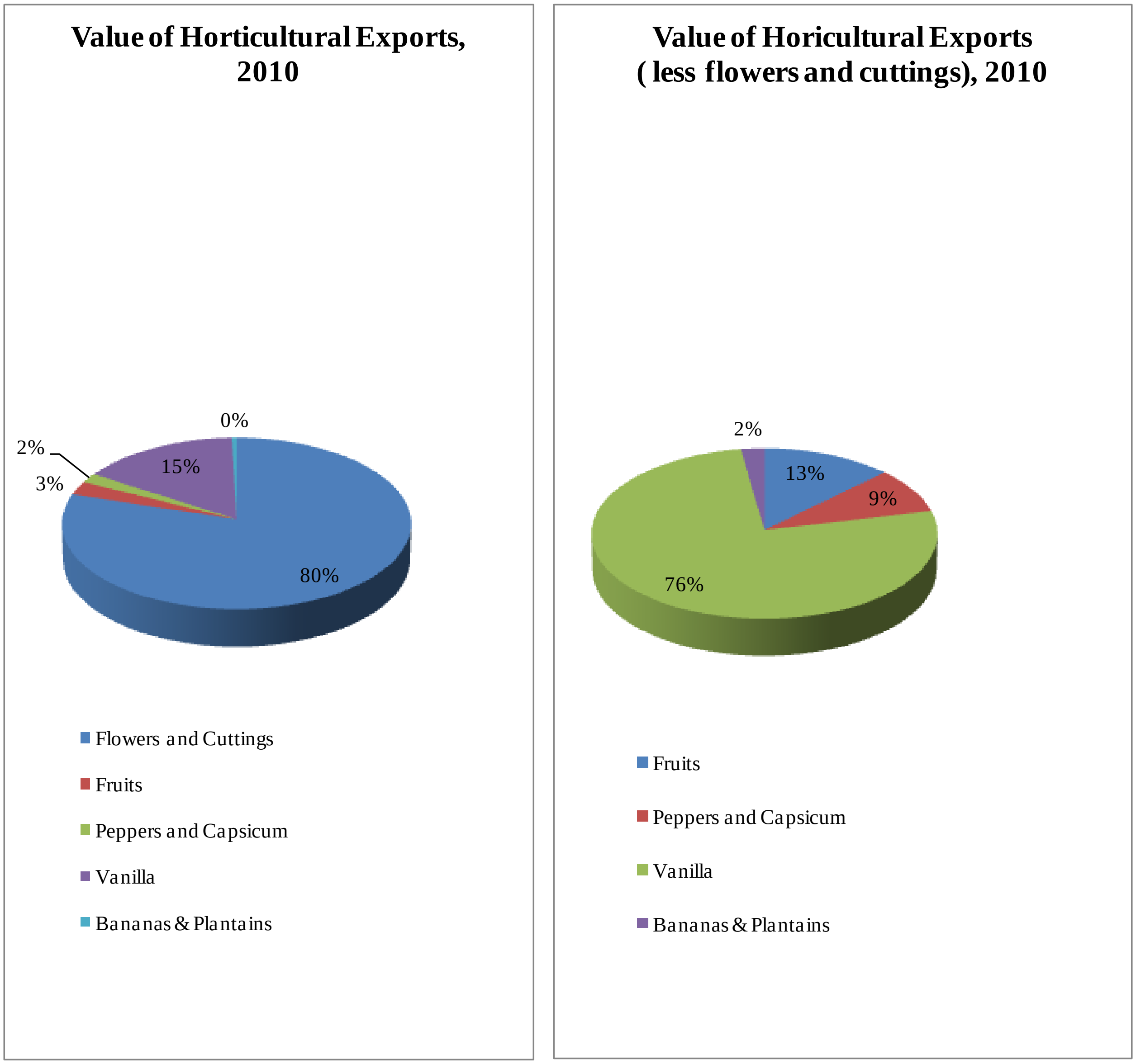


Fig. 2.8 Share of horticultural exports, 2010.
Source: UEPB

2.1.6 Management of Lake Victoria

Lake Victoria is managed by a variety of organisations, this section introduces a few of those organisations. Historical and current research and management projects were described by Muyodi et al. (2010). Ministry of Water and Environment is the government organisation responsible for the general management of water resources and the environment in Uganda. The divisions include Directorate of Environmental Affairs, Water Development and Water Resource Management.

National Environment Management Authority (NEMA) is entrusted with the role of public communication and implementing programs that protect the environment. NEMA handles Environmental Impact Assessments of projects and also issues permits for environmentally sound development projects.

With its history starting in 1947 first as the East African Fisheries Research Organisation (EAFRO), now National Fisheries Resources Research Institute (NaFIRRI), was established by the National Agricultural Research Act of 2005. The institute is part of the National Agricultural Research Organisation (NARO). NAFFIRI programs are focused on fisheries and aquacultural sector development.

Lake Victoria Environment Management Program (LVEMP) a collaborative program which started in 1997 with second phase expected to end in 2015, has the objectives to strengthen collaborations for management of Lake Victoria within East Africa and reduce environmental stress to Lake Victoria.

Established in 2002, Lake Victoria Research (VicRes) Initiative is a multidisciplinary research initiative within East Africa under the Inter-University Council for East Africa (IUCEA). The phases of activities for the initiative started in 2003 and the current phase closing in 2014. The program is funded by the Swedish Government (SIDA). In management of Lake Victoria and associated natural resources, research funds are granted to researchers for projects within the Lake Victoria region in Ethnobotany, fisheries and aquaculture and natural resource management, as of 2012, 102 projects were funded by SIDA.

Table 2.1 Summary of data for Lake Victoria from previous studies: Uganda[a)]

Theme and Location	Parameter & Group parameters					Author(s)
	Chl-a (μg/l)	TN (μg/l)	TP (μg/l)	Temp ^{o}c	Transparency (m)	
Current and historical status (2009 data): Offshore	5.9			25.88	2.8	Sitoki et al., 2010
LVEMP (2000-2005): Station UP2 Inshore	2.4-6.24	11.9-50 (NO_3-N)	50-160	25.9-26	2.3-2.7	Muyodi et al., 2010
Optical properties (2003-2004): Murchison Bay	36.72					Okullo et al., 2007
Phytoplankton dynamics (2003-2004): Murchison Bay	20-60	>1100	> 90	26.2	0.8-4	Haande et al., 2011
MWE (2000-2001): Murchison Bay	5.61	27	4		1.38	Unpublished

[a)] Years are shown in parentheses.
Detailed historical and recent data was presented in Sitoki et al., 2010.
LVEMP – Lake Victoria Environmental Management Project
MWE – Ministry of Water and Environment

2.2 THE CASE OF LAKE NAIVASHA

2.2.1 Overview of Lake Naivasha

Lake Naivasha a freshwater lake located in Kenya rift valley at an altitude of 1889 m with surface area of 180 km^2 and mean depth of 8 meters, the lake receives water from River Malewa and Giligal (Stoof-Leichsenring et al., 2011). Important for fishing, tourism and a source of water to Nakuru area and for irrigation which has sustained the floriculture sector in the Country.

Lake Naivasha a RAMSAR site of ecological significance has experienced problems of reduction in water level and effects of fertilizers and pesticides which are heavily used within it's catchment (Ballot et al., 2009). Historical events presented by Stoof-Leichsenring et al., (2011) show that Lake Naivasha has been experiencing drought periods which significantly reduces its water level through the centuries, however an issue of the recent past was the introduction of flower production in 1969. The catchment of Lake Naivasha which is estimated to be 3,400 km^2 is dominated not only by Flower farms but there also coffee and tea plantations and other horticultural crops. Heavy metals from agrochemicals were found to be as one of the problems affecting the lake (Mutia et al., 2012).

Similar to Lake Victoria, *Cyperus papyrus* exists within the catchment of Lake Naivasha though the area is reduced by grazers (buffalo and cattle) according to Morrison and Harper, 2009. Application of tools that highlight ecosystem services and the relevance of wetland conservation with community participation are emphasized to be better options for management of Lake Naivasha compared to studies that only focus on biological and geosciences with no social dimension (Morrison et al., 2013).

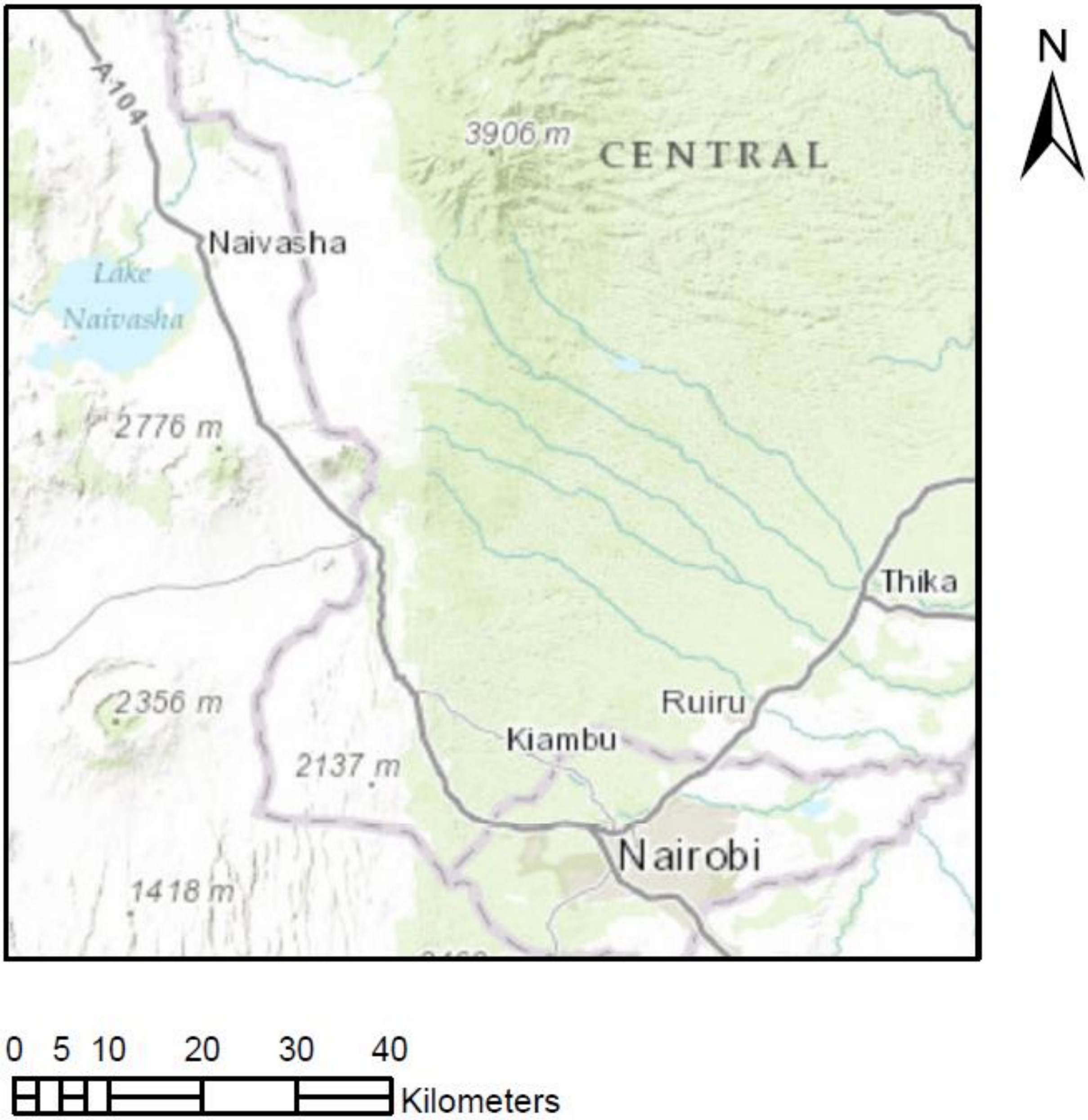
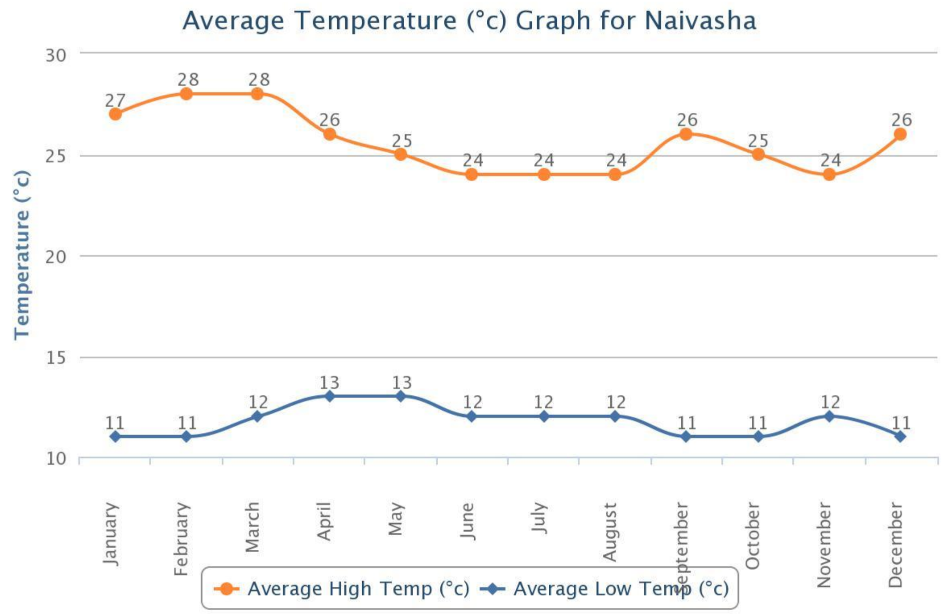


Fig. 2.9 Section of Kenya Topographic map showing Lake Naivasha and the Capital Nairobi, an extract of ArcGIS online basemap

2.2.2 Temperature and Rainfall for Naivasha

A



B

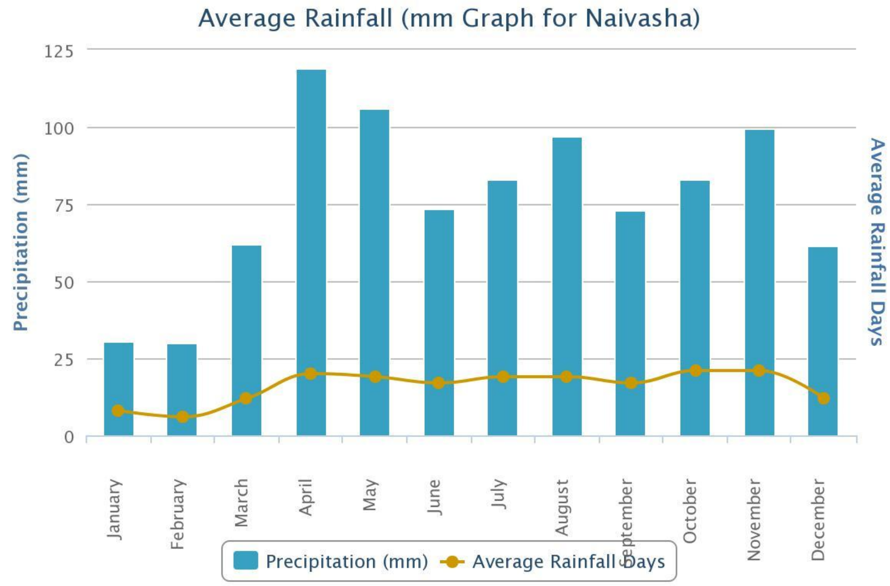


Fig. 2.10 Temperature and rainfall for Naivasha

Source: Worldweather
http://www.worldweatheronline.com/Naivasha-weather-averages/Rift-Valley/KE.aspx

2.2.3 Kenya Floriculture Industry

Flower production started in 1969 (Stoof-Leichsenring et al., 2011) and recently, 95% of the Flower production in Kenya is within Naivasha area (Mekonnen et al., 2012). The main drivers of investment in the Kenya floriculture sector are climate (low temperature), water for irrigation from Lake Naivasha, low labour and energy cost. In 2010, flowers covered an area of 3,400 hectares (Rikken, 2011).

Data shows that the Kenya flower exports have increased steadily except for 1998. As the leading exporter of roses to the European Union with 38 % of the market share, most of the roses (65 %) are sent to the Dutch Auctions. Floriculture exports were valued at KShs 42.9 billion in 2012, one of the leading horticultural exports from Kenya. The sector employs 90,000 people (KFC, 2013).

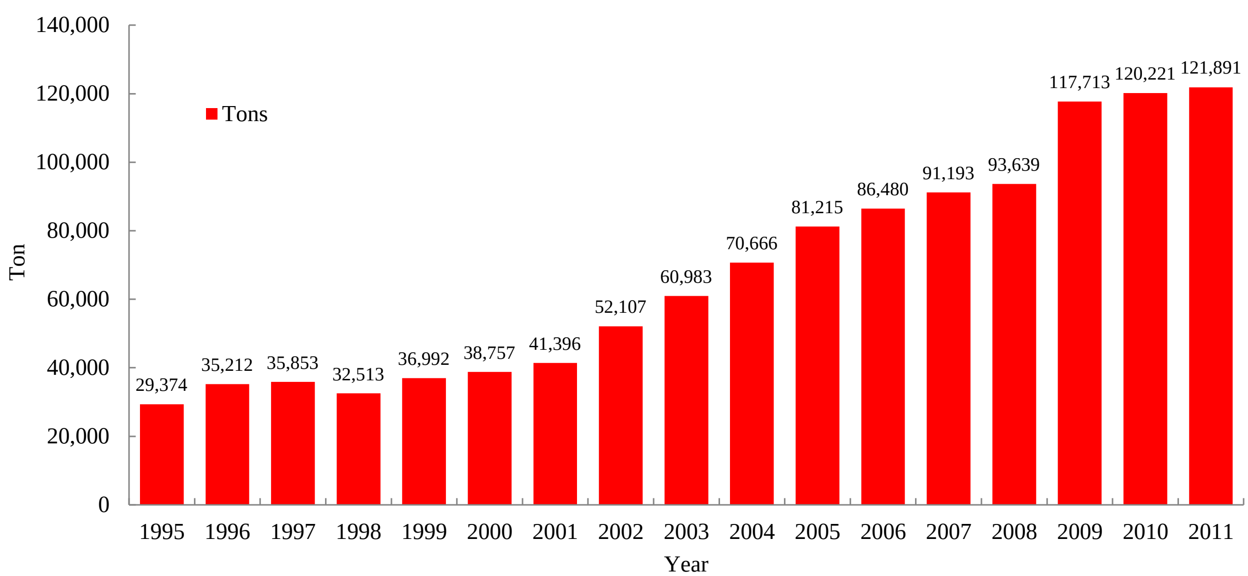


Fig. 2.11 Kenya flower exports (1995-2011)

Data Source: KFC

CHAPTER III

MONITORING POSSIBILITY USING REMOTE SENSING

3.0 REMOTE SENSING
3.1 SATELLITE DERIVED MODELS

3.1.1 Introduction

Remote sensing a process of acquiring information from objects and surfaces through analysis of light reflection and absorption within the light spectrum has found wider applications in urban planning and in assessment of water resources and terrestrial areas. The continuity of satellites at the same path and row enables monitoring of changes in Chlorophyll-a a measure of influence of environmental pressure and effectiveness of policies. Lake monitoring requires significant amount of data collected at various locations for lakes within a Country or across Countries. Issues of resource constraints or inaccessibility of lakes and sampling points are solved by remote sensing. The trans-boundary nature of many lakes makes data management challenging and lake monitoring is usually done at country level. Remote sensing offers opportunities to monitor and predict changes in Chlorophyll-a with limited amount of Chlorophyll-a and transparency data using satellite images with higher frequency per year, wider areas for analysis per single image and less susceptibility to estimation errors during cloud free days (Allan et al., 2011; Giardino et al., 2001).

Field spectrometric measurements of plant species is precise (Manevski et al., 2011), efficient for location specific studies but does not allow concurrent evaluation of features and land cover for wider regions. A single image (e.g., 170km X 185km for Landsat 7) allows spatial mapping of land cover and studying extensive utilization of natural resources. Monitoring Spatio-temporal changes in land use and land cover using remote sensing has been applied to determine the drivers of vegetation change (e.g., Maxwell and Slyvester, 2012; Otukei and Blaschke, 2010). Applications in Crop Science include mapping rice and ever cropped areas (Maxwell et al., 2013; Chen et al., 2011). The most popular satellite images used are for Landsat (5 & 7) and Moderate Resolution Imaging Spectroradiometer (MODIS), however RapidEye (5m resolution) enables mapping of land cover with greater detail compared to Landsat (30 resolution). Landsat images are freely available for download at the USGS site. Though remote sensing has been widely applied in the USA, Europe and China for over 3 decades, the studies are not very extensive in Africa.

Chlorophyll-a concentration is related with band values of satellite data for mapping it's distribution within a lake (Giardino et al., 2001). Band ratios of Landsat images were used in a regression to map Chlorophyll-a by Duan et al. (2008), such studies of remote sensing have provided reliable estimates of Chlorophyll-a (Allan et al., 2011; Zhang et al., 2011). SDT another popular indicator for assessing water resources is related with satellite band reflectance. An application is explained by Olmanson et al. (2008), who used Landsat to assess the water clarity of 10,000 lakes in Minnesota, USA. The correlations are made between natural logarithms of SDT and band values. Chlorophyll-a is more related to phosphorus and nitrogen concentrations in lakes than SDT.

3.1.2 Satellite derived Chlorophyll-a models

Table 3.1 Models of Chlorophyll-a derived from Landsat images

Model and date	Lake	Authors
In(Chl-a) = 1.9742 [In(B3)]+11.556 January, 2002	Rotorua Lakes, Lake Taupo and Rotoiti, New Zealand	Allan *et al*. (2011)
In(Chl-a) = 2.3205 [In(B3)]+13.244 October, 2002 and January, 2002		
Chl-a = 44.2-1.17(B1)-0.88(B2)+1.49(B3)+4.08(B4) November, 2005	Dajingshan reservoir, China	Xiong *et al*. (2011)
Chl-a = 7.394-0.377TM1+0.536TM2+0.732TM4 August, 2006	Lake Beysehir, Turkey	Nas *et al*. (2010)
Chla = $11.18_{pTM1} - 8.96_{pTM2}$-3.28mg/$m^3$ March, 1997	Lake Iseo, Italy	Giardino *et al*. (2001)[c]

Note: Also see Sass et al. (2007) and Nas et al. (2010) with more descriptions of other studies on Chlorophyll a and SD.

[a] ETM - Enhanced Thematic Mapper

[b] TM - Thematic Mapper.

[c] $_{pTM1\ and\ pTM2}$ are atmospherically corrected reflectance values in TM bands and 1 and 2.

3.1.3 Secchi-disk transparency (SDT) models

Table 3.2 Models of Secchi Depth derived from Landsat TM images

Model and date	Lake	Authors
InSD = 0.134TM1-0.392TM3+2.484 September, 2004	Lake Maine, USA	McCullough *et al.* (2012)
SD = -16.89+93.84(TM1/TM3)-2.162TM1 August, 2006	Lake Beysehir, Turkey	Nas *et al.* (2010)
In(SD) = 1.493(TM1/TM3)-0.035TM1-1.956 August, 2001	Lakes in Minnesota, USA	Sawaya *et al.* (2003)
SD = $8.01_{pTM1}/_{pTM2}$-8.27 March, 1997	Lake Iseo, Italy	Giardino *et al.* (2001)[a]

a) pTM1 and pTM2 are atmospherically corrected reflectance values in TM bands and 1 and 2.

3.2 SATELLITE IMAGE CLASSIFICATION AND VEGETATION DIFFERENTIATION

Image classification procedures are generally grouped as parametric or non-parametric. Parametric classifiers assume that data is normally distributed while the later are distribution free classifiers for example, Decision Trees (DT) and Support Vector Machines (SVM). The two are popularly applied to classify satellite images (Otukei and Blascke, 2010). Since DTs make no prior assumptions about the data and their capacity to manage non-linear relationships between features and classes, they are of more advantage than other classifiers (Pal and Mather, 2003). Maximum Likelihood Classification (MLC) a parametric technique with the assumption that data is derived from a normal distribution (Keuchel et al., 2003) is the most commonly used classifier (Lu and Weng, 2007), in their review, they present the advances in image classification.

In supervised classification, co-ordinates are used to create training samples for formation of classes which leads towards a more specific calculation of area occupied by the different features in the landscape. Multiple classification techniques termed classifier ensembles can be applied on the same image to compare results and improve the accuracy of image classification procedure (Waske and Braun, 2009).

The relatedness in reflectance for most plant species is a major challenge to their differentiation (Ullah et al., 2012), this has resulted into development of a diversity of indices. Currently there are over 40 vegetation indices (Peng et al., 2012). Normalized Difference Vegetation Index (NDVI), Difference Vegetation Index (DVI), Greenness Index (GI) are some of the vegetation indices (Saadat et al., 2011), NDVI being the most popular index (e.g, Chen et al., 2012; Maxwell et al., 2012). NDVI values developed from the satellite images and shapefiles are used as ancillary data for accuracy of image classification (Heinl et al., 2009). Satellite images at high resolution with high spectral definitions are analysed to differentiate plant species, RapidEye (5 m resolution) is applied in this study.

NDVI is based on the principle of light assimilation and reflectance by the objects being assessed. Chlorophyll absorbs Red-light of the electromagnetic spectrum while the mesophyll cells scatters Near Infrared (NIR) (Pettorelli et al., 2005). The range of NDVI is -1 and 1. Values in the positive range represent increasing vegetation greenness whereas values close to 0 represent soil, rock, water, ice and other objects and surfaces other than vegetation. Below is the formula for calculating NDVI.

$$\text{NDVI} = \frac{\text{NIR-R}}{\text{NIR+R}}$$

Naturally, a variety of plants exist with similar reflectance spectra (Ullah et al., 2012) which results into similar NDVI values. Ancillary data such as co-ordinates for plant species and elevation when combined to set training samples in the process of image classification improves interpretation of NDVI values (Chen et al., 2011; Manevski et al., 2011). In the section ahead, Landsat and RapidEye are analysed to assess the spatio-temporal changes in land cover for Entebbe and Kampala, Uganda.

3.3 WETLAND CHANGE

Changes in Wetlands of Lake Victoria

The dominant plant surrounding Lake Victoria is *Cyperus papyrus*. *Cyperus papyrus* in this area has been studied since the 1950s (Lind, 1956). The studies describe the biological functionality of papyrus (Kansiime et al., 2007) and socio-economic perspectives of wetland conservation. Real Estate development, Agriculture, Eucalyptus plantations (Appendix A3) are the main drivers of receding wetlands. In estimating land cover changes for the past 20 years in East Africa, Brink et al. (2014) show that Agricultural area increased by 28 %.

Remote sensing was previously applied to support strategies for management of water weeds (Cavalli et al., 2009). A similar technique for mapping the distribution of papyrus is a remaining research gap. Using the available satellite images, it is possible to map papyrus wetlands concurrently with impervious surfaces (estates, roads) for change detections and evaluating the trend of urban development.

Spatio-temporal changes in land use and land cover is essential to determine transitions between land use types over time. Since the ecological sustainability of Lake Victoria depends mostly on papyrus wetlands, their present distributions will guide landscape and urban planning efforts as Kampala the capital is located 37 Km away from Entebbe and the latter is developing further due the increase of urban population.

3.4 SPATIO-TEMPORAL CHANGES IN LAND-COVER: LAKE VICTORIA

3.4.1 Abstract

In this paper, we estimated, for Entebbe and Kampala areas on the northern shore of Lake Victoria as study areas, the present extent of *Cyperus papyrus* (papyrus) wetland and its change over time, using RapidEye (5 m resolution) and Landsat (30 m resolution) satellite images. We first estimated land cover for Entebbe area in 2010/2011 by using both RapidEye and Landsat images, second, the performance of Landsat in land cover classification was compared with that of RapidEye, and third, we identified changes in land cover in the last 15 years for the study areas by using Landsat images. The results of GIS analysis of RapidEye revealed that in 2011, 30% of Entebbe area was occupied by wetland, of which 70% was papyrus-covered. Between 1995 and 2010, the share of wetland decreased from 38% to 32% for Entebbe and from 15% to 11% for Kampala, but for both areas, the most decreases occurred in the last 5 years. The urban land use increased in Kampala from 17% to 64%, and from 9% to 23% for Entebbe, but for both areas, the type of land encroached first by the expanding urban land use was non-wetland vegetation, such as crop land, forest and green space, with relatively low encroachment on wetlands until the mid-2000s. However, the urban expansion in recent years has reached a stage to encroach wetlands.

3.4.2 Introduction

Found in areas around Lake Victoria, the second largest freshwater lake in the world, are immense wetlands with thickly grown *Cyperus papyrus* (henceforth papyrus), a popular scenery in inter-lacustrine areas in East, Central and Southern Africa but a very important ecosystem not only to the natives but also to the world, as it is important for biodiversity and environmental conservation (Lind, 1956; Lind and Visser, 1962; NFA, 2002; Kansiime et al., 2007; Maclean et al., 2011). The increase in urban population has resulted into shifts in land cover mainly to agricultural, industrial and estate developments, a phenomenon occurring in areas around Lake Victoria, Uganda, responsible for a decline in wetlands. Papyrus is a plant species that absorbs pollutants. The changes in the extent of papyrus-covered wetlands, therefore, would critically affect the water quality of Lake Victoria. Changes in land cover are evident from observations on the ground, particularly along roads and highways, but the topographic nature of wetland areas makes it difficult to know how extensive the changes have been. An accurate estimation of papyrus habitat is important in the sustainable development of the area surrounding Lake Victoria. There are some land-cover and topographical maps, such as National Forestry Authority's 'Shapefile of 1996 and 2005' (NFA, 2006) and Survey and Mapping Department's topographic map of 1998 (SMD, 1998). Though useful to know land-cover classifications, they do not give a complete estimate of wetlands, and papyrus is only mentioned among many wetland species. SMD gives data only for a single year, not enabling us to analyze changes over time.

Remote sensing techniques could be instrumental in providing information for accurately identifying papyrus-covered wetlands at present and their changes over time. Since papyrus is a perennial sedge of the Cyperaceae family which propagates by rhizomes and seeds, growing up to a height of 9 m with emergent stems and an inflorescence (Jone and Muthuri, 1985; Kansiime et al., 2007), the use of high-resolution satellite images makes it possible to identify papyrus-covered wetlands. Adam et al. (2012), Adam and Mutanga (2009) and Cavalli et al. (2009) identified papyrus-covered wetlands in some parts of the lacustrine areas in Eastern and Southern Africa using field spectral measurements and Landsat. RapidEye, one of the high-resolution satellite images of 5 m resolution, has offered an opportunity to classify papyrus. A problem of RapidEye is that its high resolution increases the number of satellite images, and therefore the cost, which are necessary to analyze a wide range of area. Landsat is another set of satellite images that are more readily available than RapidEye, but Landsat's resolution is 30 m, with which wetlands are well identified but not efficiently for papyrus. Vermeiren et al. (2012) and Abebe (2013) studied changes in land-cover patterns in Kampala using Landsat images, but their land cover classification does not include papyrus.

The purpose of this paper is to identify papyrus wetlands that face a high likelihood to be developed in future for areas along the Northern shore of Lake Victoria, through estimating the present extent of papyrus wetlands and its changes over time, using satellite images of RapidEye and Landsat. Specifically, we first identify 'papyrus wetlands' in Entebbe area by using a RapidEye image; second compare the performance of land cover classification between RapidEye and Landsat for Entebbe area; and third identify changes in wetlands in the last 15 years for Entebbe and Kampala areas based on Landsat images.

3.4.3 Materials and Methods

Study areas

We chose Entebbe and Kampala situated along the Northern coast of Lake Victoria as our study areas. Entebbe area was selected for our study to represent areas with vast papyrus wetlands facing a high likelihood of being developed in the near future. Entebbe is in Wakiso district, with the population of 1.3 million in 2011 (UBOS, 2013), total yearly rainfall of 1,507 mm, and a daily average minimum and maximum temperature of 17.5 and 25.6°C, respectively (BBC, 2012). Our Entebbe study area includes Entebbe town, the Entebbe International Airport and the surrounding areas within a square of 625 km^2 (25 km x 25 km with the coordinate of the northwest corner; 32.35 E and 0.222 N). The landscape of Entebbe town is relatively flat at an altitude of 1180 meters, while the nearby areas are mostly hilly with valleys that embrace at the bottom extensive wetlands, all of which are continuously connected with Lake Victoria on the surface or underground. As reported by Elhadi et al. (2009) for Lake Victoria and Central Africa, papyrus is the most abundant species in wetlands in the study area. Besides residential plots, agricultural fields, Eucalyptus plantations, pine and forest reserves are the main land use and land cover types. A distinct feature in the land use of the area is the concentration of a high number of export-oriented large scale flower farms (UFEA, 2010).

Kampala, the capital of Uganda, which was selected for our study to represent areas that have undergone rapid urbanization, is situated 37 km to the northeast of Entebbe, at an altitude of 1190 meters, with a population of 1.7 million (UBOS, 2013). Nearly adjacent to the Entebbe study area with only a 10 m difference in elevation from Entebbe town, the Kampala study area shares similar natural conditions as the Entebbe study area. For land cover, however, the Kampala study area has gone through tremendous changes during the past decade as one of the rapidly growing cities in Africa (Vermeiren et al., 2012), and most of the land in Kampala is occupied by buildings and roads. The Kampala study area follows the area designated as Kampala in the NFA Shapefile of 2005 (NFA, 2005), within a square of 784 km^2 (28 km x 28 km with the coordinate of the northwest corner; 32.50 E and 0.500 N).

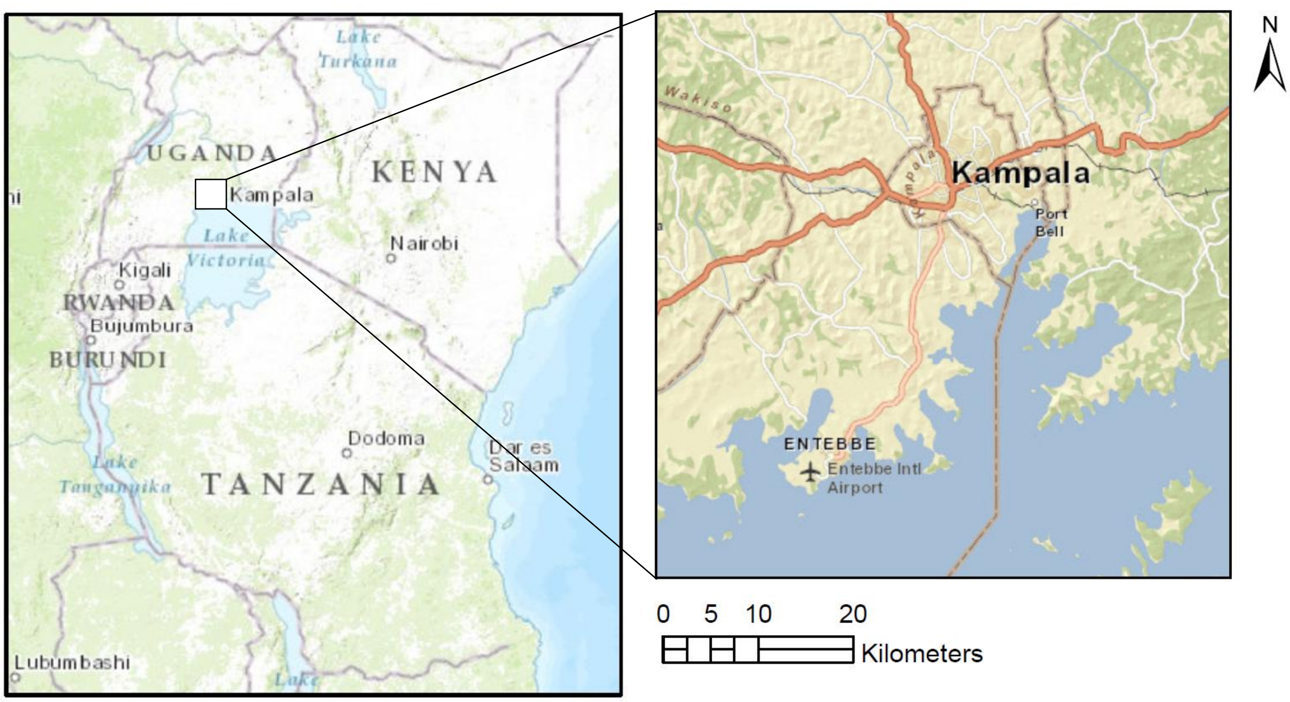


Fig. 3.1 Topographic map of East Africa (left) and street map of the study site (right), sections of ArcGIS online basemaps.

Satellite image analysis by Maximum Likelihood Classification

The procedures of image classification into various land cover are generally grouped as parametric or non-parametric. Parametric classifiers such as Maximum Likelihood

Classification (MLC) rely on the assumption that data are normally distributed (Keuchel et al. ,2003), while non-parametric classifiers are distribution free, such as Decision Trees (DT) (Otukei and Blaschke, 2010), and Support Vector Machines (SVM) (Pal and Mather, 2003). MLC is the most commonly used classifier and adopted in this study.

The satellite image analyses conducted in this study are summarized schematically in Figure 3.2. We used RapidEye high resolution satellite image of 2011 to identify the distribution of papyrus wetlands and Landsat 5 and 7 images of 1995, 1999, 2005 and 2010 to identify the distribution of wetland and its changes over-time. In order to focus on wetlands, in this study we classify land cover into four classes, papyrus-covered wetland, multi-species wetland, non-wetland vegetation and impervious surfaces for RapidEye image, and three classes, wetland (including papyrus-covered wetland), non-wetland vegetation and impervious surfaces, for Landsat images. Impervious surfaces are such land as buildings, roads and open land, and crop lands and forests are included in non-wetland vegetation.

In setting the training areas to establish classification patterns, information on actual situations (ground truths) obtained from field observation was used for RapidEye 2011 and Landsat 2010 images. For Landsat 1995 and 1999 images, the topographic map of 1998 made by the Survey and Mapping Department with the assistance from JICA (SMD, 1998) and the Shapefile of 2005 (NFA, 2005) were used. All the GIS analyses were conducted in Arcmap version 10 by Maximum Likelihood Classification.

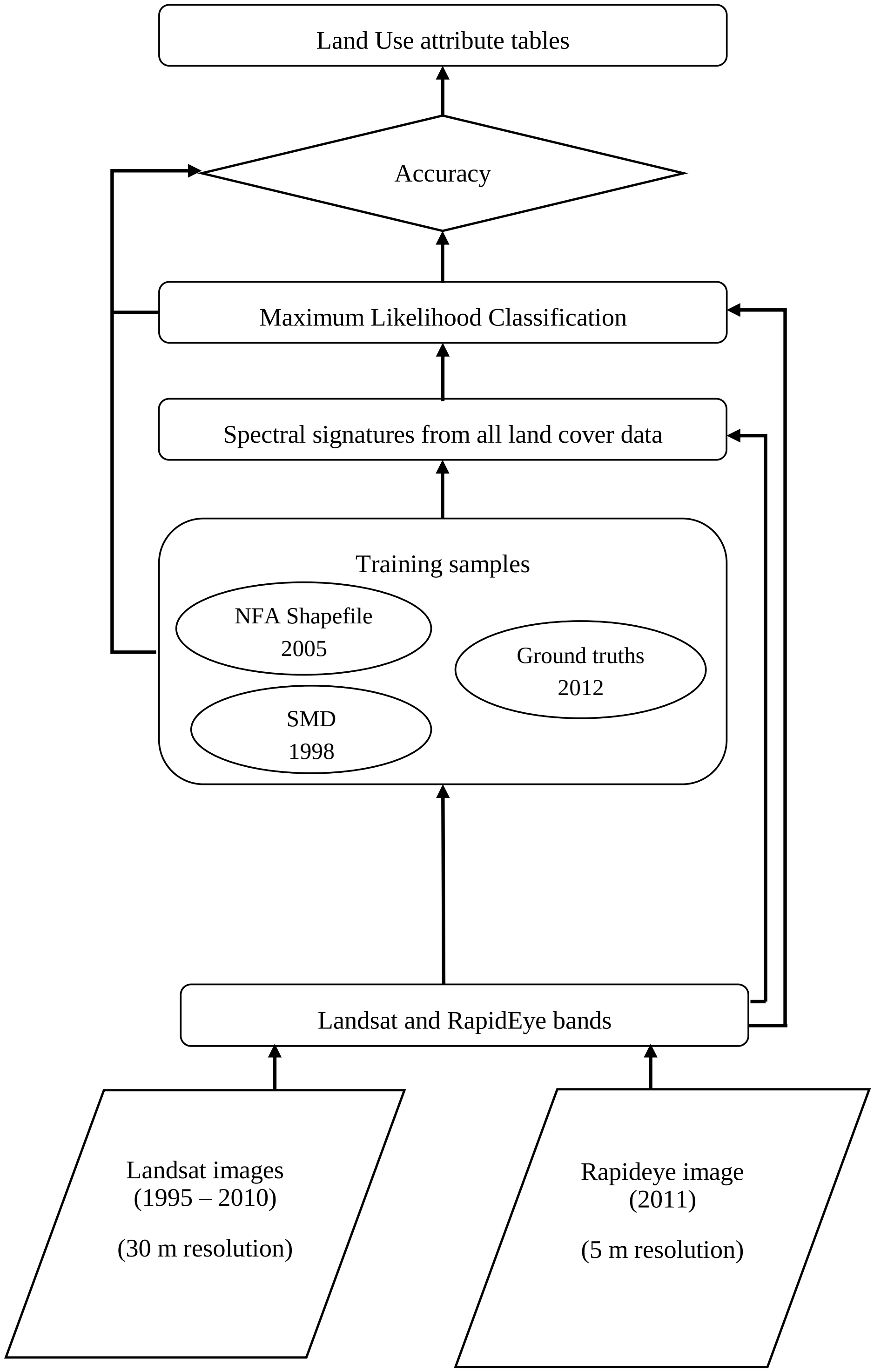


Fig. 3.2 Procedure for classification of satellite images by Arcmap 10

Satellite images

The satellite images used in this study are listed in Table 3.3. The RapidEye image was obtained from RapidEye (RapidEye, 2012). Landsat images used are Landsat 5 and 7 in Path 171 and Row 60, downloaded from the United States Geological Survey (USGS) online archive (USGS, 2012). Landsat images of 1999 and 2010 were used instead of 2000 and 2011, respectively, due to image availability in the archive, image quality and the need to estimate changes in land-cover at a 5 year interval. Landsat 7 image for 2005 used in this study was a SLC-off image due to failure of scan line corrector. The striped missing areas were excluded automatically from our analysis when we made land-cover maps. Note that Landsat image of 1999 was affected by a thick cloud cover. Since as much as 43% of the image is not visible, compelling us to confine our analysis to the visible part, the results of analysis for this year must be regarded as a reference. For the rest of the years, cloud cover is so thin or inexistent that it entails no problem for the analysis.

Table 3.3 Sensors and dates of satellite images used, RapidEye and Landsat[a]

Sensor	Cloud cover (%)	Shadow (%)	Date of image
RapidEye	0.9	0	11/02/2011
Landsat 5	na	na	14/12/2010
Landsat 7	na	na	06/01/2005
Landsat 7	37	6	24/12/1999
Landsat 5	na	na	19/01/1995

a) 'na' means that estimates are not affected by cloud cover.

Field evaluation and ground truthing

In order to collect data on actual land cover (ground truth), field evaluations were conducted between February and March 2012 in the Entebbe study area. The GPS coordinates of land-cover distributions and locations of various plant species were collected. Montana 650 (Garmin Inc) Global Positioning System (GPS) was used to capture geo-referenced images of diverse vegetation. Although papyrus dominates in wetlands in the area, diverse plant species, such as common reed and palm also occupy a certain percentage of the study area (Appendix A2). The co-ordinates of papyrus and other land-cover distributions were used to set training areas for supervised classification for the RapidEye image of 2011 and the Landsat image of 2010 by Maximum Likelihood Classification (Figure 3.2).

Results and discussion

Entebbe in 2011 (RapidEye)

The results of land cover classification for Entebbe study area based on RapidEye are summarized in Table 3.4 for the four land use land cover classes and mapped in Figure 3.3. The overall accuracy rate of the estimated classification is 83%, which is comparable to other studies (Vermeiren et al., 2012; Abebe, 2013). The accuracy rate for papyrus wetlands is 86%.

Papyrus wetlands took as much as 21% of the total area, excluding the water surface of Lake Victoria. Since multi-species wetlands took 9%, altogether 30% of the study area was occupied by wetlands, and 70% of the entire wetlands were papyrus wetlands. It is apparent in Figure 3.3 that the papyrus wetlands took a substantial share in the study area. However, it is also apparent that large, extensive papyrus wetlands were found in the central and western parts where human habitation was sparse, and in the areas stretching from the northeast corner of Figure 3.3, adjacent to Kampala metropolitan area, to the tip of Entebbe peninsular (the largest peninsular jutting out into the lake south-westerly) where urbanized land use shown as impervious surfaces dominates, papyrus wetlands were besieged by impervious surfaces. This urban-land-use-congested area is along the highway connecting Kampala with Entebbe International Airport situated at the tip of the Entebbe peninsular. Found midst of many relatively small papyrus wetlands in this area are spots of impervious surfaces, which is the sign of human encroachments.

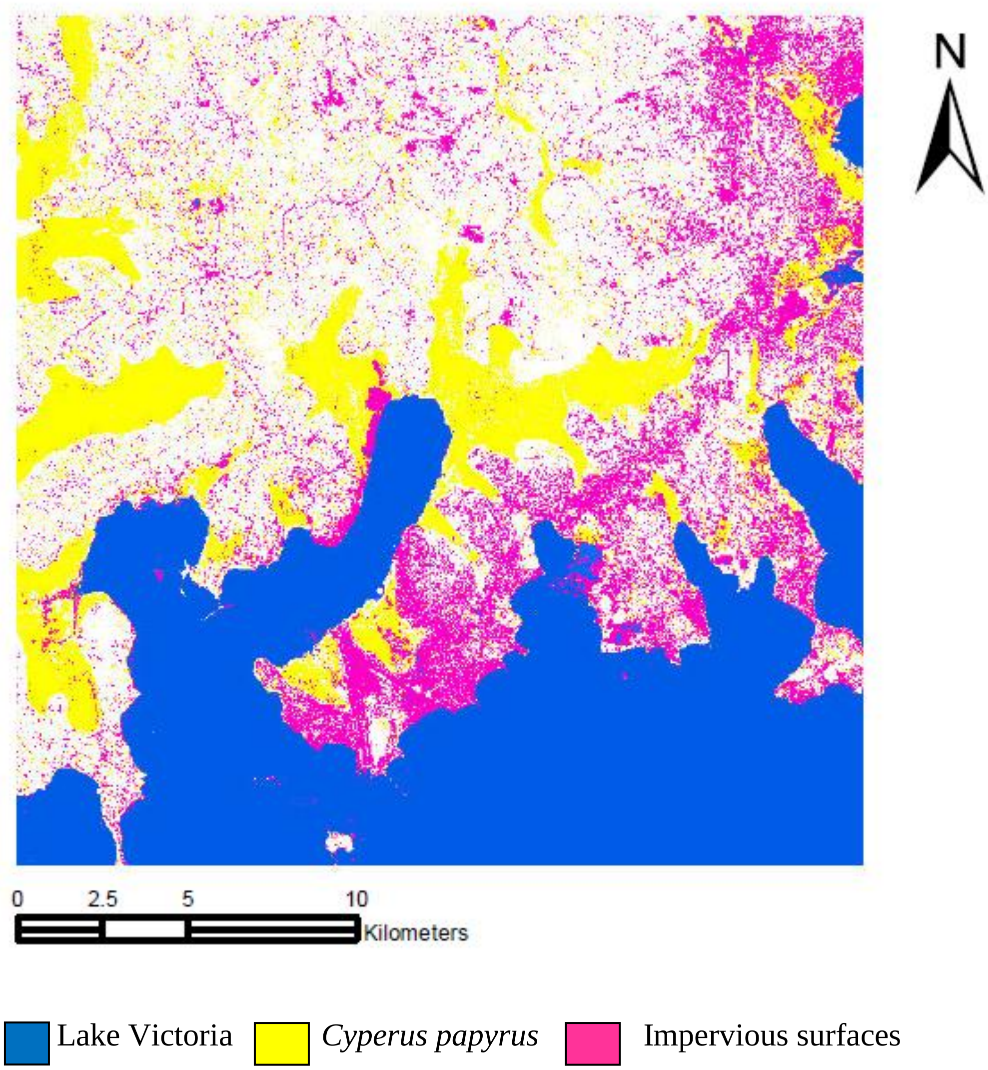


Fig. 3.3 Results of RapidEye satellite image for 2011 showing locations of *Cyperus papyrus* and impervious surfaces

Table 3.4 Results of land-cover classification estimation for Entebbe area, RapidEye 2011 and Landsat 2010[a]

	RapidEye, 2011			Landsat 5, 2010		
	Area (`000ha)	%	Accuracy (%)	Area (`000ha)	%	Accuracy (%)
Impervious surfaces	8.4	19	83	10.1	23	76
Non-wetland vegetation	22.7	51	80	20.0	45	73
Wetland:						
Cyperus papyrus	9.1	21	86			
Multi-species	3.9	9	83			
Total	13.0	30	84	14.1	32	80
Total	44.2	100	84	44.2	100	76

a) Land area without Lake Victoria.

Table 3.5 Results of land-cover estimation for Entebbe and Kampala, Landsat 1995, 1999, 2005 and 2010a

	1995		1999[b]		2005		2010	
	Area ('000 ha)	%	Area `000ha	%	Area (`000ha)	%	Area (`000ha)	%
Entebbe:								
Impervious surfaces	3.8	8.6	8.0	18	8.4	19	10.1	23
Non-wetland vegetation	23.4	52.9	20.8	47.1	19.4	44	20.0	45
Wetland	17.0	38.5	15.4	34.9	16.4	37	14.1	32
Total	44.2	100	44.2	100	44.2	100	44.2	100
Kampala:								
Impervious surfaces	12.5	16.6	25.1	33.5	32.2	42.9	48.1	64.1
Non-wetland vegetation	51.0	68	39.1	52.1	31.4	41.8	18.5	24.7
Wetland	11.6	15.4	10.8	14.4	11.5	15.3	8.4	11.2
Total	75	100	75	100	75	100	75	100

a) Lake area without Lake Victoria
b) Affected by heavy cloudcover

Comparison between RapidEye and Landsat for Entebbe

In order to examine how Landsat images perform in land cover classification, the results of land cover classification based on the Landsat 5 image of 2010 for the Entebbe study area were compared with those of the RapidEye of 2011 (Table 3.4). For Landsat with a lower resolution, three classes of land cover were distinguished: wetland, impervious surfaces and non-wetland vegetation. The overall accuracy rate of 76% for the Landsat is lower than for RapidEye, but it is 80% for wetland, higher than for other land-cover classes.

The Landsat image of 2010 identified that the area of wetlands was 14,100 ha or 32% of the entire Entebbe study area, not including the lake. Corresponding figures obtained from RapidEye were 13,000 ha and 30%; the estimation gap of 8% for the absolute area and 2% for the percentage share. These estimation gaps between RapidEye and Landsat 14% and 4% for impervious surfaces and 12% and 6% for non-wetland vegetation, respectively. There are some estimation gap between Landsat image of 30 m resolution and RapidEye image of 5 m resolution, but the degrees of the gaps are less than the degree of the resolution gap. In particular, the gap between the two estimates is relatively small for wetland. Taking note of these estimation gaps, we estimated the temporal changes in land cover using Landsat images in Entebbe and Kampala.

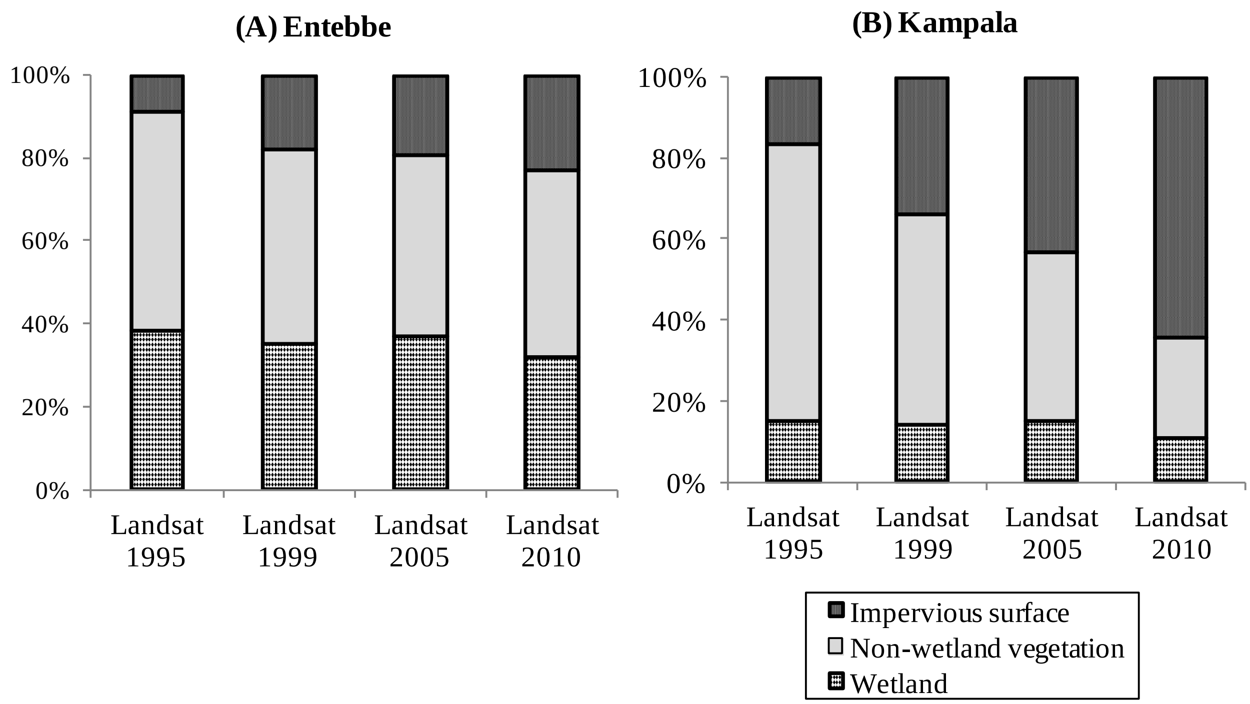


Fig. 3.4 Multi-temporal changes in land cover 1995-2010 by Landsat, Entebbe and Kampala

Changes in wetland for Entebbe and Kampala (Landsat)

The results of estimation for 1995, 1999, 2005 and 2010 based on Landsat images are summarized in Table 3.5 and shown in Figure 3.4 for Entebbe and Kampala. The changes in land cover in Entebbe study area have been relatively gradual. The share of wetlands decreased from 38% in 1995 to 32% in 2010. However, most of its decrease occurred between 2005 and 2010. As expected, the share of impervious surfaces increased markedly in the last 15 years. This increase occurred at the expense of wetlands and non-wetland vegetation, but the latter, changing from 53% in 1995 to 45% in 2010, was encroached more than the former. Comparing to Kampala area, in Entebbe with relatively less population, the pressure on changes in land cover is less. However, wetland and forest encroachment is a common problem in Entebbe area too (Baranga et al., 2009).

The changes in land cover for the Kampala study area have been far more drastic than those for Entebbe. Impervious surfaces increased tremendously from 17% in 1995 to 64% in 2010. In Kampala, too, it was non-wetland vegetation that absorbed most of the expansion of urbanized land use; the share of non-wetland vegetation decreased from 68% to 25% in the same period. In contrast, the share of wetland remained nearly at the same level of 15% until the mid-2000s, but began to decrease after 2005. This trend in wetland supports the findings of Vermeiren et al. (2012) and Abebe (2013). The changes in land cover for Kampala has been typical of rapidly developing urban centers where green spaces, crop lands, forests and wetlands have been replaced by buildings and roads for industrial and residential developments, including encroachments by new city dwellers.

A common observation for Entebbe and Kampala is that though began to decline in recent years, wetlands had remained relatively less affected during the decade between 1995 and 2005. Since it is expected that the composition of papyrus wetlands and multi-species wetland did not change rapidly, the changes in papyrus wetland would have also been small. This is in a sharp contrast to the case of the western shore of Lake Victoria where as much as 50% of papyrus wetlands were lost between 1969 and 2000 (Owino and Ryan, 2007).

However, encroachments by urban development, and by agricultural land use at a lesser extent, now reach wetlands, most of which are covered by papyrus. In 5 years from 2005, Entebbe lost 14% of wetlands and Kampala lost as much as 27%. As the Uganda economy continues to develop in future as has been in the recent past, the pressure on changes in land use will increase and encroachments of urban and agricultural land uses into papyrus wetlands will become more rampant. Whether papyrus wetlands are to be preserved from environmental points of view or to be harnessed as economic resources, whichever the case, it is now the time to establish or re-establish firm and manageable policies towards papyrus wetlands.

CHAPTER IV

RELATIONSHIP BETWEEN PHOSPHORUS, NITROGEN AND ENVIRONMENTAL INDICATOR

4.1.1 STATISTICAL MODELLING

Statistical modeling is essential to develop models and coefficients applicable to lakes with in countries and globally. There is need to forecast the status of lakes using available data, such simulations are feasible with robust models. Chlorophyll-nutrient relationships are widely published (Brown et al., 2000; Reckhow, 1993).

Panel data analysis which controls for site specific effects is applied to develop robust models for Chlorophyll-nutrient relationship. Since we incorporate cross-sectional and longitudinal data in panel data analysis, it is justifiable technique as most available data are for a few years 2 to 4 and on a limited number of lakes. Databases on multiple lakes, in many countries and for a long period are vital to develop better models.

The modelling approach used data which spans over 4 decades (45 years). These are the first panel data models to appear in this field as previous studies have shown that the most popular model was ordinary least squares (OLS) (eg.,Trevisan and Forsberg, 2007; Huszar et al., 2006). The two types of panel data models developed are the random effect and dynamic model.

4.1.2 CHLOROPHYLL-A IN LAKES

In monitoring water resources, data on parameters such as chlorophyll-a, phosphorus, nitrogen, temperature are collected routinely. The increase in phytoplankton biomass in lakes will signify the necessity of management interventions for sustainable use of a lake or reservoir. Chlorophyll-a is the most used indicator of phytoplankton biomass (Søndergaard et al., 2011). Chlorophyll-a in lakes depends mostly on phosphorus, nitrogen and temperature (Phillips et al., 2008; Brown et al., 2000), temperature is natural factor and complex to manage in a spatial context.

Sunlight on phytoplankton biomass

In relative terms, the effect of sunlight on phytoplankton is less studied compared to studies on the influence of nitrogen and phosphorus on Chl-a. However, the studies indicate that light is a major environmental factor that affects the growth of phytoplankton (Ssebiyonga et al., 2013; Zinabu, 2002). On seasonality, high summer temperature increases Chl-a concentration (Kallio, 1994), though not a general issue for temperate lakes. In addition, results of Liu et al. (2010b) show a positive correlation between Chl-a and temperature.

Depth and chlorophyll-a

Since light is essential for growth of phytoplankton, their distribution varies with water depth as influenced by reduction in light intensity (Bhutiani et al., 2009), thus, sampling depth affects the efficient estimation of Chl-a in lakes (Nõges et al., 2010). Results of a study by Carvalho et al. (2009) show that high concentration of Chl-a is found between 0 and 2 meters. Variations exist among deep and shallow lakes with a tendency of deep lakes having high subsurface concentrations of Chl-a during lake stratification compared to shallow lakes, commonly known as deep chlorophyll maximum (DCM) (Camacho, 2006; Hamilton, et al., 2010).

Regressions

Regression relationships are made using data for single or multiple lakes and the most common of these relationships is one between Chlorophyll-a and phosphorus, followed by Chlorophyll-a and nitrogen and, Chlorophyll-a and temperature (e.g., Lv et al., 2011; Trevisan and Forsberg, 2007; Liu et al., 2010b). Distributions of Chlorophyll-a, total phosphorus and total nitrogen have been spatially presented (Bachmann et al., 2012; Arhonditsis et al., 2003). Details of panel data regression are discussed in the next two sections (4.2.1 and 4.2.2).

4.2 REGRESSION ANALYSIS USING EURO-LAKES DATA

4.2.1 RANDOM-EFFECT MODEL

Abstract

Chlorophyll-a (Chl-a) is widely used as a water quality indicator. Sampling locations within lakes are visualized by spatial analytical tools for decision making in effective management of water resources. With available limnological data, a model can be developed to enable application of its coefficients in other areas where data collection is not feasible. We developed a model which explains the variation of Chl-a in lakes by adopting panel data analysis with water quality data of European lakes and verified the robustness of the model by predicting Chl-a for lakes in United Kingdom, Japan, Australia and the USA. The amount of Chl-a in lakes mainly depends on Total Phosphorus (TP) and Total Nitrogen (TN). In addition to nutrients, we included a land-use dummy as a variable to account for the effects of land-use type on lake Chl-a. In the fixed- and random-effect models, TP and TN were significant ($p < 0.05$). The random effect model was selected for the simulations on the basis of the Hausman test. The R^2 between the predicted and actual Chl-a using samples from different countries was between 0.83 and 0.66, except for the database of Florida lakes in the USA, which indicates we succeeded at developing a functional model for prediction of Chl-a from phosphorus and nitrogen.

Introduction

The increasing interference in ecosystems has resulted into diverse ecological and public health problems. The most common problem is deterioration of water quality. Chlorophyll-a (Chl-a) is among the parameters used to indicate phytoplankton biomass and is related with the concentrations of nutrients in water bodies (Søndergaard et al., 2011; Kasprzak et al., 2008; Greisberger and Teubner, 2007). In addition to Chl-a, secchi disk transparency and reflectance data from satellite images are used in remote sensing studies to assess water clarity (Olmanson et al., 2008; Hedger et al 2002; Pepe et al., 2001). Geographical Information Systems (GIS) enables the combination of point and spatial water quality data to identify areas affected by environmental problems using land use maps, watershed networks and digital elevation models. The use of Chl-a in remote sensing is cost-effective and enables the assessment of data for trans-boundary and multiple lakes.

In this paper, we used Chl-a as a waterquality indicator because phytoplankton growth depends on the aggregative effect of nutrients, environmental and geological factors. In the past studies, Jeppesen et al. (2005) concluded that reduction in external total phosphorus resulted in low Chl-a concentration. Many previous studies indicated that phytoplankton growth was mostly phosphorus-limited (Arvola et al., 2011; Lv et al., 2011; Wang et al., 2007; Brown et al., 2000). However, Trevisan and Forsberg (2007) found that it was mostly nitrogen-limited. Liu et al. (2010b) used multivariate analysis to understand the factors which influence Chl-a and found that temperature and phosphorus were important determinants.

Conventional Chl-a and nutrient relationships have singled out phosphorus and nitrogen, in that order, as important determinants of Chl-a. However, most of these studies dealt with the scale of a single lake, a group of lakes in the same area, country (Sondergard et al., 2011), region (Seip et al., 2000) or a combination of tropical and sub-tropical lakes (Huszar et al., 2006). If the relationship among Chl-a, phosphorus and nitrogen in lakes could be developed with cost-effective Geographical Information Systems and modelled across many countries, regions and even globally, such a model would be invaluable in assessing water quality and environment around lakes.

The purpose of this study was to examine if such a comprehensive and robust model could be constructed. First, we built a model in which lake Chl-a was determined by Total Phosphorus (TP) and Total Nitrogen (TN), on the basis of data for 198 lakes in 10 countries across the European continent. We used the panel data regression method to eliminate heteroscedasticity caused by large variations inherent in widely scattered observation points. Second we tested the model by applying it to lakes in United Kingdom, Japan, Australia and Florida in the USA to predict lake Chl-a.

Methodology and data

Analytical models

The relationship between trophic state indices is estimated by panel data regression models, which have the advantage of controlling the heterogeneity associated with individual units. An observation of the data set is identified along two dimensions of time ($t = 1, \ldots, T$) and water quality monitoring site ($i = 1, \ldots, N$).

$$Chl\text{-}a_{it} = \alpha + \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i + \varepsilon_{it}$$

where $\boldsymbol{x}_{it}$ is a vector of TP, TN and other explanatory variables, α and $\boldsymbol{\beta}$ are unknown parameters to be estimated, u_i is the unit-specific residual which differs among units, but the value is constant for any particular unit, and ε_{it} is the error term with the usual properties of mean 0 and homoscedastic variance, σ_ε^2. We tried out two models. The fixed effect model considers u_i as an intercept specific for each unit, while it is treated as a random variable with 0 mean and σ_u^2 for the random effect model, in which the Generalized Least Squares (GLS) method is required since the covariance of σ_ε^2 and σ_u^2 is not equal to zero for each unit. The GLS estimation is performed by transforming the variables using θ, as follows:

$$\left(y_{it}-\theta\overline{y}_i\right)=\left(1-\theta\right)\alpha+\left(x_{it}-\theta\overline{x}_i\right)\beta+\left\{\left(1-\theta\right)\overline{v}_i+\left(\varepsilon_{it}-\theta\overline{\varepsilon}_i\right)\right\}$$

where $\overline{y}_i=\sum_t y_{it}/T$, $\overline{x}_i=\sum_t x_{it}/T$, $\overline{\varepsilon}_i=\sum_t \varepsilon_{it}/T$ and $\theta=1-\sqrt{\sigma_\varepsilon^2\big/\left(T\sigma_u^2+\sigma_\varepsilon^2\right)}$. In case $\theta=0$, the equation is exactly the same as the one of fixed effect model. Adequacy of the random effect model is judged by the Hausman specification test,

$$H=\left(\beta_{Fixed}-\beta_{Random}\right)'\left(V_{Fixed}-V_{Random}\right)^{-1}\left(\beta_{Fixed}-\beta_{Random}\right)$$

where V denotes the covariance matrix of the estimator. Selection between a random effects regression and a simple pooled regression is determined by the Breusch and Pagan Lagrange multiplier test, in which the null hypothesis is that cross-sectional variance components are zero. Furthermore, the overall-, between- and within-R^2s are equal to the squared correlation coefficients corresponding to the equations $\hat{y}_{it}=\hat{\alpha}+x_{it}\hat{\beta}$, $\hat{\overline{y}}_i=\hat{\alpha}+\overline{x}_{it}\hat{\beta}$ and

$\hat{y}_{it} - \hat{\bar{y}}_i = (x_{it} - \bar{x}_i)\hat{\beta}$, respectively, except that the within-R^2 is directly calculated from residuals in the fixed effect model.

Data

Model estimation dataset. The data used to estimate the model were accessed from the online database of the European Environment Agency. The specific data selected from the database were Chl-a, TP and TN to constitute a balanced panel for 2003 to 2005, comprising of 199 stations and 198 lakes selected from lakes in Switzerland (CH), Germany (DE), Denmark (DK), Hungary (HU), Italy (IT), Lithuania (LT), Latvia (LV), Netherlands (NL), Poland (PL) and Sweden (SE). The sample lakes, observation stations and years for which data are taken for analysis are shown in Table 4.1 and appendix A7. Year-averages were used in our analysis to eliminate localised noises due to seasonal variations. Additionally, we prepared a Land-Use Dummy (LUD) to associate the effect of land use with changes in actual lake Chl-a. Each lake station co-ordinate was projected in Google Earth (Google Inc.) to examine and capture 199 images of watersheds for classification based on land use type. LUD takes the value of 1 or 0: 1 for land occupied by artificial structures and farmland and 0 for natural area which includes shrubs, forests, bare land and mountains. Images that had roads through watersheds were also classified as 0. Examples of images for typical code 1 and 0 are shown in appendix A8.

Temperature is among the environmental factors which influence the concentration of Chl-a, low temperature being associated with reduction in Chl-a concentration (Haande et al 2011, Liu et al 2010b). However, we had to omit it from the model because temperature data were missing for many observation stations. Altitude is considered as a good proxy for temperature, but it is not included in the analysis due to it's high correlation with LUD.

Simulation datasets

To test the comprehensive applicability of the model across different locations in different countries, we performed simulations using data from lakes in the United Kingdom, Japan, Australia and the USA, which are selected for lakes located at high latitudes, in urban area, estuary, low and wetland, respectively. All the simulations were based on year averages. The Lakes used for simulation are shown in Table 4.1. Lake Kasumigaura and Ibanuma are located northeast of Tokyo, since the lake is closer to the metropolitan area, the latter lake is one of the most affected lakes in Japan. Lake Albert and Alexandrina are located south of Australia near the estuary of Murray River. Lakes in the United Kingdom are part of the general database for European lakes. Data on Florida lakes were obtained from the FloridaLAKEWATCH online database, covering 57 counties. Most of the Florida lakes have surface area of approximately 4 ha and depth of less than 5 m and 70% of them are without any surface inlet or outlet (Bachmann et al., 2012), a representative sample of lakes with less than 5 m depth often called "shallow lakes" whose trophic states vary considerably with different dynamics from deep lakes. The predictions of Chl-a for the Florida lakes was made for a sub-set of 8 lakes and for the entire database.

Table 4.1 Country, year, lakes and stations for European sample and simulation datasets.

Country	Year	Lake(s)	Station(s)
Europe (10 Countries)	2003-2005	198 (Appendix A)	199 (Appendix A)
			Stations as indicated in the parentheses
			after the lake name in their order are Agard (Furdeto),
			Balatonakali, Fertorakos, Gardony, Pakozd,
			373 - Piave - Lago Di Alleghe - Alleghe,
			369 - Po - Lago Di Garda - Brenzone,
			371 - Po - Lago Di Garda - Bardolino,
			374 - Piave - Lago Di Misurina - Auronzo Di Cadore,
			361 - Piave - Lago Di Santa Croce - Farra D`Alpago,
			Lago Di Scanno, East Part, East Part,
			Burtnieku - Vidusdala, Kanieris Z - Dala,
			Kisezers - Riga Preti Milgravja Caurtekai,
			Liepajas Ez. Pie Bartas Grivas,
			Raznas - Hidropostenis Kaunati, Vrouwenzand,
			Ketelmeer West, Markermeer Midden,
			Haringvlietsluis, Wolderwijd,
			Veluwemeer Midden, Eemmeerdijk 23,
			Jez. Biale Wlodawskie, Jez. Sremskie,
			jez. Tarnowskie Duze P - 01, j. Wuksniki st. 01,
			Skarsjon (ID 1149) and Skarsjon (ID SE633344-130068)
UK	2007-2010	Loch Ore	Loch Ore - Yacht Jetty Near Outlet
		Rescobie Loch	Rescobie Loch(Bankside)
		Loch Osgaig	Loch Osgaig Swcl
		Loch Scarmclate	Loch Scarmclate Swcl
		Loch a Bhraoin	Loch A Bhraoin (Wfd), W Of Boat House, Garve
		Loch Tarff	Loch Tarff (Wfd) - South Shore,
			S Of Eilean Ban Island, Fort Augustus
		Loch Strathbeg	Loch Of Strathbeg - S Shore Boat House,
			W Of Rattray
Japan	2000-2009	Ibanuma	Omonenomunekyou
			Jousuidoushusuiguchika
			Ichihonmatsushita
			Kitainbanumachuuou
		Kasumigaura	Kakeuma, Kihara, Ushigome, Takasaki,
			Tamatsukuri, Koshin, Nishinoshuu, Asou,
			Takei, Kamaya, Jinguubashi, Itako,
			Kamaya, Juiguubashi, Itako, Sotonasakaura,
			Ikisu, Hasaki and Yasujikawa
Australia	2008-2010	Albert	Opening, Meningie and SouthWest Point
		Alexandrina	Milang, Poltalloch Currency 1 and Currency 2
USA, Florida	2000-2009	Little Orange	
(Sub-set)		Wauberg	
		Beauclair	
		Dorr	
		Yale	
		Carlton	
		Giles	
		Holden	
USA, Florida	1986-2011	1988 lakes	
(Entire database)			

Results and discussion

Descriptive statistics

Table 4.2 Descriptive statistics for the variables.[a]

Sample	Chl-a (µg/l)		TP (mg/l)		TN (mg/l)		LUD[c]	
	Mean	SD[b]	Mean	SD	Mean	SD	Mean	SD
European Sample lakes[1]	11.231	16.727	0.038	0.052	0.893	0.820	0.819	0.385
UK Lakes[1]	23.888	27.163	0.054	0.077	0.888	0.799	0.714	0.488
Japan, Ibanuma and Kasumigaura[2]	69.424	21.944	0.108	0.019	1.389	0.696	1	0
Australia, Albert and Alexandrina[3]	57.759	34.833	0.160	0.075	2.936	0.947	1	0
USA, Florida Lakes (Subset)[4]	69.152	63.774	0.060	0.045	1.791	1.143	1	0
USA, Florida Lakes (Entire database)[4]	18.852	43.013	0.045	0.111	0.832	0.741	UC[d]	UC[d]

[a] All for the entire period of observation.

[b] SD – Standard Deviation.

[c] LUD – Land-Use Dummy, 1 for lakes which are surrounded by artificial structures and 0 otherwise.

[d] UC – Unclassified lakes.

Data sources

[1] Available at

http://www.eea.europa.eu/data-and-maps/data/waterbase-lakes-8

[2] Available in Japanese at

http://www.pref.chiba.lg.jp/suiho/kasentou/koukyouyousui/data/ichiran-koshou.html

http://www.ktr.mlit.go.jp/kasumi/kasumi00014.html

[3] Available at

http://www.epa.sa.gov.au/environmental_info/water_quality/lower_lakes_monitoring/reports_and_data

[4] Available at

http://lakewatch.ifas.ufl.edu/Lakewatch_County_Data.HTM

The range of annual mean TP and TN for the analysed panel at different stations in the European sampled lakes was between 0.002 and 0.334 mg/l with a mean of 0.04 mg/l for TP and between 0.13 and 4.883 mg/l with a mean of 0.89 for TN, whereas Chl-a concentration was between 0.2 and 121.7mg/l with a mean of 11.23 μg/l (Table 4.2).

Positive correlations were observed between Chl-a and TP and, Chl-a and TN, as shown in figure 4.1 and 4.2. TP had a higher correlation with Chl-a than TN. These results are in line with other studies on correlations in Chl-a-TP-TN (e.g., Liu et al., 2010b; Brown et al., 2000). In many lakes, phosphorus is the limiting nutrient for primary production (Paerl et al., 2011; Bechmann et al., 2005; Arhonditsis et al., 2003), but nitrogen is also a limiting nutrient for phytoplankton growth in some lakes (Gunkel and Casallas, 2002). Our goal was to explore if a solid model could be constructed to describe the relationship among Chl-a, TP and TN. It is noteworthy that the large variations in the three concerned variables in our dataset were mainly arising from variations among stations located in different countries, as shown by the coefficient of variation in Table 4.3. The variations among countries were half of the variations among stations for all the three variables. Variations among years were low. This heterogeneity between stations and countries requires the use of panel regression in estimating the model.

Table 4.3 Standard deviations of Chl-a, TP and TN by station, country and year.[a]

Variables	Variation among Stations	Countries	Years	Overall
Chl-a (μg/l)	15.973	10.896	0.380	16.727
	(142)	(76)	(3.4)	(149)
TP (mg/l)	0.051	0.042	0.003	0.052
	(132)	(69)	(7.3)	(136)
TN (mg/l)	0.805	0.584	0.032	0.820
	(90)	(45)	(3.5)	(92)

[a] The figures in parentheses are coefficients of variation (%).

Regression

Table 4.4 Panel regression of Chl-a determinants, 2003-05 (N=597).[a]

Variables	Fixed effect I Coef.	SE.		Random effect II Coef.	SE.		Random effect III Coef.	SE.	
TP	180.9	16.99	***	179.2	12.91	***	177.8	12.94	***
	(0.612)			(0.606)			(0.602)		
TN	4.776	1.365	***	4.523	0.893	***	4.365	0.898	***
	(0.379)			(0.359)			(0.347)		
LUD							2.802	1.986	
							(0.204)		
Const.	0.012	1.332		0.302	1.021		-1.794	1.802	
σ_u	10.55			10.14			10.11		
σ_ε	5.346			5.346			5.346		
θ				0.709			0.708		
R^2 :									
within	0.2557			0.2556			0.2556		
between	0.5641			0.5641			0.5688		
overall	0.5359			0.5360			0.5402		

[a] The figures in parentheses are the elasticity at the means. ***, ** and * denote the significance of the coefficients at the 0.01, 0.05 and 0.1 critical level, respectively. The random effect model is estimated by GLS regression. Hausman test of the fixed- and random-effects (model I vs. II) cannot reject the null hypothesis of difference in coefficients (p-value = 0.9668). Breusch and Pagan Lagrangian multiplier test for model II shows that the null hypothesis of Var(u) = 0 is rejected (p-value<0.000).

Table 4.4 presents results of the regression analysis generated by STATA 12.1 statistical package. The analysis was performed for the fixed and random-effect models. Elasticity of each estimated coefficient is shown in parenthesis below the coefficient. Results show that TP and TN were all significant ($p < 0.05$). LUD had a coefficient of 2.802 in Model III, but not significant (p-value = 0.158). The elasticity for TP and TN was about 0.60 and 0.35 in all the estimated models respectively. Our results of the multiple regressions are consistent with the results reported by most previous studies that used data from temperate regions. Variations exist, however, for the results of the studies that include tropical lakes, such as studies by Trevisan and Forsberg (2007) and, Huszar et al. (2006) where the coefficient of TN is reported to be larger than TP. The result of the Hausman test (p-value = 0.9668) for the fixed and random effect models was in favour of the latter, so we used the random effect model in the simulations.

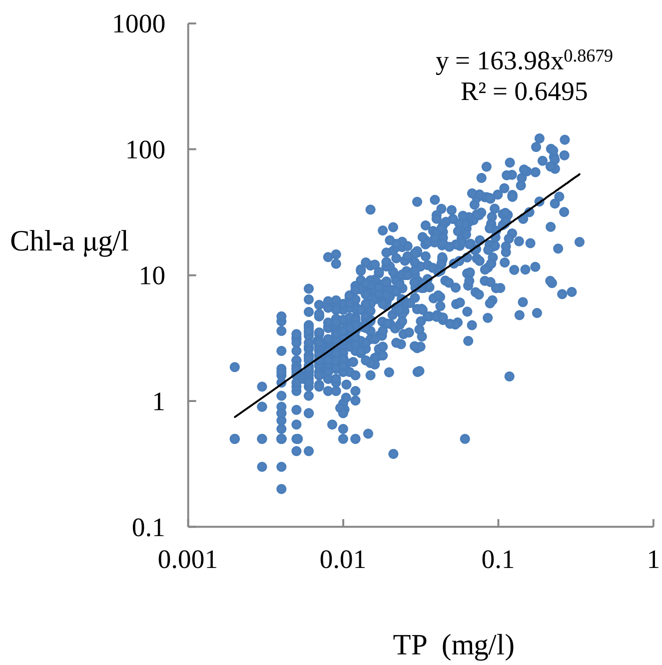


Fig. 4.1 Correlation between Chl-a and TP for European sample lakes

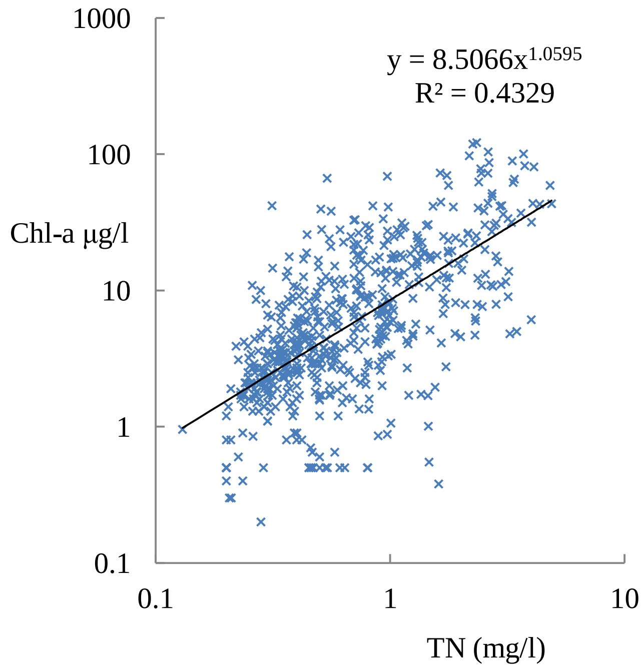


Fig. 4.2 Correlation between Chl-a and TN for European sample lakes

Simulations

The results of simulations to predict the level of Chl-a from the levels of TP and TN of various lakes in different countries using our Model II are shown in figure 4.3 to 4.6, the graphs depict the predicted Chl-a (vertical axis) against the actual data (horizontal axis) in standardized units. The data used for the simulations are presented in Table 4.1 by country.

The test samples from lakes in the United Kingdom provided the best prediction results in the simulation tests with an R^2 of 0.83. The R^2 of the simple regression between the predicted and actual Chl-a values for Lake Ibanuma and Kasumigaura in Japan is 0.63. A similar result is obtained for combined data of Lake Albert and Alexandrina in Australia (R^2 = 0.65). For these three cases, our model predicts lake Chl-a reasonably well, the variation of predicted Chl-a ‘explains’ more than 60% of the variation of the actual Chl-a.

Our model that uses only two explanatory variables has unexplained part, which could be explained by other variables that influence Chl-a, though not as strongly as TP and TN, yet evidently. Temperature is an important variable (Liu et al., 2010b; Ogbebo et al., 2009) but it is not included in our model due to lack of temperature data for many lakes in the database. Lake geochemistry, which is difficult to study on a wide scale, is another variable that influences the Chl-a – TP - TN relationship (Trolle et al., 2009). Annual precipitation (Liu et al., 2010a) and atmospheric deposition of nitrogen and phosphorus are also factors that may lead to changes in the predicted Chl-a (Muyodi et al., 2010; Morales-Baquero et al.,2006).

The results for Florida lakes in the USA are different from other simulations. The model predictions for a subset of the overall data give results similar to the other three cases with R^2 of 0.66 (Figure 4.6, Table 4.1). However, the application of the model to the entire database which includes 1988 lakes is not successful with a low R^2 of 0.24 due to the heterogeneous nature of the variables in this database (FloridaLAKEWATCH) with varying concentrations. All Florida lakes are situated in densely populated areas. There were no significant changes in the predicted Chl-a using model III, partly due to the negative correlation between the LUD and TP. This correlation reflects the decline in TP level for more than 30 years with an increasing population (Terrell et al., 2000). Differences in geology and hydrology in the Florida lakes as described by Bachman et al. (2012) also explain the low predictability of our models.

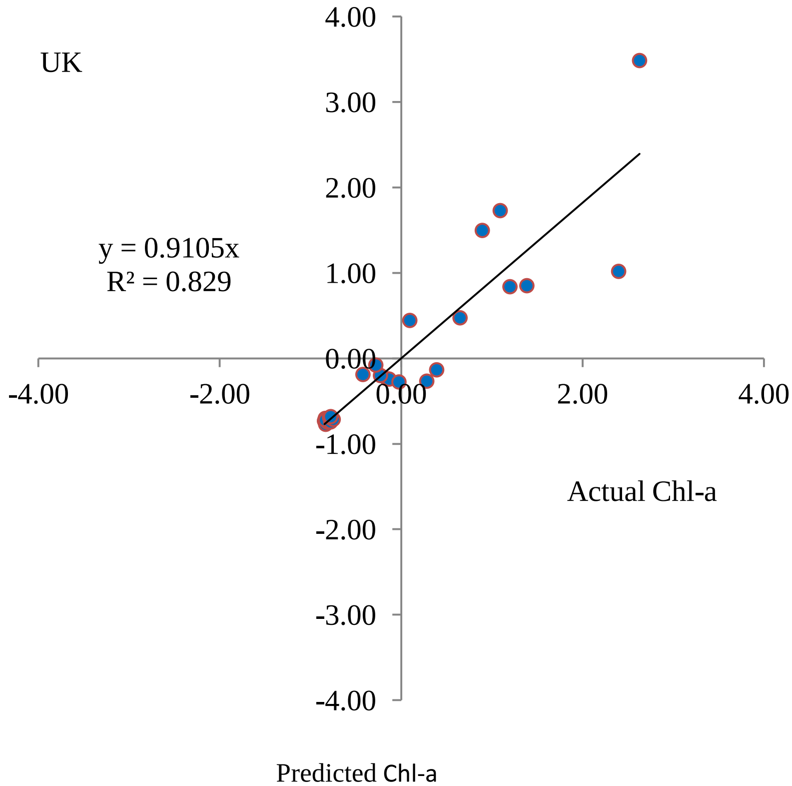


Fig. 4.3 Simulations results of the random-effect model for sample lakes in the United Kingdom

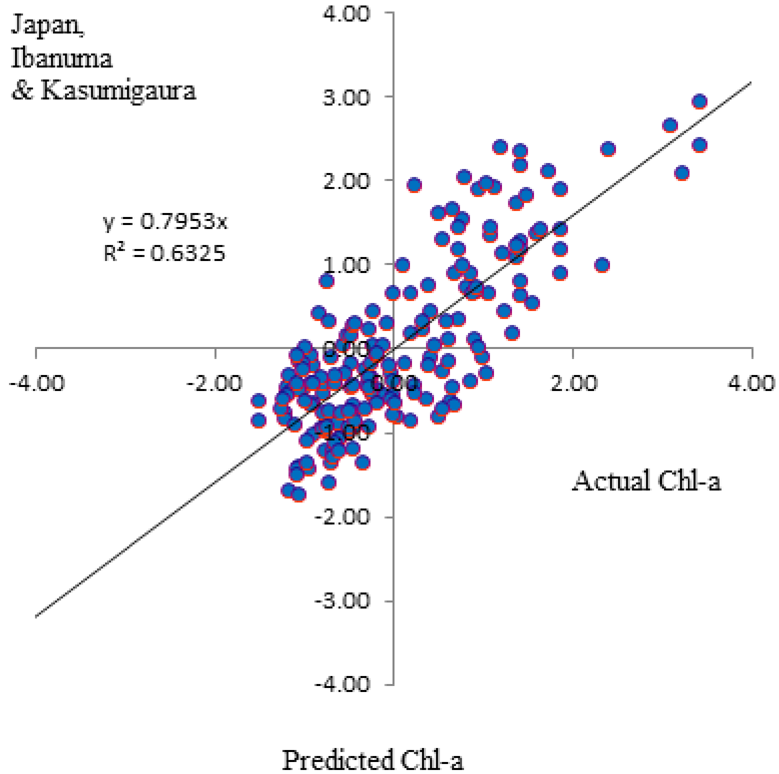


Fig. 4.4 Simulation results of the random-effect model for Lake Ibanuma and Lake Kasumigaura

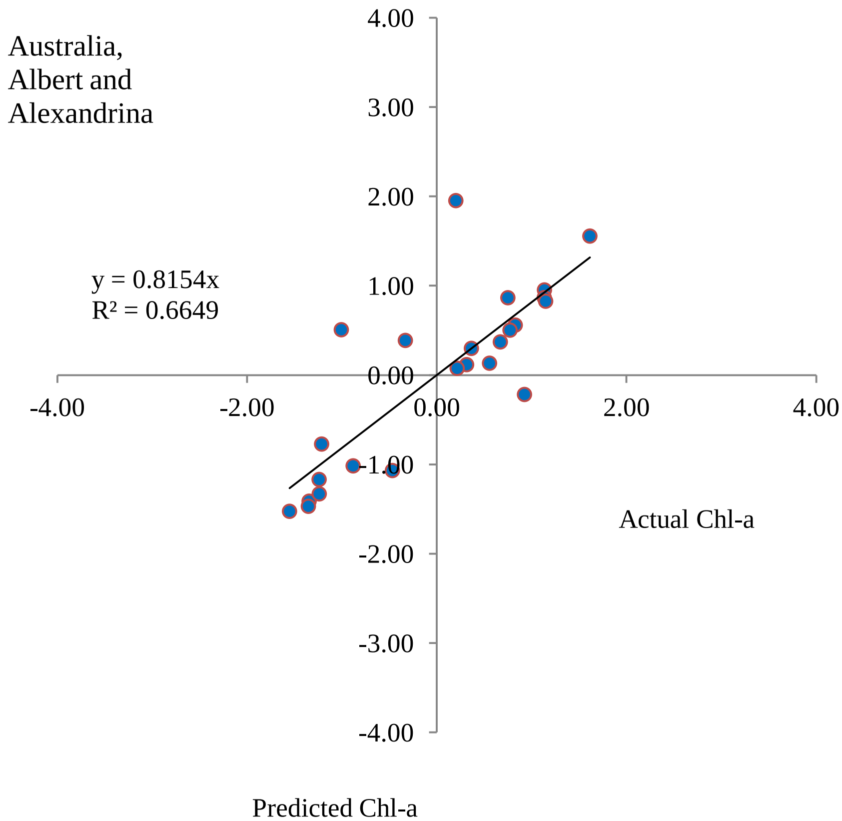


Fig. 4.5 Simulation results of the random-effect model for Lake Albert and lake Alexandrina, Australia using standardized values.

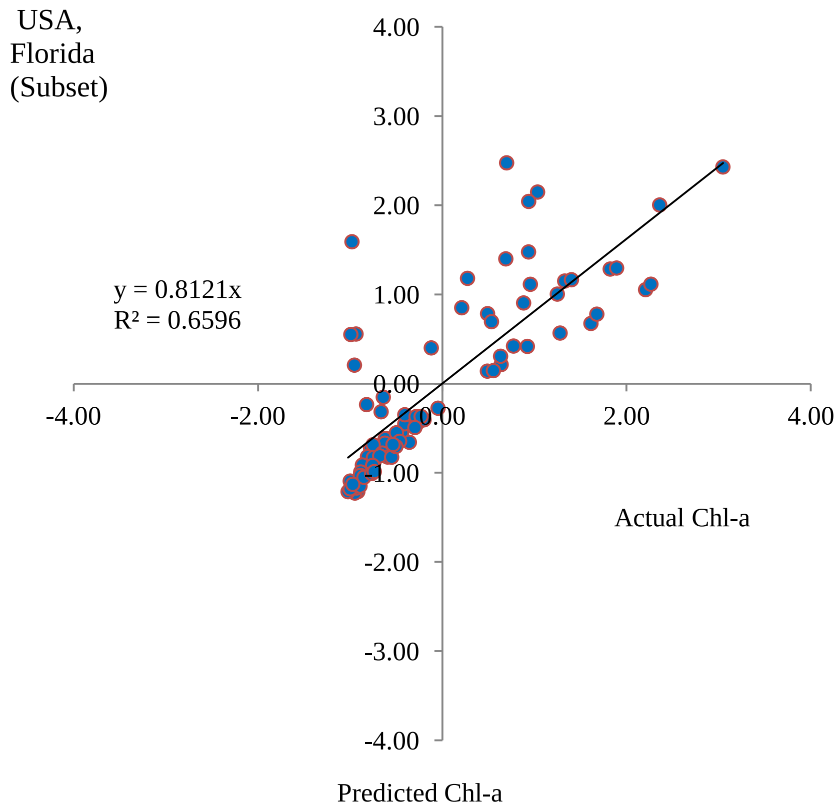


Fig. 4.6 Simulation results of the random-effect model for Subset of Florida Lakes, USA

The four figures in the simulations of random –effect model are for standardized values.

Conclusion

In this paper, we have developed a comprehensive and robust model to describe the correlations among Chl-a, TP and TN, using a multi-country dataset including 198 lakes in 10 European countries. The model was applied to lakes in other countries and continents with reasonably satisfactory results. Our model evidently shows that TP and TN are good predictors of Chl-a and will supplement existing models. Further investigation is needed to include the impacts of other factors, such as temperature, light intensity and water residence time. These variables incorporated, the model would become a powerful tool to study and regulate water quality of lakes in a wider perspective.

4.2.2 DYNAMIC PANEL MODEL

Abstract

The major objective of this paper is to show how the estimated parameters in empirical models to predict Chlorophyll-a (Chl-a) in lakes by the concentration of total phosphorous (TP) and total nitrogen (TN) could be different from those estimated by using the conventional ordinary least squares (OLS) method, if station-specific effects and auto-regressive effects of Chlorophyll-a are taken into account in the regression estimation, by applying respectively, the random effect panel estimation for the former and the dynamic panel estimation for the latter. Our estimation based on water quality data from European lakes reveals that the OLS estimation gives comparable parameters to those of many earlier studies, in which both TP and TN are significant determinants of Chlorophyll-a and the elasticity of TP is much larger than that of TN. The application of the non-conventional estimation methods alters this parameter structure radically. The station-specific effects being controlled, TN/TP is not a significant factor in determining the Chlorophyll-a concentration, and the elasticity of TP is far smaller than in the conventional estimation. The inclusion of auto-regressive effects, in addition, makes TN insignificant, leaving TP as the only significant parameter. These results suggest a strong need to take station-specific effects as well as previous concentrations of Chlorophyll-a into consideration in the study of the Chl-a = f (TP, TN) relationship.

Introduction

Chlorophyll-a (Chl-a as a variable name) is the most widely used measure of phytoplankton biomass in lakes, the relationships between Chlorophyll-a, total phosphorus (TP) and total nitrogen (TN) in water bodies have been intensively studied (Søndergaard et al., 2011; Brown et al., 2000; Reckhow, 1993) since it was first introduced by Sakamoto (1966). It is well established that the concentrations of TP and TN, together or in isolation, affect the concentration of Chlorophyll-a, though the Chl-a = f (TP, TN) relationship varies according to factors such as latitude, altitude, depth and stoichiometric characteristics of lakes (Abell et al., 2012; Gunkel and Casallas, 2002). A comprehensive model of this relationship that could be applied to lakes with different characteristics would be invaluable for developing effective policies to control lake water quality (Abell et al., 2012; Philips et al., 2008).

Various attempts have been made to establish a reasonable empirical model for the Chl-a = f (TP, TN) relationship, controlling other factors by using regression techniques (Table 4.5). Each of the authors in Table 4.5 adopted the ordinary least squares (OLS) method, except Reckhow (1993), who applied the random coefficient method. The data used for the regression analyses in this field of study are observations of Chlorophyll-a, TP, TN, and other control factors, collected from stations (sampling sites) in many lakes at a time point or, more often, over time. One problem of applying OLS to this type of data is failure to control station-specific differences within a lake, or between different lakes in a country or even in different countries. Another, even more serious problem stems from the time-series nature of the data: the concentration of Chlorophyll-a at a certain point in time depends not only on TP, TN and the factors mentioned above, but also on previous concentration. Failing to adequately address these two problems makes it difficult to accurately estimate a statistical relationship for Chl-a =

f (TP, TN). To the best of our knowledge, no study to date has taken these problems into account in a regression analysis.

Using lake water quality data provided by the European Environment Agency, this paper estimates the parameters of the Chl-a = f (TP, TN) relationship in three types of regressions: pooled OLS, random effect panel that deals with station-specific variations, and dynamic panel that takes into account the auto-regressive effects of phytoplankton biomass as well as the station-specific variations.

Table 4.5 Coefficients of TP and TN for the Chl-a = f(TP, TN) relationship estimated in previous studies.

Author(s)	Methods of estimation	Data sources	Lakes and study period[a]	Location	Coefficients[b] TP	TN
Abell et al. (2012)	OLS (pooled)	European Environment Agency and previous studies	1316 1965-2007	Between 70^0S and 83^0N, lakes in over 30 countries	*0.81*	*0.84*
Bachmann et al. (2012)	OLS(pooled)	Florida LAKEWATCH and other reports	1001 30 years	Florida, USA	0.71 *0.93*	0.52 *1.38*
Sondergaard et al. (2011)	OLS (pooled, time series)	Danish local government	440 1989-2008	Denmark	0.80 *0.95*	0.35 *1.01*
Phillips et al. (2008)	OLS (pooled)	National data archives	1138 1988-2004ǂ	Europe (16 Countries)	0.79 *1.03*	0.32 *1.36*
Trevisan and Forsberg (2007)	OLS (cross-section)	Field samples	20Ŧ	Amazonia basin, Brazil	0.31 *0.72*	0.63 *0.005*
Huszar et al. (2006)	OLS (pooled)	Published articles and unpublished data	136 lakes and 56 reservoirs 1980-2004	Between 31^0 N and 30^0 S	0.41 *0.70*	0.50 *0.94*
Brown et al. (2000)	OLS (pooled)	Florida LAKEWATCH	273 1986-1997	Florida, USA	0.91 *1.05*	0.32 *1.21*
Reckhow (1993)	GLS - Random Coefficient (cross-section)	North Carolina	63 1981T	North Carolina, USA	0.78 *0.71*	0.32 *0.58*

a) ǂ Results represent nutrient-chlorophyll-a relationships for summer.
Ŧ Sampling period: November 1999 - January 2000.
T Of the parameters for 63 lakes presented, the median is shown here.
b) Italics indicate coefficients for single predictor regression.

Methods

Regression models

The following three models are used to examine the Chl-a = f(TP, TN) relationship:

Model I Pooled OLS: $Y_i = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2\mathbf{X}_i + u_i$

Model II Random effect Panel: $Y_{it} = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2\mathbf{X}_{it} + e_i + u_{it}$

Model III Dynamic panel: $Y_{it} = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2\mathbf{X}_{it} + \boldsymbol{\beta}_3\mathbf{K}_{it} + \boldsymbol{\beta}_4\mathbf{L}^{j}.\mathbf{Chl\text{-}a}_{it} + u_{it}$

where i and t are used to identify the station and time, Y is the concentration of Chlorophyll-a, **X** is a vector of variables which affect Chlorophyll-a, and u is a conventional error term with $N(0,\sigma_u)$. In Model I, observations obtained at different time points are pooled without regard to time. In Model II, which is a "random effect panel model" of panel data analysis, the term e captures station-specific effects and is assumed to be a random variable with $N(0,\sigma_e)$. In Model III, **X** is a vector of the endogenous variables, **K** is a vector of the exogenous variables, **L.Chl-a** is a vector of the lagged variables of Chlorophyll-a, and j refers to a vector with Chl-a_{t-1}, $\text{Ch-a}_{t-2,}$ ………, Chl-a_{t-j}. The endogenous variables can be correlated with the error term, while the exogenous variables are determined outside the model. The variables were transformed to logarithm before statistical analysis. A detailed explanation of the random effect panel analysis used in this study can be found in Hsiao (2007) and Oscar (N.d.), and further information on the dynamic panel analysis is available in Arellano and Bond (1991) and Roodman (2009). All regression analyses were carried out using STATA 12.

Data set

The data used in our analysis were obtained from the online database of the European Environment Agency, which is one of the most detailed and longest online databases on the status of European lakes. We selected data on yearly averages of Chlorophyll-a, TP and TN and altitude for the period between 1965 and 2009 (Fig. 4.7) for countries within the European continent. For lakes with no altitude data, we obtained the data from Google Earth. The original data set thus compiled consists of 3055 observations for 454 stations in 396 lakes with an average observation period of 6.7years per station. We then computed the partial auto-correlation coefficient for Chlorophyll-a (Fig. 4.8) and found that it correlates positively with the previous level, i.e, the first-order lag. We therefore decided to model the first order auto-regression process, AR(1), for dynamic panel analysis. Using the first-year lag for the sample lakes reduces the observations to 2601 (3055-454) for the period 1965-2009, and this data set is used for our regression estimation for the three models. For Model III, the data for the first and second years are used as the instrumental variables, based on the results of our preliminary regressions (Arellano and Bond, 1991).

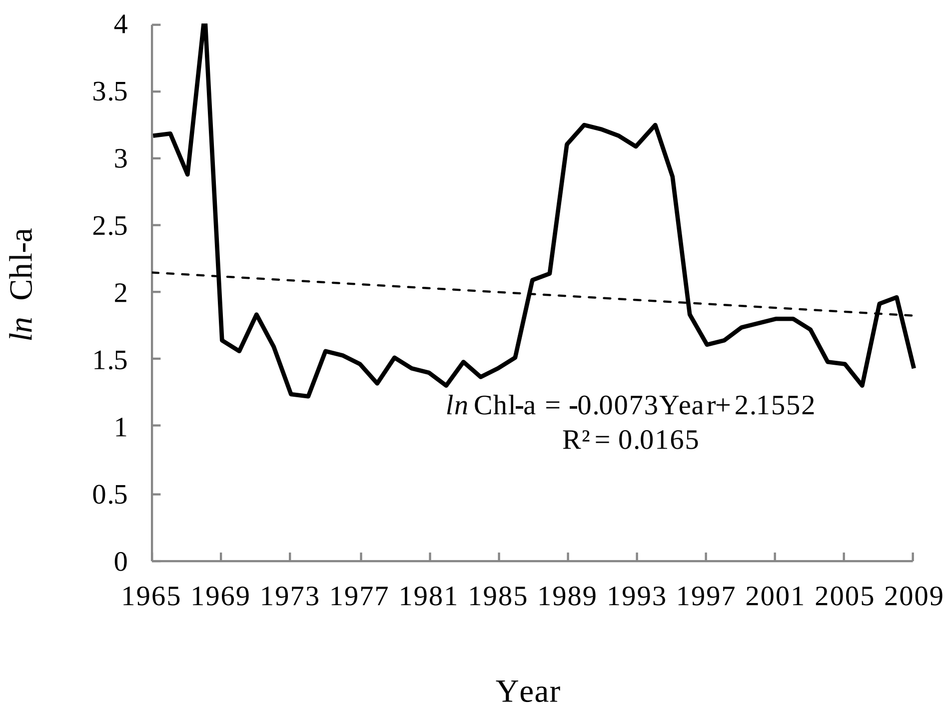


Fig.4.7 Changes in mean Chlorophyll-a, 1965-2009 (N=2601)

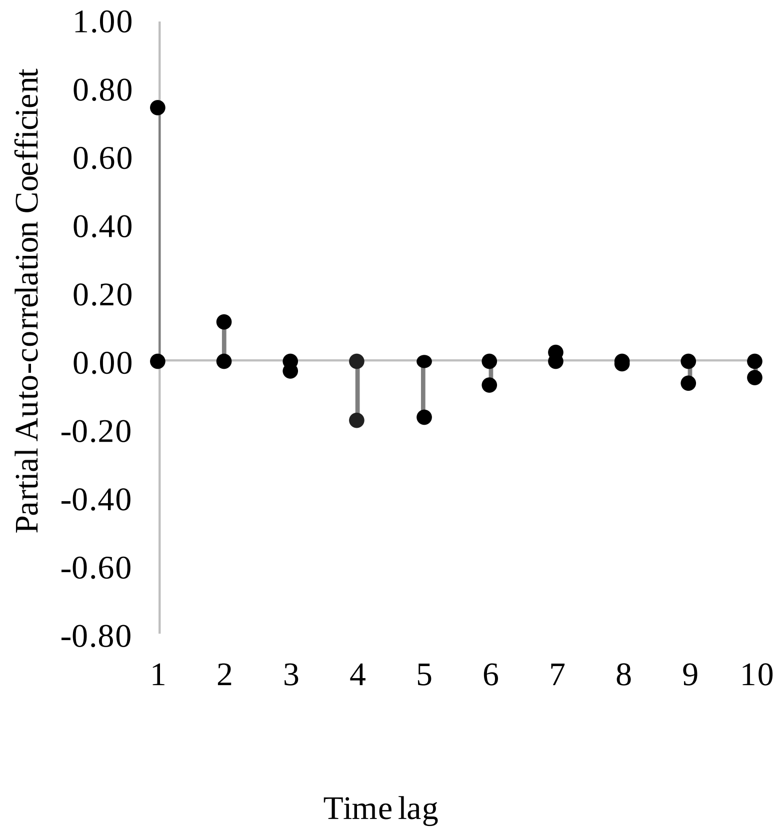


Fig. 4.8 Partial autocorrelation coefficient and time lag for *ln* (Chl-a), 1965-2009 (N=2601)

Explanatory variables

TP, TN, TN/TP, altitude and year are the explanatory variables used in our regression analyses. All of these variables are treated as **X**-variables in Models I and II. In Model III, altitude and year are treated as **K**-variables, but TP, TN and TN/TP are treated as **X**-variables, i.e., the endogenous variables. In addition to these variables, one lagged dependent variable is included as an explanatory variable in Model III. The variable TN/TP is included in the analysis because it has been widely used as an index to identify limiting nutrients (Bachmann and Hoyer, 2003). Altitude is introduced to control the effect of temperature on the concentration of Chlorophyll-a (Carvalho et al., 2009), since temperature data are missing for many lakes. Year is included to examine the existence of time trend in the variation of Chlorophyll-a even after other determinants are controlled.

Results

The means and standard deviations of Chlorophyll-a, TP, TN, TN/TP and altitude, and the simple correlation coefficients among them are shown in Table 4.6. In the over-all variation, Chlorophyll-a varied more than the explanatory variables. The overall variation is divided into variations ‘between stations’ and variation ‘within stations’. For TN, in particular, more than 70% of variation stems from variation between the stations. The simple correlation coefficients between Chlorophyll-a and TP and, between Chlorophyll-a and TN are shown in fig. 4.9 and 4.10), and the elasticity of TP with respect to Chl-a, i.e., ∂ln Chl-a/∂ln TP, is larger than that of TN. In these respects, our data set shares a similar structure with many past studies (e.g., Abell et al., 2012). The time-series of mean Chlorophyll-a values for the sample European lakes shows large fluctuations over time and a slight downward trend (Fig. 4.7).

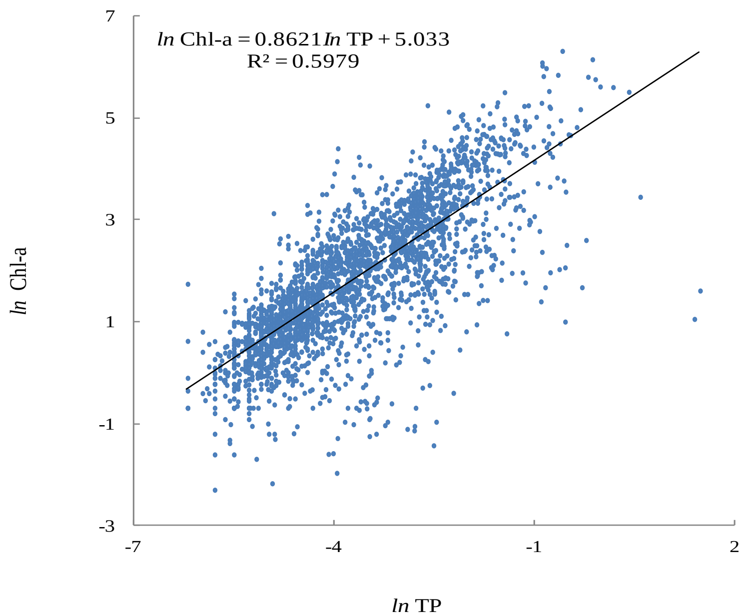


Fig. 4.9 Correlation between Chlorophyll-a and TP (N=2601)

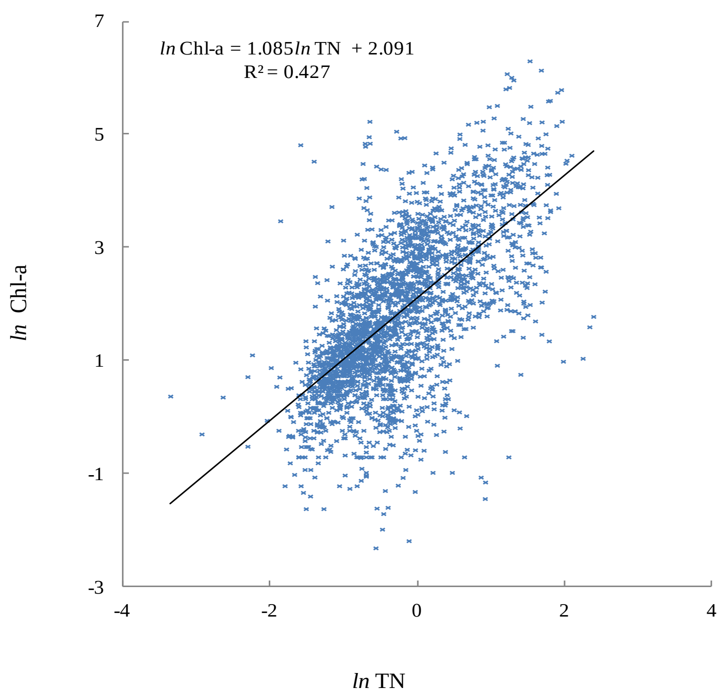


Fig. 4.10 Correlation between Chlorophyll-a and TN (N=2601)

Table 4.6 Statistics for the variables (N=2601)[a]

Variables	Unit	Stat.	Overall	Between	Within	Correlation coefficients				
						Chl-a	TN	TP	TN/TP	Altitude
Chl-a	μg/l	Mean	16.27			1.00				
		Std.Dev.	36.22	43.26	16.81					
TN	mg/l	Mean	1.11			0.48	1.00			
		Std.Dev.	1.14	1.060	0.360					
TP	mg/l	Mean	0.05			0.69	0.65	1.00		
		Std.Dev.	0.08	0.08	0.04					
TN/TP		Mean	41.86			-0.25	-0.09	-0.36	1.00	
		Std.Dev.	33.3	38.55	16.3					
Altitude	1000m	Mean	0.17			-0.18	-0.24	-0.21	0.25	1.00
		Std.Dev.	0.26	0.39	N.A.					

a) For 2601 observations of 454 stations for an average time period of 5.7years per station. "Between" standard deviation is calculated for mean values of stations, while "within" standard deviation is calculated for each station.

Table 4.7 Regression results of Chl-a determinants (N=2601)[(a]

Variables	Model I (OLS)		Model II (Random effect panel)		Model III (Dynamic panel)	
	Coefficient	Robust S.E.	Coefficient	Robust S.E.	Coefficient	Robust S.E.
ln (TP)	0.655	0.056 ***	0.386	0.04 ***	0.361	0.044 ***
ln (TN)	0.203	0.067 ***	0.211	0.045 ***	N.S.	
TN/TP	-0.002	0.001 **	N.S.		N.S.	
Altitude	-0.889	0.094 ***	-0.923	0.155 ***	-0.495	0.135 ***
Year	N.S.		N.S.		N.S.	
ln ($Chl\text{-}a_{t-1}$)	N.A.		N.A.		0.527	0.047 ***
Constant	14.85	5.66 ***	3.282	0.165 ***	2.29	0.219 ***
R^2 [(b]	0.66		0.64		0.83	
Elasticity:						
TP	0.699		0.386		0.361	
TN	0.158		0.211		N.A.	

a) *** and ** denotes the significance of the coefficients at the 1% and 5% critical level, respectively. N.S. denotes "not significant" and N.A. denotes "Not Applicable".

b) R^2 is simple correlation coefficients between predicted and actual values.

Table 4.7 presents the results of the estimations. For all the models, coefficients that are significant at 5% critical level or higher are shown. For the standard errors of the coefficients, we use robust standard errors, which produce reliable results for t-tests even when the error term has heteroscedasticity (Greene, 2011).

For the pooled OLS model, all of the coefficients of the explanatory variables are statistically significant at 5% except year. The coefficients of *ln* TP and *ln* TN are positive, while those of TN/TP and altitude are negative. The partial elasticity of TP, ∂ln Chl-a/∂ln TP, is estimated to be 0.699 and that of TN is 0.158. For the random effect panel model, the coefficients of *ln* TP, *ln* TN and altitude are significant, but the coefficient of TN/TP is not. For the dynamic panel model, the coefficient of *ln* TP remains significant, but that of *ln* TN becomes insignificant, as does TN/TP. Altitude produces a significant coefficient. The lagged dependent variable also results in a positive significant impact on the level of Chlorophyll-a.

R^2, which is the squared correlation coefficient between predicted and actual values, is shown for the purpose of comparing the statistical performance among the models. It is particularly high for the dynamic panel data model.

Discussion

Most remarkable in the results of our regression estimation is that although all the three models reveal that TP is a significant determinant of Chlorophyll-a, the magnitude of its impact varies considerably across the models: in terms of elasticity, it is reduced almost in half in the random effect panel data and the dynamic panel models from the level estimated in the OLS model.

We included two control variables in the estimation of the Chl-a = f (TP, TN) relationship, altitude and year. Altitude is included in our analysis as a proxy for temperature. Carvalho et al. (2009) and Liu et al. (2010b) reported that temperature is an important determinant of Chlorophyll-a: higher temperatures increase the Chlorophyll-a concentration in lakes. Since altitude and temperature are inversely correlated, our results in Table 4.7 confirm the results of these earlier studies. Few studies in this field have incorporated temperature and altitude in linear models, but the impact of temperature as shown by the coefficient of altitude, is too large to ignore. The coefficient of year is not significant for any of the models, which is in contrast to the results of simple time series analysis shown in Fig. 4.7. Søndergaard et al. (2011), Acker et al. (2009) and Räike et al. (2003) found some significant time trends in Chlorophyll-a for lakes and rivers, but their results were obtained using estimation methods based on a simple correlation between Chlorophyll-a and time as in Fig. 4.7.

The most important variables in the estimation of the Chl-a = f (TP, TN) relationship are TP, TN and TN/TP. Most previous studies, including those listed in Table 4.5, found that both TP and TN are significant determinants for Chlorophyll-a. Some studies pointed out that phytoplankton growth was TP-limited (e.g., Arvola et al., 2011; Lv et al., 2011; Wang et al., 2007; Bechmann et al., 2005; Jeppesen et al., 2005; Arhonditsis et al., 2003; Brown et al., 2000), while others indicated that TN is a limiting factor (Trevisan and Forsberg, 2007; Gunkel and Casallas, 2002). It has also been argued that the relative ratio of these nutrients, i.e., TN/TP, is an important factor, in conjunction with TP and TN (Jin and Hongjuan, 2010; Guildford and Hecky, 2000), as an indicator of nutrient limitation. According to Guilford and Hecky (2000), a high TN/TP indicates phosphorus limitation and a lower ratio indicates nitrogen limitation.

Our OLS estimation indicates that TP, TN and TN/TP are all significant factors in determining the concentration of chlorophyll-a and that the magnitudes of the estimated elasticity for TP and TN are in line with estimates from the previous studies shown in Table 4.5. From this evidence, one might conclude that the concentration of Chlorophyll-a in European lakes is determined by both TP and TN and by their relative abundance, and that the estimated OLS model can be used to predict the concentration of Chl-a for the management of the water quality of the lakes.

However, adopting a random effect panel regression or especially a dynamic panel regression radically alters the estimated parameter structure. Once the station-specific effects are controlled in Model II, TN/TP is not a significant factor in determining the concentration of Chlorophyll-a, which means that the effect of TN/TP is easily affected by station-specific conditions and that the elasticity of TP is far less than in the OLS model. Taking into account the effect of the previous Chlorophyll-a and the station-specific effects, Model III shows that TP is the only significant factor in determining Chlorophyll-a, and that the elasticity of TP is as small as that estimated in the random effect panel model.

In Danish lakes, Søndergaard et al. (2011) found that the elasticities of TP and TN are large and comparable to those of our sample European lakes (Figs. 4.9 and 4.10), when these nutrients are used as the single predictor. However, when both of them are included in a multiple OLS regression, their elasticities, particularly TN's, become considerably smaller (Table 4.5). Further refinements in the regression models change the parameter structure and lower the elasticities for these nutrients. If the better-performing Model III is to be used for prediction and management of lake water quality, the scenarios obtained could be very different from those of the simple predictor model or the pooled OLS model.

Conclusions

In this paper, we examined empirical models used to predict Chlorophyll-a in lakes by the concentration of total phosphorous and total nitrogen. The estimated parameters differ from those estimated using the conventional ordinary least squares (OLS) method when station specific-effects and auto-regressive effects on Chlorophyll-a are taken into account. Our estimation based on water quality data from European lakes reveals that the OLS method gives parameters comparable to those of many earlier studies in which both TP and TN were found to be significant determinants of Chlorophyll-a.

Appling the non-conventional estimation methods radically alters this parameter structure. If station-specific effects are controlled, TN/TP is not a significant factor in determining the Chlorophyll-a concentration, and the inclusion of auto-regressive effects in addition to the station-specific effects makes TN insignificant leaving TP as the only significant parameter. The elasticity of TP estimated by non-conventional methods is far smaller than by the conventional estimation. In order for the Chl-a = f (TP, TN) relationship to be used as an operational tool to predict the concentration of Chlorophyll-a in lakes, it is imperative to clarify the bio-chemical mechanism behind the relationship. Our study shows that further research beyond the conventional OLS model is necessary for this relationship to be used in developing effective policies to control lake water quality.

4.3 FORECASTING: A CASESTUDY OF JAPANESE LAKES

4.3.1 Abstract

The goal of the paper is to test the performance of Ordinary Least Squares (OLS) and dynamic panel models for predicting Chlorophyll-a in lakes. Dynamic panel model is a new proposed model developed from a sample of 396 European lakes with a pronounced difference that only phosphorus is the only significant variable in determining Chlorophyll-a in lakes when panel and station-specific effects are considered contrary to earlier studies using OLS which show that phosphorus and nitrogen significantly determine chlorophyll-a. The coefficients of OLS and those of the new model are applied. Performance of the models was simulated using annual mean data (2000-2009) from 8 Japanese lakes in three regions. The lakes were grouped based on altitude and chlorophyll-a concentrations. The two models show higher performance measured by R^2 values with variations in certain cases. The highest R^2 was 0.94 and 0.89 for dynamic panel and OLS models respectively. Simulation results show that the dynamic model is superior to OLS model and it is observed that the dynamic model is more applicable for spatial estimation of Chlorophyll-a in multiple lakes.

4.3.2 Introduction

The purpose of this section is to test the performance of the Ordinary Least Squares (OLS) model (Model A) and the Dynamic panel model (Model B) in forecasting Chlorophyll-a (Chl-a) for Japanese lakes. The most popular indicators assessed when monitoring water resources are chlorophyll-a (Søndergaard et al., 2011) and secchi-disk transparency (Olmanson et al., 2008). Models and model coefficients are necessary where there are insufficiencies in data for phosphorus and nitrogen or Chlorophyll-a that may prevent whole lake or multiple lake assessments or in cases where forecasting is necessary from data samples. Chlorophyll-a is related with reflectance data of satellite images, a relationship which enables mapping of Chlorophyll-a for single or multiple lakes in remote sensing as decision support tool (Allan et al., 2011; Oyama et al., 2009; Östlund et al., 2001).

In Chlorophyll-nutrient relationships, it is known that phosphorus and nitrogen are the main predictors of Chlorophyll-a (Bachmann et al., 2012; Huszar et al., 2006; Reckhow, 1993), other factors such as temperature and latitude also influence the level of Chlorophyll-a (Abell et al., 2012; Liu et al., 2010b). The studies reporting coefficients of variables in these relationships apply OLS, in many, expect a few, the coefficients for phosphorus are higher than that of nitrogen. The parameter structure changes when station-specific and auto-regressive effects of chlorophyll-a are considered in the dynamic panel model.

4.3.3 Method and data

The heterogeneity inherent in data due to station-specific effects if not controlled results in parameters which may not really portray the data structure from multiple lakes across varying geographical areas, panel data analysis is applied to address this problem. OLS and the dynamic panel model, a new model which is based on auto-regressive effects for the first lag in this case, were developed from 396 European sample lakes (N=2601). The regression results show that all the variable, phosphorus, nitrogen, altitude and year were significant in OLS model. The results indicated that the dynamic model was superior to pooled OLS based on the R^2 of 0.83 which is the squared value of correlation coefficient between predicted and actual chlorophyll-a. Only phosphorus was the significant variable and, altitude as a proxy for temperature in the dynamic panel model which changed the parameter structure.

We include other lake characteristics which might have influence on performance of the models. Models are tested for groups of lakes based altitude and for the combined data set. The coefficients from OLS and dynamic panel model are used to estimate Chlorophyll-a (Table 4.7).

Data for simulation is from 9 lakes in four prefectures, a representative sample of Japanese lakes. Simulations here are made for Japanese lakes using data for a 10year period using yearly mean values for 2000-2009. Lake Kasumigaura and Ibanuma are located close to Tokyo, Lake Biwa the largest lake in Japan in central while Nagano lakes in the west and mountainous areas of Japan (Table 4.8). The predictions were made using coefficients of Model A and B from Table 4.7. Figures are presented as relationships between predicted *ln* Chl-a and actual *ln* Chl-a and R^2 is a correlation coefficient to show the performance of each model. It should be noted however, that the coefficient for year in Model A was significant at 10%, in the earlier paper where the first estimates of Model A are made, only variables significant at 5% were reported, including this coefficient sets the predicted values to be comparable with estimates of Model B.

4.3.4 Results

Table 4.8 shows lake characteristics and mean for the variables used in simulations. The mean values show that Nagano lakes had the lowest concentrations of total phosphorus with exception of Lake Suwako and Tateshinako whereas Lake Kasumigaura and Inbanuma had the highest concentrations of Chlorophyll-a. Biwako which was divided into two sections i.e., North and South for the purpose of estimations also had low levels of total phosphorus and hence minimal Chlorophyll-a.

All models had high performance scores for data of high altitude lakes in Nagano prefecture (Fig. 4.11), the R^2 of 0.89 was the highest performance value shown by Model A, however, in this case Model A performed better compared to other results of the same model even above the general case. Figure 4.13 shows the estimates for the combined data of Japanese lakes, both models show higher R^2 values. Model B was superior to Model A with an R^2 of 0.94 and 0.82 respectively, for Model B, this was the highest performance level.

Table 4.8 Characteristics and mean values for the variables, 2000-2009, for the Japanese lakes used for simulation

Data sources:
Lake Imba: http://www.pref.chiba.lg.jp/suiho/kasentou/koukyouyousui/data/ichiran-koshou.html
Lake Kasumigaura: http://www.ktr.mlit.go.jp/kasumi/kasumi00145.html
Lakes in Nagano: https://www.pref.nagano.lg.jp/kankyo/mizutaiki/suishitsu/ent.htm
Lake Biwa: http://www.lberi.jp/asp/bkkc/Suishitsu/bkkcKeinenKListSearch.asp

Region	Prefecture	Lake	Altitude (m)	Surface area (km^2)	Max. depth (m)	TN (mg l^{-1})	TP (mg l^{-1})	TN/TP	Chl-a (μg l^{-1})
Kanto	Chiba	Inba	1	11.6	2.5	1.782	0.101	17.7	86.00
(Eastern)	Ibaraki	Kasumigaura	1	171.0	7.3	1.040	0.107	9.8	62.27
Chubu	Nagano	Aoki	822	1.9	58	0.242	0.006	49.4	1.68
(Central)		Kizaki	764	1.4	29.5	0.229	0.008	32.1	4.88
		Nojiri	654	4.6	38.5	0.103	0.005	21.2	2.29
		Suwa	759	13.3	7.6	0.820	0.044	18.8	41.35
		Tateshina	1250	0.1	5.0	0.166	0.036	4.7	17.11
Kansai	Shiga	Biwa - North	85	618.0	103.5	0.259	0.007	36.7	2.97
(Western)		Biwa - South	85	50.7	5.0	0.300	0.014	21.5	4.66

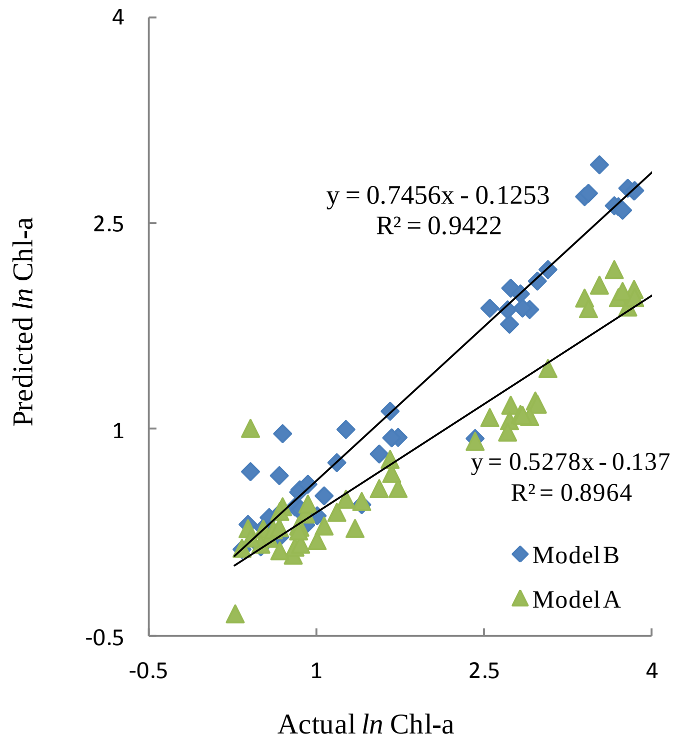


Fig. 4.11 Relationship between predicted and actual Chlorophyll-a for Nagano lakes, Japan

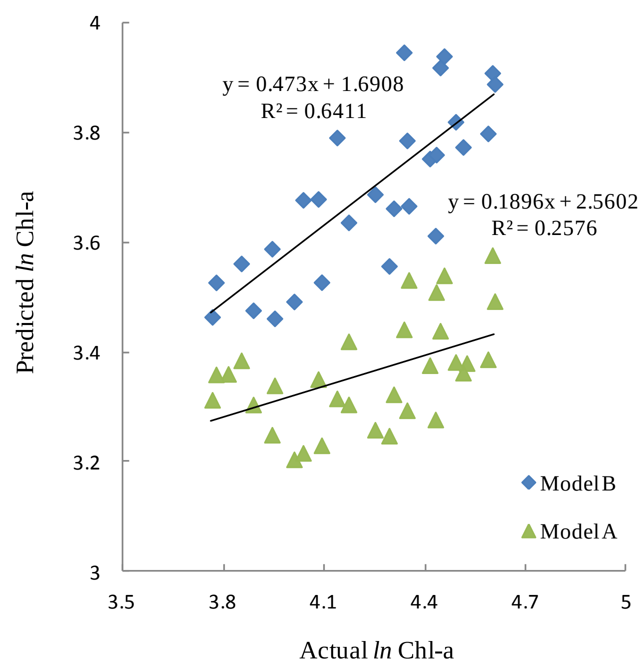


Fig. 4.12 Relationship between predicted and actual Chlorophyll-a for Lake Biwa, Japan

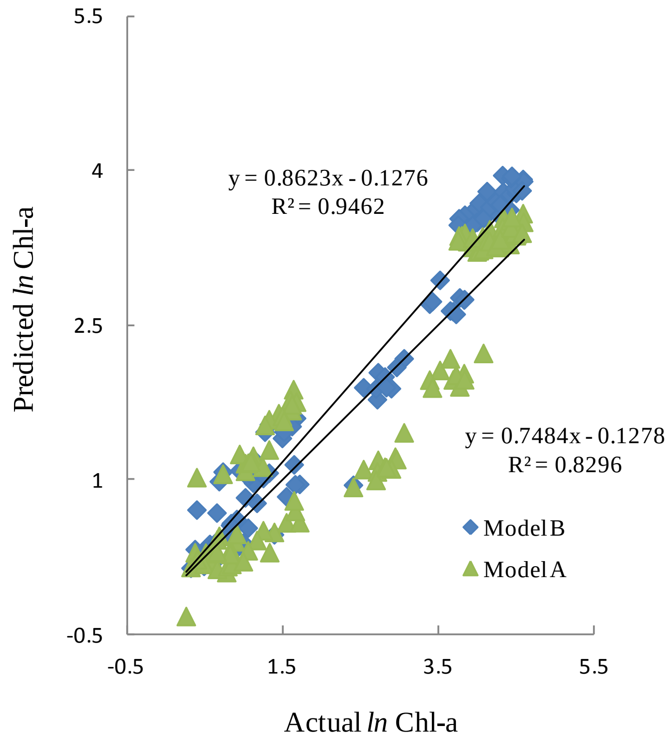


Fig. 4.13 Relationship between predicted and actual Chl-a for Japanese lakes (2000-2009)

4.3.5 Discussion

The results indicate that there are performance variations for the two models. The models give satisfactory results except for Lake Kasumigaura and Inbanuma where both models show low R^2 values, with Model A giving the least prediction. The variation in performance is partly attributed to higher concentrations of total phosphorus and total nitrogen and Chlorophyll-a in these lakes compared to majority of high altitude lakes in Nagano, the issue of high Chlorophyll-a concentrations in these two lakes has been reported by other authors (Oyama et al., 2009; Iwashita et al., 2004). Altitude i.e., temperature is also a major factor that determines Chlorophyll-a (Liu et al., 2010b), most high altitude lakes had low concentrations of

Chlorophyll-a. Differences in lake depth as shown by high Chlorophyll-a level in shallow lakes compared to deep lakes is reflected in the performance of the two models.

Inferring from the results of Figure 4.13, the two models were successful in predicting Chlorophyll-a. It is imperative that panel and auto-regressive effects be considered in modelling since Model B with three variables, phosphorus, altitude and first lag of chlorophyll-a outperformed Model A.

4.3.6 Conclusions

Results indicate that OLS and the dynamic panel model are both essential in predicting Chlorophyll-a, the dynamic panel model show a more predictive power compared to the OLS as shown by the high R^2 value. Since policies are formulated for application at wider geographical areas for management of multiple lakes within a country or across countries, we argue that the dynamic panel model is more applicable in such cases.

CHAPTER V

CONCLUSIONS AND FUTURE RESEARCHES

5.1 CONCLUSIONS

Earlier authors have done significant amounts of work to model the Chla = f(TP,TN) relationship using OLS. Their coefficients for total phosphorus and total nitrogen on average were commonly 0.7 and 0.4 respectively. The results of estimations for OLS in this thesis are consistent with those in previous studies. In simulations, OLS was successful in predicting Chlorophyll-a expect for one case of high concentration of Chlorophyll-a of which also the performance for the dynamic model in this specific scale was not as other simulations, therefore OLS remains as an important model in estimating Chlorophyll-a.

OLS does not consider station-specific and auto-regressive effects of Chlorophyll-a which are necessary considerations for the dynamic model, drawing conclusions from the performance of the dynamic model, it is concluded that the dynamic model was superior to OLS.

Results indicate that altitude a proxy for temperature had a high coefficient implying that temperature is a significant factor influencing Chlorophyll-a, it was among the factors that led to variation in Chlorophyll-a for the samples lakes in Japan used in the testing OLS and dynamic panel model.

As per the results of the dynamic model, it originates from this thesis that the parameter structure of Chlorophyll-nutrient relationships has changed from phosphorus and nitrogen as the significant variables to only phosphorus when station-specific effects are controlled and autoregressive effects of Chlorophyll-a are included in the estimation.

The adjusted coefficient of determination for dynamic model surpassed those of OLS, it is therefore recommended that the dynamic model should be applied in estimations of Chlorophyll-nutrient relationships.

For cases of sustainability, as phosphorus is a major predictor of Chlorophyll-a, adjusting production systems to hydroponics will reduce phosphorus use and hence a reduction of environmental enrichment.

Further investigation is necessary to estimate coefficients of the dynamic model using data from other regions and simulate Chlorophyll-a values other than the European dataset, this will provide insights for studies in limnology in other continents which have a high number of lakes like the United States.

Literature review has shown that coefficients of total phosphorus and total nitrogen vary in a few studies, further investigation is necessary to resolve this controversy and clarify on the underlying factors and conditions for search parameters.

The study has also revealed by proxy that temperature is a major factor that influences Chlorophyll-a based on the high coefficient of altitude. A combination of the article written by Abell et al. (2012) which focuses on Chlorophyll-a changes with variation in latitude and a related paper on global Chlorophyll-a changes in lakes along temperature and altitude gradients will be useful in understanding Chlorophyll-nutrient relationships globally.

5.2 RESEARCHES AHEAD

5.2.1 APPLICATION OF PANEL DATA MODELS TO OTHER LAKES

In this thesis, the new panel data models were tested using cross-country samples (Japan, UK, Australia and Sub-sample of Florida) for the random effect model in Chapter 4. The dynamic model was tested using sample data for Japanese lakes, the former as the superior model. Further application of these models to lakes from other continents will extend the application of these results.

5.2.2 SPATIAL DISTRIBUTION OF CHL-A IN LAKES USING REMOTE SENSING

Chlorophyll-a and secchi disk transparency (SDT) maps are usually developed from point data (Olmanson et al., 2008; Nas et al., 2010; Allan et al., 2011). A review of application of remote sensing in this regard was presented later in Chapter 3. It is based on chlorophyll maps that regions of a lake or group of lakes are evaluated. Spatial distribution of Chlorophyll-a for Lake Victoria is necessary in this regard.

REFERENCES

Abebe, G. A. 2013. Quantifying urban growth pattern in developing countries using remote sensing and spatial metrics: A case study in Kampala, Uganda. MSc thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente.

Abell, J. M., Ozkundakci, D., Hamilton, D. P., Jones, J. R. 2012. Latitudinal variation in nutrient stoichiometry and chlorophyll-nutrient relationships in lakes: A global study. Fundam. Appl. Limnol. 181: 1-14.

Acker, G. J., McMahon, E., Shen, S., Hearty, T., Casey, N. 2009. Time-series analysis of remotely-sensed seawifs chlorophyll in river-influenced coastal regions. EARSeL eProceedings 8: 114-139.

Adam, E. and O. Mutanga 2009. Spectral discrimination of papyrus vegetation (*Cyperus papyrus* L.) in swamp wetlands using field spectrometry. ISPRS. J. Photogramm. Remote Sens. 64: 612-620.

Adam, E. M., O. Mutanga, D. Rugege, and R. Ismail 2012. Discriminating the papyrus vegetation (*Cyperus papyrus* L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP. Int. J. Remote Sens. 33: 552-569.

Allan, M. G., Hamilton, D. P., Hicks, B. J., Brabyn, L. 2011. Landsat remote sensing of chlorophyll a concentrations in central North Island lakes of New Zealand. Int. J. Remote Sens. 32: 2037-2055.

Al-mulali, U. 2012. Factors affecting CO_2 emission in the Middle East: A panel data analysis. Energy 44: 564-569.

Arellano, M., Bond, S. 1991. “Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations,” Rev. Econ. Stud. 58: 277-297.

Arhonditsis, G., Brett, M. T., Frodge, J. 2003. Environmental control and limnological impacts of a large recurrent spring bloom in lake Washington, USA. Environ. Manage. 31: 603-618.

Arvola, L., Järvinen, M., Tulonen, T. 2011. Long-term trends and regional differences of phytoplankton in large Finnish lakes. Hydrobiologia 660: 125-134.

Bachmann, R. W., Bigham, D. L., Hoyer, M. V., Canfield, D. E Jr. 2012. Factors determining the distributions of total phosphorus, total nitrogen, and chlorophyll a in Florida lakes. Lake Reserv. Manage. 28: 10-26.

Bachmann, R. W., Hoyer, M. V. 2003. Predicting the frequencies of high chlorophyll levels in Florida lakes from average chlorophyll or nutrient data. Lake Reserv. Manage. 19: 229-241.

Ballot, A., Kotut, K., Novelo, E., Krienitz, L. 2009. Changes of phytoplankton communities in Lakes Naivasha and Oloidien, examples of degradation and salinization of lakes in the Kenyan Rift Valley. Hydrobiologia 632: 359-363.

Baranga, D., Chapman, C. A., Kasenene, J. M. 2009. The structure and status of forest fragments outside protected areas in central Uganda. Afric. J. Ecol. 47: 664-669.

BBC (British Broadcasting Corporation) 2012. Weather. [Online] http://www.bbc.co.uk/weather/232422 . (browsed on 21st September 2012).

Bechmann, M. E., Berge, D., Eggestad, H. O., Vandsemb, S. M. 2005. Phosphorus transfer from agricultural areas and its impact on the eutrophication of lakes - two long-term integrated studies from Norway. J. Hydrol. 304: 238-250.

Bhutiani, R., Khanna, D. R., Chandra, K. S. 2009. Light-limited population dynamics of phytoplankton: modeling light and depth effects. Environmentalist 29: 93-105.

Brink, A. B., Bodart C., Brodsky, L., Defourney, P., Ernst, C., Donney, F., Lupi, A., Tuckova, K. 2014. Anthropogenic pressure in East Africa-Monitoring 20 years of land cover changes by means of medium resolution satellite data. International Journal of Applied Earth Observation and Geoinformation. 28: 60-69.

Brown, C. D., Hoyer, M. V., Bachmann, R. W., Canfield, DE Jr. 2000. Nutrient-chlorophyll relationships: An evaluation of empirical nutrient-chlorophyll models using Florida and north-temperate lake data. Can. J. Fish. Aquat. Sci. 57: 1574-1583.

Camacho, A. 2006. On the occurrence and ecological features of deep chlorophyll maxima (DCM) in Spanish stratified lakes. Limnetica, 25: 453-478.

Carvalho, L., Solimini, A. G., Phillips, G., Pietiläinen, O. -P., Moe, J., Cardoso, A. C., Solheim, A. L., Ott, I., Søndergaard, M., Tartari, G., Rekolainen, S. 2009. Site-specific chlorophyll reference conditions for lakes in Northern and Western Europe. Hydrobiologia 633: 59-66.

Cavalli, R. M., Laneve, G., Fusilli, L., Pignatti, S., Santini, F. 2009. Remote sensing water observation for supporting Lake Victoria weed management. J. Environ. Manag. 90: 2199-2211.

Chen, J., Huang, J., Hu, J. 2011. Mapping rice planting areas in Southern China using the China environment satellite data. Mathematical and Computer Modelling 54: 1037-1043.

Choi, J. J., Jeon, B. N. 2007. Financial factors in foreign direct investments: A dynamic analysis of international data. Research in International Business and Finance 21: 1-18.

EEA (European Environment Agency). 2013. Water quality data for European lakes. http://www.eea.europa.eu/data-and-maps/data/waterbase-lakes-7

Elhadi, M. I. A., Mutanga, O., Rugege, D., Ismail, R. 2009. Field spectrometry of papyrus vegetation (*Cyperus papyrus* L.) in swamp wetlands of St Lucia, South Africa.

Preceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), 4 IV260-IV263.

Giardino, C., Pepe, M., Brivio, P. A., Ghezzi, P., Zilioli, E. 2001. Detecting chlorophyll, secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. Sci. Total Environ. 268: 19-29.

Greene, W. H. 2011. Econometric analysis. 7th ed. Prentice Hall.

Greisberger, S., Teubner, K. 2007. Does pigment composition reflect phytoplankton community structure in differing temperature and light conditions in a deep alpine lake? an approach using hplc and delayed fluorescence techniques. J. Phycol. 43: 1108-1119

Guildford, S. J., Hecky, R. E. 2000. Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: Is there a common relationship? Limnol. Oceanogr. 45: 1213-1223.

Gunkel, G., Casallas, J. 2002. Limnology of an equatorial high mountain lake - lago san pablo, ecuador: The significance of deep diurnal mixing for lake productivity. Limnologica 32: 33-43.

Haande, S., Rohrlack, T., Semyalo, R. P., Brettum, P., Edvardsen, B., Lyche-Solheim, A., Sørensen, K., Larsson, P. 2011. Phytoplankton dynamics and cyanobacterial dominance in murchison bay of Lake Victoria (Uganda) in relation to environmental conditions. Limnologica 41: 20-29.

Hamilton, D. P., O'Brien, K. R., Burford, M. A., Brookes, J. D., McBride, C. G. 2010. Vertical distributions of chlorophyll in deep, warm monomictic lakes. Aquat. Sci. 72: 295-307.

Hedger, R. D., Olsen, N. R. B., Malthus, T. J., Atkinson, P. M. 2002. Coupling remote sensing with computational fluid dynamics modelling to estimate lake chlorophyll-a concentration. Remote Sens. Environ. 79: 116-122.

Heinl, M., Walde, J., Tappeiner, G., Tappeiner, U. 2009. Classifiers vs. input variables-the drivers in image classification for land cover mapping. International Journal of Applied Earth Observation and Geoinformation 11: 423-430.

Hsiao, C. 2007. Panel data analysis – advantages and challenges. Test 16: 1-22.

Huszar, V. L. M., Caraco, N. F., Roland, F., Cole, J. 2006. Nutrient-chlorophyll relationships in tropical-subtropical lakes: Do temperate models fit? Biogeochemistry 79: 239-250.

Iwashita, K., Kudoh, K., Fujii, H., Nishikawa, H. 2004. Satellite analysis of water flow of Lake Inbanuma. Adv. Space Res. 33: 284-289.

Jeppesen, E., Søndergaard, M., Jensen, J.P., Havens, K.E., Anneville, O., Carvalho, L., Coveney, M.F., Deneke, R., Dokulil, M.T., Foy, B., Gerdeaux, D., Hampton, S.E., Hilt, S., Kangur, K., Köhler, J., Lammens, E.H.H.R., Lauridsen, T.L., Manca, M., Miracle, M.R., Moss, B., Nõges, P., Persson, G., Phillips, G., Portielje, R., Romo, S., Schelske, C.L., Straile, D., Tatrai, I, Willén, E., Winder, M. 2005. Lake responses to reduced nutrient loading - an analysis of contemporary long-term data from 35 case studies. Freshwater Biol. 50, 1747-1771.

Jin, L., Hongjuan, W. 2010. The effects of TN:TP ratios on the phytoplankton and colonial cyanobacteria in eutrophic shallow lakes. Available at http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5516982&tag=1 Accessed on 11th May, 2013.

Jone, M. B., Muthuri, F. M. 1985. The canopy structure and microclimate of papyrus (*Cyperus papyrus*) swamps. J. Ecol. 73: 481-491.

Kallio, K. 1994. Effect of summer weather on internal loading and chlorophyll a in a shallow lake: a modeling approach. 275/276: 371-378.

Kansiime, F., Kateyo, E., Oryem-Origa, H., Mucunguzi, P. 2007. Nutrient status and retention in pristine and disturbed wetlands in Uganda: Management implications. Wetlands Ecol. Manage. 15: 453-467.

Kasprzak, P., Padisák, J., Koschel, R., Krienitz, L., Gervais, F. 2008. Chlorophyll a concentration across a trophic gradient of lakes: An estimator of phytoplankton biomass? Limnologica 38: 327-338.

Kateregga, E., Sterner, T. 2007. Indicators if an invasive species: Water hyacinths in Lake Victoria. Ecological Indicators 7: 362-370.

Keuchel, J., Naumann, S., Heiler, M., Siegmund, A. 2003. Automatic land cover analysis for tenerife by supervised classification using remotely sensed data. Remote Sens. Environ. 86: 530-541.

KFC (Kenya Flower Council) 2013. The Flower Industry in Kenya. Available at www.kenyaflowercouncil.org

Law, S. H., Habibullah, M. S. 2009. The determinants of financial development: Institutions, openness and financial liberalisation. S. Afr. J. Econ. 77: 45-58.

Lind, E. M. 1956. Studies in Uganda swamps. Uganda J. 20: 166-76.

Lind, E. M., Visser, S. A. 1962. A study of a swamp at the north end of Lake Victoria. J. Ecol. 50: 599-613.

Liu, W., Zhang, Q., Liu, G. 2010a. Lake eutrophication associated with geographic location, lake morphology and climate in China. Hydrobiologia 644: 289-299

Liu, Y., Guo, H., Yang, P. 2010b. Exploring the influence of lake water chemistry on chlorophyll a: A multivariate statistical model analysis. Ecol. Model. 221: 681-688.

Lu, D., Weng, Q. 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28: 823-870.

Lundy, M. E., Spencer, D. F., Kessel, C. V., Hill, J. E., Linquist, B. A. 2012. Managing phosphorus fertilizer to reduce algae, maintain water quality, and sustain yields in water-seeded rice. Field Crops Research 131: 81-87.

Lv, J., Wu, H., Chen, M. 2011. Effects of nitrogen and phosphorus on phytoplankton composition and biomass in 15 subtropical, urban shallow lakes in Wuhan, China. Limnologica 41: 48-56.

Maclean, I. M. D., Boar, R. R., Lugo, C. 2011. A review of the relative merits of conserving, using, or draining papyrus swamps. Environ. Manag. 47: 218-229.

Manevski, K., Manakos, I., Petropoulos, G. P., Kalaitzidis, C. 2011. Discrimination of common mediterranean plant species using field spectroradiometry. International Journal of Applied Earth Observation and Geoinformation 13: 922-933.

Marques, A. C., Fuinhas, J. S. 2011. Drivers promoting renewable energy: A dynamic panel approach. Renewable and Sustainable Energy Reviews 15: 1601-1608.

Mateev, M., Poutziouris, P., Ivanov, K. 2013. On the determinants of SME capital structure in Central and Eastern Europe: A dynamic panel analysis. Research in International Business and Finance 27: 28-51.

Maxwell, S. K., Sylvester, K. M. 2012. Identification of "ever-cropped" land (1984-2010) using landsat annual maximum NDVI image composites: Southwestern Kansas case study. Remote Sens. Environ 121: 186-195.

McCullough, I. M., Loftin, C. S., Sader, S. A. 2012. Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity Remote Sens. Environ. 123: 109-115.

Mekonnen, M. M., Hoekstra, A, Y., Becht, R. 2012. Mitigating the water footprint of export cut flowers from the Lake Naivasha basin, Kenya. Water Resour. Manage. 26: 3725-3742.

Morales-Baquero, R., Pulido-Villena, E., Reche, I. 2006. Atmospheric inputs of phosphorus and nitrogen to the southwest mediterranean region: Biogeochemical responses of high mountain lakes Limnol Oceanogr. 51: 830-837.

Morrison, E. H. J., Harper, D. M. 2009. Ecohydrological principles to underpin the restoration of *Cyperus papyrus* at Lake Naivasha, Kenya. Ecohydrol. Hydrobiol. 9: 83-97.

Morrison, E. H. J., Upton, C., Pacini, N., Odhiambo-K'oyooh, K., Harper, D. M. 2013. Public perceptions of papyrus: community appraisal of wetland ecosystem services at Lake Naivasha, Kenya. Ecohydrol. Hydrobiol. 13: 135-147.

Mutia, T. M., Virani, M. Z., Moturi, W. N., Muyela, B., Mavura, W. J., Lalah, J. O. 2012. Copper, lead and cadmium concentrations in surface water, sediment and fish, *C. Carpio*, samples from Lake Naivasha: effect of recent anthropogenic activities. Environ. Earth Sci. 67: 1121-1130.

Muyodi, F. J., Bugenyi, F. W. B., Hecky, R. E. 2010. Experiences and lessons learned from interventions in the Lake Victoria basin: The Lake Victoria environmental management project. Lakes and Reservoirs: Research and Management 15: 77-88.

Nangia, V., Gowda, P. H., Mulla, D. J. 2010. Effects of changes in N-fertilizer management on water quality trend at the watershed scale. Agr. Water Manage. 97: 1855-1860.

Nas, B., Ekercin, S., Karabörk, H., Berktay, A., Mulla, D. J. 2010. An application of Landsat-5TM image data for water quality mapping in Lake Beysehir, Turkey. Water Air Soil Pollut. 212: 183-197.

NFA (National Forestry Authority) 2005. Shapefile: land use and land cover map of 1996 and 2005. National Forest Authority, Kampala, Uganda.

Nõges, P., Poikane, S., Kõiv, T., Nõges, T. 2010. Effect of chlorophyll sampling design on water quality assessment in thermally stratified lakes. Hydrobiologia 649: 157-170.

Ogbebo, F. E., Evans, M. S., Brua, R. B., Keating, J. J. 2009. Limnological features and models of chlorophyll-a in 30 lakes located in the lower Mackenzie river basin, northwest territories (Canada). J. Limnol. 68: 336-351.

Okullo, W., Ssenyonga, T., Hamre, B., Frette, O., Sørensen, K., Stamnes, J.J., Steigen, A., Stamnes, K. 2007. Parameterization of the inherent optical properties of Murchison Bay,Lake Victoria. Appl Opt. 46:8553-61.

Olmanson, G. L., Kloiber, S. M., Bauer, M. E., Brezonik, P. L. 2001. Image Processing Protocol for Regional Assessment of Water Quality. University of Minnesota Water Resources Center, Public Report Series No. 14.

Olmanson, L. G., Bauer, M. E., Brezonik, P. L. 2008. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. Remote Sens. Environ. 112: 4086-4097.

Opande, G. O., Onyango, J. C., Wagai, S. O. 2004. Lake Victoria: The water hyacinth (*Eichhornia crassipes* [MART.] SOLMS), its socio-economic effects, control measures and resurgence in the Winam gulf. Limnologica 34: 105-109.

Oscar, T. N.d. Panel data analysis, fixed and random effects (using stata 10.x) (ver. 4.1). Princeton University. Available at http://www.princeton.edu/~otorres/Panel101.pdf Accessed on 11th May, 2013.

Östlund, C., Flink, P., Strömbeck, N., Pierson, D., Lindell, T. 2001. Mapping of water quality of Lake Erken, Sweden, from Imaging spectrometry and Landsat Thematic mapper. The Science of the Total Environment 268: 139-254.

Otukei, J. R., Blaschke, T. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. J. Appl. Earth Observ. 12 (SUPPL. 1): S27-S31.

Owino, A. Ryan, P. 2007. Recent papyrus swamp habitat loss and conservation implications in western Kenya. Wetlands Ecol. Manag. 15: 1-12.

Oyama, Y., Matsushita, B., Fukushima, T., Matsushige, K., Imai, A. 2009. Application of spectral decomposition algorithm for mapping water quality in a turbid lake (Lake Kasumigaura, Japan) from Landsat TM data. ISPRS J. Photogramm. Remote Sens. 64: 73-85.

Pal, M., Mather, P. M. 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sens. Environ. 86: 554-565.

Peng, J., Liu, Y., Shen, H., Han, Y., Pan, Y. 2012. Vegetation coverage change and associated driving forces in mountain areas of Northwestern Yunnan, China using RS and GIS. Environ. Monit. Assess. 184: 4787-4798.

Pepe M, Giardino C, Borsani G, Cardoso AC, Chiaudani G, Premazzi G, Rodari E, Zilioli E. 2001. Relationship between apparent optical properties and photosynthetic pigments in the sub-alpine lake iseo. Sci. Total Environ. 268 31-45

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. -., Tucker, C. J., Stenseth, N. C. 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology and Evolution 20, 503-510.

Pfister, S., Bayer, P., Koehler, A., Hellweg, S. 2011. Projected water consumption in future global agriculture: Scenarios and related impact. Science of the Total Environment 409, 4206-4216.

Phillips, G., Pietiläinen, O.-P, Carvalho, L., Solimini, A., Solheim, A. L., Cardoso, A. C. 2008. Chlorophyll-nutrient relationships of different lake types using a large European dataset. Aquat. Ecol. 42, 213-226.

Räike, A., Pietiläinen, O.-P., Rekolainen, S., Kauppila, P., Pitkänen, H., Niemi, J., Raateland, A., Vuorenmaa, J. 2003. Trends of phosphorus, nitrogen and chlorophyll a concentrations in Finnish rivers and lakes in 1975-2000. Sci. Total Environ. 310, 47-59.

RapidEye 2012. http://www.RapidEye.com/products/ortho.htm [Online] (browsed on 1st December, 2012.

Reckhow, K. H. 1993. A random coefficient model for chlorophyll-nutrient relationships in lakes. Ecol. Model. 70, 35-50.

Rikken, M. 2011. The global competitiveness of the Kenyan Flower Industry. Technical Paper Prepared for the Fifth Video Conference on the Global Competitiveness of the Flower Industry in Eastern Africa. December 2011.

Roodman, D. 2009. “How to do xtabond2: An introduction to “difference” and “system” GMM in Stata. Stata Journal 9, 86-136.

Saadat, H., Adamowski, J., Bonnell, R., Sharifi, F., Namdar, M., Ale-Ebrahim, S. 2011. Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 608-619.

Sakamoto, M. 1966. Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. Arch. Hydrobiol. 62,1-28.

Santos, G. F. 2013. Fuel demand in Brazil in a dynamic panel data approach. Energy Economics 36, 229-240.

Sass, G. Z., Creed, I. F., Bayley, S. E., Devito, K. J. 2007. Understanding variation in trophic status of lakes on the Boreal Plain: A 20 year retrospective using Landsat TM imagery. Remote Sens. Environ. 109, 127-141.

Sawaya, K. E., Olmanson, L.F., Heinert, N. J., Brezonik, P.L., Bauer, M. E. 2003. Extending satellite remote sensing to local scales : land and water resource monitoring using high-resolaution imagery. Remote Sens. Environ. 88, 144-156.

Seip, K. L, Jeppesen E, Jensen, J. P., Faafeng B. 2000. Is trophic state or regional location the strongest determinant for chl-a/TP relationships in lakes? Aquat. Sci. 62 195-204.

Sitoki, L., Gichuki, J., Ezekiel, C., Wanda, F., Mkumbo, O, C., Marshall, B, E. 2010. The Environment of Lake Victoria (East Africa): Current status and historical changes. Internat. Rev. Hydrobiol. 95: 209-223.

SMD (Survey and Mapping Department) 1998. The topographic map of 1998. Survey and Mapping Department. Entebbe, Uganda.

Søndergaard, M., Larsen, S. E., Jørgensen, T. B., Jeppesen, E. 2011. Using chlorophyll a and cyanobacteria in the ecological classification of lakes. Ecol. Indic. 11, 1403-1412.

Sebilo, M., Mayer, B., Nicolardot, B., Pinay, G., Mariotti, A. 2013. Long-term fate of nitrate fertilizer in agricultural soils. Proc. Natl. Acad. Sci. USA. 110: 18185-18189.

Ssebiyonga, N., Erga, S. R., Hamre, B., Stamnes, J. J., Frette, O. 2013. Light conditions and photosynthetic efficiency of phytoplankton in Murchison Bay, Lake Victoria, Uganda, Limnologica - Ecology and Management of Inland Waters 43: 185-193.

Stoof-Leichsenring, K. R., Junginger, A., Olaka, L. A., Tiedemann, R., Trauth, M. H. 2011. Environmental variability in Lake Naivasha, Kenya, over the last two centuries. J. Paleolimnol. 45: 353-367.

Takagi, S., Shi, Z. 2011. Exchange rate movements and foreign direct investment (FDI): Japanese investment in Asia, 1987–2008. Japan and the World Economy 23, 265-272.

Terrell, J. B, Watson D L, Hoyer M V, Allen M S and Canfield D E Jr. 2000. Temporal water chemistry trends (1967-1997) for a sample (127) of Florida waterbodies, Lake and Reservoir Management, 16 177-194

Trevisan, G. V., Forsberg, B. R. 2007. Relationships among nitrogen and total phosphorus, algal biomass and zooplankton density in the central Amazonia lakes. Hydrobiologia 586, 357-365.

Trolle, D., Zhu G, Hamilton D, Luo L, McBride C, Zhang L. 2009. The influence of water quality and sediment geochemistry on the horizontal and vertical distribution of phosphorus and nitrogen in sediments of a large, shallow lake Hydrobiologia, 627 31-44

Turner, G. F., Seehausen, O., Knight, M. E., Allender, C. J., Robinson, R. L. 2001. How many species of cichlid fishes are there in African lakes? Mol. Ecol. 10, 793-806.

UBOS (Uganda Bureau of Statistics) 2013. Projected mid-year population for Kampala and Wakiso. Kampala, Uganda

UEPB (Uganda Export Promotion Board) 2013. Volume and Value of exports for fish and fish products.

UFEA (Uganda Flower Exporter`s Association) 2010. Meeting at UFEA on floriculture in Uganda. 5th November, 2010. Lugogo, Kampala.

Ullah, S., Schlerf, M., Skidmore, A. K., Hecker, C. 2012. Identifying plant species using mid-wave infrared (2.5-6μm) and thermal infrared (8-14μm) emissivity spectra. Remote Sens. Environ 118, 95-102.

UNESCO (United Nations Educational, Scientific and Cultural Organisation) 2013. Water resources: Statistics. Available at http://www.unwater.org/statistics_res.html (Accessed on 12th June, 2013).

USGS (United States Geological Survey) 2012. Online database of satellite data. http://landsat.usgs.gov/band_designations_landsat_satellites.php (Accessed on 15th July, 2012).

Vermeiren, K., A. Van Rompaey, M. Loopmans, E. Serwajja, and P. Mukwaya. 2012. Urban growth of Kampala, Uganda: Pattern analysis and scenario development. Landscape Urban Plan. 106: 199-206.

Wang, X., Lu, Y., He, G., Han, J., Wang, T. 2007. Multivariate analysis of interactions between phytoplankton biomass and environmental variables in Taihu Lake, China. Environ. Monit. Assess. 133, 243-253.

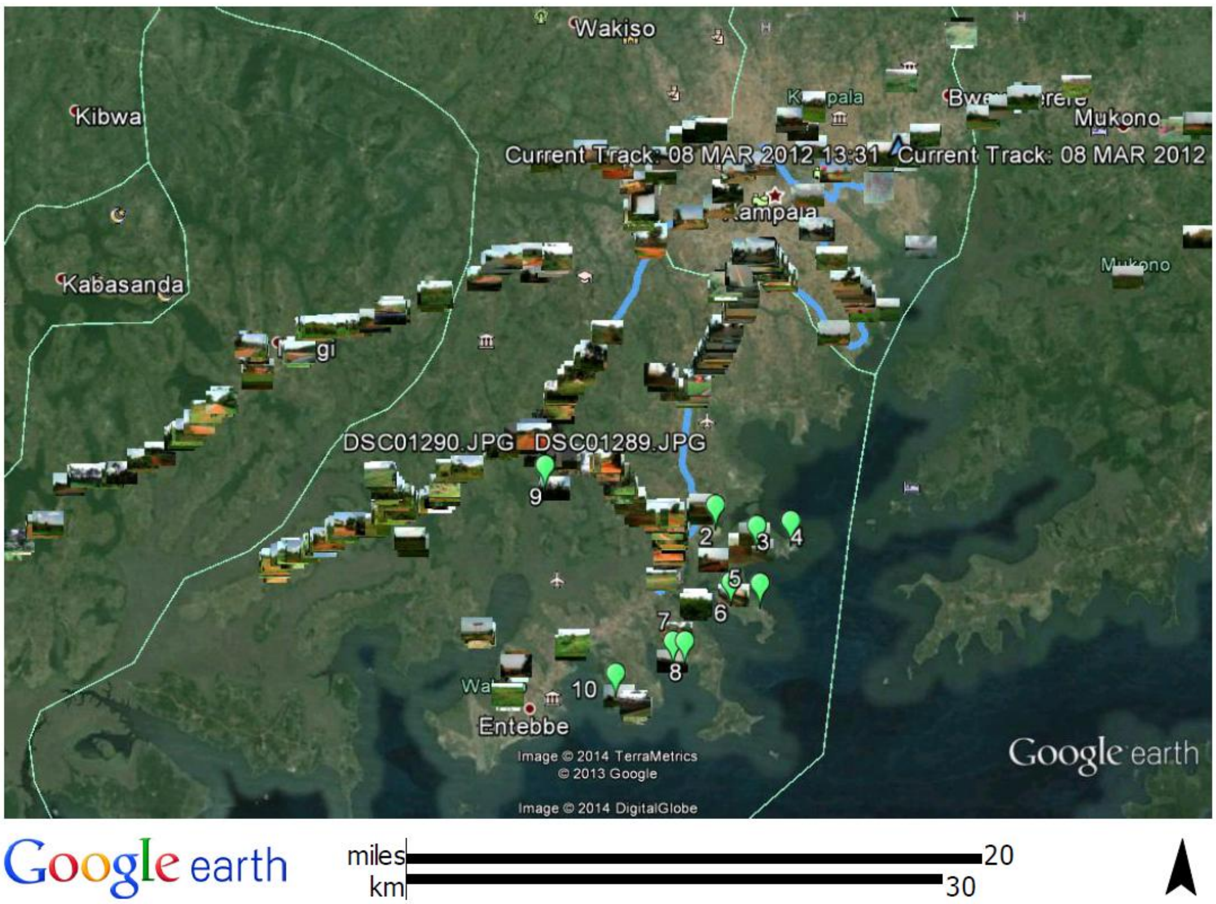
Waske, B., Braun, M. 2009. Classifier ensembles for land cover mapping using multitemporal SAR imagery. ISPRS Journal of Photogrammetry and Remote Sensing 64, 450-457.

Xiong, Y., Chen, X., Qiu, G, Y., Tan, S. Mapping of chlorophyll-a concentration of reservoir Dajingshan, Zhuhai, China, from LandSat thematic mapper. Geoscience and Remote Sensing Symposium (IGARSS). 2011 IEEE International, p. 3366 – 3369.

Zhang, Y.C, Lin, S., Qian, X., Wang, Q.G., Qian, Y., Liu, J.P., Ge, Y. 2011. Temporal and spatial variability of chlorophyll a concentration in Lake Taihu using MODIS time-series data. Hydrobiologia 661, 235-250.

Zinabu, G-M. 2002. The effects of wet and dry seasons on concentrations of solutes and phytoplankton biomass in seven Ethiopian rift-valley lakes. Limnologica - Ecology and Management of Inland Waters 32: 169-179.

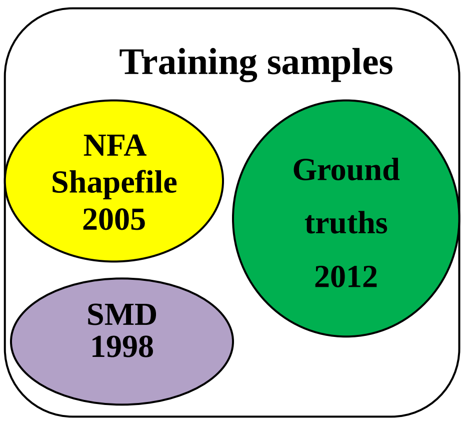
APPENDICES



A1. Ground truth (photographs and GPS route) for Entebbe and Kampala: Feb – Mar 2012

Groundtruth:
Sample Photographs

A



B

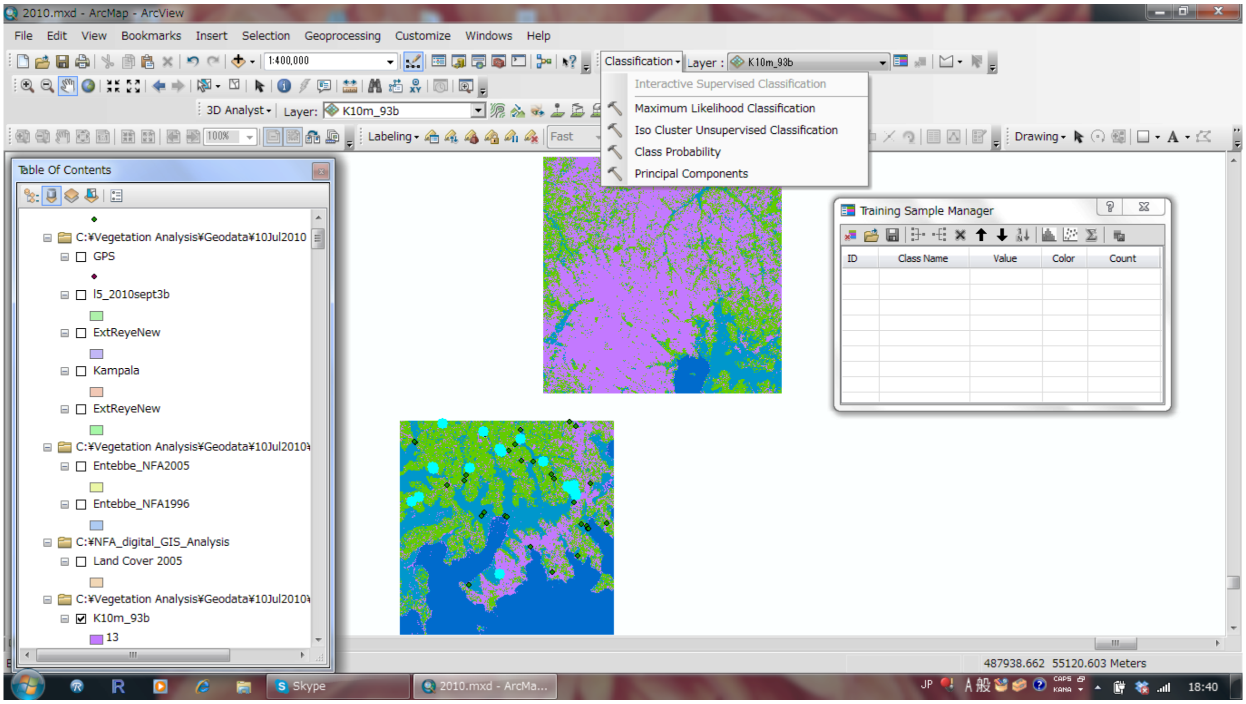
A2. Plant species images and co-ordinates taken during February and March 2012 for classification of RapidEye satellite image.
[**A**] *Cyperus papyrus*
0^0 06' 26"N 32^0 32' 50"E,
[**B**] Multi-species wetland
(*Cyperus papyrus*, Palm and other plant species)
0^0 08' 26"N 32^0 32' 04"E.

A3. Eucalyptus, a common tree species mostly planted in wetlands.
0^0 10' 58.89"N 32^0 26' 46.48"E

A4. Grassland expanse in Wakiso, Uganda. The area was classified among Non-wetland vegetation. A significant area in the study site was occupied by savannah vegetation.

0^0 7' 17.256"N 32^0 20' 5.842"E

A5. Residential area in Kampala, Uganda
0^0 15' 52.5248"N 32^0 37' 29.4235"E



A6. Satellite image analysis in ArcMap 10

A7. European lakes and stations consisting of 198 lakes and 199 stations, for developing the **random effect model,** two stations 05-369 and 05-371 originate from Lago Di Garda. Stations names are shown in the parenthesis where (-do-) after the lake name implies similar name for the lake and station while (number) are station names included in **Table 4.1** in their respective order as shown in this appendix, Lake NL1 and NL2 were missing geographically.[a]

Country	Lake and station	Country	Lake and station
CH	Bielersee (-do-)	HU	Balatonfenyves (-do-)
CH	Brienzersee (-do-)	HU	Keszthely (-do-)
CH	Burgaschisee (-do-)	HU	Retkozi To (-do-)
CH	Greifensee (-do-)	HU	Szigliget (-do-)
CH	Pfaffikersee (-do-)	HU	Velencei-To (Agard Molo) - (4)
CH	Thunersee (-do-)	HU	Velencei-To (Nemet Tisztas) - (5)
CH	Turlersee (-do-)	IT	Alleghe (6)
DE	Gro-er Muggelsee (-do-)	IT	Avigliana Piccolo (-do-)
DE	Kummerower See (-do-)	IT	Barbellino Ii O Pian Barbellino (-do-)
DE	Muritz (Au-enmuritz) (-do-)	IT	Cancano (-do-)
DE	Muritz (Binnenmuritz) (-do-)	IT	Candia (-do-)
DE	Sacrower See (-do-)	IT	Fusine (-do-)
DE	Schweriner See (Au-ensee) (-do-)	IT	Lago Di Garda (7) and (8)
DE	Schweriner See (Innensee) (-do-)	IT	Maggiore-Ghiffa (-do-)
DE	Stechlinsee (-do-)	IT	Maggiore-Lesa (-do-)
DE	Steinhuder Meer (-do-)	IT	Mergozzo (-do-)
DK	Arreskov So Amtst.4 (-do-)	IT	Mezzo (-do-)
DK	Arresoen (-do-)	IT	Misurina (9)
DK	Bryrup Langso, Bry 1 (-do-)	IT	Monte Spluga (-do-)
DK	Engelsholm So St Amt (-do-)	IT	Santa Croce (10)
DK	Furesoen Dybeste Sted Amt St (-do-)	IT	Scais (-do-)
DK	Gundsomagle So St Amt (-do-)	IT	Scanno (11)
DK	Hinge So (-do-)	IT	Truzzo (-do-)
DK	Holm So (-do-)	IT	Val Di Lei (-do-)
DK	Hornum So (-do-)	LT	Alnis (-do-)
DK	Kvie So (-do-)	LT	Lukstas (12)
DK	Magleso Amt St 1 (-do-)	LT	Rubikiai (13)
DK	Nors So St 1 (-do-)	LT	Sventas (-do-)
DK	Ravn So, Midt -- Rav1 (-do-)	LV	Akacis (-do-)
DK	Soby So, Midtjylland (-do-)	LV	Burtnieku (14)
DK	Sogard So St Amt (-do-)	LV	Kanieris (15)
DK	Soholm So St Amt (-do-)	LV	Kishezers (16)
DK	Store Sogard So Amt St 1 (-do-)	LV	Liepajas (17)
DK	Tisso Amt St 1 (-do-)	LV	Raznas (18)
DK	Utterselv Mose, Ostlige Bassin (-do-)	LV	Usmas (-do-)
DK	Vesterborg So (-do-)	NL	Ijsselmeer (19)
HU	Agard (1)	NL	Ketelmeer + Vossemeer (20)
HU	Balaton (Siofoki Medence) (2)	NL	Markermeer (21)
HU	Balaton (Szemesi Medence) (3)	NL	NL1 (22)

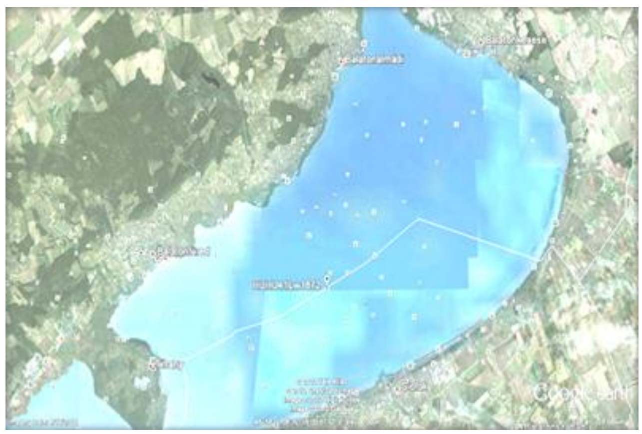
A7 Continued.

Country	Lake and station	Country	Lake and station	Country	Lake and station
NL	NL2 (23)	SE	Grissjon (-do-)	SE	Pahajarvi (-do-)
NL	Randmeren-Oost (24)	SE	Gryten (-do-)	SE	Rammsjon (-do-)
NL	Randmeren-Zuid (25)	SE	Hagasjon (-do-)	SE	Remmarsjon (-do-)
PL	Biale Wlodawskie (26)	SE	Hallsjon (-do-)	SE	Rotehogstjarn (-do-)
PL	Sremskie (27)	SE	Hallvattnet (-do-)	SE	Rundbosjon (-do-)
PL	Tarnowskie Duze (28)	SE	Harasjon (-do-)	SE	S. Bergsjon (-do-)
PL	Wuksniki (29)	SE	Harsvatten (-do-)	SE	Sangen (-do-)
SE	Abiskojaure (-do-)	SE	Hinnasjon (-do-)	SE	Sannen (-do-)
SE	Algarydssjon (-do-)	SE	Hjartsjon (-do-)	SE	Siggeforasjon (-do-)
SE	Algsjon (-do-)	SE	Hokesjon (-do-)	SE	Skaravattnet (-do-)
SE	Allgjuttern (-do-)	SE	Holmeshultasjon (-do-)	SE	Skargolen (-do-)
SE	Alsjon (-do-)	SE	Horsan (-do-)	SE	Skarsjon (1) - (30)
SE	Baen (-do-)	SE	Humsjon (-do-)	SE	Skarsjon (2) - (31)
SE	Baste Trask (-do-)	SE	Jutsajaure (-do-)	SE	Spjutsjon (-do-)
SE	Bergtrasket (-do-)	SE	Krankesjon (-do-)	SE	St. Envattern (-do-)
SE	Betarsjon (-do-)	SE	Langsjon (-do-)	SE	St. Lummersjon (-do-)
SE	Bjorken (-do-)	SE	Larkesholmssjon (-do-)	SE	Stensjon (-do-)
SE	Branntrasket (-do-)	SE	Latnjajaure (-do-)	SE	Stora Skarsjon (-do-)
SE	Brunnsjon (-do-)	SE	Lilla Oresjon (-do-)	SE	Storasjo (-do-)
SE	Bysjon (-do-)	SE	Lillesjo (-do-)	SE	Stor-Backsjon (-do-)
SE	Dagarn (-do-)	SE	Lillsjon (-do-)	SE	Storsjon (-do-)
SE	Dagtorpssjon (-do-)	SE	Limmingsjon (-do-)	SE	Svanshalssjon (-do-)
SE	Degervattnet (-do-)	SE	Louvvajaure (-do-)	SE	Svartesjon (-do-)
SE	Djupa Holmsjon (-do-)	SE	Malaren. Bjorkfjarden (-do-)	SE	Svartsjon (-do-)
SE	Edasjon (-do-)	SE	Malaren. Ekoln (-do-)	SE	Svinarydsjon (-do-)
SE	Ekholmssjon (-do-)	SE	Malaren. Galten (-do-)	SE	Tangerdasjon (-do-)
SE	Ellestadssjon (-do-)	SE	Masen (-do-)	SE	Tangersjo (-do-)
SE	Fagertarn (-do-)	SE	Mossjon (-do-)	SE	Tarnan (-do-)
SE	Faglasjon (-do-)	SE	N. Yngern (-do-)	SE	Tomeshultagolen (-do-)
SE	Fiolen (-do-)	SE	Navarn (-do-)	SE	Torrgardsvattnet (-do-)
SE	Fjarasjo (-do-)	SE	Njalakjaure (-do-)	SE	Tvaringen (-do-)
SE	Forsjon (-do-)	SE	Norrsjon (-do-)	SE	Ulvsjon (-do-)
SE	Fracksjon (-do-)	SE	Ojsjon (-do-)	SE	V. Rannobodsjon (-do-)
SE	Fyrsjon (-do-)	SE	Oljaren (-do-)	SE	Valasjon (-do-)
SE	Fysingen (-do-)	SE	Orsjon (-do-)	SE	Valkeajarvi (-do-)
SE	Gipsjon (-do-)	SE	Orvattnet (-do-)	SE	Vanern, Tarnan (-do-)
SE	Glimmingen (-do-)	SE	Oversjon (-do-)	SE	Vastra Solsjon (-do-)
SE	Gosjon (-do-)	SE	Overudsjon (-do-)	SE	Vikasjon (-do-)
SE	Gransjon (-do-)	SE	Ovre Fjatsjon (-do-)	SE	Vuolgamjaure (-do-)
SE	Granvattnet (-do-)	SE	Ovre Skarsjon (-do-)	SE	Ymsen (-do-)

[a] Data source; European Environment Agency (EEA).
Available online at http://www.eea.europa.eu/data-and-maps/data/waterbase-lakes-8

A

B



A8. Typical examples of the 199 Google Earth images used to code Land-Use Dummy.
[**A**] Co-ordinates: 10.053771E 46.062865N
Country: Italy
National Station ID: 12
Elevation:1868m
Land-Use Dummy: 0
[**B**] Co-ordinates: 18.00545E 46.94603N
Country: Hungary
National Station ID: HU41Lw1872
Elevation: 105m
Land-Use Dummy: 1

A9. Country Codes and Country Names for the 396 lakes and 454 stations used to develop the **dynamic panel model**

Country Code	Country
BE	Belgium
CH	Switzerland
DE	Germany
DK	Denmark
FR	France
GB	Great Britain
HR	Croatia
HU	Hungary
IT	Italy
LT	Lithuania
LV	Latvia
NL	Netherlands
PL	Poland
PT	Portugal
RO	Romania
RS	Serbia
SE	Sweden
SI	Slovenia