DEPARTMENT OF GEOSCIENCES AND GEOGRAPHY A5

ENHANCED PROCESSING OF SPOT MULTISPECTRAL SATELLITE IMAGERY FOR ENVIRONMENTAL MONITORING AND MODELLING

BARNABY JOHN FORRESTER CLARK

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE

Enhanced processing of SPOT multispectral satellite imagery for environmental monitoring and modelling

Barnaby John Forrester Clark

Department of Geosciences and Geography
Faculty of Science
University of Helsinki
Finland

Academic dissertation

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium XII, Main Building, Fabianinkatu 33, on 23rd June 2010, at 12 o'clock.

Helsinki 2010

Supervisor: Dr. Petri Pellikka
Professor of Geoinformatics
Department of Geosciences and Geography
University of Helsinki
Finland

Pre-examiners: Dr. Michael Schaepman
Professor of Remote Sensing
Department of Geography
University of Zurich
Switzerland

Dr. Markus Holopainen
Docent
Department of Forest Resource Management
University of Helsinki
Finland

Opponent: Dr. Philip Lewis
Professor in Remote Sensing
Department of Geography
University College London
University of London
UK

Publisher:
Department of Geosciences and Geography
Faculty of Science
PO Box 64, FI-00014 University of Helsinki
Finland

ISSN-L 1798-7911
ISSN 1798-7911 (print)
ISBN 978-952-10-6305-3 (paperback)
ISBN 978-952-10-6306-0 (PDF)
http://ethesis.helsinki.fi

Helsinki University Print
Helsinki 2010

To my family

Three passions, simple but overwhelmingly strong, have governed my life: the longing for love, the search for knowledge, and the unbearable pity for the suffering of mankind.

The Autobiography of Bertrand Russell, Prologue

ABSTRACT

The Taita Hills in southeastern Kenya form the northernmost part of Africa's Eastern Arc Mountains, which have been identified by Conservation International as one of the top ten biodiversity hotspots on Earth. As with many areas of the developing world, over recent decades the Taita Hills have experienced significant population growth leading to associated major changes in land use and land cover (LULC), as well as escalating land degradation, particularly soil erosion. Multi-temporal medium resolution multispectral optical satellite data, such as imagery from the SPOT HRV, HRVIR, and HRG sensors, provides a valuable source of information for environmental monitoring and modelling at a landscape level at local and regional scales. However, utilization of multi-temporal SPOT data in quantitative remote sensing studies requires the removal of atmospheric effects and the derivation of surface reflectance factor (ρ_s). Furthermore, for areas of rugged terrain, such as the Taita Hills, topographic correction is necessary to derive comparable ρ_s throughout a SPOT scene. Reliable monitoring of LULC change over time and modelling of land degradation and human population distribution and abundance are of crucial importance to sustainable development, natural resource management, biodiversity conservation, and understanding and mitigating climate change and its impacts.

The main purpose of this thesis was to develop and validate enhanced processing of SPOT satellite imagery for use in environmental monitoring and modelling at a landscape level, in regions of the developing world with limited ancillary data availability. The Taita Hills formed the application study site, whilst the Helsinki metropolitan region was used as a control site for validation and assessment of the applied atmospheric correction techniques, where multiangular ρ_s field measurements were taken and where horizontal visibility meteorological data concurrent with image acquisition were available. The proposed historical empirical line method (HELM) for absolute atmospheric correction was found to be the only applied technique that could derive ρ_s within an RMSE of < 0.02 ρ_s in the SPOT visible and near-infrared bands; an accuracy level identified as a benchmark for successful atmospheric correction. A multi-scale segmentation/object relationship modelling (MSS/ORM) approach was applied to map LULC in the Taita Hills from the multi-temporal SPOT imagery. This object-based procedure was shown to derive significant improvements over a uni-scale maximum-likelihood technique. The derived LULC data was used in combination with low cost GIS geospatial layers describing elevation, rainfall and soil type, to model degradation in the Taita Hills in the form of potential soil loss, utilizing the simple universal soil loss equation (USLE). Furthermore, human population distribution and abundance were modelled with satisfactory results using only SPOT and GIS derived data and non-Gaussian predictive modelling techniques. The SPOT derived LULC data was found to be unnecessary as a predictor because the first and second order image texture measurements had greater power to explain variation in dwelling unit occurrence and abundance. The ability of the procedures to be implemented locally in the developing world using low-cost or freely available data and software was considered. The techniques discussed in this thesis are considered equally applicable to other medium- and high-resolution optical satellite imagery, as well the utilized SPOT data.

Keywords: atmospheric correction, HELM, LULC change, MSS/ORM, population distribution and abundance modelling, SPOT, USLE.

ACKNOWLEDGMENTS

First of all I want to sincerely thank Kaisa Tuovinen, without whom none of this work would ever have happened. It was a wonderful opportunity to have a reason, and be able to come to Finland to carry out this PhD. It has been a truly life changing experience for me, for which I will be eternally grateful to her for. Kaisa, I wish you all the best. Furthermore, I would especially like to thank my supervisor Professor Petri Pellikka for accepting me into, and providing me with an opportunity to study and work in, the Geoinfomatics Research Group at the Department of Geography. Most especially, I am thankful to Petri for "keeping the faith" in me whilst this thesis was painstakingly finished. I am also very grateful for the independence and friendly working environment he has provided. Petri is also a very good man to undertake fieldwork in Africa with, as he is calm and collected and can successfully deal with the problems, technical or otherwise, which inevitably seem to crop up! I am also grateful to Petri for allowing me to use his photography from the Taita Hills on the cover of this thesis.

Many thanks go to Professor Michael Schaepman and Docent Markus Holopainen for agreeing to be my pre-examiners and reviewing this work, and to Professor Philip Lewis for agreeing to be my opponent. I would also like to thank the Head of Department Professor John Westerholm for his support. I am particularly grateful to Juha Suomalainen for his sustained assistance with FIGIFIGO data collection and processing, and for the informative discussions on the physics of multiangular reflectance.

The friendly atmosphere in the University of Helsinki's Department of Geography has been much appreciated and has positively contributed to my work and my enjoyment of my time in Helsinki. For that I say thank you to all the other PhD students and the staff, particularly Tino Johansson and Jari-Pekka Mäkiaho, and especially Gareth Rice for the tortuous gym sessions and enlightening discussion on politics and the arts. Furthermore, a big big thank you to Dr G for proof reading my synopsis; no small work! Johanna Jaako and Airi Töyrymäki have been very helpful for all administrative matters and I am grateful for the IT support which I have received from Tom Blom and Hilkka Ailio. Also thanks to Matti Mõttus for his very useful and informative discussions with me right at the end of this thesis research, and to Juhani Virkanen who assisted in the analysis of sediments. Antero Keskinen was very helpful in assisting me with the spectrometer field measurements we made in the Taita Hills, and also being my chauffeur a couple of times! Thanks also to the visiting scholars at the Department of Geography, University of Helsinki, for the enjoyable times, notably Hakki Erdogan.

Particular gratitude goes to Tuuli Toivonen for the discussions and encouragement on my thesis when it was most needed. Also, thanks to her for having two children whilst I have been working on my PhD so that, indirectly, I was able to get some funding through the "trickle down" of finances!! Financial support from the Academy of Finland through the TATIA and TAITATOO projects, and the Department of Geography at the University of Helsinki, was gratefully appreciated, as was the Helsinki SPOT data provided by Spot Image under OASIS research grant 51.

The completion of this thesis has been a long and winding road for me, and where would I be without the "Taita team"… well probably a good deal more stressed out than I already am… so you guys need a special mention!! Mika Siljander (a.k.a Chilli) it has been an honour to work with you. I will never forget our enjoyable times working and travelling together in Kenya and all the long discussions we have had on

all sorts of topics, but most especially travelling. I have never met a man who is so knowledgeable about places, or as well travelled as you. You are a true inspiration to me. Also, through seeing your continued hard work and efforts, I think I have understood what "sisu" really is! Nina Himberg (a.k.a Princess) I respect you very much and it is always a joy to talk with you and that kept the office life lively, having discussions on all and any subject. You always laugh at my jokes as well, which is great! Alemu Gonsamo (a.k.a Moozo) you have earned my highest respect with the way that you worked so hard and consistently throughout your PhD. Our scintillating discussions over the years have shown me what a truly knowledgeable and intellectual man you are, with a very great depth of understanding about the workings of the world (for better and for worse!). You truly deserve all the good things that come to you, and I wish you well in what is sure to be a flourishing academic career. Eduardo Maeda, I have been very impressed with the speed and efficiency that you have extended and deepened the environmental modelling research in the Taita Hills. I wish you all the best in finishing up your PhD in record time! Janne Heiskanen, it was a pleasure to work with you and I have always valued your considered opinions. I wish you and your family all the best. Thanks also to Jan Hjort for all his interesting chats and for putting up with sharing an office with me for all those years! I wish you all the best in what I am sure will continue to be a very successful academic career.

I would also like to thank Andrew Harrison, formerly of the Department of Geography at the University of Bristol, for those early (in terms of my GIS/RS career, for what it is!) discussions on remote sensing and for the recommendation to Steve Hurcom, which got me started working in the geoinformatics area. Or maybe, actually I should BLAME it all on you!! Heartfelt thanks also go to my excellent former boss Paul Sill, and to all the other members of the MSI/Marconi Geodata Team; you know who you are! It was a great privilege to work with you all over those years, and immense fun as well as a great learning experience in the "real world" of business, which is a very useful compliment to my years now spent in academic research.

There are also lots of other people, too numerous to mention in name, that have contributed towards my PhD studies and to my time in Helsinki in many different and wonderful ways, and to whom I would like to say thank you! I wish everyone who I have befriended here in Helsinki all the best and I will always look back fondly on the years I spent in Finland.

Last but certainly not least, very special thanks go to my family for their continued understanding, support and encouragement. They have given me their warmly appreciated unconditional love, knowing that I would be away from "home" for all these long years. I have missed you all many many times.

Punavuori, Helsinki, 4th February, 2010

Barnaby John Forrester Clark

EARTH

"***The Blue Marble***" is a famous photograph of Earth taken on December 7, 1972 by the crew of the Apollo 17 spacecraft at a distance of about 29,000 kilometres (18,000 miles). It shows Africa and the Arabian Peninsula, Madagascar, and Antarctica. It was an important image in opening the eyes of the world to Earth as a single planetary entity, to the connectedness and oneness of our home, to Earth as a living organism - Gaia, and to the beauty and vulnerability of Earth within the overwhelming dark vastness of space.

This is a NASA public domain image.

LIST OF ORIGINAL ARTICLES

This thesis consists of a review and summary of the following six research articles, which are referred to in the text by their Roman numerals:

I. **Clark B**, J Suomalainen & P Pellikka. The selection of appropriate pseudo-invariant ground targets for use in empirical line spectral calibration of SPOT satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, submitted for publication.

II. **Clark B**, J Suomalainen & P Pellikka. An historical empirical line method for the retrieval of surface reflectance factor from multi-temporal SPOT HRV, HRVIR and HRG multispectral satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, submitted for publication.

III. **Clark B**, J Suomalainen & P Pellikka. A comparison of methods for the retrieval of surface reflectance factor from multi-temporal SPOT HRV, HRVIR and HRG multispectral satellite imagery. *Canadian Journal of Remote Sensing*, submitted for publication.

IV. **Clark BJF** & PKE Pellikka (2009). Landscape analysis using multi-scale segmentation and object-oriented classification. *In* Röder A & J Hill (eds.): *Recent Advances in Remote Sensing and Geoinformation Processing for Land Degradation Assessment*, 323-341. Taylor & Francis, London.

V. Erdogan HE, P Pellikka & **B Clark**. Modelling the impact of land cover change on potential soil loss in the Taita Hills, Kenya, between 1987 and 2003 using remote sensing and geospatial data. *International Journal of Remote Sensing*, in press.

VI. Siljander M, **BJF Clark** & PKE Pellikka. A predictive modelling technique for human population distribution and abundance estimation using remote sensing and geospatial data in a rural mountainous area in Kenya. *International Journal of Remote Sensing*, in press.

The articles are reprinted with kind permission of the publishers. Some of the articles contain colour figures, which have been printed below in greyscale.

AUTHOR'S CONTRIBUTION

I was solely responsible for writing research Papers I, II, III and IV. Juha Suomalainen was responsible for FGI's FIGIFIGO field measurements and for providing model fitted BRDF data and the 3D anisotropy factor "cake" plots. For Paper V, I contributed the text for the remote sensing and study area sections, and processed the remote sensing and elevation data, whilst Hakki Erdogan conducted the USLE analysis and Petri Pellikka formulated the article. During the revisions process, I rewrote parts of the article. For Paper VI, I wrote the remote sensing sections and processed the remote sensing data, whilst Mika Siljander conducted the data analysis and formulated the article.

SYMBOLS AND ABBREVIATIONS

θ_V	Sensor view zenith incidence angle (applied to SPOT imagery)
θ_{VZ}	Sensor view zenith angle
θ_Z	Solar zenith angle
λ	Wavelength
μm	Micrometre
π	A mathematical constant (Pi) whose value is the ratio of any circle's circumference to its diameter in Euclidean space (3.14159…)
ρ	Reflectance
ρ_s	Surface reflectance factor
ρ_{SAT}	Top of atmosphere at-satellite reflectance
τ	Atmospheric optical thickness or depth (AOD)
φ_V	Sensor azimuth angle
φ_Z	Solar azimuth angle
φ_r	Relative azimuth between sensor azimuth and solar principal plane
Ω	Projected solid angle of cone of radiance
ω_r	Conical solid angle of sensor IFOV
cos *i*	Cosine of the solar incidence angle
Gt	Gigatonne
ha	Hectares
i	Subscript denoting incident radiance
km	Kilometre
L_p	Path radiance
L_{SAT}	Top of atmosphere at-satellite radiance (W m^{-2} sr^{-1} μm^{-1})
ln	Natural logarithm
m	Metre
nm	Nanometre
r	Subscript denoting reflected radiance
T_V	Atmospheric transmittance from ground target to sensor
T_Z	Atmospheric transmittance from Sun to ground target
W m^{-2} sr^{-1} μm^{-1}	Watts per square metre per steradian per micrometre

AOD	Atmospheric optical depth
AREONET	Aerosol Robotic Network
ASD	Analytical Spectral Devices Inc.
ASPRS	American Society for Photogrammetry and Remote Sensing
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
AWiFS	Advanced Wide Field Sensor
BRDF	Bidirectional reflectance distribution function
BRF	Bidirectional reflectance factor
CBERS	China-Brazil Earth Resources Satellite
CCD	Charge-coupled device
CEOS	Committee on Earth observation satellites
CEOS-WGCV	CEOS Working Group for Calibration and Validation
CNES	French Centre National d'Etudes Spatiales
COST	Cosine of solar zenith angle, T_Z estimation technique
DDV	Dark dense vegetation
DEM	Digital elevation model

DN	Digital numbers
DOS	Dark object subtraction
ECV	Essential Climate Variables
EL	Empirical Line
EMR	Electromagnetic Radiation
ENVISAT	Environmental Satellite
EOS	Earth Observing System
ERS	European Remote Sensing
ERTS-1	Earth Resources Technology Satellite-1
ESA	European Space Agency
ETM+	Landsat Enhanced Thematic Mapper Plus
ETM+ SLC	Landsat Enhanced Thematic Mapper Plus, Scan-line corrector
FAO	Food and Agriculture Organization of the United Nations
FIGIFIGO	Finnish Geodetic Institute Field Goniospectrometer
FAPAR	Fraction of absorbed photosynthetically active radiation
FGI	Finnish Geodetic Institute
FMI	Finnish Meteorological Institute
FOV	Field-of-view
FWHM	Full width at half maximum
GAM	Generalized additive model
GCOS	Global Climate Observing System
GCP	Ground control point
GEOSS	Global Earth Observation System of Systems
GIFOV	Ground instantaneous field-of-view
GIS	Geographical Information Systems
GLCM	Gray Level Co-occurrence Matrix
GLM	Generalized linear models
GOLD	Global Observations of Land Cover Dynamics
GOOS	Global Ocean Observing System
GPS	Global Positioning System
GRASP	Generalized regression analysis and spatial prediction
GTOS	Global Terrestrial Observing System
GPWv3	Global Population of the World Version 3
HCRF	Hemispherical-conical reflectance factor
HDRF	Hemispherical-directional reflectance factor
HELM	Historical Empirical Line Method
HPD	Hierarchical patch dynamics
ICSU	International Council for Science
IFOV	Instantaneous field-of-view
IGBP-DIS	International Geosphere-Biosphere Programme Data and Information System
IGN	Institut Geographique National
IPCC	Intergovernmental Panel on Climate Change
IPCC-AR4	Fourth Assessment Report
IRS	Indian Remote Sensing satellites
ITCZ	Inter-Tropical Convergence Zone
IUCN	International Union for Conservation of Nature
KARI	Kenya Agricultural Research Institute
KMI	Kenya Meteorological Institute
LAI	Leaf area index
LCCS	Land Cover Classification System

LIDAR	Light detection and ranging
LISS	Linear Imaging Self-Scanning Sensor
LULC	Land use and land cover
MAD	Mean absolute difference
MERIS	Medium Resolution Imaging Spectrometer
MISR	Multiangular Imaging SpectroRadiometer
ML	Maximum-likelihood classification
MMU	Minimum Mapping Unit
MODIS	Moderate Resolution Imaging Spectroradiometer
MSS	Multi-spectral scanner
MSS/ORM	Multi-scale segmentation/object relationship modelling
NASA	National Aeronautics and Space Administration
NDVI	Normalized difference vegetation index
NIR	Near infrared
NOAA	National Oceanic and Atmospheric Administration
OO	Object-oriented
PIF	Pseudo-invariant features
PRA	Participatory Rapid Appraisal
PSF	Point spread function
RADAR	Radio detection and ranging
RGB	Red/Green/Blue
RMSE	Root mean squared error
RTM	Radiative Transfer Model
SAR	Synthetic Aperture Radar
SNSB	Swedish National Space Board
SSTC	Belgian Scientific, Technical and Cultural Services
SPOT	Satellite Pour l'Observation de la Terre
SPOT HIVIR	High Resolution Visible Infrared (SPOT 4)
SPOT HRG	High Resolution Geometric (SPOT 5)
SPOT HRV	High Resolution Visible (SPOT 1, 2 & 3)
SPOT-VGT	SPOT VEGETATION
SWIR	Shortwave infrared
TEEB	The Economics of Ecosystems and Biodiversity
TM	Landsat Thematic Mapper
UN	United Nations
UNEP	United Nations Environment Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFCCC	United Nations Framework Convention on Climate Change
USLE	Universal Soil Loss Equation
UTM	Universal Transverse Mercator
WIST	Warehouse inventory search tool
WGS-84	World Geodetic Survey 1984
WMO	World Meteorological Organization
6S	Second Simulation of the Satellite Signal in the Solar Spectrum

CONTENTS

PREFACE

"*Activities that devastate the environment and societies continue unabated. Today we are faced with a challenge that calls for a shift in our thinking, so that humanity stops threatening its life-support system. We are called to assist the Earth to heal her wounds and in the process heal our own – indeed, to embrace the whole creation in all its diversity, beauty, and wonder. This will happen if we see the need to revive our sense of belonging to a larger family of life, with which we have shared our evolutionary process.*"

Dr. Wangari Maathai
2005 Nobel Peace Prize acceptance speech

As I write this I am sitting at my cluttered desk in my diminutive flat in Helsinki on a bitterly cold Sunday afternoon in late January 2010. Out of my window, through the fading dim grey light, I can see a flurry of snowflakes swirling through the blustery air, desperate in their futile battle to stay aloft and resist joining their brethren already congealing on the frozen ground. My mind begins to wonder and I start to think of the huge variations between the Finnish seasons, of Nature's rhythms and cycles, of the inexorable passage of time, but moreover of the fate of the future for our incomprehensibly beautiful and precious, but threatened planet.

It was an early start to a bright Canadian summer's day in St John, New Brunswick, on the 23rd July 1917, then a small town nestled amongst the seemingly endless boreal forest, when Elizabeth Maud Goodeve Hutchings came into the world. She is my wonderful grandmother, and she is a respectable 92 years old at the time of writing. It is both awe-inspiring and deeply disturbing to consider, even for a fleeting moment, the changes that have occurred on Earth during her lifetime; so much good, so much bad, so many astonishing technical advances, so many irreplaceable losses. More than anything, the twentieth century witnessed extraordinary human population growth from 1.65 billion to 6 billion, and world population is expected to reach 6.9 billion in 2010 (UN Population Division 1999; UN Population Division 2008). The highest rates of population growth, averaging 2% per year, were experienced during the late 1960s, and the largest annual increments to world population of approximately 86 million persons each year occurred in the late 1980s (UN Population Division 1999).

Today, whilst citizens of the world's developed countries are privileged to live in a technologically advancing consumerist age, within the timeframe of a few human lifetimes, the unabated growth of human population, geographically uneven rapid economic development and industrialization, the ever increasing exploitation of the world's resources, and the continued global disparities and inequalities in income and access to resources, have led to an escalation of environmental problems and pressures to the extent that they now threaten sustainability at a global scale. It is easy for those in the food secure developed countries to overlook that, according to the Food and Agriculture Organization of the United Nations (FAO 2009), in 2009 1.02 billion people were undernourished, that is to say were starving. That is approximately 1 in 6 people on Earth, and 100 million more individuals than in the previous year, 2008.

It is difficult to overstate the severity of the current situation, and a summary listing of major environmental threats makes for very grim reading indeed: ¼ of the Earth's land surface is undergoing degradation, ½ the World's wetlands have been drained, ¾ of fishing grounds are fully or overexploited, and some predictions state

that, due to rising sea temperatures and acidification, coral reefs could have completely disappeared by 2050 (UNEP 2010a). The 2007 UN Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) unequivocally stated that humans have significantly changed the composition of the atmosphere and that, as a result, Earth's climate is changing and operating as an accelerator for associated environmental disturbances such as melting of snow and ice, rising sea level, flooding, droughts and wildfires. In addition to direct emissions of greenhouse gases, other human actions such as land use change, landscape fragmentation and overexploitation of natural resources are acting as drivers of climate change. For increases in global average temperature exceeding 1.5 to 2.5°C, there are projected to be major changes in ecosystem structure and function, species' ecological interactions and shifts in species' geographical ranges, with predominantly negative consequences for biodiversity and ecosystem goods and services, e.g. water and food supply. Approximately 20 to 30% of plant and animal species are likely to be at increased risk of extinction, and the resilience of many ecosystems is likely to be exceeded this century (IPCC-AR4 2007). As noted in the Climate Change Science Compendium (UNEP 2009a), the growth in carbon dioxide (CO_2) emissions from energy and industry has exceeded even the most fossil-fuel intensive scenario developed by the IPCC at the end of the 1990s. Global emissions were growing by 1.1% each year from 1990-1999 and this accelerated to 3.5% per year from 2000-2007.

Globally, deforestation continues at an alarming speed, with around 13 million hectares of forest lost per annum from 1900 to 2005 (FAO 2007), mainly from conversion into agricultural land. That is a loss in forest cover equivalent to the size of Bangladesh every year. Furthermore, forests act as vital carbon sinks and it is estimated that the World's 4 billion hectares of forests (FAO 2007) store 283 gigatonnes (Gt) of carbon in their biomass and that this stock decreases by 1.1 Gt annually because of deforestation and forest degradation (FAO 2006). Deforestation and forest degradation accounted for 17% of total anthropogenic greenhouse gas emissions in 2004 in terms of CO_2 equivalents (UN-MDG 2009), for example. Globally c. 36% of forest areas are primary forest, that is forests of native species in which there are no clearly visible indications of human activity and where ecological processes are not significantly disturbed. About 6 million hectares of these forests have been lost or modified each year since 1990 (FAO 2007). The continuing decline in primary forests in humid tropical countries especially is a matter of most serious concern, not only because of the impact on the global carbon budget, but also due to the loss of important natural habitats and rich biodiversity (Melillo *et al*. 1996). While there are currently insufficient data to determine globally the level of forest biological diversity decline, there is nonetheless a clear downward trend in key countries in which primary forests are under pressure from growing populations, expansion of agriculture, poverty and logging (FAO 2007). From 2000 to 2005 nine of the ten countries that account for more than 80% of the World's primary forest lost at least 1% of this area, led by Indonesia (13% loss in just five years), Mexico (6%), Papua New Guinea (5%) and Brazil (4%) (FAO 2007).

Humid tropical forest clearing from 2000 to 2005 was estimated by Hansen *et al*. (2008) to be 27 million hectares, which represented a 2.4% reduction in the area of humid tropical forest. 55% of total biome clearing occurred within only 6% of the biome area, emphasizing the presence of forest clearing "hotspots." Forest loss in Brazil accounted for 48% of total biome clearing, nearly four times that of the next highest country, Indonesia, which accounted for 13%. Over three-fifths of clearing occurred in Latin America and over one-third in Asia. Africa contributed 5% to the

estimated loss of humid tropical forest cover, reflecting the absence of current agro-industrial scale clearing in humid tropical Africa. However, considering all forest types, Africa played a significant part in global deforestation. Over the period 2000 to 2005 the largest net losses of forests occurred in Latin America and the Caribbean, with c. 4.3 million hectares of deforestation annually, but this was closely followed by Africa with c. 4.0 million hectares lost annually (FAO 2006; UN-MDG 2009). The estimated total forest area for Africa in 2005 was 635 million hectares, accounting for c. 16% of global forest area.

The UN Global Biodiversity Outlook 2 study estimates that 130 species become extinct every day, more than 1000 times the natural extinction rate (Secretariat of the Convention on Biological Diversity 2006). Whilst it is thought there are 90 million or more species on Earth, only around 2 million species are known to science (UNEP 2010a). Therefore, we know little of what biodiversity we are irrevocably losing every year. According to the International Union for Conservation of Nature (IUCN) 2009 red list, 12% of birds, 21% of mammals, 28% of reptiles, 30% of amphibians, 35% of invertebrates, 37% of freshwater fish, and 70% of all plants are threatened with extinction. In response to the severity of this situation, the UN has declared 2010 as the International Year of Biodiversity. To date, the international community has failed to reverse the rate of loss of biodiversity, despite that nearly all countries signed the 1992 UN Convention on Biological Diversity. Economies everywhere continue to dismantle the productive life-support systems of Earth. The most recent UNEP Economics of Ecosystems and Biodiversity (TEEB) study estimates that up to US$5 trillion-worth of natural or nature-based capital is lost annually.

The true value of ecosystem products and services flowing into the world economy have not previously been understood or successfully integrated into the world economy. As Dr. Achim Steiner, the current executive director of the UN Environment Programme (UNEP), stated in a 2010 radio interview with the BBC World Service, “the invisibility of biodiversity in economic terms has been one of the tragedies of the 20th century.” Biodiverse ecosystems not only provide essential goods to humanity, such as fresh water, food, fibre, and medicines, but also irreplaceable services, including purification of air and water, and regulation of pests, disease and soil erosion. Biodiverse forested watersheds in particular are a sustainable source of clean water, especially important with the ever increasing urbanization of humanity.

Apart from nature’s immediate usefulness to humankind, many - including myself - argue that all life has an intrinsic right to exist, and thus deserves protection. We must also recognize the right of future generations to inherit, as we have, a healthy planet thriving with life, and that continues to provide opportunities to harvest the economic, cultural and spiritual benefits of nature. However, the direct causes of biodiversity loss throughout many of the varied ecosystems of the Earth’s biosphere, such as habitat change, the introduction of invasive alien species, and climate change, show no signs of abating. Furthermore, the contributions of ecosystems to human societies promise to become all the more apparent and important as environmental change accelerates. Biodiverse ecosystems tend to be more resilient, and can therefore better cope with an increasingly unpredictable world (Secretariat of the Convention on Biological Diversity 2006). Climate change will bring more extreme weather events, from which intact ecosystems can offer physical protection. Higher levels of pollution will call for more detoxification processes, a service provided by healthy wetlands for example. Sadly, those already suffering from poverty will be most affected by biodiversity loss as the rural poor rely on ecosystems for their daily needs, and to sustain them through challenging times. When the services provided by

ecosystems are disrupted, the disadvantaged also lack the means to replace them. With proper sustainable management, however, ecosystems could potentially offer a path out of poverty.

I was born in 1975 and therefore environmental remote sensing, using civilian Earth observation satellite imagery, as a distinct scientific academic discipline is not much older than me! After leaving university the first time around, having obtained a BSc in Geography from Bristol, I joined the "rat race" and went to work on the 15th floor of a towering, shiny blue-glass office block in London's post-industrial reconceptualized and revitalized Docklands. It was an intimidating new office world for me on my first day as a sprightly, enthusiastic and somewhat naive 21 year old. That first morning at 9 o'clock sharp, as the lift doors opened with a ping and I subconsciously adjusted my already pinching taut tie before stepping forward over the threshold, and into the bright neon-like lighting of the reception, I began my initiation into corporate life and into the ins and outs of supplying and generating geospatial data for the ever changing and cutthroat telecoms market. The work was hard and the hours where long, but at least, I thought, I was on the way.

The view from my desk out of the tinted window was engrossing, drawing my gaze in the infrequent moments when I was able to take a break from the incessant phone calls and e-mails. It offered a vista across the entire sprawling grey mass of south and central London; through tantalizing glimpses of the slivery snaking Thames making its way, almost ignored, through the gritty heart of London, to the distant promise of the rural idyll in the far off, but just visible on a clear day, North Downs of Kent and Surrey. In the back of my mind I was always aware that there was a whole planet out there, a whole different world from London office life waiting to be explored.

Though the job was good, honing my skills in remote sensing and GIS, and the people I worked with were great, gradually, over the years I began to wonder if this really was what life was all about; the 9 to 5; eating takeaway jacket potatoes with cheese for lunch at your desk with one hand, whilst typing out an urgent quote for the overpaid sales team with the other – indigestion guaranteed; reading about the latest celebrity hairdo in the Metro every morning whilst waiting for the perpetually delayed 8:16 to Waterloo… then one day it happened. I had my epiphany. I got to the office, logged in, and opened my e-mail inbox. Straight to work, there was no Facebook in those days! Another request for mapping data pricing, "no problem" I thought, "I have covered that before, although a while ago now", so I opened up the previous quote and updated the info, but… I had totally forgotten what year it was… never mind the day or the week! That was when it hit me like a train, my life had become stayed; there was no differentiation; there was no real meaning; one year *was* the same as another. What else then, but time for a change, a real change, a significant change. At that point, having been lucky enough to have met and be in a relationship with a beautiful young Finnish woman, whom I had chanced upon in a heaving laughter filled bar in London's atmospheric Covent Garden, I made a decision there and then to pack it all in and head off on a new path in life. Having worked for 7 years with the various stresses and strains of London's hyper-competitive office scene, I have indeed, therefore, been a very fortunate human being to be able to take the opportunity to go back to "studying" and to have the time once again to really think more deeply about matters.

"*The wisdom of a learned man cometh by opportunity for leisure: and he that hath little business shall become wise.*
How can he get wisdom that holdeth the plough, and that glorieth in the goad, that driveth oxen, and is occupied in their labours, and whose talk is of bullocks?"

The Bible, Ecclesiasticus, 38:24-25.

Within this thesis I have covered a lot of academic ground in terms of the subject areas that I have delved into - ranging from a consideration of the hydrocarbon absorption spectra of an asphalt car park in Espoo (Finland) to the indigenous forest core habit distribution of the critical endangered East African Taita Thrush *Turdus helleri* - throughout all of which I have strived continuously not to talk 'bullocks'! I do, nonetheless, accept any errors and omissions within this thesis as my own responsibility. Through necessity of the specifics of the subject matter, this thesis is very methodological and technically oriented and thus offers little in the way of deep meaningful insights into the background causes and drivers of the environmental changes that have been mapped and modelled. However, I want to stress that these issues are of course fundamental and essential in addressing and hopefully resolving environmental degradation. Were it not for the specific aims of this thesis, a lack of detailed discussion on these issued would evidently be a fundamental oversight. Within the TAITA and TATATOO projects, however, there are others whose research considers the socio-economic and cultural aspects behind land use practices in the Taita Hills.

By offering insights into reliable processing methods for SPOT data - which are just as viable applied to any other type of medium or high resolution optical satellite imagery - it is my sincere desire that this thesis may be of some interest and use to the work of others in accurately monitoring and modelling environmental changes, hopefully with the aim of conserving natural resources and implementing sustainable development. Especially, I hope this thesis is of some assistance to those living and working in developing countries with limited financial, data and software resources. Ways of implementing the outlined methodologies in such circumstances are discussed throughout this thesis.

"*You must be the change you wish to see in the world.*"

Mahatma Gandhi, Indian political and spiritual leader (1869 - 1948)

The failure of the 2009 Copenhagen UN Climate Change Conference to reach consensus shows both the complexity of global environmental issues and a general lack of political will to move forward in properly addressing these problems. Are politicians' real roles to keep the status quo or to change it? As the French theorist Michel Foucault notes, knowledge is power. Therefore a better education and understanding for all human beings not only enables better local management of the environment, but can also give real weight to a grassroots change in approach and a chance to implement truly sustainable development. It is the duty of all forward thinking citizens of Earth to further a higher level of environmental consciousness and to do everything they can to work towards true sustainability in all aspects of their life. It's not easy to do, and it will take time and effort, but as my grandmother used to tell me, if you take care of the pennies the pounds will take care of themselves.

A good friend and colleague of mine, Mika Siljander, wrote in the Preface to his PhD thesis:

> "I strongly believe that we need to take into account abiotic, biotic and anthropogenic factors and processes behind environmental problems before we can resolve them. We need to have a broad knowledge for all these study fields, not just one. We need to have a holistic understanding for environmental problems and we need to combine more than just one entity to form something new — we need synthesis. And because geography is a discipline of synthesis; "*There has always been a view of geography as a discipline of synthesis. Holism has been there the whole time…*" (Holt–Jensen 1988), I therefore strongly believe that geographers can contribute fighting against environmental problems in a very special way as we geographers have the skill to see the "The Whole Picture"."

I think this also very well describes my opinion about the potential of geography. Geographers are often belittled as the "jack of all trades, and the masters of none", but this power of oversight may actually prove to be a crucial skill in successfully addressing the myriad of environmental issues that will no doubt continue to confront us in the years to come.

1. INTRODUCTION

1.1 Remote sensing based environmental monitoring of the Earth's land surface

Earth is currently experiencing very serious environmental problems and threats at all scales from the local level right through to the global level. Figures from the United Nations (UN) Population Division (2008) show that world population rose from approximately 2.5 billion in 1950 to 6.5 billion in 2005, and is expected to reach 6.9 billion in 2010. Moreover, world population is projected to continue growing quickly and, for example, to increase to 9.1 billion by 2050. This unprecedented rapid population growth, and the consequent escalating requirements for natural resources to meet even the most basic human needs of water, food, and shelter, are without doubt one of the main drivers of land cover change and increasing environmental degradation (Preston 1996).

The largest rates of population growth have been, and are expected to continue to be, in the developing countries of Africa and Asia, and are having significant social, economic, and environmental impacts on these regions. Africa's population has grown from c. 227 million in 1950 to about 1 billion people in 2010. With the highest birth-rate of any continent – 2.3% (2005-2010) compared to a world average of 1.1% (2005-2010) – the population is projected to grow to 2 billion by 2050 (UN Population Division 2008). Africa's total land area is about 29.6 million km^2, which represents 20% of the Earth's surface. Over 65% of Africa is arid or semi-arid (of which 43% is desert), c. 22% is forest, and only ~20% is suitable for cultivation (FAOSTAT 2006). Pressure on Africa's environment, which is mainly the result of population growth, has intensified and in many areas has exceeded the capacity of natural resources to meet expanding human needs with the current technology. This has resulted in loss of forests and biodiversity, land degradation, increasing water shortages and declining water quality (Eva *et al*. 2006).

Land degradation is a composite term that describes the negative impacts of a large number of naturally occurring but human influenced processes on an environment (Stocking & Murnaghan 2000). It implies a reduction in the productive capacity of an ecosystem and in its value as an economic resource (UNEP 1992). At a localized landscape level there is a strong cyclical connection between land use and degradation. Changes in land use over time, such as agricultural expansion, influence degradation processes and this in turn affects the future utilization of the land. Across many parts of Africa, conversions of land to agriculture associated with poor land management practices have caused considerable land degradation and erosion. About 25% of Africa is subject to water erosion and 22% to wind erosion, and desertification processes affect over 45% of the land area, of which 55% is at high or very high risk (UNEP 1992). Water stress and scarcity are endemic in approximately a quarter of all African countries, and generally water quality is deteriorating while water-related diseases are increasing (Eva *et al*. 2006). Africa has only 9% of global renewable freshwater resources and the average water availability per person is far less than the global average and falling (FAO AQUASTAT 2006). It is estimated that around 6,000 people die in Africa every day as a result of contaminated water and poor hygiene, and about 3 million annually as a result of water-related diseases (UNEP 2005). In many regions, especially northern and eastern Africa, land degradation and water stress are exacerbated by the high variability and unreliability of rainfall, often

resulting in prolonged droughts followed by periods of intense rainfall, which regularly causes severe flooding (Eva *et al*. 2006).

The last 40 years have seen unprecedented land cover and land use changes occurring in sub-Saharan Africa. These have been a consequence of a series of both naturally occurring and man-made factors and major disturbances; notably high rates of population increase, civil wars and resultant population movements, economic development and globalization, as well as droughts and increasingly significant climate change (Brink & Eva 2009). These factors have acted to continuously erode natural resources and ecosystems in the region. The conversion of natural vegetation to agriculture by the expansion of croplands is the major land cover dynamic in sub-Saharan Africa. For example, based on a stratified random sampling procedure with Landsat Thematic Mapper (TM) satellite imagery, Brink & Eva (2009) found that, over the period 1975 to 2000, agricultural areas increased by 57% in sub-Saharan Africa from c. 200 million hectares (ha) to nearly 340 million ha, an average annual change rate of 2.3%. This resulted in a 5% loss of woodlands and grasslands (60 million ha) and a 16% loss of forests (71 million ha), although it should be noted that the largest forest losses occurred outside the humid forests of the Congo basin, which accounted for only 16% of the total forest loss, despite accounting for more that 80% of sub-Saharan Africa's forest area. Assuming linear change over time, this indicates a yearly deforestation rate of 0.7%, meaning that the region has been losing nearly 3 million ha of forests per annum. Further, barren areas increased by 15%, amounting to 6.5 million ha. Every year from 1975 to 2000, then, over c. 50 000 km^2 of natural vegetation in sub-Saharan Africa was converted to croplands.

Such decreases in natural vegetation lead not only to a fragmentation and loss of habitat, biodiversity and stored carbon, but also to a degradation or loss of ecosystem services such as sources of freshwater, pastures, forest resources, fuel wood and bush meat, with potential long-term impacts on climate, socio-economic stability, sustainable food production, food security and human welfare (Foley *et al*. 2005; Eva *et al*. 2006). A consequence of this, for example, can be to make protected areas or reserves some of the few remaining zones within a region where fuel wood, richer pastures and game resources are more abundant or easily available, hence attracting and encouraging illegal and unmanaged resource use activities in such areas (Clerici *et al*. 2006). Furthermore, where expansion in and management of agricultural areas is poorly organized and administered, as is often the case in sub-Saharan Africa, changes in the soil water cycle, vegetation cover and nutrient depletion, can give rise to feedback mechanisms in localized land degradation processes leading to increased soil erosion, as well as natural hazards such as floods and landslides (Brink & Eva 2009).

The population of eastern Africa increased from c. 65 million in 1950 to 287 million in 2005, and is projected to reach 711 million by 2050 (UN Population Division 2008). In Kenya, both the total fertility rate of 5.0 and a growth rate of 2.8% are extremely high. Kenyan population has grown from c. 6 million in 1950 to 31.5 million in the 1999 census, and is projected to be 85.5 million by 2050 (UN Population Division 2008). Many land cover change studies have reported intensified land use, expanding agricultural areas at the expense of natural vegetation cover, increased environmental pressures, and forest loss, fragmentation and degradation throughout Kenya's ecologically important regions. For example, using aerial photography from 1958 and 1985 and SPOT satellite image from 1995, Imbernon (1999) reported an almost total loss of bushland in the upper Embu area, with a 75% increase in the cultivated area between 1985 and 1995, and although there was only a

minor reduction in forest cover, there was a significant change in the species composition of the forests from native tree species to exotic species, such as grevillea and eucalyptus. Using Landsat TM satellite imagery from 1986, 1995, and 2003, Baldyga *et al*. (2007) studied land use changes in the Mau forest area, the largest indigenous forest in East Africa and Kenya's most vital water tower (UNEP 2010b), and identified significant deforestation between 1995 and 2003 related to the expansion of small-scale agriculture, shared almost equally between exotic plantations and indigenous forests. Lung & Schaab (2006) studied land cover changes and forest fragmentation and disturbance in the vicinity of Kakamega forest from 1972 to 2001, based on the analysis of seven Landsat satellite images, and they too identified severe deforestation and fragmentation increasing over time, especially in the southern part and edges of the forest, related to accessibility and local population distribution. A general overview of environmental degradation for a number of key Kenyan ecosystems is given in the UNEP Kenya: Atlas of Our Changing Environment report (UNEP 2009b).

In the ecologically important and environmentally sensitive Taita Hills in southeastern Kenya, which are the main study area of this thesis, local population in the Taita-Taveta district as a whole has increased from 90 000 in 1962 to 246 671 inhabitants counted in the 1999 census (Republic of Kenya 2001). The population is particularly concentrated in the fertile Taita Hills themselves, where population density is 400-500 persons per km^2 in the best agricultural areas and over 900 persons in some sub-locations (Republic of Kenya 2001). There has been an increase in the area under cultivation in the Taita Hills and due to poor agricultural management, erodible soils and the rugged topography, the area is subject to land degradation and accelerated soil erosion (KARI 2005), especially in the foothills. The Taita Hills form the northernmost part of Africa’s Eastern Arc Mountains, which have been identified by Conservation International as one of the top ten biodiversity hotspots in the world. Of particular scientific and conservation interest are the remnant indigenous forest patches, which are home to many rare or endangered endemic animals and plants (Githiru & Lens 2007). Identified threats to the these forest patches include encroachment, over extraction of firewood and building materials, charcoal burning, illegal logging, lack of awareness among the communities living adjacent to forests, fires (both deliberate and naturally occurring) and colonization by suppressive and fast growing exotic tree species (EAWLS 2005). It can be seen, therefore, that land use and land cover (LULC) change, land degradation and population growth are key issues in the Taita Hills which require detailed study.

It can clearly be seen, then, that throughout Africa, and indeed throughout the developing world in general, there is an ongoing and pressing requirement for better spatial databases at local, national and regional levels covering a range of natural and socio-economic variables; most especially population, and LULC and LULC change over time (Eva *et al*. 2006). Further, these databases need to be regularly updated and should be standardized and interoperable, so that comparisons with and utilization within other studies can be undertaken. Spatially explicit up-to-date information on population is very useful for areas of the developing world experiencing rapid changes in population distribution and abundance, because the traditional nationwide census is often a painstaking error-prone operation, normally undertaken only once a decade. Additionally, demographic information is usually provided in national or administrative units of which the sub-national reference units can be of vastly different size and shape (Li & Weng 2005; Mubareka *et al*. 2008). For spatial analysis, however, it is often preferable to record human population estimates in

regular analysis grids, such as is the case with the LandScan database (Dobson *et al*. 2000). Moreover, the problem with existing population models is their coarse scale, for example LandScan has a grid resolution of 1 km^2, thus generalizing and obscuring the internal variability of population data. Therefore, at a local scale, there is great demand for cost-efficient applications to generate spatially explicit human population geospatial databases and distribution maps at finer resolutions.

Increasing population pressure and agricultural expansion, together with insufficient attention to natural resources management and environmental concerns is rapidly undermining Africa's environmental sustainability, particularly affecting the poorest countries and peoples. There is an urgent requirement, therefore, to develop and implement sustainable natural resources management policies, programs and planning (Eva *et al*. 2006). Whilst specific erosion processes and their resultant landforms, such as gully formation through soil erosion by water, can be usefully studied at a micro level, land degradation cannot be directly assessed at a macro level through any single measure. Rather, use must be made of indicator variables that demarcate the likely occurrence of degradation. At a landscape level, these indicators are changes in LULC patterns over time such as loss of vegetation cover, including deforestation. Crucial to an understanding of landscape changes is, then, to accurately map and quantitatively assess LULC and LULC changes over time at periodic intervals. As Brink & Eva (2009) note, further studies are required to provide information not just on the magnitude and nature of LULC changes but also to spatially localize them.

All these factors unequivocally show the significant and ongoing pressing requirement for quantitative, scientifically based, repeatable and accurate environmental monitoring and modelling from the local level through to global scales; an issue to which spaceborne satellite remote sensing of the Earth's surface is uniquely placed to provide information and insight. Satellite remote sensing offers the possibility for routinely monitoring the environment over regional, as well as national, continental, and global scales, and makes it feasible to collect synoptic and spatially explicit data of inaccessible and extensive areas, with information available in spectral, spatial, angular and temporal resolutions and polarization domains. Consequently, remote sensing has played a substantial pivotal role in the development of understanding and modelling of a myriad of environmental processes, as well as in the monitoring of LULC (De Jong *et al.* 2004), from global down to local scales since the first civilian Earth observation satellites were launched in the early 1970s.

Land cover is defined as the observed (bio)physical cover on the Earth's terrestrial surface and includes vegetation and man-made features, as well as bare rock, bare soil and inland water surfaces (DiGregorio 2005). The primary units for characterizing land cover are thematic categories (e.g. Forest) or continuous variables classifiers (e.g. fraction of Tree Canopy Cover). Secondary outputs include surface area of land cover types, land cover change, and observational by-products such as field survey data and geometrically and atmospherically corrected satellite imagery products. Categories and classifiers must be defined consistently in order to identify land cover changes over time, and standardization of classification schemes is an important issue if greater use and understanding of digital mapping products is to be facilitated. However, existing inconsistencies between different land cover mapping products and change monitoring systems have complicated the ability to successfully synthesize land cover assessments on regional and global scales (GTOS-T9 2009). Consequently, the Land Cover Classification System (LCCS) software was developed by FAO and UNEP (DiGregorio 2005), which offers a standardized *a priori*

classification system that can be used for any land cover mapping exercise regardless of the methodology, scale, source material and geographic location (DiGregorio 2005). The LCCS classifiers provide a comprehensive and flexible framework for thematic land cover characterization, and enable compatibility to be achieved between existing datasets and for future monitoring systems. However, it is only recently that LCCS has become recognized and widely used to provide a basic level of thematic land cover standardization (GTOS-T9 2009).

In current practice, many research institutes and observational programmes do not explicitly distinguish between land cover and land use. Land use characterizes the arrangements, activities and inputs people have undertaken on a certain land cover type to produce, change or maintain it (DiGregorio 2005). However, because of the implicit or explicit role of humans in land use characterization and mapping, DiGregorio (2005) argues it should be considered distinct and dealt with separately from land cover type, thus ensuring internal and external consistency and comparability. Land cover and land use transitions may be interoperable, such as a change from natural vegetation cover to cropland agriculture, but this relationship does not hold in all circumstances because land use characterization includes considerations that go beyond land cover. Nevertheless, in this thesis discussion and in thesis Paper IV, the term LULC is used in place of land cover where appropriate as, by including the Cropland, Plantation Forest and Built-up Areas classes in the mapping of the Taita Hills, some inference is being made about the land use, even if in the strictest sense the classes have been derived from LCCS. However, in the modelling Papers V and VI the mapping data is referred to specifically as land cover.

LULC and vegetation have a central role in the Earth's climate and hydrological systems, and in biogeochemical cycling. Vegetation meditates ca. 90% of the gaseous exchange between the terrestrial biosphere and the atmosphere (Ozanne *et al.* 2003); consequently, changes in LULC that modify vegetation alter the rate of carbon exchange (Houghton 2003) and can have a significant impact on atmospheric carbon dioxide (CO_2) concentrations (Hilker *et al.* 2008). Furthermore, the emergence of a number of global scale environmental problems has led to many policy-driven needs, notably binding international agreements, which have motivated the production of land cover information for biodiversity assessments (Secretariat of the Convention on Biological Diversity 2006), climatic modelling, and quantification of carbon cycling (DeFries & Belward 2000; Rosenqvist *et al.* 2003), for reporting and scientific modelling purposes. The 2007 UN Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) unequivocally stated that humans have significantly modified the composition of the atmosphere and that, as a result, Earth's climate is changing and operating as an accelerator for associated environmental disturbances, such as melting of snow and ice, rising sea level, flooding, droughts and wildfires (IPCC-AR4 2007). The precise quantification of the rate of climate change is a divisive issue, but is very important in determining whether feedback or amplification mechanisms, in which the terrestrial surface plays an important role, are operating within the climate system. Consequently, the issue of climate change in particular has generated significant science- and policy-driven requirements for globally consistent sets of observational data, which are needed to attribute the causes of climate change, analyse the potential impacts, and evaluate the adaptation options (GTOS-52 2008). The climate observing system in the terrestrial domain is, however, still poorly developed, while at the same time there is increasing significance being placed on terrestrial data for impact, adaptation and mitigation activities.

The Global Climate Observing System (GCOS) Second Adequacy Report noted that the many difficulties encountered with regards to terrestrial observations, including the lack of homogeneous datasets and internationally agreed validation protocols and benchmarks, could be resolved by the creation of an intergovernmental technical commission for terrestrial observations, similar to those that existed for the atmospheric (GCOS) and oceanic domains (Global Ocean Observing System, GOOS). As a result, the Global Terrestrial Observing System (GTOS) was initiated by FAO, UNEP, WMO, UNESCO, ICSU and other stakeholders, to support the terrestrial observational requirements of the United Nations Framework Convention on Climate Change (UNFCCC). In particular, in collaboration with the Global Earth Observing System of Systems (GEOSS) Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV), GTOS was tasked with establishing standardization (guidelines, benchmarks, acceptable data processing/generation methodologies) and developing international coordination for the 13 terrestrial Essential Climate Variables (ECVs) originally identified by the GCOS Implementation Plan: consisting of Land cover and Albedo, as well as Biomass, Fire disturbance, Fraction of absorbed photosynthetically active radiation (FAPAR), Glaciers and ice caps, Groundwater, Lake levels, Leaf area index (LAI), Permafrost and seasonally-frozen ground, River discharge, Snow cover, and Water use; observations of which are currently measured by numerous organizations for a variety of purposes.

A range of space agencies, research institutes, and national agencies currently carry out the operational production of multi-temporal global, regional and national land cover and LULC datasets. Trade-off between the spatial and temporal resolutions is a key issue in the selection of satellite imagery for LULC mapping (Aplin 2006), which can be divided between the requirement of mapping at “coarse” resolutions (250 m-1 km) for determining land cover type globally or over continental scales, and mapping at “medium” resolutions (5-50 m) for determining type and detecting land cover change (Achard *et al*. 2008; GOFC-GOLD 2008). In “medium” and “fine” ($\leq$ 5 m) resolution studies, the spatial extent of the data coverage is relatively small and the temporal resolution of the revisit time and the acquisition of cloud-free imagery is poor, whilst in the “coarse” resolution studies the data cover very large areas with high temporal resolution (Cihlar 2000); typically global coverage every 1-2 days. Additionally, *in situ* data are acquired for the monitoring of land cover, and as reference for calibration and validation of satellite derived land cover and land cover change measurements. Over the last twenty years a number of satellites have been launched that are specifically designed with capabilities to monitor land cover and land cover changes. For example, the MODIS, ASTER and MISR sensors of NASA’s Earth Observing System (EOS) offer a wide range of high level preprocessed data products, such as global land cover (Friedl *et al*. 2002) or LAI (Myneni *et al*. 2002), and with low cost or free web dissemination. Following the launch of very high resolution optical satellites such as IKONOS (4 m multispectral resolution) and QuickBird (2.44 m multispectral resolution) at the start of the last decade, sensors that have a spatial resolution of 20 to 30 m, such as SPOT and Landsat TM/ETM+ (30 m), are now often referred to as “medium” resolution (Blaschke *et al*. 2004), whereas previously they were considered high resolution. Somewhat confusingly, the term “moderate” resolution is used variously to apply to both Landsat type data and 250-500 m resolution MODIS type data. In this thesis discussion, satellite imagery having a pixel size larger than the ~80 m of Landsat MSS data are considered as being “coarse” resolution.

As is detailed in Table 1 and illustrated in Figure 1 below, even though there is a current strong emphasis in the remote sensing community on the development and standardization of operational global scale products, there is a permanent requirement for LULC mapping at a variety of scales, especially at the medium resolutions appropriate for detailed analysis of LULC and LULC change. The relevance and benefit of the research efforts into global level products, for the utilization of medium resolution optical satellite imagery data for environmental monitoring and modelling, is that it has driven forward standardization in data processing and validation. This allows small scale localized research projects and studies, such as the work carried out for this thesis, to contribute more widely, and the generated datasets can be utilized if detailed metadata on the preprocessing, processing and product validation are given. For example, by adhering to the GTOS developed standards and guidelines, and using LCCS for deriving the classification scheme, LULC data are interoperable with data from other scales and study areas.

This thesis concentrates on the utilization of multispectral SPOT data. The 20 m pixel spatial resolution of the HRV and HRVIR sensors, and the 10 m resolution of the HRG sensors, the 60 km by 60 km scene area (at nadir view), combined with spectral information covering the green, red, near infrared (NIR) and shortwave infrared (SWIR) wavelengths, make images captured by these instruments suitable for local and regional scale studies at what can be termed a landscape level. A landscape is actually a complex term that has changed meaning through time, although landscapes are increasingly regarded as complex systems composed of a large number of spatially heterogeneous components that interact in non-linear ways and exhibit emergence, self-organization and adaptive properties through time (Wu & Marceau 2002; Hay & Marceau 2004; see Section 4.6 for further discussion). Within the discipline of landscape ecology, diverse interacting patches are central to a conceptualisation of a landscape (Wu & Hobbs 2007). In this thesis, the term "landscape level" is taken to mean at a scale that coincides with human vision and perception of the environment, where humans differentiate a landscape into components such as fields, woodlands, built-up areas, etc, which relate to patches.

The landscape features in the Taita Hills are heterogeneous, small scale and dominated by cultivation, forming a patchwork of human settlements, small-holder subsistence plots, and small stands of plantation and native tree species. The average farm size is 2 ha (Soini 2005a; Ruotsalainen 2008) and the largest indigenous forest patches occur in a scatter of three hilltop remnants; Mbololo (c. 179 ha), Ngangao (c. 136 ha) and Chawia (c. 94 ha) (Lens *et al*. 2002). However, most of the other remnant indigenous forest patches are very much smaller; for example Yale was 15.7 ha, Fururu 8.1 ha, Macha 2.5 ha and Mwachora 2.3 ha in 2004 (Pellikka *et al*. 2009). There are also a number of patches formed from plantation forests. This landscape can be differentiated in the $\leq$ 20 m spatial resolution SPOT data, whereas in the c. 30 m Landsat TM/ETM+ imagery the resolution is insufficient to identify many of the smaller landscape features. Landsat data has additional spectral bands in the blue and SWIR wavelengths, but SPOT data has sufficient spectral information to enable the differentiation of major land cover types. SPOT temporal resolution is also appropriate for inter-annual LULC change studies (see Table 1 and Figure 1). The SPOT data therefore represents the coarsest "medium" resolution imagery that can be easily utilized for detailed LULC mapping and LULC change analysis in the Taita Hills, and this was a major reason why SPOT data was chosen. The SPOT data were acquired at low cost through the French Centre National d'Etudes Spatiales (CNES) ISIS programme.

The GTOS ECV-T9 guidelines for 30 m land cover type data requirements and suggested technical approach are also a useful basis as a general standard for landscape level LULC mapping for environmental monitoring and modelling studies in regions of the developing world with limited ancillary data availability:

Data requirements (adapted from GTOS-T9 2009):

1. Classification scheme based on LCCS with categories relevant for assessing a wide range of environmental applications. In particular, attention should be given to classes that are poorly represented in coarse-resolution output products, and those classes reflecting human land use (e.g. urban types, agricultural types, impervious surfaces).
2. A spatial resolution of ≤ 30 m with temporal updates every 3-5 years.
3. Overall and regional accuracies exceeding 90% at the highest level of aggregation.
4. Validation should be based on the use of a probability-based sampling strategy.

Technical approaches (adapted from GTOS-T9 2009):

The use of computer-assisted methods enables a cost-effective approach to creating accurate, high resolution products. Automated methods are preferable where possible as the interpretation is repeatable and efficient. Even in a fully automated process, visual inspection of the result by an analyst familiar with the region should be carried out to ensure correct interpretation. Validation must be statistically rigorous, although obtaining suitable sources of validation can be problematic; high resolution satellite imagery and aerial photography are comparable to *in situ* data, but may be costly; see Table 1.

Table 1*. The role of medium resolution optical satellite imagery, such as SPOT data, within land cover monitoring studies at multiple scales (after GTOS-T9 2009).*

Sensor & Resolution	Examples of Current Sensors	Common MMU	Update Frequency	Cost of Data Acquisition	Utility for Land Cover Mapping
Coarse Optical (250 m-1 km)	SPOT-VGT (1998-) Terra-MODIS (2000-) Envisat-MERIS (2004-)	10-100 ha	Annual	Free Internet dissemination or very low cost	Consistent global annual monitoring to identify phenological pattern, basic land cover types, large changes and locate "hotspots" for further analysis with finer resolution data
Medium Optical (5-50 m)	Landsat MSS (1972-1982) Landsat TM/ETM+ (1982-) SPOT (1986-) IRS AWiFs or LISS (1988-) CBERS (1999-)	0.5-5 ha	1-5 years	Some free, otherwise <$0.001/km^2 for historical data $0.02/ km^2 to $0.5/ km^2 for recent data	Primary tool to map major LULC and LULC changes and associated area estimates
Fine Optical (≤ 5 m)	IKONOS (1999-) QuickBird (2001-) Aerial photos (post WWII-) (Also *in situ* data)	< 0.1 ha	1-10 years	High to very high $2-30 km^2	Detailed surveys and mapping, validation of results from coarser resolution analysis, and training of algorithms

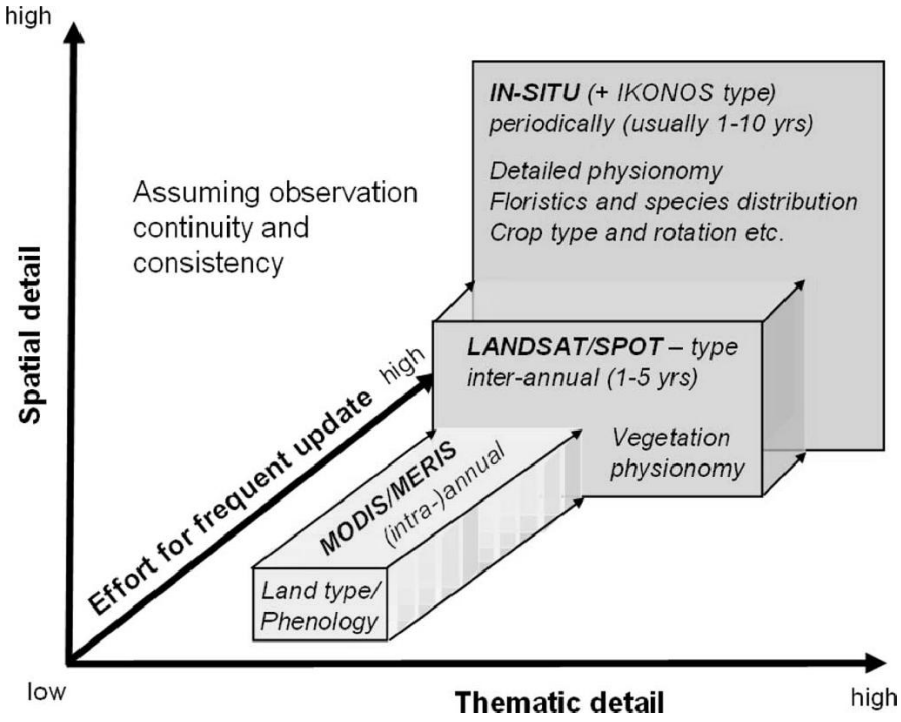


Figure 1. Conceptual location of medium resolution Landsat/SPOT type data within a framework for integrated observations of land cover and vegetation for environmental monitoring and modelling studies (from Herold *et al*. 2008).

It can be seen, then, that for many developing regions of the world, LULC change, land degradation and population growth are key issues. During the last few decades the growing number of remote sensing data sources, the development of remote sensing and geographical information systems (GIS) data analysis software, along with very significant improvements in computing resources, and the science- and policy-driven requirements for environmental data, have created major opportunities for improvements in LULC characterization (DeFries & Belward 2000) and have allowed for more sophisticated geospatial environmental applications to be developed. Geospatial environmental modelling using remote sensing, GIS and more recently spatial statistical techniques have deepened understanding of the mechanisms and driving forces behind environmental degradation. However, there is a need for further studies using standardized classification schemes, and data processing and validation procedures, to provide information on the magnitude and nature of LULC changes, and to spatially localize them. When also supplied with detailed metadata, this allows for interoperability and greater data usability. Moreover, there is still a requirement for improved methods of environmental monitoring and modelling utilizing remote sensing and geospatial data, in order to derive higher accuracy products.

1.2 Objectives of the thesis

This thesis contributes to knowledge on the utilization of medium resolution optical satellite imagery data for environmental monitoring and modelling at a landscape level, in regions of the developing world with limited ancillary data availability. This is based on the application of multi-temporal SPOT multispectral data to the Taita Hills study site in Kenya. Environmental monitoring in the Taita Hills involved the derivation of a time-series of surface reflectance factor (ρ_s) retrievals, as well as LULC maps and the analysis of LULC change over time. Together with a further SPOT imagery database and extensive multiangular ρ_s field measurements for the Helsinki metropolitan region, which was utilized as a control site, research was conducted into the best performing applicable atmospheric correction methodology to both accurately retrieve ρ_s and provide radiometric stability within a set of multi-temporal SPOT scenes. Also, in order to derive as accurate as possible mapping of the complex heterogeneous LULC in the Taita Hills, an advanced classification technique based on multi-scale segmentation/object relationship modelling (MSS/ORM) was applied and compared to a standard maximum-likelihood (ML) procedure. Environmental modelling in the Taita Hills involved the application of the remote sensing derived data along with other geospatial information to consider potential soil loss due to changes in land cover. Further, predictive modelling techniques were applied to map the distribution and abundance of human population based on SPOT and GIS derived geospatial data as predictor variables. The conceptual overlaps between the various stages in the preprocessing/processing of, and environmental monitoring and modelling with, the SPOT data are summarized in Figure 2.

Paper I examines the characteristics of spectrally pseudo-invariant surface types suitable for use in the application of empirical line (EL) spectral calibration techniques. Based on goniometer ρ_s measurements taken in the Helsinki metropolitan region control site study area, the spectral spatial, temporal, and multiangular reflectance variability of various asphalt, sand, gravel, grass, and fake turf targets are detailed. Consideration is given to the physical causes of the different behaviours and the most appropriate surface types, and the requirements for suitable site selection for EL calibrations, are discussed.

Paper II outlines the proposed *historical empirical line method* (HELM) for the retrieval of ρ_s from multi-temporal SPOT multispectral data. The application of HELM to SPOT imagery datasets covering the Taita Hills and the Helsinki metropolitan area is detailed, and the ρ_s retrieval accuracy results are compared to the partially corrected at-satellite reflectance (ρ_{SAT}). Calibration to nadir only ρ_s is denoted as the HELM-1 approach, whilst calibration to ρ_s modelling the exact illumination and view geometries of the imagery is termed the HELM-2 approach. The estimated error in applying HELM-1 to the control site SPOT imagery database, given the ± 31° range in the SPOT sensor view incidence angle (θ_V), is modelled.

Paper III utilizes nadir ρ_s collected in the Taita Hills application site, as well the multiangular ρ_s field data from the Helsinki metropolitan region control site, to undertake a comparative assessment of the absolute atmospheric correction techniques applicable in local and regional landscape level remote sensing studies, in circumstances where no detailed overpass concurrent atmospheric measurements or meteorological data are available. Namely, in addition to HELM, dark object subtraction (DOS) and the Second Simulation of the Satellite Signal in the Solar Spectrum (6S) radiative transfer model (RTM) atmospheric correction methods are applied to the multi-temporal SPOT imagery databases covering both the Taita Hills

application site and the Helsinki metropolitan region control site. Performance is assessed both on the ability of each methodology to accurately retrieve ρ_s and also to provide radiometric stability within the multi-temporal datasets. 6S is applied using modelled atmospheres and general estimates of atmospheric optical depth (AOD) at 0.55 μm for the Taita Hills dataset, but is applied using meteorological data for the Helsinki control dataset. Together, Papers I, II and III give a complete description and verification of HELM.

Paper IV examines mapping LULC and LULC changes for environmental monitoring at a landscape level in the Taita Hills, utilizing an object-oriented (OO) MSS/ORM classification procedure applied to the SPOT imagery. This enables the integration of the mapping process into the heart of a conceptual framework for landscape analysis based on the theory of hierarchical patch dynamics (HPD), within which the interaction of ecological processes operating at different scales can be understood. This represents a significant theoretical advancement over standard uni-scale per-pixel classification techniques. The OO and ML classifications are compared and assessed based mainly on ground reference data collected in the Taita Hills. As a preprocessing step before classification, the HELM atmospheric correction developed in Papers I, II and III is applied to the data, along with topographic corrections. To enable the multi-disciplinary utilization of the generated LULC data, classification is based on the FAO LCCS.

Paper V uses the universal soil loss equation (USLE) and the SPOT derived LULC data generated by the MSS/ORM classification procedure, as described in Paper IV, to model potential soil loss in the Taita Hills, and to determine the areas of highest potential soil loss risk. In accordance with working in regions of the developing world with limited ancillary data availability, and considering the ability of the procedures to be replicated locally using a minimum of resources, low cost and freely available GIS data were utilized. Local rainfall records of the Kenyan Meteorological Department were interpolated into a 20 m grid, a 20 m digital elevation model (DEM) was generated from Survey of Kenya 1:50,000 scale topographic paper maps, and the 1:1,000,000 scale Explanatory Soil Map of Kenya was digitized to derive a generalized digital soil map.

Paper VI utilizes ρ_s derived in Paper II and LULC maps generated by the MSS/ORM classification procedure applied in Paper IV, as well as calculated SPOT imagery texture measurements and other GIS derived geospatial data (topography, climate and distance), as predictor variables to model the distribution and abundance of human population within the Taita Hills. Dwelling units were interpreted and digitized from airborne image mosaics and presence-absence of dwelling units in a 100 m analysis grid was used as a response variable. Prediction models were created using a generalized regression analysis and spatial prediction (GRASP) method that utilizes the generalized additive model (GAM) regression technique. The population abundance modelling results were compared with two existing global population data sets: Global Population of the World Version 3 (GPWv3) and LandScan 2005. With local population growing quickly in the Taita Hills, but a census being taken only once every decade, the ability to adequately map population distribution and abundance solely from GIS and remote sensing based variables is of significant importance. Again, the methodology was developed considering the ability of the procedures to be replicated locally using a minimum of resources.

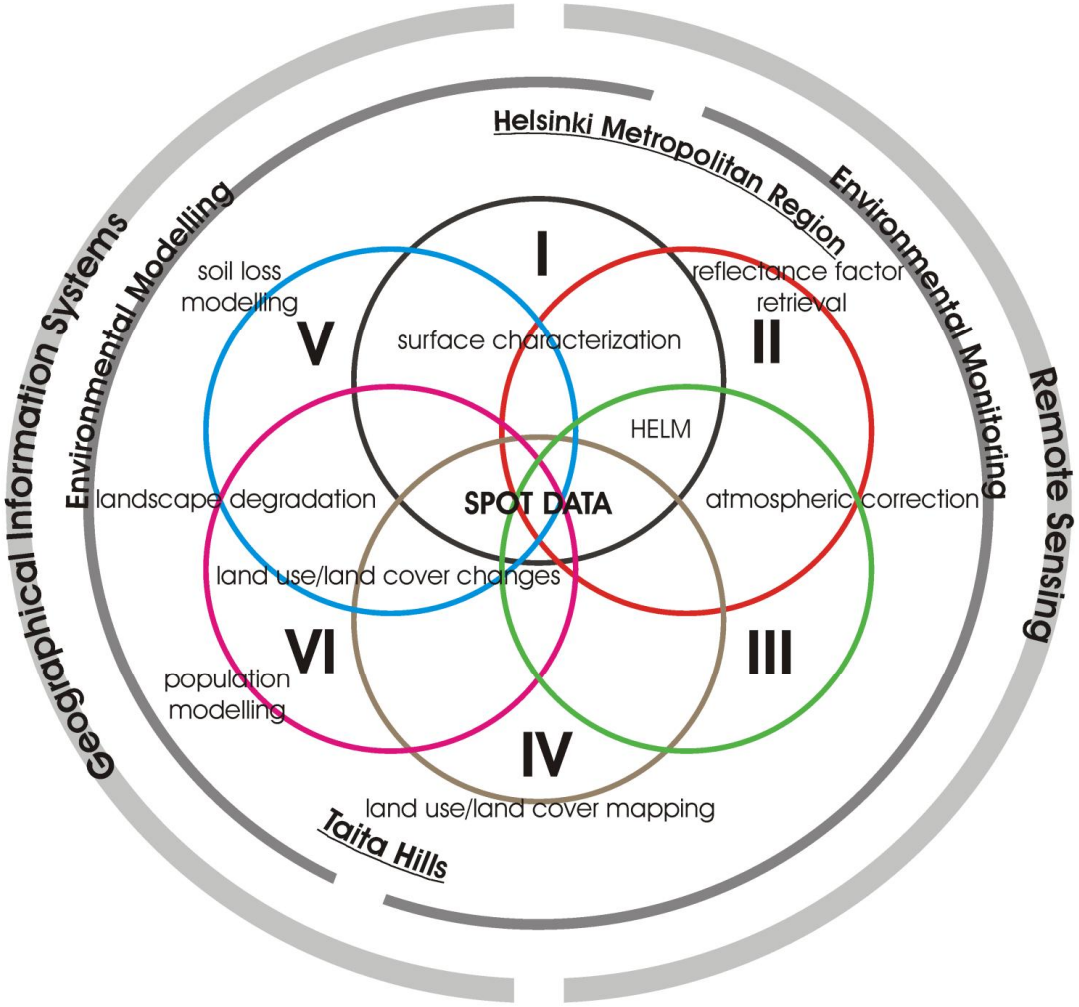


Figure 2. Conceptual synthesis of the environmental monitoring and modelling approaches, and geographical areas, applied in Papers I–VI, centred around the utilized SPOT multispectral satellite imagery data.

2 BACKGROUND

2.1 Key properties of optical remote sensing of the Earth's land surface

In general terms, remote sensing refers to obtaining information about an object or an area by utilizing electromagnetic radiation (EMR) without being in direct contact with the object or area (De Jong *et al.* 2004). In 1988 the American Society for Photogrammetry and Remote Sensing (ASPRS) adopted a formal definition of remote sensing as: "*the art, science, and technology of obtaining reliable information about physical objects and the environment, through the process of recording, measuring and interpreting imagery and digital representations of energy patterns derived from non-contact sensor systems*" (Colwell 1997). Remote sensing is a scientific discipline as the sensors used to make measurements are designed and built using physical principles, and the collected data are processed using mathematical and statistical

analyses (Jensen 2000). Remote sensing is, then, integrated with other mapping sciences, such as geoinformatics and GIS. However, remote sensing can also be considered an art as usually, at some point in the imagery interpretation process, skilled human judgement comes into play. For example, in the LULC data generated for the Taita Hills utilizing the MSS/ORM approach, as detailed in Paper IV of this thesis, the selection of training areas for the classifiers, the identification of appropriate segmentation parameters, the derivation of semantic rules for object classification, and the final manual editing of the maps to remove apparent errors, were all tasks undertaken directly by the remote sensing analyst. Consequently, the appropriateness and accuracy with which these tasks were achieved relied on the knowledge, understanding and skills of the analyst. The total experience, scientific knowledge and background understanding that the analyst has obtained throughout their lifetime are brought to bear on the interpretation process. Such learning cannot easily be measured or programmed, or completely understood (Jensen 2000), thus remote sensing is as much an art as it is a science.

The majority of remote sensing instruments measure changes in the properties and amount of EMR, although other force fields may also be utilized, for example sound waves in *sound navigation and ranging* (sonar), which is used underwater (Jensen 2000). Remote sensing devices are not in physical or intimate contact with the object of interest, but may be very close such as in the case of *in situ* field measurements of radiance made with spectrometers. Remote sensing for Earth observation most commonly utilizes airborne and especially spaceborne sensor platforms. In terrestrial remote sensing, sensors collect reflected, emitted and scattered EMR from visible to microwave wavelengths, with the aim of inferring information on the physical, chemical and biological conditions of the Earth's surface (Campbell 2002). Passive sensors, which form the majority of Earth observation remote sensing instruments, detect naturally occurring EMR that is emitted or reflected by the observed surface, with reflected EMR originating from the Sun being a very commonly utilized source of the radiation. Passive sensors include charge-coupled devices (CCD), radiometers, and traditionally aerial photography film. Active sensors, on the other hand, emit their own EMR in order to scan objects and areas whereupon the sensor then detects and measures the EMR reflected or backscattered from the target. Examples include radio detection and ranging (RADAR) and the increasingly popular light detection and ranging (LIDAR).

The focus of this thesis is on satellite based optical remote sensing in the visible to SWIR spectral regions, related to the spectral resolutions of the SPOT HRV, HRVIR, and HRG sensors. "Optical" remote sensing refers to wavelengths of EMR that can be manipulated by mirrors, lenses, and prisms, to reflect, refract, disperse, absorb, and polarize the EMR for data collection. The optical region in Earth observation is generally considered to range from visible to thermal infrared wavelengths, i.e. from 0.4–14 micrometres (μm) (Liang 2004), due to the restrictions imposed by atmospheric absorption. In terrestrial remote sensing the Earth's surface is the object of interest, from which perspective the atmosphere is usually considered to be an impediment which distorts the EMR signal received from the surface, rather than an object of investigation in its own right (Goetz *et al.* 1985). The atmosphere attenuates EMR both on its path from the Sun to the ground surface, and from the surface to the satellite sensor, though the effects are heavily wavelength dependent and also vary with changes in the atmospheric conditions at the time of image acquisition. Sensors designed to image the Earth's surface operate in specific spectral wavebands in which the atmospheric transmission is high, known as atmospheric

windows. This dissertation deals in detail with the removal of atmospheric effects in SPOT satellite imagery, covered in Papers I, II, and III.

EMR reaching the Earth's surface from the Sun is either reflected (scattered), transmitted or absorbed. The amount of reflected EMR varies as a function of wavelength, view and illumination directions (angles), time, polarization, and surface type, which enables inference of surface properties from the measured reflectance (Barnsley 1999). A basic axiom of remote sensing is that specific surface types or covers, such as water with varying levels of suspended impurities, soils of various types, or vegetation of various species, have an individual and characteristic manner of interacting with incident EMR. The spectral response (i.e. the wavelength variability) of the reflectance is traditionally the most utilized information source for remote sensing of land surfaces (see Figure 3 below).

Considering the spectral response of vegetation, which forms the major component of the Earth's land surface, a characteristic spectral curve for healthy green vegetation is shown in Figure 3. Vegetation typically has low reflectance in the visible spectrum (0.4–0.7 μm), particularly in the blue and red wavelengths, a very steep increase in reflectance around 0.7 μm, known as the "red edge", and high reflectance in the NIR (0.7–1.3 μm). The primary chemical and physical determinants of leaf optical properties are plant pigments, notably chlorophylls a and b, carotenoids and xanthophylls, and leaf mesophyll structure and water content (Tucker & Sellers 1986). Leaf reflectance in the visible wavelengths is controlled by the pigments, particularly chlorophyll, which has high absorptance and low reflectance and transmittance in the blue and far red portions of the visible spectrum. The absorption involves electronic transitions in the chlorophyll molecule centred on the magnesium component of the photoactive site (Goetz *et al.* 1985). Absorption in the blue wavelength region is also a consequence of electronic transitions in carotenoid pigments. In the NIR, the dominant feature of high leaf reflectance is associated with leaf cell structure and cellular arrangement within leaves, and hydration state (Slaton *et al.* 2001). As leaf structure is highly variable between plant species, NIR reflectance is useful in discriminating vegetation types. Further, plant stresses will alter reflectance in this spectral region (Lillesand & Kiefer 1994). Reflectance in the middle-infrared or SWIR regions (1.3–3.0 μm) is mainly dominated by the presence of water in the leaves. Generally speaking, leaf reflectance in the NIR is affected primarily by leaf structure, whereas reflectance in the visible is determined mostly by photosynthetic pigments, the transition between the two around 0.7 μm giving the red edge (Jensen 2000). The aggregations of plants and leaves that occur as vegetation cover on the Earth's surface are highly variable, and can form complex canopies, such as in forest covers, which have complicated and varying scattering properties. Multiple leaf layers in a canopy allow multiple scattering, and consequently NIR reflectance increases with a greater number of canopy layers (Lillesand & Kiefer 1994). Reflectance of vegetation cover is typically highly anisotropic, i.e. variable depending on the viewing and illumination direction (Sandmeier *et al.* 1998b), determined by the optical properties of the canopy components, canopy structure, and also the local topography (Asner *et al.* 1998).

In contrast, soil and other bare surfaces, such as the human built environment, display a less variable spectral response, which typically increases steadily across the visible and NIR (VIS/NIR) spectrum (Figure 3). Further, reflectance from water surfaces is more or less zero at wavelengths beyond the red (Tso & Mather 2001) (Figure 3). The reflectance of land surfaces can vary very substantially as a function of time. An obvious case is a field of crops as they develop from the leafing stage, through growth to maturity and, finally, to harvest and senescence. The topographic location of a target area, in terms of slope orientation, and the atmospheric conditions also effect reflectance.

Remote sensing makes it possible to collect data of inaccessible and extensive areas with information available in spectral, spatial, angular and temporal resolutions and polarization domains. Consideration of polarization, however, is outside the scope of this thesis. Optical satellites typically collect multispectral imagery, i.e. radiance is measured in a number of wavebands, which are described by the spectral sensitivity functions. The spectral resolution of a sensor refers to the number and bandwidth of the wavebands, whilst the radiometric resolution refers to the sensitivity of the sensor to detect variation in the received radiance. The range of the view and illumination angles over which satellite imagery acquisition occurs is determined by the satellite orbital characteristics and the sensor viewing geometry (Barnsley *et al*. 1994). The pixel size of a satellite image denotes the area on the ground covered by a single pixel. The spatial resolution, however, refers to the level of spatial detail that is discernable within an image (Aplin 2006), and is determined by the ground instantaneous field-of-view (GIFOV) and spatial response function of the sensor. The temporal resolution of a satellite refers to the average revisit period to a particular location on Earth (Aplin 2006). This is dependent upon various factors such as the swath width, orbital altitude and inclination, the sensor view angle and off-nadir viewing capabilities, and the latitude of the location itself. In the case of optical satellite imagery, the probability of acquiring cloud free imagery is also related to the latitude of the location as well as the temporal resolution of the satellite.

Remote sensing in Earth resource analysis can be applied in the physical, natural, and social sciences (Jensen 2000), for example in geography, soil science, geology, hydrology, urban planning, agriculture, and forestry. The greatest limitation of remote sensing is, perhaps, that its utility is often overestimated or overstated (Jenson 2000). There is a huge amount of ‘power’ embedded in digital mapping products, in the same way that there is with topographic maps (Wood & Fels 1992; Pickles 1995), and often end users will unquestioningly assume that all the contained information is correct and represents the “truth” (whatever that may be). By those outside of remote sensing circles there is, then, perhaps a propensity to see remote sensing and GIS as a *panacea* for providing *all* the required information to address and resolve spatially oriented problems. The use of standardized nomenclatures, generation and validation procedures, and detailed metadata will assist end users in comprehending the realities and usage limitations of remote sensing derived data products.

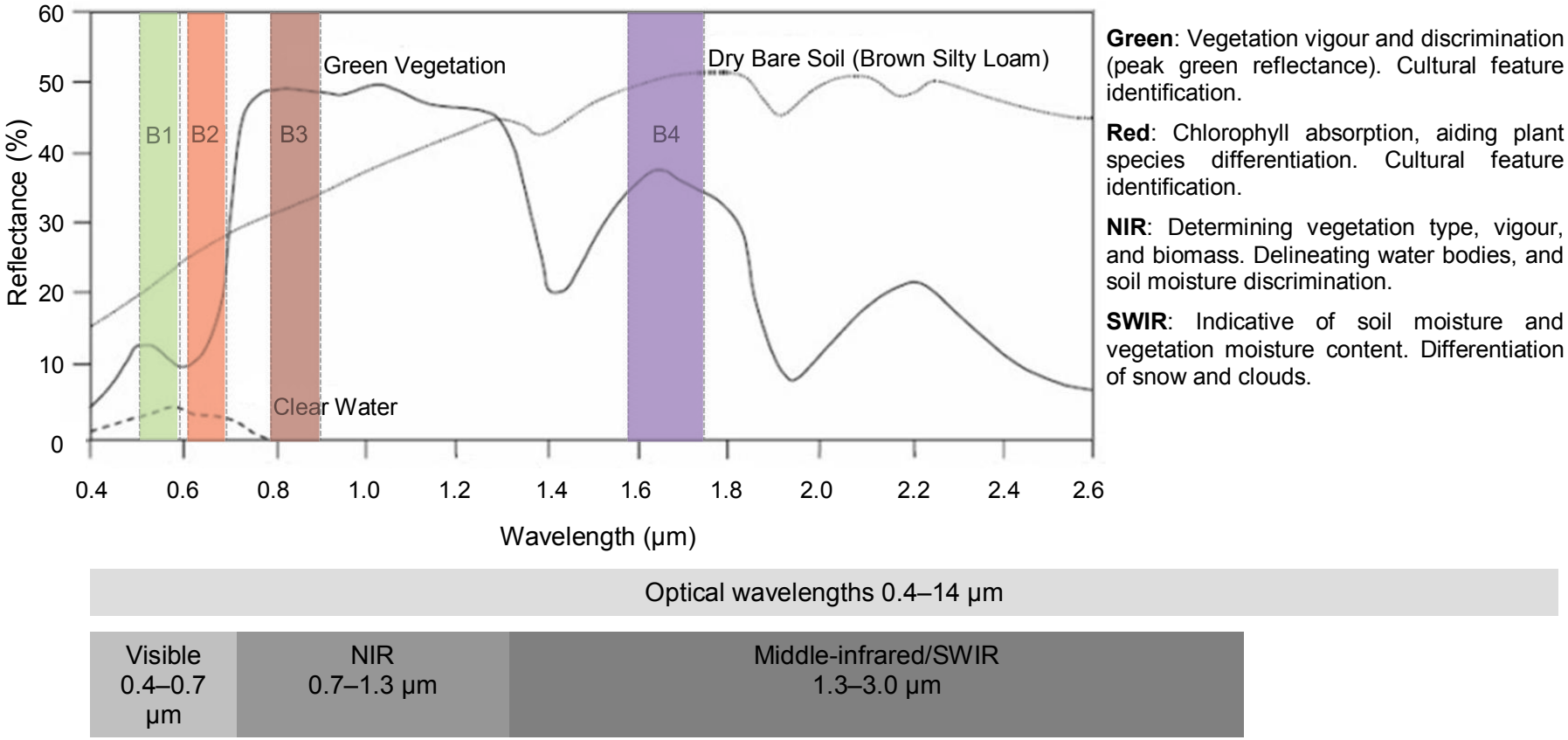


Figure 3. Characteristic reflectance spectrums of common Earth surface materials in the visible and near to middle infrared wavelengths. The full width at half maximum (FWHM) positions of the SPOT spectral green (B1), red (B2), NIR (B3) and SWIR (B4) bands are indicated along with the general usage applications of the spectral data for Earth surface remote sensing studies (adapted from Lillesand & Kiefer 1994).

Remote sensing of the Earth's land surface is generally considered to have commenced in 1859 when Gaspard Tournachon took an oblique aerial photograph of a village near Paris from a balloon (De Jong *et al.* 2004). Many years later, in July 1972, the United States of America successfully launched the Earth Resources Technology Satellite-1 (ERTS-1) into orbit and ushered in a revolutionary new era of civilian spaceborne Earth observation. ERTS-1 was, in fact, a Nimbus weather satellite modified to carry two types of sensor systems, but it proved able to provide high quality images from the four waveband multi-spectral scanner (MSS), which collected data in two visible and two NIR bands at a 68 m x 83 m spatial resolution over a 185 km by 185 km scene area, and three return beam vidicon (RBV) television cameras. Most importantly, these MSS images gave widespread worldwide scientific recognition of the capabilities and possibilities of remote sensing, and an appreciation of its significant value as a tool in environmental monitoring (Harper 1983). The main advantages recognized at the time were the global, repetitive, and multispectral coverage of the data, the low cost, and the lack of the previously pervasive and stifling political, security and copyright restrictions (Curran 1985). In January 1975 the programme name was changed to Landsat, and would prove to be of critical significance throughout the history of spaceborne environmental remote sensing, both in the high points and the low points of the programme. Up until the failure of the Landsat 7 ETM+ scan-line corrector (SLC) on the 31st May 2003, which has severely limited the data quality and usability of subsequent imagery, Landsat imagery had consistently been by far the most widely utilized satellite data source in the Earth sciences (De Jong *et al.* 2004).

Following the lead of the Landsat programme, many other countries decided to develop their own Earth observation missions. In 1978 the French government initiated the SPOT (Satellite Pour l'Observation de la Terre) programme, which included a series of, for that time, high-resolution optical imaging Earth observation satellites and ground control resources for satellite control and programming, image production, and distribution. The satellites were developed by the French Centre National d'Etudes Spatiales (CNES) in cooperation with the Belgian Scientific, Technical and Cultural Services (SSTC) and the Swedish National Space Board (SNSB) partner agencies. The SPOT programme was designed to improve the knowledge and management of the Earth's resources, but was from the start a commercial operation. In 1982 Spot Image, a public limited company created by CNES, IGN, and various private contractors, was established in Toulouse, France, to run the operation. However, data has always been obtainable for free or at low cost for scientific research purposes. For example, in June 2007 Spot Image launched the non-profit Planet Action initiative (www.planet-action.org), along with major geospatial industry partners such as ESA and ESRI, through which free imagery can be obtained for environmental monitoring projects with a core emphasis on climate change impacts, causes and solutions.

2.2 Overview of SPOT data

This thesis intentionally concentrates specifically on the utilization of medium resolution 20 m and 10 m multispectral SPOT multi-temporal imagery for landscape level environmental monitoring and modelling applications. Focusing solely on SPOT multispectral imagery allows for adequate consideration of the peculiarities of the data, most notably the ability of SPOT sensors to view off-nadir. It is therefore useful

to provide a brief overview of the characteristics of the SPOT sensors and the derived imagery data.

The SPOT series of satellites have been providing high quality consistent optical imaging of the Earth's surface since SPOT 1 was launched into an 822 km high near-polar sun-synchronous orbit on the 22nd February 1986. To date, five SPOT satellites have been successfully launched into orbit. SPOT 1 provided several innovations in civilian Earth resources remote sensing at that time; notably the use of two independent across track off-nadir pointable linear array 'pushbroom' sensors to collect 10 m spatial resolution panchromatic, and 20 m spatial resolution multispectral, imagery. SPOT 2 joined SPOT 1 in orbit on the 22nd January 1990 and SPOT 3 followed on the 26th September 1993 (Table 2). The design of these first three satellites was exactly the same; each of their payloads consisting of two identical High Resolution Visible (HRV) optical imaging instruments, magnetic tape data recorders, and a telemetry system for transmitting the images to the ground-based receiving stations.

Each HRV sensor could acquire images in either a panchromatic (P) mode, with a single wide band covering the visible part of the electromagnetic spectrum, or in a multispectral (XS) mode with three bands covering the green, red, and NIR parts of the spectrum (Table 2). The two HRV sensors could function in either XS or P mode, both simultaneously or independently. At nadir view, each HRV sensor imaged an area of 60 km x 60 km with a 3 km overlap, giving a total swath width of 117 km. This is greater than the maximum inter-track distance of 108 km occurring at the Equator, given the orbital characteristics of the SPOT satellites with an 822 km altitude, 98.7° inclination, 101.4 minute orbital period, and a local Equatorial crossing time of 10:30 a.m. in descending node. If HRV was a nadir view only sensor, then the revisit time for any given point on the Earth would be 26 days. This is obviously far too coarse a temporal resolution to allow for adequate phenological/land cover change detection studies or disaster mapping and monitoring, especially considering the issue of cloud cover. Consequently, the HRVs were designed so that the position of each entrance mirror could be tilted off-nadir across track by ground control through ± 27° in 0.6° increments. This allowed for the observation of regions of interest over a 950 km wide strip centred on the ground track, and not just those areas directly beneath the satellite. The off-nadir viewing capability also gave the HRV sensors the ability to acquire across track stereoscopic pairs of images that can be used in DEM generation.

An individual HRV image swath width therefore varied from 60 km at nadir view, to 80 km at maximum off-nadir view, and the temporal resolution of the satellite was increased from 26 days to a revisit interval of 1 to 4 days depending on the latitude. At 45° latitude a given area could be imaged on 11 occasions during the 26-day orbital cycle and, at the Equator, a given point observed 7 times. Above 40° N or S any point could be observed each day, whereas at the Equator a thin approximately 250 km wide strip (out of the 2,823 km separating the two adjacent orbital tracks) remained inaccessible on any given day.

SPOT 4 and SPOT 5 share the same orbital characteristics as SPOT 1, 2 and 3, and also carry a pair of imaging sensors with the same off-nadir viewing capabilities, but the design specifications in both cases are upgraded compared to the earlier satellites. SPOT 4 was launched on the 24th March 1998, the principal improvements being the addition of a 20 m resolution SWIR band B4 covering 1.58-1.75 μm and the onboard registration of all the spectral bands by replacing the panchromatic band with band B2 (0.61-0.68 μm) operating in both a 10 m resolution so-called monospectral (M) mode, as well as in a 20 m resolution multispectral mode (XI) (Table 2).

Improved knowledge of ground reflectance, acquired by SPOT 1 and 2, was used to introduce variable sensor gains according to landscape type and season, allowing better utilization of the 8-bit dynamic range. Greater onboard data storage capacity was also added and the design lifetime was increased from 3 to 5 years, although all the SPOT satellites have long outlived their design life. As a result of these modifications, the instruments were renamed High Resolution Visible Infrared (HRVIR) sensors.

SPOT 5 was launched on the 4th May 2002 and carries two High Resolution Geometric (HRG) instruments, which are a further development of the HRVIR sensors on SPOT 4. The spatial resolution of the bands in the multispectral mode (HI) has been increased to 10 m, although the SWIR B4 is actually imaged at 20 m and is supplied by Spot Image resampled to 10 m to match the spatial resolution of the other bands. The panchromatic band (HM) was changed back to a green-red bandwidth (0.48-0.71 μm), similar to the HRV P mode, but with an increased spatial resolution of 5 m. Furthermore, each HRG sensor can operate in a so-called Supermode where two panchromatic images are acquired simultaneously by two dedicated arrays of CCD detectors that are vertically and horizontally offset by half a pixel (2.5 m) in the focal plane. Utilizing a patented three-phase process, a 2.5 m resolution black and white image can then be generated from these 5 m images. Also, 5 m or 2.5 m resolution colour imagery can be derived operationally by combining the panchromatic information with simultaneously acquired 10 m multispectral (HI) data.

Additionally, SPOT 4 and SPOT 5 both also carry identical wide angle VEGETATION 1 and 2 sensors, respectively, designed for continental and global scale vegetation mapping and the production of daily and ten-day global synthesis products for utilization in monitoring long-term environmental changes (Duchemin & Maisongrande 2002). Furthermore, SPOT 5 also carries a High Resolution Stereoscopic (HRS) imaging instrument for acquiring simultaneous 10 m resolution panchromatic along track image stereopairs, specifically for DEM generation. However, description of these instruments is beyond the scope of this thesis because the research presented here concentrates on the utilization of multi-temporal images derived in the multispectral modes (XS, XI, and HI) of the HRV sensors on SPOT 1, 2, and 3, the HRVIR sensors on SPOT 4, and the HRG sensors on SPOT 5.

Currently only SPOT 4 and 5 are in orbit and fully operational. SPOT 2 ceased operations on 30th June 2009 and was deorbited one month later. SPOT 1 was withdrawn on the 31st December 1990 and it broke up in the atmosphere after deorbiting in 2003. SPOT 3 stopped working after it suffered a stabilization system failure on the 14th November 1997. The continuity of data collection in future will be ensured by the 8 m multispectral resolution SPOT 6 (Astroterra) satellite due for launch in 2012 and SPOT 7 planned for 2014. Furthermore, SPOT Image is involved in the PLEIADES program (http://smsc.cnes.fr/PLEIADES/), which in 2010 should see the launch of Pleiades 1 (PHR1A), forming the first of a constellation of new-generation mini-satellites that are smaller, cheaper, more agile, and offer better overall performance than the previous SPOT sensors, collecting visible and NIR (VIS/NIR) data at 2.8 m nadir resolution and with an off-nadir viewing capability of ± 45° in any direction. However, at such large off-nadir viewing angles the differences in reflectance to nadir for the observed Earth surface materials are very likely to be significant. This issue is investigated in detail in this thesis and in Paper I, considering the view angle range of the SPOT HRV/HRVIR/HRG sensors.

***Table 2**. SPOT satellite series technical data.*

	SPOT 1, 2 & 3	SPOT 4	SPOT 5
Instrument Payload	2 x HRV	2 x HRVIR [a]	2 x HRG [b]
Spectral Bands & Spatial Resolution (m)	Panchromatic (P) 10 m Multispectral (XS) 20 m	Panchromatic (M) 10 m Multispectral (XI) 20 m	Panchromatic (HM) 5 m Multispectral (HI) VIS/NIR 10 m, SWIR 20 m
Spectral Resolution (µm)	P: 0.50-0.73 B1 (Green): 0.50-0.59 B2 (Red): 0.61-0.68 B3 (NIR): 0.78-0.89	M: 0.61-0.68 B1 (Green): 0.50-0.59 B2 (Red): 0.61-0.68 B3 (NIR): 0.78-0.89 B4 (SWIR): 1.58-1.75	HM: 0.48-0.71 B1 (Green): 0.50-0.59 B2 (Red): 0.61-0.68 B3 (NIR): 0.78-0.89 B4 (SWIR): 1.58-1.75
Image Swath Width	60 km (nadir), 80 km at max off-nadir view	60 km (nadir), 80 km at max off-nadir view	60 km (nadir), 80 km at max off-nadir view
Radiometric Resolution	8 bit	8 bit	8 bit
Angle of Incidence	± 31.06°	± 31.06°	± 31.06°
Launch Date (Retirement Date)	SPOT 1: 22.02.1986 (Retired: 31.12.1990) SPOT 2: 22.01.1990 (Retired: 29.07.2009) SPOT 3: 26.09.1993 (Retired: 14.11.1997)	24.03.1998	04.05.2002
Revisit Capability	1 to 4 days, dependent on latitude	1 to 4 days, dependent on latitude	1 to 4 days, dependent on latitude

[a] Note that SPOT 4 also carries the VEGETATION 1 sensor.
[b] Note that SPOT 5 also carries the HRS and VEGETATION 2 sensors.

In the metadata provided with SPOT imagery, the off-nadir view of the sensor is defined by the angle of incidence (θ_V), which is the angle between the normal to the reference spheroid passing through the scene centre, and the instrument look direction for the same point (Spot Image Corp. 1997). The angle is denoted as right (R), or a negative value, when the sub-satellite point passes to the east of the scene centre, and left (L), or a positive value, when the sub-satellite point passes to the west of the scene centre, considering the solar illuminated descending orbit. Because of the curvature of the Earth, whilst the sensor view angle (θ_{VZ}) itself has a ± 27° angular range, θ_V has a ± 31.06° range. The scene orientation angle is defined as the clockwise angle between the centre line of the raw scene and the meridian passing through the centre of the raw scene (Spot Image Corp. 1997). The sensor azimuth (φ_V) was therefore calculated as the orientation angle plus 90° for right (negative) off-nadir views, and the orientation angle plus 270° for left (positive) off-nadir views. The relative azimuth (φ_r) between the sensor azimuth and the solar principal plane can be defined following the convention as outlined by Sandmeier *et al*. (1998a), where the solar position is taken as 180° azimuth. Therefore, 0° relates to a sensor view of the forward scattering and 180° to backscattering; 0° to 360° is clockwise from the forward scattering position (Figure 4). This coordinate system was also applied to the field measured multiangular ρ_s utilized throughout this thesis. Therefore, note that whilst negative θ_V relates to SPOT viewing from east of a scene centre, - θ_{VZ} relates to forward scattering viewing.

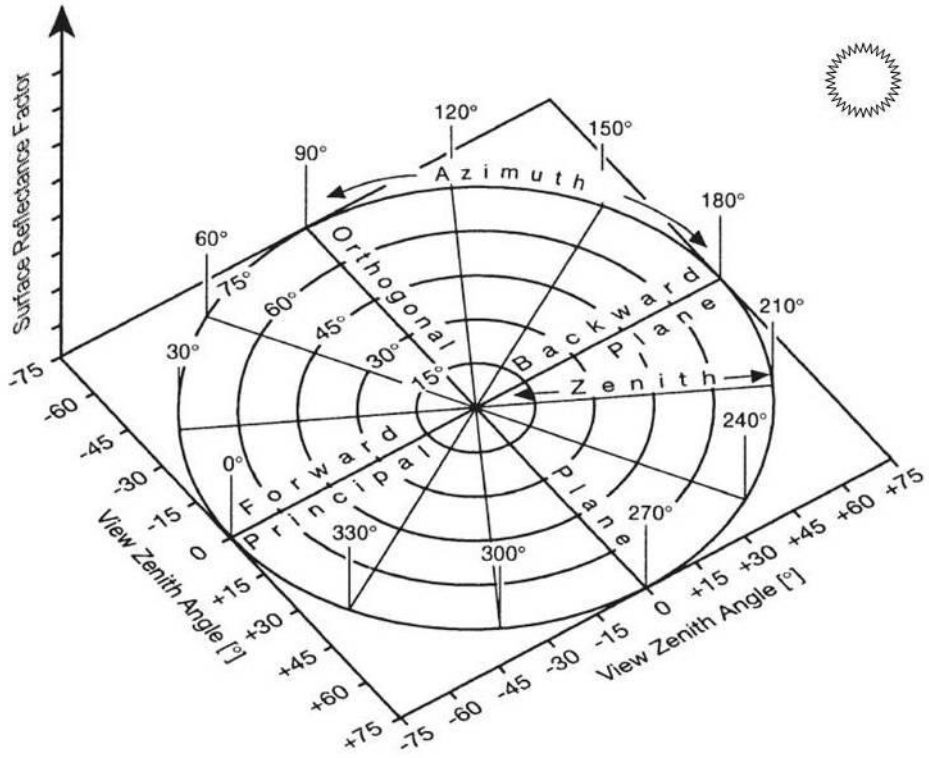


Figure 4. Coordinate system for multiangular surface reflectance factor (ρ_s) (after Sandmeier et al. 1998a).

The SPOT program has been designed to improve the knowledge and management of the Earth's resources and numerous research projects have utilized multi-temporal HRV, HRVIR or HRG multispectral imagery in various environmental monitoring studies. For example, mapping deforestation and forest degradation in the Brazilian Amazon (Alves & Skole 1996; Souza *et al*. 2003); determining African tropical forest structure characteristics (de Wasseige & Defourny 2002); detecting change in the wetlands of the US (Houhoulls & Mlchener 2000); mapping and modelling the spatial distribution of forest fires in Andhra Pradesh, India (Vadrevu *et al*. 2006); discriminating crops in the Saga Plains, Japan (Murakami *et al*. 2001); mapping and modelling urbanization in the Tunis Metropolitan Area (Weber & Puissant 2003); and local validation of the thematic accuracy of the International Geosphere-Biosphere Programme Data and Information System (IGBP-DIS) 1 km Global Land-Cover Data Set (Scepan 1999).

SPOT data also makes an important contribution in multi-temporal, multi-sensor environmental studies, especially following the SLC failure on Landsat 7 ETM+ in May 2003. SPOT imagery has been used, for example, in combination with Landsat data for vegetation change detection in the Brazilian Amazon (Lu *et al*. 2008), and for monitoring land use changes in the Hangzhou urban area in China (Deng *et al*. 2008); in combination with ASTER and Landsat imagery for change detection over a forested landscape in Canada (Wulder *et al*. 2008); in conjunction with IRS and Landsat data for Boreal forest leaf area index (LAI) estimation in central Finland (Stenberg *et al*. 2008); in data fusion with ERS Synthetic Aperture Radar (SAR) imagery and Landsat data for mapping land cover in a Mediterranean landscape (Chust *et al*. 2004); in combination with IKONOS data for inventorying landslides (Nichol & Wong 2005); and SPOT XS data has been integrated with panchromatic aerial photography for land cover change detection (Petit & Lambin 2001).

2.3 Utilization of SPOT data

As Song *et al.* (2001) note, by far the most common use of optical satellite imagery is in mapping LULC and LULC change over time through image classification and change detection methodologies. There are now more than 20 million archived SPOT scenes covering the time span since SPOT 1 was flown in 1986 available (source: www.cnes.fr/web/CNES-en/1419-a-wealth-of-applications.php) for use in multi-temporal studies, and the number of archived images grows continuously. Unfortunately, the raw digital numbers (DN) recorded in a satellite image are not an accurate measure of change over time because they are a function not only of surface conditions but also of the diurnally variable atmospheric conditions, the seasonally variable Earth-Sun distance, the solar zenith angle (θ_Z), the sensor view incidence angle (θ_V), and the sensor calibration (Moran *et al.* 2001). Sensor calibration can be achieved utilizing the gain coefficients supplied with the SPOT data that allow the conversion of the 8-bit DN into top of atmosphere at-satellite radiance (L_{SAT}), in units of Watts per square metre per steradian per micrometre (W m^{-2} sr^{-1} μm^{-1}), on a spectral band basis. It is then possible to convert these radiances into top of atmosphere at-satellite reflectance (ρ_{SAT}), which normalizes for variations due to the Earth-Sun distance and the solar zenith angle. This then leaves the contribution of the atmosphere and the effect of an off-nadir sensor view angle to be accounted for.

A substantial part of this thesis is concerned with the removal of atmospheric effects from SPOT imagery, dealt with specifically in Papers I, II and III, and also covered in Paper IV. Consequently, it is necessary to justify why such a large amount of effort was invested in determining the best methods for accurately retrieving p_s, and to consider in which circumstances and why absolute atmospheric correction should be applied, and indeed also to consider when it might be unnecessary. Further, it is also pertinent to give a brief general overview of the effects that the atmosphere has on optical satellite imagery, as well as on the SPOT spectral bands wavelengths specifically.

This thesis deliberately concentrates exclusively on the utilization of medium resolution multispectral SPOT multi-temporal imagery for landscape level environmental monitoring and modelling applications, allowing full consideration of the particularities of the data. However, it is acknowledged that in general it is much more useful to utilize optical satellite imagery from multi-sensor sources in environmental studies. This not only improves the chances of cloud-free image acquisition, but also increases the temporal, spectral, spatial, and possibly angular information that can be gathered for the challenging task of environmental monitoring. Furthermore, the failure of the Landsat 7 ETM+ SLC in May 2003, which has severely limited the data quality and usability of subsequent imagery, combined with the threat of a possible future data gap in the Landsat programme (Goetz 2007), clearly demonstrates the inherent pitfalls of relying too heavily on any one single satellite system. It is clear, therefore, that long-term landscape level environmental remote sensing studies will increasingly become based on data from multiple sensors, with medium resolution satellites other than the historically most heavily utilized Landsat series, such as SPOT, ASTER, IRS and CBERS, gaining greater importance.

Further, it is likely that an increased utilization of multi-scale image segmentation and classification approaches will lead to greater numbers of studies making better and more theoretically integrated use of multiple resolutions of imagery covering a particular landscape or ecosystem. Although there are significant financial costs involved in the use of SPOT data for commercial applications, imagery is

usually free or considerably cheaper for research purposes. The consistency of the imagery quality, the relatively long-term imagery achieve dating back to 1986, and the planned continuation of the programme, makes SPOT a viable source of medium resolution data for environmental monitoring applications

However, with the likely continuation of the trend for the satellite data providers to supply optical imagery processed to higher level specific reflectance products, there will be an increasing requirement for the integration of historical and newly acquired multi-sensor imagery into existing databases. All these factors are major drivers for the need to have optical satellite imagery pixel information processed to a common comparable physical unit of measurement, which is the *surface reflectance factor* (ρ_s). This can only be achieved through the application of reliable absolute atmospheric correction.

2.3.1 Atmospheric effects: in what circumstances are corrections necessary?

Setting aside, for a moment, the argument above that the long-term integration of optical imagery requires the application of an absolute atmospheric correction procedure, whether or not it is actually necessary to atmospherically correct a multi-temporal image dataset basically depends on the research application. As Song *et al*. (2001) point out, if every image is to be classified individually and the derived LULC maps used in post-classification change detection analysis, or if the change analysis is based on image differencing with no *a priori* assumption on the thresholds, then radiometric correction is unnecessary. If, however, the classification is to be based on a common set of derived training signatures, or the change detection is based on a methodology which assumes a zero mean for stable classes in the difference image, then radiometric normalization (i.e. a relative atmospheric correction) is needed. Relative atmospheric correction is an empirical approach based on an assumption of a linear relationship between image spectral bands across time, which can be determined by identifying spectrally stable and spatially well defined pseudo-invariant features (PIFs) present in all the scenes (Yuan & Elvidge 1996). For example, Schott *et al*. (1988) developed a method that estimates the slope and intercept of the assumed linear relationship from the mean and standard deviation of the PIFs' DN values, normalizing the different image dates together and accounting for the purturbative factors. Chavez & MacKinnon (1994) developed a hybrid method whereby a selected Landsat MSS scene from a multi-temporal dataset was calibrated to field measures of ρ_s, taken concurrently with the image acquisition, and the other images where then normalized to this radiometric master scene.

In general, therefore, for applications where a common radiometric scale is assumed among the multi-temporal imagery, then at least a relative correction is necessary. As Song *et al*. (2001) note, however, if multi-temporal imagery from multiple satellite sensors is to be used, a relative normalization approach is complicated and an absolute correction is better. Furthermore, if the classification training data or the change detection methodology are to be extended through space and time, i.e. utilized or compared with additional imagery from outside the original study area (for example in a multi-sensor, multi-resolution project), then a full absolute atmospheric correction becomes necessary. Moreover, if there is a requirement to derive any biophysical parameters, such as biomass or LAI, or vegetation indices, such as the Normalized Difference Vegetation Index (NDVI), and to compare them with derivations from other scene dates, sensors, or geographic areas, then the retrieval of ρ_s is mandatory. As noted by Moran *et al.* (2001), ρ_s has

become the basic measurement required for most remote sensing applications and models.

It can be argued, therefore, that the application of absolute atmospheric corrections is a prerequisite for quantitative remote sensing studies. Accurate retrieval of ρ_s from optical satellite imagery, such as SPOT data is, then, a highly desirable and important preprocessing step. Numerous approaches to absolute atmospheric correction have been developed, but fundamentally they all consist of two major steps (Liang *et al*. 2001): firstly, the estimation or determination of the various atmospheric parameters required for the proposed calculation methodology and, secondly, the prediction of ρ_s, or the process of *reflectance factor retrieval* (RFR) as it is also commonly known.

2.3.2 Retrieval of surface reflectance factor from SPOT multispectral data

Considering attenuation of electromagnetic radiation travelling through the atmosphere, due to the combined effects of scattering and absorption by gases and aerosols, the amount of radiant energy transmitted, relative to that for no atmosphere, is known as the atmospheric transmittance (T) and can be approximated as:

$$T = e^{-\tau / \cos(\theta)} \quad (1)$$

where τ is the atmospheric optical thickness, T can represent either T_V the atmospheric transmittance from ground target to sensor or T_Z the atmospheric transmittance from sun to ground target, and θ can be either θ_V or θ_Z respectively. This relationship holds for atmospheric transmittance with scattering and weak absorption when θ_V and θ_Z values are < 70° (Moran *et al.* 1992).

Scattering is typically the most dominant atmospheric effect in optical satellite imagery (Chavez 1988; Chavez 1989; Song *et al*. 2001) and adds brightness to the visible bands of the SPOT data, such that the recorded L_{SAT} of dark-objects will be much higher than their true surface values (Moran *et al*. 1992). Conversely, atmospheric absorption acts to subtract brightness from the longer SPOT NIR and SWIR wavelength bands (Moran *et al*. 1992). Consequently, it can be considered that scattering has an additive, and absorption a multiplicative, effect on the remotely sensed signal (Song *et al*. 2001).

Rayleigh (gaseous) scattering is the main type of scattering occurring in clear conditions (Chavez 1989), such as those at the time of cloud-free SPOT scene acquisition. It occurs throughout the visible spectrum, but its effect decreases rapidly with increasing wavelength (λ^{-4}) (Gao *et al.* 2009). Further, there is no blue band on the SPOT sensors, where Rayleigh scattering effects are greatest. Fortunately, Rayleigh scattering and absorption by ozone and oxygen can be well characterized because of the relative stability of the gaseous composition of the atmosphere (Liang *et al*. 2001). More problematic, however, is scattering and absorption due to tropospheric aerosols, because of the high temporal and spatial variability of atmospheric aerosol loadings (Kaufman 1993; Teillet & Fedosejevs 1995; Liang *et al*. 2001). Thus, this constitutes the greatest difficulty in the atmospheric correction of optical satellite imagery, such as SPOT data (Liang *et al*. 2001). Here the total amount of aerosol (Mie) scattering is greater than Rayleigh scattering and also affects longer wavelengths, as the effect of aerosol scattering decreases with increasing wavelength at a slower rate that Rayleigh scattering (typically λ^{-2} to λ^{-1}) (Gao et al. 2009). The

role of Mie scattering, however, only becomes more dominant in hazy and very hazy atmospheric conditions (Chavez 1989).

A useful generalization is that the NIR and SWIR wavelengths (> 0.7 μm) are largely free from atmospheric scattering (Moran *et al*. 2003), whilst the visible region (0.4 – 0.7 μm) is heavily affected. Atmospheric gaseous absorption is principally due to water, carbon dioxide, oxygen and ozone, and primarily affects the NIR and SWIR wavelengths (Vermote *et al*. 1997b). The SPOT spectral bands are deliberately placed in portions of the spectrum relatively unaffected by gaseous absorption. Nevertheless, absorption needs to be accounted for in radiometric correction (Chavez 1996).

The theoretical framework, definitions and nomenclature for reflectance quantities in optical remote sensing have been developed from those originally specified by Nicodemus *et al*. (1977). The physically based terminology outlined by Nicodemus *et al*. (1977) was adapted to the remote sensing case by Martonchik *et al*. (2000) and further extended and updated by Schaepman-Strub *et al*. (2006). The development of operational reflectance products from NASA's MODIS and MISR sensors has also contributed to the adoption of more uniform and physically consistent reflectance terminology within the remote sensing community (Schaepman-Strub *et al*. 2006). However, ambiguous or erroneous terminological usage remains, and description of the physical conditions of measurement are often lacking, even within the peer reviewed literature (Schaepman-Strub *et al*. 2006). The surface reflectance (ρ) is defined as the ratio between the radiant exitance (M [W m^{-2}]) and the irradiance (E [W m^{-2}]) at the surface within a specific spectrum. The *surface reflectance factor* (ρ_s) is the ratio of the radiant flux reflected by a surface to that reflected into the same reflected-beam geometry and wavelength interval by an ideal (i.e. non-absorbing and non-transmitting) Lambertian standard surface under identical conditions of illumination (Schaepman-Strub *et al*. 2006). This standard surface is commonly approximated in field measurement circumstances by a Spectralon® panel. Under typical field conditions viewed by spaceborne, airborne, or ground based sensors, with a GIFOV formed from a conical solid angle of observation, the ambient sky forms a hemisphere of illumination and radiation diffuses downward with variable intensity from different parts of the sky. This situation can be approximated by modelling the total irradiance as the superposition of both direct solar illumination, from a specific direction, and anisotropic diffuse irradiance (Martonchik *et al*. 2000; Schaepman-Strub *et al*. 2006). The ratio of direct to diffuse irradiance is a function of wavelength, with decreasing Rayleigh and (to a lesser extent) aerosol scattering with increasing wavelength, which strongly influences the spectral dependence of directional reflectance from the surface (Martonchik *et al*. 2000).

Typical observation circumstances are, then, most accurately described by what could be termed the "in-field" hemispherical-conical reflectance factor (HCRF), following the overall naming convention originally established by Nicodemus *et al*. (1977) that the angular characteristics of the illumination are mentioned first followed by the angular characteristics of the reflected radiance; the addition of the term "in-field" acknowledges that the hemispherical diffuse irradiance component is not isotropic like a theoretical HCRF. If the instantaneous field-of-view (IFOV) solid angle of observation of a sensor is very small, then the reflected radiance may be near constant over the full cone angle, *approximating* a directional reflectance; a *directional* reflectance meaning the conceptual scattering of a collimated beam of light into a specific direction within the hemisphere (Nicodemus *et al*. 1977), and which cannot be measured directly. The multiangular ρ_s field measurements taken for this thesis were made with the Finnish Geodetic Institute Field Goniospectrometer

(FIGIFIGO; Suomalainen *et al*. 2009; see Section 4.1.2 below), utilizing foreoptics with a 3° field-of-view (FOV). As noted by Suomalainen (2009), the difference between a conical and a directional reflectance factor can only be seen if there are significant second derivative angular effects at the scale of the optics opening angle. Away from the direct backscatter hotspot region, where it is not possible to measure in any case, real surfaces or samples tend to have quite smooth multiangular reflectance properties. Therefore, in such circumstances, the measured in-field HCRF can be assumed to be equivalent to the in-field hemispherical-directional reflectance factor (HDRF). However, the SPOT satellites have an IFOV with a full conical angle of 4.18°, which can consequently only be considered as an approximation of HDRF. Moreover, HDRF or HCRF depends on the angular distribution of the illumination and the proportion of the diffuse to direct irradiance, as well as the scattering properties of the surface itself. The amount and spectral distribution of diffuse irradiance is dependent on atmospheric conditions, local topography and the reflectance properties of the adjacent ground surface (Martonchik *et al*. 2000). Consequently, as Lyapustin & Privette (1999) note, multiangular ρ_s measurements made under ambient sky conditions show significant shape differences relative to the reflectance characteristics of the actual surface itself.

In an ideal situation and circumstances, therefore, the most desired reflectance quantity to be retrieved would be the bidirectional reflectance factor (BRF) because it is a function only of the intrinsic scattering properties of the surface and consequently describes the directional reflectance characteristics of the surface itself. The BRF also relates to the theoretical bidirectional reflectance distribution function (BRDF), which describes the scattering of a parallel beam of incident light from one specific direction in the hemisphere into another explicit direction in the hemisphere (Schaepman-Strub *et al*. 2006). Omitting the spectral dependence, the BRDF (f_r [sr^{-1}]) can be defined as (Schaepman-Strub *et al*. 2006):

$$f_r(\theta_i, \varphi_i; \theta_r, \varphi_r) = \frac{dL_r(\theta_i, \varphi_i; \theta_r, \varphi_r)}{dE_i(\theta_i, \varphi_i)} \quad (2)$$

where θ_i, φ_i; θ_r, φ_r are the zenith and azimuth angles of the direction of illumination and reflection, respectively; dE_i (W m^{-2}) is the irradiance from the illumination direction; and dL_r (W m^{-2} sr^{-1}) is the radiance reflected into the differential solid angle at θ_r, φ_r (Schaepman-Strub *et al*. 2006). Being a ratio of infinitesimal quantities, the BRDF is conceptual and cannot be measured, but it can be theoretically estimated with a level of uncertainty from HCRF and/or biconical measurements. The BRF is the ratio of the reflected radiant flux $d\Phi_r$ from the surface area dA in a specific direction to the reflected flux of the same dA from an ideal Lambertian surface $d\Phi_r^{LAM}$ irradiated under identical conditions and, unlike BRDF, it is a unitless quantity. Ignoring the spectral dependence, BRF can be denoted as (Schaepman-Strub *et al*. 2006):

$$\mathrm{BRF}(\theta_i, \varphi_i; \theta_r, \varphi_r) = \frac{d\Phi_r(\theta_i, \varphi_i; \theta_r, \varphi_r)}{d\Phi_r^{\mathrm{LAM}}(\theta_i, \varphi_i)} \quad (3)$$

The view zenith and azimuth angles are omitted from $d\Phi_r^{LAM}$ because there is no angular dependence for an ideal Lambertian surface, which reflects equally in all

viewing directions as thus has a BRDF of $1/\pi$. The BRF of a surface is therefore numerically equivalent to its BRDF multiplied by π (Schaepman-Strub *et al*. 2006):

$$\mathrm{BRF}(\theta_i,\varphi_i;\theta_r,\varphi_r) = \pi \cdot f_r(\theta_i,\varphi_i;\theta_r,\varphi_r) \quad (4)$$

Omitting the spectral dependence, the HDRF can be defined as (Schaepman-Strub *et al*. 2006):

$$\mathrm{HDRF}(\theta_i,\varphi_i,2\pi;\theta_r,\varphi_r) = \frac{d\Phi_r(\theta_i,\varphi_i,2\pi;\theta_r,\varphi_r)}{d\Phi_r^{\mathrm{LAM}}(\theta_i,\varphi_i,2\pi)} \quad (5)$$

where, in this instance, θ_i, φ_i is the solar position and 2π [sr] is irradiance of hemispherical extent. Further, again omitting the spectral dependence, the HCRF can be defined as (Schaepman-Strub *et al*. 2006):

$$\mathrm{HCRF}(\theta_i,\varphi_i,2\pi;\theta_r,\varphi_r,\omega_r) = \frac{\int_{\omega_r}\int_{2\pi} f_r(\theta_i,\varphi_i;\theta_r,\varphi_r)\cdot L_i(\theta_i,\varphi_i)\cdot d\Omega_i \cdot d\Omega_r}{(\Omega_r/\pi)\cdot\int_{2\pi} L_i(\theta_i,\varphi_i)\cdot d\Omega_i} \quad (6)$$

where ω_r is the conical solid angle of the sensor IFOV and Ω is the projected solid angle of the cone, the subscripts *i* and *r* denoting the incident and reflected radiance.

It can be seen, then, that remote sensing measurements do not coincide directly with conceptual bidirectional reflectance quantities, rather they are approximations which can be used as the basis for further diffuse irradiance corrections and modelling to estimate BRF, or the BRDF. Consequently BRDF estimation is also dependent on the model accuracy, as well as the quality of the measurement data. Moreover, however, BRF is *not* retrievable in any meaningful way from single observation angle satellite imagery, such as SPOT data. Nonetheless, by accounting for the atmospheric effects in SPOT imagery, it is possible to retrieve ρ_s described as the in-field HCRF, which is equivalent to spectroradiometric ρ_s measurements made at the surface; i.e., that is to say an *absolute* atmospheric correction is necessary.

Numerous approaches to absolute atmospheric correction of optical satellite imagery for RFR have been developed. The choice of the technique applied will depend mostly on the availability and quality of atmospheric or meteorological data coinciding with the image acquisition dates and geographic areas. Where there are detailed overpass concurrent measurements of atmospheric properties available, notably the atmospheric optical depth (AOD), it is possible to make full use of radiative transfer models (RTMs), such as MODTRAN (Berk *et al*. 2000) or 6S (Second Simulation of the Satellite Signal in the Solar Spectrum; Vermote *et al*. 1997b), which require such parameters as inputs for RFR. RTMs represent the most mathematically advanced absolute correction methodologies, and can output very accurate predictions of ρ_s. For example, Moran *et al*. (1992) found that using RTMs in conjunction with overpass concurrent measures of AOD and water vapour, ρ_s could be retrieved with an accuracy of ± 0.02 (absolute reflectance units) for the Landsat TM VIS/NIR bands.

However, it is expensive and logistically difficult to obtain detailed atmospheric measurements concurrent with the image acquisition, and it may be impossible to obtain such data for archived imagery. Good quality historical meteorological data and generalized assumptions about the atmospheric composition, based on modelled

atmospheres, allows for a wider application of RTMs. Nevertheless, in many areas of the world there is a paucity of meteorological data available that is detailed enough, and has an appropriate spatial and temporal frequency, to allow for the accurate application of RTMs to multi-temporal imagery datasets. Moreover, there are also many regions of the world where little or no meteorological data is available at all. Unfortunately, often these areas are also the places where the most rapid and significant changes in LULC are occurring, and where the need for accurate historical and up-to-date mapping and environmental monitoring is the greatest.

Such were the circumstances in the TAITA research project undertaken by the Department of Geosciences and Geography, at the University of Helsinki (www.helsinki.fi/science/taita/taitahills.html), which required the absolute atmospheric correction of multi-temporal SPOT imagery to study LULC changes in the ecologically important and environmentally sensitive Taita Hills in Kenya. A research issue that consequently arose was to determine the most appropriate procedure, which entailed an accuracy assessment and comparison of the various applicable methods, the results of which are detailed in this thesis.

Firstly, it is actually still possible to utilize an RTM for RFR by using inputs derived from model atmospheres and estimates of AOD. Several image-based methods have been suggested for predicting AOD based on the expected relationships between reflectance in the mid-infrared region around 2.2 μm and the blue and red spectrums over areas of dark dense vegetation; the DDV approach (Kaufman *et al*. 1997). Application of similar AOD retrieval techniques to SPOT data is *possible* (Lin *et al*. 2002) but, however, matters are severely complicated by the lack of an appropriate middle-infrared band or a blue band on the HRV/HRVIR/HRG sensors (see Table 2), and also by the off-nadir θ_V range (Lin *et al*. 2002). Consequently, in the study conducted for the TAITA project and detailed in this thesis, the AOD at 0.55 μm for the Taita Hills application study site was approximated as the average of the AREONET measurement for southern Africa, as reported by Remer *et al*. (2005), and used in conjunction with standard atmosphere and aerosol models; see Section 4.3.4 below.

Alternatively to RTMs, there are also simplified absolute atmospheric correction methodologies available that do not require *in situ* atmospheric/meteorological data, and which can be used operationally for RFR. For example, image-based corrections utilize only information derived entirely from the satellite scene itself. Estimates of the atmospheric path radiance (L_p) (i.e. the sensor recorded radiance contributed by the atmosphere itself) and downwelling diffuse irradiance can be made for each of the SPOT bands based on the recorded radiance for within-scene areas of assumed very low reflectance; such as areas of clear water or topographic shadow. Such approaches are general termed *dark-object subtraction* (DOS) techniques (Chavez 1988), as an additive scattering component is estimated for each band and subtracted from every pixel in the image. However, the atmospheric transmittance also has a multiplicative effect and further procedures, such as the COST method (Chavez 1996), have been developed to improve the accuracy of DOS corrections. The DOS approaches utilized in this thesis are detailed in Section 4.3.3 below.

Furthermore, where the opportunity exists to visit the study area, which is the case in most local or regional scale remote sensing projects (for example to collect LULC training and ground reference test data), and a spectrometer or goniometer is available for use or can be borrowed, then it is possible to make measurements of ρ_s in the field. So-called *empirical line* (EL) methods align L_{SAT} data to field measurements of the ρ_s of dark and bright spectrally stable within-scene calibration sites (Smith &

Milton 1999), utilizing a standard linear regression equation in the form $y = ax + b$; where a is the slope of the regression line, representing the atmospheric attenuation, and b is the intercept with the x-axis, representing the atmospheric path radiance (L_p) (see Figure 5). A separate correction is derived for each spectral band in the data.

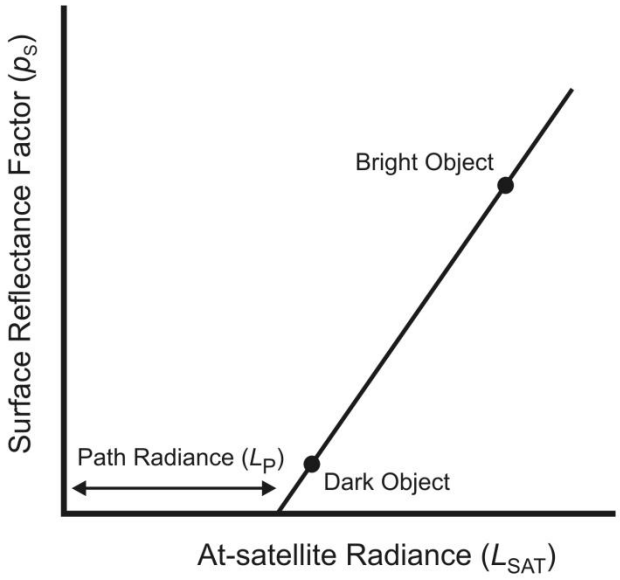


Figure 5. The derivation of a prediction equation from two-within scene calibration targets using the empirical line (EL) atmospheric correction method.

Previous researchers have successfully retrieved ρ_s from remotely sensed data utilizing EL approaches (e.g., Smith & Milton 1999; Perry *et al*. 2000; Moran *et al*. 2001; Karpouzli & Malthus 2003; Moran *et al*. 2003; Xu & Huang 2006). The two main assumptions are that (i) the atmosphere is approximately homogenous throughout the image area, and (ii) that there is a linear relationship between L_{SAT} and ρ_s. As noted by Chavez (1989), assuming atmospheric stability across a scene area could leave local errors due to non-homogeneity. However, as a SPOT scene is a relatively small 60 km by 60 km area, compared to the 180 km by 180 km Landsat TM scene area for example, this is unlikely to be a significant source of error unless there are clouds present, which indicates that there is localized spatial variability in the atmospheric properties. Furthermore, as Moran *et al*. (1990) note, although the relationship between L_{SAT} and ρ_s is quadratic for the full range of reflectance, it is sufficiently linear over the range 0–0.7 ρ_s to allow linear interpolation with negligible error. Few naturally occurring surface types have a reflectance greater than 70%.

Research efforts into EL methods have enabled the minimization of required field measurements. For example, in outlining their *refined empirical line* (REL) method for Landsat data, Moran *et al*. (2001) showed that, because of the near-linear relationship between L_{SAT} and ρ_s, an accurate estimation of the correction lines could be obtained using only two reflectance targets (Figure 5): firstly, detailed field measurements of ρ_s for *one* appropriate within-scene bright calibration target and, secondly, an estimate of L_{SAT} for $\rho_s = 0$ derived using an RTM and "reasonable" water and aerosol models, or measurements of atmospheric conditions on a typical cloud-free day. The *historical empirical line method* (HELM) proposed and detailed in this thesis is similar to the REL approach and is based on the derivation of a linear ρ_s prediction equation from two within-scene calibration targets. In HELM, however, the

estimate of path radiance (L_p) is derived from identified within-scene dark-objects (such as areas of clear water or topographic shadow) directly from the imagery, by making assumptions about their intrinsic reflectance (similar to DOS approaches), and therefore negates the requirement to utilize an RTM and estimate atmospheric parameters. Consequently, as a minimum, ρ_s field measurements of only one calibration target are required, although it is good practice to collect data for additional sites in order to validate and assess the accuracy of the correction lines. Moreover, as Smith & Milton (1999) note, if the calibration and validation targets are spectrally pseudo-invariant over time, which they should be if correctly chosen, then the measurement of ρ_s need not coincide with the image data acquisition.

The purpose of HELM is, therefore, to (re)construct the historical linear relationship between L_{SAT}, as recorded in the multi-temporal SPOT imagery, and ρ_s for the spectrally pseudo-invariant pixels (PIPs) as measured in the field. Compared to a nadir viewing only sensor, such as Landsat, matters with SPOT multispectral data are complicated by the off-nadir viewing capability of the HRV/HRVIR/HRG sensors, which gives a ± 31° possible range in θ_V of an off-nadir SPOT scene. Ideally the multiangular ρ_s behaviour of the calibration site should be measured in the field and modelled. This obviously entails the use of some kind of goniometer. However, given that any form of reliable multiangular ρ_s measurements will give useful information about the target, this need not be a very advanced expensive automated instrument, such as FIGOS (Sandmeier & Itten 1999), and could be a simple manually manoeuvrable device with a limited zenithal arc measurement range. For example, the goniometer utilized by Jackson *et al.* (1990) consisted of a 2.3 m long tubular aluminium arm affixed next to a metal protractor and guide, which had notches every 5° between –45° and +45° zenith into which the arm could be manually positioned. This enabled the measurement of 19 view angles that could be achieved over approximately a 5-minute time period. In this instance the spectrometer was held at the top of the arm, but it is just as practical to have the spectrometer at ground level and position the foreoptics at the top of the arm, connected to the spectrometer by fibre optic cable. It is within the budget and capabilities of most remote sensing groups to construct a simple goniometer.

Nevertheless, it is accepted that in some usage circumstances it may only be possible, for a variety of reasons, to collect nadir ρ_s measurements of the calibration target. For this thesis it was possible to utilize the Finnish Geodetic Institute Field Goniospectrometer (FIGIFIGO; Suomalainen *et al.* 2009). Consequently, FIGIFIGO was used to take detailed multiangular ρ_s measurements and investigate the effect of HELM calibration to nadir ρ_s on RFR accuracy. Calibration to nadir ρ_s is denoted as the HELM-1 approach, whilst calibration to ρ_s data modelling the exact illumination and view geometry of the SPOT imagery is termed the HELM-2 approach.

HELM was developed in response to the TAITA research project requirement to perform absolute atmospheric correction of multi-temporal SPOT imagery based on the limited ancillary data availability of working in the Taita Hills, in common with many areas of the of the developing world. Consequently, HELM is designed for use in absolute atmospheric correction in circumstances where there are no overpass concurrent atmospheric or detailed meteorological measurements available, but there is access to a spectrometer (and preferably a goniometer) and field access to the research site(s). For practical reasons, HELM is limited to regional scale and local landscape level studies, as opposed to countrywide or continental areas, because of the requirement for a calibration target to be found within each scene area and measured in the field. Because it is an empirical statistical method, HELM gives no

estimation of atmospheric parameters; rather, a direct prediction of ρ_s is derived for given input L_{SAT} values, and it is against these that the accuracy of HELM can be assessed. Further details of HELM are given in Section 4.3.2 below, and are summarized in Paper II.

Often assessments of atmospheric correction methodologies are based on improvements in the accuracy of derived parameters such as classifications and landscape metrics (Mahiny & Turner 2007), Jeffries-Matusita distances (Song *et al*. 2001), estimates of timber volume (Norjamäki & Tokola 2007), or the ability of the approaches to normalize the reflectance throughout multi-temporal images (Schroeder *et al*. 2006). Arguably, though, the best method is to compare RFR with ρ_s of targets as measured in the field. For SPOT data this does, however, necessitate detailed multiangular measurements of validation sites be made, which is not always possible. Therefore, whilst the Taita Hills provided an application site representing the typical circumstances of a region of the developing world with limited ancillary data availability, it was also necessary to have a control site, where full meteorological data records were available and where the multiangular reflectance behaviour of the ground targets could be measured. For this thesis study it was possible to utilize FIGIFIGO and the Helsinki metropolitan region was taken as the control site.

Based on the work of Moran *et al*. (1992), Schroeder *et al*. (2006) state that a benchmark for establishing successful absolute atmospheric correction of optical satellite imagery in the VIS/NIR bands is an overall accuracy of ± 0.02; i.e. 2% absolute reflectance units. Liang *et al*. (2002) were able to atmospherically correct Landsat ETM+ data with an absolute RMSE of 0.009 to 0.015 ρ_s in the visible bands and 0.027 to 0.041 ρ_s in the NIR, based on a comparison with spectrometer field measurements. This represented a relative error of ~10% throughout the VIS/NIR bands, because of the larger absolute reflectance values in the NIR. Using the REL approach, Moran *et al*. (2001) were able to retrieve ρ_s with a mean absolute difference (MAD) of ≤ 0.01 for the TM/ETM+ VIS/NIR bands, assessed relative to spectrometer field measurements. Sellers *et al*. (1995) identified the accuracy level required for reflectance inputs into, and/or validation data for, land surface-atmosphere models and climate models as ± 0.02 reflectance units. It should be noted, however, that climate models usually require broadband albedo, and consequently the ρ_s derived from the SPOT data would not be a suitable input. Nonetheless, the stated required accuracy levels are indicative of what can be considered as a useable reflectance product. ± 0.01-0.02 absolute ρ_s accuracy was referenced as being achievable in the VIS/NIR bands of Landsat and SPOT data using RTMs and overpass concurrent radiosonde profiles (Hall *et al*. 1992), although the TM SWIR bands were found to be problematic with ± 0.06 absolute accuracy (Hall *et al*. 1991). Based on these previous studies, therefore, an absolute atmospheric correction methodology should be expected to achieve VIS/NIR RFR within ± 0.02 ρ_s absolute accuracy, derive SWIR absolute accuracy better than the ρ_{SAT} estimates, and achieve 10% overall relative accuracy in all spectral bands, in order to be considered an effective and usable atmospheric correction methodology.

3. STUDY AREAS

This thesis focuses on the application of multi-temporal SPOT multispectral data to environmental monitoring and modelling at a landscape level in the Taita Hills, Kenya. This is representative of utilizing medium resolution optical satellite imagery for environmental studies in regions of the developing world with limited ancillary data availability. Further, the Helsinki metropolitan region was taken as a control site where full meteorological data records were available and where the FIGIFIGO goniometer could be used to measure the multiangular ρ_s behaviour of ground targets. This enabled the investigation of HELM calibration errors, and also an assessment of the RFR accuracy of both the proposed HELM technique, and a comparative assessment with other absolute atmospheric correction methodologies that are applicable in the intended HELM usage circumstances.

3.1 Taita Hills, Kenya

The Taita Hills are located in the Taita-Taveta District of southeast Kenya at latitude 3º 25´ S, longitude 38º 20´ E (Figure 6). The hills cover an area of approximately 1000 km^2 and are surrounded by the semi-arid Acacia/Commiphora shrubland and dry savannah of the Serengeti Plains, some of which falls within sections of Tsavo National Park. The area of mapping utilized in Paper IV covers 89 220 ha from 3° 31′ 27″ S to 3° 16′ 46.5″ S and from 38° 14′ 21.6″ E to 38° 22′ 11″ E (Figure 6 (a)). Whilst the surrounding plains are at an elevation of 600-700 m a.s.l., the Taita Hills rise abruptly in a series of ridges with the highest peak of Vuria at 2208 m, although the average elevation of the hills is 1 500 m. The hills are steep due to the hardness of quartzite caps overlaying softer metamorphic rocks, including gneisses. For the population modelling in Paper VI, the elevation threshold to separate the hills from the lowlands was set to 1 100 m a.s.l., based on the gradient change from the plains to foothills, resulting in a study area of 326 km².

The climate of this region is influenced by the Inter-Tropical Convergence Zone (ITCZ) which leads to a bi-modal rainfall incidence, with a longer rainy season during March-May/June and short rains in October-December. However, the annual variability of precipitation is high, especially in the semi-arid shrubland surrounding the hills. According to records from the Kenya Meteorological Institute (KMI), average annual precipitation in the years 1986 to 2003 was 1 132 mm at the Mgange station, located in the hills at 1 768 m a.s.l., and 587 mm at the Voi station, located in the lowlands at 560 m a.s.l. (Figure 6 (b)). Annual variation was strong, as yearly maximums at Mgange reached ca. 2 000 mm and minimums were as low as 700 mm. The variation between rainy seasons was also strong. In Mgange, the short rains in October-December in the years 1986 to 2005 resulted in precipitation between 161 and 1 352 mm, with a mean of 462 mm (KMI).

In addition to the measured rainfall, mist and cloud precipitation usually occur throughout the year and provide a significant moisture component to the areas over 1 500 m a.s.l. Despite that the Taita Hills lie approximately 150 km inland from the coast, orographic rainfall plays an important part in the local climate as the hills form the first significant barrier which moisture laden air from the Indian Ocean encounters. Due to the orographic rainfall pattern, the southern and eastern slopes receive more precipitation than northern or western slopes, and a ‘rain shadow’ effect is discernable especially on the northwestern side of the hills, with the distinctive

Euphorbia candelabrum and more commonly *Euphorbia bussei* var. *kibwezensis* growing in the drier conditions. The annual mean average temperature ranges from 20ºC in the hills to 25ºC in the lowlands (Jaetzold & Schmidt 1983).

The agroecological zones follow closely the precipitation zones. In general, the farming in the hills is intensive small-scale subsistence farming. In the lower highland zone and in the upper midland zone over 1 400 m with annual rainfall on average 1 100 mm, the typical crops are maize, beans, cassava, peas, potatoes, cabbages, tomatoes, cassava, mango and banana (Jaetzold & Schmidt 1983, Soini 2005a). On the slopes and lower parts of the hills with average annual rainfall between 600 and 900 mm, early maturing maize species and sorghum and millet species are cultivated. In the lower midland zones with average rainfall between 500 and 700 mm, dryland maize types and onions are cultivated, among others. The two growing seasons, totalling to 150-170 days, coincide with the long and the short rains (Jaetzold & Schmidt 1983). The land is prepared during the dry season, and the crops are seeded prior to the short rains and long rains. Harvesting takes place after the end of the rainy seasons (Gachimbi *et al*. 2005). Maize is the most characteristic crop in the landscape and is seeded in August-September and harvested in January-February, and seeded again in March and harvested in June-July.

Terracing is widespread on the slopes, although the steepest slopes are unsuitable for crop growing. In addition to terracing, soil and water conservation measures include contour farming, grass banks, trash banks, planting of agro-forestry trees, mulching, and crop rotation (De Bie 2005), indicating that the local population in the hills are generally aware of the problem of soil erosion. Furthermore, local studies have also found awareness of these issues amongst the farmers and shown that they take the overexploitation of natural resources and falling productivity seriously; see, for example, the Participatory Rapid Appraisal (PRA) carried out by the Kenya Agricultural Research Institute (KARI) in the Wusi sub-location of the Taita Hills (Lekasi *et al*. 2005). Nevertheless, despite the soil conservation measures, visible soil erosion (rill, sheet and gully erosion) is evident in the croplands (Gachimbi *et al*. 2005).

The Taita Hills form the northernmost part of Africa's Eastern Arc Mountains, which run from southern Tanzania (Lovett & Wasser 1993) and boast an extremely high diversity of flora and fauna and high levels of endemism (Githiru & Lens 2007), and have consequently been identified by Conservation International as one of the top ten biodiversity hotspots in the world. Of particular scientific and conservation interest are the indigenous forest patches that are home to many rare or endangered endemic animals and plants. Today, only a small amount of native forest remains, occurring in a scatter of three larger hilltop remnants; Mbololo (c. 179 ha), Ngangao (c. 136 ha) and Chawia (c. 94 ha) as reported by Lens *et al*. 2002, from a field survey (see Figure 6 (b)), and further much smaller fragments embedded in a mosaic of human settlements, small-holder cultivation plots (known locally as 'shambas'), and plantations of exotic tree species such as *Cupressus lusitanica, Pinus* spp., *Eucalyptus* spp., and *Grevillea robusta* that were established in the hills between the 1950s and 1980s (Pellikka *et al*. 2009). The indigenous forest cover has been termed upland moist or mist forest by Beentje & Ndiang'ui (1988), but is also referred to as montane forest or cloud forest by other workers. It is mapped in Papers IV and Paper V as closed canopy broadleaved forest (see Table 8 below). The characteristic tree species include *Newtonia buchananii, Tabernaemontana stapfiana, Macaranga conglomerata, Albizia gummifera, Phoenix reclinata, Strombosia scheffleri, Cola greenwayi, Podocarpus* spp., *Ochna holstii,* and *Millettia oblate* (Beentje &

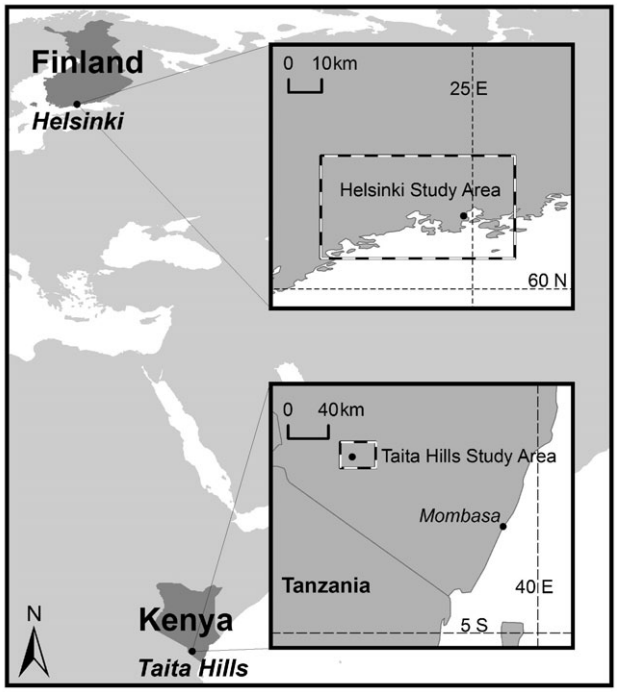
Ndiang'ui 1988). These indigenous forest patches also play an important role in both capturing additional moisture and storing the precipitation on the hilltops.

The population of the whole Taita-Taveta district has grown from 90 000 (1962) to 246 671 inhabitants in 1999 (Republic of Kenya 2001), with a 1.8% annual growth rate between 1989 and 1999 (Msagha 2004). Population is concentrated in the fertile Taita Hills, Sagala Hills and the town of Taveta close to Mt. Kilimanjaro, and in trading centres, such as Voi and Mwatate. The population within the Taita Hills is concentrated in the best agricultural areas close to fertile river valleys in areas receiving rainfall more than 1 000 mm annually. The district capital Wundanyi is the only town in the hills with ca. 4 500 inhabitants in Wundanyi sub-location in 1999, while the rural area consists of villages of varying size. The population density in the best agricultural areas is between 400 to 500 inhabitants per km^2, but some sub-locations have a density of more than 900 persons according to the 1999 Kenya census. The sub-locations consisting of areas in the hills and the lowlands have a density between 100 to 200 people per km^2, while the rural lowlands have a density between 5 to 30 people. The population in Wundanyi, Mwatate, Tausa and Mwambirwa divisions was 135 000 based on the 1999 census, in an area of 3000 km^2, resulting in an average population density of 253 per km^2. The least populated areas are the western and northern plains and northwestern parts of the Taita Hills due to the rain shadow effect. Most people in the hills live on small farms at elevations between 1300 and 1800 m a.s.l., where the annual long-term rainfall varies between 360 mm to 1935 mm and the mean daytime temperature is 20 °C. The average size of a farm is 2 ha (Soini 2005a; Ruotsalainen 2008).

The population growth has been a driving factor behind rising environmental pressure on the Taita Hills. Consequent intensification and increases in the area under cultivation for subsistence farming has led to a scarcity of available land in the hills, which has contributed to clearance of new agricultural land in the lowlands and a movement of population towards the lowlands, giving dynamic changes in the land use patterns. Due to poor agricultural management, erodible soils and the large relative height differences in the hills, the foothills especially are subject to escalating land degradation (KARI, 2005) in the form of accelerated soil erosion, loss of natural vegetation cover, and lowering of water tables (Pellikka *et al*. 2005).

Identified threats to the remnant indigenous forest patches include encroachment (for settlement, agriculture and livestock grazing), over extraction of firewood and building materials, charcoal burning, poor enforcement of government policies and regulations, illegal logging, lack of awareness among the communities living adjacent to forests, fires (both deliberate and naturally occurring) and colonization by suppressive and fast growing exotic tree species (EAWLS 2005). There is, therefore, a pressing requirement for environmental monitoring and modelling in the Taita Hills to allow for planning for the sustainable use of natural resources. The production of a digital geospatial database for the Taita Hills, and the mapping and analysis of LULC changes over time, have been major aims of the TAITA research project undertaken at the Department of Geosciences and Geography at the University of Helsinki.

(a)



(b)

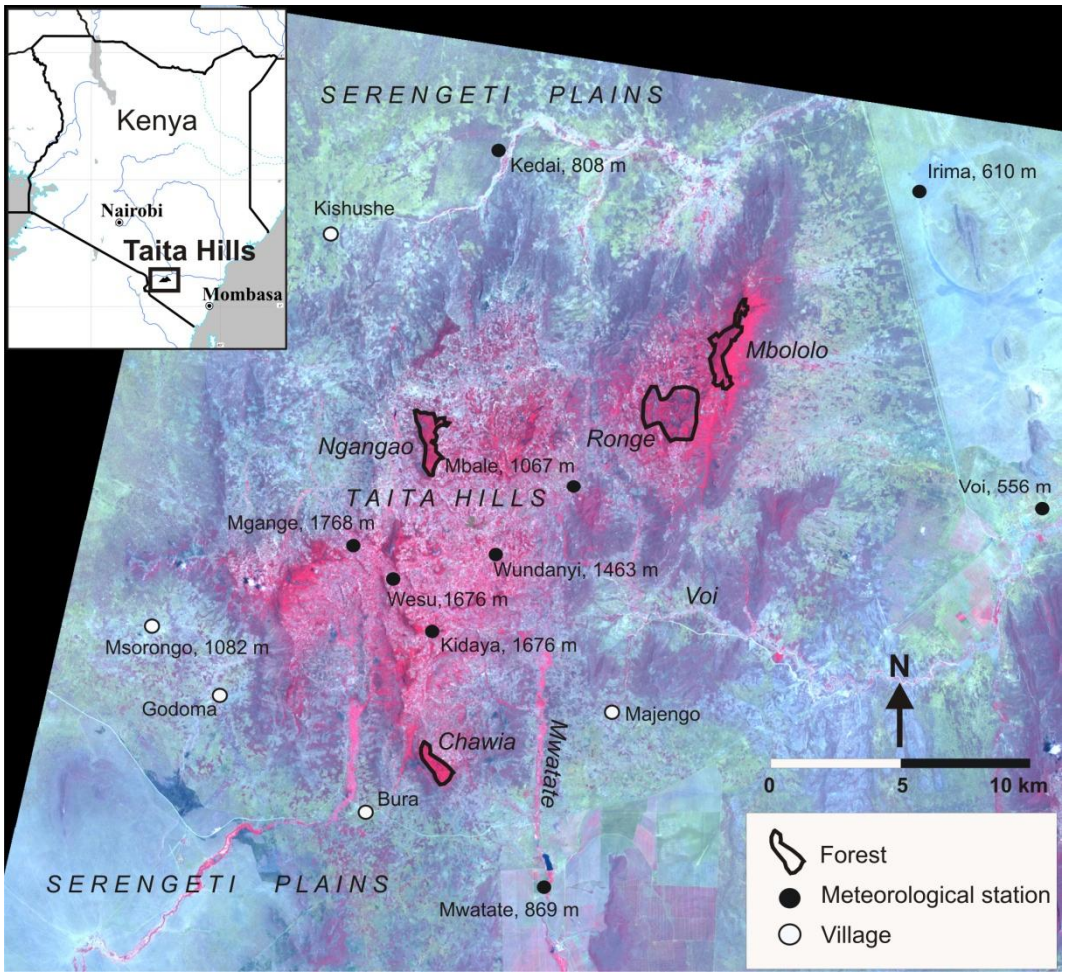


Figure 6. (a) Locations and extents of the Helsinki metropolitan region, Finland, and the Taita Hills, Kenya, study areas of Papers I-VI, and (b) Taita Hills as depicted in the 15.10.2003 SPOT 4 HRVIR 1 scene, with the main indigenous forest patches, discussed in Paper IV, denoted.

3.2 The Helsinki metropolitan region control site, Finland

The Helsinki metropolitan region in Finland was chosen as the control site for assessing HELM as it provided field access from both the University and the Finnish Geodetic Institute for utilising the FIGIFIGO goniometer to measure the angular ρ_s behaviour of the calibration and validation ground targets. Additionally, detailed meteorological data records were available. The Helsinki metropolitan region is situated at approximately 60° 12' N, 24° 56' E, along the Baltic coastline of southern Finland (Figure 6 (a)). The area is topographically flat, with a near sea-level elevation, and has extensive urbanization and development as well as agricultural areas, lakes and forests. Being located along the coastline, the climate is consequently transitional between maritime and continental and, due mainly to the Gulf Stream, is not as cold as other regions at the same latitude north, such as southern Greenland or Siberia. Average temperatures fall below 0°C from December until March and the spring arrives at the start of April (Finnish Meteorological Institute: www.fmi.fi). Only in May is there much green vegetation present and summer lasts until mid-September. SPOT scene selection for the Helsinki control dataset was consequently limited to images acquired during the northern hemisphere late spring and summer period because of the requirement for snow free data.

4. MATERIALS AND METHODS

4.1 Data

The following sections describe the details of the datasets utilized in this thesis, and the procedures applied in their generation and processing, whilst the preprocessing of the SPOT satellite imagery is covered separately in Section 4.2. Table 3 summarises the remote sensing and *in situ* datasets used in the dissertation and the corresponding research papers.

Table 3. *In situ, airborne and satellite remote sensing datasets used in the original Papers.*

Data	Paper
SPOT satellite imagery, Taita Hills, Kenya	II; III; IV; V; VI
SPOT satellite imagery, Helsinki metropolitan region, Finland	I; II; III
Airborne remote sensing: colour digital imagery mosaics covering parts of Taita Hills	IV; VI
Field measured nadir HCRF data, Taita Hills, Kenya	I; II; III
FIGIFIGO field measured multiangular HDRF, Helsinki metropolitan region, Finland	I; III
In situ land use/land cover training and ground reference test data, Taita Hills, Kenya	IV; V

4.1.1 SPOT HRV, HRVIR and HRG satellite imagery

Whilst the objective of the image selection for the Taita Hills application site was to choose dry season cloud free SPOT scenes most suitable for LULC mapping, the objective for the control dataset was to provide as wide a range as possible in the scene sensor and solar geometries; particularly in the scene incidence angle (θ_V),

which is the key variable of interest for verifying HELM. For reference purposes within this thesis, the Helsinki metropolitan region control SPOT imagery dataset is referred to as Dataset 1 and the Taita Hills application site imagery dataset is referred to as Dataset 2. As detailed in Table 4, for the control Dataset 1 five available multi-temporal scenes dating from 1993 through 2005, and seasonally from the 8th May to the 13th July throughout the late spring and summer, were chosen. These provided a range of off-nadir θ_V from L29.7° to R28.7° (which represents near-maximum off-nadir views in both directions), a range of solar zenith angles (θ_Z) from 37.5° to 43.8°, a range of solar azimuth angles (φ_Z) from 161° to 182.7°, and a variation in the local overpass time from 12:26 for SPOT 5 through 13:26 for SPOT 3. 1993 to 2005 was also a long enough time span to cover significant changes in the LULC in the area, in this instance through the process of urbanization, and thus provided a further challenge to the application of HELM in terms of the identification of a persistent ground calibration target.

***Table 4**. Details of the SPOT satellite imagery in the study Datasets 1 and 2.*

Dataset 1. Control Site: Helsinki Metropolitan Region, Finland, Lat 60º 12' N, Long 24º 56' E								
Image Date	Local Time [a]	Path/ Row	SPOT Sensor	θ_V	φ_V [b]	θ_Z	φ_Z	φ_r [c]
11.07.1993	13:01:29	073/226	1 HRV 1	L 9.3	289.7	38.2	170.6	299.1
08.05.1994	13:25:56	073/226	3 HRV 2	L 29.7	294.4	43.5	182.7	291.7
11.06.2002 [d]	12:53:58	073/226	4 HRVIR 1	L 2.5	288.3	37.5	170.1	298.2
10.05.2003	12:26:06	073/226	5 HRG 1	R 28.7	103.4	43.8	161.8	121.6
13.07.2005	12:34:18	073/226	5 HRG 2	R 16.1	106.2	39.2	161.1	125.2
Dataset 2. Application Site: Taita Hills, Kenya, Lat 3º 25' S, Long 38º 20' E								
Image Date	Local Time [a]	Path/ Row	SPOT Sensor	θ_V	φ_V [b]	θ_Z	φ_Z	φ_r [c]
01.07.1987	10:48:32	143/357	1 HRV 1	R 10.4	98.9	36.3	41.4	237.5
25.03.1992	10:49:52	142/357	2 HRV 1	R 13.8	98.9	26.5	79.0	199.9
25.03.1992	10:49:51	143/357	2 HRV 2	R 9.3	98.9	26.0	78.7	200.2
12.02.1999	10:57:14	143/357	4 HRVIR 2	L 4.2	278.7	27.7	113.7	345.0
02.06.2002	11:03:19	142/357	4 HRVIR 1	L 20.2	278.6	32.4	36.9	61.7
15.10.2003	10:49:36	143/357	4 HRVIR 1	R 10.4	98.8	21.0	104.3	174.5

θ_Z is solar zenith angle; φ_Z is solar azimuth angle; θ_V is sensor view incidence angle; φ_V is sensor azimuth angle, and φ_r is the relative azimuth between sensor azimuth and the solar principal plane.

[a] Helsinki images are UTC +3 hours (accounting for daylight saving time [DST] during summertime period); DST is not used in Kenya, so Taita Hills images are all local times UTC +3 hours.

[b] Sensor azimuth is calculated as the orientation angle + 90 for right (negative, east) off-nadir view and the orientation angle + 270 for left (positive, west) off-nadir view.

[c] The relative azimuth between sensor azimuth and the solar principal plane follows the convention of Sandmeier *et al*. (1998a); the sun is taken as 180° azimuth. Therefore 0° relates to a sensor view of the forward scattering and 180° to backscattering; 0° to 360° is clockwise.

[d] The 2002 image suffers saturation in the green, red, and SWIR bands, and calibration error in the red band.

As detailed in Table 4, the application Dataset 2 consisted of five multi-temporal scenes dating from 1987 through 2003, and from the 12th February to the 15th October. Despite being acquired at various times of the year and falling at the end or start of potential rainy seasons, all these images have similar dry moisture and phenological conditions, as no significant rainfall occurred in the month of (or before) scene capture in each instance, according to the local rainfall records of the Kenyan Meteorological Department. This reflects the unreliability of rainfall in this region of east Africa. There is a range of off-nadir θ_V from L20.2° to R13.8°, a range of θ_Z from

21° to 36.3°, a range of φ_Z from 36.9° to 113.7°, and a variation in the local overpass time from 10:48 through to 11:03.

As an application site for HELM, the Taita Hills represent a challenge because the rural agricultural and savannah make-up of the area offers very limited homogeneous manmade or vegetation free areas that could be taken as calibration targets. There is also an absence of any significant areas of clear standing water; the only major water body in the area being a heavily silted reservoir, that has varying sediment load (and therefore reflectance properties) and water-levels throughout the year. This area may, then, present difficulties for the identification of a dark-object for the HELM and DOS methods.

As can be seen from Table 4 and Figure 7, in the imagery in Dataset 1, because of the northerly latitude of Helsinki, the sun remains relatively low in the sky (smallest θ_Z is 37.5°) and in the south and eastern side of the hemisphere. The exception is the 1994 image where, due to the later overpass time (because θ_V = L29.7°), the sun just makes it into the southwestern sector of the sky (φ_Z = 182.7°). Note that the geometrical situation would be similar at high latitudes in the southern hemisphere, although here the sun would remain in the northeastern sector of the sky and the local solar time would be earlier. In the 2005 and 2003 images the sensor was viewing the backscattering, but approximately 55° off from the solar principal plane. For the other years, the sensor was viewing the forward scattering, but approximately 65° off from the principal plane. It is well established that ρ_s anisotropy effects are usually most pronounced along the principal plane (Sandmeier *et al*. 1998b; Sandmeier & Itten 1999), so multiangular effects are likely to be weaker at off-principal plane relative azimuth angles (φ_r), such as in Dataset 1. Countering this, however, a lower sun position in the sky enhances multiangular effects relative to a smaller θ_Z along the same plane.

Conversely, as can be seen from Table 4 and Figure 7, the difference of being in the Tropics is clear in the Dataset 2 imagery as, although the sun remains to the east of the scene centre because of the mid-morning overpass time of SPOT satellites at the Equator, the sun is higher in the sky and falls both in the north and southern half of the hemisphere, depending on the time of year. Consequently, a similarity or coincidence of the sensor azimuth angle and the principal plane can occur and multiangular effects in the imagery are likely to be more prominent. Indeed, in the 2003 image the sensor was viewing the backscattering approximately 5° off from the principal plane whilst in the 1999 scene the sensor was viewing the forward scattering approximately 15° off from the principal plane. In the 1992 image the sensor was viewing the backscattering 20° from the principal plane, and in the 1987 and 2002 scenes the sensor backscattering view was 57.5° and 61.7° from the principal plane, respectively. Because of the different latitudinal locations, and the differences in overpass times, there is no overlap in the θ_Z or φ_Z between Dataset 1 and Dataset 2. The two sites together, therefore, provide and allow for a thorough assessment of HELM in handling various geometrical circumstances.

The Taita Hills SPOT data was acquired through the TAITA project as part of the CNES ISIS program. The Helsinki SPOT data was provided by Spot Image under OASIS research grant 51. All the images were supplied as Level 2A scenes, with radiometric correction of distortions due to variations in the sensors' CCD sensitivity and geometric corrections for systematic effects (such as the Earth's rotation and curvature and variations in the satellites' orbital altitude), as well as rectification to a UTM/WGS-84 projection system based on ephemeris data (i.e. without ground control points). The stated average accuracy of the rectification for SPOT 1 to 4

scenes was only 350m (30m for SPOT 5 images), and consequently the imagery required further geometric processing before it was useable; see Section 4.2.1 below.

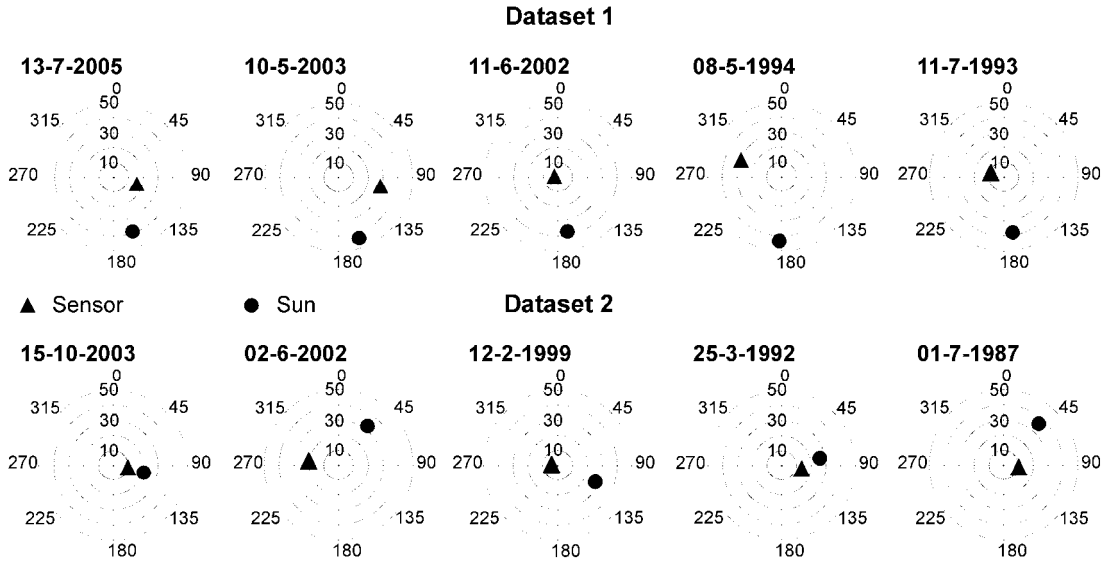


Figure 7. Polar plots of the solar and sensor positions for the SPOT scenes in Datasets 1 and 2.

4.1.2 Field measurements of surface reflectance factor

For the Helsinki metropolitan region control site study area, FIGIFIGO (Suomalainen *et al.* 2009) was used to take detailed multiangular daylight ρ_s measurements of different surface types as candidates for HELM calibration/validation sites. Following preliminary identification in Google Earth™, eight sites were visited in the field and deemed suitable for data collection and utilization (Table 5). The major constraints in data acquisition were the limited northern hemisphere spring/summer measurement window and the requirement for sustained cloud-free weather conditions, as well as other usage requirements for the goniometer itself in various different summer acquisition campaigns around Finland. This made planning and execution of a coherent sampling strategy very difficult and consequently data was collected over a period of 3 summers from 2005 to 2007. As shown in Table 5, the measured surface types included sand, gravel, asphalt, artificial turf, and managed real turf.

FIGIFIGO utilizes an ASD FieldSpec® Pro FR spectrometer with a spectral range from 350 – 2500 nm and an output spectral resolution of 1 nm, but a true observation resolution of 3 nm at 700 nm, and 10 nm at 1400 and 2100 nm. In order to enable the approximation of HDRF, foreoptics with a 3° FOV were mounted on the goniometer measurement arm, giving a 10 cm diameter GIFOV at nadir. FIGIFIGO automatically measures zenith angles (θ_{VZ}) up to ± 70° with either a 2.5° or 5° interval. Self-shadowing was present only over an area of 5° diameter around the exact backscattering direction. The full hemispherical view range in azimuth was achieved by rotating the whole instrument around the target in 10° to 30° steps. The view geometry was defined with a digital high-precision inclinometer and compass, and the solar position was calculated using GPS time signal and coordinates. To allow for the compensation for any variation in the amount of illumination occurring during measurement, incident irradiance was continuously monitored with a pyranometer. Measurement of 3 to 6 zenith arcs between 0° and 90° azimuth from the principal

plane was considered to be sufficient for description of the target ρ_s, and therefore full hemispherical measurement took 15 to 45 minutes depending on the desired level of sampling. FIGIFIGO ρ_s have an estimated relative accuracy of 1 – 5% depending on wavelength, sample properties, and measurement conditions (Suomalainen *et al*. 2009). For each target a BRDF model based on the Lommel-Seeliger law (Suomalainen 2006; Hapke 1993) was fitted to the FIGIFIGO ρ_s measurements to allow interpolation and extrapolation of the data.

Table 5*. Site details of the field data collected in the Helsinki metropolitan region control site study area for utilization in HELM.*

#	Site Name	Site Location	Surface Type	Measurement Dates	Site Usage
1	Hietsu Beach	60° 10' 28" N 24° 54' 28" E	Medium sand [a]	13.09.2005 17.07.2006 08.06.2007	Calibration/ Validation
2	Vermo Car Park	60° 12' 52" N 24° 50' 22" E	Asphalt	20.05.2005 25.05.2005 05.07.2005	Validation
3	Pasila Sports Ground	60° 12' 29" N 24° 56' 39" E	Very fine gravel [a]	07.06.2005 13.09.2005	Validation
4	Pasila Velodrome	60° 12' 10" N 24° 56' 35" E	Artificial turf	04.06.2007 07.06.2007	Validation
5	Malmi Airfield	60° 15' 00" N 25° 02' 41" E	Asphalt	04.06.2007	Validation
6	Töölö Soccer Pitch Turf	60° 10' 28" N 24° 54' 28" E	Managed grass	19.07.2006 05.06.2007	Validation
7	Kumpula Sports Ground	60° 12' 50" N 24° 58' 09" E	Very coarse sand [a]	13.09.2005	Validation
8	Pasila Baseball Ground	60° 12' 23" N 24° 56' 31" E	Very coarse sand [a]	07.06.2005 13.09.2005	Validation

[a] Aggregate classes as described by the Wentworth scale

For the Taita Hills application site, following inspection of the Dataset 2 SPOT imagery and field visits in the Taita Hills themselves, only five locations were deemed suitable for ρ_s measurement, and it was clear that only the roadside quarry (see Table 6) could be taken as a HELM calibration target. This was ~60 m wide and 200 m long and formed of light-grey calcareous coarse sand, underlain by bare rock and some angular pebbles of the same material. In order to include some validation data, it was necessary to utilize sub-optimal locations. As can be seen in Table 6, nadir ρ_s measurements were made of a sandy school playground, a compacted red soil road area, and an area of asphalt hard-standing, all of which were ~60 m by 60 m in extent. At each site GPS was used to record the location of the target centre so its position could be accurately determined within the imagery. Differential GPS was not employed, as the appropriate equipment was unavailable during this particular field trip, but the device was left averaging until the reported estimated accuracy was ≤ 5 m. Across each of the sites the nadir ρ_s was measured 15 times, each recorded measurement itself being the average of 15 spectrum samples, utilizing an ASD

FieldSpec® Handheld VNIR (325 – 1075 nm, 3.5 nm spectral resolution) spectrometer. At the start of each measurement set the spectrometer was calibrated to a Spectralon® white reference panel. No pyranometer was available to allow corrections for variation in illumination conditions during measurement. The spectrometer was handheld in a nadir view position at 1.2 m height facing towards the sun, with a 25° bare-head optic giving a GIFOV of 53 cm in diameter. Consequently, the measured ρ_s was the nadir "in-field" HCRF.

***Table 6**. Details of the field data collected in the Taita Hills application site study area.*

#	Site Name	Site Location	Surface Type	Measurement Dates	Site Usage
1	Roadside Quarry	03° 30' 19" S 38° 15' 46" E	Calcareous bare rock/ /Pebbles/Sand	25.01.2005	Calibration/ Validation
2	Roadside Hard Standing	03° 23' 52" N 38° 32' 56" E	Asphalt	27.01.2005	Validation
3	Bare Red Soil	03° 29' 50" S 38° 19' 25" E	Bare lateritic soil	27.01.2005	Validation
4	School Playground Sand	03° 21' 01" S 38° 20' 29" E	Medium sand	26.01.2005	Validation

4.1.3 Field data for classification training and accuracy assessment

During field visits to the Taita Hills in January 2005 and 2006 both training areas for the LULC classes to be classified (see Section 4.6.1 and Table 8 below), and ground reference test data to enable accuracy assessment of the 2003 SPOT scene classifications, were obtained. The ground reference test data were collected using stratified random road sampling (points falling in areas visually identified to have changed land cover relative to the 2003 image were discarded and regenerated), and from 0.5 m resolution true-colour digital aerial photography (see Section 4.1.4 below) flown in January 2004, 3 months after the SPOT scene acquisition, using stratified random sampling. The photography is limited to 8 mosaic areas covering ca. 12% of the Taita Hills, with three large mosaics covering ca. 30% of the upland area specifically (see Paper VI: Figure 1), and whilst the road sampling extended into the lowlands, less reference points were collected in the field because of logistical and financial constraints. A minimum statistically valid class sample size of 60 was calculated based on the multinomial distribution approach outlined by Plourde & Congalton (2003). Lesser numbers of points were collected for the spatially limited classes, such as Water, and more for the spatially extensive classes, like Cropland. The ephemeral Burned Area and Cloud/Shadow classes could not be sampled, and Bare Rock was not assessed as it could not be automatically mapped.

4.1.4 Airborne remote sensing data

A NIKON D1X true colour (RGB) digital camera, equipped with a 14 mm lens producing a 78 degree opening angle, was mounted on a light aircraft and used to acquire airborne imagery with 60% overlap and 30% sidelap between flight-lines. The

camera was utilized as part of the EnsoMOSAIC system, consisting of flight planning software, navigation software, a triggering unit, GPS and a power source (Holm *et al.* 1999). Data acquisition took place between 8 and 9 am on January 25 2004, and between 12 and 1 pm on January 27 2004, at altitudes between 2100 and 2700 m above the land surface, resulting in an approximate ground resolution varying between 27 and 40 cm for the study area. The imagery was processed on a frame-by-frame basis to remove brightness variations due to light-falloff and bi-directional effects, using the correction methods developed by Pellikka (1998). The frames were then mosaicked together using the EnsoMOSAIC software (Holm *et al.* 1999). The resulting mosaics were orthorectified, projected to a Transverse Mercator projection with a Clarke 1880 spheroid and Arc 1960 datum, and resampled to 0.5 m ground resolution. The geometric accuracy of the mosaics was verified during field work to be ± 2 m, using GPS Trimble® GeoXT™ GPS with differential correction base (reference) data.

4.1.5 Taita Hills Digital Elevation Model (DEM)

The digital elevation model (DEM) for the Taita Hills study area was derived from nine adjacent scanned Survey of Kenya 1:50 000 scale topographic maps, from which the 50-feet (~15 m) contours were automatically captured and then merged together in ArcGIS. These contours were interpolated into a 20 m planimetric resolution raster DEM utilizing the TOPOGRID method based upon the ANUDEM program (Hutchinson 1989). The method applies a discretised thin plate spline technique, in which the roughness penalty has been modified to allow the fitted DEM to follow abrupt changes in relief, such as streams and ridges, which is useful in rugged terrain. The heights were converted from feet to metres to be compatible with other metric measurements. Based on a digitisation accuracy of ± 1 mm and a general estimate of combined source mapping and interpolation error of half the contour interval, the DEM planimetric accuracy was estimated as ± 50 m and the altimetric accuracy as ± 8 m. Further, as spot height information on the scanmaps was not utilised in the interpolation process, these heights were used to assess the altimetric root mean square error (RMSE) of the DEM, which was also ~8 m.

4.1.6 Taita Hills rainfall geospatial data layer

Mean monthly precipitation surfaces were produced from monthly rainfall records of 11 stations within the Taita Hills from 1987 to 2005 obtained from the Kenya Meteorological Institute (KMI); see Paper V Table 4. Only eight years of rainfall data were used due to the poor availability of records of monthly precipitation before the year 1987. Mean annual precipitation was interpolated on to a 20 m resolution grid using the procedures from the ANUSPLIN package (Hutchinson 1991), which generates good results in areas of rugged topography (Cole & Arundel 2004), such as the Taita Hills. This technique uses tri-variate functions of longitude, latitude, and elevation in kilometres to fit thin plate spline functions, which can be viewed as a generalization of standard multi-variate linear regression. The degree of smoothness of the fitted function is usually determined automatically from the data by minimizing a measure of predictive error of the fitted surface given by the generalized cross

validation (Craven & Wahba 1979). The DEM was then used to generate rainfall layers in with a 20 m resolution.

4.1.7 Data for human population modelling in the Taita Hills

The Taita Hills geospatial database, created as part of the TAITA and TAITATOO projects, was used as the main data source for the human population occurrence and abundance modelling in Paper VI. Dwelling units were mapped using on-screen digitization from airborne digital camera data acquired during January 2004 covering 30% of Taita Hills highlands, as described in Section 4.1.4 above. The 20 m DEM, described in Section 4.1.5 above, was used to calculate the mean elevation, slope and aspect. The topographical wetness index (ω) was derived utilizing a custom-made ArcGIS geoprocessing model. In addition, irradiance (kWh/m^2/month) was calculated from the DEM using an ARC/INFO AML macro (shortwavc.aml) (Kumar *et al.* 1997; Zimmermann 2000). The precipitation grid layers, as described in Section 4.1.6 above, were also utilized in Paper VI. Vector map layers for main roads and rivers were digitized from the Kenya 1:50 000 scale topographic maps, and were used for the Euclidean distance grid calculations, undertaken in ArcGIS. Two existing global population datasets, Gridded Population of the World (GPWv3) at 5 km resolution and LandScan 2005 (Dobson *et al.* 2000) at 1 km resolution, and Kenyan 1999 census data (Republic of Kenya 2001) were used for the human population abundance model comparison.

The remote sensing predictor variables derived from the 15.10.2003 SPOT 4 HRVIR 1 20 m satellite image covering the Taita Hills consisted of (i) the mapped land cover, as also utilized in Papers IV and V and detailed in Section 3.6.1 below, (ii) ρ_s retrieved using HELM, as also utilized directly in Papers II and III, and (iii) image texture measurements made from the finally processed HELM atmospherically corrected and topographically corrected scene (see Section 4.5 below). Only four of the mapped LULC classes (Croplands, Thicket, Woodland and Plantation Forest) were used due to the relatively low prevalence or high correlation with other predictors of the other land cover classes. The percentage of spatial coverage for different land cover classes in each 100 m analysis square was calculated using the *summarize* function in ArcGIS. The first-order image statistics were the mean ρ_s of SPOT bands 2 and 3 (red and NIR). The green band 1 ρ_s was excluded from the modelling because of high correlation with the red band ($r > 0.95$). For the second-order image texture measurements based on the Gray Level Co-occurrence Matrix (GLCM), the angular second moment, contrast, correlation, sum of squares variance, inverse difference moment and entropy were calculated, as they were the most relevant texture measures according to Baraldi & Parmiggiani (1995). Three different sizes of moving windows were tested using the GRASS GIS freeware, 3×3, 7×7 and 15×15, with the result that the 3×3 window gave a ”salt and pepper” effect and the 15×15 window gave very strong smoothing effect. Consequently, the 7×7 window size was employed (as did Shaban & Dikshit 2001). The texture measures were calculated in four directions (0°, 45°, 90° and 135°) and averaged, as suggested by Haralick *et al.* (1973). The GLCM second-order image texture measures were calculated for both the red and NIR bands but, due to high correlation, only texture measures for the red band were used in the final regression analysis. Furthermore, all the red band second-order image texture measures had strong correlation with each other, with the exception of the angular second moment (ASM) and correlation

(COR) ($r < 0.1$); see Equations 7 & 8 below. Therefore, only these two measures were utilized in the final regression analysis. To summarize these image texture measures in each 100 m analysis grid square, the mean of pixel values from the texture images were calculated.

$$\text{ASM} = \sum_{i=0}^{quant_K} \sum_{j=0}^{quant_K} h_c(i, j)^2 \tag{7}$$

$$\text{COR} = \sum_{i=0}^{quant_K} \sum_{j=0}^{quant_K} \frac{(i-\mu)(j-\mu)h_c(i, j)^2}{\sigma^2} \tag{8}$$

where $quant_k$ = quantization level of band k (e.g., for SPOT imagery 2^8 = 0 to 255); $h_c(i, j)$ = the (i, j)th entry in one of the angular spatial-dependency matrices, and:

$$\mu = \sum_{i=0}^{quant_K} \sum_{j=0}^{quant_K} i \times h_c(i, j) \tag{9}$$

$$\sigma^2 = \sum_{i=0}^{quant_K} \sum_{j=0}^{quant_K} (i-\mu)^2 \times h_c(i, j) \tag{10}$$

4.2 SPOT data preprocessing

4.2.1 Geometric correction

The first step in processing the Level 2A SPOT imagery was to make the multi-temporal scenes spatially comparable through geometric correction. The images in the Helsinki metropolitan area control Dataset 1 were rectified to a 1:20,000 scale digital scanmap with a 2 m pixel resolution. All rectifications were better than ½ pixel, which is ~10 m for the HRV and HRVIR 20 m multispectral data and ~5 m for the HRG 10 m multispectral data. For the Taita Hills application Dataset 2, because of the rugged terrain in the area, the imagery was orthorectified utilizing the generated 20 m DEM and the SPOT geometric correction model available in ERDAS IMAGINE software. This model uses manually identified ground control points (GCPs) in the image and a geometric control source, the DEM, and a single frame space resection technique to complete the orthorectification for a single SPOT scene, processed individually (ERDAS 2005). Firstly, the 2003 image was orthorectified to the 1:50,000 scale scan-maps and then the 1987 and 1992 scenes were orthorectified based on this geometric master scene, deriving an inter-scene agreement of 0.45 pixels RMSE. With both imagery datasets, a nearest-neighbour resampling technique was employed to ensure that the original pixel values were preserved.

4.2.2 Conversion of SPOT image DN into at-satellite radiance

The first radiometric processing step applied to the SPOT imagery for a HELM atmospheric calibration is to convert the 8-bit DN values into top of atmosphere at-satellite radiance (L_{SAT}) in W m^{-2} sr^{-1} μm^{-1}. The conversion is implemented using the band-specific absolute calibration gain coefficients (G) supplied in the SPOT image metadata and the simple equation:

$$L_{SAT} = DN / G \tag{11}$$

4.3 Atmospheric correction of SPOT data in ancillary data limited circumstances

4.3.1 Identification of within-scene dark-object radiance values

For the HELM atmospheric correction detailed in Paper II and the DOS (dark object subtraction) techniques applied in the correction comparisons in Paper III, it was necessary to estimate path radiance (L_P) from the SPOT scenes themselves. Usually within a SPOT image, some pixels will either be formed from a surface material with very low reflectance, such as clear water, or will be in complete topographic shadow. Such pixels are known as dark-objects and the L_{SAT} recorded in the VIS/NIR can be considered to be composed primarily of the atmospheric upwelling L_P, assuming that the areas are large enough to counter the adjacency effects of surrounding land cover types (Chavez 1996). A dark-object reflectance of 0.01 (1%) (Moran *et al*. 1992; Chavez 1996) was considered appropriate for the SPOT green and red bands. However, in this study, based on reflectance measurements of clear lake water, it was felt that a nominal value of 0.001 (0.1%) was a better assumption for the NIR, as the reflectance from water surfaces is more or less zero at wavelengths beyond the red (Tso & Mather 2001). Furthermore, because the amount of scatter is negligible in the SWIR (Moran *et al*. 2003), it was considered scattering could effectively be ignored for SPOT band 4, which is primarily attenuated by absorption by atmospheric water vapour (Moran *et al*. 2003).

The utilized dark-object radiance value identification procedure was to visually inspect both the histogram and spatial location of the darkest pixels in a scene. This was done both to ensure that the dark-object was a meaningful feature, such as an area of topographic shadow and not an edge pixel, and also to ensure that the values were similar to those of adjacent pixels, as this indicated that the value was not a dropout. This, therefore, also allowed values with a small frequency of occurrence to be selected.

For both Dataset 1 and 2 imagery it was found that, whilst the spatial location of the dark-object for the green and red bands was the same, the location of the NIR dark-object was always different. Furthermore, the NIR dark-objects identified in both datasets were, without exception, areas of water, justifying basing the assumption of a 0.1% nominal NIR dark-object reflectance on the characteristics of water. Therefore, another factor in need of consideration was the possibility that the identification of separate dark-objects for the visible and NIR wavelengths from different spatial locations could incorrectly alter the relative relationships between the bands. However, as Chavez (1988) noted in a comparison of a simple histogram based and an improved mathematically correlated selection technique, both methods generated

very similar haze values when sufficient topography existed; that is to say, when there were truly very low reflectance dark-objects existing within every scene.

Topographically flat desert regions are the only likely problematic application area, as here there may be a complete absence of water or topographic shadow within a scene. In such circumstances, however, the atmosphere is very stable and dry and it would be possible, for example, to utilize the REL (Moran *et al*. 2001) approach of estimating L_{SAT} for $\rho_s = 0$ using an RTM and standard atmosphere and aerosol models.

In this study, then, the same input dark-object radiance values were used for the HELM and DOS approaches, although each technique varies slightly in the method used to estimate L_P from these inputs. Note also these approaches assume a homogenous atmosphere throughout the scene area and consequently that the same spectral band specific additive scattering components can be subtracted from every pixel in the image.

4.3.2 The Historical Empirical Line Method (HELM) for atmospheric correction

The full details of HELM are outlined in Paper II, whilst an assessment of the spectral spatial, temporal, and multiangular reflectance variability of various asphalt, sand, gravel, grass, and fake turf potential pseudo-invariant ground calibration targets is given in Paper I. HELM is designed for absolute atmospheric correction of SPOT imagery in local landscape level and regional scale remote sensing studies where no detailed meteorological data are available, but where there is field access to the research site(s) and a goniometer or spectrometer is available. The purpose of HELM is to (re)construct the historical linear relationship between L_{SAT}, as recorded in the multi-temporal SPOT imagery, and ρ_s for spectrally pseudo-invariant features (PIFs), as measured in the field.

Ideally, the HELM calibration target multiangular reflectance behaviour can be modelled from ρ_s goniometer measurements. This enables the derivation of spectral band specific ρ_s equivalent to geometrical circumstances of existing and future SPOT imagery within a database, which can be used for a HELM-2 calibration for RFR. Where a goniometer is unavailable, however, it is necessary to undertake a HELM-1 calibration to nadir ρ_s field measurements. The induced calibration error in doing so, relative to the possible ± 31° θ_V range of SPOT data, was investigated for control Dataset 1 and the results are detailed in Paper II. In the atmospheric comparison study, as the emphasis was on applying corrections in data limited circumstances, HELM was applied as a HELM-1 calibration to nadir ρ_s.

The HELM sampling strategy is straightforward and there are two main objectives that can be fulfilled within one day of good weather: (1) quantifying the spectral stability of the nadir ρ_s across the site, and (2) capturing the variation in the multiangular reflectance characteristics (or, if a goniometer is not being used, the nadir ρ_s) at a single location within the target, with the range of θ_Z corresponding to the overpass times that will be experienced by the SPOT sensors throughout the entire year, or the season(s) of interest.

Site multiangular ρ_s behaviour can be investigated by calculating the HDRF anisotropy factor ($HDRF_{ANIF}$) by normalizing the multiangular HDRF data to the nadir HDRF ($HDRF_O$). This allows for the direct comparison of multiangular effects from different spectral bands. For a specific wavelength or wavelength interval (λ), $HDRF_{ANIF}$ can be defined as (after Sandmeier & Itten 1999):

$$\mathrm{HDRF}_{\mathrm{ANIF}}(\theta_i,\varphi_i;\theta_r,\varphi_r;\lambda)=\frac{\mathrm{HDRF}(\theta_i,\varphi_i;\theta_r,\varphi_r;\lambda)}{\mathrm{HDRF}_{\mathrm{O}}(\theta_i,\varphi_i;\lambda)} \quad (12)$$

In the limiting circumstances where only nadir ρ_s measurements are available, the assessment of target ρ_s stability for a HELM-1 application can be simplified to determining the relationship between nadir ρ_s and θ_Z and computing the mean absolute difference (MAD) between the measured and modelled values for each SPOT band. A site with a small MAD across all wavelengths would likely be a suitable calibration target (Moran *et al*. 2001). This relationship can be modelled using a simple second order polynomial, where *b* and c_i are fitted constants:

$$\rho_s = b + c_1\theta_Z + c_2\theta_Z^2 \quad (13)$$

Successful application of HELM relies on the identification of a spectrally homogenous and bright calibration target within each SPOT scene area, as required to cover the study site(s). Previous localized studies that have applied small sized sites in an EL correction have utilized concrete, asphalt, beach sand, 'packed earth' bare soil, bare rock, managed sports ground turf and artificial turf sports fields as calibration targets (Smith & Milton 1999; Moran *et al*. 2001; Karpouzli & Malthus 2003; Xu & Huang 2006).

Where very large pseudo-invariant surfaces occur, which is mainly in the semi-arid regions of the world, they have been used extensively both in regional studies and for the operational on-orbit calibration and validation of satellite sensors. For example, de Vries *et al*. (2007) used field measurements of ρ_s from three claypan sites ranging in size from 23,000 to 44,000 m^2 in Queensland, Australia to calibrate multi-temporal Landsat TM/ETM data. The on-orbit calibration of the TM/ETM sensors themselves is based on the extensive alkali gypsum flats at Chuck Site, White Sands, New Mexico, USA (Slater *et al*. 1987; Thome *et al*. 1997), where the target site is in excess of 50 km^2 (Thome 2001). Teillet *et al*. (2001) proposed a generalized vicarious calibration procedure for multiple satellite sensors based on airborne hyperspectral and ground based spectrometer ρ_s measurements of a dry lake playa in Nevada, USA, and a rangeland area in Alberta, Canada, from which 7 km by 7 km 'prime' sites were chosen. The procedure was successfully applied to provide on-orbit calibration of NOAA-14, AVHRR, OrbView-2 SeaWiFS, SPOT-4 VGT, SPOT-1/2 HRV and Landsat-5 TM data.

The on-orbit calibration of SPOT HRV data was based on the La Crau test site in southeastern France, which is a 400 m x 400 m area (160,000 m^2) formed of sandy clay soil with pebbles (Santer *et al*. 1992; Rondeaux *et al*. 1998), and SPOT data have also been validated against measurements from the White Sands test site (Teillet *et al*. 2001; Gellman *et al*. 1993). Such spatially extensive sites are, however, generally very rare and for the worldwide application of HELM in localised or regional studies the utilisation of much smaller calibration targets is necessary. Nevertheless, it is still potentially problematic to identify spectrally homogenous calibration targets that are large enough to counter the radiometric 'contamination' from adjacency effects and from the point spread function (PSF) of the GIFOV of the SPOT sensors.

Karpouzli & Malthus (2003) state that targets need to be at least three times the pixel size to derive a central L_{SAT} pixel value allowing a good estimate of the target ρ_s. This represents a minimum size requirement of 60 x 60 m for 20 m resolution SPOT

HRV/HRVIR data, and 30 x 30 m for 10 m HRG imagery. Based on a mathematical consideration of sensor PSF and atmospheric blurring, however, Richter (1997) calculated that a target size seven times the GIFOV of the sensor is necessary so that the central nine pixels can be sampled effectively. More conservatively, Moran *et al*. (2001) argue that, even where the target is bright and surrounded by a darker surface, a ratio of sensor resolution to target size of 1:8 is required to ensure that at least four central pixels remain uncontaminated. This, then, translates to a minimum target size of 160 x 160 m for 20 m and 80 x 80 m for 10 m imagery, a more stringent size constraint that can be applied to the requirements for a HELM *calibration* target, as opposed to a validation site. Evidently, the larger the spatial extent of the site the better, in terms of obtaining an accurate L_{SAT} representation of the target ρ_s from the SPOT imagery, as more pixels can be included into the target AOI (area of interest).

As discussed by Smith & Milton (1999), Thome (2001), Karpouzli & Malthus (2003), and de Vries *et al*. (2007) there are several critical characteristics that must be considered in pseudo-invariant site selection, especially given it may be necessary to make HELM calibrations to nadir only ρ_s measurements. A calibration target should be (1) spatially extensive, (2) have near-Lambertian reflectance characteristics to minimize ρ_s effects due to changes in solar and view geometry, and (3) be a homogeneous bare surface devoid of temporally variant features, such as vegetation cover. Furthermore, the site should also (4) be located on flat and level terrain so there are no topographic illumination variations present. The calibration site also needs (5) to be spectrally bright enough in all the SPOT bands to enable a good estimate of the correction lines, given that dark-object selection identifies the lower radiance values. In order that measurement of ρ_s need not coincide with the image data acquisition, the calibration target also needs (6) to have stable or predicable ρ_s over time, although it is acknowledged that this is the most time-consuming factor to assess. Finally, (7) the field accessibility of the site is also a consideration, especially if a goniometer is to be utilized. These requirements are not insurmountable, however, and indicate that manmade/artificial surface features, or vegetation free natural surfaces, are likely to be the most appropriate calibration targets.

In order to assess the accuracy of the derived correction lines, ideally further measurements of ρ_s for a number of validation ground targets with varying reflectance should also be collected. Given the SPOT sensors' GIFOV, it may be necessary to utilize sub-optimal sites for some validation targets, depending on the specifics of the study area. Prior to field measurements, an estimate of spectral stability across a potential calibration/validation site at the scale of the sensor GIFOV can be obtained by examining the range and standard deviation (SD) of the L_{SAT} pixel values of the target AOI in the SPOT satellite image itself. Care should be taken to avoid the selection of mixed pixels at the periphery of the target. Additionally, for areas where very high-resolution imagery is available online, the Google Earth™ or other similar mapping websites are a good way of visually interrogating possible calibration/validation sites before visiting them in the field. Consideration should be given to the constituent parts of a site that will not be visible in the SPOT imagery because they cannot be resolved at a 20 m or 10 m pixel resolution. For example, as noted by Milton *et al*. (1997), large asphalt parking lots or areas of hard standing are often used as PIFs, but they usually contain painted lines and markings and may also contain areas of surface contaminants, such as oil stains, which may be overlooked by the human observer. Nevertheless, the reflectance of such features will be integrated into the GIFOV recorded by the SPOT sensor. Another, seemingly obvious, point to

consider is the likely presence of motor vehicles, which may not be visible in the SPOT imagery.

The temporal spectral stability of a target is another area where assumptions are more often made than tested, which may be a critical oversight. For example, Anderson & Milton (2006) measured the surface of a disused weathered concrete runway and found that there were both diurnal and seasonal variations in ρ_s. It was the scale of the annual ρ_s variation, for a target type that is often used for calibration, which was most surprising. There was seasonal darkening of the surface by a factor of two in the visible spectrum caused by the growth of a biological material (algal lichen) during the springtime, and an increase in ρ_s throughout the rest of the year. They also measured a non-systematic diurnal reduction in ρ_s over the 400-1000 nm spectral range during a day in June 2003 that was attributed to the activation of algal lichen photosynthetic activity brought about by the arrival at the site of an air mass with high humidity. Additionally, as might be expected, a systematic daily change in ρ_s due to the variation in θ_Z was also noted, with a general trend of decreasing nadir ρ_s with increasing θ_Z and vice-versa.

Changes in surface moisture conditions at a vegetation-free site, related to recent rainfall and its intensity or – in some instances – related to fluctuations in the height of the local water table, can drastically alter its reflectance characteristics as wetness lowers ρ_s (Wheeler *et al*. 1994; Moran *et al*. 2001). However, drying times may be fairly rapid in cloud-free and moderately windy conditions. For example, Moran *et al*. (2001) found that a packed earth target in a semi-arid area artificially saturated to a depth of 2 cm dried within 15 minutes. Given that cloud-free SPOT images are usually chosen for utilization, it is relatively unlikely that target wetness will be a major factor in the temporal spectral stability. Nevertheless, it may be prudent to consult any available rainfall records for the study area in the period immediately prior to the image acquisition dates, to ascertain if there had been any precipitation and if so how much. Bare soil targets may also show a gradual brightening of their reflectance over a season due to rain compaction (Moran *et al*. 2001).

Vegetated targets are tantalizing as calibration sites because of their high reflectance in the NIR band, a characteristic that other surface materials do not often share. Moran *et al*. (2001) argued that sites with > 90% cover and very high LAI could be utilized, and they gave a measured example of an irrigated field of mature cotton. Similarly, Karpouzli & Malthus (2003) used the managed turf of a sports pitch as an EL calibration target. Nonetheless, vegetated sites are problematic for HELM because they are likely to show at least some temporal variability, and also are likely to show strong ρ_s variability with the sensor view angle (Sandmeier & Itten 1998b).

As detailed in Paper II, site 1 (see Table 5 above), Hietsu beach, was chosen as the Dataset 1 HELM calibration target because of its temporal persistence, large size, relative brightness, spectral homogeneity and limited multiangular reflectance properties. To apply HELM to the Dataset 1 imagery, the average L_{SAT} values for each spectral band for the calibration and validation sites were derived from each L_{SAT} image. Care was taken to avoid the selection of mixed pixels at the periphery of the targets. The utilization of a spectrally homogenous calibration site, covering a number of uncontaminated pixels, meant that the spatially averaged site mean reflectance spectrum could be directly equated with the support of the SPOT sensor response. Next, models describing the multiangular ρ_s behaviour, based on the Lommel-Seeliger function (Hapke 1993; Suomalainen 2006), were fitted to the FIGIFIGO ground target measurements (Table 5) to allow interpolation of the data. Then, for each spectral band, both nadir and geometrical ρ_s relating to each scene were synthesized based on

the wavelength intervals and relative sensitivities of the utilized sensors. This allowed for the application of both HELM-1 and HELM-2 calibrations, in order to study the effects on RFR accuracy, as well as the derivation of the best possible ρ_s estimates of the validation sites for utilization in the accuracy assessment. Then the within-scene dark-object radiance values for the VIS/NIR were identified using the methodology discussed in Section 4.3.1 above. Based on the assumptions of HELM, these L_{SAT} values relate to ρ_s of 0.01 in the green and red, and 0.001 in the NIR. For each spectral band in each image, a separate HELM-1 and HELM-2 correction line was calculated utilizing a standard linear regression equation of the form $y = ax + b$; where the SPOT L_{SAT} was taken as the independent variable and nadir ρ_s as the dependent variable, and a is the slope of the regression line, representing atmospheric attenuation, and b is the intercept with the x-axis, representing L_P. Finally, the derived HELM-1 and HELM-2 correction equations for each band and image were applied to compute an output ρ_s file from each input Dataset 1 L_{SAT} SPOT image.

A 60 m wide and 200 m long roadside quarry was taken as the Dataset 2 Taita Hills application site HELM calibration target (see Table 6 above). To apply HELM to Dataset 2, the average L_{SAT} values for each spectral band for the calibration and validation sites, and the dark-objects, were identified in each image. The nadir HCRF field measurements were then processed to simulate the response of the spectral bands of the various SPOT sensors that captured the imagery (see Table 4 above). Finally, HELM correction equations were derived for each spectral band for each image, and applied to compute an output ρ_s file from each input Dataset 2 L_{SAT} SPOT image. The calibration to nadir HCRF represented a HELM-1 application.

4.3.3 Image-based atmospheric correction methods

In ancillary data limited circumstances, the ability to make an absolute atmospheric correction based only on information derived from the SPOT data itself represents a significant advantage. Consequently, in Paper III four different DOS approaches were implemented and compared with the 6S and HELM corrections. Ignoring atmospheric polarization, refraction and adjacency effects, and assuming isotropic sky irradiance in a homogenous cloud free atmosphere and Lambertian reflectance from a flat, uniform ground surface, the relationship between L_{SAT} and surface reflectance (ρ_s) can be characterized as (Song *et al.* 2001):

$$L_{SAT} = L_P + \frac{\rho_s \cdot \left(E_O \cdot \cos\theta_Z \cdot T_Z + E_{DOWN}\right) \cdot T_V}{\pi \cdot \left(1 - S \cdot \rho_s\right)} \qquad (14)$$

where L_P is the path radiance (W m^{-2} sr^{-1} μm^{-1}); S is the fraction of up-welling radiation back-scattered by the atmosphere to the surface and is small enough that it can be omitted (Song *et al.* 2001); T_V is the atmospheric transmittance from ground target to sensor; T_Z is the atmospheric transmittance from sun to ground target; E_O is the exoatmospheric solar constant (W m^{-2} μm^{-1}); E_{DOWN} is the downwelling diffuse irradiance (W m^{-2} μm^{-1}); and cos θ_Z is the cosine of the solar zenith angle. To retrieve ρ_s, Equation 14 can be rearranged as (Chavez 1996; Song *et al.* 2001):

$$\rho_s = \frac{\pi \cdot \left(L_{SAT} - L_P\right)}{T_V \cdot \left(E_O \cdot \cos\theta_Z \cdot T_Z + E_{DOWN}\right)} \quad (15)$$

$E_O = E \div d$; where E is the SPOT sensor and band specific equivalent solar irradiance in W m^{-2} μm^{-1} (obtained from the SPOT website: www.spotimage.fr/html/_167_224_229_.php), and d is the date corrected Earth-Sun distance in astronomical units ($d = au^2$) estimated as (Tso & Mather 2001):

$$d = (1 - 0.01673 \cdot \cos(0.9856 \cdot (JD - 4)))^2 \quad (16)$$

where JD denotes the Julian Day. Equation 15 forms the basis for the implementation of the image based atmospheric correction approaches applied in this study and it can be seen that there are four unknown atmospheric correction variables to be estimated: L_P, T_V, T_Z and E_{DOWN}. Omitting these completely derives the at-satellite reflectance (ρ_{SAT}), which normalizes for variation in the Earth-Sun distance and θ_Z:

$$\rho_{SAT} = \frac{\pi \cdot L_{SAT}}{E_O \cdot \cos\theta_Z} \quad (17)$$

This partially corrected case was included in the study as a comparator with the DOS, 6S and HELM full correction methods. Similar to the work of Song *et al*. (2001), four DOS approaches were implemented, each of which estimates ρ_s based on different simplifying assumptions for the calculation of L_P, T_V, T_Z and E_{DOWN}, as summarized in Table 7 below. In all four DOS methods the same procedure as that adopted for HELM was utilized in identifying the darkest within-scene pixels in each spectral band from which to estimate L_P (Section 4.3.1). Consequently, the same input dark-object radiance values were derived and the same assumptions for the surface reflectance of the dark-objects were applied. However, because of the different assumptions in the calculation of T_V, T_Z and E_{DOWN} in each of the four approaches, they all derived slightly different estimates of L_P. Assuming a dark-object surface reflectance of 1% for the SPOT green and red bands 1 and 2, and 0.1% for the NIR band 3, L_P can be calculated as:

$$L_{pVIS} = L_{DOS} - 0.01\left(\left[E_O \cdot \cos\theta_Z \cdot T_Z + E_{DOWN}\right]T_V / \pi\right) \quad (18)$$

$$L_{pNIR} = L_{DOS} - 0.001\left(\left[E_O \cdot \cos\theta_Z \cdot T_Z + E_{DOWN}\right]T_V / \pi\right) \quad (19)$$

where L_{DOS} is the radiance value of the identified darkest within-scene object. It was considered unnecessary to apply a correction to remove L_P for SWIR band 4.

As detailed in Table 7, DOS1 is the simplest approach where T_V and T_Z for all spectral bands are assumed to be 100% and E_{DOWN} is ignored (Chavez 1989). Calculation of L_P for bands 1 to 3 is based on Equations 18 and 19 above, applying these simplified DOS1 assumptions for T_V, T_Z and E_{DOWN}. This is a basic haze removal approach. As no scattering is assumed for band 4, SWIR ρ_s prediction is identical to ρ_{SAT}. DOS1 is likely to produce unacceptable results, especially for higher

surface reflectance values (Chavez 1996), because it does not correct for the multiplicative effect of transmittance.

Table 7. *Parameterisations utilized in the four implemented DOS approaches.*

Method	T_V	T_Z	E_{DOWN}
DOS1	1	1	0
DOS2 (COST)	$\cos(\theta_V)$	$\cos(\theta_Z)$ [a]	0
DOS3	$e^{-\tau r/\cos(\theta v)}$	$e^{-\tau r/\cos(\theta z)}$	6S
DOS4	$e^{-\tau/\cos(\theta v)}$	$e^{-\tau/\cos(\theta z)}$	πL_p

[a] for green, red and NIR bands, and unity for SWIR.

Chavez (1996) argued that for Landsat images with an atmospheric optical thickness between 0.08 – 0.3 and a solar zenith angle of 30° - 50° (which were considered as usual viewing conditions) T_Z can be approximated to a first order by the cosine of the solar zenith angle, cos θ_Z. Thus, a simple multiplicative DOS correction can be derived by equating T_Z to cos θ_Z for the VIS/NIR bands, which Chavez (1996) termed the COST approach, and which for a test dataset in a semi-arid environment was shown to derive as accurate estimates of ρ_s as corrections based on RTMs and *in situ* atmospheric measurements (Chavez 1996). COST forms the basis of the DOS2 approach in this study although, as a further step, variation in T_V due to differences in the SPOT imagery off-nadir θ_V was accounted for by equating T_V to cos θ_V for all bands. In DOS2, SWIR T_Z is assumed 100%, therefore the only difference in ρ_s prediction compared to ρ_{SAT} is the correction for θ_V. Note, however, that both DOS1 and 2 ignore E_{DOWN}, which according to Moran *et al.* (1992) can contribute as much as 25% of the spectral radiance received at the surface, even for relatively clear atmospheres. Consequently, this is likely to induce error in ρ_s estimation. Calculation of L_P for DOS2 is based on Equations 18 and 19 above, applying the given DOS2 assumptions for T_V, T_Z and E_{DOWN}.

Following Song *et al.* (2001), DOS3 was a simplified Rayleigh scattering only atmosphere model approach whereby the effect of aerosols is ignored. The optical thickness for Rayleigh scattering (τ_r) was estimated on a per-band basis using Equation 20 (Kaufman 1989), where λ is wavelength in μm, taking the median value of the spectral channel width as the band centre wavelength in each SPOT band (this being the same for SPOT HRV/HRVIR/HRG sensors):

$$\tau_r = 0.008569 \cdot \lambda^{-4} \cdot \left(1 + 0.0113 \cdot \lambda^{-2} + 0.00013 \cdot \lambda^{-4}\right) \qquad (20)$$

Following the standard approximation for atmospheric transmittance (Moran *et al.* 1992), T_V and T_Z for the Rayleigh scattering only atmosphere model were estimated as $e^{-\tau r/\cos(\theta v)}$ and $e^{-\tau r/\cos(\theta z)}$. Further, E_{DOWN} for a Rayleigh atmosphere was estimated from the 6S RTM defined as zero aerosol optical depth at 0.55 μm, based on the midlatitude summer standard atmosphere model for Dataset 1 and the tropical standard atmosphere model for Dataset 2. E_{DOWN} was taken as scene-specific atmospheric diffuse irradiance at ground level reported by 6S. Calculation of L_P was based on Equations 18 and 19, iterating on the 6S estimates of E_{DOWN}.

In DOS4, an attempt was made to account for the additional effects of atmospheric aerosols and T_V and T_Z were calculated as $e^{-\tau/\cos(\theta v)}$ and $e^{-\tau/\cos(\theta z)}$. Given the assumption of isotropic sky radiance, $4\pi L_P$ was considered as an estimate of

exoatmospheric irradiance loss and the optical thickness of the atmosphere can therefore be estimated by (Song *et al.* 2001):

$$T_Z = e^{-\tau / \cos(\theta_Z)} = 1 - \frac{4\pi L_p}{E_O \cdot \cos\theta_Z} \quad (21)$$

Solving for τ and substituting with Equation 18 for estimating L_P for a 1% dark reflector in the visible wavelengths derives (Song *et al.* 2001):

$$\tau = -\cos\theta_Z \cdot \ln\left(1 - \frac{4\pi\left(L_{DOS} - 0.01\left[E_O \cdot \cos\theta_Z \cdot T_Z + E_{DOWN}\right]T_V / \pi\right)}{E_O \cdot \cos\theta_Z}\right) \quad (22)$$

The difference for the NIR is the assumption of a 0.1% dark-object reflectance. In DOS4, all four atmospheric correction variables (L_P, T_V, T_Z and E_{DOWN}) are initially unknowns before τ is estimated. Consequently it was necessary to derive a solution iteratively. Initial calculations were made based on defining L_P as the uncorrected radiance values of the identified within-scene dark-objects, deriving E_{DOWN} directly from these approximations of L_P, and by setting T_V = T_Z = 1. After the initial calculation of τ, new values of T_V and T_Z were obtained and these were fed back into the calculations again until the estimate of τ stabilized.

There are limitations in the assumptions and simplifications of these DOS approaches, nevertheless such DOS methodologies have the significant benefit that they are totally image based and can consequently be implemented easily on historical multi-temporal satellite imagery, with no need for additional information. This makes them a valid alternative to HELM corrections.

4.3.4 Second Simulation of the Satellite Signal in the Solar Spectrum (6S) RTM

In Paper III, the 6S RTM (Vermote *et al.* 1997a, 1997b) was utilized as it is robust, universally applicable, freely available, and widely utilized and it offers a viable alternative to HELM and DOS correction methods if used with atmospheric models and reasonable aerosol estimates. 6S requires inputs for the geometrical conditions of the scene acquisition, a model of the gaseous composition of the atmosphere, a description of the concentration and type of aerosols, details of the spectral sensitivity of the satellite band, and the type and spectral variation of the ground reflectance (Vermote *et al.* 1997b). Various typical atmospheric scenarios are already implemented in the software and are available for the user.

Definition of the geometrical conditions at the time of scene capture was straightforward, based on the supplied SPOT metadata. 6S does not, however, account for the SPOT θ_V. For all images in the Taita Hills application site Dataset 2, the 6S tropical standard atmosphere model (which determines the gaseous composition) and the continental aerosol model were selected as the closest approximation. To determine the concentration of aerosols, the user can either specify the AOD at 0.55 μm or the horizontal visibility in km at the time of scene acquisition (which 6S uses to estimate the AOD based on the standard aerosol profile). For Dataset 2, because of the lack of horizontal visibility measures, the AOD at 0.55 μm was approximated as the

average of the AREONET measurements for southern Africa, 0.19, as reported by Remer *et al*. (2005, Table 8, p.965). This compares to a global average AOD at 0.55 μm for land of 0.18 (Remer *et al*. 2005).

For the Helsinki metropolitan region control site Dataset 1, the midlatitude summer standard atmosphere model and the continental aerosol model were selected. In order to facilitate a comparison of the possible differences in the output ρ_s, 6S corrections were applied to Dataset 1 using both overpass concurrent measurements of horizontal visibility, taken as the average UTC09:00 (12:00 local time) visibility of the five weather stations falling in the Helsinki area (WMO 02795, 02974, 02978, 05105, 05194; obtained from the Finnish Meteorological Institute: www.fmi.fi), and also a general estimate of AOD at 0.55 μm of 0.17 derived as the average of the AREONET measurements for Western Europe, as reported by Remer *et al*. (2005, Table 8, p.965).

The spectral bandwidths and sensitivities of the SPOT 1 HRV1 and HRV2 sensors are implemented in 6S, but the sensitivities for the other SPOT sensors had to be calculated and defined as required in 0.0025 μm steps. Input values to be atmospherically corrected were given as L_{SAT}. For both Dataset 1 and 2, the same Hietsu beach and roadside quarry HELM calibration sites were also utilized for 6S calibration. For every image, the average L_{SAT} value for each spectral band within the site AOI polygon was taken as the input. Within 6S the surface type was defined as sand in all cases, this being correct for Hietsu beach and the nearest approximation for the roadside quarry surface. As the assumption is being made that detailed knowledge of the BRDF of the calibration sites is *unknown* at the time of atmospheric correction, the sand surfaces were defined as homogenous with no directional effects (i.e. Lambertian). The elevation of the calibration sites was also required for 6S, this being sea level for Hietsu beach and 960 m above sea level for the roadside quarry.

The *a*, *b* and *c* correction coefficients output from 6S can be used in conjunction with Equations 23 and 24 below to apply the correction to the whole SPOT image, give an assumption of a homogenous atmosphere throughout the scene area. For a specified spectral band:

$$\rho_s = \frac{y}{1 + c \cdot y} \tag{23}$$

$$y = a \cdot L_{SAT} - b \tag{24}$$

Several test runs made with different measurement locations within the same scene taken as the 6S calibration site (and therefore with varying L_{SAT} values as inputs) derived identical *a*, *b* and *c* transformation coefficients as outputs. This is because all the input parameters for 6S, other than site elevation and L_{SAT}, were held constant throughout the image area. Consequently, as a basic implementation, it was appropriate to base a correction of the whole image on 6S outputs derived only from the calibration site.

4.4 Atmospheric correction accuracy assessment methodology

For both Papers II and III, the performance and accuracy of the atmospheric correction methodologies applied to control Dataset 1 and application Dataset 2 were assessed using two approaches. Firstly, consideration was given to the difference between the predicted ρ_s and the field measured ρ_s for the verification sites, the precision of both of which was derived to 0.001 ρ_s. The HELM calibration sites for Datasets 1 and 2 were included in the accuracy assessments as validation targets, as in this way the HELM induced error in calibration to nadir ρ_s was included into the assessment. For control Dataset 1, the field estimates of ρ_s were defined as those relating to the specific illumination and θ_V geometry of the 2005, 2003 and 2002 SPOT scenes, as modelled from goniometer measurements. For Dataset 2, only nadir ρ_s were available and consequently the 2003 image, which was nearest in time to the field measurements and had a relatively small θ_V of R10.4°, was chosen for assessment. Both the absolute and relative root mean square error (RMSE and $RMSE_r$) and bias (Bias and $Bias_r$) (see Equations 25-28) of the different atmospheric correction approaches in predicting ρ_s in each spectral band for each site, compared to that derived from field measurements, were calculated and the summary mean average RMSE, $RMSE_r$ and bias for all bands and all measurement sites were compared. The utilized equations were as follows (Heiskanen 2006):

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2}{n}} \tag{25}$$

$$\mathrm{RMSE}_\mathrm{r} = \frac{\mathrm{RMSE}}{\bar{y}} \times 100 \tag{26}$$

$$\mathrm{Bias} = \frac{\sum_{i=1}^{n} \hat{y}_i - y_i}{n} \tag{27}$$

$$\mathrm{Bias}_\mathrm{r} = \frac{\mathrm{Bias}}{\bar{y}} \times 100 \tag{28}$$

where $\hat{y}_i$ is the estimate, y_i is the field observation, $\bar{y}$ is the mean of the observations, and n is the number of measurement sites. The statistical significance of the bias with n-1 degrees of freedom, based on repeated dependent measurements, was tested using:

$$t = \mathrm{Bias} \div \left(\sigma \Big/ \sqrt{n} \right) \tag{29}$$

where σ is the standard deviation of the residuals $\left(\hat{y}_i - y_i\right)$. The level of statistical significance of the bias in a spectral band was considered by noting when t exceeded the t values corresponding to a probability of both $p < 0.05$ and $p < 0.01$.

Secondly, the ability of each atmospheric correction methodology to maintain or reduce the variation in the average ρ_s between all the multi-temporal SPOT scenes within each dataset, compared to the partially corrected ρ_{SAT}, was assessed. Theoretically, the removal of atmospheric effects should reduce the spectral variation between images (Schroeder *et al*. 2006). The spectral band specific absolute difference in the mean average ρ_s of the scene between every image pair in each dataset was calculated and the mean average absolute difference for each band, as well as the overall average for all bands combined, were compared. Because of differences due to the presence of clouds at the edges of some of the images in both datasets, and due to shifts along track in Dataset 1 imagery, a subset area was utilized for each dataset in calculating the mean ρ_s.

4.5 Topographic correction

For mountainous and rugged terrain regions, such as the Taita Hills, topographic correction is at least as important as atmospheric correction, if comparable ρ_s are to be derived throughout the area (Liang 2004). Therefore, after applying the HELM atmospheric corrections to the SPOT imagery, it was necessary to further remove the slope-aspect effects. Use was made of a topographic correction method based on the cosine function described by Teillet *et al*. (1982), with band specific '*c*' correction factors calculated for general vegetation classes identified using a novel method. Due to the requirement to identify vegetation classes in the imagery *before* a topographic correction and classification were applied, Normalized Difference Vegetation Indexes (NDVI) were derived for each scene and were then clustered using the automated ISODATA algorithm in ERDAS IMAGINE software. NDVIs were used because they are a ratio between the red and NIR bands and are consequently relatively unaffected by topographic effects, i.e. both bands respond to the variation in illumination in a similar way. The ISODATA algorithm is more usually utilized to automatically identify data clusters in multi-spectral feature space, but here it is used in a one-dimensional manner to capture naturally occurring frequency clusters in the NDVI.

Illumination can be defined as the cosine of the solar incidence angle (cos *i*), representing the proportion of direct solar radiation hitting a pixel within a SPOT image. The amount of illumination is therefore dependent on the relative orientation of the pixel toward the Sun's actual position during image acquisition, as determined from a DEM of the area, and can be calculated as:

$$\cos i = \cos\theta_Z \cos\theta_S + \sin\theta_Z \sin\theta_S \cos(\varphi_Z - \varphi_S) \qquad (30)$$

where θ_Z is the solar zenith angle, φ_Z is the solar azimuth, and θ_S is the slope and φ_S is the aspect of the pixel. If the surface is flat then the aspect is undefined and i is simply θ_Z.

Following heuristic experimentation, ten NDVI cluster classes were initially derived for each image and, for each spectral band and vegetation class combination, the reflectance values for pixels with a slope greater than 5° were regressed against the corresponding cos *i* values using standard linear regression. Where there was not a significant relationship, which occurred with some of the least vegetated NDVI clusters, the classes were merged and re-regressed until groupings with a usable stronger coefficient of determination were generated. A *c* correction factor for each utilized spectral band and vegetation class combination was then calculated by

dividing the intercept of the finalized regression line by its slope; see Equation 31. No *c* factors were applied to slopes less than 5° in the implementation of the correction, as the cosine correction here is very slight in any case. Removal of slope-aspect effects from the HELM corrected SPOT imagery was implemented as:

$$\rho_{\mathrm{H}} = \rho_{\mathrm{T}} \frac{\cos\theta_{\mathrm{Z}} + c}{\cos i + c}$$
$$c = \frac{b}{m} \qquad (31)$$

where p_{H} is the surface reflectance of a corrected pixel, ρ_{T} is the surface reflectance of an uncorrected pixel, and *b* is the *y*-intercept and *m* is the gradient of the linear regression line of *cos i* against ρ_{T} for a specific spectral band and vegetation cover type combination. This model does, however, have limitations; for example, it does not account for irradiance reflected from surrounding terrain. However, several studies have found that good correction results can be obtained using the c-correction model (see, for example, Riaño *et al*. 2003).

4.6 Multi-scale segmentation and object-oriented classification of SPOT data

A crucial first step in an understanding of landscape changes is to accurately quantify and map LULC over time. Remote sensing offers the most efficient methodology for routinely monitoring at a landscape level over local and regional scales. However, as noted by Burnett & Blaschke (2003), in common with all observation of reality, remotely sensed images are an imperfect capturing of patterns, which are themselves an imperfect mirror of ecosystem processes. The GIFOV of a sensor, which is realized as the pixel resolution of the imagery, is actually a complex phenomenon determined by technical constraints as much as by mapping requirements. Moreover, there is not one individual scale that is appropriate for mapping a landscape if it is accepted that reality is formed of a mosaic of process continuums. Drawing on the work of Koestler (1967), Wu & Loucks (1995) and Wu (1999) argue that by breaking down ecological complexity through a hierarchical scaling strategy, so called hierarchical patch dynamics (HPD) can provide a conceptual framework within which the interaction between ecological processes operating at different scales may be understood. An important characteristic of complex systems is that they take the form of a nested hierarchy, termed a holarchy; thus hierarchies are composed of interrelated subsystems known as holons, each of which are made of smaller subsystems, until a lowest level is reached. The interwoven patterns of heterogeneity and homogeneity have as their basic units the landscape element or *patch*. Patches may be defined as conceptual groupings of spatial homogeneity, that are ubiquitous and which vary at different scales (Wu 1999).

Within the formal framework of Hierarchy theory, a hierarchically organized entity can be conceived as a three-tiered nested system in which levels corresponding to slower behaviour are at the top (Level +1), while those reflecting successively faster behaviour are seen as a lower level in the hierarchy (Level –1). The level of interest is referred to as the focal level (Hay *et al*. 2002). These systems exhibit instability at lower levels, but possess meta-stability at higher levels as, in general, small scale processes tend to be more stochastic. O'Neill *et al*. (1986) recommend the

use of three hierarchical levels as a minimum in analytical studies. This multi-scale analysis conceptualises a landscape as a holarchy, see Figure 8 below; i.e. as a spatially nested patch hierarchy where larger patches are formed of smaller, functional patches. When translating hierarchy theory to landscape ecology, holons are synonymous with patches: the ecological unit at a particular scale. Patches interact with other patches at the same and at higher and lower levels of organization through loose horizontal and vertical coupling. The hierarchical structuring of a landscape is defined at various critical levels of organization where interactions are stronger within levels than between them, and where each level operates at specific spatial and temporal scales (Hay & Marceau 2004). Stronger gradients in these flux rates results in more apparent boundaries, or local heterogeneity (Burnett & Blaschke 2003).

The concept of scale then becomes central to an understanding and analysis of landscapes (Levin 1992). Scale represents the 'window of perception', the filter, or measuring tool, with which a system is viewed, quantified and analysed; consequently real-world objects only exist as meaningful entities over a specific range of scales (Hay *et al*. 2002). Landscape ecologists define scale as having grain and extent, where grain refers to the smallest intervals in an observation set and extent refers to the range over which observations at a specific grain are made (O'Neill & King 1998). In remote sensing, grain is equivalent to the spatial, spectral and temporal resolutions of the image pixels, whilst the extent represents the geographic area, combined spectral bandwidths and temporal duration covered by an image as a whole (Hay *et al*. 2001). Scale may be measured in absolute units or relative to the phenomenon of interest; i.e. the focal scale.

As Hay *et al*. (2001) note, the type of information obtained from remote sensing data is largely determined by the relationship between the actual size of objects in the scene, and the size, or resolution, of the analysis. Considering a specific real world object of interest, as discussed by Strahler *et al*. (1986), there are two fundamental types of resolution present within a remotely sensed image: in the high-resolution (H-resolution) case, the objects of interest vary at a lower spatial frequency than image sampling and features can be resolved. Theoretically, traditional uni-scale per-pixel classification processes are most suited to the H-resolution case, as an object is described by a number of pixels and these pixels are more likely to have homogenous spectral attributes related to the nature of the object itself. In the low-resolution (L-resolution) scenario, the objects of interest are smaller than the pixel size and therefore the spectral response recorded for a single pixel is an integration of many smaller image objects. This results in mixed pixels (mixels) and it should be noted that there are always mixels present in a remotely sensed image as different objects of different spatial extents will be in H- and L-resolution within the same scene from a particular satellite image; i.e. there are multiple scales of objects represented within a single image. Furthermore, different images representing the same objects will be have H- and L-resolution cases depending of the pixel resolution of a particular image. This, then, relates to the general point already made that real-world objects only exist as meaningful entities over a certain range of scales.

Standard image classification procedures, such as the maximum-likelihood (ML) classifier, work on a uni-scale pixel-by-pixel basis and therefore ignore both useful spatial information surrounding the pixel and multi-scale information within the image. Class assignment is based solely on the principle that pixels of the same land cover type will be close in multi-spectral feature space. As Burnett & Blaschke (2003) themselves note, this does not hold true for complex environments. Rather, Burnett & Blaschke (2003) propose a multi-scale segmentation/object relationship

modelling (MSS/ORM) methodology for landscape analysis based on HPD theory and suggest that more accurate analysis can be derived through the application of this technique, especially for heterogeneous landscapes. Central to this methodology is the generation of meaningful image objects relating to landscape patches by multi-scale segmentation, where a search is made for the gradient of flux zones within and between patches. Critically, because MSS/ORM is a move away from pixel-based to object-based image analysis, it is possible to explore multiple scales of objects within the *same* image (Figure 8).

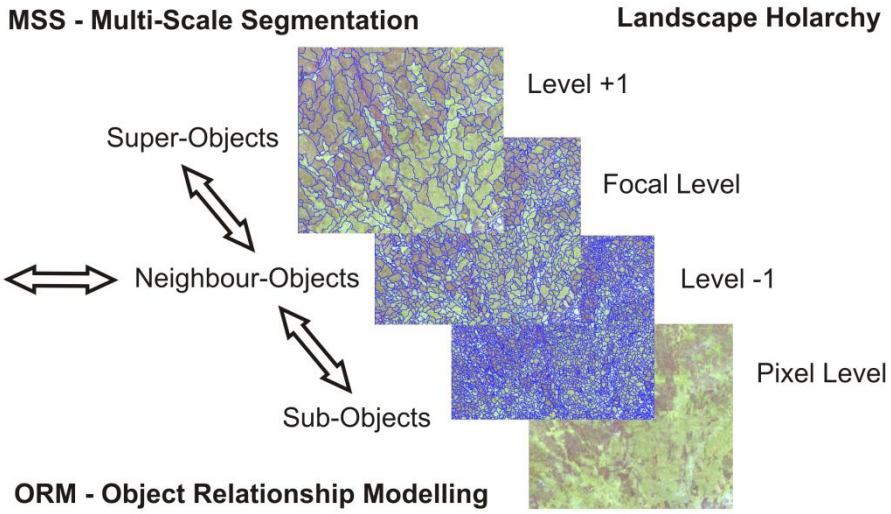


Figure 8. MSS/ORM applied to SPOT imagery covering the Taita Hills to derive a landscape holarchy of 3 hierarchical levels.

Segmentation is the division of remotely sensed images into discrete regions or objects formed of aggregations of pixels that are homogenous with regard to spectral and/or spatial characteristics. Homogenous in this instance refers to the fact that the within-object variance is less than the between-object variance. Research into segmentation techniques is not new, see for example Haralick (1983) and Haralick & Shapiro (1985), and there are a large number of possible methodologies, but the availability of operational software is a development of the last decade. In the Paper IV study, all MSS/ORM work was implemented with the eCognition software (now known as "Definiens"), which utilizes a fractal net evolution approach (FNEA) to multi-scale segmentation. Further details are given in Section 2 and Section 3.2.1 of Paper IV itself, and the full details of the FNEA methodology and of the workings of eCognition's object-oriented fuzzy analysis and classification are covered in depth in Baatz & Schäpe (2000) and Benz *et al*. (2004). FNEA is the most appropriate segmentation methodology for enabling the implementation of the desired MSS/ORM conceptual approach to landscape analysis. Additionally, a major advantage of FNEA is that the heuristics do not evaluate the absolute heterogeneity of a region, but rather evaluate the change in heterogeneity over a merge. This has the desirable effect of enabling relatively homogeneous image segments to remain separate, even if the mean values of adjacent regions are similar. This is important in terms of deriving ecologically meaningful image segments from medium resolution data, such as SPOT imagery, where there are large numbers of pixels formed of mixed land cover types and general spectral overlaps, i.e. there is low spectral contrast.

Following a multi-scale segmentation, a model of the hierarchical relationships between the image objects was built up using both directly calculable properties, such as the mean spectral values or the number of sub-objects, and by the derivation of semantic rules requiring the input of a human expert on the landscape in question. This can be considered as a training phase in an object-oriented (OO) LULC classification methodology. For the OO classification of the Taita Hills, segmented image objects covering the same sample areas that were used in the ML classification were used as training polygons for the fuzzy nearest neighbour classifier implemented in eCognition. The OO classification of the 13 different LULC classes for the Taita Hills, see Section 4.6.1 below, was based on a large number of hierarchical and spatial fuzzy rules, as well as object spectral properties. For full details of the implemented MSS/ORM approach and the OO classification, refer to Paper IV.

As noted by Blaschke (2010), whilst earlier work in the field utilized the term 'object-oriented' image analysis (for example, Blaschke & Hay 2001 and Benz *et al*. 2004), more recently many authors have switched to the term 'object-based' (with or without the hyphen) as 'oriented' was seen as too close to the object-oriented computer programming paradigm (see Hay & Castilla 2008 for discussion). However, others have continued to use the phrase 'object-oriented' (e.g. Navulur 2007), as was the case with the terminology utilized in Paper IV. Moreover, there has also been a recent and ongoing debate within the multi-disciplinary object based image analysis (OBIA) community as to whether or not geographical space should be included in the name of this concept (Blaschke 2010). Hay & Castilla (2008) argue that the most appropriate name is 'geospatial' or 'geographical' object based image analysis (GEOBIA), so as to clearly identify the area as a sub-discipline of GIScience. This debate has yet to be resolved, although it is interesting to note that the most substantial work published in the field to-date (Blaschke *et al*. 2008) was entitled 'Object-based image analysis'.

In accordance with working in regions of the developing world with limited resources, there is a requirement for low cost or preferably freely available data and software. Although in this study the land cover was generated using the commercial Definiens eCognition software, there are freeware alternatives; for example SPRING (www.dpi.inpe.br/spring/english/index.html), which is a product of Brazil's National Institute for Space Research and is capable of image segmentation. However, the classification capabilities of the software may be limited, but the segmentation classification could be conducted in other freeware tools such as WEKA (www.cs.waikato.ac.nz/ml/weka/), which is open-source data mining software in which utilization of the expectation maximization clustering algorithm would be appropriate.

4.6.1 LCCS classification nomenclature

Standardization of LULC classification schemes is an important issue if greater use and understanding of digital mapping products is to be facilitated. The LCCS software developed by FAO and UNEP (DiGregorio 2005) has emerged as a widely accepted conceptual system and was consequently utilized to derive the LULC classes for the TAITA project, as shown in Table 8. The LCCS is a comprehensive, standardized *a priori* classification system that can be used for any mapping exercise regardless of the methodology, scale, source material and geographic location (DiGregorio 2005), and is available at:

www.glcn-lccs.org & www.fao.org/docrep/008/y7220e/y7220e00.htm.

As well as logically guiding the user through the derivation of appropriate mutually exclusive LULC classes, as is shown in Table 8, the LCCS software generates unique codes (third column in Table 8) and Boolean formulas (fourth column in Table 8) for each class which allows other users to precisely reconstruct the detailed definitions utilized. This is very useful as previously it would not necessarily be certain what was meant when a map contained a LULC class name, such as 'Thicket'.

In LCCS terminology, a 'classifier' is one of many measurable diagnostic characteristics that are used in the definition of a land cover class, such as vegetation cover and height. Additionally, a 'modifier' is a further optional refinement to a classifier which helps specify the exact properties of a class. Thanks to the code and Boolean formula in LCCS, the user is free to call the class any general or colloquial name desired and retain interoperability. The classes listed in Table 8 were developed based on extensive inspection of the SPOT imagery and fieldwork knowledge to determine what was feasible to map. Note that burned areas, and clouds and cloud shadows are not included as possible classes in the LCCS software but they occur in the imagery and are consequently specified in the Taita Hills map legend.

Table 8. *LCCS nomenclature adopted for SPOT imagery LULC mapping of the Taita Hills.*

ID	User Land Cover Name	LCCS Code	LCCS Boolean Formula (Classifiers)
1	Cropland	11251 - 12699	A3B2XXC2D1 - C4C10C17C13C17
2	Shrubland (20% to 70% Cover)	20373	A4A11B3XXXXXXF1
3	Thicket (Closed Shrubland >70% Cover with Emergent Trees)	20354 - 13554	A4A10B3XXXXXXF2F5F10G2F1 - B9G7
4	Woodland	20013	A3A11
5	Plantation Forest	10001-S1002S1003W7	A1-S1002S1003W7
6	Broadleaved Closed Canopy Forest	20088 -13152	A3A10B2XXD1 - B5
7	Grassland with scattered shrubs and trees	20412 -104774	A2A10B4XXXXXXF2F5F10G2F2F6F10G3 - B12G7G9
8	Bare Soil & Other Unconsolidated Material	6005	A5
9	Built-up Area	5001	A1
10	Bare Rock	6002-1	A3 - A7
11	Water	8002-5	A1B1 - A5
12	Burned Area	*Not Available*	-
13	Cloud/Cloud Shadow	*Not Available*	-

4.7 Soil loss modelling using the Universal Soil Loss Equation (USLE)

In Paper V, multi-temporal land cover maps of 1987, 1992 and 2003, derived from the SPOT imagery using the MSS/ORM methodology as detailed in Paper IV, were utilized in combination with compiled geospatial layers of soil, rainfall and topography to model soil erosion in the Taita Hills. In a simple approach, based on the Universal Soil Loss Equation (USLE), potential soil loss was estimated quantitatively as (Wischmeier & Smith 1978; Lal 1996; Morgan & Davidson 1991):

$$A = R \cdot K \cdot L \cdot S \cdot C \cdot P \qquad (32)$$

where A is mean annual soil loss due to water erosion per unit area per year (t ha^{-1} yr^{-1}). R is rainfall erosivity factor (MJ mm ha^{-1} h^{-1} yr^{-1}), K is soil erodibility factor (t ha h ha^{-1} MJ^{-1} mm^{-1}), L is slope length factor, S is slope steepness factor, C is cover management factor, and P is support practice factor. Even though soil conservation practices do exist in the Taita Hills, it was not possible to assess the P-factor for the whole study area, and therefore it was not varied in the calculations.

In accordance with working in regions of the developing world with limited ancillary data availability, and considering the ability of the procedures to be replicated locally using a minimum of resources, low cost and freely available GIS data and a simplified methodology based on USLE, were utilized. A 20 m DEM was generated from cheaply obtained Survey of Kenya 1:50,000 scale topographic paper maps, see Section 4.1.5 above, and inexpensive local rainfall records of the Kenyan Meteorological Department were interpolated into 20 m grid layers, as detailed in Section 4.1.6 above, whilst the freely available 1:1,000,000 scale Explanatory Soil Map of Kenya was utilized as a generalized digital soil map. Although in this study the land cover was generated using the commercial Definiens eCognition software, there are freeware alternatives, as discussed in Section 4.6 above. All the required GIS processing for this simplified methodology could be carried out in the well-known GRASS (http://grass.itc.it/) freeware GIS.

Slope-length factor (LS) is calculated from the slope steepness (percentage) and length (m) of a given slope. In Paper V, the steepness layer was derived from the DEM and slope length was assumed to be a fixed 150 m for each pixel (Ogawa *et al.* 1997). The Exploratory Soil Map of Kenya at a scale of 1:1 000 000 (Sombroek *et al.* 1980) was used for the derivation of the soil erodibility factor (K-factor). The map contained 12 major soil mapping units using the FAO classification based on soil survey measurements. By assuming that the resolution of the K-factors matches that of the soil mapping units, the mean K-factors were obtained according to Kassam *et al.* (1991) for each soil class and their texture classes. The rainfall erosivity factor (R-factor), defined as the potential ability of rain to cause erosion and given as the product (EI_{30}) of the total energy of rainstorm (E) and the maximum 30-min intensity (I_{30}) (Wischmeier *et al.* 1958; Foster *et al.* 1981), is often determined from rainfall intensity. However, since detailed data on storm intensity was unavailable for the Taita Hills, it was necessary to use a proximate quadratic regression method based on annual rainfall (Millward & Mersey 1999). In the Taita Hills, R-factor (MJ mm ha^{-1} h^{-1} yr^{-1}) was determined using the annual rainfall amount (mm) as suggested by Morgan & Davidson (1991) as:

$$R = p \cdot 0.5 \qquad (33)$$

where p = mean annual rainfall (mm) and R = rainfall erosivity factor (MJ mm ha^{-1} h^{-1} yr^{-1}). Crop management factor (C) depends on the vegetation cover, which dissipates the kinetic energy of the raindrops before impacting the soil surface. Therefore, vegetation cover and cropping systems have a large influence on runoff and erosion rates. In principle, the more vegetation cover or LAI there is over the soil the more protective is the vegetation cover. C-factors were derived from the land cover layers produced for 1987, 1992 and 2003 using the values described by Wischmeier & Smith (1978). For a full description of the implemented simple methodology for estimating potential soil loss, refer to Paper V.

4.8 Human population modelling

To help manage population growth and the related environmental pressures, it is important to have precise estimates of the number of inhabitants and spatially explicit information of the human population distribution (Sutton *et al*. 1997). In developed countries, the population numbers and distribution are well known, and data collection and dissemination are computerized and accessible, but in many developing countries traditional census methods are still utilized. Such censuses are often time-consuming, costly, error-prone, difficult to update and the census interval is often too long for many applications (Li & Weng 2005). Therefore, in addition to conventional census calculations, other methods have been used to derive geospatial data on population. Aerial photography interpretation has been the traditional remote sensing based method to estimate population and to map the population distribution at local or regional scales. Large-scale aerial photographs have been used since the 1950s for dwelling unit counting, e.g. Porter (1956) used a rural dwelling unit count in Liberia, and Lo (1986) was able to estimate population for the city of Athens, Georgia USA, using aerial photographs at the census tract level with high accuracy. However, the aerial photography based dwelling unit count method is itself time-consuming and expensive, and problematically for developing countries where it is potentially most useful, it requires abundant up-to-date high resolution aerial photographs to cover large areas (Lo 1989).

For a relatively long period, therefore, satellite based remote sensing has been considered as a potentially adequate option for population estimation. A number of studies have reported satisfactory results by combining satellite data with regression techniques for population estimation, despite the limitations of the medium spatial resolutions of the sensors used. For example, Forster (1983) used Landsat MSS data and developed a multi-regression equation with standard deviation of separate bands and various reflectance ratios to predict housing density, and Lo (1995) linked radiance values of image pixels with residential densities for the Kowloon metropolitan area of Hong Kong using a multispectral SPOT image. In addition to local and regional scale population estimations, global scale population databases have been developed using geospatial data modelling to create gridded population models. The most commonly utilized global population models are the Gridded Population of the World (GPWv3) with 5 km resolution (CIESIN 2005), LandScan at 1 km resolution (Dobson *et al*. 2000) and the UNEP/Global Resource Information Database at 1–degree resolution (UNEP/GRID 2006). However, these existing global population datasets have limitations at a local or regional level as they are relatively too coarse in resolution and consequently generalize and obscure the internal variability of population when considered for use at these scales. Therefore, Paper VI concentrates on the development of an enhanced probability based application that can be used for population distribution mapping and for population abundance estimation at local or regional scales.

In Paper VI, therefore, human population distribution and abundance in the Taita Hills were modelled using the Generalized Regression Analysis and Spatial Prediction (GRASP) modelling framework, which uses generalized additive models (GAMs) for model calibration (Lehmann *et al*. 2002). The response variable (dwelling unit presence - absence) was derived from airborne imagery mosaics covering ca. 30% of Taita Hills through on-screen digitizing in GIS. Geospatial GIS- and remote sensing-based map layers were used as predictors. Prior to modelling, the full data set (n = 10488, 100 m analysis squares) was randomly divided into model

calibration 70% (n = 7342) and model evaluation 30% (n = 3146) datasets following the split sample approach (Guisan & Zimmermann 2000). The outcome of the dwelling unit prediction models were extrapolated to cover the whole Taita Hills area (n = 34143, 100 m analysis squares) and a human population distribution map was created for the Taita Hills semi-automatically. Additionally, the human population abundance model was compared with two existing global population datasets, GPWv3 and LandScan 2005, and the Kenyan census data for 1999. Refer to Paper VI for a full description of the applied predictive modelling methodology used in Taita Hills dwelling unit occurrence and abundance prediction. Figure 9 below details the methodological flowchart of the modelling steps.

4.8.1 Applied statistical model calibration techniques

Linear regression techniques have been traditionally used in predictive modelling, e.g. in ecological and biogeographical research (Guisan & Zimmermann 2000). The basic linear regression model has the form:

$$Y = \alpha + X^T\beta + \varepsilon \quad (34)$$

where Y is the response variable, α is a constant called the intercept and $X = (X_1, ..., X_p)$ is a vector of p predictor variables, $\beta = \{\beta_1, ..., \beta_p\}$ is the vector of p regression coefficients (one for each predictor), and ε is the error. However, when using linear regression several assumptions should be met: (i) linearity of the relationship between dependent and independent variables, (ii) independence of the errors, (iii) homoscedasticity (constant variance) of the errors, and (iv) normality of the error distribution (Zar 1999). Violation of one or more of these multiple linear regression assumptions may lead to incorrect or misleading results. However, these assumptions are seldom all met when utilizing geospatial datasets. To overcome these violations, new prediction methods have been developed which allow non-Gaussian error distributions and non-linear relationships between response and predictor variables (Guisan & Zimmermann 2000). In modelling human population distribution and abundance in the Taita Hills in Paper VI, use was made of GAMs (Hastie & Tibshirani 1990), which are non-parametric extensions of generalized linear models (GLMs) (McCullagh & Nelder 1989).

4.8.2 Generalized linear models (GLM) and generalized additive models (GAM)

Generalized linear models (GLMs) are an extension of classical multivariate linear regression, allowing non-normal response variables to be modelled (McCullagh & Nelder 1989). In GLMs, the predictor variables X_j (j = 1, ..., p) are combined to produce a linear predictor LP which is related to the expected value $\mu = E(Y)$ of the response variable Y through a link function $g()$, such as:

$$g(E(Y)) = LP = \alpha + X^T\beta, \quad (35)$$

where α, X, β are those described in Equation 34 above. The model is written for the generic variables X and Y; the corresponding terms for the ith observation in the sample is:

$$g(\mu_i) = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} \quad (36)$$

Unlike classical linear models, which presuppose a Gaussian (i.e. normal) distribution and an identity link, the distribution of Y in GLMs may be any of the exponential family distributions (e.g., Gaussian, Poisson or binomial) and the link function may be any monotonic differentiable function (like logarithm or logit). GLMs do not force data into unnatural scales and therefore they allow non-linearity and non-constant variance structure in the data (McCullagh & Nelder 1989; Collet 2003). GLM models are usually built using a full stepwise approach, in which explanatory variables are included or excluded from the full model using Akaike Information Criterion (AIC) (Akaike 1974) and changes in scaled deviance (McCullagh & Nelder 1989; Venables & Ripley 2002).

Generalized additive models GAMs (Hastie & Tibshirani 1990; Yee & Mitchell 1991) support non-Gaussian error distributions and non-linear relationships between response and predictor variables. GAMs are non-parametric extensions of GLMs model regressions that apply nonparametric smoothers to each predictor and additively calculate the component response. GAMs are data-driven rather than model driven and allow consideration of more complex response shapes than those possible through GLMs (Yee & Mitchell 1991). A GAM model is expressed by:

$$g(E(Y)) = \alpha + s_1(X_{1i}) + s_2(X_{2i}) + \ldots s_p(X_{pi}) \quad (37)$$

where g is the link function that relates the linear predictor with the expected value of the response variable Y, X_{pi} is a predictor variable and s_p a smoothing function. In Paper VI the response variable was linked to the set of predictor variables through a logit link function for dwelling unit occurrence and log link function for dwelling unit abundance.

4.8.3 Evaluation of predictive models

In Paper VI the predictive models were evaluated as follows: (i) by using the percentage of explained deviance as an indicator of model explanatory power (D^2), obtained by dividing the difference between null and residual deviance by the null deviance (Guisan & Zimmerman 2000); (ii) by using the area under the curve (AUC) from the receiver operating characteristic plot to indicate the model predictive power (ROC, Fielding & Bell 1997). As a general rule, an AUC between 0.5–0.7 indicates a poor discriminate capacity, 0.7–0.9 indicates reasonable capacity, and 0.9 or higher indicates a very good capacity (Swets 1998); (iii) calculating Cohen's Kappa (Cohen 1960) and cross-validated Kappa; (iv) calculating the contribution for each predictor, giving an indication of the contribution of the variable within the selected model and corresponding to the possible range of variation on the scale of the linear predictor; and also (v) visual interpretation of the prediction maps.

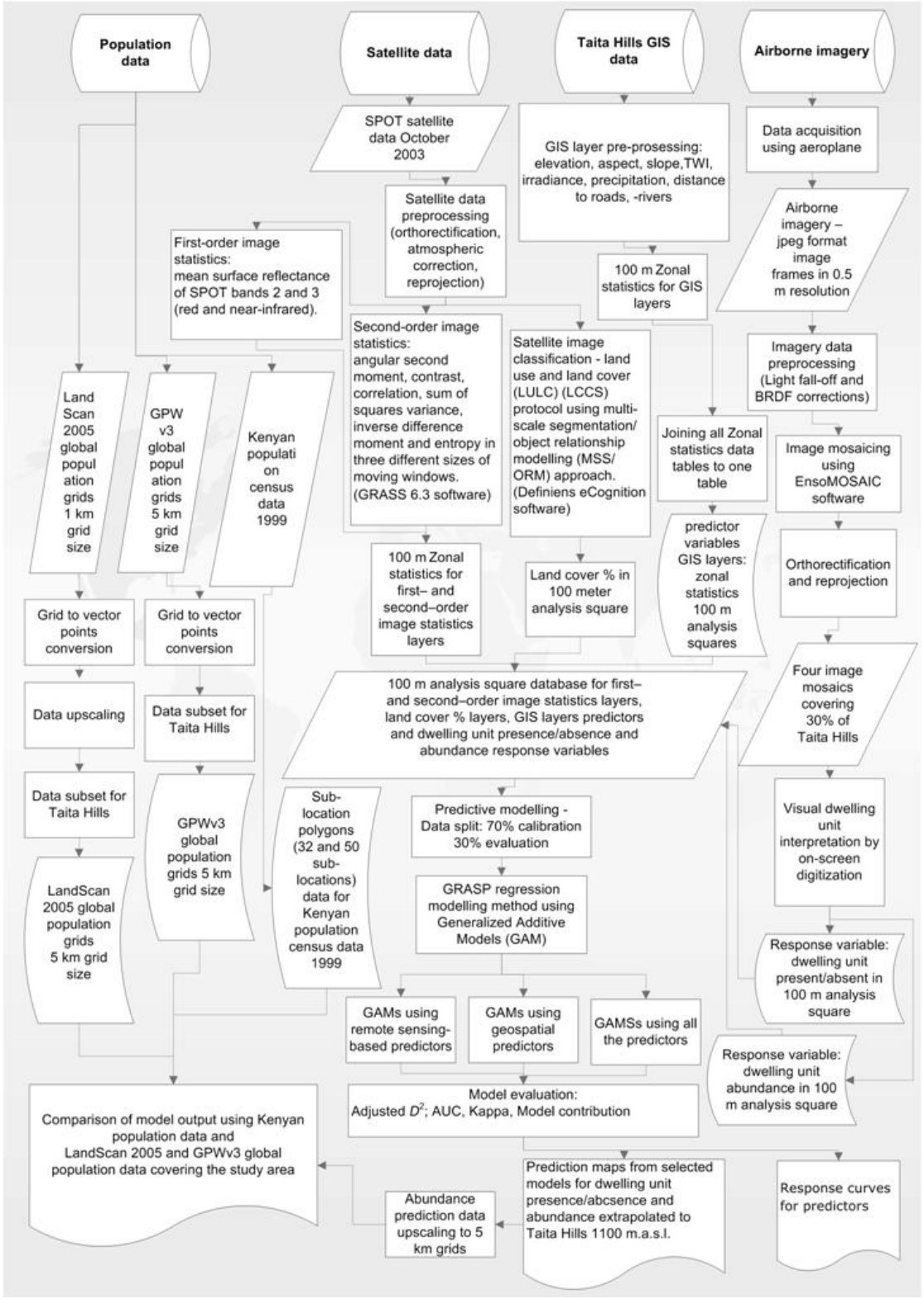


Figure 9. Methodological flowchart of the modelling steps for human population distribution and abundance modelling (source: Siljander 2010).

Consideration should also be given to the ability of the procedures to be replicated by local specialists in the developing world using a minimum of resources,

low cost or freely available data and software, and a reproducible methodology. In the Paper VI study, GRASP was implemented with commercial software S-PLUS 6.2 (Insightful Corp.), but there is also a GRASP package for the freeware R-program (R Development Core Team 2008: http://www.r-project.org/). Similarly, the required GIS processing for the creation of geospatial datasets could be carried out in the GRASS (http://grass.itc.it/) freeware GIS.

5. RESULTS AND DISCUSSION

5.1 Selection of appropriate HELM calibration targets

As noted by Moran *et al*. (1995), the accuracy of calibrations to field measured ρ_s and of validation based on these targets depends critically on the accuracy of ρ_s measurements and the derived BRDF estimates, and errors are minimized for targets that are spectrally bright and uniform with near-Lambertian scattering properties. Gu *et al*. (1992) evaluated the BRDF of the La Crau SPOT calibration site in France and found that reflectance measurements could vary by ± 10% over a range of viewing angles from 0° to 30°. Similarly, based on the BRDF of gypsum sand measured at the White Sands site in New Mexico, USA, Jackson *et al*. (1990) found variation in the BRF from nadir of ± 10% for SPOT maximum off-nadir views.

Based on goniometer ρ_s measurements taken in the Helsinki metropolitan region control site study area, Paper I examined the spectral spatial, temporal, and multiangular reflectance characteristics of various potential pseudo-invariant surface types suitable for use in the application of empirical line (EL) spectral calibration techniques, such as HELM, to SPOT imagery. As detailed in Table 5 above, the surface types consisted of asphalts, sands, gravel, grass, and fake turf targets. Consideration was given to the physical causes of the different behaviours of the measured surfaces, and the most appropriate surface types and the requirements for suitable site selection for EL calibrations, were discussed.

The Paper I results showed that all real world surfaces exhibit some degree of non-Lambertian multiangular scattering properties, particularly along the solar principal plane, as well as spatial and temporal spectral variability. This has implications for the accurate application of EL spectral calibration procedures for optical satellite imagery, especially data like SPOT that can be collected with off-nadir view angles. As the multispectral spatial resolution of high resolution next generation civilian satellites continues to improve, so too does the agility and off-nadir viewing capabilities of the sensors. This has two implications: firstly, the size requirement for a calibration target is reduced and their likely occurrence increases, so selection therefore becomes easier. This is likely to boost interest in the application of EL approaches in future. For example, the forthcoming SPOT 6 and SPOT 7 satellites have a planned 8 m multispectral spatial resolution, which would give a minimum calibration target size requirement of 64 by 64 m.

Secondly, however, the multiangular reflectance properties of the calibration site become increasingly important and can have a significant impact on the error in EL calibrations if not accounted for properly. For example, the forthcoming PLEIADES mini-satellites program, due for launch in 2010, will collect visible/NIR data at 2.8 m nadir resolution and will have an off-nadir viewing capability of ± 45° in any direction. Based on the field measurements made even for pseudo-invariant targets, at such large off-nadir viewing angles the difference in reflectance to nadir are very likely to be significant.

Ideally, therefore, the multiangular reflectance characteristics of a calibration target should be captured using a goniometer. In the HELM atmospheric correction, developed specifically for SPOT data, this enables the derivation of spectral band specific ρ_s equivalent to geometrical circumstances of existing and future SPOT imagery within a database, which can be used for a HELM-2 calibration for RFR. Currently, however, it will more usually be necessary to undertake a HELM-1 calibration to nadir ρ_s field measurements, where it is imperative that appropriate pseudo-invariant ground targets are carefully selected. Of the sand, gravel, asphalt, grass turf, and artificial turf surface types investigated in Paper I, a medium grained beach sand offered the most appropriate characteristics for use in the HELM-1 or HELM-2 absolute atmospheric correction of multi-temporal SPOT imagery.

Vegetation-free sites make better calibration targets as the multiangular reflectance properties, within the ± 30° θ_{VZ} range relating to the SPOT sensors, are generally more limited and are similar throughout the spectrum, unlike vegetation; compare Figures 10 and 11 below. The Hietsu beach sand was found to be the most spectrally stable site over time as here weathering changes are a very long term process. The most desirable calibration target sand properties are a well-sorted and finer grained material because this will have less surface roughness, and consequently less sensitivity to θ_Z and θ_{VZ} variations, and lesser multiangular ρ_s anisotropy. Further, it is also likely to have increased volumetric scattering leading to a brighter spectrum, which is useful in better determining HELM correction lines.

New asphalt is too spectrally dark to be a calibration target, but weathers quickly so that within the space of a few years the ρ_s spectrum resembles that of the exposed constituent gravel aggregates (Paper I: Figures 7 & 8 (D)). The presence of structural damage from usage will further increase multiangular ρ_s anisotropy, because of a consequent significant increase in surface roughening and the increased presence of shadowing. As well as temporal variance, asphalt use can be problematic because of the presence of painted markings and motor vehicles, as well as patch repairs, which all increase spatial spectral heterogeneity (Paper I: Figure 6).

Results from the site 6 grass turf showed that vegetation is too spectrally dark in the visible bands, and has too strong wavelength dependent multiangular ρ_s properties to be a viable HELM calibration target, see Figure 10, even if managed sites may actually be relatively spatially and temporally spectrally stable (Paper I: Figure 11 (B) & (C)). However, a managed vegetated site may be useful as a validation target for the NIR band because of its high ρ_s if field measurements are made very close to the imagery acquisition time. The same could also be said for asphalt targets.

Where angular ρ_s field measurements are unavailable, it is unwise to attempt EL calibrations to nadir ρ_s for SPOT imagery with viewing geometries near to the solar principal plane, as here multiangular ρ_s variations will be the greatest. In all cases, field measurements should be made within the θ_Z range within which SPOT overpass will occur, and it is recommended that detailed knowledge and measurements of the calibration site be collected and a minimum of assumptions be made. It should be noted that these findings are also applicable to any EL atmospheric correction methodology that is applied to SPOT data, or any other medium or high resolution optical remote sensing imagery.

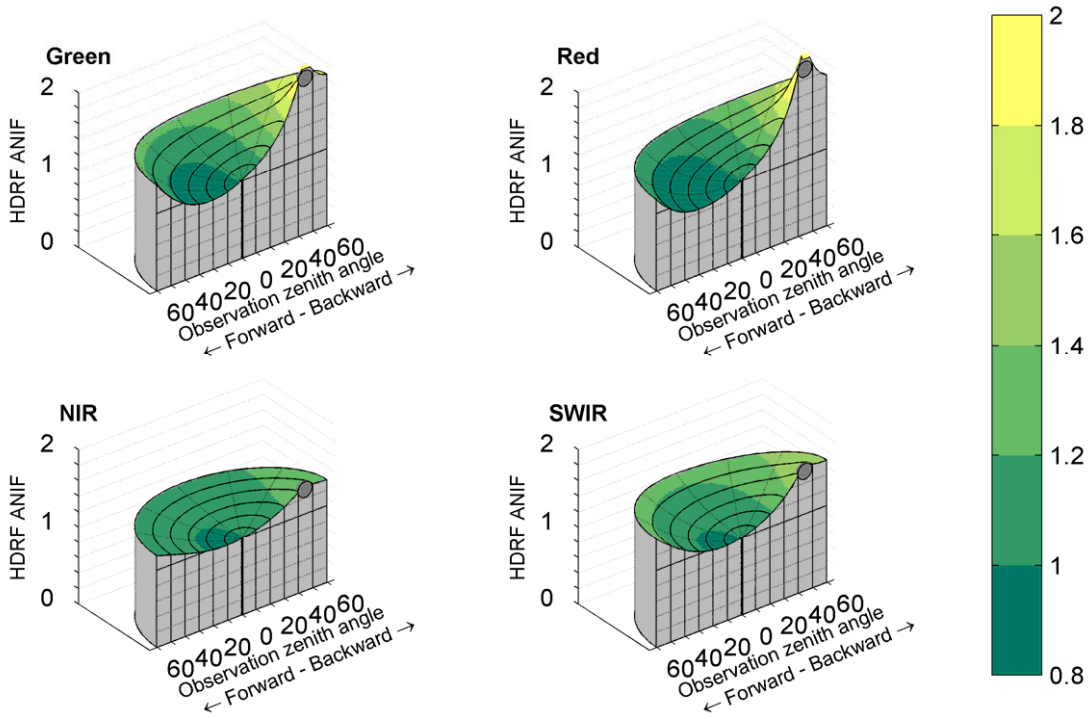


Figure 10. Site 6 managed grass turf SPOT spectral band specific HDRF anisotropy factor ($HDRF_{ANIF}$); mean $\theta_Z = 44°$.

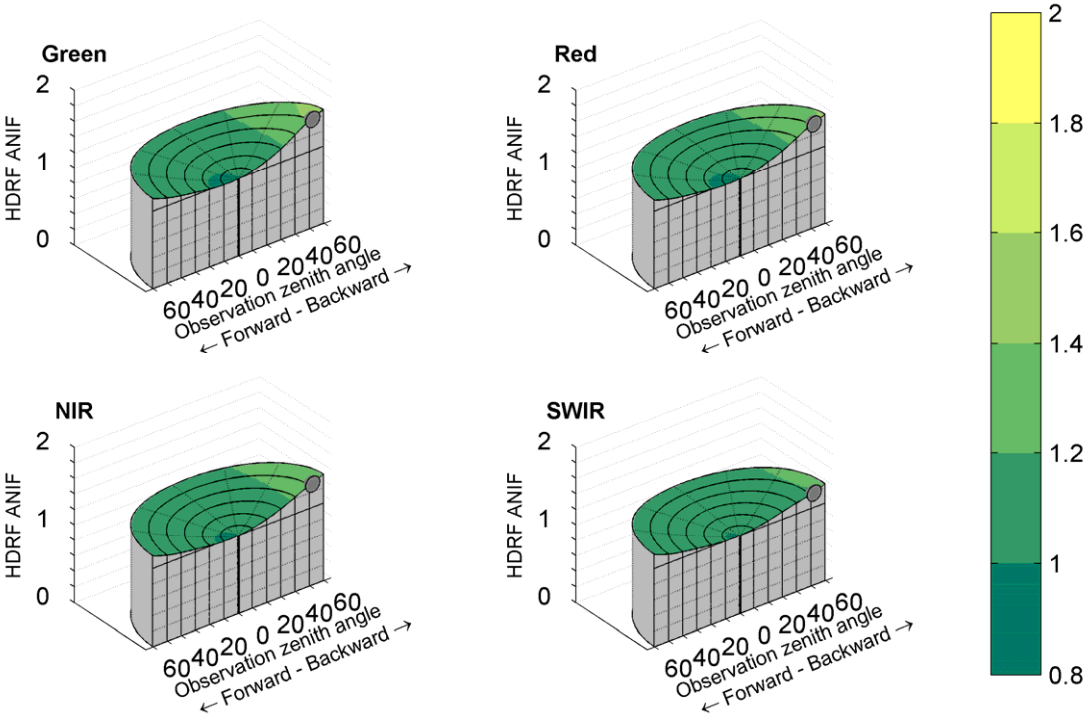


Figure 11. Site 1 Hietsu beach medium sand SPOT spectral band specific $HDRF_{ANIF}$; mean $\theta_Z = 51°$.

5.2 HELM calibration error

Paper II outlined the proposed HELM absolute atmospheric correction technique for the retrieval of ρ_s from multi-temporal SPOT multispectral data. Calibration to nadir only ρ_s is denoted as the HELM-1 approach, whilst calibration to ρ_s modelling the exact illumination and view geometries of the SPOT imagery is termed the HELM-2 approach. In Paper II the application of HELM-1 and HELM-2 to the Helsinki metropolitan region control site Dataset 1 SPOT imagery, and the application of HELM-1 to the Taita Hills application study area Dataset 2 SPOT imagery, were detailed and the RFR accuracy results were compared to the partially corrected ρ_{SAT} (Paper II: Sections 3.2 & 3.3). Given the ± 31° range in the SPOT sensor view incidence angle (θ_V), a general error model for HELM-1 calibration to nadir ρ_s for SPOT imagery was derived based on FIGIFIGO measurements from the Helsinki control site study area. Further, the estimated calibration error in applying HELM-1 to the control site Dataset 1 SPOT imagery was modelled based on the multiangular ρ_s behaviour of the site 1 beach sand calibration target (Paper II: Section 3.1).

5.2.1 HELM-1 nadir calibration error estimates

Given the stated ± 0.02 ρ_s VIS/NIR benchmark for successful atmospheric correction, HELM-1 nadir calibration uncertainty also needs to be within this range. The actual error encountered will depend on the multiangular ρ_s characteristics of the utilized surface and the SPOT image geometry, as covered in detail in Paper I. Based on an assessment of the Helsinki control site study area FIGIFIGO measurements for the non-vegetated site 1 sand and the site 2 and 5 asphalts, a general error model for HELM-1 calibration to nadir ρ_s for SPOT imagery was identified in Paper II as ≤ 2% (0.02) ρ_s across all bands. Imagery viewing in the forward scattering direction is likely to have very small calibration error, even at maximum 31° θ_V in the principal plane. However, as illustrated in Figure 12 below, the exception is when θ_V is ≥ R20° in the backscattering direction within ± 55° azimuth of the principal plane, assuming azimuthal symmetry in the calibration target multiangular ρ_s behaviour. This relates to a φ_r range of 110° between 125°–235°. SPOT imagery with viewing geometry falling within these limits could be HELM-1 atmospherically corrected, but the error in the calibration may then exceed 0.02 ρ_s. Alternatively, it is possible to avoid utilizing such scenes by establishing the illumination and view geometries in the SPOT SIRIUS online image catalogue before ordering. These HELM-1 limitations demonstrate that, if at all possible, multiangular ρ_s measurements of the calibration target (at least along the principal plane in the backscattering direction) should be collected and are useful information. Nevertheless, for a large range of view geometries, the nadir calibration error for targets with near-Lambertian reflectance properties is likely to be ≤ 0.02 ρ_s.

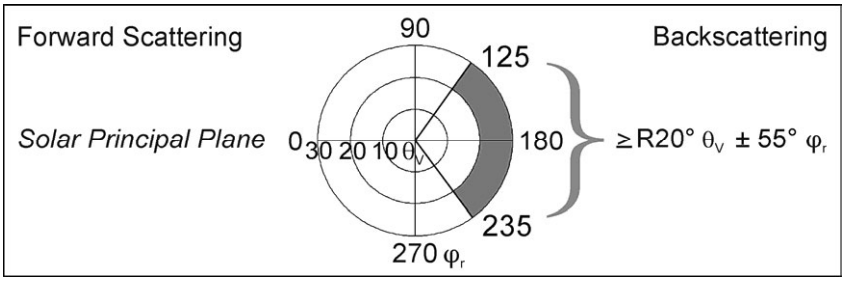


Figure 12. Generalized error model for the application of HELM-1 nadir calibration to SPOT imagery with a possible ± 31° range in view incidence angles (θ_V). Grey shading indicates viewing angles where the difference between nadir ρ_s and angular ρ_s may exceed 0.02 at backscatter viewing $\theta_V \geq 20°$ within ± 55° azimuth of the solar principal plane, assuming azimuthal symmetry in the target multiangular characteristics. Because of the SPOT satellites' orbit geometry, imagery viewing the backscattering will always be denoted in the metadata by R (right, negative, east) θ_V.

The calibration error in applying HELM-1 to the control site Dataset 1 SPOT imagery was estimated from the modelled geometric ρ_s derived for the site 1 beach sand calibration target from the FIGIFIGO measurements. As shown in Paper II Table 6, the overall average calibration RMSE for all bands and all years was 0.014 ρ_s and, except for the 2002 image, band specific angular ρ_s was always greater than the nadir ρ_s; i.e. nadir calibration was always an underestimation. The band specific RMSE for all years also increased with wavelength. For the forward scattering viewing images, average RMSE for all bands and all years was 0.008, but for backscattering viewing images 0.019. As expected, the largest differences between the nadir and image geometry ρ_s occurred in the 2003 scene, which had a R28.7° θ_V at 43.8° θ_Z, with a band specific error of -0.019 in the green, -0.023 in the red and NIR, and -0.029 in the SWIR. In contrast, although the 1994 scene had a L29.7° θ_V at 43.5° θ_Z, because of the forward scattering view the RMSE was -0.011 in the green, -0.014 in the red, and -0.015 in the NIR. The 2005 image was viewing the backscattering at R16.1° θ_V, and had an RMSE for all bands of 0.012. There was negligible error for the 1993 scene, which was viewing the forward scattering at L9.3° θ_V, and the nadir and geometric ρ_s were essentially identical for the 2002 scene because the image was viewing the forward scattering at only L2.5° θ_V.

5.2.2 Surface reflectance factor prediction accuracy

The application of HELM to SPOT imagery Datasets 1 and 2 covering the Taita Hills and the Helsinki metropolitan area is detailed in Paper II, and the ρ_s retrieval accuracy results are compared to the partially corrected ρ_{SAT}. Application of HELM-1 and HELM-2 to the Dataset 1 2002, 2005 and 2003 images allowed for the assessment of RFR accuracy from scenes with similar θ_Z, but variations in θ_V of L2.5°, R16.1°and R28.7°. These scenes were used as they occurred closest in time to field measurements of the validation targets, taken over the summers of 2005–2007 (Table 5). The HELM procedures were also successfully applied to the 1994/1993 images. However, the long length of time from field measurement added too much uncertainty to include them into the accuracy assessment. Overall, as can be seen from Paper II Tables 7, 8 and 9, examination of the results revealed several main points. Firstly, as expected, there was an increase in HELM-1 RFR error with an increase in θ_V in the backscattering direction. Secondly, contrary to expectation, the nadir calibrated

HELM-1 approach derived slightly higher accuracy RFR than the HELM-2 geometric calibration. Thirdly, HELM-1 VIS/NIR RFR RMSE was within 0.02 ρ_s for all scenes and bands, except 0.21 for the 2003 image red band. However, HELM-2 VIS/NIR RFR RMSE exceeded 0.02 ρ_s in all bands in the 2003 image. Fourthly, for HELM-1 the overall relative RMSE ($RMSE_r$) was 10% or better, with the exception of the red band which was slightly higher at 13.5%. The HELM-2 overall relative $RMSE_r$ were higher and exceeded 10% in both the red (15.9%) and SWIR (13.3%) bands (Table 7). In both cases the NIR $RMSE_r$ was substantially lower. These values are mostly near the desired 10% $RMSE_r$ and are within the ~10–15% range. Fifthly, and most importantly, both HELM-1 and HELM-2 derived substantially more accurate RFR than the ρ_{SAT} estimates, most significantly in the SWIR band where this was the basis of successful RFR.

The reason the HELM-1 approach derived slightly higher accuracy RFR than the HELM-2 geometric calibration was because the Hietsu calibration target ρ_s were slightly higher than those of the validation targets, relative to the input L_{SAT} values (Paper II: Figure 9 [A]). As a result, calibration to nadir ρ_s already led to HELM-1 correction lines that slightly overestimated the modelled image geometry ρ_s for the validation targets, even if the correction lines were a slight underestimation of the calibration target geometric reflectance itself. As detailed in Paper II Section 3.1, Hietsu beach sand ρ_s increased with increasing off-nadir θ_V in the backscattering direction. Consequently, as the L_p estimates remained the same, calibrating HELM-2 to higher image geometry ρ_s increased the slope of the correction lines slightly and actually made the RFR overestimation slightly worse in this instance. There was undoubtedly uncertainty and error propagation in predicting image geometry specific ρ_s of targets from fitted models, and matching them to a single linear prediction line for each band, but this alone may not have accounted for the observed HELM overestimation. By way of scattering, atmospheric aerosols reduce the apparent reflectance of bright targets and increase it for dark-objects leading to a loss of information (Song *et al*. 2001). This lost information cannot be fully recovered by a HELM correction because the number of levels of L_{SAT} values will not exceed that originally captured in the image DN. Furthermore, the spectrally bright beach sand was surrounded by dark sea water on two sides. It was considered that this derived a strong adjacency effect, giving lower L_{SAT} relative to the targets completely over land where greater scattering was likely. Because of short path lengths, adjacency effects in ρ_s field measurements are negligible (Richter *et al*. 2006). Consequently, site 1 L_{SAT} were relatively lower and derived higher ρ_s estimates (Paper II: Figure 9 [A]).

Nevertheless, overall, Dataset 1 results indicated Hietsu beach sand was a suitable HELM calibration target, even if HELM-1 predictions were actually better than estimates based on the modelled image geometry ρ_s. As illustrated in Figure 12 above, the Dataset 1 imagery viewing geometries were within the ranges identified as being suitable for application of HELM-1 with minimal error. In circumstances where SPOT image geometry is viewing backscattering at near-maximum off-nadir θ_V at relative azimuth angles nearer to the principal plane than was the case in Dataset 1 (~50°–58°), it is likely that there will be variation between nadir and geometric ρ_s > 0.02. In such instances, HELM-1 calibration error becomes significant and HELM-2 RFR performance should be better. Further, where the calibration target ρ_s is not relatively brighter than the validation sites, HELM-2 should give improved RFR over HELM-1 in any case. It should be remembered, however, that the difference between HELM-1 and HELM-2 RFR was generally very small, ~0.01 ρ_s in all bands. Moreover, both HELM-1 and HELM-2 gave significantly better ρ_s estimates than

ρ_{SAT}, indicating they were effective in reducing atmospheric effects in the SPOT scenes. Based on the Dataset 1 RFR results, a generalized statement was made in Paper II that HELM-1 performance was ± 0.02 ρ_s in the VIS/NIR and ± 0.03 ρ_s in the SWIR, whilst HELM-2 performance was ± 0.03 ρ_s in the VIS/NIR and ± 0.04 ρ_s in the SWIR. This represented band specific overall relative errors of ~10–15%.

As can be seen from Paper II Table 10, nadir calibrated HELM-1 prediction accuracy results for Dataset 2 were similar to those for Dataset 1. Based on the 2003 scene with an off-nadir θ_V of R10.4°, HELM-1 derived an average VIS/NIR RMSE of 0.018 in predicting ρ_s for the three verification targets. As with Dataset 1, this value is within the desired 0.02 ρ_s benchmark and 10% relative accuracy. The partially corrected ρ_{SAT} gave a higher average RMSE of 0.039. There was no SWIR assessment as an ASD FieldSpec® Handheld VNIR spectrometer was used in Taita Hills. Overall, then, both sets of RFR results indicated HELM was effective in reducing atmospheric effects in SPOT data relative to ρ_{SAT}.

5.3 Comparisons of atmospheric correction procedures for SPOT data

The study reported in Paper III utilized field measured nadir ρ_s collected in the Taita Hills application site (see Table 6), as well FIGIFIGO measured multiangular ρ_s field data from the Helsinki metropolitan region control site (see Table 5), as validation data to undertake a comparative assessment of the absolute atmospheric correction techniques for SPOT data applicable in local and regional landscape level remote sensing studies, in circumstances where no detailed overpass concurrent atmospheric measurements or meteorological data are available. Namely, in addition to HELM, which was applied as a HELM-1 nadir calibration, the four DOS and the 6S RTM absolute atmospheric correction methods were applied to the multi-temporal SPOT imagery databases. Performance was assessed both on the ability of each methodology to accurately retrieve ρ_s and also to provide radiometric stability within the multi-temporal datasets. 6S was applied using modelled atmospheres and general estimates of AOD at 0.55 μm for the Taita Hills dataset, but was applied using horizontal visibility meteorological data for the Helsinki control dataset.

5.3.1 Surface reflectance factor retrieval accuracy

Considering the application of the atmospheric correction methodologies to the Helsinki metropolitan region control site Dataset 1 SPOT imagery, the overall average RFR accuracy assessment results for the 2005, 2003 and 2002 SPOT scenes are detailed in Paper III Table 10. As noted in Table 4 above, the 2002 image suffered problems with sensor calibration in the red band, which affected L_{SAT} values for all validation sites, and saturation in the green, red, and SWIR bands, which affected several of the spectrally brighter validation targets as well as the Hietsu beach calibration site for HELM. Consequently, the individual RFR accuracy results for 2002 were not reported. Nevertheless, the averaged results were coherent and consistent with the results obtained individually from the 2003 (Paper III: Table 11) and 2005 (Paper III: Table 12) scenes. It can be noted from Tables 11 and 12 in Paper III that results for the 2005 and 2003 images were similar in that they derived the same ranked order in the RFR accuracy performance of the applied techniques: HELM, DOS2 (COST), 6S Met Data, 6S AOD 0.17, DOS4, DOS3, and DOS1.

As can be seen from Paper III Table 10, and Figure 13 (a) and Figure 14 (a) below, based on the 2005, 2003, and 2002 control site scenes, HELM derived the least overall RMSE in predicting ρ_s for the six validation targets, with an overall average RMSE for all SPOT spectral bands of 0.017, a VIS/NIR value of 0.014 and a SWIR value of 0.025. This VIS/NIR error value is within the desired 0.02 ρ_s benchmark. Also, HELM SWIR performance was significantly better than ρ_{SAT} (0.064), which was taken as the basis of successful SWIR correction. The average overall $RMSE_r$ for all bands was 8.5%, and the band specific overall $RMSE_r$ were ~10% or less for all bands, with the NIR $RMSE_r$ being the lowest.

Most importantly, none of the other applied methodologies were within the desired VIS/NIR absolute accuracy limit or equalled HELM SWIR performance. Next most accurate were DOS2 (COST), with an overall average RMSE for all bands of 0.034, and the 6S approaches, both with an overall average RMSE of 0.036. DOS1 was the least accurate method, with an overall RMSE of 0.058, which was worse than applying no atmospheric correction at all, as ρ_{SAT} gave an overall RMSE of 0.047. DOS3, with an overall RMSE of 0.049, was also slightly worse than ρ_{SAT}, whilst DOS4, with an RMSE of 0.046, was marginally better than ρ_{SAT}.

Additionally, as illustrated in Figure 13 (b) below, HELM had a small overall average bias (0.005), although DOS2 (COST) was the smallest overall (-0.004). Moreover, HELM was the only method that did not underestimate ρ_s overall. DOS1, DOS3 and DOS4 all had larger negative overall biases than the -0.021 of ρ_{SAT}, although the bias of the 6S approaches were smaller but still negative, with both deriving -0.018.

Considering each spectral band separately, it can be seen from Paper III Table 10 and Figure 14 below that HELM derived the lowest average RMSE in every spectral band, and of these the average SWIR RMSE was the highest, as was the case with DOS2, DOS3 and DOS4 and ρ_{SAT}. Futher, ρ_{SAT} had negative bias except in the green band, caused by the greater amount of scattering at shorter wavelengths. DOS1, 3 and 4, and the 6S approaches had negative bias in all bands, whilst DOS2 had overall positive bias in the red and NIR. HELM had a slight negative bias of -0.001 in the green band, no bias in the NIR, and positive bias in the red and SWIR.

The results of the Paper III atmospheric correction comparisons showed that all the applied methods, except HELM consistently underestimated ρ_s for both the Taita Hills and the Helsinki study areas (Paper III: Tables 10, 11, 12 and 13). Furthermore, in SPOT data with near-maximum off-nadir view incidence angles (± 31° θ_V), RFR error increased for all the applied corrections. Most importantly, HELM derived the best and most consistent RFR performance, and was the only approach that achieved RFR in the VIS/NIR bands with a RMSE within the desired 0.02 ρ_s benchmark. Also, HELM SWIR performance was significantly better than the other techniques as well as the partially corrected ρ_{SAT}.

DOS1 corrections did not account for the multiplicative effects of atmospheric attenuation and consequently, following subtraction of L_P for the VIS/NIR bands, ρ_s was consistently underestimated and led to less accurate RFR than ρ_{SAT}. Because the applied DOS2 (COST) method derives estimates of atmospheric transmittances based on the cosine of θ_V and the solar zenith angle (θ_Z), its RFR accuracy was found to be too dependent on the SPOT scene illumination and view geometries, which are not related to atmospheric conditions at the time of image acquisition. It is a well-known problem with COST that the cosine function for T_Z generally underestimates T_Z for larger θ_Z, while overestimating T_Z for smaller θ_Z (Wu *et al*. 2005; Chavez 1996). DOS3 models a Rayleigh scattering only atmosphere and does not consider the role of

aerosols, which were clearly demonstrated in Paper III as fundamental in determining the radiometric attenuation properties of a real atmosphere. Consequently, DOS3 ρ_s was consistently underestimated for all the spectral bands. DOS4 derived inconsistent RFR results, and was found to be sensitive to the estimates of L_P derived from the identification of within scene dark-objects. Application of 6S using standard atmosphere and aerosol models, even with the utilization of horizontal visibility meteorological data, although better overall than the DOS methods, was not accurate enough to meet the desired RFR accuracy requirements.

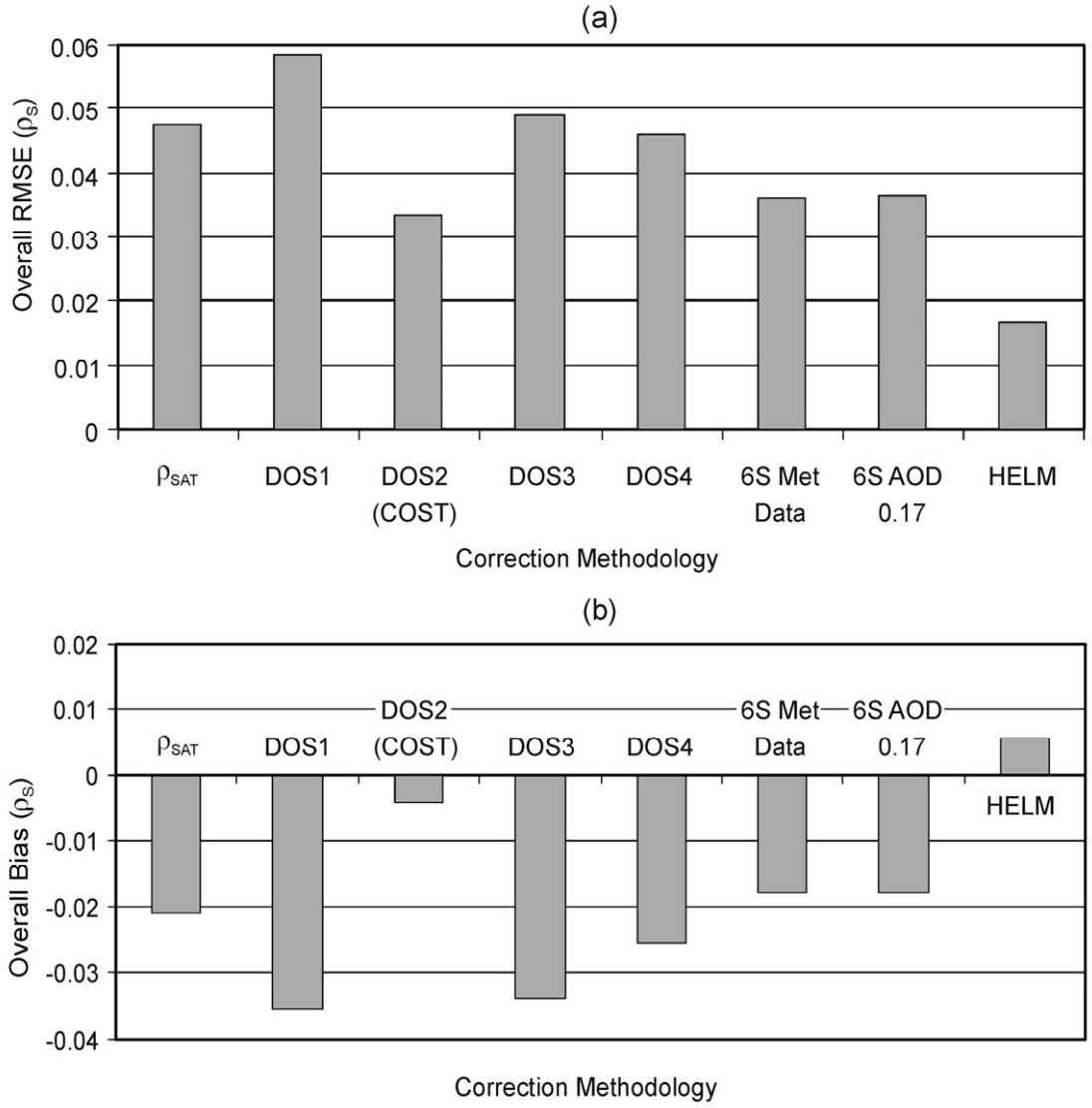


Figure 13. (a) Surface reflectance factor (ρ_s) overall average RMSE and (b) ρ_s overall average bias of all spectral bands for each atmospheric correction methodology applied to the 2005, 2003 and 2002 SPOT scenes covering the Helsinki metropolitan region control site study area.

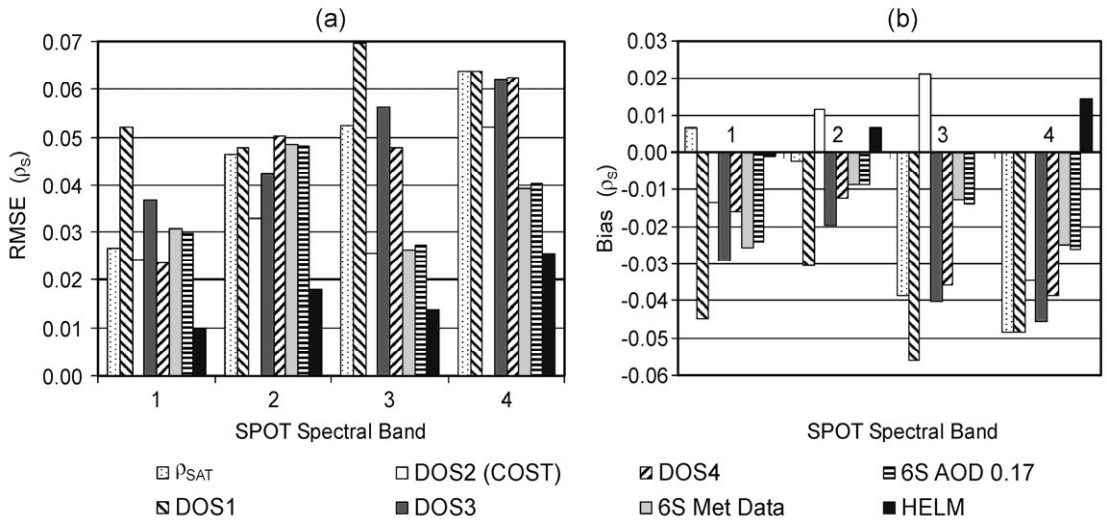


Figure 14. (a) Average RMSE (ρ_s) and (b) average bias (ρ_s) of each SPOT spectral bands for each atmospheric correction methodology applied to the 2005, 2003 and 2002 SPOT scenes covering the Helsinki metropolitan region control site study area.

5.3.2 Time series stability

As noted by Schroeder *et al*. (2006), theoretically the removal of atmospheric effects should reduce spectral variation between satellite images in a multi-temporal dataset. In a study by Song *et al*. (2001), DOS3, closely followed by DOS1 and a relative normalization method, were found to be the best techniques in minimizing reflectance variations within a multi-temporal Landsat dataset, compared to DOS2 (COST), DOS4, and dark dense vegetation (DDV) approaches. Similarly, Schroeder *et al*. (2006) found that relative normalization derived the greatest reflectance stability within a multi-temporal Landsat dataset compared to DOS3, 6S, and DDV absolute correction approaches and ρ_{SAT}. As neither study had field measured ρ_s data, they were unable to assess absolute atmospheric correction accuracy, but both showed that the more complicated correction methodologies actually increased the spectral variation between images.

Similar results were observed in the results reported in Paper III. As can be seen from Paper III Table 14, overall HELM derived the lowest sum total variation in average ρ_s between the multi-temporal SPOT scenes for Datasets 1 and 2 combined. However, considering each dataset separately, HELM showed a mixed ability to reduce variation compared to ρ_{SAT}. Whilst in Dataset 2 the mean average absolute difference for all bands combined was 2.78 (% ρ_s) for ρ_{SAT} and 2.23 for HELM, in Dataset 1 it was 2.43 for ρ_{SAT} and 2.60 for HELM. ρ_{SAT} derived lower variation than the DOS and 6S approaches in both Dataset 1 and 2, except for DOS1 in Dataset 1. Overall, therefore, it can be seen that the more complicated methods in predicting ρ_s did not lead to reduced variability between scenes, with DOS1 and DOS3 giving lesser variation than 6S or DOS4, in both Dataset 1 and 2. It should also be noted, however, that the differences between the methods were very small and have to be reported to $1/100^{th}$ of 1% in order to differentiate between the approaches.

Considering the individual spectral bands for the control site Dataset 1 imagery, as can be seen in Figure 15 below, the largest variation in average ρ_s between scenes was in the NIR and the least variation in the green for all correction methods. This is

most likely because the NIR, although it has little atmospheric scattering, is directly affected by atmospheric absorption and slightly differing amounts of vegetative photosynthetic activity between the scenes due to, for example, phenological or plant moisture differences. Because of the high NIR reflectance of vegetation compared to the green and red bands, these variations in vegetative conditions derive the observed higher variability in ρ_s. The very small variation seen in ρ_{SAT} green band average ρ_s for the Dataset 1 imagery suggests that the contribution from atmospheric scattering was consistent and also that, as was noted by Song *et al*. (2001), because atmospheric aerosols reduce the apparent reflectance of bright targets and increase it for dark-objects, scattering may even be acting to smother the signal coming from the Earth's surface itself. This is contrary to the expectation that variations in atmospheric scattering in the shortest wavelength SPOT band would act to add to the variability in average ρ_s between scenes. As consequence of this stability in the ρ_{SAT} imagery, as can be seen in Figure 16 below, with the exception of DOS1 which derived a slight reduction in variability, all the other correction methods added slightly to the variation in mean ρ_s. The difference between all methods was minimal, however, only ranging from 2.36 for DOS1 to 2.92 for DOS2, a 0.56 difference.

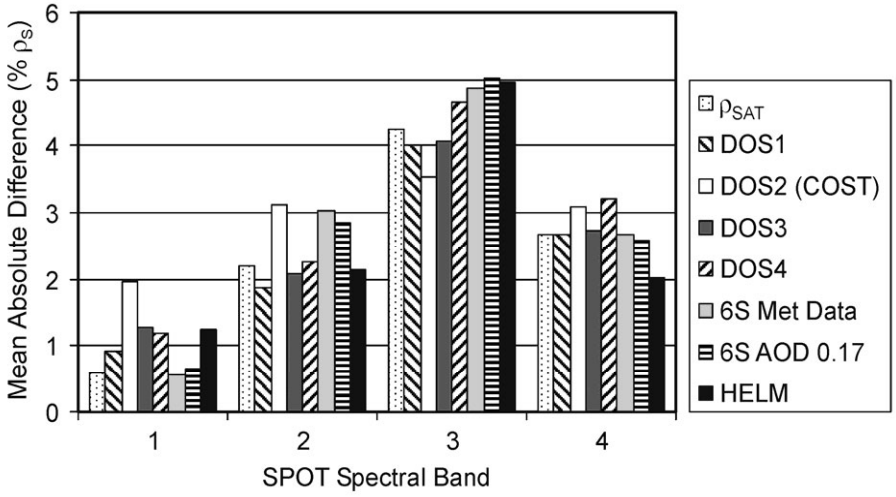


Figure 15. Mean absolute difference (% ρ_s) of each spectral band for each atmospheric correction methodology applied to the Helsinki metropolitan region control site Dataset 1 SPOT imagery.

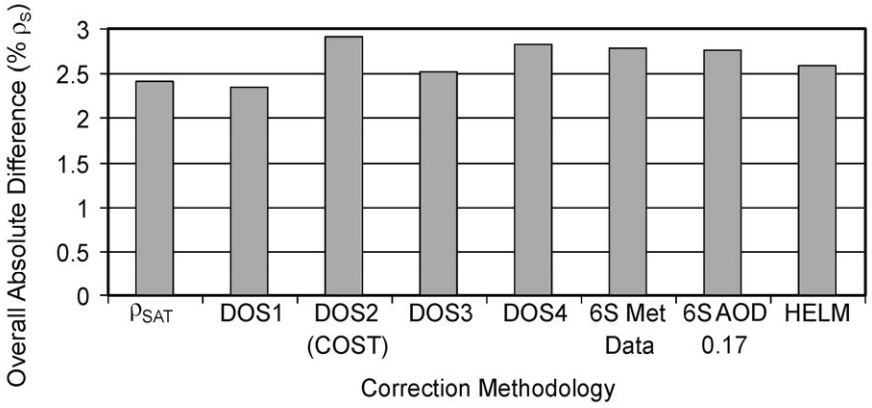


Figure 16. Overall mean absolute difference (% ρ_s) for all spectral bands for each atmospheric correction methodology applied to the Helsinki metropolitan region control site Dataset 1 SPOT imagery .

What can be stated in general, therefore, is that the HELM variations in average ρ_s are commensurate with ρ_{SAT}. Overall, HELM did not significantly add to or reduce the variability in mean ρ_s between multi-temporal SPOT scenes, although the effect is wavelength dependent. It should be noted, however, that HELM has the advantage over the other applied corrections that performance is equivalent in imagery time-series stability, but significantly better in predicting ρ_s accurately. For HELM to be effective requires the careful selection of suitable dark-objects and the identification of one appropriate pseudo-invariant spectrally bright ground calibration target per SPOT scene area. Use of appropriate HELM calibration targets is important to enable up-scaling of averaged ρ_s field measurements to the SPOT sensors' GIFVO and to counter adjacency effects. Given the results of the Paper III study, in circumstances where no suitable HELM targets can be identified or where there is no field access within a particular scene area, then the application of 6S with standard atmosphere and aerosol models, and general estimates of AOD at 0.55 μm, would likely give the next best RFR performance. Furthermore, if the spectrometer used to make ρ_s measurements for HELM does not cover the SWIR, 6S is again likely to offer the next best ρ_s estimates. However, the 6S prediction error in the VIS/NIR bands is likely to exceed the desired 0.02 ρ_s benchmark and, depending on the specifics of the SPOT scene illumination and viewing geometry, could be more than twice the expected HELM RFR RMSE.

5.4 LULC data from multi-scale segmentation and object-oriented classification

In Paper IV, a multi-scale segmentation/object relationship modelling (MSS/ORM) approach was applied to map LULC at a landscape level in the Taita Hills, Kenya, from multi-temporal SPOT multispectral satellite imagery. This object-oriented procedure was shown to derive improvements over a uni-scale maximum-likelihood (ML) technique in this area of complex heterogeneous land cover. This was in terms of both an increase in the assessed overall accuracy of the 11 LULC class classification from 65.6% to 73.5%, and in a Kappa Index of Agreement from 0.6 to 0.66 (Paper IV: Table 4), but also more significantly in the derivation of visually superior land cover maps based on meaningful homogeneous landscape patches and free from the 'salt-and-pepper' classification noise effect typical of maximum-likelihood results (Paper IV: Figure 3). This is due to the theoretical advancements possible when conceptualizing a landscape and its depiction in a remotely sensed image as a spatially nested patch hierarchy definable at various critical levels of organization operating at specific spatial and temporal scales. Useful spatial information surrounding each pixel and multi-scale information within the image are incorporated into the classification process by the MSS/ORM approach, where a search is made for apparent boundaries in the gradient of flux zones within and between landscape patches identifiable through local heterogeneity. In particular, the fractal net evolution approach (FNEA) to multi-scale segmentation was successful at capturing image objects relating to ecologically meaningful landscape patches identifiable in the SPOT data. Segmentation studies are usually focused on high resolution imagery, such as digital aerial photography or IKONOS data, so it was interesting to note the successful application of the MSS/ORM approach to data derived from a medium resolution sensor such as SPOT. Despite improvement in classification performance with the MSS/ORM approach, it was nevertheless still

necessary to undertake manual editing, by reference to the original imagery, in order to increase the accuracy of the LULC classifications to a level suitable for further utilization in multi-disciplinary applications. This level was identified as a validated classification accuracy of ≥ 90% at the highest level of class type aggregations by the GTOS ECV-T9 guidelines for 30 m land cover type data requirements (GTOS-T9 2009). Manual editing increased the Taita Hills LULC mapping overall accuracy to 89%, and the Kappa Index of Agreement to 0.87, with the individual users' and producers' accuracies both being > 90% with the exception of three cover types (Paper IV: Table 4).

Blaschke (2010) gave a recent review article of the state-of-the-art in object-based image analysis (OBIA) based on analysis of over 820 articles, including 145 journal papers and 84 book chapters. Blaschke (2010) identifies OBIA as a growing major trend in remote sensing and GIScience over the last decade, with Gamanya *et al*. (2009) even claiming that object-oriented processing is becoming more popular that traditional pixel-based image analysis. Not all OBIA utilizes multi-scale segmentation and classification, though, and Lang & Langanke (2006) demonstrated that for specific cases a one level representation (OLR) may be sufficient and more straightforward. In either case, however, remote sensing information is used as a proxy for phenomena or processes (Blaschke 2010), and enables improvements in image analysis. As was the case in the Paper IV study, overcoming the 'salt-and-pepper' effect of per-pixel type image analysis has been widely recognized in research work as a major benefit of OBIA (Blaschke 2010).

Furthermore, a very large number of studies have reported improvements in OBIA land cover mapping and change analysis accuracies compared to per-pixel based approaches (Blaschke 2010). For example, Shackelford & Davis (2003) derived a fuzzy logic OBIA methodology to map urban land cover from IKONOS data incorporating image texture measures and a length-width contextual measure. Compared to ML approach results of 79% accuracy for urban areas and 87% accuracy for suburban areas, the OBIA technique derived 8% and 11% improvements in accuracy, respectively. It should be noted that this reported ~10% increase in classification accuracy was similar to the results obtained in Paper IV. Im *et al*. (2008) compared change detection techniques based on per-pixel change classification, object/neighbourhood correlation, and image segmentation and found that the object based change classifications had a higher Kappa Index of Agreement of up to 0.9, compared to 0.8–0.85 for the traditional approaches. Platt & Rapoza (2008), in a comparison of a ML classification with an OBIA approach for mapping land cover around Gettysburg, Pennsylvania, noted that there are at least four components in OBIA not typically used in per-pixel analysis which contribute to the enhanced performance: (1) the segmentation procedure, (2) the nearest neighbour classifier, (3) the integration of expert knowledge, and (4) feature space optimization. Similarly, all these components were utilized in the Paper IV study.

The final maps derived in the Paper IV study were used to identify major landscape changes that occurred in the Taita Hills over the 1987 to 2003 research period. As discussed in Paper IV, the mapped changes where overwhelming in the form of the clearance of natural vegetation for agricultural expansion, with croplands increasing by 10,478 ha, mainly into shrubland (6,414 ha) and thicket (closed canopy shrublands; 3,300 ha) areas (Paper IV: Tables 6 & 7), especially on the lowlands and foothills. This was similar to trends detected in other parts of Kenya and East Africa (Imbernon 1999; Soini 2005b; Baldyga *et al*. 2007), and also over sub-Saharan Africa as whole (Eva *et al*. 2006). Increasing environmental pressure on the Taita Hills was

implicated by other mapped landscape level changes. Although strong seasonal variations in water levels are acknowledged for the East African environment, there was a major 77% reduction in the extent of standing water from 84 to 19.5 ha, despite that reference to local meteorological data shows that all three images were acquired during similarly dry periods. There was also a large 145% (546 ha) increase in the area of bare soil, mainly from cropland and shrubland areas (Paper IV: Table 7), and a 34% increase in built-up areas. Given this, it was heartening to report that the extent of the core area of the remnant indigenous forest patches had remained stable from 1992 (Paper IV: Table 5), although no attempt was made to assess the health of the forests from the SPOT imagery, which is an important factor for habitat quality.

However, there was a 10% reduction in the total area of indigenous forest (mapped as Broadleaved Closed Canopy Forest) over the study period, dropping from 774 ha in 1987, to 741 ha in 1992 and 694 ha in 2003 (Paper IV: Table 6). In combination with this, the number of patches increased slightly between 1987 and 1992 whilst the mean size reduced (Paper IV: Table 5), indicating fragmentation, but then by 2003 the total number of patches fell while the mean size increased, reflecting a loss of smaller patches. Most of this loss (54 ha) was to plantation forest and woodland, with only a relatively minor 22 ha being mapped as converted to croplands (Paper IV: Table 7). Although there were large scale conversions of shrublands and thicket to agriculture, it therefore appears that there has been little deforestation in the Taita Hills during the studied period, as both woodland and plantation forest cover increased (9% and 1.5% respectively). This contrasts with a study by Ward *et al*. (2004) based on an analysis of Landsat imagery from 1987 and 1999 utilizing unsupervised classification methods, which reported a large 37% decrease of indigenous forest in the Taita Hills, but mapped for both 1987 and 1999 with an erroneous over-estimation of cover with nearly the entire upland areas of the hills depicted as forest. In reality the remnant patches present during this time period were very small and only covered less than 1% of the total area, as indicated both from field measurements by Lens *et al*. 2002 and from the results of this study. It is, however, acknowledged that the three-date 'snap-shot' study reported in Paper IV forms a minor temporal sampling of a complex environmental system in constant flux. Nevertheless, in an area of the world where detailed accuracy assessed landscape level digital mapping and change analysis is sparse, this information derived using the standardized UN FAO LCCS nomenclature is useful in many applications, as well as an indicator for the likely presence and extent of land degradation processes occurring in the Taita Hills.

Further, Paper IV also demonstrated a novel method of implementing a "*c*" correction based topographic normalization (Teillet *et al*. 1982) to SPOT imagery (Paper IV: Section 3.1.2). For mountainous and rugged terrain regions, such as the Taita Hills, topographic correction is at least as important as atmospheric correction, if comparable ρ_s values are to be derived throughout the area (Liang 2004). This is critical both for traditional classification techniques and for image segmentation procedures, where changes in reflectance should relate solely to differences in land cover and not to variation in illumination conditions. To identify general vegetation classes, from which to derive *c* factors, *before* a topographic correction and classification were applied, NDVIs were derived and cluster classes were identified within them using the automated ISODATA algorithm. This is unusual as the ISODATA algorithm is normally utilized to automatically identify data clusters in multispectral feature space, but here it was used in a one-dimensional manner to capture frequency clusters in the NDVI. It was considered that the use of *c* factors

accounted for both diffuse irradiance and the non-Lambertian reflectance behaviour of the vegetation within each generalized group, and it also had the effect of limiting the overcorrection of weakly illuminated pixels. Visual inspection and re-regression of the corrected reflectance values for each class area against the cosine of the cos *i* values, which derived no relationships, demonstrated that topographic effects had been successfully removed from the SPOT imagery.

5.5 Modelling potential soil loss using GIS data and SPOT derived LULC maps

In Paper V, multi-temporal land cover maps of 1987, 1992 and 2003, derived from SPOT imagery using an MSS/ORM methodology as detailed in Paper IV, were utilized in combination with compiled geospatial layers of soil, rainfall and topography to model potential soil loss in the Taita Hills, Kenya. In accordance with working in regions of the developing world with limited ancillary data availability, and considering the ability of the procedures to be replicated locally using a minimum of resources, low cost and freely available GIS data and a simplified methodology based on the Universal Soil Loss Equation (USLE), were utilized. Local rainfall records of the Kenyan Meteorological Department were interpolated into a 20 m grid layer, a 20 m DEM was generated from Survey of Kenya 1:50,000 scale topographic paper maps, and the 1:1,000,000 scale Explanatory Soil Map of Kenya was utilized as a generalized digital soil map.

Even considering the minimized data requirements of the applied methodology, the lack of a detailed soil map and of a support practice factor (*P*-factor) layer (i.e. a geospatial layer denoting the application of soil conservation measures, such as terracing and agro-forestry, which have been implemented in certain areas of the Taita Hills) could be considered as limitations of the Paper V study. Nevertheless, the implemented procedure demonstrated that, in a GIS environment, it was possible to quickly and easily collate and process geospatial data to enable adequate quantitative modelling of potential soil loss in data limited circumstances. A simple USLE model was applied to identify the areas of highest potential soil loss risk in the Taita Hills, and this was related to the mapped changes in land cover in order to predict long term erosion hazard. The availability of the mapped USLE factors and the quantitative spatial information concerning annual potential soil loss supports the design and development of soil conservation programs in an area suffering from land degradation and localized severe soil loss, such as gully formation, as detailed in Sirviö *et al*. (2004) and evidenced in the field.

Land cover change in the Taita Hills was significant over the studied period from 1987 and 2003, with population growth leading to a shortage of agricultural land and the movement of people into the lowlands, evidenced by a 39% (9.3 km^2) increase in croplands from 30% to 41% of the study area (Paper V: Table 7). Agricultural expansion took place mostly in the foothills and lowlands surrounding the hills, at the expense of natural shrubland and thicket (closed canopy shrublands), but also occurred to lesser extent within the higher areas of the hills themselves (Paper V: Figure 3). These land cover changes were shown to have a strong impact on modelled annual potential soil loss, which was classified into low (0–2 t h^{-1} yr^{-1}), moderate (2–10 t h^{-1} yr^{-1}), high (10–30 t h^{-1} yr^{-1}) and very high (> 30 t h^{-1} yr^{-1}) soil loss classes (Paper V: Figure 5). The USLE modelling results showed a 60% (4 km^2) increase in the area of very high potential soil loss, from 5 758 ha to 9 656 ha and from 7% of the study area in 1987 to 12% in 2003, due mainly to the expansion of

croplands, accounting for a 2 085 ha increase in very high soil loss potential (Paper V: Figure 5 & Table 9). However, the relative proportion of very high soil loss potential within croplands remained 20% both in 1987 and 2003, indicating that newly cleared agricultural lands with vulnerable soils were the most at risk areas.

Overall, the slope steepness and length factor (*LS*-factor) and the rainfall erosivity factor (*R*-factor) for croplands decreased because agricultural expansion has encroached into lowlands with more gentle topography and lower rainfall (Paper V: Table 8). The opposite trend can be seen with shrubland, where the *LS*- and *R*-factors increased since shrublands were cleared and the remaining areas were overall on steeper slopes. Another significant change to be detected was the substantial increase in burned areas (Paper V: Table 7) located more on the lowlands at the edges of agricultural expansion areas (Paper V: Figure 3), thus decreasing the *LS*- and *R*-factors significantly (Paper V: Table 8). This related to the local practice for establishing new fields in areas of shrublands by clear cutting the boundaries then burning the encircled area to clear the vegetation and add some nutrients to the soil. In the hills themselves, the soil loss potential generally decreased due to increased vegetation cover within the croplands, a consequence of continued adoption and implementation of agroforestry practices. However, agricultural areas occurring on the steeper slopes still remained as zones of very high soil erosion potential (Paper V: Figure 6). The parts of the study area most severely at risk of soil erosion were found to be the newly cleared croplands in the southern, southwestern and northern foothills and lowlands.

The land cover changes mapped in the Taita Hills, and discussed in Paper V, are similar to trends detected throughout many parts of East Africa (Imbernon 1999; Soini 2005b), and in fact throughout sub-Saharan Africa as whole (Brink & Eva 2009). It is therefore considered likely that potential soil loss due to water and wind erosion has increased throughout many environmentally pressured parts of the region, as a direct consequence of rapid and unmanaged land cover changes overwhelming in the form of the clearance of natural vegetation for agricultural expansion (Eva *et al*. 2006).

5.6 The use of SPOT derived predictor variables in human population prediction

In Paper VI, population distribution and abundance were modelled for the rural upland subsistence farming area of the Taita Hills, Kenya, using dwelling unit data (presence-absence) and population count data (abundance) as the response variable and geospatial and remote sensing based data as predictors. The prediction models were created with the GRASP method, which utilizes the GAM regression technique. The Paper VI results showed that population distribution models explained 19% to 31% of variation in the dwelling unit occurrence data, indicating a fair explanatory power, and that the predictive power for population distribution models was good with AUC of 0.80 to 0.86. The abundance models explained 28% to 47% of the variation in human population abundance throughout the study area. Combining the geospatial and remote sensing based predictors gave the overall best modelling results when compared with only remotely sensed or geospatial based predictor models. The best individual predictors for modelling the variability in human population distribution when using combined predictors were: angular second moment image-texture measurement, precipitation, mean elevation, surface reflectance for SPOT red and NIR bands, correlation image-texture measurement and distance to roads, respectively (Paper VI: Tables 3, 4, 5 and 6). The fairly poor performance of the land

cover classes utilized as predictors in the human population predictive modelling indicated that it is not necessary to make use of classified LULC data. Instead, the first and second order image-texture measurements derived from the SPOT satellite image gave higher contributions to the models. Second-order image texture measures have been shown to be important factors in previous urban population density analysis studies (for example Shaban & Dikshit 2001; Li & Weng 2005).

A further benefit of using the first and second order image-texture measurements is the fact that they can be calculated quickly and easily, and can be used as a proxy indicator to quantify the variation in vegetation as a continuous variable in statistical modelling. Whilst in the MSS/ORM classification discussed in Paper IV, it was found that it was not possible to differentiate between croplands and shrublands at any level in the segmentation hierarchy, based on the image texture of segmented image primitives, this finding is limited to a consideration of these specific land covers. This analysis was motivated by the difficulty in distinguishing between lowland croplands and shrublands in the classification, because of the close spectral similarly between the two within the SPOT data. Image texture was found to be useful in modelling population as here image wide texture measures evidently varied with more substantial variations in vegetation type and cover related to the likely occurrence of dwelling units. That land cover data may not necessarily be needed for successfully modelling population distribution and abundance is an important finding, considering the laborious and costly classification work required to derive and assess accurate LULC maps from satellite imagery.

This study also showed that modelling using only GIS derived geospatial predictors had the lowest performance in population models, and therefore it is suggested that they should not be used in isolation as a predictors for dwelling unit distribution and abundance modelling. However, it should be noted that the important distance to road predictor was generated from 1:50,000 scale topographic mapping. In the Taita Hills, in common with many other rural areas in the developing world, the majority of homestead dwellings are in fact only accessed by footpaths. Consequently, the availability of a more detailed road network GIS layer that also included footpaths, and its use as a geospatial predictor, might have improved the model performance.

GAMs were chosen for modelling because in various ecological species distribution studies they have outperformed conventional linear regression techniques (Yee & Mitchell 1991; Thuiller *et al.* 2003). Moreover, GAMs are more suitable for geospatial data modelling as environmental predictors are often non-Gaussian with non-constant variance. However, in previous human population modelling studies using geospatial data, it is mainly linear regression techniques that have been utilized (for example, Lo 1995; Schnaiberg *et al.* 2002; Gustafson *et al.* 2005; Li & Weng 2005). To the authors' knowledge, this study is the first time that GAM models have been used for human occurrence and abundance prediction using geospatial predictors. Thus, the good modelling results should encourage other predictive human population and abundance studies to consider using the GAM technique alongside, or instead of, more traditional regression methods.

The GAM abundance model was extrapolated for the whole Taita Hills study area and the model was capable of discriminating between inhabited and uninhabited areas. For example, in the vicinity of the Ngangao indigenous forest (Paper VI: Figure 4) an absence of dwelling units was predicted within the forest itself and low dwelling unit probability was modelled for the cultivated fields, whilst high population concentrations were correctly identified in and around the villages. When abundance

models were compared with two existing global population datasets, GPWv3 and LandScan 2005, the results showed that there was statistically significant correlation between the combined and the remote sensing based models and the GPWv3 product ($r > 0.8$) but the correlation was non-significant with the geospatial based model ($r = 0.19$). For LandScan 2005 the correlations were lower (Paper VI: Table 7). The correlation between Kenyan census data for 1999 and predicted population abundance models were high for the remote sensing based ($r = 0.71$) and combined models ($r = 0.51$) when only sub-locations over 1100 m elevation ($n = 32$) were used. For the geospatial model the correlation was non-significant. However, there was low correlation ($r = 0.34$) between the remotely sensed based population abundance models and the Kenyan census data for 1999 for the sub-locations also extending into the dry lowland areas ($n = 50$) surrounding the Taita Hills, where the population density is highly variable, and no correlation for combined and geospatial models (Paper VI: Table 8). This is mainly due to the inadequate description provided by the census sub-locations for these areas, which vary in size and shape and generalize and obscure the internal variability of population relative to the spatially explicit combined models, which have 100 m grid size.

In summary, then, it can be stated that the predictive models using combined geospatial and SPOT derived predictors were found to be more accurate than the global datasets and also correlated well with the Kenyan 1999 census data. However, it must be kept in mind that in general the modelling performance can be affected by different factors such as the analysis scale, the possible occurrence of spatial autocorrelation, the selected predictors, and the applied modelling techniques and model parameterization.

6. CONCLUSIONS AND FUTURE PROSPECTS

The Taita Hills in southeastern Kenya form the northernmost part of Africa's Eastern Arc Mountains, which have been identified as one of the top ten biodiversity hotspots on Earth. As with many areas of the developing world, over recent decades the Taita Hills have experienced significant population growth. This has led to associated major changes in land use and land cover (LULC), as well as escalating land degradation, particularly in the form of soil erosion. Multi-temporal medium resolution multispectral optical satellite data, such as imagery from the SPOT HRV, HRVIR, and HRG sensors, provides a valuable source of information for environmental monitoring and modelling at a landscape level over local and regional scales. Satellite remote sensing represents the most efficient methodology for routinely collecting synoptic and spatially explicit data of extensive and inaccessible areas, with information available in spectral, spatial and temporal resolutions.

However the utilization of multi-temporal SPOT data, or any other multispectral medium or high resolution satellite imagery, in quantitative remote sensing studies requires the removal of atmospheric effects and the derivation of surface reflectance factor (ρ_s). Furthermore, for mountainous regions or areas of rugged terrain, such as the Taita Hills, it is necessary to apply topographic correction to remove slope-aspect effects to derive comparable ρ_s throughout a scene area. Reliable monitoring of LULC and LULC change over time and modelling of land degradation and human population distribution and abundance are of crucial importance to sustainable development, natural resource management, biodiversity conservation, understanding

ecosystems and biogeochemical cycling, and understanding and mitigating climate change and its impacts.

The main purpose of this thesis was to develop and validate enhanced processing of SPOT satellite imagery for use in environmental monitoring and modelling, in regions of the developing world with limited ancillary data availability. The Taita Hills formed the application study site (Papers II, III, IV, V, VI), whilst the Helsinki metropolitan region was used as a control site for validation and assessment of the applied atmospheric correction techniques, where multiangular ρ_s field measurements could be taken and where horizontal visibility meteorological data concurrent with image acquisition were available (Papers I, II, III). The main findings of this study, which can be considered as methodological enhancements in the processing and utilization of multi-temporal SPOT multispectral satellite imagery for environmental monitoring and modelling at a landscape level over local or regional scales, are as follows:

- The application of the proposed *historical empirical line method* (HELM) for absolute atmospheric correction of SPOT data requires the identification of a spectrally pseudo-invariant calibration site within each scene area. Considering the ± 30° view zenith angle (θ_{VZ}) range relating to the SPOT sensors, vegetation-free sites make better calibration targets than vegetated ones. This is because the multiangular reflectance properties are more limited and are similar throughout the spectrum, unlike vegetated surfaces. Of the measured sands, gravel, asphalts, and artificial turf vegetation-free surface types, sand was found to be the most appropriate as a HELM calibration site. Sand is spectrally stable over time, as changes due to weathering are a very long term process. The most desirable sand calibration target properties are a well-sorted and finer grained material because this will have less surface roughness, and consequently less sensitivity to solar zenith angle (θ_Z) and θ_{VZ} variations, and lesser multiangular ρ_s anisotropy. Further, finer grained sands are also likely to have increased volumetric scattering leading to a brighter spectrum, which is useful in better determining HELM correction lines.

- Of the seven absolute atmospheric correction methodologies applied and compared in this thesis - HELM, four dark-object subtraction (DOS) methods, and the 6S radiative transfer model (RTM) applied with general estimates of atmospheric optical depth (AOD) at 0.55 μm and with horizontal visibility meteorological data - HELM derived the most accurate ρ_s retrieval, with HELM-1 giving an overall absolute accuracy of $< 0.02\ \rho_s$ for the SPOT visible and near-infrared (VIS/NIR) bands, and $< 0.03\ \rho_s$ for the SPOT shortwave infrared (SWIR) band. This represented a relative RMSE of ~10% or less for all bands. Most importantly, HELM was the only applied procedure that achieved VIS/NIR ρ_s retrieval with a RMSE within the desired 0.02 ρ_s benchmark identified as a measure of successful atmospheric correction. Further, HELM SWIR ρ_s retrieval was better than the other applied methodologies as well as the partially corrected at-satellite reflectance (ρ_{SAT}), which was taken as the basis of successful SWIR atmospheric correction in this study. HELM was also able to maintain radiometric stability within multi-temporal SPOT imagery datasets comparable with the ρ_{SAT}, and better than the other applied techniques. This is a significant enhancement for the retrieval of ρ_s in areas of the world there is a paucity of meteorological data available that is detailed enough, and has an appropriate spatial and temporal

frequency, to allow for the full application of RTMs to multi-temporal satellite imagery datasets. These areas are also often the places where the most rapid and significant changes in LULC are occurring, and where the need for environmental monitoring is greatest.

- 6S applied with horizontal visibility meteorological data, and with general estimates of AOD at 0.55 μm, utilizing standard atmosphere and aerosol models, derived almost identical accuracy and bias in predicting ρ_s for the validation targets in the Helsinki control site from SPOT data. This is important as it shows 6S can be applied in areas where no meteorological data are available with equal accuracy to those areas where horizontal visibility meteorological data are available. Further, in the atmospheric correction comparison conducted for this thesis, 6S was found to give the next best ρ_s retrieval performance after HELM and derived better accuracy than the applied DOS approaches. Therefore, in circumstances where no suitable HELM targets can be identified or where there is no field access within a particular scene area, then the application of 6S with standard atmosphere and aerosol models, and general estimates of AOD at 0.55 μm, would likely give the next best RFR performance. Also, if the spectrometer used to make ρ_s measurements for HELM does not cover the SWIR, 6S is again likely to offer the next best ρ_s estimates. However, in such circumstances, the 6S prediction error in the VIS/NIR bands is likely to exceed the desired 0.02 ρ_s benchmark and, depending on the specifics of the SPOT scene illumination and viewing geometry, could be more than twice the expected HELM ρ_s retrieval RMSE.

- For mountainous areas, such as the Taita Hills, topographic correction is important if comparable ρ_s values are to be derived across SPOT scenes. This thesis showed NDVIs and automated data clustering algorithms can be utilized to identify general vegetation classes, from which to derive "*c*" factors, *before* a topographic correction and classification are applied. This enables the successful application of a "*c*" correction based topographic normalization to SPOT data, to remove the slope-aspect effects in the imagery. It is argued that the application of the "*c*" factors derived in this manner accounts for both diffuse irradiance and the non-Lambertian reflectance behaviour of the vegetation within each generalized group, and it also has the effect of limiting the overcorrection of weakly illuminated pixels.

- A multi-scale segmentation/object relationship modelling (MSS/ORM) approach was applied to map LULC at a landscape level in the Taita Hills, from the multi-temporal SPOT imagery. This object-based procedure was shown to derive improvements over a uni-scale maximum-likelihood technique in this complex heterogeneous area, both in terms of an increase in the assessed overall accuracy of the classification from 65.6% to 73.5%, and in a Kappa Index of Agreement from 0.6 to 0.66, but also more significantly in the derivation of visually superior land cover maps based on meaningful homogeneous landscape patches and free from the 'salt and-pepper' classification noise effect typical of maximum-likelihood results. This is due to the theoretical advancements possible when conceptualizing a landscape and its depiction in a remotely sensed image as a spatially nested patch hierarchy definable at various critical levels of organization operating at specific spatial and temporal scales. Useful spatial information

surrounding each pixel and multi-scale information within the image are incorporated into the classification process by the MSS/ORM approach, where a search is made for apparent boundaries in the gradient of flux zones within and between landscape patches identifiable through local heterogeneity. In particular, the fractal net evolution approach (FNEA) to multi-scale segmentation was successful at capturing image objects relating to ecologically meaningful landscape patches identifiable in the SPOT data. However, manual editing of the final LULC maps was still necessary to bring the data up to an accuracy level (90%) that was appropriate for the application of the information to environmental monitoring and modelling in the Taita Hills.

- In regions of the developing world with limited ancillary data availability, such as the Taita Hills, simplified and easily implemented methodologies for modelling land degradation at a landscape level are useful in supporting local efforts in sustainable land use planning and soil conservation programmes. LULC and LULC changes mapped from SPOT imagery can be used in combination with low cost GIS derived geospatial layers describing elevation, rainfall and soil type, to model degradation in the form of potential soil loss, utilizing the simple universal soil loss equation (USLE). Utilization of LULC mapped from SPOT imagery allows for the derivation of pertinent cover management factors (*C*-factors) for the USLE application, and changes in LULC can be related to variation in the modelled potential soil loss to identify areas at significant risk from soil erosion.

- Human population distribution and abundance can be modelled with satisfactory results using only GIS and SPOT derived data and non-Gaussian predictive modelling techniques applied using the Generalized Regression Analysis and Spatial Prediction (GRASP) modelling framework. The SPOT imagery derived land cover data was found to be unnecessary as a predictor because the first and second order image texture measurements had greater power to explain variation in dwelling unit occurrence and abundance. This makes the application of human population modelling significantly easier, since the derivation and validation of accurate land cover information from SPOT data is a relatively time consuming and costly process. Local scale predictive human population abundance models were more suitable than the existing coarser scale GPWv3 and LandScan 2005 global population datasets in estimating the abundance and mapping the distribution of population in the Taita Hills. SPOT derived predictor variables utilized on their own did not derive satisfactory modelling results, as was also found to be the case with GIS derived geospatial predictor variables utilized in isolation. However, combining the geospatial and remote sensing based predictors gave the best overall modelling performance, and demonstrated that human population distribution and abundance could be modelled with satisfactory results.

The methods outlined in this thesis utilized multi-temporal SPOT satellite imagery; however it is considered that they are equally applicable to any medium or high resolution multispectral optical remote sensing data. ρ_s was derived within 0.02 ρ_s absolute accuracy (~10% relative accuracy), a benchmark identified as a measure of successful atmospheric correction. To enable interoperability and more widespread data usability, LULC was mapped using a nomenclature generated with the FAO LCCS, and the applied data processing and validation procedures were in line with the Global Terrestrial Observing System (GTOS) guidelines for deriving land cover at

medium resolution mapping scales as an essential climate variable (ECV-T9). Despite the application of an advanced MSS/ORM approach in deriving LULC, manual editing of the final maps was still necessary to bring the data up to an accuracy level (90%) that was appropriate for the application of the information to environmental monitoring and modelling in the Taita Hills. Consequently, it can be seen that the semi-automatic derivation of accurate LULC in complex heterogeneous landscapes, such as those of the Taita Hills, remains a challenging task and needs further research. This is a reason why the GTOS guidelines recommend that, in all cases, LULC products are inspected by a remote sensing analyst familiar with the mapped region. When utilizing the MSS/ORM approach, automated derivation of the scale parameters for an FNEA image segmentation, rather than the application of heuristic rules, would be a significant step forward in extending the application of the approach. This could probably be achieved by calculating the scale-space appearance and persistence of landscape patches, but this is still currently a research area.

This thesis has focused on the methodological aspects of monitoring and modelling the key environmental issues of LULC change, land degradation and population growth, based on the Taita Hills application study site. During the last few decades the development of remote sensing and GIS data analysis software, along with very significant improvements in computing resources, the growing number of remote sensing data sources, and the science- and policy-driven requirements for environmental data, have generated major improvements in LULC characterization and have allowed for more sophisticated geospatial environmental applications to be developed. Throughout the developing regions of the world, there is an ongoing requirement for further environmental studies at multiple scales using standardized classification schemes, and data processing and validation procedures, supplied with detailed metadata, to allow interoperability and greater usability of information on the magnitude and nature of LULC changes. The most effective environmental monitoring and modelling procedures will be those which can be successfully implemented locally in circumstances of limited financial, data and software resources.

However, a full understanding of the mechanisms and driving forces behind environmental changes requires the consideration of socio-economic, cultural, and political, as well as physical, factors. Environmental monitoring and modelling using remote sensing and GIS derived geospatial data and spatial statistical techniques are, therefore, only tools, albeit powerful ones, applied to the study of what are complex real world problems. Integrated research on environmental changes requires the collaboration of earth scientists such as physical geographers, ecologists and climatologists, and social scientists such as human geographers, anthropologists and economists, as well as remote sensing scientists, and is thus truly interdisciplinary as a research paradigm. As all these issues are inherently socio-spatial, the discipline of geography is well placed to make significant contributions to such studies, and to offer the benefit of oversight over a wide range of pertinent factors. The complex nature of the problems requires high-levels of cooperation. By pulling together, though, research efforts can be concentrated and offer a real chance to move towards scientifically determined actions for conserving natural resources and implementing truly sustainable development.

REFERENCES

Achard F, G Grassi, M Herold, M Teobaldelli & D Mollicone (2008). *Use of satellite remote sensing in LULUCF sector*. GTOS GOFC-GOLD Report n. 33. Land Cover Project Office, Jena University. 25 p. Available at: www.fao.org/gtos/gofc-gold/series.html.

Akaike H (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–722.

Alves DS & DL Skole (1996). Characterizing land cover dynamics using multi-temporal imagery. *International Journal of Remote Sensing* 17(4), 835–839.

Anderson K & EJ Milton (2006). On the temporal stability of ground calibration targets: implications for the reproducibility of remote sensing methodologies. *International Journal of Remote Sensing* 27(16), 3365–3374.

Aplin P (2006).On scales and dynamics in observing the environment. *International Journal of Remote Sensing* 27, 2123–2140.

Asner GP, BH Braswell, DS Schimel & CA Wessmen (1998). Ecological research needs from multiangular remote sensing data. *Remote Sensing of Environment* 35, 161–173.

Baatz M & A Schäpe (2000). Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. *In* Strobl J & T Blaschke (eds.): *Angewandte geographische Informationsverarbeitung* Vol. XII, 12–23. Wichmann, Heidelberg, Germany.

Baldyga TJ, SN Miller, KL Driese & CN Gichaba (2007). Assessing land cover change in Kenya's Mau Forest region using remotely sensed data. *African Journal of Ecology* 46, 46–54.

Baraldi A & F Parmiggiani (1995). An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33, 293−304.

Barnsley MJ, AH Strahler, KP Morris & J-P Muller (1994). Sampling the surface BRDF: 1. Evaluation of current and future satellite sensors. *Remote Sensing Reviews* 8, 1937–1960

Barnsley MJ (1999). Digital remotely sensed data and their characteristics. *In* Longley PA, MF Goodchild, DJ Maquire & DW Rhind (eds.): *Geographical Information Systems: principles, techniques, applications and management*, 567–580. John Wiley & Sons, New York.

Beentje HJ & N Ndiang'ui (1988). An ecological and floristic study of the forests of the Taita Hills, Kenya. *Utafiti* 1(2): 23–66.

Benz UC, P Hoffmann, G Willhauck, I Lingenfelder & M Heynen (2004). Multi-resolution, object oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing* 58, 239–258.

Berk A, GP Anderson, PK Acharya, JH Chetwynd, LS Berstein, EP Shettle, MW Matthew & SM Adler-Golden (2000). *Modtran4 User's Manual*. US Air Force Research Laboratories, Hanscom AFB, MA.

Blaschke T & GJ Hay (2001). Object-oriented image analysis and scale-space: Theory and methods for modelling and evaluating multi-scale landscape structure. *International Archives of Photogrammetry and Remote Sensing* 34 (Part 4/M5), 22–29.

Blaschke T, C Burnett & A Pekkarinen (2004). Image Segmentation Methods for Object-based Analysis and Classification. *In* de Jong SM, FD van der Meer (eds.): *Remote Sensing Image Analysis: Including the Spatial Domain*, 211–236. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Blaschke T, S Lang & G Hay (2008). *Object Based Image Analysis.* 817 p. Springer, Heidelberg, Berlin, New York.

Blaschke T (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 2–16.

Burnett C & T Blaschke (2003). A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling* 168, 233–249.

Brink AB & HD Eva (2009). Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach, *Applied Geography* 29(4), 501–512.

Campbell JB (2002). *Introduction to remote sensing* (3rd ed.). 621 p. Taylor & Francis, London.

Chavez PS Jr. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multi-spectral data. *Remote Sensing of Environment* 24, 459–479.

Chavez PS Jr. (1989). Radiometric calibration of Landsat Thematic Mapper multispectral images. *Photogrammetric Engineering and Remote Sensing* 55(9), 1285-1294.

Chavez PS Jr. (1996). Image-based atmospheric corrections – revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62(9), 1025–1036.

Chavez PS Jr. & DJ MacKinnon (1996). Automatic detection of vegetation changes in the southwest United States using remotely sensed images. *Photogrammetric Engineering and Remote Sensing* 60(5), 571–583.

CIESIN (Center for International Earth Science Information Network), Columbia University, & Centro Internacional de Agricultura Tropical (CIAT) (2005). *Gridded population of the world version 3 (GPWv3): Population grids*. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.

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

Chust G, D Ducrot & JLL Pretus (2004). Land cover discrimination potential of radar multitemporal series and optical multispectral images in a Mediterranean cultural landscape. *International Journal of Remote Sensing* 25(17), 3513–3528.

Cohen J (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 37–46.

Cole KL & ST Arundel (2004). Modeling the Climatic Requirements for Southwestern Plant Species. *Proceedings of the Twenty first Annual Pacific Climate Workshop*. State of California, Department of Water Resources.

Collett D (2003). *Modelling Binary Data* (2nd ed.). 387 p. Chapman and Hall/CRC, Boca Raton.

Colwell RN (1997). History and place of photographic interpretation. *In* Philipson WR (ed.): *Manual of Photographic Interpretation* (2nd ed.). 33–48. American Society for Photogrammetry & Remote Sensing, Bethesda.

Craven P & G Wahba G (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerical Mathematics* 31, 377–403.

Curran PJ (1985). Principles of Remote Sensing. 282 p. Longman Group Limited, London.

De Bie CAJM (2005). Assessment of soil erosion indicators for maize-based agro-ecosystems in Kenya. *Catena* 59, 231–251.

DeFries RS & AS Belward (2000). Global and regional land cover characterization from satellite data: an introduction to the special issue. *International Journal of Remote Sensing* 21, 893–916.

De Jong SM, FD van der Meer & JGPW Clevers (2009). Basics of Remote Sensing. *In* de Jong SM & FD van der Meer (eds.): *Remote Sensing Image Analysis: Including the Spatial Domain,* 1–16. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Deng J-S, K Wang, J Li & Y-H Deng (2008). Urban Land Use Change Detection Using Multisensor Satellite Images. *Pedosphere* 19(1), 96–103.

de Vries C, T Danaher, P Scarth & S Phinn (2007). An operational radiometric calibration procedure for the Landsat sensors based on pseudo-invariant target sites. *Remote Sensing of Environment* 107, 414–429.

de Wasseige C & P Defourny (2002). Retrieval of tropical forest structure characteristics from bi-directional reflectance of SPOT images. *Remote Sensing of Environment* 83(3), 362–375.

DiGregorio A (2005). *Land Cover Classification System (LCCS): Classification Concepts and User Manual*. *Software version 2*. 190 p. FAO Environment and Natural Resources Service Series, No. 8. FAO, Rome.

Dobson, JE, EA Bright, PR Coleman, RC Durfee & BA Worley (2000). LandScan: a global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing* 66, 849–857.

Duchemin B & P Maisongrande (2002). Normalization of directional effects in 10-day syntheses derived from VEGETATION/SPOT: I. Investigation of concepts based on simulation. *Remote Sensing of Environment* 81, 90–100.

EAWLS (2005). East African Wildlife Society: *Stakeholders Workshop on the Conservation and Management of Taita Hills Forests* — Workshop Summary Report.

ERDAS (2005). *ERDAS Field Guide*. Leica Geosystems Geospatial Imaging, LLC, Norcross, Georgia, USA.

Eva HD, AB Brink & D Simonetti (2006). *Monitoring land cover dynamics in sub-Saharan Africa*. Luxembourg: Office for Official Publication of the European Communities, EUR 22498 EN.

FAO (2006). *Global Forest Resources Assessment 2005: Progress towards sustainable forest management*. Food and Agriculture Organization of the United Nations, Forestry Paper 147, Rome.

FAO (2007). *State of the World's Forests 2007*. 144 p. Food and Agriculture Organization of the United Nations, Rome.

FAO (2009). *The State of Food Insecurity in the World 2009*. 56 p. Food and Agriculture Organization of the United Nations, Rome.

FAOSTAT (2006). *2006 database*. Food and Agriculture Organization, Rome. Available at: http://faostat.fao.org/.

FAO AQUASTAT (2006). *2006 database*. Food and Agriculture Organization, Rome. Available at: http://faostat.fao.org/

Fielding A & J Bell (1997), A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38–49.

Foley JA, R Defries, GP Asner, C Barford, G Bonan, SR Carpenter *et al*. (2005). Global consequences of land use changes. *Science* 309, 570–574.

Forster B (1983). Some urban measurements from Landsat data. *Photogrammetric Engineering and Remote Sensing* 49, 1693–1707.

Foster GR, LJ Lane, JD Nowlin, JM Laflen & RA Young (1981). Estimating erosion and sediment yield on field-sized areas. *Transactions of America Society of Agricultural engineering* 24, 1253–1262.

Friedl MA, DK McIver, JCF Hodges, XY Zhang, D Muchoney, AH Strahler, CE Woodcock, S Gopal, A Schneider, A Cooper, A Baccini, F Gao & C Schaaf (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment* 83, 287–302.

Gachimbi LN, PT Gicheru, MK Nyangw'ara, JK Lekasi, IV Sijali & J Kimigo (2005). *Agricultural Productivity and Sustainable Land Management in Kenya (APSLM): Smallholder farming, rural livelihoods, biodiversity and copping mechanisms in Taita Taveta District: Baseline Survey Report*. Kenya Agricultural Research Institute, Nairobi. Technical Report No. 1.

Gamanya R, R de Maeyer & M De Dapper (2009). Object-oriented change detection for the city of Harare, Zimbabwe. *Expert Systems with Applications* 36(1), 571–588.

Gao B-C, MJ Montes, CO Davis & AFH Goetz (2009). Atmospheric correction algorithms for hyperspectral remote sensing of land and ocean. *Remote Sensing of Environment* 113, S17–S24.

Gellman DI, SF Biggar, MC Dinguirard, PJ Henry, MS Moran, KJ Thome & PN Slater (1993). Review of SPOT-1 and -2 calibrations at White Sands from launch to the present. *Proceedings of SPIE*, 1938-13, 118–125.

Githiru M and L Lens (2007). Application of fragmentation research to conservation planning for multiple stakeholders: An example from the Taita Hills, southeast Kenya. *Biological Conservation* 137, 271–278.

Goetz AFH, JB Wellman & WL Barnes (1985). Optical remote sensing of the earth. *Proceedings of the IEEE* 73, 950−969.

Goetz S (2007). Crisis in Earth observation. *Science* 315, 1767.

GOFC-GOLD (2008). *Reducing greenhouse gas emissions from deforestation and degradation in developing countries: a sourcebook of methods and procedures for monitoring, measuring and reporting*, GOFC-GOLD report version COP13-2. Available at: www.gofc-gold.uni-jena.de/redd.

GTOS-52 (2008). Terrestrial Essential Climate Variables for Climate Change Assessment, Mitigation and Adaptation: Biennial report supplement. 44 p. Global Terrestrial Observing System, Food and Agriculture Organization of the United Nations, Rome.

GTOS-T9 (2009). *Assessment of the status of the development of the standards for the Terrestrial Essential Climate Variables, T9: Land Cover; Version 16*. 26 p. Global Terrestrial Observing System, Food and Agriculture Organization of the United Nations, Rome.

Gu X, G Guyot & M Verbrugghe (1992). Evaluation of measurement errors in ground reflectance for satellite calibration. *International Journal of Remote Sensing* 13, 2531–2546.

Guisan A & NE Zimmermann (2000). Predictive habitat distribution models in ecology. *Ecological Modelling* 135, 147–186.

Gustafson EJ, RB Hammer, VC Radeloff & RS Potts (2005). The relationship between environmental amenities and changing human settlement patterns

between 1980 and 2000 in the Midwestern USA. *Landscape Ecology* 20, 773–789.

Hall FG, PJ Sellers, DE Strebel, ET Kanemasu, RD Kelly, BL Blad, BJ Markham, JR Wang & F Huemmrich (1991). Satellite remote sensing of mass balance: results from surface energy and FIFE. *Remote Sensing of Environment* 35, 187–199.

Hall FG, KF Huemmrich, PJ Goetz, PJ Sellers & JE Nickeson (1992), Satellite remote sensing of the surface energy balance: success, failures and unresolved issues in FIFE, *Journal of Geophysical Research* 97(D17), 19061–19090.

Hansen MC, SV Stehman, PV Potapov, TR Loveland, RJG Townshend, RS DeFries, KW Pittman, B Arunarwati, F Stolle, MK Steininger, M Carroll & C DiMiceli (2008). Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 105(27), 9439–9444.

Hapke B (1993). *Theory of reflectance and emittance spectroscopy*. Cambridge, UK: Cambridge University Press.

Haralick R, K Shanmugan & I Dinstein (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* 3(1), 610–621.

Haralick RM (1983). Decision making in context. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5(4), 417–428.

Haralick RM & L Shapiro (1985). Survey: image segmentation techniques. *Computer Vision, Graphics, and Image Processing* 29, 100–132.

Harper D (1983). *Eye in the Sky. Introduction to Remote Sensing*. 252 p. Multiscience Publications Ltd., Montréal.

Hastie T & R Tibshirani (1990). *Generalized Additive Models*. Chapman and Hall, London.

Hay GJ, DJ Marceau, P Dubé & A Bouchard (2001). A multi-scale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology* 16(6), 471–490.

Hay GJ, P Dubé, A Bouchard & DJ Marceau (2002). A scale-space primer for exploring and quantifying complex landscapes. *Ecological Modelling* 13, 27–49.

Hay GJ & DJ Marceau (2004). Multiscale Object-Specific Analysis (MOSA): An integrative approach for multiscale landscape analysis. *In* de Jong SM & FD van der Meer (eds.): *Remote Sensing Image Analysis: Including the Spatial Domain*, 71–92. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Hay GJ & G Castilla (2008). Geographic Object-Based Image Analysis (GEOBIA). *In* Blaschke T, S Lang, G Hay (eds.): *Object Based Image Analysis*, 93–112. Springer, Heidelberg, Berlin, New York.

Heiskanen J (2006). Tree cover and height estimation in the Fennoscandian tundra-taiga transition zone using multiangular MISIR data. *Remote Sensing of Environment* 103, 97–114.

Herold M, CE Woodcock, TR Loveland, J Townshend, M Brady, C Steenmans & C Schmullius (2008). Land Cover Observations as part of a Global Earth Observation System of Systems (GEOSS): progress, activities, and prospects. *IEEE Systems* 2, 3, 414–423.

Hilker T, NC Coops, MA Wulder, AT Black & RD Guy (2008). The use of remote sensing in light use efficiency based models of gross primary production: a review of current status and future requirements. *Science of the Total Environment* 404, 411–423.

Holm M, A Lohi, M Rantasuo & S Väätäinen (1999). Creation of large mosaics of airborne digital camera imagery. In *Proceedings of the 4th International Airborne Remote Sensing Conference and Exhibition* / 21st Canadian Symposium on Remote Sensing, Vol II, 520–526. 21-24 June 1999. Ottawa, Ontario, Canada.

Holt-Jensen A (1988). *Geography. History and concepts* 2nd ed. 162 p. Chapman, London.

Houghton RA (2003). Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000. *Tellus* B 55, 378–390.

Houhoulls PF & WK Mlchener (2000). Detecting wetland change: A rule-based approach using NWI and SPOT-XS data. *Photogrammetric Engineering and Remote Sensing*, 66(2), 205–211.

Hutchinson MF (1989). A new procedure for griding elevation and streamline data with automatic removal of spurious pits. *Journal of Hydrology* 106, 211–232.

Hutchinson MF (1991). The application of thin plate splines to continent-wide data assimilation. In J.D. Jasper (Ed), *Data Assimilation Systems*. 104–113, BMRC Res. Rep. No. 27, Bureau of Meteorology, Melbourne.

Im J, JR Jenson & JA Tullis (2008). Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing* 29(2), 399–423.

Imbernon J (1999). Pattern and development of land use changes in the Kenyan highlands since the 1950s. *Agriculture, Ecosystems and Environment* 76, 67–73.

IPCC-AR4 (2007). *Climate Change 2007: Synthesis Report.* UN Intergovernmental Panel on Climate Change (IPCC), Fourth Assessment Report (AR4), 79 p. Geneva: www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html

Jackson RD, PM Teillet, PN Slater, G Fedesojevs, MF Jasinski, JK Aase & MS Moran (1990). Bidirectional measurements of surface reflectance for view angle corrections of oblique imagery. *Remote Sensing of Environment* 32, 189–202.

Jaetzold R & H Schmidt (1983). *Farm Management Handbook of Kenya*. Vol. II. East Kenya. Ministry of Agriculture, Kenya.

Jensen JR (2000). *Remote Sensing of the Environment: An Earth Resource Perspective*. 544 p. Prentice Hall, NJ.

KARI (2005). *Assessment of land degradation and its impacts on land use sustainability in Taita Taveta catchment*. EM Muya & PT Gicheru (Eds), Kenya Agricultural Research Institute, Miscellaneous, Paper No. 63.

Karpouzli E & T Malthus (2003). The empirical line method for the atmospheric correction of IKONOS imagery. *International Journal of Remote Sensing* 20(13), 2653–2662.

Kassam AH, HT van Velthuizen, AJB Mitchell, GW Fischer & MM Shah (1991). *Agro-Ecological land resources assessment for agricultural development planning: a case study of Kenya — Technical Annex 2: Soil Erosion and Productivity*. Land and Water Development Division, FAO, Rome, Italy.

Kaufman YJ (1989). The atmospheric effect on remote sensing and its correction. *In* Asrar G (ed.), *Theory and Application of Optical Remote Sensing*. 341 p. Wiley, New York.

Kaufman YJ (1993). Aerosol optical thickness and atmospheric path radiance. *Journal of Geophysical Reviews* 98: 2677–2692.

Kaufman YJ, A Wald, LA Remer, B Gao, R Li & L Flynn (1997). The MODIS 2.1 μm channel—correlation with visible reflectance for use in remote sensing of aerosol. *IEEE Transactions on Geoscience and Remote Sensing* 35, 1–13.

Koestler A (1967). *The Ghost in the Machine*, Vol. 4. 34–64, Random House, New York.

Kumar L, AK Skidmore & E Knowles (1997). Modelling topographic variation in solar radiation in a GIS environment. *International Journal for Geographical Information Science* 11, 475–497.

Lal R (1996). Deforestation and land-use effects on soil degradation and rehabilitation in Western Nigeria. I. Soil physical and hydrological properties. *Land Degradation & Development* 7, 19–45.

Lang S, T Langanke (2006). Object-based mapping and object-relationship modeling for land use classes and habitats. *Photogrammetrie, Fernerkundung, Geoinformation* 10(1), 5–18.

Lehmann A, JM Overton & JR Leathwick (2002). GRASP: generalized regression analysis and spatial predictions. *Ecological Modelling* 157, 189–207.

Lekasi KV, IV Sijali, P Gicheru, L Gachimbi, MK Nyangw'ara (2005). *Agricultural Productivity and Sustainable Land Management Project Report on Participatory Rapid Appraisal for Wusi Sub-location in the Taita Hills Catchment*. Kenya Agricultural Research Institute, Nairobi. Technical Report No. 10.

Lens L, S Van Dongen, K Norris, M Githiru & E Matthysen (2002). Avian Persistence in Fragmented Rainforest. *Science* 298, 1236–1238.

Levin SA (1992). The problem of pattern and scale in ecology. *Ecology* 73(6), 1943–1967.

Li G and Q Weng (2005). Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, USA. *Photogrammetric Engineering & Remote Sensing* 71, 947–958.

Liang S, H Fang & M Chen (2001). Atmospheric correction of Landsat ETM+ surface imagery – Part I: Methods. *IEEE Transactions on Geoscience and Remote Sensing* 39(11), 2490–2498.

Liang S, H Fang, JT Morisette, M Chen, CJ Shuey, CL Walthall & CST Daughtry (2002). Atmospheric correction of Landsat ETM+ land surface imagery – Part II: Validation and applications. *IEEE Transactions on Geoscience and Remote Sensing* 48(12), 2736–2746.

Liang S (2004). *Quantitative remote sensing of land surfaces*. 534 p. John Wiley & Sons, Inc., New Jersey.

Lillesand TM & RW Kiefer (1994). Remote sensing and image interpretation (3rd ed.). 750 p. John Wiley & Sons, New York.

Lin TH, AJ Chen, GR Liu & TH Kuo (2002). Monitoring the atmosphere aerosol optical depth with SPOT data in complex terrain. *International Journal of Remote Sensing* 23(4), 647–659.

Lo CP (1986). Accuracy of population estimation from medium-scale aerial photography. *Photogrammetric Engineering and Remote Sensing* 52, 1859–1869.

Lo CP (1989). A Raster Approach to Population Estimation Using High-Altitude Aerial and Space Photographs. *Remote Sensing of Environment* 27, 59–71.

Lo CP (1995). Automated Population and Dwelling Unit Estimation from High-Resolution Satellite Images: A GIS Approach. *International Journal of Remote Sensing* 16, 17–34.

Lovett JC & SK Wasser (eds.) (1993). *Biogeography and Ecology of the Rainforests of Eastern Africa*. Cambridge University Press, Cambridge.

Lu D, M Batistella & E Moran (2008). Integration of Landsat TM and SPOT HRG images for vegetation change detection in the Brazilian Amazon. *Photogrammetric Engineering and Remote Sensing* 74(4), 421–430.

Lung T & G Schaab (2006). Assessing fragmentation and disturbance of west Kenyan rainforests by means of remotely sensed time series data and landscape metrics. *African Journal of Ecology* 44 (4), 491–506.

Lyapustin AI & JL Privette (1999). A new method of retrieving surface bidirectional reflectance from ground measurements: Atmospheric sensitivity study. *Journal of Geophysical Research-Atmospheres* 104, 6257−6268.

Mahiny AS & BJ Turner (2007). A comparison of four common atmospheric correction methods. *Photogrammetric Engineering and Remote Sensing* 73(4), 361–368.

Martonchik JV, CJ Bruegge & A Strahler (2000). A review of reflectance nomenclature used in remote sensing. *Remote Sensing Reviews* 19, 9–20.

McCullagh P & JA Nelder (1989). *Generalized Linear Models* (2nd ed.). Chapman and Hall, New York.

Melillo JM, RA Houghton, DW Kicklighter & AD McGuire (1996). Tropical deforestation and the global carbon budget. *Annual Review of Energy & Environment* 21, 293–310.

Milton EJ, KP Lawless, A Roberts & SE Franklin (1997). The effects of unresolved scene elements on the spectral response of calibration targets: an example. *Canadian Journal of Remote Sensing* 23, 252–256.

Millward AA & JE Mersey (1999). Adapting the RUSLE to model soil erosion potential in a mountainous tropical watershed. *Catena* 38, 109–129.

Moran MS, RD Jackson, GF Hart, PN Slater, RJ Bartell, SF Biggar, DI Gellman & RP Santer (1990). Obtaining surface reflectance factors from atmospheric and view angle corrected SPOT-1 HRV Data. *Remote Sensing of Environment* 32, 203–214.

Moran MS, RD Jackson, PN Slater & PM Teillet (1992). Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. *Remote Sensing of Environment* 41, 169–184.

Moran MS, RD Jackson, TR Clarke, J Qi, F Cabot, KJ Thome & BL Markham (1995). Reflectance factor retrieval from Landsat TM and SPOT HRV data for bright and dark targets. *Remote Sensing of Environment* 52, 218–230.

Moran MS, R Bryant, K Thome, W Ni, Y Nouvellon, MP Gonzalez-Dugo, J Qi & TR Clarke (2001). A refined empirical line approach for reflectance retrieval from Landsat-5 and Landsat-7 ETM+. *Remote Sensing of Environment* 78, 71–82.

Moran MS, R Bryant, K Thome, CD Holifield & S McElroy (2003). Refined empirical line approach for retrieving surface reflectance from EO-1 ALI images. *IEEE Transactions on Geoscience and Remote Sensing* 41(6), 1411–1414.

Morgan RPC & DA Davidson (1991). *Soil Erosion and Conservation*. Longman Group, U.K.

Msagha J (2004). Population development in the Taita Hills, Kenya. *In* Pellikka P, J Ylhäisi & B Clark (eds.), *Taita Hills and Kenya, 2004 – seminar, reports and journal of a field excursion to Kenya*. Expedition reports of the Department of Geography, University of Helsinki 40, 39–46. ISBN 952-10-2077-6, Helsinki 2004, 148 pp.

Mubareka S, D Ehrlich, F Bonn & F Kayitakire (2008). Settlement location and population density estimation in rugged terrain using information derived from Landsat ETM and SRTM data. *International Journal of Remote Sensing* 29, 2339–2357.

Murakami T, S Ogawa, N Ishitsuka, K Kumagai & G Saito (2001). Crop discrimination with multitemporal SPOT/HRV data in the Saga Plains, Japan. *International Journal of Remote Sensing* 22, 1335–1348.

Myneni RB, S Hoffman, Y Knyazikhin, JL Privette, J Glassy, Y Tian, Y Wang, X Song, Y Zhang, GR Smith, A Lotsch, M Friedl, JT Morisette, P Votava, RR Nemani & SW Running (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment* 83, 214−231.

Navulur K (2007). *Multispectral image analysis using the object-oriented paradigm*. CRC Press, Boca Raton, FL.

Nichol JE & Wong MS (2005). Satellite remote sensing for detailed landslide inventories using change detection and image fusion. *International Journal of Remote Sensing* 26(9), 1913–1926.

Nicodemus FE, JC Richmond, JJ Hsia, IW Ginsberg & T Limperis (1977). *Geometrical considerations and nomenclature for reflectance*, Institute for Basic Standards, National Bureau of Standards: Washington, DC, USA.

Norjamäki I & T Tokola (2007). Comparison of atmospheric correction methods in mapping timber volume with multitemporal Landsat images in Kainuu, Finland. *Photogrammetric Engineering and Remote Sensing* 73(2), 155–163.

Ogawa S, G Saito, N Mino, S Uch Da, NM Khan & M Shafiq (1997). Estimation of soil erosion using USLE and Landsat TM in Pakistan. *Proceedings of 18th Asian Conference on Remote Sensing*, 20–24 October, 1997, Malaysia, (http://www.gisdevelopment.net/ aars/acrs/1997/ps3/ps3015.asp).

O'Neill RV, DL DeAngelis, JB Waide & TF Allen (1986). *A hierarchical concept of ecosystems*. Princeton University Press, Princeton.

O'Neill RV & AW King (1998). Homage to St. Michael; or, why are there so many books on scale. *In* Peterson DL & VT Parker (eds.): *Ecological scale theory and applications*, 3–15. Columbia University Press, New York.

Ozanne CMP, D Anhuf, SL Boulter, M Keller, RL Kitching, C Korner, FC Meinzer, AW Mitchell, T Nakashizuka, PL Silva Dias, NE Stork, SJ Wright & M Yoshimura (2003). Biodiversity meets the atmosphere: a global view of forest canopies. *Science* 301, 183–186.

Pellikka P (1998). Development of correction chain for multispectral airborne video camera data for natural resource assessment. *Fennia* 176, 1–110.

Pellikka PKE, BJF Clark, T Sirviö & K Masalin (2005). Environmental change monitoring applying satellite and airborne remote sensing data in the Taita Hills, Kenya. *Proceedings of the 1st International Conference on Remote Sensing and Geoinformation Processing in the Assessment and Monitoring of Land Degradation and Desertification*. Trier, Germany, September 7–9, 2005 (RGLDD).

Pellikka P, M Lötjönen, M Siljander & L Lens (2009). Airborne remote sensing of spatiotemporal change (1955–2004) in indigenous and exotic forest cover in the Taita Hills, south-east Kenya. *International Journal of Applied Earth Observation and Geoinformation* 11, 221–232.

Perry EM, T Warner & P Foote (2000). Comparison of atmospheric modelling versus empirical line fitting for mosaicking HYDICE imagery. *International Journal of Remote Sensing* 21(4), 799–803.

Petit C & Lambin EF (2001). Integration of multi-source remote sensing data for land cover change detection. *International Journal of Geographical Information Science* 8(15), 785–803.

Pickles J (1995). *Ground Truth: The Social Implications of Geographic Information Systems*. 248 p. The Guilford Press, New York.

Platt RV & L Rapoza (2008). An evaluation of an object-oriented paradigm for land use/land cover classification. *The Professional Geographer* 60(1), 87–100.

Plourde L & RG Congalton (2003). Sampling method and sample placement: how do they affect the accuracy of remotely sensed maps. *Photogrammetric Engineering and Remote Sensing* 69, 289–297.

Porter PW (1956). *Population Distribution and Land Use in Liberia*. PhD thesis, London School of Economics and Political Science, London.

Preston S (1996). The effect of population growth on environmental quality. *Population Research and Policy Review* 15, 95–108.

Remer LA, YJ Kaufman, D Tanré, S Mattoo, DA Chu, JV Martins, R-R Li, C Ichoku, RC Levy, RG Kleidman, TF Eck, E Vermote BH & Holben (2005). The MODIS Aerosol Algorithm, Products and Validation. *Journal of the Atmospheric Sciences*, Special Issue 62, 947–973.

Republic of Kenya (2001). *The 1999 Population & Housing Census*. Central Bureau of Statistics, Ministry of Planning and National Development, Kenya.

Riaño D, E Chuvieco, J Salas & I Aguado (2003). Assessment of different topographic corrections in Landsat TM data for mapping vegetation types. *IEEE Transactions on Geoscience and Remote Sensing* 41(5), 1056–1061.

Richter R, M Bachmann, W Dorigo & A Müller (2006). Influence of the adjacency effect on ground reflectance measurements. *IEEE Geoscience and Remote Sensing Letters* 3(4), 565–569.

Richter R (1997). On the in-flight absolute calibration of high resolution spaceborne sensors using small ground targets. *International Journal of Remote Sensing* 18(13), 2827–2833.

Rondeaux G, MD Steven, JA Clark & G Mackay (1998). La Crau; a European test site for remote sensing validation. *International Journal of Remote Sensing* 19(14), 2775–2788.

Rosenqvist Å, A Milne, R Lucas, M Imhoff & C Dobson (2003). A review of remote sensing technology in support of the Kyoto Protocol. *Environmental Science and Policy* 6, 441–455.

Ruotsalainen A (2008). *Enhancing local livelihoods in Taita Hills, Kenya: indigenous tree species as part of farmers´ livelihoods and environmental rehabilitation*. MSc thesis, University of Helsinki, Faculty of Science, Department of Geography.

Sandmeier S, C Müller, B Hosgood & G Andreoli (1998a). Sensitivity analysis and quality assessment of laboratory BRDF data. *Remote Sensing of Environment* 64, 176–191.

Sandmeier S, C Müller, B Hosgood & G Andreoli (1998b). Physical mechanisms in hyperspectral BRDF data of grass and watercress. *Remote Sensing of Environment* 66, 222–233.

Sandmeier SR & KI Itten (1999). A field goniometer system (FIGOS) for acquisition of hyperspectral BRDF data. *IEEE Transactions on Geoscience and Remote Sensing* 37, 978−986.

Santer R, XF Gu, G Guyot, JL Deuzé, C Devaux, E Vermote & M Verbrugghe (1992). SPOT calibration at the La Crau test site (France). *Remote Sensing of Environment* 41, 227–237.

Scepan J (1999). Thematic Validation of High-Resolution Global Land-Cover Data Sets. *Photogrammetric Engineering and Remote Sensing* 65(9), 1051–1060.

Schaepman-Strub G, M Schaepman, TH Painter, S Dangel & JV Martonchik (2006). Reflectance quantities in optical remote sensing - Definitions and case studies. *Remote Sensing of Environment* 103, 27–42.

Schnaiberg J, J Riera, MG Turner & PR Voss (2002). Explaining human settlement patterns in a recreational Lake District: Vilas County, Wisconsin, USA. *Environmental Management* 30, 24–34.

Schott J R, C Salvaggio & WJ Volchok (1988). Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment* 26(1), 1–16.

Schroeder TA, WB Cohen, C Song, MJ Canty & Z Yang (2006). Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. *Remote Sensing of Environment* 103, 16–26.

Secretariat of the Convention on Biological Diversity (2006) *Global Biodiversity Outlook 2*. 88 p. Montreal.

Sellers PJ, BW Meeson, FG Hall, G Asrar, RE Murphy, RA Schiffer, DE Strebel, FP Bretherton, RE Dickenson, RG Ellingson, CB Field, KF Huemmrich, CO Justice, JM Melak, NT Roulet, DS Schimel & PD Try (1995). Remote sensing of the land surface for studies of global change: Models – Algorithms - Experiments. *Remote Sensing of Environment* 51, 3–26.

Shaban MA & O Dikshit (2001). Improvement of classification in urban areas by the use of textural features: the case study of Lucknow city, Uttar Pradesh. *International Journal of Remote Sensing* 22, 565–593.

Shackelford AK & CH Davis (2003). A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing* 41(9), 1920–1932.

Siljander M (2010). *Geospatial environmental data modelling applications using remote sensing, GIS and spatial statistics.* PhD Thesis, Department of Geosciences and Geography, University of Helsinki, Finland.

Sirviö T, A Rebeiro-Hargrave & P Pellikka (2004). Geoinformation in gully erosion studies in the Taita Hills, SE-Kenya, preliminary results. *Proceedings of the 5th AARSE conference* (African Association of Remote Sensing of the Environment), 18–21 October, 2004, Nairobi, Kenya. CD-Publication, no page numbers.

Slater PN, SF Biggar, RG Holm, RD Jackson, Y Mao, MS Moran, JM Palmer & B Yuan (1987). Reflectance and radiance-based methods for the in-flight absolute calibration of multispectral sensors, *Remote Sensing of Environment* 22, 11–37.

Slaton MR, ER Hunt Jr. & WK Smith (2001).Estimating near-infrared leaf reflectance from leaf structural characteristics. *American Journal of Botany* 88, 278−284.

Smith GM & EJ Milton (1999). The use of the empirical line method to calibrate remotely sensed data. *International Journal of Remote Sensing* 24(5), 1143–1150.

Soini E (2005a). *Livelihood capital, strategies and outcomes in the Taita hills of Kenya*. 48 p. ICRAF Working Paper no. 8. World Agroforestry Centre, Nairobi, Kenya.

Soini E (2005b). Land use change patterns and livelihood dynamics on the slopes of Mt. Kilimanjaro, Tanzania. *Agricultural Systems* 85, 306–323.

Sombroek WG, HMH Braun & BJA Van der Pouw (1980). *The exploratory soil map and agro-climatic zone map of Kenya*. Report No. E1; Kenya Soil Survey, Nairobi.

Song C, CE Woodcock, KC Seto, M Pax Lenney & SA Macomber (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment* 75, 230–244.

Souza C, L Firestone, LM Silva & D Roberts D (2003). Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment* 87(4), 494–506.

Spot Image Corporation (1997). *DIMAP format description*. http://www.spotimage.com/automne_modules_files/standard/public/p555_fileLINKEDFILE1_cap.pdf (accessed 01.12.2009).

Stenberg P, M Rautiainen, T Manninen, P Voipio & M Mõttus (2008). Boreal forest leaf area index from optical satellite images: model simulations and empirical analyses using data from central Finland. *Boreal Environment Research* 13, 433–443.

Stocking M & N Murnaghan (2000). Land Degradation—Guidelines for Field Assessment. UNU/UNEP/PLEC Working Paper. Overseas Development Group, University of East Anglia, Norwich, (Available online at: www.unu.edu/env/plec/l-degrade/index-toc.html

Strahler AH, CE Woodcock & JA Smith (1986). On the nature of models in remote sensing. *Remote Sensing of Environment* 20, 121–139.

Suomalainen J (2006). *Multiangular spectrometry and optical properties of debris covered surfaces*. M.S. thesis, Department of Physics, University of Helsinki, Available at: http://ethesis.helsinki.fi/julkaisut/mat/fysik/pg/suomalainen/

Suomalainen J, T Hakala, J Peltoniemi & E Puttonen (2009). Polarised multiangular reflectance measurements using the Finnish Geodetic Institute Field Goniospectrometer. *Sensors* 9, 3891–3907.

Sutton P, C Roberts, C Elvidge & H Meij (1997). A comparison of night-time satellite imagery and population density for the continental united states. *Photogrammetric Engineering and Remote Sensing* 63, 1303–1313.

Swets K (1988). Measuring the accuracy of diagnostic systems. *Science* 240, 1285–1293.

Teillet PM, B Guindon & DG Goodenough (1982). On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing* 8, 84–106.

Teillet PM & G Fedosejevs (1995). On the dark target approach to atmospheric correction of remotely sensed data. *Canadian Journal of Remote Sensing* 21, 373–387.

Teillet PM, G Fedosejevs, RP Gauthier, NT O'Neil, KJ Thome, SF Biggar, H Ripley & A Meygret (2001). A generalized approach to the vicarious calibration of multiple Earth observation sensors using hyperspectral data. *Remote Sensing of Environment* 77, 304–327.

Thome K, B Markham, J Barker, P Slater & S Biggar (1997). Radiometric calibration of Landsat. *Photogrammetric Engineering and Remote Sensing* 63(7), 853–858.

Thome KJ (2001). Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method. *Remote Sensing of Environment* 78, 27–38.

Thuiller W, MB Araújo S & Lavorel (2003). Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science* 14, 669–680.

Tso B & PM Mather (2001). *Classification methods for remotely sensed data*. Taylor & Francis, London.

Tucker CJ & PJ Sellers (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing* 7, 1395–1416.

United Nations Population Division (1999). Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, *The World at Six Billion*,

United Nations Population Division (2008). Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, *World Population Prospects: The 2008 Revision*, http://esa.un.org/unpp.

UNEP (1992). *World Atlas of Desertification.* London: Edward Arnold.

UNEP (2005). *One Planet Many People, Atlas of Our Changing Environment.* Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi.

UNEP (2009a). *Climate Change Science Compendium 2009*, 68 p. United Nations Environment Programme.

UNEP (2009b). *Kenya: Atlas of Our Changing Environment*, 168 p. United Nations Environment Programme. Available at:
(www.unep.org/publications/contents/pub_details_search.asp?ID=4014)

UNEP (2010a). 2010 International Year of Biodiversity, United Nations Environment Programme, Multimedia Resources: www.unep.org/newscentre/multimedia.

UNEP (2010b). Press Release January 2010: *Tree planting in Kenya's Mau Complex signals new beginnings for a critical ecosystem.* Available at:
(www.unep.org/Documents.Multilingual/Default.asp?DocumentID=608&ArticleID=6447&l=en)

UNEP/GRID (2006). *UNEP/GRID spatial data clearinghouse*. Available at: http://grid2.cr.usgs.gov/datasets/datalist.php (accessed 9 October 2009).

UN-MDG (2009). *The Millennium Development Goals Report 2009*, 56 p. Department of Economic and Social Affairs of the United Nations Secretariat, New York.

Vadrevu KP, A Eaturu & KVS Badarinath (2006). Spatial distribution of forest fires and controlling factors in Andhra Pradesh, India using Spot satellite datasets. *Environmental Monitoring and Assessment* 123, 75–96.

Venables WN & BD Ripley (2002). *Modern Applied Statistics with S* (4th ed.). 495 p. Springer–Verlag, New York.

Vermote EF, D Tanré, JL Deuzé, M Herman & JJ Morcette (1997a). Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Transactions on Geoscience and Remote Sensing* 35(3), 675–686.

Vermote EF, D Tanré, JL Deuzé, M Herman & JJ Morcette (1997b). *6S User Guide* (Version 2). Available from: ftp://kratmos.gsfc.nasa.gov/pub/6S/ (accessed 01.2010).

Ward J. C Dull, G Hertel, J Mwangi, S Madoffe & K Douce (2004). Monitoring for Sustainable Forestry and Biodiversity in the Eastern Arc Mountains of Tanzania and Kenya. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34, Part XXX. (see also: USDA Forest Service,

1999, Monitoring Changes in Forest Condition, Land Fragmentation and Conversion: Eastern Arc Mountains of Tanzania and Kenya http://www.easternarc.org/landchange/arcfinal1100_files/frame.htm.

Weber C & A Puissant (2003). Urbanization pressure and modelling of urban growth: example of the Tunis Metropolitan Area. *Remote Sensing of Environment* 86, 341–352.

Wheeler RJ, SR LeCroy, CH Whitlock, GC Purgold & JS Swanson (1994). Surface characteristics for the alkali flats and dunes regions at White Sands Missile Range, New Mexico. *Remote Sensing of Environment* 48, 181–190.

Wischmeier WH, DD Smith & RE Uhland, (1958). Evaluation of factors in the soil-loss equation. *Agricultural Engineering* 39, 458–462.

Wischmeier WH & DD Smith (1978). *Predicting rainfall erosion losses: a guide to conservation planning*. 58 p. Agricultural Handbook, No. 587. Science and Education Administration, US Department of Agriculture, Washington, DC, USA.

Wood D & J Fels (1992). *The Power of Maps*. The Guildford Press, New York

Wu J & DJ Marceau (2002). Modelling complex ecological systems: An introduction, *Ecological Modelling* 153, I-2: 1–6.

Wu J (1999). Hierarchy and scaling: extrapolating information along a scale ladder. *Canadian Journal of Remote Sensing* 25, 367–380.

Wu J & R Hobbs (Eds) (2007). *Key Topics in Landscape Ecology*. Cambridge University Press, Cambridge.

Wu J & OL Loucks (1995). From the balance-of-nature to hierarchical patch dynamics: a theoretical framework shift in ecology. *Quarterly Review of Biology* 70, 439–466.

Wulder MA, CR Butson & JC White (2008). Cross-sensor change detection over a forested landscape: Options to enable continuity of medium spatial resolution measures. *Remote Sensing of Environment* 112, 796–809.

Wu J, D Wang & ME Bauer (2005). Image-based atmospheric correction of QuickBird imagery of Minnesota cropland. *Remote Sensing of Environment* 99, 315–325.

Xu, J & J Huang (2006). Refined empirical line method to calibrate IKONOS imagery. *Journal of Zhejiang University SCIENCE A* 7(4), 641–646.

Yee TW & ND Mitchell (1991). Generalized Additive Models in Plant Ecology. *Journal of Vegetation Science* 2, 587–602.

Yuan D & CD Elvidge (1996). Comparison of relative radiometric normalization techniques. *ISPRS Journal of Remote Sensing* 51, 117–126.

Zar JH (1999). *Biostatistical Analysis* (4th ed.). 929 p. Prentice-Hall, London.

Zimmermann NE (2000) *Shortwavc.aml*. Available online at: www.wsl.ch/staff/niklaus.zimmermann/programs/aml1_1.html (accessed January 2009).

Working on this thesis has led me on a fascinating journey to Finland and to East Africa, and has further deepened my appreciation of what a wondrous and precious planet we inhabit. It has been a privilege to have the opportunity to work on environmental issues, and I hope this experience is the start of many more fruitful projects to come. I know I will never stop learning and caring passionately about our home, the Earth.

Department of Geosciences and Geography A
ISSN-L 1798-7911
ISSN 1798-7911 (print)
ISBN 978-952-10-6305-3 (paperback)
ISBN 978-952-10-6306-0 (PDF)
http://ethesis.helsinki.fi

Helsinki University Print
Helsinki 2010

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE