Mapping Vegetation Types in a Savanna Ecosystem in Namibia: Concepts for Integrated Land Cover Assessments

Christian Hüttich

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Friedrich-Schiller-Universität Jena

Mapping Vegetation Types in a Savanna Ecosystem in Namibia: Concepts for Integrated Land Cover Assessments

Dissertation

zur Erlangung des akademischen Grades *doctor rerum naturalium*
(Dr. rer. nat.)

vorgelegt dem Rat der Chemisch-Geowissenschaftlichen Fakultät
der Friedrich-Schiller-Universität Jena

von Dipl.-Geogr. Christian Hüttich
geboren am 12. Mai 1979 in Jena

Gutachter:

1. Prof. Dr. Christiane Schmullius

2. Prof. Dr. Stefan Dech (Universität Würzburg)

Tag der öffentlichen Verteidigung: 29.06.2011

Acknowledgements

This dissertation would not have been possible without the help of a large number of people. First of all I want to thank the scientific committee of this work: Prof. Dr. Christiane Schmullius and Prof. Dr. Stefan Dech for their always present motivation and fruitful discussions.

I am extremely grateful to my scientific mentor Prof. Dr. Martin Herold, for his never-ending receptiveness to numerous questions from my side, for his clear guidance for the definition of the main research issues, the frequent feedback regarding the structure and content of my research, and for his contagious enthusiasm for remote sensing of the environment.

Acknowledgements are given to the former remote sensing team of "BIOTA Süd": Dr. Ursula Gessner (DFD-DLR), Dr. Rene Colditz (CONABIO, Mexico), Manfred Keil (DFD-DLR), and Dr. Michael Schmidt (CONABIO, Mexico) for extremely fruitful discussions, criticism, and technical support during all phases of the dissertation. I am very thankful for a number of impressive field campaigns in Namibia and South Africa. I also have to thank a lot Dr. Jens Oldeland (University of Hamburg) for his motivation and valuable feedback to my manuscripts and very forthcoming brainstorming.

I am very grateful to Ben Strohbach and Franziska Kangombe from the National Botanical Research Institute of Namibia, for their extremely helpful discussions on vegetation ecology in Namibia, for valuable refinements of the manuscripts, for the provision of the vegetation data, and the guidance during our field campaigns.

I really appreciate Dr. Christopher Conrad for scientific support and to ensure a very open, friendly, and inspiring working atmosphere. I am really thankful to my colleagues from the Institute of Geography and Geology's Remote Sensing Unit in Würzburg: Sarah Asam, Anna Cord, Sebastian Fritsch, Vitus Himmler, Dr. Doris Klein, Pattrick Knöfel, Dr. Tobias Landmann, Sylvia Lex, Fabian Löw, Dr. Miriam Machwitz, Dr. Noellie Yahou Rüth, Gunther Schorcht, Martin Schmidt, Dr. Matthias Schramm, Michael Thiel, Yvonne Walz, Dr. Martin "Oggi" Wegmann, Julian Zeidler. Special thanks to all of you for reviewing and improving the earlier manuscripts of the dissertation.

I am really grateful to work in the 'Biotop' with Ursula, Miriam, and Oggi, Finn, and later with Anna and Noellie. Without their fellowship and the valuable help to a large number of specific problems this dissertation would not have been finished jet. Acknowledgements are also given to my student assistants: Sylvia Lex, Michael Gottschling, Michael Wiesner, and Peter Zellner. They did a great job during hours of remote sensing data pre-processing.

I am deeply grateful to my parents Conny and Frank Hüttich. Without their motivation and financial support during my studies this dissertation would not have been realised. I am very thankful to my brothers Georg and Martin Hüttich, my grandmothers Marianne Frenzl and Renate Hüttich, and my friends for their tolerance and demonstrating me from time to time that there is a life besides remote sensing.

This dissertation is dedicated to my wife Franziska, for her love, patience, and her understanding for too long office hours, my frequent absence during the field conference trips, and for her mental support.

This dissertation has been carried out in the framework of the "BIOTA Southern Africa" project funded by the Federal Ministry of Education and Research (BMBF).

Jena, April 2011 Christian Hüttich

Abstract

Mapping Vegetation Types in a Savanna Ecosystem in Namibia: Concepts for Integrated Land Cover Assessments

CHRISTIAN HÜTTICH

The characterisation and evaluation of biodiversity and land-cover in Southern Africa's Savannas is a major prerequisite for suitable and sustainable land management and conservation purposes. However, mechanisms for frequent update of the status and trends of biodiversity and land change processes are still missing. The knowledge of the spatial distribution of vegetation types is an important information source for all social benefit areas. Remote sensing techniques are essential tools for mapping and monitoring of land-cover. The development and evaluation of concepts' for integrated land-cover assessments attracted increased interest in the remote sensing community since evolving standards for the characterisation of land-cover enable an easier access and inter-comparability of earth observation data.

Regarding the complexity of the savanna biome in terms of the spatiotemporal heterogeneity of the vegetation structure and rainfall variability, the main research needs are addressing the assessment of the capabilities and limitations of using satellite data for land-cover and vegetation mapping purposes. While integrating Moderate Resolution Imaging Spectroradiometer (MODIS) time series data for mapping vegetation types in Namibia, the temporal characteristics of semi-arid life-forms types were used for the classification of vegetation types in Namibia. The Random Forest framework was applied and evaluated for classifying vegetation types using MODIS time series metrics as input features. The study region comprised the Kalahari in the north-eastern communal lands of Namibia.

Regarding of the evolving global standardisation process for land-cover, there is the need to report the capabilities and limitations of the FAO and UNEP Land Cover Classification System (LCCS) in regional case studies. LCCS is evaluated in terms of the applicability in open savanna ecosystems and as ontology for the semantic integration of an *in-situ* vegetation database in a coarse scale mapping framework based on MODIS data. Further, the capabilities of the methodological setups of global land-cover mapping initiatives are assessed while using the results of the integrated vegetation type mapping framework. In order to assess the existing accuracy uncertainties of mapping savannas at global scales, the effects of composite length and varying observation periods were compared in terms of mapping accuracy. The implications for global monitoring were discussed. The determinants of precipitation amount and mapping accuracy were evaluated by comparing MODIS and Tropical Rainfall Measuring Mission (TRMM) time series data.

Integrating multi-scale land-cover information, such as life form, cover, and height of vegetation types (*in-situ*), vegetation physiognomy and local patterns (Landsat), and phenology (MODIS) in an ecosystem assessment framework, resulted in a flexible land-cover map including a broad structural-physiognomic and a phytosociological legend. The principle of classifiers and modifiers in LCCS proved to be applicable in dry savanna ecosystems and can be confirmed as overarching land-cover ontology. Analyses of time series classifications showed that mapping accuracy increases with increasing observation period. Small composite period lengths lead to increased mapping accuracies. The relationship between mapping accuracy and observation period was observed as a function of precipitation input and the magnitude of change between land-cover stages.

The integration of *in-situ* data in a multi-scale framework leads to improved knowledge of the regionalisation of Namibian vegetation types. On the one hand, the case study in the north-eastern Kalahari showed that multi-data mapping approaches using *in-situ* to medium resolution MODIS time series data bear the potential of the wall-to-wall update of existing vegetation type maps. On the other hand, the global remote sensing community can extend the reference databases by integrating regional standardised biodiversity and ecotype assessments in calibration and validation activities. The studies point on the uncertainties of mapping savannas at global scales and suggest possible solutions for improvements by adapting the remotely sensed feature sets, classification methods, and integrating dynamic processes of semi-arid ecosystems in the mapping framework.

Keywords: Vegetation types, land-cover, savanna, MODIS time series, Random Forest, LCCS, TRMM, Namibia

Deutsche Zusammenfassung

Vegetationskartierung in der Trockensavanne Namibias unter Nutzung botanischer Felddaten und multitemporaler Satellitendaten: Konzepte zur integrativen Erfassung der Landbedeckung

CHRISTIAN HÜTTICH

Die Erfassung und Bewertung des Zustandes der Landbedeckung und der damit zusammenhängenden Biodiversität im südlichen Afrika stellt eine Grundvoraussetzung für ein nachhaltiges Landmanagement und dem Schutz natürlicher Ressourcen dar. Die Kenntnis der räumlichen Verbreitung von Vegetationstypen ist eine wichtige Informationsgrundlage in Namibia, die zahlreiche gesellschaftliche Bereiche betrifft (z.B. Weidequalität, Tourismus). Existierende Kartierungen der Verbreitung von Vegetationstypen auf nationaler Ebene basieren auf Arbeiten der 70er Jahre und bedürfen folglich einer flächendeckenden Aktualisierung. Methoden der Fernerkundung erweisen sich dabei als wichtiges Werkzeug. Einflüsse der intra- und interannuellen Variabilität der Vegetationsdynamik semi-arider Savannen auf die Qualität fernerkundungsbasierter Kartierungen auf regionalem und globalem Maßstab wurden bisher unzureichend untersucht. Die Entwicklung und Bewertung skalenübergreifender integrativer Konzepte zur Erfassung und Bewertung der Landbedeckung ist in das Zentrum des wissenschaftlichen und politischen Interesses in den Bereichen der Geowissenschaften und Ökosystemforschung gerückt. Das *Land Cover Classification System* der FAO und UNEP bietet eine Schnittstelle zur technischen und semantischen Integration von Erdbeobachtungsinformationen und ist in der Internationalen Organisation für Normung (ISO) als internationaler Standard für Landbedeckung unter Begutachtung. In diesem Zusammenhang sind Erfahrungen und Schnittstellen zur technischen und semantischen Integration von Erdbeobachtungsinformationen in den Savannen des südlichen Afrikas bisher unzureichend dokumentiert.

Basierend auf dem Forschungsbedarf lassen sich folgende Forschungsfragen ableiten:

- Wie lassen sich Klassifikationsmethoden von Satellitenzeitreihen an semi-aride Savannen anpassen, um der hohen räumlichen und zeitlichen Variabilität der Vegetationsdynamik gerecht zu werden?
- Wie können Erdbeobachtungsdaten von *in-situ* bis regionalem Maßstab standardisiert und in zur Regionalisierung zur Verbreitung von Vegetationstypen in Namibia beitragen?
- Wo liegen die Stärken und Schwächen des *Land Cover classification System* (LCCS) der FAO und UNEP unter Anwendung zu Landbedeckungskartierung in semi-natürlichen semi-ariden Savannen?
- Wie kann die globale Fernerkundungsgemeinschaft von regionalen Fallstudien zur Erfassung der Biodiversität und Vegetationsverbreitung profitieren?

Hinsichtlich der räumlichen und zeitlichen Heterogenität der Vegetationsstruktur und Niederschlagsmuster in den Savannen des südlichen Afrikas war die im Rahmen einer **integrativen Vegetationstypenkartierung** die Potentialanalyse einer integrativen Nutzung von *in-situ*-Daten, hochaufgelösten Satellitendaten (Landsat) und mittel aufgelösten Satellitenzeitreihen (MODIS) zur Kartierung von Vegetationstypen in Namibia ein wesentliches inhaltliches Ziel dieser Studie. Im Weiteren wurden die Methoden der *Random Forest* Klassifikation für die Kartierung von Vegetationstypen in einem Testgebiet in der Kalahari im Nordosten Namibias unter der Verwendung von MODIS Zeitreihenmassen evaluiert. Hinsichtlich der Notwendigkeit der **Evaluierung des *Land Cover Classification System* (LCCS)** der FAO und UNEP als internationaler Standard zur Klassifikation der Erdoberfläche in regionalen Fallstudien sollte in dieser Arbeit die Eignung von LCCS zur Klassifikation der Landbedeckung in Savannen unter Verwendung von multitemporalen Fernerkundungsdaten getestet werden. Weiterhin wurde das Potential von LCCS als Ontologie zur semantischen Integration von *in-situ* Vegetationsdaten in ein hierarchisches Kartierschama bewertet. Das Ergebnis der integrierten Nutzung von *in-situ*-,

Landsat-, und MODIS Daten in einem standartisierten Kartierungsschema wurde weiterhin genutzt, um die Möglichkeiten und Grenzen der technischen Optimierung der Landbedeckungskartierung in Savannen im Rahmen globaler Kartierungsinitiativen zu evaluieren. Der dritte inhaltliche Schwerpunkt lag auf der Methodenentwicklung zur zeitlichen Übertragbarkeit des Kartierschemas und der Analyse der **Implikationen für ein großflächiges Monitoring** semi-arider Savannen. Um die existierenden Ungenauigkeiten bei der Kartierung von Vegetionsstruktur in Savannen zu erklären, wurden die Auswirkungen von Kompositlänge und variierenden Beobachtungszeiträumen auf die Kartengenauigkeit analysiert und hinsichtlich der Auswirkungen auf das globale Monitoring von Savannen diskutiert. Die Einflussgrösse des Niederschlagseintrages und dessen Einfluss auf die Klassifikationsgenauigkeiten bei der Landbedeckungsklassifikation unter Verwendung von MODIS-Zeitreihen wurden unter Einbezug von Niederschlagszeitreihen des satellitengetragenen Sensors *Tropical Rainfall Measuring Mission* (TRMM) bewertet.

Die synergetische Nutzung multi-skaliger Landbedeckungsinformationen wie Wuchsform, flächenhafte Bedeckung und Höhe verschiedener Vegetationstypen (auf *in-situ* Ebene), lokale Muster der Vegetationsphysiognomie (Landsat) und Informationen zur Vegetationsdynamik und Phänologie (MODIS) zur Vegetionskartierung resultierte in einer ‚flexiblen' Landbedeckungskarte. Dies ermöglicht die Nutzung der Landbedeckungskarte unter Anwendung eines standardisierten hierarchischen Legendensystemes, einer Legende zur Vegetationsphysiognomie und einer Legende verschiedener Pflanzengemeinschaften. Das Prinzip der LCCS- spezifischen *Classifier* und *Modifier* erwies sich als anwendbar und kann in den Trockensavannen Namibias als Klassifikationssystem zur Beschreibung der Landoberfläche bestätigt werden. Die Analysen der Klassifikation von Zeitreihen zeigten einen Anstieg der Kartierungsgenauigkeiten mit steigender Beobachtungsdauer. Weiterhin führten geringere Kompositlängen zu höheren Klassifikationsgenauigkeiten. Der Zusammenhang zwischen Kartierungsgenauigkeit und Beobachtungszeitraum wurde als Funktion des Niederschlagseintrages und der Variabilität des Landbedeckungstypes beschrieben.

Die Integration von *in-situ*-Daten in ein multi-skaliges Kartiervorhaben führte zu einer Optimierung hinsichtlich der Regionalisierung von Vegetationstypen in Namibia. Einerseits bestätigte die Fallstudie in der Kalahari, dass multi-skalige Kartieransätze unter Einbezug von *in-situ*-Daten bis hin zu ‚grob' aufgelösten MODIS Zeitreihen das Potential für eine Neuauflage der nationalen Vegetationskarte Namibias besitzen. Andererseits zeigten diese Studien, dass globale Fernerkundungsprojekte durch fortschreitende Datenstandards profitieren können, indem regionale Datensätze zur Biodiversitätserfassung und der biogeographischen Kartierung in großräumige Kalibrierungs- und Validierungsakivitäten einbezogen werden. Letztendlich zeigen die Studien noch bestehende Ungenauigkeiten in der Kartierung von Savannen auf globalem Maßstab auf. Lösungsansätze wurden vorgeschlagen, die darauf abzielen, die fernerkundlichen Eingangsvariablen und Klassifikationsmethoden auf die dynamischen Prozesse der Landbedeckung im semi-ariden Raum anzupassen.

List of Tables

List of Figures

Glossary

AEZ	Agro-Ecological Zones
ANOVA	Analyses of Variance
AVHRR	Advanced Very High Resolution Radiometer
AWiFS	Advanced Wide Field Sensor
B-distance	Bhattacharyya Distance
BIOTA	Biodiversity Monitoring Transect Analysis
CART	Classification and Regression Trees
CBD	Convention on Biological Diversity
CEOS	Committee of Earth Observation Satellites
DCA	Detrended Correspondence Analysis
DOY	Day of the Year
DMP	Desert Margin Project
DT	Decision Trees
EGS	End of the Growing Season
EOS	End of Season
ETM	Enhanced Thematic Mapper
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organisation
GEO	Group on Earth Observation
GEO-BON	Group on Earth Observation - Biodiversity Observation Network
GEOSS	Global Environmental Observation System of Systems
GIMMS	Global Inventory Modelling and Mapping Studies
GIS	Geographic Information System
GLCC	Global Land Cover Characterisation
GLCN	Global Land Cover Network
GLC2000	Global Land Cover 2000
GOFC-GOLD	Global Observation of Forest Cover – Global Observation of Land Dynamics
GPS	Global Positioning System
IGOL	Integrated Global Observations for Land
IGBP	International Geosphere-Biosphere Programme
IGBP-DIS	International Geosphere-Biosphere Programme Data and Information Systems Initiative
IRS	Indian Remote Sensing Satellites
ISO	International Standardisation Organisation
ITC	Intertropical Convergence Zone
LCCS	Land Cover Classification System
LCML	Land Cover Meta Language
LDCM	Landsat Data Continuity Mission
LPC	Land Product Validation Subgroup
LGS	Length of the Growing Season
LOS	Length of Season
LUCC	Land Use and Land Cover Change
NBRI	National Botanical Research Institute of Namibia
NDVI	Normalised Difference Vegetation Index
MAP	Mean Annual Precipitation
MGS	Mid Position of the Growing Season
MERIS	Medium Resolution Imaging Spectrometer
MIR	Middle Infrared
MODIS	Moderate Resolution Imaging Spectroradiometer
MVC	Maximum Value Composite
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration
NPP	Net Primary Production
OOB	Out-of-Bag

UMD	University of Maryland
UN	United Nations
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
USGS	United States Geological Survey
SGS	Start of the Growing Season
SOS	Start of Season
SOTER	Global Soil and Terrain Database
SPOT-VGT	SPOT Vegetation
TiSeG	Time Series Generator
TM	Thematic Mapper
TRMM	Tropical Rainfall Measurement Mission
TWINSPAN	Two Way Indicator Species Analysis
VCF	Vegetation Continuous Field

Table of Content

1 Introduction

1.1 Research motivations and questions

Land-cover, describing the physical state of the earth's surface, is a key information source for a variety of social sectors. Geographic information products describing the spatial distribution of land-cover are key management resources for environmental and socio-economic planning but also constitute an essential parameter for environmental modelling at all scales.

Substantial progress in all technical engineering disciplines yielded to increasing accessibility of environmental geo-information data. Satellite and airborne technologies play a crucial role for the measurement, provision, and processing of land-cover information. As a result, the user and producer community of remotely sensed land-cover data increased during the last decades, subsequently the implementation of earth observation technologies found its way into a wider range of scientific disciplines. Information exchange implicates the association of different conceptual perceptions, academic paradigms, and thematic definitions on the same research subject. For the scientific community the consequence was an increase of available land-cover information and the need to create interfaces between land-cover datasets from *in-situ* to global scale observations.

As a result, extensive earth observation data archives were established, such as datasets of 35 years of Landsat imagery (White *et al.*, 2008), the long-term Normalised Difference Vegetation Index (NDVI) data from the Global Inventory Monitoring and Modelling Study (GIMMS, Pettorelli *et al.*, 2005) derived from the Advanced Very High Resolution Radiometer (AVHRR), or ten years of satellite time series imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS). Improved earth observation data archives were made available for integrated studies on different spatial scales.

One of the main challenges for the operationalisation of land-cover characterisation is the creation of mechanisms for the generation of "compatible and comparable" map products. Effort has been made for the harmonisation and standardisation of different conceptual perceptions of land-cover. On the one hand, harmonisation follows a bottom-up process through the removal of inconsistencies and the assessment of similarities of existing classification systems. On the other hand, standardisation is a top-down process, and thus more rigid, since different existing land-cover definitions were transferred to one common definition. The Land Cover Classification System (LCCS, Di Gregorio, 2005), developed by the UN Food and Agriculture Organisation (FAO) has evolved as the most accepted classification system and acts as a "common language for land cover characterisation" (Herold *et al.*, 2006). LCCS is being proved as an international standard in the International Organisation for Standardisation (ISO, Herold *et al.*, 2006). The LCCS-based standardisation of the global land-cover databases has been proven successful, but challenging (Herold *et al.*, 2008).

In general, there is a need to foster and improve the vertical harmonisation process which means, for example, the integration of the best available land-cover data on various scales (e.g. field plots, fine resolution satellite data, and satellite time series). Further on, it has to be analysed what the benefits of an increasing cooperation for different user communities are. There is very little experience in the vertical harmonisation at regional scales. Experiences have shown that the application of LCCS-based land-cover mapping is problematic in spatiotemporally dynamic savanna ecosystems (Cord *et al.*, 2010). The capabilities and limitations of LCCS have not been proven in practice, at regional scales, and in the savanna biome.

Comparative studies of the global 1 km land-cover datasets derived from different mapping initiatives (IGBP DISCover, Loveland *et al.*, 2000; University of Maryland, Hansen *et al.*, 2000; MODIS, Friedl *et al.*, 2002; and Global Land Cover 2000, Bartholomé & Belward, 2005) showed limited agreements of the class-specific accuracies. Despite of differing legends among the global datasets, comparisons of land-cover types related to the savanna biome show low spatial agreement and comparatively low mapping accuracies with 60.3 % for herbaceous vegetation and 65.8 % for shrublands (Herold *et al.*, 2008). One of the reasons might be the fact that savanna ecosystems are characterised by zones of fuzzy transitions of different life forms. Related to that, a main problem for remote sensing applications is the highly heterogeneous spatial distribution of trees, shrubs, and grasses in savannas. Each life form has specific phenological characteristics, and underlies different environmental cues (e.g. soil moisture, precipitation, and temperature, Childes, 1989; Archibald & Scholes, 2007; Privette *et al.*, 2004; Sankaran *et al.*, 2005). Since all global land-cover maps are based on the statistical classification of spectral and temporal features, the tree-grass coexistence and the phenological characteristics of different plant communities can influence the classification result in terms of accuracy and stability. Regarding the distinct inter-annual variability of the vegetation dynamics, the effects of such environmental constraints on the mapping performance have poorly been analysed for arid and semi-arid ecosystems.

Considering the training database generation (supervised classification) and class labelling process (unsupervised classification), all of the above mentioned coarse scale mapping initiatives are based on broad top-down approaches which hinders the possibility to analyse effects of environmental cues and constraints on the classification result. Moreover, the assembly of predictor variables derived from remotely sensed time series traditionally follows a 'broad' approach capturing the characteristics of global bioregions. Little research has been carried out to adapt the remotely sensed feature space to the specific ecosystem dynamics. Considering the variations of the spatial agreement of savanna land-cover types and the considerably low mapping accuracies, research has to be addressed in order to adapt time series classification techniques to the temporal variations in seasonal environments, such as phenological cycling of certain vegetation types.

A number of international panels aim at the development and implementation of standards and integrated concepts for global environmental monitoring (Townshend *et al.*, 2008). Examples are given with the Global Observations of Forest Cover and Land Dynamics (GOFC-GOLD), the Global Land Cover Network (GLCN) of the FAO and UNEP, the Global Earth Observation System of Systems (GEOSS), or the Group of Earth Observations Biodiversity Observation System (GEO-BON). The goal that these programs and panels have in common is the integration of *in-situ*, regional, and global land-cover observations which is closely linked with harmonisation and standardisation tasks. However, currently there is a lack of practical experience and implementation for up-, and down-scaling of earth observation data in specific case studies. Research is needed in order to report the positive feedbacks of integrated concepts to the included research and user communities.

Ecological applications focus on the assessment of ecosystem structure, function, biogeochemistry, at the patch, landscape and regional scale, or on the calibration and validation programs for satellite-derived biophysical parameters. Such research objectives were realised in numerous research programs in sub-Saharan Africa, such as the Biodiversity Transect Monitoring program (BIOTA Southern Africa, Krug *et al.*, 2006), the southern African regional Science initiative (SAFARI 2000, Swap *et al.*, 2002), and studies along the Kalahari Transect (Shugart *et al.*, 2004). Thus, the number, value, and importance of ecological applications based on satellite observations have increased during the last decades (Leimgruber *et al.*, 2005). The central issues being investigated in arid and semi-arid ecosystems are the assessment of land degradation and the related loss of biodiversity, which is a major ecosystem service in the savanna biome. Examples for the value of functional diversity are given by (a) the balanced coexistence of tree and grass species (to avoid bush encroachment) and maintain a sustainable carrying capacity for livestock, (b) the economic value of a diverse and healthy plant and animal life since large proportions of the local business are based on tourism, and (c) the knowledge of the spatial distribution of numerous savanna vegetation types that are in particular important for large-scale nature conservation and rangeland management purposes.

More than 60 % of Namibia's semi-natural land is not surveyed in terms of vegetation distribution. The spatial-explicit knowledge of plant communities is key decision support information for the resource management and land use planning on the administrative level (Burke & Strohbach, 2000; Strohbach *et al.*, 2004; Strohbach & Petersen, 2007). Currently, there is a lack of wall-to-wall geo-information of the distribution of vegetation types in Namibia. The dominating land use in the semi-arid savanna in Namibia is livestock farming (Figure 1-1). Most parts of the arid lands of the Namib Desert are wildlife- and nature reserves. These land uses are depending on the availability of natural resources such as plant diversity, animal wildlife, and the geological situation. Some areas are being studied intensively, such as the Etosha National Park (Du Plessis, 1999).

The use of remote sensing techniques has been proven to be successful in a number of mapping projects that were conducted at national and regional scales in Namibia in order to improve existing vegetation maps or extrapolate botanical field data. The former aimed at the re-mapping of the current vegetation map of Namibia after Giess (1971) using annual MODIS time series. By using phenological and spectral features, transition zones between broad vegetation units could be refined (Colditz *et al.*, 2007). The second task aimed at the spatial extrapolation of *in-situ* plot data of plant communities. Regional projects were realised in parts of Namibia, typically based on single Landsat images (Verlinden & Laamanen, 2001; Strohbach *et al.*, 2004; Vogel, 2005). The problems arising from the previous mapping activities can be seen in a lack of interoperability, linkages and operational workflows between satellite and *in-situ* data covering local to national scales. The potential of the integrated use of *in-situ* and satellite data with different spatial and temporal characteristics (e.g. Landsat and MODIS) has not been assessed in the complex dynamic ecosystems of Namibia.

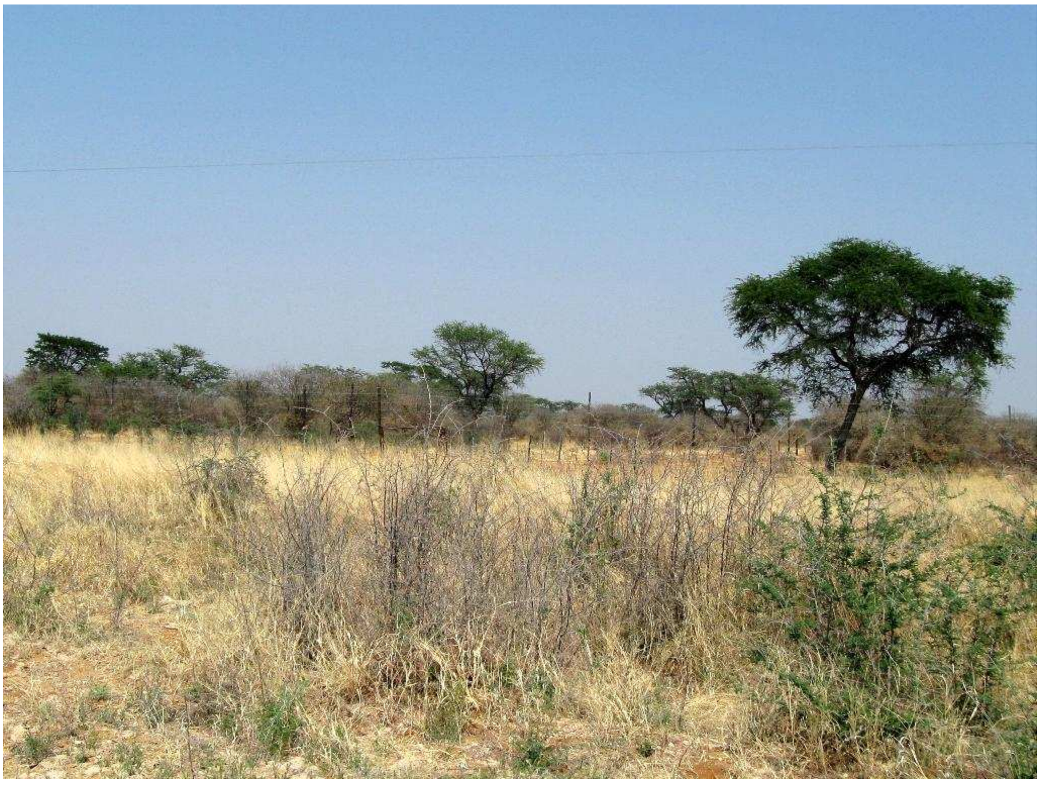


Figure 1-1 The image, taken in the dry season (September 2008) in north-eastern Namibia, shows the typical landscape of the Kalahari, characterised by the coexistence of woody and herbaceous life forms with differing phenological cycles. Major parts of these semi-natural savannas are fenced and used for large-scale cattle and goat farming.

Savannas are focussed research areas for analysing the complexity of biological systems, while comprising complex mechanismas determining the coexistence of trees and grasses. Intensive studies on the biocomplexity of savannas have been conducted by Hanan *et al.* (2006) in the Kruger National Park, South Africa. The concept of 'biocomplexity' is defined as *"properties emerging from the interplay of behavioural, biological, chemical, physical, and social interactions that affect, sustain, or are modified by living organisms, including humans"* (Michener et al., 2001). As shown in Figure 1-1, the savannas of the Kalahari combine the characteristic coexistence of trees and grasses which is secondary formed by human land use (rangeland and game farming).

Understanding these semi-natural dynamic systems requires a high degree of information integration across spatial, temporal, and thematic scales. The generation of reliable vegetation maps is a precondition. For vegetation mapping purposes this increasing synthesis efforts across scales are required in a bottom-up framework to combine the available *in-situ* information, covering species occurrence, habitat description, spatial patterns of plant communities, and temporal land dynamics. The challenge can be seen in the way how this multi-scale information can be harmonised and transferred to operational systems being applicable for ecosystem monitoring and modelling. The research needs address the semantic implementation of multi-scale land-cover information in ecosystem assessment initiatives.

Another important research need addresses the technical implementation of multi-scale land-cover data in an integrated mapping framework. This means that (a) up-scaling processes from *in-situ* to regional scales are challenging in very heterogeneous landscapes, (b) the generation of meaningful reference databases for land-cover classification are problematic in remote and inaccessible drylands, and (c) the use of inter-annual time series represents the spatiotemporal variations of the precipitation patterns that influence the classification result and mapping accuracy.

Based on the research motivations mentioned above, major research questions are:

How can satellite data classification techniques be adapted to semi-arid environments in order to account for the temporal requirements in mapping seasonal semi-arid environments?

How can earth observation information from in-situ to regional scales be harmonised into a standardised framework for the regionalisation of ecosystem diversity assessments?

What are the strengths and weaknesses of the FAO and UNEP Land Cover Classification System (LCCS), if applied in a bottom-up mapping framework in a semi-natural semi-arid environment?

How can the global remote sensing community learn from regional biodiversity and vegetation type assessments and vice versa?

1.2 Research framework and assumptions

Considering the principal research questions above, the research framework in this dissertation addresses three key issues, as displayed in Figure 1-2. Aiming at a contribution for the overarching research goal addressing ***integrated land cover monitoring mechanisms***, the research includes and links earth observation assessments from different spatiotemporal scales and the implementation in remote sensing applications.

Following a bottom-up approach, the first research issue focuses on ***integrated vegetation type mapping***, where local scale measurements on ecotype and biodiversity regionalised within a multi-scale mapping framework. Data mining is conducted on order to collect all available land-cover observation data (from *in-situ* plots to remote sensing imagery) and integrate them into a vegetation type mapping framework. The assumption is that measurements on the phenological state of the vegetation and measurements of the geographic context (environmental attributes such as lithology, soils, and topography) derived from earth observation satellites can significantly improve the regionalisation process and the quality of ecosystem diversity databases. In particular, time series imagery covering the annual and inter-annual phenological state of vegetation and their environmental attributes are considered being useful predictors for the classification of semi-arid vegetation types. Further on, high resolution (Landsat-type) imagery has to act as an intermediate step for up-scaling processes from local observations to data with moderate spatial resolutions. In the context of vegetation type classification using multi-temporal imagery (as they often represent very similar and thus 'weak' predictor variables in semi-arid ecotypes), it is assumed that the use of non-parametric ensemble machine learning algorithms is necessary for a proper representation of high dimensional predictor variables in a classification algorithm.

In connection with scale-bridging environmental assessments the second research focus is set on the ***standardisation of earth observation data*** as a key integrative process in an environmental assessment framework. Therefore, it is assumed that existing comprehensive *in-situ* measurements, covering vegetation occurrence and distribution, and environmental attributes such as soils, topography, are basically underestimated in large-area land cover assessments. The incorporation of the most detailed information level including the basic and generic land cover entities, such as classifiers and modifiers in LCCS (Di Gregorio, 2005) or *"dimensions and or measurements that describe the processes under investigation"* (Comber, 2008), can improve the accuracy of vegetation and land cover maps and make them utilisable for a wider range of users that are related to the geographic region. Further, the application of LCCS-based environmental assessments in the framework of bottom-up case studies will show up the strengths and weaknesses of standards of the set of classifiers and the class boundary definitions as implemented in the current LCCS version. Moreover, detailed class descriptions (from broad structural to detailed floristic and environmental attributes) will help to understand general classification requirements for mapping arid and semi-arid vegetation types using multi-temporal predictors.

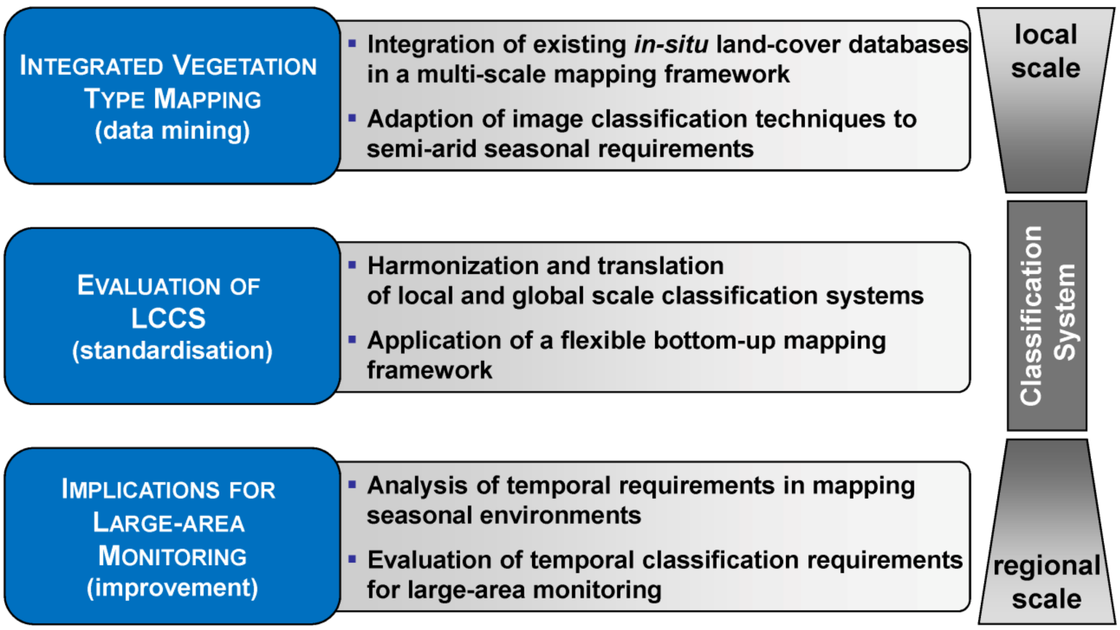


Figure 1-2 Overview of the framework of research.

The transfer of ground-based reference information on a mapping framework based on 'coarse' scale MODIS time series imagery (as used in some global land cover mapping initiatives) leads to several ***implications for improved global monitoring***. Semi-arid land cover classes at global scales are still insufficiently mapped and their mapping accuracies differ strongly between the global land cover databases. It is hypothesised that acquisition time (dry vs. rainy season) and observation period have significant effects on the land cover classification accuracies in semi-arid savannas. Due to the increased spatiotemporal variability of vegetation activity and phenological cycling, ecosystem-adapted feature assemblies of phenological and multispectral time series predictors will significantly increase the performance, accuracy and resilience of semi-arid land cover information. Such kinds of "customised" assemblies of feature sets can be useful towards a reliable spatial representation of land-cover types with a distinct dynamic component. The analyses of temporal requirements in mapping seasonal environments and the evaluations of the technical adoptions will help to identify existing gaps and needs for the technical design of global scale land cover mapping initiatives and satellite sensor design for arid lands.

1.3 Objectives and structure of the thesis

The dissertation is organised in seven main chapters. In the following the chapter one **Introduction** describes the objectives and structure of research based on the mentioned research framework. Section 1.4 includes the description of the climate, geomorphology, hydrology, vegetation, soils, and land use of the Kalahari test site. Chapter two provides a **research review** of the status of land-cover and biodiversity in southern African savannas, classification systems, and methodological aspects of land-cover mapping based on satellite time series. Following the described assumptions and research hypotheses, the research presented in the chapters three to six is structured in the three main parts: ***Data mining***, ***standardisation***, and ***improvement***. These overarching research tasks are organised and presented in form of research articles, which have been published or submitted to peer-reviewed scientific journals and international scientific conferences, as listed in Appendix A.

Chapter three describes the ***data mining*** and focuses on analysing the **suitability of MODIS time series metrics to map vegetation types** in dry savanna ecosystems, applied at the Kalahari test site presented in section 1.4. The two main research objectives are:

i. Application of an integrated concept for vegetation mapping in a dry savanna ecosystem based on local scale *in-situ* botanical survey data with high resolution (Landsat) and coarse scale (MODIS) satellite time series data.

ii. Analysis of the suitability of intensity-related and phenology-related metrics derived from MODIS time series for single annual and long-term inter-annual classifications from 2001 to 2007.

Chapter four is related to the ***standardisation*** part and aims at the **integration of earth observation data by using LCCS** at different spatial and temporal scales for mapping in southern African savannas, covering *in-situ* botanical field plots, local vegetation distribution patterns derived from Landsat imagery, and inter-annual MODIS time series. The research focus is set on the assessment and reporting of the experiences of the applicability of LCCS for land-cover mapping in the Kalahari test site. The research objectives are:

i. To test the applicability of the concept of the LCCS classifiers in semi-arid ecosystems and to demonstrate the benefits and limitations for the scientific communities related to remote sensing and biodiversity.

ii. To apply a flexible legend of typical Kalahari savanna vegetation types using the UN-FAO Land Cover Classification System (LCCS).

iii. To present a bottom-up ecosystem assessment framework of a dry semi-arid ecosystem by integrating local scale *in-situ* botanical survey data with Landsat imagery and coarse scale satellite time series data.

The options for the ***improvement of large-area monitoring of savannas*** in terms of accuracy and data integration are presented in chapter five by analysing the **effects of temporal**

compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems. The implications for the global monitoring of the savanna biome are discussed based on the temporal requirements for different feature assemblies derived from inter-annual MODIS time series. The research objectives are aiming at analysing and evaluating the:

i. Effects of temporal compositing on classification accuracies indicating (inter-)annual temporal requirements and environmental constraints in mapping semi-arid cover types.
ii. Effect of varying observation periods on land-cover classification accuracies considering the inter-annual rainfall variability.
iii. Implications for global monitoring initiatives in order reveal options for improving and adapting the mapping design on the actual ecosystem-specific requirements.

In chapter six the potential of MODIS time series for bottom-up vegetation mapping in the Kalahari test site are presented and discussed. Building upon the findings of the potential and requirements of annual and inter-annual MODIS time series features for the standardised LCCS-based mapping of physiognomic-structural land-cover types and phytosociologic vegetation types, this section discusses the potential of integrating *in-situ* vegetation databases for fuzzy vegetation type mapping. The options for the ***improvement of the regionalisation of plant communities in Namibia*** are discussed with exemplary vegetation types as an example. The main objectives are evaluation of:

i. Effects of temporal compositing on classification accuracies indicating (inter-annual) temporal requirements and environmental constraints in mapping semi-arid cover types.
ii. Effects in the framework of global monitoring initiatives in order reveal options for improving and adapting the mapping design on the actual ecosystem-specific requirements.

Chapter seven reports a summary and **synthesis** discussion of the results and conclusions of chapter three to six. The discussions and concluding remarks will point on the capabilities and pathways for future research tasks in the framework of integrated and interdisciplinary ecosystem monitoring initiatives in southern Africa.

1.4 Description of the study area

The study area is situated in the north-eastern communal area in Namibia, comprising the administrative districts of Otjinene, Omaheke, and Okakara with a spatial coverage of about 57,500 km². The definition of the investigation area is based on the reconnaissance survey of the landscapes, soils, and vegetation, realised in the spring 2004, in the framework of a collaboration of the Desert Margins Program (DMP) and the National Botanical Research Institute of Namibia (NBRI, Strohbach *et al.*, 2004). The field data in this dissertation are based on the results of the DMP project. The next sections provide an overview of the climate, the physical geographical situation, and the land use of the study area.

1.4.1 Climate, geomorphology, and hydrology

Climate - The arid to semi-arid climate in Namibia is influenced by the subtropical high pressure zone causing westerly winds with cold dry air with very low cloud coverage. The average long-term rainfall ranges from 325 mm to 450 mm (Figure 1-3) with a decreasing precipitation rate from north to the south (Strohbach *et al.*, 2004). Due to an increased influence of the Intertropical Convergence Zone (ITC) associated with warm and moist air masses in the summer months, a distinct seasonality of dry winters and wet summers is evident. The rainy season is characterised by local and heavy rainfall rates due to convective rainfall events. A distinct inter-annual spatiotemporal variability of the rainfall distribution is evident leading to an increased severity of droughts (Leser, 1972; Botha, 1999). The temperatures range from 30 – 32 °C during the hottest month (December) and 2 – 6 °C during the coldest month (July). The average number of frost days per year ranges from 5 – 30 days (Mendelsohn *et al.*, 2002; Mendelsohn & Obeid, 2002).

Geomorphology and Hydrology – The study area is located in the western part of the Kalahari basin. This basin is characterised by deep sand deposits (Kalahari sands), partly covered by clay and silt sediments. The topography is mainly flat, partly undulating to rolling. Slopes with gradients between 6° - 9° can be found at so-called Omuramba valleys (incised river beds) or fossil longitudinal dunes. Water surface runoff is non-existent in these permeable sands, except of some occasional incised rivers. High infiltration rates and retention capacity is evident due to the massive sand layer, reaching depths of 5 m at the transition zones to 100 m at the centre of the basin (Thomas & Shaw 1991). The geomorphologic units of the study area as described by Strohbach *et al.* (2004) are shown in Table 1-1.

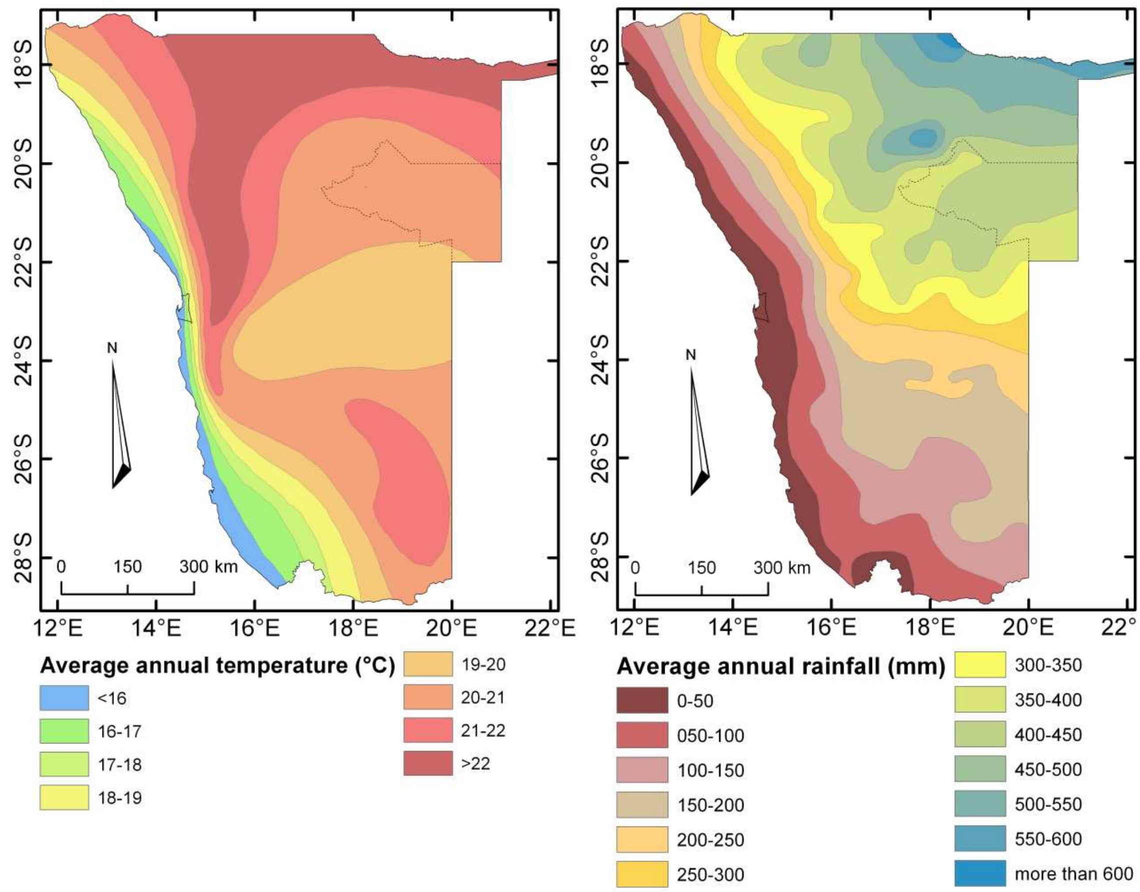


Figure 1-3 Long-term average annual temperature and rainfall in Namibia (Mendelsohn *et al.*, 2002). Note the extent of the study area indicated by the dashed line.

1.4.2 Soils

The semi-arid savannas of the Kalahari are characterised by a low soil moisture and water intake levels combined with often very heavy rainfall events. As a consequence the aeolian and fluvial soils in the study area are generally poorly developed. A soil survey was conducted in the framework the Desert Margins Program (Strohbach *et al.*, 2004). The soil characteristics are described below stratified after the defined geomorphologic units.

The interdunal depressions and pans are dominated by two soil types: petric Calcisols and haplic Arenosols. The petric Calcisols are often overlaying a hard calcrete or petrocalcic horizon at a depth of up to 50 cm. The haplic Arenosols are characterised by a yellow to brown colour, a weak horizon development, and deep sand layer of fine grained to loamy fine sands with a slight clay content.

Table 1-1 Geomorphologic units in the study area.

Geomorphologic unit	***Description***
Interdunal depressions	Sandy interdunal depressions and pans, isolated pans and dune streets
Kalahari sand dunes	Undulating to rolling longitudinal sand dunes; deep to very deep reddish, unconsolidated coarse-grained sedimentary sands
Kalahari sand plains or sand deposit	Flat to almost flat sand plains; white to reddish sandy textured soils; covers most of the study area
Flat calcrete plains with pans	Patchy distribution of calcrete plains; moderately deep soils with a sandy to loamy texture
Drainages or omiramba	Fossil drainage lines formed during the wetter phase predominating the early-to-mid Pleistocene age
Kalahari sand plains with shallow underlying calcrete	Shallow calcrete pans and reworked whitish loose unstructured Kalahari sands; situated at the lower elevations than the Kalahari sand plains
Floodplains and areas subjected to regular flooding	Repeated seasonal flooding due to relatively low relief, patchy distribution of pans and watercourses
Plateau with Karstveld or limestone	Flat to gently undulating topography; hard and soft calcrete and limestone outcrops
Eroded surface of the plateau	Undulating to rolling topography; part of the slopes off the Waterberg Mountains

The soils of the Kalahari sand dunes can be described as loamy sands with very low agricultural potential. At the dune crest and dune foot ferralic Arenosols are found with a red to brown soil colour due to higher contents of sesquioxides. Similar to the haplic Arenosols, the ferralic Arenosols show a weak horizon development. As shown in Figure 1-4 (right) ferralic Arenosols are the dominant soil type of the Kalahari basin.

Different pedogenic processes are apparent in the soils of the sand plains and ridges of the Kalahari. On the sand ridges haphic Arenosols are found, whereas the Kalahari sand plains are characterised by ferralic Arenosols. The soils of the calcrete plains are classified as haphic Calcisols on the flat calcrete plains and petric and haphic Calcisols. Weathered calcrete causes the comparatively dark soil colours and pronounced textures and horizon development. Haphic Calcisols are typically occurring on the slopes of the omiramba.

In the drainage lines or omiramba petric Calcisols, deep skeletic Regosols, shallow lithic Leptosols on sandstone and calcrete, and relatively fertile calcaric and arenic Fluvisols are found. The soils of the Karstveld and ridges are dominated by petric Calcisols and contain a top layer of weathered material and high gravel content (Mendelsohn *et al.*, 2002; Strohbach *et al.*, 2004).

1.4.3 Vegetation

The study area is subdivided into four broad vegetation type units. Following the national vegetation type map after (Giess, 1971), the central and north-eastern part of the study area is covered by *Forest savannah and woodland* of the *Northern Kalahari*, followed by the *Camelthorn savanna* of the *Central Kalahari* in the southern part of the area. Smaller areas of the northern part are covered by the vegetation type of *Mountainous savannah and Karstveld*, whereas the most western area is classified as *Tree and bush savanna* of the *Thornbush savanna* (Figure 1-4, left).

A phytosociological analysis of the field plots during the DMP survey resulted in a classification of plant species associations by discriminating *Characteristic species*, *Differentiating species*, and *Typical species*. The following description of the vegetation distribution in terms of dominant and most common species in the study area was derived from a phytosociological classification of the vegetation survey data. A synoptic table of vegetation types after Strohbach *et al.* (2004) is presented in Appendix B.

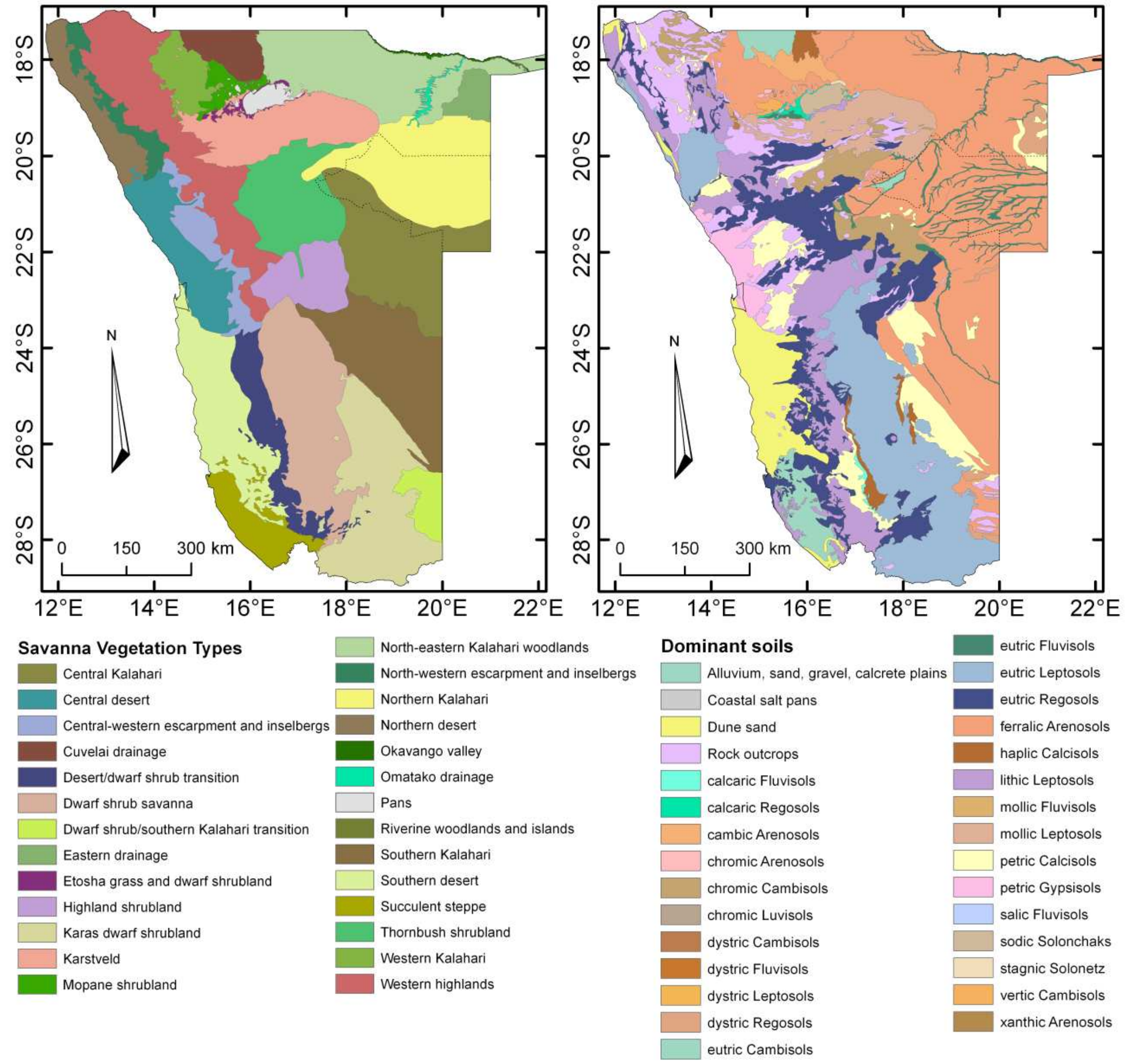


Figure 1-4 Vegetation types after Giess (1971) and the dominant soils in Namibia (Mendelsohn *et al.*, 2002).

Plant associations of broadleaved deciduous shrub- and bushlands, dominated by *Terminalia – Combretum* savannas, are the most widespread vegetation types in the study area and represent the sand plains and deposits of the *Northern Kalahari*. The most widely distributed vegetation type is the association of *Terminalia sericea - Combretum collinum* moderately closed shrub- and bushland. This type is associated to deep sand habitats such as sandy plains and longitudinal dunes. In terms of spatial coverage, the second-largest vegetation type are *Acacietea*, as this vegetation type gradually replaces the *Terminalia – Combretum* savannas in the transition zones from deep to shallow sand habitats (loamy sand to loam) in the western part of the study area.

The *Acaciatea* are characterised by the presence of different *Acacia* species, typically occurring in the *Central Thornbush Savanna* and the *Camelthorn Savanna* after Giess (1971). Beside the transition zones from deep Kalahari sands to the mountainous regions, *Acacia* species are found along drainage channels and small depressions and pans. The most prominent association are described as *Acacia mellifera – Stipagrostis uniplumis* tall to high moderately closed shrublands (Figure 1-5f), occurring on plains, interdunal valleys, and sand drift plains at an altitude of 1330 m (± 88 m), typically on shallow to deep sandy soils, calcretes, and soil crusts. This vegetation type is associated with a high degree of bush encroachment. The sub-association of *Enneapogon desvauxii – Acacia hebeclada* tall, semi-open to moderately closed shrublands (Figure 1-5g), is typically found along the omiramba drainage lines and indicate the occurrence of clay-enriched soils with comparatively high pH values. *Acacia luederitzii – Ptycholobium biflorum* bushlands (Figure 1-5h) are typically found on the former

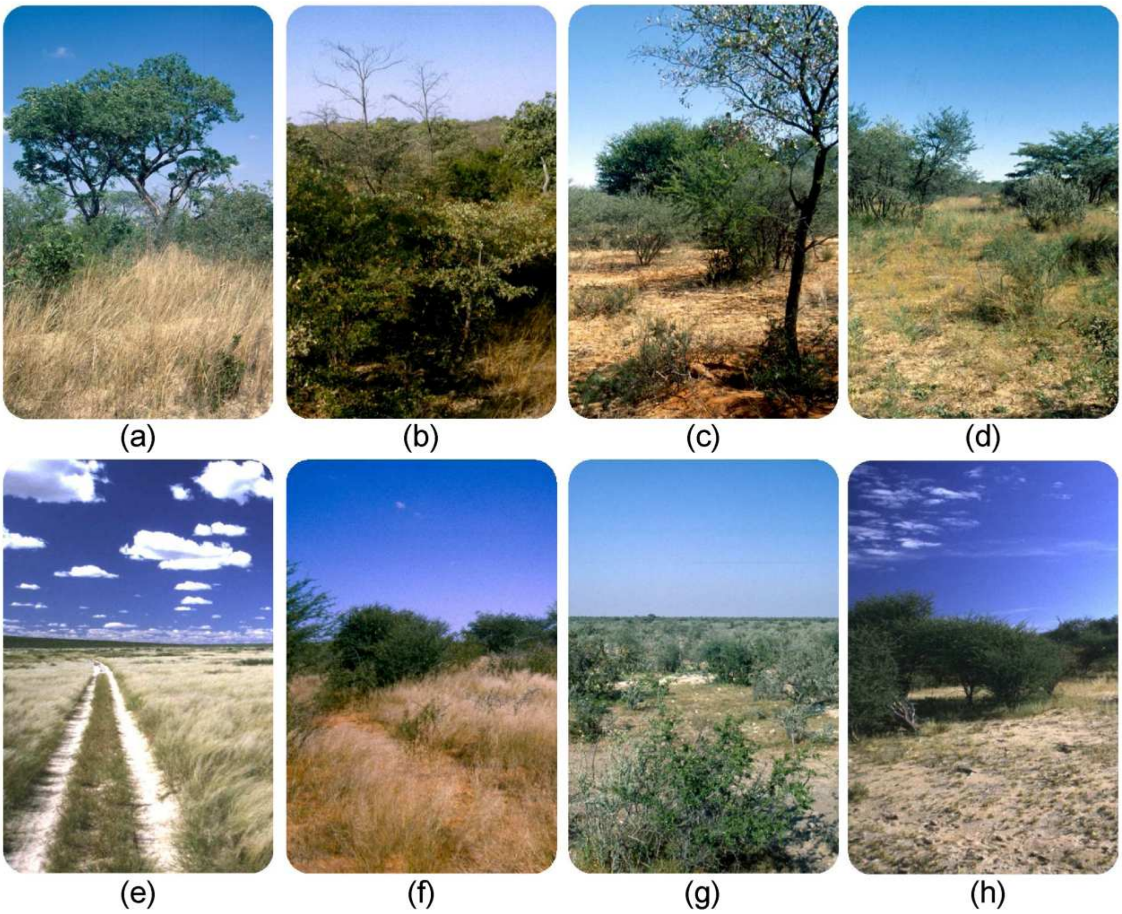


Figure 1-5 Vegetation types of the Kalahari in north-eastern Namibia: (a) *Pterocarpus angolensis* - *Burkea africana* tall, moderately closed bushland, (b) *Terminalia sericea* - *Combretum collinum* moderately closed shrub- and bushland, (c) *Terminalia sericea* - *Acacia mellifera* bushland, (d) The *Terminalia sericea* – *Acacia erioloba* short, moderately closed bushland, (e) *Eragrostis rigidior* – *Urochloa brachyura* tall, semi-open shrubland, (f) *Acacia mellifera* – *Stipagrostis uniplumis* tall to high moderately closed shrubland, (g) *Enneapogon desvauxii* – *Acacia hebeclada* tall, semi-open to moderately closed shrubland, (h) *Acacia luederitzii* - *Ptycholobium biflorum* short, moderately closed bushland. The photos were taken by Ben Strohbach during a field survey in 2004 (Strohbach *et al.*, 2004).

floodplains of the Omatako Omuramba. These floodplains are characterised by shallow and crusted soils.

The *Pterocarpus angolensis* – *Burkea africana* tall, moderately closed bushland association (Teak woodlands, Figure 1-5a) is typically found on the northern sandy plains with sand depths of more than 150 cm and precipitation rates between 400 and 500 mm. This vegetation type is part of the dry forests and tall woodlands in north-eastern Namibia and can reach a tree cover up to 12 %.

1.4.4 Land use

Rangeland farming is the main land use in the semi-natural savannas of Namibia. The carrying capacity and thus the livestock density follow a NE-SW gradient of vegetation productivity. The north-eastern communal areas are at the higher end of the spectrum with a carrying capacity of 30 – 40 kg/hectare. In the north-eastern part of the study area small scale agriculture on communal land is found. The smaller southern and south-western regions are used for livestock farming and tourism on freehold land (Mendelsohn *et al.*, 2002).

In the framework of the Agro-Ecological Zoning (AEZ) Program initiated by the FAO, the study area was classified into eight agro-ecological zones that can be defined as so-called "land entities" with uniform climatic conditions, and geomorphologic and pedologic characteristics (De Pauw, 1996; De Pauw *et al.*, 1999). The dominating AEZ in the study area are Kalahari sand plains, sand dunes, and sand ridges. The AEZ evaluation showed that crop production is not suitable due to short dependable growing periods, low water holding capacity, a and low nutrient status. The full list of the agro-ecological zoning of the study area is described in Appendix C. Maps of land-use and the agro-ecological zones are presented in Appendix D.

2 Research Review

2.1 Status of land-cover and biodiversity modelling and monitoring in southern Africa

The knowledge of the climatic and human impacts on the sub-Saharan environments and the related impact on the current status and trends of biodiversity and major land modifications foster the need for developing operational earth observation mechanisms for the terrestrial biodiversity and ecosystems (Muchoney, 2008). In the following sections an overview of the status of the land-cover and biodiversity monitoring initiatives is given.

2.1.1 Perspectives and challenges for global terrestrial monitoring

Regarding the increasing environmental pressures leading to distinct needs for the tracking and reporting of land change processes, increasing data archives and multilateral earth observation initiatives expended the perspectives for global terrestrial monitoring. The perspectives are strengthened global cross-linkage of monitoring systems and scientific disciplines. This results into increased challenges of the standardisation of methods and databases. The following sections provide a summary of the main global emphases on land-cover and biodiversity monitoring.

2.1.1.1 Biodiversity

As a result of numerous international science meetings on earth observation in the last decades the principle goal of a *"wider use of earth observation technologies"* (Group on Earth Observation, 2005) in the main areas of social benefit has been stated. Recently, the Group on Earth Observations (GEO) and the Global Environmental Observation System of Systems (GEOSS) was established in order to implement earth observation techniques for nine subjects of social relevance (Group on Earth Observation, 2005; Muchoney, 2008). Besides *Disasters*, *Human Health*, *Energy Management*, *Climate*, *Water Cycle*, *Weather*, and *Agriculture*, the social benefit areas of *Protection of Ecosystems* and *Conserving Biodiversity* directly address the long-term implementation of remote sensing techniques for biodiversity and ecosystem monitoring. Comparably important objectives were defined by the Convention on Biological Diversity (CBD) that could not have been realised without satellite-based observations, such as the characterisation and monitoring of *trends in extend of selected biomes, ecosystems and habitats* and *connectivity and fragmentation of ecosystems* (CBD, 2005).

Following the demands of both CBD and GEO, one of the core components of GEOSS was the establishment of the GEO Biodiversity Observation Network (GEO BON). The main task of GEO BON is to provide a facility for the *"combination of top-down measures of ecosystem integrity from satellite observations with a host of bottom-up measures of ecosystem processes, population trends of key organisms"* (Scholes *et al.*, 2008), and to guide data collection, standardisation, and information exchange. The structure of the workflow from measuring biodiversity on species-, community-, and ecosystem scale and the proposed provision of important user-defined products is visualised in Figure 2-1.

Global assessments of the drivers of biodiversity loss, as done in the CBD 'Programs of Work' (CBD, 2006) identified habitat change as the most important driver of change in the last 50 years. In the sub-Saharan sub-tropical forests, grasslands, and savannas, land-use change was identified as major driver of habitat change, with *highly increasing impact* on biodiversity. Other globally observed processes such as climate change, invasive species, and pollution were classified with *very rapid increase of impact.* The global observation of land-cover plays therefore an important role for the establishment of ecosystem-level monitoring on global scales (Pereira & Cooper, 2006).

2.1.1.2 Land cover

The global monitoring of land-cover addresses important issues in all societal benefit areas (e.g. ecosystems, agriculture, and biodiversity), as implemented and emphasised in the GEO process. Land-cover was recently declared as one of the essential climate variables (ECV, Herold 2009a; Herold 2009b; Arndt *et al.*, 2010). However, even with increasing numbers of global terrestrial observation products, recent comparative studies highlighted the existence of major inconsistencies between the land-cover products in mapping savannas at global scales, since they were embedded in different project structures, classification systems, and aims of research (Herold *et al.*, 2008; Jung *et al.*, 2006). Figure 2-2 shows the spatial agreement between four global land-cover datasets. The comparison with the global distribution of tropical and sub-tropical savannas highlights large disagreements of class assignments in these ecosystems at the main transition zones between the biomes.

The coexistence and spatially heterogeneous distribution of different life-form layers of woody and herbaceous vegetation (Frost, 1996; Sankaran *et al.*, 2005; Privette *et al.*, 2004; Scholes *et al.*, 2002) may be a key factor for low mapping accuracies of savanna land-cover classes, as shown in several comparative assessment studies (Giri *et al.*, 2005; Hansen & Reed, 2000). Synergistic approaches of combining the available land-cover data sources in order to detect and reduce class-specific inconsistencies were conducted. An example is given by Jung *et al.*, (2006) where existing land-cover products (GLCC, GLC2000, MODIS) are merged into a common legend (SYNMAP) based on a 'best estimate' fuzzy agreement.

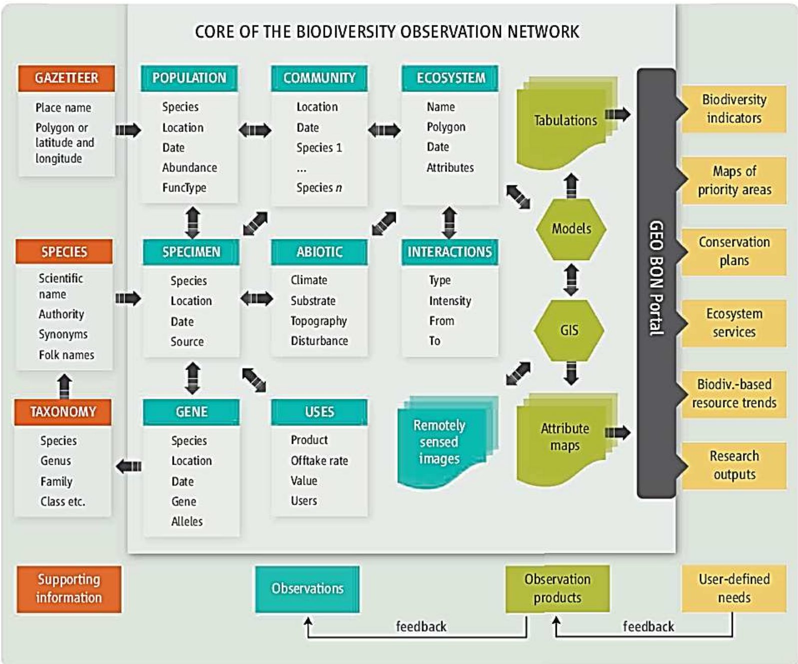


Figure 2-1 Structure of an integrated biodiversity observation system showing observations and derived products on different scales (population, community, and ecosystem) and the different user-defined information products. Note the increased importance of satellite-based observations and GIS and modelling components for the generation of most of the products (e.g. conservation plans, maps of priority areas, and biodiversity indicators). Adapted from Scholes *et al.* (2008).

The work of different globally acting research communities in the framework of the United Nations Framework Convention on Climate Change (UNFCCC), the Integrated Global Observations for Land (IGOL), and the GEO process revealed that there (still) exists a lack of operational global land-cover observations, which emphasised the need for "building technical creditability". One major issue addresses the need for integrated land cover assessments by combining *in-situ* data and satellite observations from high- and moderate resolution data. Tasks were defined to move towards a (more) operational situation and to define standards for global and regional land-cover observations, such as the development of standards for land-cover characterisations, and the development of standardised methods for the accuracy assessment of land-cover datasets (Herold *et al.*, 2008; Herold *et al.*, 2009).

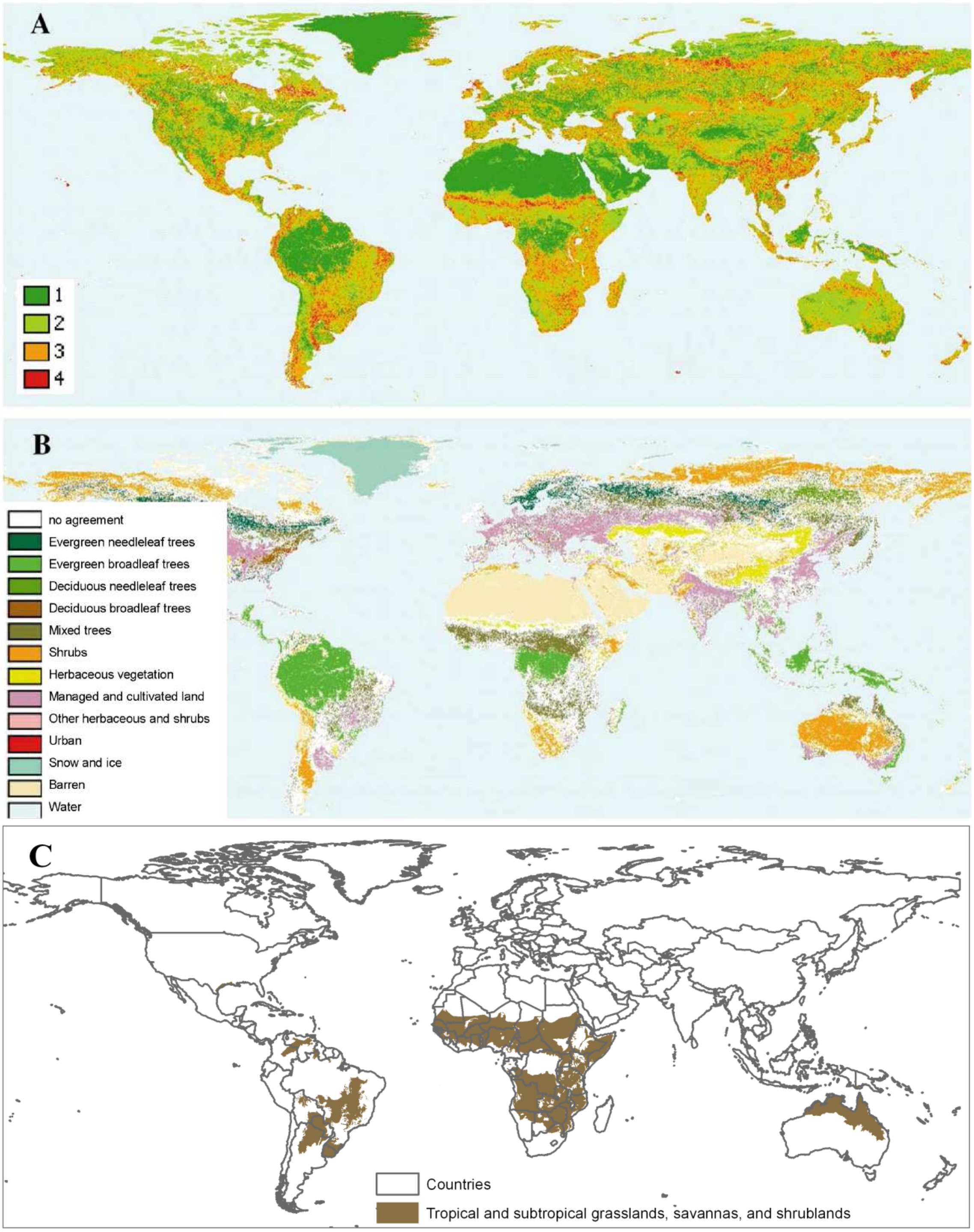


Figure 2-2 Spatial agreement between four global land cover datasets (A) and land cover characteristics for areas where 3 of the 4 global datasets agree (B) given a generalised legend (Herold *et al.*, 2008). (C) shows the spatial distribution of tropical and subtropical grasslands, savannas, and shrublands (Olson *et al.*, 2001).

Results from international harmonisation and standardisation activities postulated (a) a general lack of harmonisation activities, such as *a priori* legend definitions as well as *a posteriori* inter-comparisons of existing legends, (b) a lack of experience for regional harmonisation activities for different ecosystems, and (c) a general lack of integrated ecosystem assessment approaches which is a pre-condition for effective harmonisation results and acceptance for the involved user communities. International panels and user groups provide the technical requirements, data, and communication infrastructure for the definition and generation of 'flexible' geo-information data on land-cover and land surface processes. In the framework of the ISO TC211 process, the standardisation process of geographic information and geomatics of the Technical Committee 211 (TC211) of the International Organisation of Standardisation (ISO), the Land Cover Classification System (LCCS, Di Gregorio 2005) was submitted as an international land-cover standard. However, the experiences of the scientific community in terms of the feasibility of a range of ecosystems (based on regional case studies) are still outstanding. (Townshend *et al.*, 2007; Herold *et al.*, 2006; Herold *et al.*, 2009; Herold *et al.*, 2008).

2.1.2 Savanna system analyses: concepts, processes, and applications

The semi-arid savannas of southern Africa are characterised by a distinct biocomplexity which makes field-based ecosystem assessment and remote sensing projects challenging. Examples are given with the determinants of vegetation distribution (e.g. tree-grass interactions). In the following sections a review of scientific literature on concepts, processes and applications of savanna system analyses is provided.

2.1.2.1 Biocomplexity

In order to understand the coupled nature of human and natural systems, the knowledge of the different interactions and processes in ecology were introduced as a resource for the future conservation of natural ecosystems and biodiversity hotspots (Callicott *et al.*, 2007; Cadenasso *et al.*, 2006; Green *et al.*, 2005; Walters, 2007). The main models are the concepts of *biodiversity* (species richness in a biotic community, Wilson, 1988), *biocomplexity* (the process of discovering *"the complex chemical, biological, and social interactions in our planet's systems"*, Colwell, 1998), and *connectivity* (Colwell, 1999).

A fundamental requirement to move from the metaphorical term of biocomplexity to quantification and model construction is the definition, assessment and integration of the scales of measurement (time and space) and the definition and specification of the components of structural complexity (Figure 2-3). The main fields of research were defined by Michener *et al.* (2001) in order to understand the ecological dynamics of biocomplexity, which comprises a synthesis of information from across temporal, thematic, and spatial scales:

- *"Add knowledge about the environment, ranging from the genetic diversity of microorganisms to global climate change".*
- *"Learn about human influences on natural processes and of natural processes on human behaviour".*
- *"Develop new methods and computational strategies to model and manage complex systems".*
- *"Use biologically or biocomplexity-inspired design strategies to discover new materials, sensors, engineering processes, and other technologies".*

The scale of measurement is in particular important for analysing semi-arid savanna ecosystems, since vegetation patterns are primarily controlled by precipitation and temperature, but also by non-linear processes such as fire, herbivory and human activities (Bucini & Hanan, 2007; Hanan *et al.*, 2006). The determinants of the different system components are insufficiently analysed for ecosystem monitoring and modelling purposes in savannas and requires a highly interdisciplinary analysis network. In the following, an overview of the main monitoring and modelling activities of sub-Saharan savannas is given.

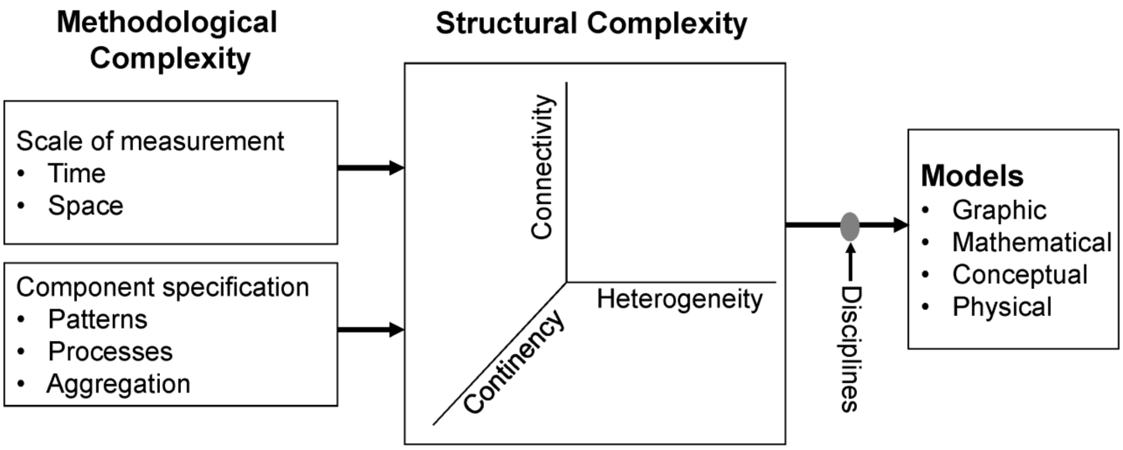


Figure 2-3 Conceptual framework for the organisation of the components of biocomplexity. The complex components of a system can be structured and thus quantified along spatial, temporal and organisational axes, as used by ecological, social, and physical disciplines for model generation (Cadenasso *et al.*, 2006, modified).

2.1.2.2 Tree-grass interactions

Savannas are one of the major terrestrial biomes, cover an area of about 33 million km², and contain one fifth of the world's population providing basic ecological services such as rangelands and livestock (Beerling & Osborne, 2006). They are characterised by a coexistence of woody and herbaceous life forms with a continuous transition of contrasting life-form compositions (ranging from large grasslands, continuous grasslands with scattered trees, to closed woodlands) following a climatic gradient with distinct rainy and dry seasons. The key determinants of mostly non-linear processes (such as climate, topography, soils, geomorphology, fire, herbivory, land use) and interactions among them, defining the spatial distribution of tree-grass mosaics, are not well explored (Archer, 1995; Scholes & Archer, 1997; Jeltsch *et al.*, 1998; Jeltsch *et al.*, 2000; Richter & Rethman, 1999; Ludwig, *et al.*, 2001). The contrasting life form compositions are split into different photosynthetic pathways, where savanna grasses increase their photosynthetic activity in hot conditions (C_4), and trees follow the C_3-type, which is adapted to a temperate climate (Ehleringer & Bjorkman, 1977). Studies of interactions between woody life forms and grasses were conducted in order to analyse the effect of woody plants on grasses and *vice versa*. The main factors influencing these interactions were defined as (a) physiological characteristics (canopy structure and rooting system), (b) photosynthetic pathway and phenology (evergreen, deciduous), (c) inter-annual variability of resources (light, water, and nutrients), (d) grazing intensity, (e) and the intensity of fire disturbances. These processes are analysed using three types of models: niche separation models, balanced competition models (allow coexistence), and disequilibrium models (assume no stable equilibrium). Keeping the possible contributions of satellite earth observation in mind, a main influence is assigned to niche separation by phenology, since the alternation of warm dry and hot wet conditions determine seasonal growth patterns and thus opportunities for niche building (Scholes & Archer, 1997; Eamus *et al.*, 1999).

2.1.2.3 Determinants of vegetation distribution

Rainfall **-** The determinants of the coverage of woody and herbaceous life forms were analysed on continental scale based with the utilisation of coarse scale satellite observations. Synergistic analyses of station data covering mean annual precipitation (MAP), temperature, and soil characteristics coupled with remotely sensed burnt area maps based on eight years of AVHRR observations showed that African woody cover is mainly a function of MAP in arid and semi-arid savannas. Based on a piece-wise linear regression, MAP of 650 ± 134 mm was estimated at which the maximum tree cover is attained. A minimum of 101 mm MAP is required for the occurrence of trees (Sankaran *et al.*, 2005). Similar results were obtained by Bucini & Hanan (2007) who compared the MODIS 500 m tree cover product (VCF, Hansen *et al.*, 2003) with MAP. Compared to the quintile regression line representing the upper boundary of woody cover as presented by Sankaran *et al.* (2005), the comparison of MODIS with MAP achieved the best functional responses with a sigmoid relationship. Based on the sigmoid response to MAP, process types of savannas were classified: Arid savannas (MAP < 400 mm) are less affected by perturbations and have thus little impact on tree cover, whereas MAP is a major controlling factor in semi-arid and mesic savannas (MAP 400 - 1600 mm).

Fire – MAP determines the upper limit of the occurrence of woody cover, whereas fire disturbances control the distribution of woody life forms below the MAP-determined boundary (Sankaran *et al.*, 2005). Fire is the main process linking climate feedbacks (droughts, regional precipitation, and lightning) and vegetation feedbacks (tree mortality, C_4 grass cover, and forest cover) that determinates and balances the tree-grass mixture (Beerling & Osborne, 2006). Fire induced disturbance is a key ecological buffering mechanism since woody plant densities are reduced and maintained at lower levels (Jeltsch *et al.*, 1998; Higgins *et al.*, 2000; Beckage *et al.*, 2009). Recent studies on fire regime modelling indicated that the intermittent and stochastic nature of fire occurrences is one of the main drivers for tree-grass coexistence. Decreased fire frequencies can lead to bush encroachment, whereas increasing fire rated can lead to grassland conversions (Trollope, 1998; D'Odorico *et al.*, 2006; D'Odorico *et al.*, 2007). Applied studies on the historic anthropogenic use of fires support these findings, since controlled fires were used to reduce large-area bush encroachment in north-eastern Namibia and South Africa (Trollope, 1998; Sheuyange *et al.*, 2005). Satellite-based applications proved to be effective for the long-term characterisation of the fire regime in savannas. In this context, Verlinden & Laamanen (2006) analysed the relationship between fire frequency, rainfall, land-cover, and land management. A number of studies applied satellite time series analyses in various studies in Africa to assess large-scale fire patterns (Justice *et al.*, 1998; Dougill & Trodd, 1999; Justice *et al.*, 2002; Roy *et al.*, 2002; Giglio, 2003; Laris, 2005; Sheuyange *et al.*, 2005).

Geology and soil properties – Sankaran *et al.* (2005) and Bucini & Hanan (2007) found that soil properties and the underlying geological setting are important in terms of the distribution of the tree-grass matrix, as proved by slight positive correlations between woody cover and percentage of clay and nitrogen mineralisation potential. The local soil type occurrence is a controlling factor triggering the vegetation phenology (Scholes, 1990). Studies of woody deciduous Kalahari sand vegetation and in the central lowlands of South Africa showed that soil moisture is a key parameter for the phenological patterns (Childes, 1989; Shackleton, 1999; Shackleton, 2002; Dougill & Thomas, 2004). Soil moisture was found to be the “synthesizing” variable of climate, soil, and vegetation interactions, and is jet poorly understood in semi-arid ecosystems. Analyses along a north-south rainfall gradient in the Kalahari showed that the soil moisture contrast between canopy and inter-canopy increases with increasing aridity, leading to positive feedbacks in terms of state and transitions, even when small changes of one of the environmental variables were observed (D'Odorico *et al.*, 2007).

Land use – The (over-)use of the natural resources in water-limited ecosystems leads to degradation processes on large scales. In Namibia, overstocking (cattle, sheep, and goats) was identified as the main driver for land degradation, often initiated by traditional environmental perceptions, as outlined in a number of integrated studies in Namibia and South Africa (Sullivan, 1999; Thomas & Twyman, 2004; Reed, Dougill & Taylor, 2007; Allsopp *et al.*, 2007). Considering the temporal behaviour of the determinants of vegetation distribution, it has to be noted that rainfall and fire can be observed at short to periodic (rainfall) and medium to episodic (fire) timescales. The impact of land use affected influences on the vegetation cover can mostly be detected at longer time periods, e.g. by the characterisation of zones of herbivore intensity or monitoring of changes in vegetation structure. Typically, intensive grazing leads to decreased cover of herbaceous vegetation. Along with the reduction of the grass layer, bush encroachment can widely be observed with an exceeding of the natural ‘carrying capacity’, as observed in semi-natural rangelands but also in national parks (Trodd & Dougill, 1998; Lambin *et al.*, 2001; Augustine, 2003; Asner *et al.*, 2009).

In central Namibia, the main bush encroachment species is *Acacia mellifera* leading to large-area bush thickening. Sequent above-average rainy seasons can stimulate bush encroachment, whereas increased fire activity counteracts the succession of woody vegetation (Joubert *et al.*, 2008). Beside the application of state-and-transition models the use of multi-scale satellite data proved to be effective for the spatiotemporal characterisation of bush encroachment, often based on Landsat imagery (Dougill & Trodd, 1999; Strohbach & Petersen, 2007; Verlinden & Laamanen, 2001). The application of fuzzy classification approaches by combining vegetation indices and constrained ordination on hyperspectral imagery proved to be useful for mapping savanna vegetation types. The combination of seasonal information (dry and rainy season image) helped to detect species with bush encroaching characteristics (Oldeland *et al.*, 2010a; Oldeland *et al.*, 2010b). Long-term AVHRR NDVI time series were used to assess the effects and magnitude of human-induced land degradation. However, the quantification of degradation rates and the distinction between natural variations and anthropogenic induced land degradation is challenging, since significant trends in the vegetation activity can often result from both, environmental cues and human impact (Wessels *et al.*, 2004; Wessels *et al.*, 2007).

2.1.2.4 Monitoring of vegetation phenology and dynamics using satellite time series

The previous section described the main determinants of the vegetation distribution in the savanna biome. All of these factors have effects on the phenological patterns of different life forms or plant communities. This section gives a broad overview of the recent advances in monitoring arid and semi-arid vegetation phenology and vegetation dynamics in southern Africa.

Phenology - The monitoring of plant phenology is a key methodology for the quantification of global climate change impacts. Shifts in plant phenology, in particular spring green-up, were observed at species level, based on satellite remote sensing, and based on the monitoring of carbon dioxide (CO_2) concentrations (Cleland *et al.*, 2007; Bertin, 2008). Remote sensing enables the observation of phenology at ecosystem- and global scales and is a widely used tool to assess ecosystem-specific responds to climatic change. It comprises a number of different methodological approaches, such as global and local NDVI thresholds or conceptual-mathematical amplitude divergences. However, currently no generally accepted definition for the satellite derived phenological metrics exists (White *et al.*, 2009). The derivation of phenological time-related metrics such as start-of-season (SOS), end-of-season (EOS), and length-of-season (LOS) was analysed in a number of studies under different sensor settings, such as MODIS (Zhang *et al.*, 2003; Fisher & Mustard, 2007; Funk & Budde, 2009; Zhang *et al.*, 2009; Hird & McDermid, 2009) or AVHRR and SPOT-VGT time series (Tateishi & Ebata, 2004; Delbart *et al.*, 2006).

Measurements of vegetation phenology in rainfall-driven dry tropical and arid ecosystems were used to assess effects of precipitation and soil water potential on specific deciduous vegetation types (Jolly & Running, 2004). Inter-annual observations based on MODIS time series in combination with metrological and soil information were used to assess environmental cues that are responsible for the green-up of grasses and trees. One of the main results was that trees have a lower variability in their inter-annual phenological pattern than grasses (Archibald & Scholes, 2007). Comparable patterns of rainfall-phenology linkages were achieved at continental scales using MODIS and Tropical Rainfall Measurement Mission (TRMM) data. Sub-Saharan phenology shifts in a north-south direction, following mainly the onset of the rainy season. The atmospheric controls on the phenology were found to be much more complex since characteristic phenological pattern vary between ecosystems (Zhang, 2005; Zhang *et al.*, 2009). A lagged relationship between annual NDVI and annual rainfall was observed in southern African Miombo woodlands. This was explained by a one-year lag-impact of precipitation and vegetative activity (Camberlin *et al.*, 2007). Further research is needed to clarify the possibly non-linear response of photosynthetic activity to previous rainfall years, as described by Martiny *et al.* (2005), since the determinants of the seasonal phenological response has to be understood before inter-annual relationships can be clarified.

Vegetation dynamics – Large-scale monitoring of vegetation dynamics has been conducted in the framework of the BIOTA Southern Africa project (Keil *et al.*, 2010). Time series of the enhanced vegetation index (EVI, derived from the MODIS product MOD13Q1) have been analysed to map patterns of temporal variations in vegetation productivity indicating the different rainfall regimes of summer-rainfall in the northern part of Namibia and winter-rainfall in the Western Cape region (Figure 2-4). Anomalies of vegetation activity from the long-term mean are shown in Figure 2-5. The maps indicate the NE-SW gradient of vegetation productivity and spatial patterns of annual variations in vegetation activity as a function of spatiotemporal rainfall patterns in the semi-arid savannas and transition zone towards the desert.

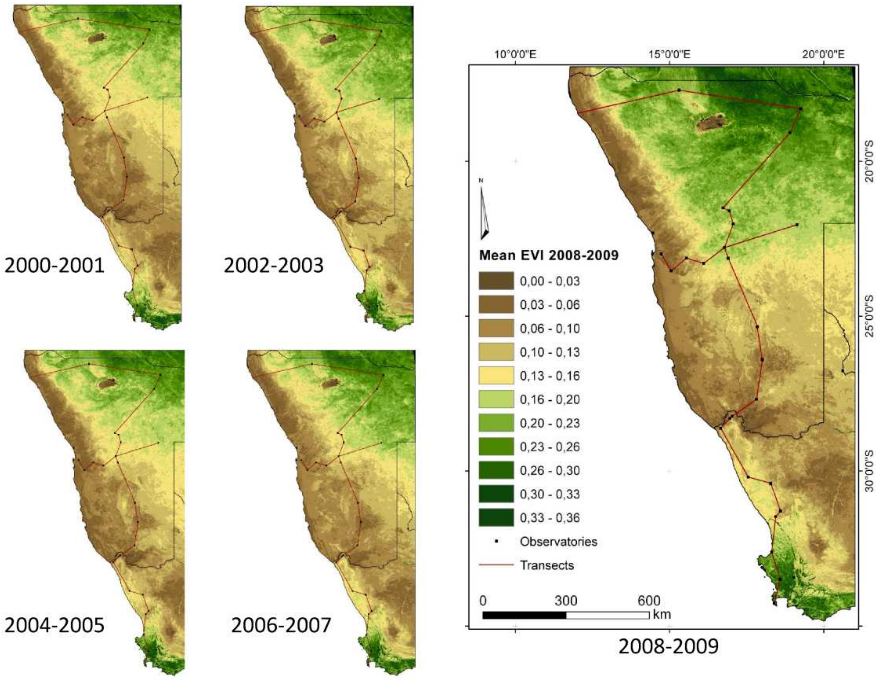


Figure 2-4 Annual mean EVI showing vegetation productivity for the growing seasons of 2000/01, 2002/03, 2004/05, 2006/07, and 2008/09 over the "BIOTA Southern Africa" research area (Keil *et al.*, 2010, modified).

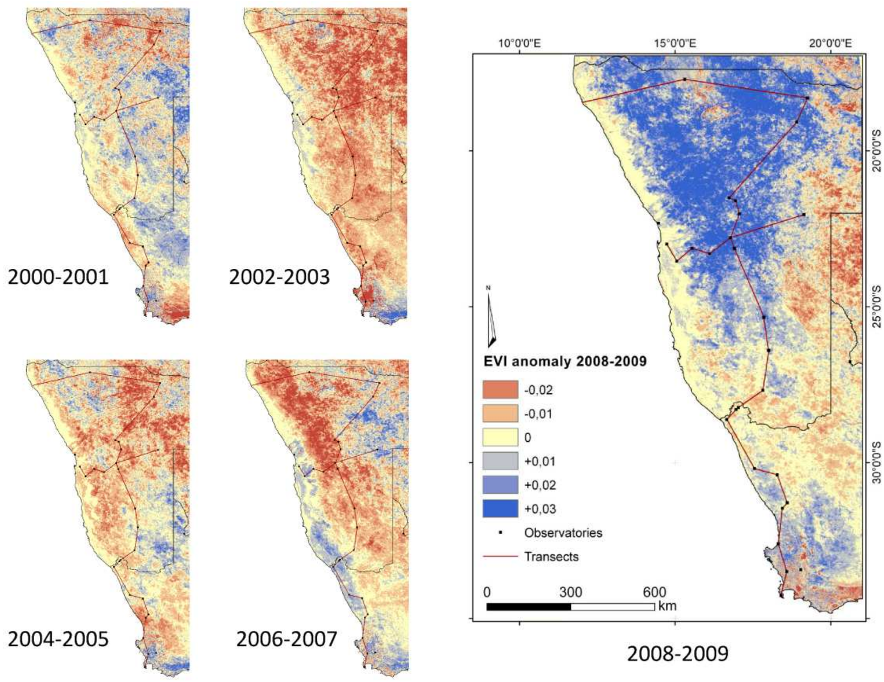


Figure 2-5 Annual variation of vegetation activity from the long-term mean (2000 to 2009) for the growing seasons of 2000/01, 2002/03, 2004/05, 2006/07, and 2008/09 over the "BIOTA Southern Africa" research area (Keil *et al.*, 2010, modified).

2.2 Classification of vegetation data: concepts and methods

Methods for the thematic classification of vegetation data are determined by the scale of analysis, and thus, the thematic content of the data source. Different classification systems exist in the scientific disciplines of botany, biogeography, and remote sensing. Regarding the integrated use of vegetation classification systems and databases in this dissertation, an overview and evaluation of common classification methods is provided in the following sections.

2.2.1 Scale issues

The availability of satellite imagery and derived products has increased in recent years with the consequence that environmental datasets, such as satellite-based vegetation maps or local plot information, have to be harmonised regarding their thematic content. Recent studies emphasised that the most classification uncertainties (e.g. overlapping class definition boundaries and mixed pixel problems) were found in highly heterogeneous landscapes (Thompson, 1996; Jung *et al.*, 2006). The integrated use of multi-scale land-cover information is a precondition for environmental assessments in spatiotemporally dynamic landscapes. The technical and semantic integration of land-cover features comprises the use of data from all available spatial, temporal and thematic scales.

An overview of the linkage of spatial, temporal, and thematic resolution of different land-cover data is provided in Figure 2-6, visualising the information content of *in-situ* data, Landsat-, and MODIS data in terms of their spatial and thematic detail versus the effort for frequent update.

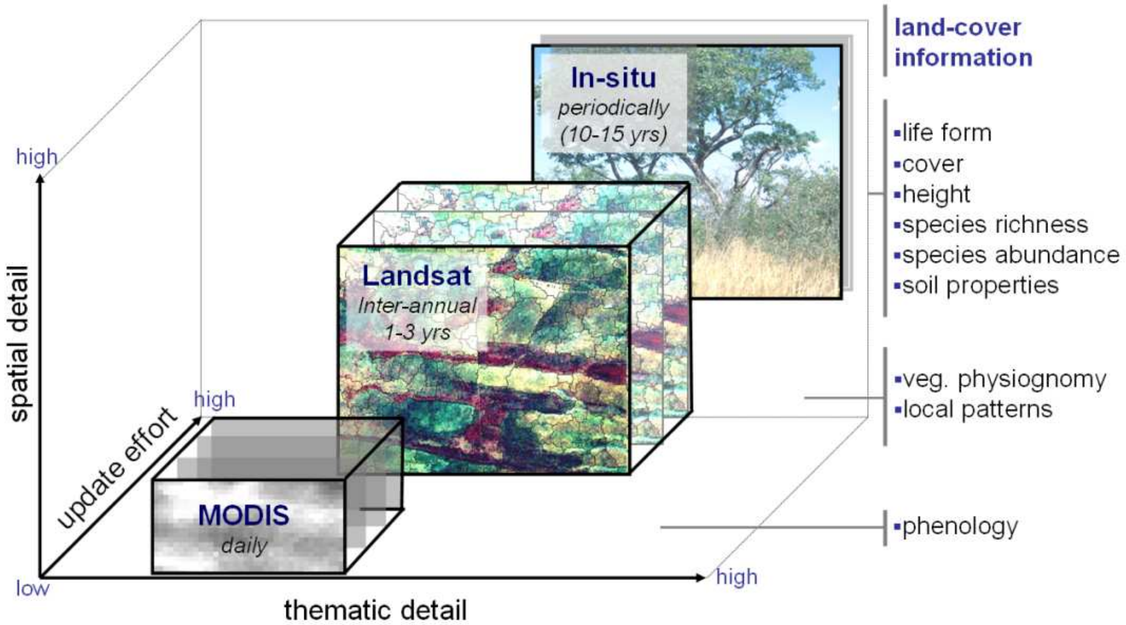


Figure 2-6 Relationship between spatial vs. thematic detail and update effort in terms of land-cover information for three types of earth observation data (*in-situ*, Landsat, and MODIS) in an integrated ecosystem assessment framework (Herold *et al.*, 2008, modified).

The *in-situ* level provides the highest spatial and thematic detail (e.g. life form, cover, height, species, and soil characteristics). The high resolution Landsat-type imagery can act as an intermediate data level since it provides information on physiognomic-structural land-cover information and local patches of herbaceous and woody vegetation components. The advantage of MODIS-type satellite time series is the high repetition rate and a daily global coverage. MODIS emerged to one of the most important satellite time series sensors for the assessment of vegetation phenology (Zhang *et* *al.*, 2001; Zhang *et al.*, 2006). However, the disadvantage of such coarse satellite time series in heterogeneous environments is the lower spatial resolution (250 m – 1,000 m).

The integration of land-cover observations from the *in-situ* level to medium and coarse satellite observations in a comparable and standardised manner is challenging. Despite of data availability, one of the main reasons for the still existing lack of coupled systems of *in-situ* and remotely sensed data is the multidimensional nature of land cover information in terms of their geographical and thematic information content. Thereby, the spatial and thematic dimensions are closely linked. An example is given in the following by a comparison of the concepts of 'habitat' and 'land-cover'.

The scientific meaning of habitat can be defined after Hall *et al.* (1997) as *"resources and conditions present in an area that produce occupancy, including survival and reproduction, by a given organism, and, as such, imply more than vegetation and vegetation structure. A habitat is the sum of the specific resources that are needed by an organism".* The term land-cover comprises the observed bio-physical description of the earth's surface (Di Gregorio & Jansen, 2000). It has to be considered that the description of a habitat is a species-specific concept, while land-cover, from the remote sensing perspective, is limited to the vegetation structure component, often used as an overarching level in remotely sensed map product. The distribution of habitats can directly be derived from land-cover, or indirectly by integrating auxiliary environmental data. On the other hand, land-cover products can be improved by integrating fine scale information of the on-site habitat descriptions. Often, the synergistic use of both concepts is hampered by the lack of clearly defined information needs of the differing user groups (e.g. remote sensing specialists and resource managers, McDermid *et al.*, 2005).

The observation of multi-scale vegetation structure requires a clear distinction and categorisation of the observed phenomena. Many classification systems have a nested hierarchical structure. A hierarchical model describing patterns of matter and energy flux was introduced by Woodcock & Harward (1992) for a forest application. A distinction is drawn between trees/gaps, stands, forest types, and scenes. The stand (or patch) is referred to a contiguous area of a distinct species composition, vegetation cover, and plant distribution. Trees/gaps are parts of stands. A similar categorisation of image objects concerning their resolution was introduced by Strahler *et al.* (1986) with the 'scene' model. The scene model describes the relationship between the spatial resolution of the image and the target object size. It distinguishes between H-resolution (pixel size smaller than the object) and L-resolution (pixels larger than the object).

The synergistic use of all (available) data sources is a two-sided benefit for both local and ecosystem-scale observations. This is particularly important for the monitoring of semi-arid landscapes due to constraints caused by limited infrastructure, accessibility, and observation networks. In the case of the Kalahari this means that coupled systems of *in-situ*- and remote sensing data can improve the local habitat descriptions by involving the temporal component. Currently, the annual and inter-annual vegetation dynamics and phenological cycles of plant communities are still not well understood. Area-wide phenological observatories and *in-situ* observations are still missing. On the other hand, detailed descriptions of the local soil and vegetation structure patterns are required to understand the remote sensing variables used for ecosystem-scale studies, regarding feature selection and reduction for classification, derivation of phenological metrics, or monitoring and assessing degradation states. The inclusion of high resolution imagery in combination with image segmentation techniques has proved to become an intermediate step for such up- and down-scaling processes (Burnett, 2003).

2.2.2 Vegetation classification systems

As described above, the use of the classification system and class nomenclature is depending on the scale of observation (species vs. ecosystems). Classification systems at the species level are 'subjective' in terms of the individual ecological requirements of the observed biocenosis, whereas ecosystem- and global classification systems are distinguished by the underlying class definitions. An overview of methods that are relevant for this dissertation from the scientific disciplines of plant sociology to global land-cover assessments is given in the following.

2.2.2.1 Phytosociological analyses

The assessment and description of biocenoses is one of the key issues in plant sociology. Beside the quantification of biodiversity and gradient analyses, phytosociology aims at quantifying species composition patterns among communities (Ewald, 2003). One major field of application is the regionalisation of biogeographical boundaries between communities, such as zones of rapid change (ecotones), gradual vegetation gradients (ecoclines), mosaics, and transition zones, as visualised in Figure 2-7.

Ordination and classification techniques are the main methods for the description of plant communities. Ordination techniques enable the grouping of species communities according to their similarity, often combined with correlation analyses with abiotic factors (e.g. soils, slope and elevation) to develop verifiable hypotheses of the determinants of species communities. Classification is defined as the grouping of biocenosis with similar composition, resulting in discrete classes (Begon *et al.*, 1998).

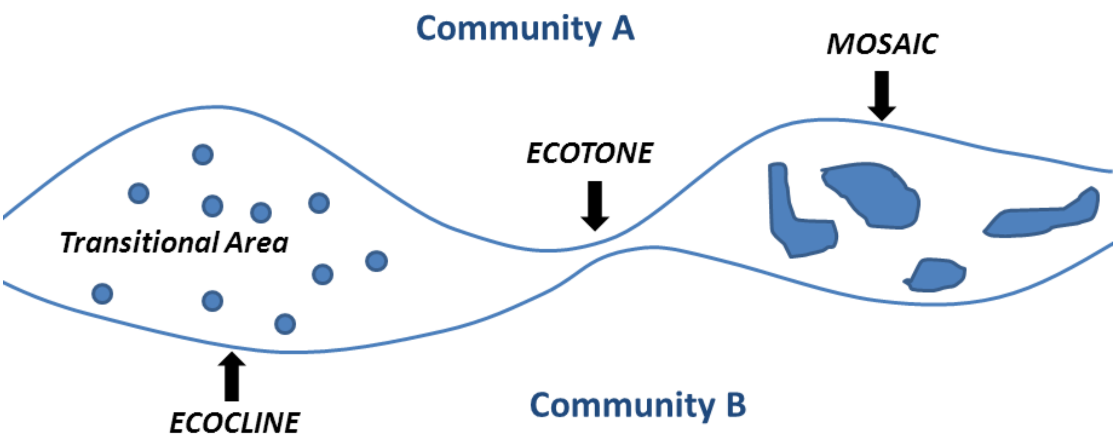


Figure 2-7 The varying extends of ecotones and ecoclines, and mosaics are demonstrated with linear representation of two species communities (Kent *et al.*, 1997, modified).

Widely applied cluster methods in the field of numerical classification of species data are the minimum variance method (*Ward's*, Ward, 1963) and the two way indicator species analysis (TWINSPAN, Hill, 1979). TWINSPAN has widely been used for the classification of species data. Since the result of TWINSPAN is a two-way ordered table, a one-dimensional structure of often more complex species data can be represented, which has been critically reviewed and discussed by Kent (2006). Ordination techniques are widely used methods for gradient analyses. For analysing environmental gradients with a high species turnover, Detrended Correspondence Analysis (DCA, Hill, 1979b) is recommended. An example of the application of DCA is shown in Figure 2-8 for the ordination of two common shrub families of the *Terminalia* and *Acaciatea* in the Kalahari. The presence of an ecocline between those two plant communities is indicated by overlapping plot positions.

The most applied methods for the basic vegetation survey are the transect method and the Braun-Blanquet method. The Braun-Blanquet approach is used to systematically describe plant communities. A precondition for the plot description is the presence of homogeneous and "representative" vegetation stands. After the definition of homogeneous sites, the percentage cover of each species will be estimated using a seven-level scale under consideration of abundance-dominance criteria (Dierssen, 1990). Since the selection of transitional areas and mosaics is avoided using the Braun-Blanquet system, the ecological importance of transition zones tend to be underrepresented in classical phytosociological assessments (Kent *et al.*, 1997).

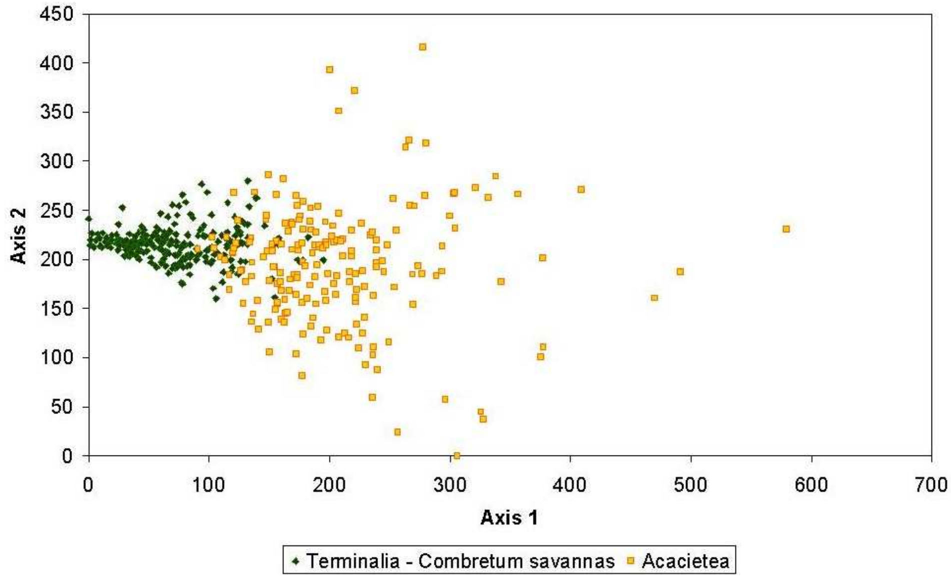


Figure 2-8 Scatterplot of the ordination analysis result based on Detrended Correspondence Analyses (DCA). The ordination analysis was performed on all relevés (N=422) collected in the framework of the DMP project. Gradual transitions between the dominating vegetation types in the Kalahari (*Terminalia* and *Acaciatea*) are indicated by overlapping relevé positions (Strohbach *et al.*, 2004).

2.2.2.2 Land cover classification schemes in southern African savannas

The way how names are given to vegetation classes may represent different vegetation type specific characteristics, e.g. physiognomic-structural (cover, height), dominant growths forms or floristic composition, habitat characteristics, aspects of seasonality (e.g. seasonal change of the ground layer), or combinations of them. Due to often fuzzy and gradual nature of semi-arid ecosystems, special attention is given to the naming of savanna classes. The term 'savanna' can particularly lead to confusion since it is used in different manners. As summarised in Eiten (1992), different studies refer to the term savanna as:

- *"a general structural category applicable to any vegetation from the poles to the equator, of scattered trees and/or shrubs, and other large persistent plants like palms, over a more or less continuous and permanent (live or dry, when not burned) ground layer".*
- *"a category where the presence of a dominating herbaceous ground layer (or forbs, or dwarf shrubs) is required".*
- *"a vegetation type class (or a related series of narrower vegetation types) in which, besides physiognomy, also environment and floristic composition determine the definition".*

Some of the mentioned definition approaches of savanna types are not applicable for integrating them in remote sensing initiatives. For example, floristic aspects and characteristics of the sub-canopy layer cannot be measured by coarse optical satellite sensors. As discussed by Thompson (1996), physiognomic-structural criteria are the most important aspects for a remote sensing based separation of semi-arid landscape units. The schematic representation of the structural components life form, height, and cover of the standard classification scheme after Thompson (1996) is shown in Figure 2-9.

A regional community-based classification scheme including the agricultural potential (so-called Veld Types) was developed by Acocks (1988) for southern Africa. A second widely used physiognomic structural and hierarchical classification scheme has been developed by Edwards (1983). The advantage of Edwards' structural scheme is its independence of geographical location, since no local site descriptions and habitat names are used. Further classification schemes comprise the LCCS-based legend used for the AFRICOVER Project (Mayaux *et al.*, 2004; Mayaux *et al.*, 2003) and the legend used for the Zimbabwean National Woody Cover Mapping Project (Thompson, 1996).

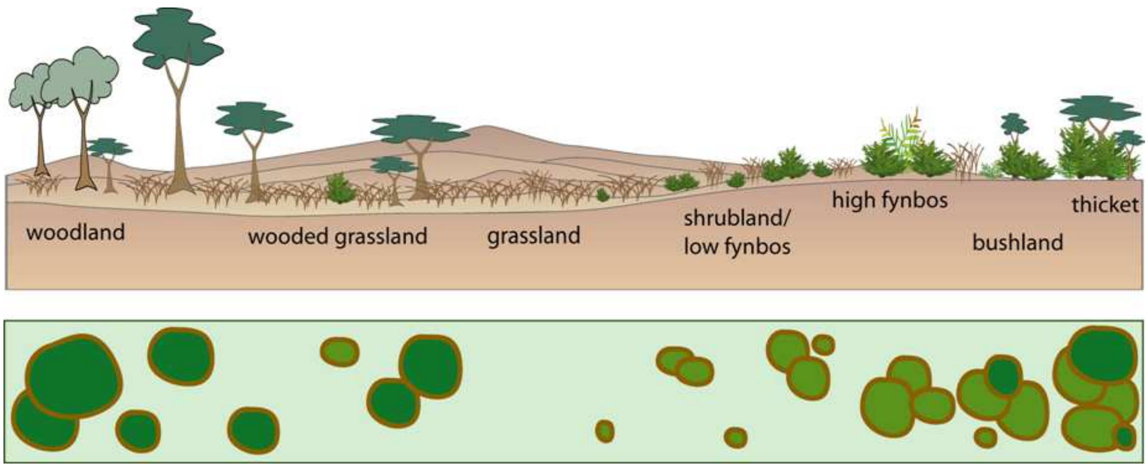


Figure 2-9 Schematic representation of the physiognomic-structural components of land-cover types in Southern African semi-arid landscapes after Thompson (1996, modified). The upper illustration shows the vertical layering system of the vegetation structure. The lower image visualises the top view of different life forms, as measured by satellite data.

Comparisons of the class-specific definition boundaries of tree cover for these classification schemes indicate the highest deviations in the class definitions related to grasslands. Further, the different nomenclatures make a comparison of the developed classification products difficult, as shown in Figure 2-10. The (hypothetical) upper limit of percentage tree cover is used for drawing the class boundaries as defined in the different classification systems. Note that the hypothetical value of 100 % tree cover is not given in natural forests.

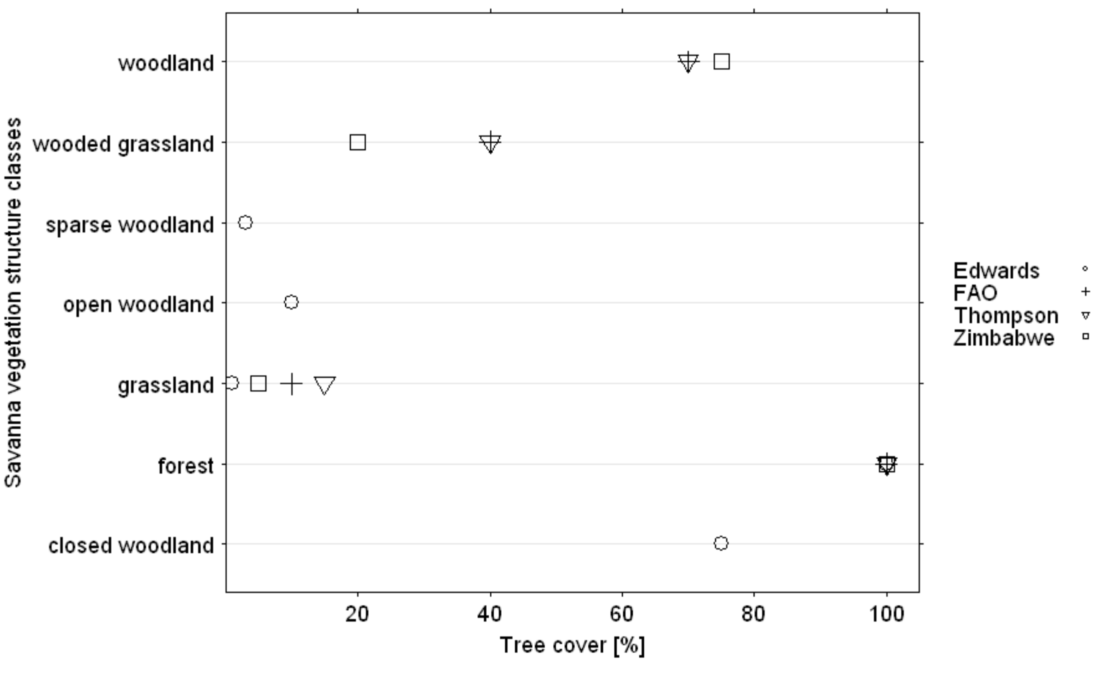


Figure 2-10 Comparison of four savanna vegetation classification schemes showing the class-specific upper boundary definition of percentage tree cover after Thompson (1996, modified).

2.2.2.3 Land Cover Classification System (LCCS)

Traditional classification systems were developed based on broad biogeographic units as used for global mapping projects and environmental studies, as developed for the International Geosphere-Biosphere Programme Data and Information systems initiative (IGBP-DIS, Loveland *et al.* 2000). The global scale IGBP-DISCover classification system includes 17 land-cover classes describing basic land-cover and broad biogeographic units (e.g. *Urban and Built-Up*, *Cropland/Other Vegetation Mosaic*, and *Open Shrublands*, and *Savannas*). Regional to continental scale classification schemes often include a mixture of physiognomic or biogeographic characteristics, and land-use specific descriptors (Eiten, 1968; Anderson *et al.*, 1976). The major points of criticism of these classification systems are that they often describe a mosaic of different land surface types within one class (e.g. *Cropland/Vegetation Mosaic*), and they include fuzzy class definition criteria, such as *Open Shrubland* and *Savannas*. Most of these classification systems include a mix of 'land-cover' and 'land-use' terminology, as discussed by Jansen & Gregorio (2002) and Ahlqvist (2007).

Discussions on the growing needs of comparability and compatibility of spatial land-cover information in the early 1990's led to the development of standardised map legends, applied in a number of mapping initiatives, such as CORINE Land Cover (Bossard *et al.*, 2000) or GLC2000 (Bartholomé & Belward, 2005).

As discussed in Comber *et al.* (2005) the inconsistencies of semantic definitions of land-cover nomenclatures (e.g. land-cover / land-use, or mosaic classes) with bio-physical properties of certain land-cover types lead to problems when applied on satellite based mapping initiatives. They argue that the semantic harmonisation will be unresolved as long as the class attributes are delivered in simple text descriptions.

The development of the Land Cover Classification System plays a key role in the evolving standardisation process of land-cover databases since it addresses all kinds of classification approaches independent of geographic scale and data source. LCCS is an *a priori* classification system which means that the classes represent 'abstract conceptualisations' of real world features. Thus, LCCS uses a set of hierarchical arranged pre-defined criteria (classifiers). As described in Di Gregorio (2005), the eight major land-cover categories in the hierarchical dichotomous phase are:

- Cultivated and Managed Terrestrial Areas
- Cultivated Aquatic or Regularly Flooded Areas
- (Semi) Natural Terrestrial Vegetation
- (Semi) Natural Aquatic or Regularly Flooded Vegetation
- Artificial Surfaces and Associated Areas
- Bare Areas
- Artificial Water Bodies, Snow and Ice
- Natural Water Bodies, Snow, and Ice

Each specific land-cover category is further described in a second modular phase using a set of classifiers (and modifiers) and environmental attributes. The classifiers define the physiognomic-structural vegetation distribution (e.g. cover, height, and stratification), as well as detailed floristic or land-use descriptions. Further details on the classifier specification are described in section 4.5.1.

Any user-defined land-cover class is linked to a Boolean formula, which includes class-specific classifiers and modifiers. A standard class name is generated based on the string of classifiers (LCCS level), as shown for *natural and semi-natural terrestrial vegetation* and *cultivated and managed terrestrial areas* in Table 2-1. Note the increasing level of detail in the standard class name with increasing number of classifiers used. Using LCCS classifiers the initial process of the classification scheme definition can be made comprehensible and comparable, since each class definition entity is stored in the LCCS code. Examples are given in Table 2-1. For example, 'A3' defines the presence of a tree layer, whereas 'A4' represents a dominant herbaceous layer.

Recent advances in the development and communication process of the harmonisation and standardisation of land-cover data resulted in a common acceptance of LCCS as the international standard classification system for land-cover applications (Herold *et al.*, 2006). Moreover, LCCS evolved to the 'common land-cover language' as demonstrated in a number of studies, e.g. for comparisons of existing legends (e.g. CORINE Land Cover and GLC2000 Neumann *et al.*, 2007), legend translations, synergistic analyses of (dis-)agreement (Jung *et al.*, 2006), standardised regional land-cover assessments (Cord *et al.*, 2010), and global mapping initiatives (such as GlobCover and MODIS Land Cover, Defourny *et al.*, 2009, Friedl *et al.*, 2010).

Table 2-1 Definition of standard land-cover classes based on the set of LCCS classifiers (adapted from Jansen & Gregorio, 2002).

Classifiers used	***LCCS level***	***Standard class name***
Natural and semi-natural terrestrial vegetation		
Life form and cover	A3.A10	Closed forest
Height	A3.A10.B2	High closed forest
Spatial distribution	A3.A10.B2.C1	Continuous closed forest
Leaf type	A3.A10.B2.C1.D1	Broadleaved closed forest
Leaf phenology	A3.A10.B2.C1.D1.E2	Broadleaved deciduous forest
Cultivated and managed terrestrial areas		
Life form	A4	Graminoid crop(s)
Field size	A4.B1	Large-to-medium sized field(s) of graminoid crop(s)
Field distribution	A4.B1.B5	Continuous large-to-medium sized field(s) of graminoid crop(s)
Crop combination	A4.B1.B5.C1	Monoculture of large-to-medium sized field(s) of graminoid crop(s)

The LCCS methodology is embedded in the stand-alone software LCCS 2.4 (LCCS2) including the functionalities of classification, legend generation, and legend translation. Further improvements of LCCS-2 resulted in the release of the Land Cover Meta Language (LCML) implemented in the LCCS-3 software. Further details and technical specifications of LCCS-2 and LCCS-3 are provided by the Global Land Cover Network (GLCN, FAO, 2009). The research presented in chapter four is based on the LCCS-2 software.

2.2.3 Classification of satellite time series

Satellite time series were already mentioned as a key data source for biogeographic mapping initiatives at regional and global scales. They provide spatiotemporally continuous observations of land-cover (Gillespie *et al.*, 2008). The high temporal resolution is a major advantage for the statistical differentiation of vegetation types. In the following, an overview of the main data and product characteristic is given. Non-parametric classification techniques emerged as the most widely used methods for satellite time series, as described in the following sections.

2.2.3.1 Data and databases

Global satellite-based terrestrial assessments were conducted by different research initiatives and programs and provide the following datasets with a resolution of 1 km: IGBP DISCover (Loveland *et al.*, 2000), MODIS land cover (Friedl *et al.*, 2002), University of Maryland (UMD) land cover product (Hansen *et al.*, 2000), and Global Land Cover 2000 (GLC2000, Bartholomé & Belward, 2005). Cartographic improvements towards a better spatial resolution were delivered by the GlobCover initiative with a global 300 m land cover product for the years 2005-2006 and 2008-2009 (Defourny et al., 2009), and the MODIS land cover product with a spatial resolution of 500 m, delivered by the MODIS land cover team (Friedl *et al.*, 2010) on an annual basis. The main methodological specifications of the globally available products (based on satellite time series) are listed in Table 2-2.

2.2.3.2 Classification and regression trees

The overview of the classification techniques applied for land-cover mapping indicates that, besides unsupervised learning, tree-based methods are widely used for the classification of satellite time series data. The most common methods are *Classification and Regression Trees* (CART, Breiman *et al.*, 1984). CART is also implemented in the C4.5 algorithm used for the classification of the MODIS land-cover product (Friedl *et al.*, 2002; Quinlan, 1993). For instance, the independence of input data distribution is a key property for the classification of multi-frequent data. The classifier makes no assumptions regarding the distribution of the data being classified. Goward & Prince (1995) found that semi-arid systems can benefit from that characteristic, since rainfall driven systems have a distinct spatiotemporal variability, and thus, a multi-modal frequency in the underlying satellite time series.

Table 2-2 Overview of existing land-cover databases, adapted from Herold (2009).

Dataset	***Data***	***Production approach***
IGBP DISCover	AVHRR (1992-93), 1 km	Computer-assisted image processing interpretation, multi-temporal unsupervised classification of NDVI data; classification process is not automated but more similar a traditional manual image interpretation.
University of Maryland global Land Cover Product	AVHRR (1992-93), 1 km, 8 km, 1 deg.	Developed on a continent-by continent basis; automated unsupervised classification procedure; Training data and phenological metrics that describe the temporal dynamics of vegetation over an annual cycle; metrics have the potential to be used as input variables to a global land cover classification.
MODIS Land Cover Products (MOD 12)	MODIS/ Terra (since 2000), 1 km	Supervised classification approach using a decision tree classification algorithm in conjunction with boosting; top-down approach: image classification for the whole globe; using a global suite of training sites interpreted primarily from Landsat Thematic Mapper (TM) data.
Continuous Fields Tree Cover Project MODIS	MODIS/ Terra, (2000- 2005), 500 m	Automated classification procedure using a regression tree algorithm; using high-resolution imagery (Landsat ETM+, IKONOS) to derive global training data; using training data as dependent variable, predicted by independent variables in form of annual MODIS metrics; employing these training data and phenological metrics with a regression tree to derive global percent cover; outputs from the regression tree are further modified by stepwise regression and bias adjustment.
Global Land Cover Map for the year 2000 (GLC 2000)	SPOT Vegetation (2000), 1 km	19 regional products produced by regional GLC2000 partners, with a regionally specific legend; GLC 2000 data interpreters designed individual classification procedures; harmonisation and merging of regional products to one global product with generalised legend (bottom-up approach).
GLOBCOVER	ENVISAT/ MERIS (2005/06; 2008/09) 300 m	A priori stratification to split the world in 22 equal-reasoning regions; for each region, per-pixel classification algorithm to derive homogeneous land cover classes; per-pixel temporal characterisation through temporal metrics and classification; labelling rule-based procedure using best available products and experience of an international experts network.

Decision trees split the feature space in set of sub-spaces using binary decisions. The class boundary is derived by asking a sequence of nested yes/no questions. For example, the first node (root) splits a variable X. Cases larger than the value of five follow the left branch (Figure 2-11). The remaining values go to the right branch where further splits will be performed. Terminal nodes are reached when the data arriving at a node is of a single class. At the point where splitting is no longer possible, the class label is assigned corresponding to the majority class within the terminal node. Different splitting criterions (*impurity* functions) can be chosen. These are based on *probabilities of class assignment* and can be derived using different statistical metrics measuring data heterogeneity, such as *Misclassification error*, *Gini index*, and *Cross-entropy or deviance* (Hastie *et al.* 2009). The ensemble methods applied in this dissertation are using the *Gini index* (Bankhofer & Vogel, 2008), where p_i is the probability at which an element of class i (elements in a node) is chosen, $_s$ is the total number of classes, calculated as follows:

$$G = \sum_{i=1}^{s} p_i^2$$

The *Gini index* is a heterogeneity measurement with a value range from zero to one. The value of $p_i = 1$ (and thus $G = 0$) is given at a maximum heterogeneity. The minimisation of G results in maximum homogeneity of the resulting split at a given node.

The main advantage of recursive binary trees is their interpretability. Decision trees thus emerged to the most frequently used methods for classification problems. Following Seni & Elder (2010) the properties of decision trees can be summarised as follows:

1. *"**Ability to deal with irrelevant inputs** - Nodes will be derived from the best variables. Trees naturally perform variable selection. Trees provide an variable importance score if applied in ensemble framework".*

2. *"**No data processing needed** - Trees can naturally handle binary, categorical, and numerical data".*

3. *"**Scalable computation** - Trees are fast to generate in comparison to other iterative methods".*

4. *"**Missing value tolerant** - Missing values have little impact on results of tree-based methods".*

5. *"**Off-the-shelf procedure** - There are a few tuneable parameters".*

Decision trees are discontinuous piecewise constant models and are not suitable for the prediction of linear relationships. Another limitation is that they are sensitive to variances. Slight changes in the training data can cause remarkable changes in the resulting tree model.

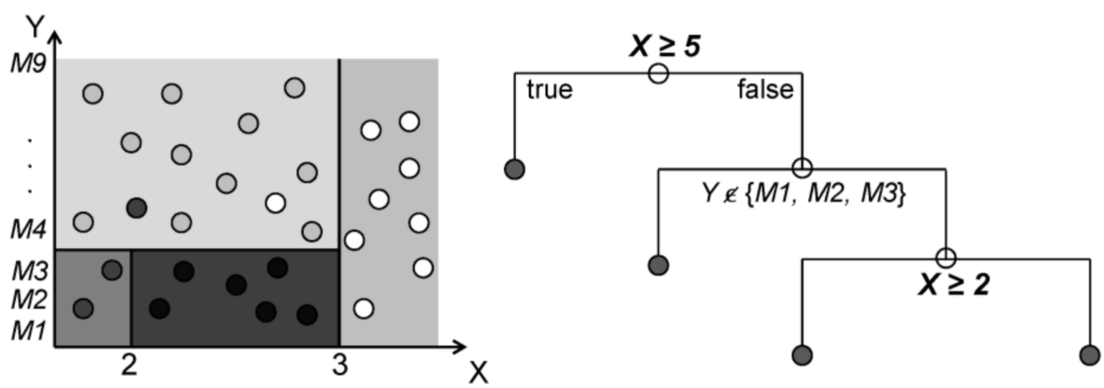


Figure 2-11 Decision tree example. The recursive decision tree shows examples of a "split" node and "terminal" node, adapted from Seni & Elder (2010).

2.2.3.3 Random Forest

Ensemble methods aim at reducing the variance of a single tree-based model prediction and, compared to single models, generally improve the accuracy of a classification. As shown by a number of studies, either with the use of decision trees (Ho, 1995) or neural networks (Brown, 2004; Hansen & Salamon, 1990), ensemble methods have produced significantly higher accuracies. Ensembles will be built by constructing varying models and the combination of their estimates. Combinations can be made by voting or through model estimate weights.

The term *bagging* defines the combination of varied decision tree models. A bagging procedure called *bootstrap aggregation* was introduced by Breiman (1999; 2001; Breiman *et al.* 2006), called Random Forest. Random Forest bootstraps the training data and takes the majority vote of their estimates (Figure 2-12). Using Random Forest a stochastic component is added to increase the "diversity" among the trees in the ensemble. Thus, Random Forests aim at reducing the variance of a prediction function, such as tree-based models with high variance and low bias. The ensemble of trees increases the significance of *weak learners* over time and causes a decreasing classification error (Figure 2-13). The bagging process in Random Forests causes a generation of *de-correlated* trees. Those will be produced by applying a random sub-sampling (with replacement) of the training data. Both regression and classification can be chosen for the prediction (ensemble of trees).

The process is described below according to Hastie *et al.* (2009):

a) For $b = 1$ to B:

1. The bootstrap sample Z^* of size N is drawn from the training data.

2. The random forest tree T_b is grown to the bootstrapped data. The following steps are recursively performed for each terminal node of the tree, until the minimum node size n_{min} is reached.

2.1 M variables are selected at random from the p variables.
2.2 The best variable/split point among the m are picked.
2.3 The node is split into two daughter nodes.

b) Output of the tree ensemble $\{T_b\}_1^B$.

The prediction at each point x is based on:

Classification: $$C_{rf}^B(x) = majority\ vote\ \{C_b(x)\}_1^B$$

Regression: $$\hat{\mathrm{f}}_{rf}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b\ (x)$$

$C_b(x)$ is the class prediction of the bth random forest tree. The majority vote is applied on the classification results C_b of the ensemble of B trees. When regression is used, the target predictions $\hat{\mathrm{f}}_{rf}^{B}$ at point x are averaged. The recommendation in terms of node size is one for regression and five for classification. Figure 2-12 demonstrates the principle of Random Forest classification. The input data is classified with an ensemble of decision trees (DT_1 - DT_n). Classification results of the terminal nodes (TN_1 - TN_n) of each tree are combined by majority voting. The model benefits from averaging by using the majority vote for the final class assignment.

Out-of-bag samples – Within the bootstrapping process in random forest, so-called out-of-bag (OOB) samples will be generated, defined as: *"For each observation $z_i = (x_i, y_i)$, construct its random forest predictor by averaging only those trees corresponding to bootstrap samples in which z_i did not appear"* (Hastie *et al.* 2009). Doing so, the OOB error estimate is equal to an N-fold cross-validation. The cross validation is being performed along with the classification process. Typically, the OOB error stabilises after 200 trees grown which indicates a probable termination of the training process (Hastie *et al.* 2009). The OOB error estimate is identically equal to the producer's accuracy in the classical way of an accuracy assessment. The rapid decrease and levelling of the classification error is shown in Figure 2-13.

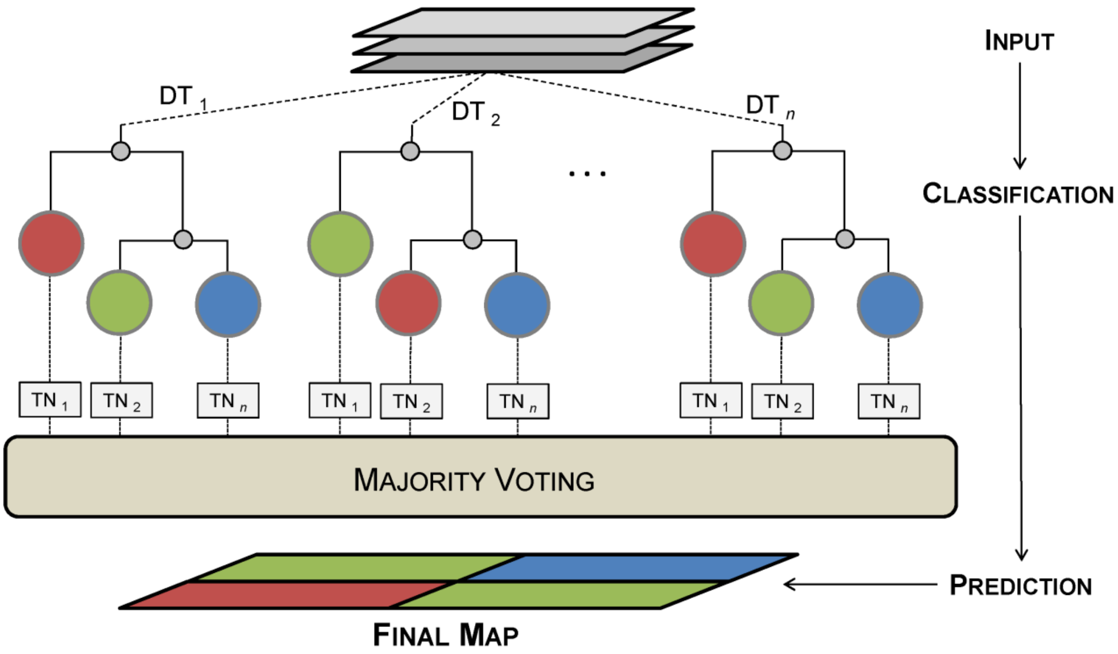


Figure 2-12 Diagram of an ensemble Random Forest classification workflow following Breiman (2001).

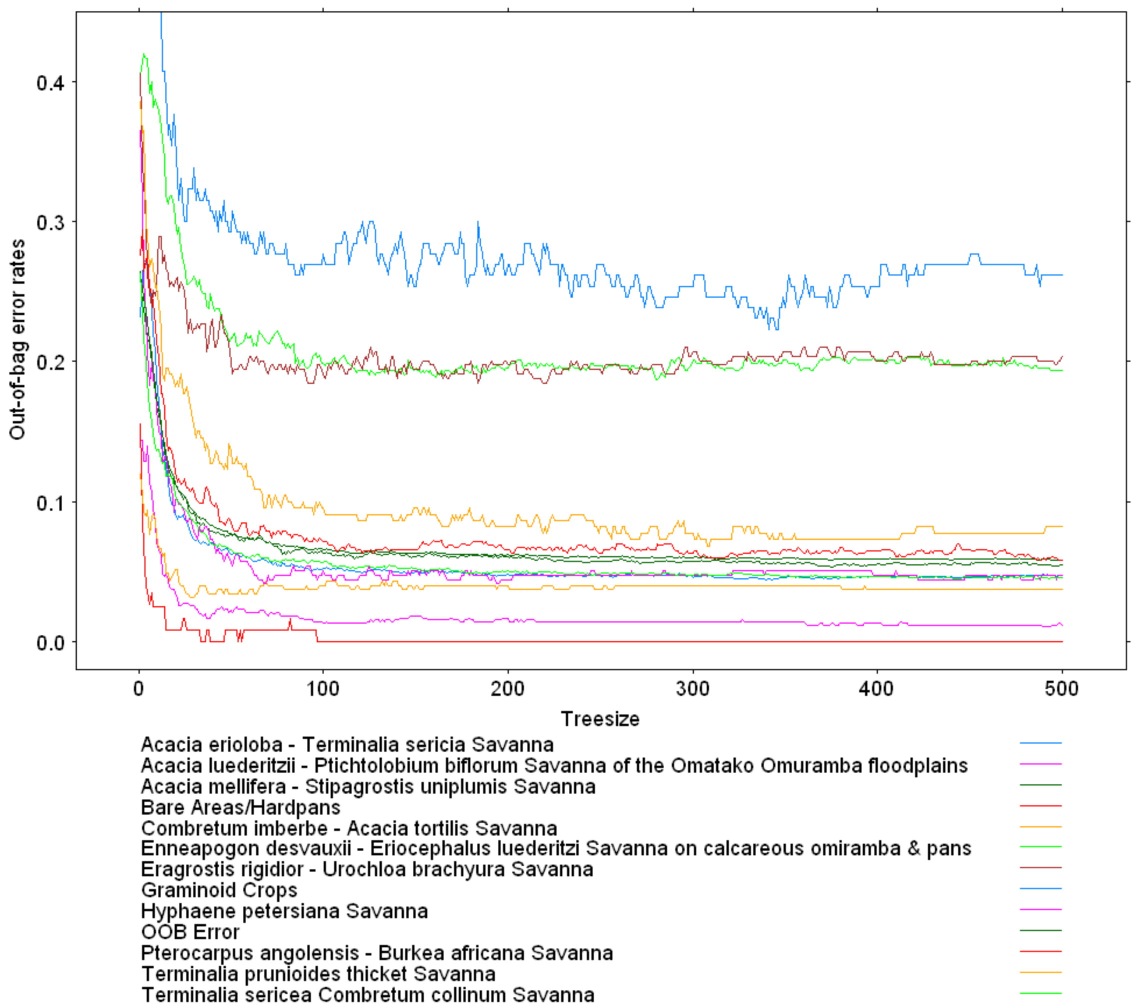


Figure 2-13 Out-of-bag error computed on a random forest classification of Kalahari vegetation types.

Variable importance – An important and forthcoming feature of the Random Forest framework is the calculation of the variable importance score. Two types of variable importance measurements were provided. The first type shows the improvement of the split criterion (*Gini index*). The improvement is attributed to the feature at each split of the tree and finally averaged over all trees. The second type makes use of the OOB samples. The prediction accuracy is calculated when the *b*th tree is grown. The accuracy is again calculated after the values of the *j*th feature are randomly permuted in the OOB samples. The result is a decrease of accuracy as a consequence of the permutation process. The decrease of accuracy is averaged over all trees. The second variable importance score indicates the prediction strength of each feature.

Variable importance assessment is a widely applied method in biological sciences. For instance molecular genetics benefit from feature importance rankings as the classifiers typically have to handle data frames with more than 6.000 predictor variables (Archer & Kimes, 2008; Kuhn, 2008; Strobl *et al.*, 2007). Random forests were widely used in ecological sciences and remote sensing (Prasad *et al.*, 2006), both for optical and Radar applications (Waske & Braun, 2009; Na *et al.*, 2010) In particular the property to deal with weak prediction variables seems to be promising to adapt the feature space of remotely sensed data towards the optimisation of a specific classification problem. In fact, applications based on a multi-dimensional feature space (e.g. classification of time series or features derived from object-oriented feature space generation) will generally benefit from tree-based ensemble methods.

2.3 Summary

The literature review summarised the recent challenges of land-cover assessments and biodiversity monitoring. One of the key issues in all recent initiatives working at global scales is data harmonisation. Vertical data harmonisation increases the thematic comparability of earth observation data at species level (biodiversity) and coarse satellite observation (land-cover and dynamics). Special attention was given to the review of system analyses of savanna ecosystems (e.g. tree-grass interactions, effects of rainfall/fire/grazing, and determinants of vegetation distribution). Savannas are characterised by a distinct spatial heterogeneity of vegetation structure and precipitation patterns. These patterns are influenced by distinct temporal dynamics, such as phenological cycling of vegetation and temporal shifts of the rainy season.

Secondly, an overview was given of the concepts and methods that are relevant to the thematic and methodological framework of this thesis. As for the classification system, there exist huge (scale-dependent) differences with direct consequences on the definition of the term ‘savanna’. This results in a number of different classification schemes being applied in remote sensing projects and vegetation surveys at regional scales. A leading role plays the FAO and UNEP Land Cover Classification System aiming at the standardisation of land-cover classification.

Thirdly, recent advances in the classification of satellite time series were shown. Special attention is given to tree-based ensemble classifiers, such as Random Forest. The application of ensemble machine learning algorithms increased recently in the remote sensing community, as they provide some advantages for the classification of hyper-dimensional satellite data.

3 On the Suitability of MODIS Time Series Metrics to Map Vegetation Types in Dry Savanna Ecosystems

A Case Study in the Kalahari of NE Namibia[1]

Abstract - The characterisation and evaluation of the recent status of biodiversity in Southern Africa's Savannas is a major prerequisite for suitable and sustainable land management and conservation purposes. This paper presents an integrated concept for vegetation type mapping in a dry savanna ecosystem based on local scale in-situ botanical survey data with high resolution (Landsat) and coarse resolution (MODIS) satellite time series. In this context, a semi-automated training database generation procedure using object-oriented image segmentation techniques is introduced. A tree-based Random Forest classifier was used for mapping vegetation type associations in the Kalahari of NE Namibia based on inter-annual intensity- and phenology-related time series metrics. The utilisation of long-term inter-annual temporal metrics delivered the best classification accuracies (Kappa = 0.93) compared with classifications based on seasonal feature sets. The relationship between annual classification accuracies and bi-annual precipitation sums was conducted using data from the Tropical Rainfall Measuring Mission (TRMM). Increased error rates occurred in years with high rainfall rates compared to dry rainy seasons. The variable importance was analyzed and showed high-rank positions for features of the Enhanced Vegetation Index (EVI) and the blue and middle infrared bands, indicating that soil reflectance was crucial information for an accurate spectral discrimination of Kalahari vegetation types. Time series features related to reflectance intensity obtained increased rank-positions compared to phenology-related metrics.

[1] Published as: Hüttich, C., Gessner, U., Herold, M., Strohbach, B. J., Schmidt, M., Keil, M., & S., Dech (2009): On the Suitability of MODIS Time Series Metrics to Map Vegetation Types in Dry Savanna Ecosystems: A Case Study in the Kalahari of NE Namibia. *Remote Sensing*, *1*(4), 620-643.

3.1 Introduction

Plant communities are basic natural resource management units and provide baseline information for ecological processes and functioning in semi-arid rangelands to evaluate dynamic tendencies and grazing capacity (Strohbach, 2001). However, there is a lack of consistent environmental geodata on a national scale in Namibia (Burke & Strohbach, 2000). In the past, a number of vegetation survey projects have been carried out in parts of Namibia. Volk (1966) and Giess (1971) conducted general descriptive vegetation surveys on a national scale. Phyto-sociological and vegetation-environmental studies have been realised in selected regions, e.g., in the Khomas Hochland south of Windhoek (Volk & Leippert, 1971), Central Namib (Robinson, 1976; Jürgens *et al.*, 1997), Waterberg (Jankowitz & Venter, 1987) and in the Etosha Pan by le Roux *et al.*, (1988) and du Plessis *et al.* (1998). Nation-wide estimations of biomass and vegetation cover have been conducted using remote sensing techniques (Strohbach *et al.*, 1996; Sannier *et al.*, 1998; Colditz *et al.*, 2007). Nevertheless, approximately 60 %–70 % of Namibia's surface has not been analyzed in terms of vegetation composition and community structure (Burke & Strohbach, 2000), which emphasises the need for implementing standardised bottom-up-approaches in the field of biodiversity assessment and vegetation mapping in Namibia.

New techniques have emerged during the past years to estimate and analyse land-use and land-cover change (LUCC), land value and land functioning, strongly supported by interdisciplinary approaches of different scientific communities on land change science, remote sensing, geoinformatics, and local scale studies (Verburg *et al.*, 2009; Herold, 2006). Satellite applications have proven to be an effective tool for land-cover mapping and monitoring, providing consistent spectral, spatially explicit and temporal up-to-date indicators of surface processes and status of biodiversity (Lambin, 1999; Nagendra, 2001).

Phenological characterisations of satellite time series of the Advanced Very High Resolution Radiometer (AVHRR) onboard the National Oceanic and Atmospheric Administration satellites (NOAA) and Moderate Resolution Spectroradiometer (MODIS) onboard Terra/Aqua have been widely used for global land-cover mapping purposes (DeFries *et al.*, 1998) and mapping of southern African Biomes and Bioregions (Steenkamp *et al.*, 2008). Statistical metrics computed on MODIS time series have been successfully used as input features to classify land-cover classes characterised through differing phenological patterns, such as Hansen *et al.* (2003) for mapping a global continuous fields tree cover and Gessner *et al.* (2009) for mapping fractional vegetation cover in Namibia on a regional scale.

Non-parametric machine learning algorithms proved to be an effective tool for land-cover mapping using a high number of input variables. Classification and regression trees (CART), as introduced by Breiman *et al.* (1984), have been applied on satellite time series reflectances and vegetation indices, such as the normalised difference vegetation index (NDVI, DeFries *et al.*, 1995; Friedl & Brodley, 1997; Hansen *et al.*, 2000). The application of ensembles of classification and regression trees resulted in increased mapping accuracies and more stable classification results (Chan *et al.*, 2001). Random Forests, a tree-based ensemble classifier (Breiman, 2001), proved to be effective for land-cover classification for mapping agricultural land-use (Pal, 2005) and to map ecotope classes based on hyperspectral imagery (Chan & Paelinckx, 2008). The applications of tree-based classification algorithms have focussed on a per-pixel-basis in the global and regional remote sensing community. Integrating land-cover related object attributes, such as shape and neighbourhood, for classification issues have been recently applied and can enhance thematic depth and mapping accuracies.

Object-oriented classification techniques have been combined with Random Forest classification for mapping agricultural lands by Watts & Lawrence (2008). Classification and regression trees (CART, Breiman *et al.*, 1984) are well known and applied for global classifications of plant functional types (Running *et al.*, 1995; Bonan *et al.*, 2002), mapping fractional vegetation cover mapping on global (Hansen *et al.*, 2003) and regional scales in Namibia (Gessner *et al.*, 2008), and classifying global and regional land-cover maps (DeFries *et al.*, 1995; Chan & Paelinckx, 2008).

Results from legend harmonisation analyses of the main global land-cover maps showed limited class agreements in highly heterogeneous landscapes, characterised by mixed classes of trees, shrubs and herbaceous vegetation. Thus, semi-arid savanna ecosystems, characterised as lands with herbaceous or woody understories and a forest canopy cover between 10 %–30 % have the lowest mapping accuracies compared with more homogeneous vegetation types (Herold *et al.*, 2008; Jung *et al.*, 2006). The most recent global land-cover map on a 300m spatial resolution developed from bi-monthly composites of the Medium Resolution Imaging Spectrometer (MERIS) onboard the ENVISAT platform is provided by the GLOBCOVER project. Validation results over classes related to savanna ecosystems such as closed to open shrubland and grassland (>15% vegetation cover, <5m height) show insufficient users accuracies (Bicheron et al., 2008), which emphasises the need for adopted regional studies to develop standardised training methods for image classification, and estimate the most important satellite-based features for classification.

Since 2000, the interdisciplinary project BIOTA-Africa (Biodiversity Transect Analysis in Africa) has collected and analyzed data on different levels of detail in Namibia, involving different observations such as plant species composition from in-situ data and coarse scale earth observation land-cover data on regional scale. Bottom-up approaches of ecosystem assessment and understanding of landscape functioning can provide basic information to develop reliable regional validated vegetation maps. High resolution satellite time series data provide consistent information on landscape dynamics, disturbances and land change processes.

Towards the development of reliable land-cover geo-information, a thorough understanding on what detail can be mapped on each scale and how land-cover information can be integrated has to be provided to make synergistic use of the multiple observation data sources. Recognizing the existing uncertainties, especially in savanna ecosystems, this paper aims to conduct a bottom-up remote sensing-based vegetation mapping to outline the capabilities for developing reliable land-cover products in dry semi-arid ecosystems. Thus, the purpose of this paper is to:

- present an integrated concept for vegetation mapping in a dry savanna ecosystem based on local scale in-situ botanical survey data with high resolution (Landsat) and coarse scale (MODIS) satellite time series data.

- analyse the suitability of intensity-related and phenology-related metrics derived from MODIS time series for single annual and long-term inter-annual classifications from 2001 to 2007.

3.2 Material and methods

3.2.1 Study region

The study region, as shown in Figure 3-1, comprises the communal areas in the Eastern Kalahari in Namibia with a geographic extent from 17°30’E to 21°E and 19°45’S to 21°45’S. The area is characterised by a sub-continental climate with summer rain sums of 350–450 mm in the long-term annual average, usually with a high variability. The geology of the study area is dominated by Aeolian Kalahari sands with sporadic rock outcrops of sandstone, limestone, schist and dolomite of the Karoo Sequence and the Damara Sequence with a mean altitude of 1,200 m (Mendelsohn & Obeid, 2002). The landscape can be grouped in three main vegetation types after Giess (1971), the Central Kalahari, the Northern Kalahari and the Thornbush shrubland in the western part of the study region. Topographically, the transition between the Central Namibian Highlands and the Kalahari Basin is apparent from the SW-NE oriented incised omirimba (shallow water courses with no visible gradients or visible water course, typical of the arid Kalahari sand plateau (King, 1963).

The area can be classified into seven agro-ecological zones (AEZ) of common land management practices, e.g., the Southern Omatako and Fringe plains of the Central Plateau, Kalkveld, pans and stabilised dunes of the Kalahari Sand Plateau. All AEZ are characterised by extensive grazing and limited cropping (De Pauw *et al.*, 1999).

3.2.2 Field survey and vegetation data processing

In this study, data of different scales were used and each data type contains scale-specific land-cover information. *In-situ* land-cover descriptions resulting from the reconnaissance survey of the landscapes, soils and vegetation of the Eastern Communal Areas provide a detailed description of land types, vegetation composition and physiognomy and habitat settings. However, extensive effort is needed for updating and the spatial alienability is limited.

A number of 422 randomly selected vegetation samples, as described in Strohbach *et al.* (2004), were taken during the late growing season from April to May 2004 by applying a stratified sampling with a standardised plot size of 20 m × 50 m following the Braun-Blanquet approach (Burke & Strohbach, 2000). The survey included GPS reading and the record of floristic composition and habitat information after Edwards (1983) using the SOTER methodology (FAO, 1993). Species and habitat data were archived in the TurboVeg database (Hennekens, 1996) of the National Botanical Research Institute (NBRI), then classified and refined using the TWINSPAN (Hill, 1979) and PHYTOTAB (Westfall *et al.*, 1997) packages.

The resulting vegetation associations were clustered after characteristic, differentiating and typical species (Strohbach *et al.*, 2004), to synergise 12 main vegetation types according to dominant species occurrence, as shown in the synoptic legend in Table 3-1. The main vegetation structure types after Edwards (1983) are moderately closed to open shrub- and bushlands followed by a thicket and a woodland class. The sample size varies among the vegetation types due to the general inaccessibility of the study area, especially the sparsely populated central and eastern areas. Two additional classes (graminoid crops and bare areas/hardpans) were added to the classification legend to properly represent the major land-cover characteristics in NE Namibia.

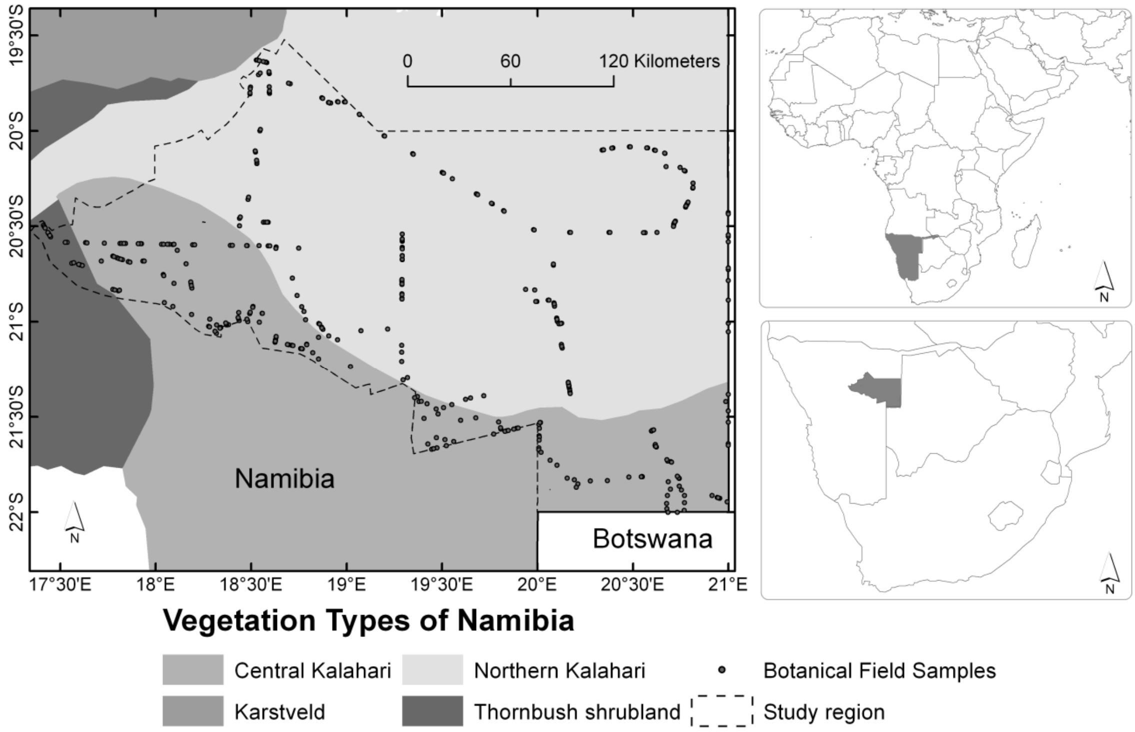


Figure 3-1 Overview of the study region in communal areas in the Namibian Eastern Kalahari showing the main savanna vegetation types after Giess (1971), overlain with the distribution of botanical field samples (Strohbach *et al.*, 2004).

Table 3-1 Synoptic vegetation type legend showing the ten vegetation types synthesised from the phyto-sociological analysis. The structural vegetation type classes after Edwards (1983) as well as mean and standard deviation (Sd) of the cover of tree, shrub and herbaceous layer as sampled in the field indicate slight differences in the life form composition. Note the variances of relevé numbers due to limited accessibility and the extended sample size after the merge process with image objects from the segmentation of Landsat imagery.

Synoptic Vegetation Type Legend	*vegetation structure types after Edwards*	*Cover tree layer [%] Mean*	*Sd*	*Cover shrub layer [%] Mean*	*Sd*	*Cover herb layer [%] Mean*	*Sd*	*No. of plots*	*Smpl. size [pixel]*
Pterocarpus angolensis - Burkea africana woodlands (Pa-Ba)	Tall moderately closed bushland	11.1	8.5	40	7.7	40	11.6	17	758
Combretum imberbe - Acacia tortilis woodlands (Ci-At)	Tall semi-open woodlands	12.3	7.5	6.7	2.9	26.7	15.3	3	582
Terminalia sericea Combretum collinum shrub- and bushlands (Ts-Cc)	Short moderately closed bushland	3.3	4.6	40.0	9.1	34.6	12.6	212	3,500
Acacia erioloba - Terminalia sericia bushlands (Ae-Ts)	Tall moderately closed bushland	4.6	7.0	36.7	10.5	30.8	14.3	59	3,480
Hyphaene petersiana plains (Hp_pl)	Short moderately closed bushlands	4.8	3.8	41.9	8.0	27.5	7.6	8	316
Acacia mellifera -Stipagrostis uniplumis shrublands, Typicum (Am-Su)	High moderately closed shrublands	2.9	4.8	40.0	11.6	33.7	14.4	87	3,834
Enneapogon desvauxii - Eriocephalus luederitzi short shrublands on calcareous omiramba & pans (Ed-El)	Tall semi-open shrublands	2.2	4.4	23.6	15.1	44.7	15.6	9	568
Acacia luederitzii - Ptichtolobium biflorum floodplains of the Omatako Omuramba (Al-Pb)	Short moderately closed bushland	10.3	5.5	31.7	14.4	21.7	9.8	6	770
Terminalia prunioides thickets (Tp_th)	Tall moderately closed thickets	18.3	7.6	38.3	10.4	28.3	12.6	3	220
Eragrostis rigidior - Urochloa brachyura grasslands (Er-Ub)	Tall semi-open Shrubland	0.4	0.7	16.1	8.3	72.9	17.5	12	304
Graminoid Crops (Gr_cr)									130
Bare Areas, Hardpans (pans)									120

3.2.3 Satellite data

Five Landsat-7 ETM+ scenes with six reflectance bands (1–5, 7) from 0.45–2.35 μm spectral range (path 176 to 178, row 74 to 75) acquired between March and June 2000 with 30m resolution were used as input for the object-oriented segmentation explained in section 3.2.4. The MODIS collection 5 product (MOD13Q1, 16-day composites in sinusoidal projection) at the original 232 m resolution was used for time series analysis. The pre-processing included subsetting to the areal extent of the vegetation survey plots and quality analysis. Low quality data were identified based on the MODIS quality assessment science data sets (Justice *et al.*, 2002) and gaps were filled by linear interpolation. The quality analysis was conducted using the Time Series Generator (TiSeG) software (Colditz *et al.*, 2008). Annual time series were calculated for EVI and the blue (459–479 nm), red (620–670 nm), nir (841–876 nm), mir (1,230–1,250 nm) spectral bands (Huete *et al.*, 2002). Bi-annual cumulative precipitation data as averaged over the study region from the Tropical Rainfall Measuring Mission (TRMM, Satake *et al.*, 1995) were used to analyse the relationship of classification error rates to bi-annual rainfall conditions.

3.2.4 Training database generation

An object-oriented segmentation was performed on five Landsat scenes to retrieve homogeneous objects based on similar reflectance settings to regionalise *in-situ* information of the training points on a coarse MODIS pixel size. Using the multi-resolution segmentation function in the Definiens Developer 7.0 software package (Definiens, 2009), three segmentation levels were created on pixel level data to locally adapt the optimisation procedure for computing homogeneous image segments.

The first segmentation level was generated using a scale parameter of 10. The higher levels were computed on increased shape and compactness values (level 2) and a scale parameter of 25 (level 3). Increasing the scale parameters caused that small scale image objects were merged in homogenous objects to achieve a 232 m pixel size. Each specific botanical field plot was intersected with the surrounded Landsat segments. The segments assigned with a training class label were used as training database for MODIS time series metrics, as visualised in Figure 3-2. The up-scaling procedure resulted in an increase of the spatial extent of the training database (compare the plot number of the botanical field survey with the sample size applied on the MODIS data in Table 3-1), a key issue to capture the phenological variability of semi-arid vegetation in the MODIS features within the training process. To avoid the inclusion of natural disturbance by fires in the training database, the MODIS burned area product (MOD45A1, Roy *et al.*, 2008) was used to exclude training areas located on burnt training sites.

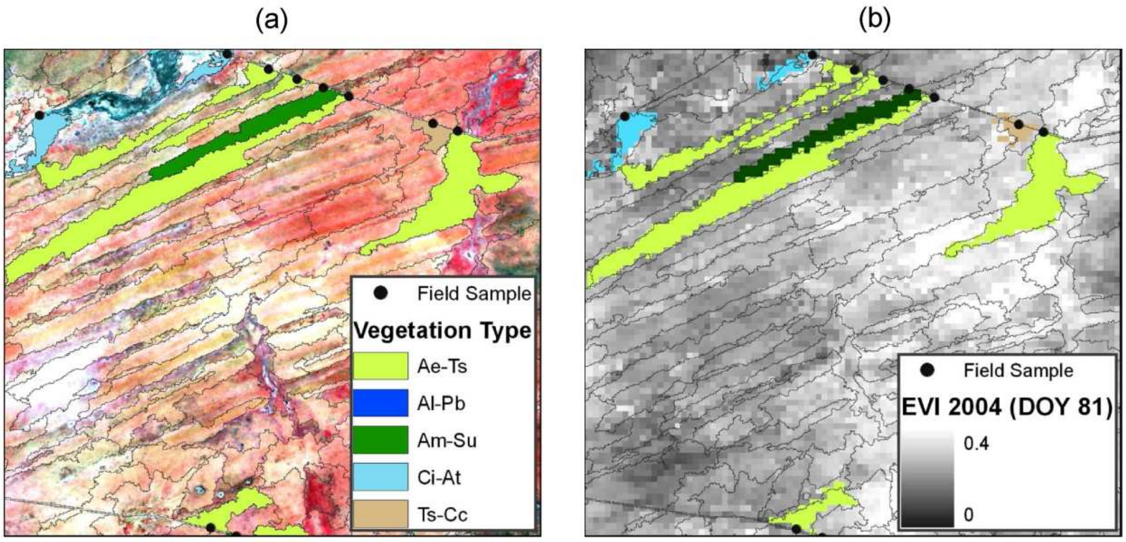


Figure 3-2 Examples for the generation of training data from *in-situ* to MODIS 232 m pixel size. Botanical field plots were intersected with homogeneous segments retrieved from Landsat imagery (2a). The training data on the 232m MODIS pixel is visualised in (2b), displayed on the MODIS image of the 81th day of the year (DOY) 2004. Note the detailed description of the vegetation type legend in Table 3-1.

3.2.5 Calculation of time series metrics

The feature set for the classification was arranged from two kinds of time series metrics, such as phenological metrics and intensity-related metrics derived from a temporal segmentation of the annual time series. The metrics were computed on an annual basis from 2001 to 2007. The resultant annual statistical features were used to derive inter-annual metrics by computing the long-term statistics for each annual segment metric, as displayed in the flowchart in Figure 3-3. Phenological metrics including information on the timing of recurring vegetation cycles (emergence and senescence of canopy) and were generated using the TIMESAT software (Jönsson & Eklundh, 2004). An adaptive Savitzky-Golay filter was applied on the upper envelope of the seasonal EVI curve to reduce atmospheric errors and enhance the EVI values to a meaningful phenological curve. The season cut-off parameter was determined to a number of one growing season per year. The start of the growing season (SGS) and the end of the growing season (EGS) were defined as the date for which the EVI value has increased to 25% between minimum and maximum EVI. Based on SGS and EGS, the length of the growing season (LGS) and mid position of the growing season (MGS) was derived. Further phenological metrics were extracted that describe the curve characteristics of the seasonal EVI time series such as maximum EVI value, base EVI value, left derivative, right derivative and amplitude. The integrated EVI over the growing season was computed, a widely used proxy for net primary production (NPP).

Metrics related to reflectance intensity values of the spectral bands of the sensor were used as single date images for the 16-day EVI composites, and additional statistical metrics were computed from the temporally segmented MODIS EVI and reflectance bands. Three temporal segments per year were derived from the annual time series following the main seasonal characteristics of vegetation activity in Namibia. The first segment (Jan.–Apr.) features the main rainy season (mid- to late summer), segment two (May–Jul.) the fall and winter season and segment three (Aug.–Dec.) the hot, dry spring and early summer. Temporal mean, standard deviation, minimum, maximum, and range (between minimum and maximum) were calculated on each temporal segment from the spectral bands of the MOD13Q1 product (EVI, blue, red, NIR, MIR). The feature generation workflow is visualised in Figure 3-3.

Classifications were performed based on the annual and long-term feature sets. The annual classification sets consist of the 75 MODIS derived annual statistical segment metrics, 10 MODIS-derived phenological date- and NPP-related metrics, and 23 Savitzky-Golay-filtered EVI 16-day composites for the analysed year (DOY 001-353). The total number of the annual feature set was 108. The 75 statistical features from the temporal segmentation for six years of MODIS time series were used to derive 375 inter-annual features.

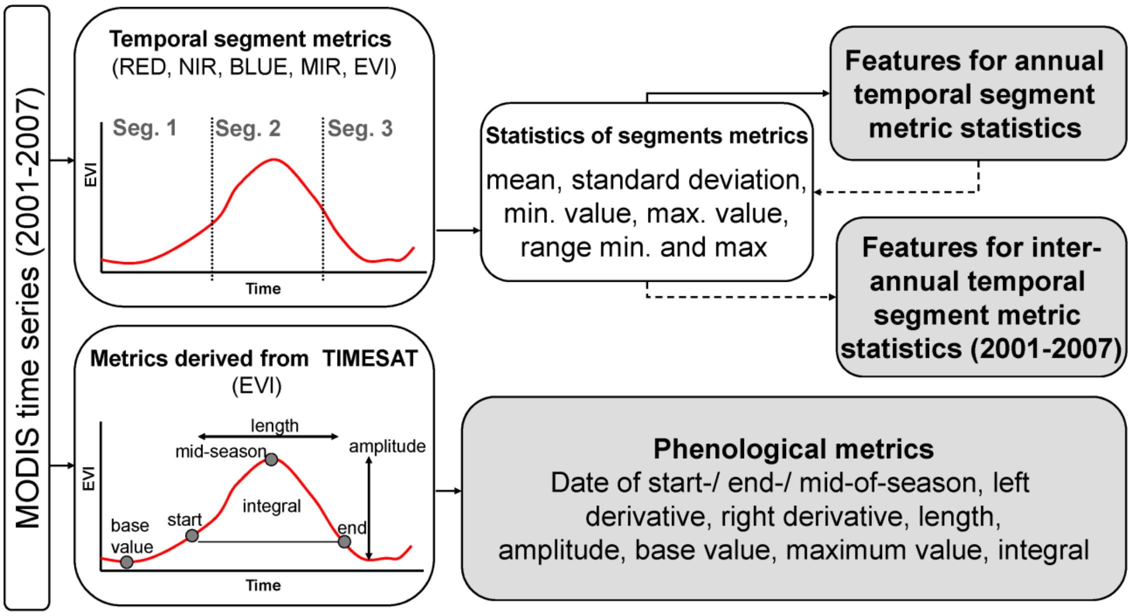


Figure 3-3 Flowchart of the extraction of intensity-related temporal segment metrics and phenological time series metrics derived from the TIMESAT software (Jönsson & Eklundh, 2004). Note the resulting feature sets for the classification in the grey boxes.

3.2.6 Separability analysis

A separability analysis was performed to estimate the suitability of the MODIS time series derivates for classifying vegetation type units. Bhattacharyya distance (B-distance) has been successfully used for time series discriminant analysis (Chaudhuri *et al.*, 1991) land-use classification and pattern recognition in urban areas (Herold, *et al.*, 2003). The B-distance coefficient indicates the probable proportion of variance reduction and class discrimination for a multi-feature space. The B-distance is implemented in MULTISPEC, a free software package designed for hyperspectral image analysis (Landgrebe & Biehl, 2001). The advantage of using B-distance compared with other common distance measures like transform divergence and Jefferies Matusita distance is its large dynamic range, which does not saturate when applied to a large set of features. The classification was performed using a Random Forests. Those non-parametric tree-based classifiers do not require a Gaussian distribution of the data. However, B-distance is a parametric separability measure. Regarding the differing statistical assumptions, a direct combination of both methods is problematic. Here, B-distance and Random Forests were applied independently. Bhattacharyya distance measures were computed on two different feature sets. The first set includes the date- and NPP-related phenological metrics and the annual seasonal segment metrics of the year 2004 as the reference year for the vegetation survey with a total number of 108 features. The second comprises a number of 375 features of long-term temporal segment metrics (2001 to 2007).

3.2.7 Random Forest classification

Decision trees (DT) are structured in simple binary decisions. They are independent of data distribution and can handle categorical variables. The hierarchical structure further allows a biogeophysical interpretation of the relationship between input features and classes and can be useful if multi-source high dimensional remote sensing data is used. Bagging (bootstrap aggregation) attains to reduce variance of a classification by training a number of weak classifiers with varying bootstrap samples from the training set and subsequently averaging the predictions. Boosting is based on classifications using different takes of weighted training sets to be combined in the resultant prediction (Hastie *et al.*, 2009). Comparisons of bagging using Random Forest and boosting based on Adaboost achieved best accuracies for Adaboost at the cost of computation time (Chan *et al.*, 2001; Chan & Paelinckx, 2008; Gislason *et al.*, 2006).

Random Forest is a tree-based classifier where multiple trees are produced and combined based on equally weighted majority voting. A randomly selected third of the original training dataset is excluded for training each particular tree. This so-called out-of-bag (OOB) bootstrap sample is randomly permuted among the input features for each tree. With the remaining 2/3 of the training data, trees are grown to their maximal depth using the impurity gini index (Breiman *et al.*, 1984) due to the fact that the random permutation of samples and features antagonises overfitting. Here, the Random Forest package implemented in the R statistics language (Liaw & Wiener, 2002) was used.

The OOB sample is used to estimate the prediction error for each permutation. Gislason *et al.* (2006) and Breiman (2001) showed that the prediction error based on the OOB sample is slightly higher than using an independent test set. The more conservative OOB error was used to estimate classification accuracies. The variable importance function implemented in Random Forest is based on the internal OOB error estimates. The prediction error is computed on every tree based on the OOB bootstrap sample. In a second step, the OOB error is computed by permuting each predictor variable. The difference in the accuracy measures are averaged over the tree ensemble. The accuracy for each variable is normalised by the standard error (Liaw & Wiener, 2002). In this study, variable importance is examined to evaluate the used features in an ecological context. A detailed overview of Random Forests is given in Breiman (2001).

3.3 Results

3.3.1 Suitability of MODIS time series metrics for semi-arid vegetation mapping

The characteristic phenological patterns of each vegetation type is represented by the Enhanced Vegetation Index (EVI) time series, shown in Figure 4, exemplary for the rainy season of the year 2004 with the dry-season transitions. Slight differences in the amplitude of green vegetation become apparent by comparing the maximum EVI values of moderately closed and semi-open shrub- and bushland classes. Semi-open vegetation types reach maximum EVI values between 0.45 and 0.5, while closed vegetation classes reach their maximum at 0.55. Distinct differences are visible in the green-up and senescence between vegetation structural types.

In general, as shown in Figure 3-4 (a and b), a significantly steep slope during green-up is apparent within the moderately closed shrub- and bushland classes groups *Acacia erioloba - Terminalia* sericea bushlands (Ae-Ts), *Terminalia sericea - Combretum collinum* shrub- and bushlands (Ts-Cc), and *Acacia mellifera - Stipagrostis uniplumis* shrublands (Am-Su). Compared with that, a slow increase in photosynthetic activity is visible in *Pterocarpus angolensis - Burkea africana* woodlands (Pa-Ba) and

Hyphaene petersiana plains (HP_pl). *Acacia luederitzii - Ptichtolobium biflorum* floodplains of the Omatako Omuramba (Al-Pb) and *Terminalia prunioides* thickets (Tp_th) have a distinct offset in greening and early senescence. For the rain period of 2003–2004, the time frame of green-up onset varies between end of November and January (day of year 337/001). The semi-open shrub- and bushland vegetation types are also characterised by a slight increase of vegetation activity and a lower peak in amplitude. An appreciable impact of the so-called small rainy season (varying in time and intensity from September to December) is visible in the *Eragrostis rigidior - Urochloa brachyura* grasslands (Er-Ub) expressed by a small peak before the main vegetative activity, which may be typical for vegetation types with a high grass cover.

The results of the separability analysis indicate the temporal and spectral capabilities of the MODIS-MOD13Q1 product to discriminate vegetation type patterns. The Bhattacharyya distance indicates a relative measure of how reliable one class can be statistically separated compared to the remaining classes. A low distance value indicates a low separability and *vice versa*.

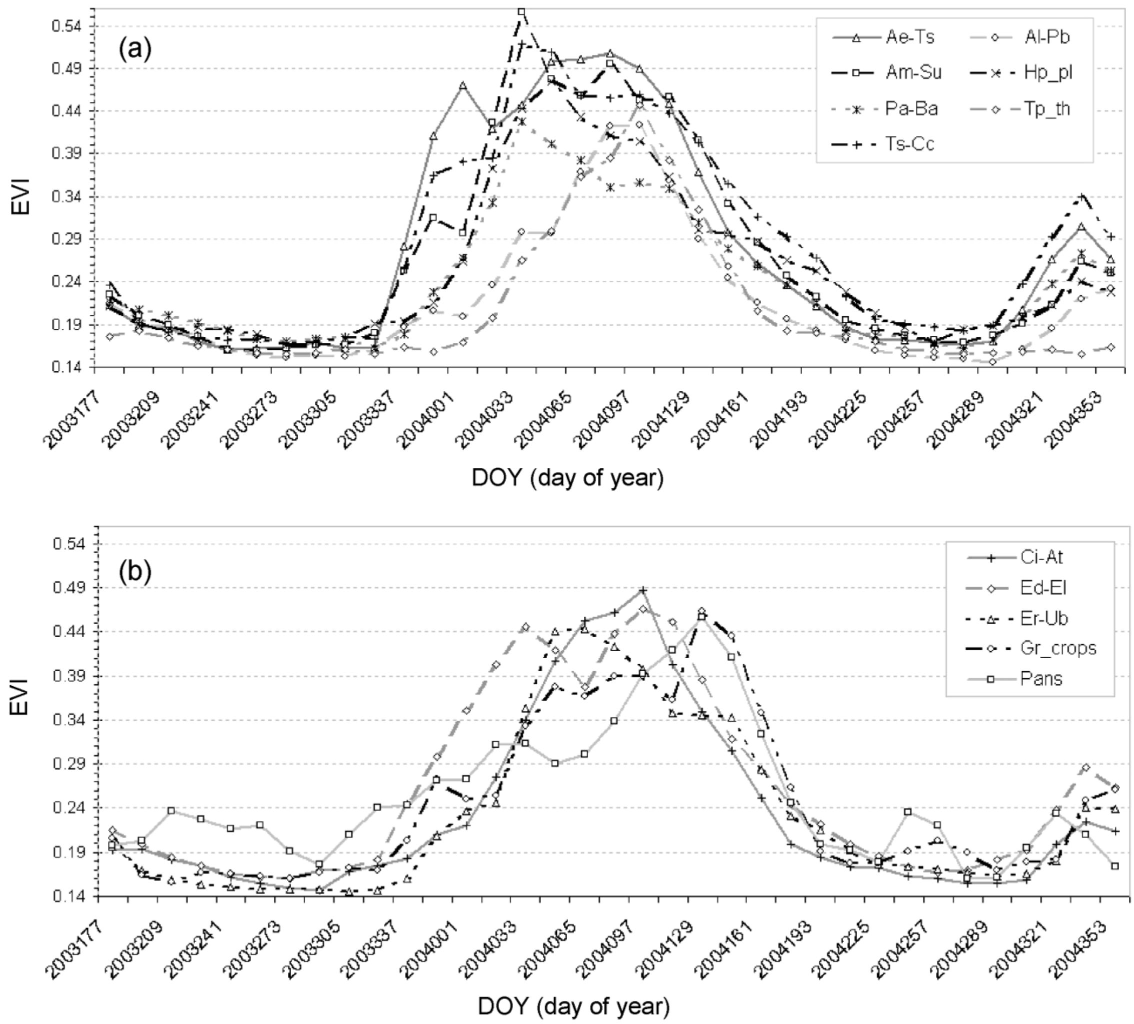


Figure 3-4 EVI time series of the year 2003–2004 averaged for the vegetation type classes with 4a moderately closed shrub- and bushland vegetation cover and 4b semi-open shrub- and bushland vegetation after Edwards (1983) and graminoid crops and pans shown in the MODIS (MOD13Q1, 232 m) Enhanced Vegetation Index (EVI) smoothed with a Savitzky-Golay filter. See Table 3-1 for class labels.

Table 3-2 showing average Bhattacharyya distance measures between Kalahari vegetation type classes, calculated on MODIS intra-annual segment metrics for the season 2003–2004 (lower left values), indicate strong spectral and temporal confusions and thus a more uncertain classification between vegetation types with similar phenology, e.g., *Acacia mellifera - Stipagrostis uniplumis* shrublands, *Acacia erioloba - Terminalia sericea* bushlands, *Terminalia sericea - Combretum collinum* shrub- and bushlands (Ae-Ts vs. Ts-Cc = 0.2, Ae-Ts vs. Am-Su = 0.31).

The highest mean inter-class B-distance values reached the Bare areas/ Pans caused by significant delay of the growing period due to late green-up onset caused by a typical flooding situation in the end of the rainy season and a higher surface albedo compared with shrub- and grasslands. The significance of bands within the visible and middle infrared spectral range for discrimination of vegetation community classes due to different soil settings is discussed later in section 3.4. Highest B-distance values reach *Combretum imberbe - Acacia tortilis* woodlands (Ci-At) and *Hyphaene petersiana* plains (HP_pl) with 44.2. The reason for that can be seen in contrary phenologies. Ci-At has a stronger slope and early onset and slight slope of decline at the end of the rainy season, whereas HP_pl is characterised by a delay in onset of green-up and fast senescence.

Table 3-2 Matrix showing average Bhattacharyya distance measures between Kalahari vegetation type classes calculated on MODIS intra-annual segment metrics for the season 2003–2004 (lower left) versus long-term inter-annual MODIS segment features (upper right, italic values) from 2001 to 2007. See Table 1 for detailed class labels. Note the bold values benchmarking examples for the lowest and highest B-distance values.

		Inter-annual MODIS segment features 2001-2007											
		Gr_cr	Ae-Ts	Al-Pb	Am-Su	Ci-At	Ed-El	HP_pl	Pa-Ba	Tp_th	Ts_Cc	Er_Ub	Pans
Annual MODIS segment features 2004	Gr_cr	-	30,6	131	31,2	191	44,8	62,3	72,2	54,5	36,9	63,8	140
	Ae-Ts	3,42	-	43,7	**2,87**	125	**7,44**	33,6	23,3	25,8	**5,78**	30	151
	Al-Pb	13,4	6,31	-	37,1	77,7	63,8	167	140	128	81,3	103	146
	Am-Su	3,49	**0,31**	5,94	-	124	**11,8**	31,8	30,5	21,8	**13,5**	38,7	142
	Ci-At	23,8	18,9	10,5	18,8	-	151	284	211	241	153	120	144
	Ed-El	7,17	**1,47**	6,73	**1,53**	21	-	30,2	37,9	31	**12,3**	44,1	155
	HP_pl	13,3	7,03	**23,1**	6,89	**44,2**	7,5	-	79,6	48,7	41	116	162
	Pa-Ba	7,52	**1,8**	11,6	2,37	**29,7**	3,28	4,64	-	59,2	16,9	31,6	159
	Tp_th	8,46	6,72	10,5	5,99	**26,6**	9,21	17,3	8,51	-	40,8	78,6	143
	Ts_Cc	3,78	**0,2**	6,25	**0,24**	18,3	**1,42**	6,43	**1,61**	6,3	-	27,4	162
	Er_Ub	7,71	5,63	8,18	6,31	8,6	6,73	21,1	11,5	13,6	5,65	-	146
	Pans	15,3	14,7	17,6	14,5	15,4	16	18,8	17,6	14	14,4	12,1	-

3.3.2 Vegetation type mapping

The vegetation type mapping is based on a vegetation survey in Namibia's eastern communal areas and describes the transition zone between Kalahari to the Otavi Mountains where calcareous geology is dominating in the NW and the mountain savanna transition zone at the SW border of the study area. The resulting vegetation type map is shown in Figure 3-5.

In addition to the Omatako Omuramba, a number of smaller omirimba valleys cross the study area in SW-NE direction. In general, the spatial distribution of vegetation types follows the major topographic units. The *Combretum imberbe - Acacia tortilis* woodlands are typical for the main river bed of the Omatako Omuramba, dominated by bright soils, as displayed in in Figure 3-5b. These open woodlands change to *Enneapogon desvauxii - Eriocephalus luederitzianus* short shrublands on calcareous omirimba and pans (Ed-El), visible in the lower river course where the *Acacia luederitzii - Ptichtolobium biflorum* floodplains of the Omatako Omuramba (Al-Pb) are disappearing. Besides the Omatako river, the smaller omirimba valleys are mainly mapped as Ed-El or *Eragrostis rigidior - Urochloa brachyura* grasslands (Er-Ub), whereas their periphery is mapped as linear patches of *Acacia* *mellifera - Stipagrostis uniplumis* shrublands (Am-Su), Figure 3-5c. *Acacia mellifera - Stipagrostis uniplumis* shrublands (Am-Su) and *Acacia erioloba - Terminalia sericea* bushlands (Ae-Ts) are the dominating vegetation types in the southern part of the study area classified as Thornbush shrubland and Central Kalahari class after Giess (1971) and mark the transition zone to the central Kalahari. *Terminalia sericea - Combretum collinum* shrub- and bushlands (Ts-Cc) are dominant on deep Kalahari sands.

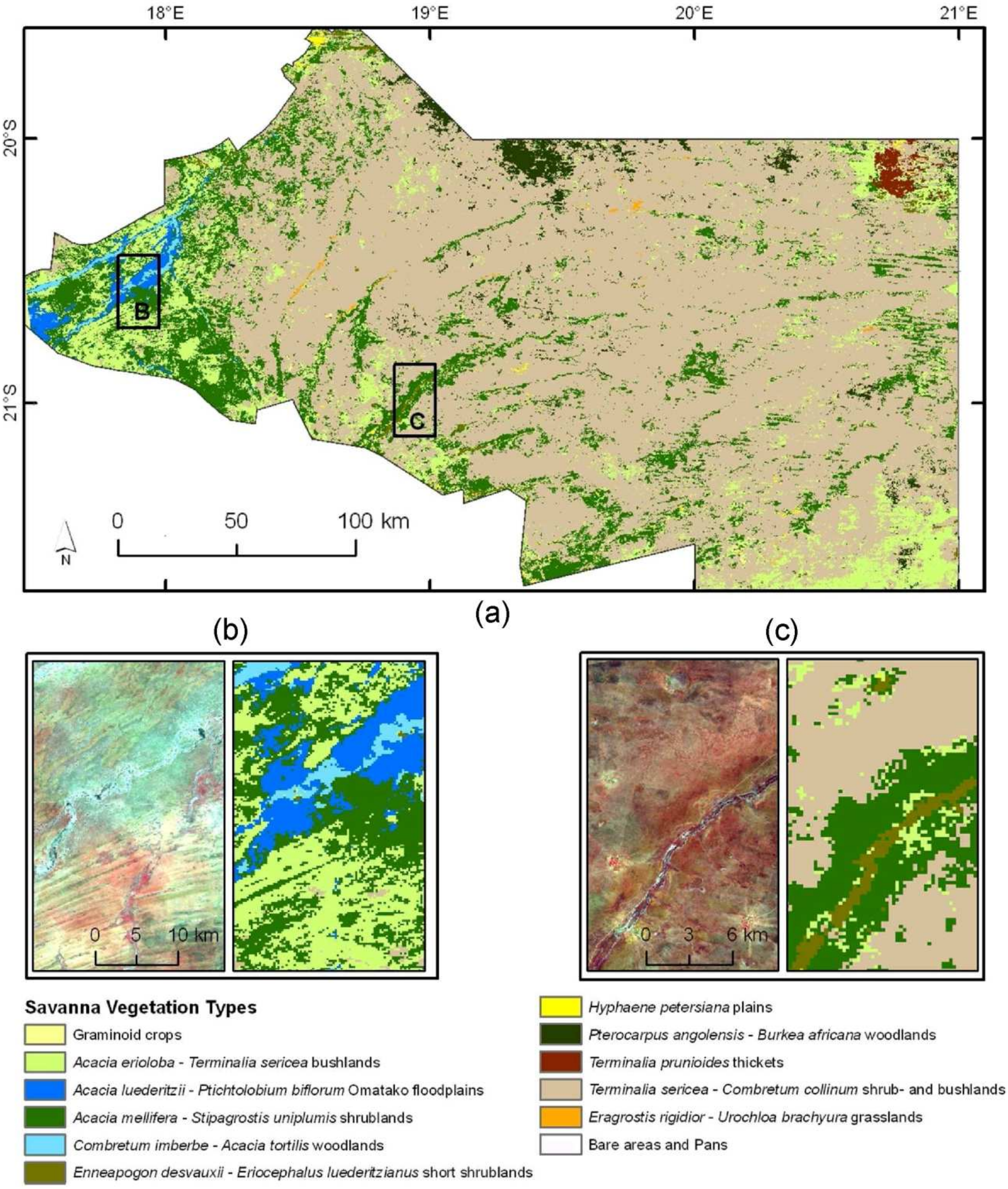


Figure 3-5 Vegetation type classification derived from inter-annual MODIS time series metrics (2001–2007) based on Random Forest classification. The vegetation type map is shown in (a). (b) and (c) show the classification result for examples of the Omatako River region (Box B) and an Omarumba valley cut deep into Kalahari sands (Box C), (b) and (c) are compared to Landsat-TM images (RGB-4-3-2).

3.3.3 Classification error assessment

The classification accuracies of the classifications performed on the six annual sets between 2001 and 2007 and the inter-annual set including the complete time series database are notably different. As described in Section 3.2.6, the map accuracy was calculated on the out-of-bag data randomly selected in each classification iteration. Every pixel in the training database will be used as reference for the estimation of the classification error.

The range of the kappa statistics for annual classification sets is between 0.87 and 0.91. The inter-annual classification set reaches an increased kappa coefficient of 0.93. The resulting maps contain a higher omission error rate with user's accuracies ranging from 83.79 % for the year 2006–2007 to 88.41 % for the season 2004–2005. The lowest commission error rate was reached also for the season of 2004–2005 with a producer's accuracy of 96.11 %, whereas the season of 2006–2007 scored lowest producer's accuracies of 94.17 %. The inter-annual classification set achieved comparatively higher producer's accuracies (97.73 %) than user's accuracies (94.86 %).

The long-term feature set resulted in good classification accuracies for all classes in a range from 85 % to 100 %, where bare areas and pans reached best results with 100% user's and producer's accuracies. Semi-open shrub- and bushland vegetation types show increased spectral confusions and thus reached higher commission errors, e.g., *Enneapogon desvauxii - Eriocephalus luederitzianus* short shrublands on calcareous omirimba and pans (Ed-El) with 85.03 %. Lower producer's accuracies reached *Acacia mellifera - Stipagrostis uniplumis* shrublands (Am-Su, 95.09 %), *Acacia erioloba - Terminalia sericea* bushlands (Ae-Ts, 96.18 %), *Enneapogon desvauxii - Eriocephalus luederitzianus* short shrublands on calcareous omirimba and pans (Ed-El, 96.60 %), and *Terminalia sericea - Combretum collinum* shrub- and bushlands (Ts-Cc, 93.48 %). More significant differences become apparent if seasonal and inter-seasonal classification sets are compared. Graminoid crops were mapped with user's accuracies between 45 % for the season 2006–2007 and 59 % in the 2003–2004 season, whereas 93.84 % user's and 100 % producer's accuracy were estimated based on the long-term feature set. Similar situations were observed for the *Enneapogon desvauxii - Eriocephalus luederitzianus* short shrublands and *Eragrostis rigidior - Urochloa brachyura* grasslands, as shown in Table 3-3.

Seasonal and inter-seasonal MODIS time series features were tested for their suitability to map plant communities of different savanna vegetation types in the Kalahari and Kalahari-transition zone. B-distance values indicate a spectral separability that is approximately ten times higher for the inter-seasonal dataset than the seasonal set of MODIS time series metrics. This result is also represented in the classification error rates. Whereas the classification error range varies between the different classes in the annual classification, a more balanced result among the classes was achieved in the long-term set, ending up in a more accurate map with a Kappa coefficient of 0.93.

Precipitation data from the Tropical Rainfall Measuring Mission (TRMM) were used to compare the annual classification error rates with annual precipitation sums. Therefore, bi-annual cumulative rainfall sums were averaged over the study area. Figure 3-6 shows the comparison of annual OOB error rates and bi-annual cumulative rainfall. Increased OOB error rates occurred in years with high precipitation sums. Compared with the previous years, the growing seasons of 2005–2006 and 2006–2007 with precipitation sums in a range of 1,000 mm to 1,160 mm were wet rainy seasons. These years were mapped with increased error rates from nine to ten percent overall OOB errors. In contrast, the seasons of 2002–2003, 2003–2004, and 2004–2005 were characterised by low rainfall below 800 mm. Comparatively low error rates below eight percent were achieved for those seasons.

3.4 Discussion

3.4.1 Requirements for rainfall amount for vegetation type mapping

To summarise the suitability of multi-temporal MODIS time series features for vegetation type mapping in dry semi-arid savannas, the "best map" could be achieved using inter-annual segment metrics. Significant characteristics from MODIS time series metrics between the mapped vegetation type associations become apparent by including in this case a number of six annual growing periods. The vegetation mapping based on annual satellite time series achieves more reliable results for dryer years.

Table 3-3 User's-, producer's and overall accuracies and Kappa coefficients for six annual and the inter-annual classifications.

class	*2001/02*		*2002/03*		*2003/04*		*2004/05*		*2005/06*		*2006/07*		*2001-2007*	
	users	***prod.***	***users***	***prod.***	***users***	***prod.***	***users***	***prod.***	***users***	***prod.***	***users***	***prod.***	***users***	***prod.***
Gr_cr	50.76	*91.66*	52.30	*93.15*	59.23	*100.0*	48.46	*98.43*	62.30	*100.0*	45.38	*100.0*	93.84	*100.0*
Ae-Ts	94.10	*92.27*	93.99	*92.63*	95.54	*93.03*	94.62	*91.14*	93.50	*89.49*	93.47	*87.42*	96.35	*96.18*
Al-Pb	95.84	*96.72*	95.97	*96.72*	95.58	*94.72*	97.14	*97.01*	97.14	*95.16*	92.33	*92.45*	98.96	*97.94*
Am-Su	92.67	*89.94*	92.98	*89.52*	94.26	*91.07*	92.93	*91.94*	92.22	*90.34*	90.32	*90.53*	96.63	*95.09*
Ci-At	90.89	*97.06*	91.75	*96.56*	91.23	*93.98*	96.90	*96.90*	95.53	*97.03*	87.62	*94.61*	96.28	*97.96*
Ed-El	75.52	*91.86*	75.00	*92.40*	77.81	*91.32*	77.81	*96.71*	69.71	*92.74*	67.78	*93.67*	85.03	*96.60*
HP_pl	92.72	*97.34*	94.30	*98.02*	90.82	*97.61*	93.67	*95.48*	92.08	*97.32*	91.13	*94.11*	97.78	*96.86*
Pa-Ba	94.19	*97.01*	93.27	*97.11*	91.95	*95.74*	96.04	*97.32*	90.50	*97.86*	92.48	*96.42*	97.22	*99.86*
Tp_th	81.36	*96.75*	81.81	*98.36*	91.36	*99.01*	96.36	*99.53*	82.27	*98.36*	90.90	*99.00*	95.90	*100.0*
Ts_Cc	93.34	*89.09*	93.11	*89.28*	91.85	*90.03*	92.22	*89.71*	91.37	*87.97*	91.88	*87.22*	95.48	*93.48*
Er_Ub	73.35	*95.70*	75.00	*95.39*	71.71	*96.88*	75.65	*99.13*	73.02	*97.36*	71.38	*96.44*	84.86	*98.85*
Pans	93.33	*98.24*	93.33	*98.24*	95.83	*99.13*	99.16	*100.0*	98.333	*100.0*	90.83	*98.19*	100	*100.0*
Overall accuracy	**85.67**	**94.47**	**86.07**	**94.78**	**87.26**	**95.21**	**88.41**	**96.11**	**86.50**	**95.30**	**83.79**	**94.17**	**94.86**	**97.73**
Kappa	**0.91**		**0.90**		**0.90**		**0.91**		**0.88**		**0.87**		**0.93**	

As shown in Table 3-1, the major life forms of the vegetation types in the Kalahari are characterised by open to closed shrublands, leading to very similar percent coverage values of the particular life form. This can be challenging in terms of an accurate statistical discrimination when coarse resolution satellite imagery is used. EVI time series are a measure of photosynthetic active vegetation due to the increased reflection of healthy vegetation in the near infrared band. Increased EVI values will be observed due to high rainfall rates. The closer the vegetation cover or the greenness of vegetation, the higher are the observed EVI values. Similar EVI observations for different classes can cause a decreased spectral separability among life form classes. For example, similar EVI values can be observed in a shrub-dominated savanna as in a grass- dominated savanna, if rainfall sums are high enough to produce a full coverage of green vegetation on the MODIS pixel. On a 232 m resolution an observation of a pure pixel for one single homogeneous life form group is improbable in savanna ecosystems due to the heterogeneous vegetation structure. A spectral separation of similar life form compositions, for instance the shrub and herbaceous layer, is therefore difficult, as is reflected by decreased classification accuracies. The similar intensity level of EVI values in the growing season of different life-form classes may be the reason for decreased classification accuracies in the wet rainy seasons of 2005–2006 and 2006–2007. On the other hand, the dry rainy seasons of 2002–2003, 2003–2004, and 2004–2005 effected an increased significance of the reflectance characteristics of the underlying soil properties. The substantial influence of soil properties in local variations in vegetation phenology is discussed by Zhang *et al.* (2005) and thus, the significance of soil-related features for mapping vegetation types in open savannas is discussed in Section 4.2. One reason for the increased accuracy of a long-term feature set can be explained by a characteristic response of each vegetation type class to precipitation conditions, as discussed in Klein & Roehrig (2006). The variability of the phenological cycle of different life-form classes of savanna vegetation was analysed by Archibald & Scholes (2007). They showed that woody vegetation is less sensitive to inter-annual variability than grasses. Memory effects of tree species, as discussed below and differing environmental cues on plant phenology, such as soil moisture for grasses and temperature for some tree and shrub species, can explain those differences. An inclusion of features in the classification process containing information on the inter-annual variability of the photosynthetic activity during the growing seasons, which is a response of spatially and temporally variable rainfall events, can help to separate different life-form classes in savannas.

However, sub-Saharan ecosystems have a high sensitivity to short-term rainfall variability, which can cause significant effects on land-cover dynamics (Vanacker *et al.*, 2005). Hence, vegetation phenology is strongly related to precipitation amounts. Further research has to be conducted to develop a comprehensive understanding of the phenology in dry savannas and their interactions with precipitation patterns.

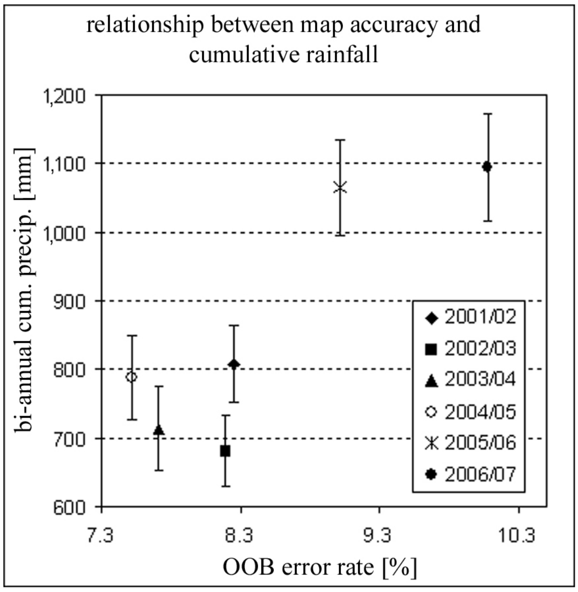


Figure 3-6 Relationship between the out-of-bag (OOB) error rates for seasonal vegetation type classifications and cumulative bi-annual rainfall (mean and standard deviation). Note the increasing OOB error rates with increasing precipitation.

3.4.2 Spectral and temporal requirements for dry savanna vegetation mapping

Remote sensing applications dealing with time series images can be time consuming in terms of image processing and memory allocation. Variable importance can be useful for a pre-selection of relevant input variables to develop an ecologically-based understanding of remote sensing parameters. Here, the variable importance is used to highlight the most useful time series metrics for mapping in an open savanna ecosystem.

Figure 3-7 displays the 15 most important features for the classification of the seasonal datasets (2003–2004) for the intensity-related and phenology-related time series metrics, showing the class-wise variable importance score. An analysis of the variable ranking shows that EVI is the most frequently used variable followed by the middle infrared band. Less scored bands are the red and near infrared bands. The list of the 15 high-ranked features indicates a major importance of the phenological information in the EVI metrics. The high-ranked middle infrared and blue bands indicate that reflectance information in the pedological context to a vegetation type is of major importance to distinguish vegetation type classes. The spatial distribution of semi-arid vegetation types is related to subjacent soil properties. Soil type and colour are represented in the visible spectral range (soil colour) and middle infrared spectral reflectance (soil properties). Regarding the similar life form composition of the vegetation type classes, as displayed in Table 3-1, the phenological patterns were observed to be very similar. Therefore, the information content of EVI may be too low for accurate vegetation type class discrimination which highlights the importance of the blue and middle infrared band in the variable ranking.

Figure 3-7 indicates the dominance of intensity-related features from the temporal segmentation (e.g., BLUE 2004 Seg-3 Max, coding the maximum value of segment three in 2004 of the blue band) followed by the occurrence of the EVI 16-day composites (e.g., EVI 2003 DOY-305, coding the EVI value of the day of the year 305 in 2003).

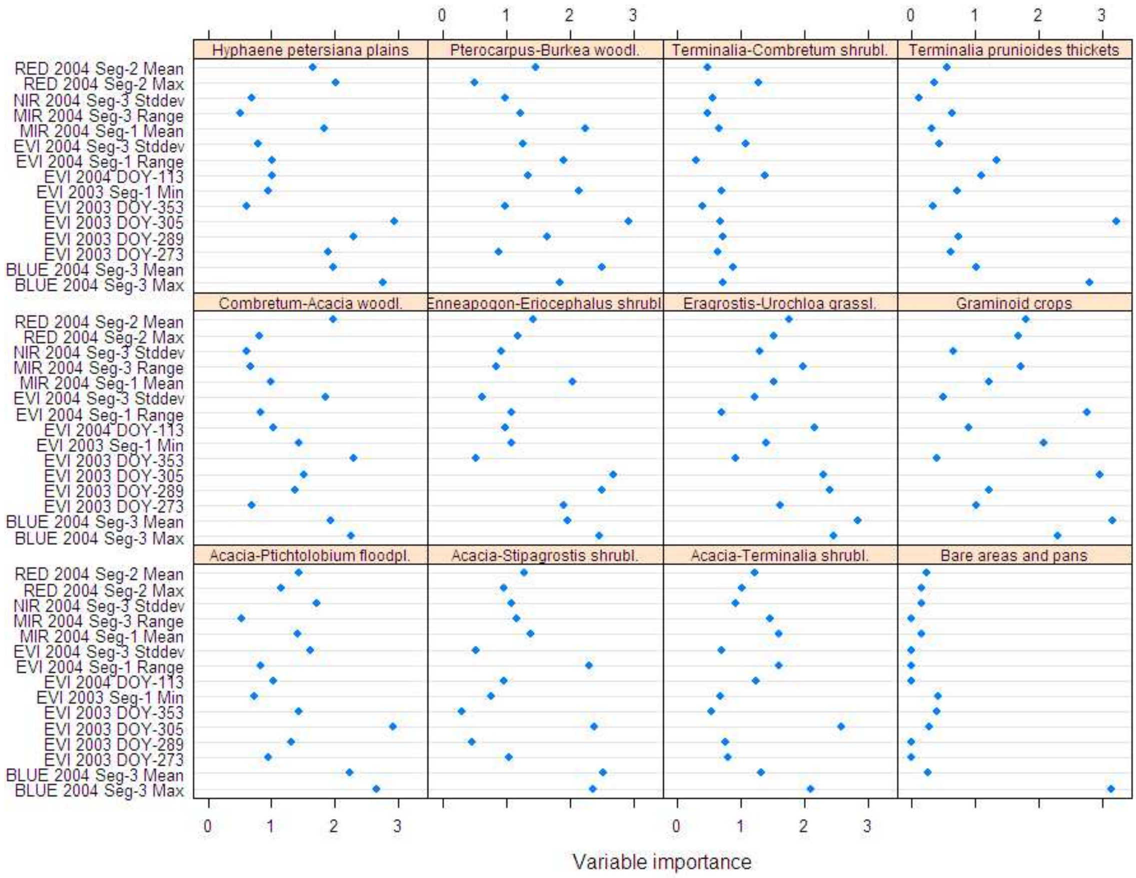


Figure 3-7 Variable importance visualizing the 15 top-ranked MODIS time series metrics for mapping Kalahari vegetation types based on the seasonal feature set 2003–2004 as decrease of OOB error.

The analysis of the variable importance on a per-class basis indicates the bio-physical relevance of each time series metric for an increase of the accuracy in the tree ensemble. An example is shown for the maximum reflectance of the blue band in the third segment of 2004 (BLUE 2004 Seg-3 Max) assigning the highest relevance for the class bare areas and pans.

Due to periodic flooding in pans, the highest variance reduction in the tree ensemble can be achieved using spectral information of the blue band. Similar patterns of increased variable importance of blue reflectance can be observed for the *Eragrostis rigidior - Urochloa brachyura* grasslands and graminoid crops. Both classes are characterised by sparse vegetation cover and assign increased importance values for the features related to soil properties. Similar findings on the response of herbaceous and woody life form groups were discussed in Archibald & Scholes (2007).

Beside the spectral information, the occurrence of date-related features such as the single EVI 16-day composite layers strengthens the hypothesis that dry savanna vegetation types can be statistically discriminated from their phenological patterns. An example is given for the open shrubland classes *Hyphaene petersiana* plains and *Enneapogon desvauxii - Eriocephalus luederitzianus* short shrublands on calcareous omirimba and pans, where EVI features of the beginning rainy season score increased importance values.

In general, date-related features of the early start of the growing period have increased relevance for discriminating open savanna vegetation types. An example is given for the Acacia- dominated savanna classes where the EVI 16-day composite of the day of the year 305 (EVI 2003 DOY-305) scores increased positions. As shown in Figure 3-4 this feature marks the immediate time before the main growing season.

3.5 Conclusions

This study highlights the importance of combining local scale botanical assessments with coarse scale remote sensing applications. It was demonstrated that even though different vegetation type associations showed similar phenological characteristics, they gave good classification results. Improved accuracies were be achieved by the integration of inter-annual time series metrics of six years, which highlights the importance for studying longer-term inter-annual dynamics in dry savanna ecosystems. The feature ranking indicated a high potential of the spectral bands for the discrimination of vegetation type classes. Beyond EVI, an integration of the the full spectral range from the visible to the middle infrared wavelength range is therefore recommended for mapping landscapes with open vegetation cover. Results of the separability analyses highlighted the capabilities and limitations for mapping savanna vegetation types. Vegetation types characterised by differing soil properties reached high Batthacharrya distance values and thus enhanced classification accuracies. Lower B-distance measures between closed shrubland and woodland classes due to similar phenology and soil properties indicated less accurate mapping results. In summary, all vegetation types were mapped with lower error rates in drier rainy seasons, which means that long-term phenological observations including the inter-seasonal precipitation variability is useful for an in-depth characterisation of semi-arid vegetation.

The Random Forest technique used in this study proved to be robust in terms of classification error, overfitting and feature analysis functionality. The use of novel data-mining methods in combination with bottom-up approaches can therefore improve the understanding of remote sensing applications in biodiversity biogeographic research. Yet, further research has to be conducted in similar ecosystems in a synergetic use of different remotely sensed time series data such as the combination of MODIS products and inter-sensor products with botanical field data. Regarding the highly heterogeneous vegetation structure of savannas, new optical sensor systems, such as the five RapidEye satellites and the Advanced Wide Field Sensor (AWiFS) onboard the IRS satellites, seem promising for accurate mapping of open savanna vegetation structure.

Regarding the optimisation of global land-cover products, a process of inter-annual time series metrics beyond one or two years may increase the low class accuracies for semi-arid shrub- and grassland classes. The application of standardised land-cover classification systems is crucial for a bottom-up transfer of botanical information on a coarse remote sensing scale. To expand the vegetation type mapping to the whole Namibia, the most challenging task will be to collect the necessary botanical field data.

A number of more than 10,000 botanical field samples (relevés) are available for Namibia and therefore a highly interdisciplinary project structure is needed to synergise these data on a coarse scale. The translation of data from the botanical survey into the FAO-UN Land Cover Classification System (LCCS) to generate flexible map products will increase the usefulness of land-cover information to a broader user community and will be focussed on in future research.

4 Integrating *in-situ*-, Landsat-, and MODIS data for mapping in Southern African savannas

Experiences of LCCS-based land-cover mapping in the Kalahari in Namibia[2]

Abstract – Integrated ecosystem assessment initiatives are important steps towards a global biodiversity observing system. Reliable earth observation data are key information for tracking biodiversity change on various scales. Regarding the establishment of standardised environmental observation systems, a key question is: What can be observed on each scale and how can land cover information be transferred? In this study, a land cover map from a dry semi-arid savanna ecosystem in Namibia was obtained based on the UN LCCS, in-situ data, and MODIS and Landsat satellite imagery. In-situ botanical relevé samples were used as baseline data for the definition of a standardised LCCS legend. A standard LCCS code for savanna vegetation types is introduced. An object-oriented segmentation of Landsat imagery was used as intermediate stage for downscaling in-situ training data on a coarse MODIS resolution. MODIS time series metrics of the growing season 2004/2005 were used to classify Kalahari vegetation types using a tree-based ensemble classifier (Random Forest). The prevailing Kalahari vegetation types based on LCCS was open broadleaved deciduous shrubland with an herbaceous layer which differs from the class assignments of the global and regional land-cover maps. The separability analysis based on Bhattacharya distance measurements applied on two LCCS levels indicated a relationship of spectral mapping dependencies of annual MODIS time series features due to the thematic detail of the classification scheme. The analysis of LCCS classifiers showed an increased significance of life-form composition and soil conditions to the mapping accuracy. An overall accuracy of 92.48 % was achieved. Woody plant associations proved to be most stable due to small omission and commission errors. The case study comprised a first suitability assessment of the LCCS classifier approach for a southern African savanna ecosystem.

[2] Published as: Hüttich, C., Herold, M., Strohbach., B., & S., Dech (2010): Integrating in-situ-, Landsat-, and MODIS data for mapping in Southern African savannas: Experiences of LCCS-based land-cover mapping in the Kalahari in Namibia. *Environmental monitoring and assessment*, *E-pub*, doi: 10.1007/s10661-010-1602-5. With kind permission of Springer Science and Business Media. Note electronic supplementary material of the online version in Appendix E.

4.1 Introduction

Spatially consistent and standardised land cover data provide key information for the monitoring of the state of ecosystems and biodiversity conservation at various scales. The integration of environmental databases and the development of interfaces for environmental data from local to global scales (e.g. *in-situ* field survey data and global archives of satellite imagery) is a forthcoming task for interdisciplinary land-cover and biodiversity monitoring studies. The delivery of spatially explicit land-cover information plays an important role for ecosystems monitoring (Hollingsworth et al., 2005). In this context the task of the Group on Earth Observations and Biodiversity Observation Network (GEO BON) is to generate a platform to support the standardisation of top-down observations from satellite systems with bottom-up measurements on ecosystem processes and species data (Scholes *et al.*, 2008). Such comparable and standardised mechanisms for land-cover data are needed for developing policy-relevant biodiversity indicators (Pereira & Cooper, 2006; Scholes & Biggs, 2005).

The Land Cover Classification System (LCCS) of the United Nation's Food and Agriculture Organisation (FAO) is the currently most accepted land-cover standard, as it is being implemented as an international standard (ISO 19144-1:2009). On global scales LCCS-based land cover legends were used within the framework of the Global Land Cover 2000 (GLC2000, Fritz *et al.*, 2004) and the 300-m ENVISAT-MERIS product (GlobCover, Defourny *et al.*, 2009). Regional Suitability analyses of LCCS-based land-cover and land-use characterisation were conducted within the Africover (Jansen & Di Gregorio, 2002) and BIOTA-Africa (Cord *et al.*, 2010) framework.

The aim of LCCS is to provide a standardised description of a certain land cover in a set of pre-defined, diagnostic, and hierarchically arranged generic criteria describing a certain land-cover (Di Gregorio, 2005). LCCS uses physiognomic-structural classifiers for the definition of primarily vegetated land cover classes, such as life form, cover, height and spatial distribution. The utilisation of the FAO Land Cover Classification system proved an effective tool for a standardised legend definition on global scales (Latifovic, 2004; Herold *et al.*, 2008). However, little experiences have been exchanged in the scientific community for adaptations of LCCS in regional case studies and to assess the suitability and capabilities of LCCS on fine thematic scales.

Results from legend harmonisation analysis of the main global land cover maps (IGBP DISCover, UMD, MODIS 1-km, GLC2000, Loveland *et al.*, 2000; Justice *et al.*, 1998; Hansen *et al.*, 2003; Fritz *et al.*, 2004; Hansen *et al.*, 2002) showed limited class agreements in highly heterogeneous landscapes (Herold *et al.*, 2008). Primarily in savanna biomes the problem of mixed classes is evident in coarse scale land-cover maps. In order to overcome existing inconsistencies of class descriptions, research has to be addressed to regional adaptations of LCCS in savanna ecosystems.

The world's semi-arid savanna ecosystems cover more than 20 % of the earth's surface. These biomes can be rated as biocomplex systems as the floral and faunal composition is controlled by nonlinear processes (e.g. grazing intensity and fire frequency). Savanna ecosystems are characterised as lands with a mixture of herbaceous and woody life forms, often appearing as a structure of a woody tree and shrub layer with herbaceous understories (Hanan *et al.*, 2006). Compared to more homogeneous vegetation types, savanna classes have the lowest mapping accuracies in the global datasets (Jung *et al.*, 2006). Spatio-temporal variability of rainfall patterns is one of the most prominent uncertainties for semi-arid land-cover mapping based on earth observation data (Scholes *et al.*, 2004; Privette *et al.*, 2004).

The data requirements on the spatial, radiometric, spectral and temporal resolution of the satellite imagery to be used for specific earth observation tasks are often determined by the scale of the study. Three main fields of applications can be stated: (a) direct mapping of individual plants and plant associations, (b) habitat mapping, and (c) the analyses of relationships between spectral radiances and *in-situ* data of species distribution patterns (Nagendra, 2001; Skidmore *et al.*, 1996; Turner, 2003; Muchoney, 2008). Satellite time series proved to be an efficient tool for monitoring and assessing the state of semi-arid ecosystems. At regional scales, time series of the Moderate

Resolution Spectroradiometer (MODIS) were successfully used to map the southern African Miombo ecosytems based on band pair difference analyses (Sedano *et al.*, 2005), characterised as open forests, thickets, and grassland formations (Frost, 1996). Dependencies of the spatial resolution of different environmental datasets were emphasised for the rangelands of the open Kalahari savannas (Trodd & Dougill, 1998). A promising concept of representing savanna vegetation structure is the mapping of the proportional cover of life forms based on satellite time series imagery as done on global (DeFries *et al.*, 2000; Hansen *et al.*, 2003) and regional scales, e.g. for Africa and Australia (Scanlon *et al.*, 2002; Guerschman *et al.*, 2009; Wagenseil & Samimi, 2007; Gessner *et al.*, 2009).

A well-known problem can be seen in the harmonisation of *in-situ* data with coarse scale satellite imagery. LCCS can be used as a common land-cover language (Neumann *et al.*, 2007). On one hand bottom-up ecosystem assessment approaches provide key information to develop reliable regionally validated land cover maps. On the other hand allow top-down satellite observations for large-area measurements of biophysical variables to be used for vegetation type assessments and biodiversity monitoring, such as species-, habitat-, and plant community mapping. In this context, the main research questions are: What can be observed on each scale and how can land cover information be transferred and included in integrated observation concepts? What is the potential of LCCS to be used as a translation tool of land cover data (*in-situ* and satellite earth observations) for savanna ecosystems? Which benefits can be achieved for the connected scientific communities related to biodiversity and remote sensing when using LCCS? The aims of this paper are to:

- assess the applicability of the concept of the LCCS classifiers in semi-arid ecosystems and demonstrate first experiences of using LCCS as a 'land-cover language' in a bottom-up mapping framework.

- present a flexible legend of typical Kalahari savanna vegetation types using the UN-FAO Land Cover Classification System (LCCS).

- demonstrate a concept for a bottom-up ecosystem assessment framework by integrating local scale *in-situ* botanical survey data with Landsat imagery and coarse scale satellite time series data.

4.2 Mapping framework

4.2.1 Study region

The mapping area comprises the eastern communal areas in the eastern Kalahari in Namibia from 17°30'E to 21°E and 19°45'S to 21°45'S. The area is characterised by a sub-continental climate with a long term annual average summer rain period of 324-450 mm and often erratic rainfall events (Mendelsohn & Obeid, 2002). The greater part of the study area is characterised by longitudinal Kalahari sand dunes and plains. Within the Kalahari sand plains, flat calcrete depressions and pans as well as fossil drainage lines (Omiramba) are found. Floodplains and areas subjected to regular flooding are apparent in the western part of the study region (Strohbach *et al.*, 2004).

4.2.2 Field data

The *in-situ* reference database includes a number of 422 botanical relevé samples taken during a reconnaissance survey of soils and vegetation of the eastern communal areas in Namibia from April to May 2004. A stratified sampling scheme was applied with a plot size of 20 m × 50 m after Braun-Blanquet (Strohbach, 2001). The sampling included a detailed description of the vegetation composition (species richness and abundance), life form composition, and habitat settings within a pedologic assessment (see Table 4-1).

Table 4-1 Overview of the land-cover information from *in-situ* data and satellite earth observation data grouped in land-cover aspects concerning vegetation structure, environmental attributes in relation to the spatial resolution and the effort for frequent update.

	Botanical field survey	*Landsat TM*	*MODIS*
Vegetation structure	Synoptic vegetation types indicating the dominant species per class (based on full species list), height class and growth form, cover abundance (crown cover in %) of every height class	local patterns of vegetation physiognomy, e.g. life form composition and coverage, and local patterns of photosynthetic active vegetation observed by the red and NIR bands	phenological characteristics of different vegetation types for the annual growing season observed from time series metrics, derived from MOD13Q1 surface reflectance values in the red band (620-670 nm) and near infrared band (NIR, 841-876 nm) and the derived enhanced vegetation index (EVI, Huete *et al.*, 2002)
Environmental attributes	slope, terrain type, aspect, stone cover estimation, lithology (parent material), erosion severity, surface sealing/crusting, disturbances	geological context, soil texture and top soil surface colour, observed from surface reflectance values in the visible and middle infrared spectral range	geological context, soil texture and top soil surface colour observed from the visible blue and MIR bands, observed from surface reflectance values in the blue band (459-479 nm) and middle infrared band (MIR 1230-1250 nm)
Spatial resolution	20 x 50 m plots (Braun-Blanquet approach)	30 m pixel size	232 m pixel size
Temporal resolution	Periodically (10 – 20 years)	Intra- and inter-annually (0 – 3 years)	16 days

A phyto-sociological analysis was conducted resulting in a synoptic vegetation type legend. The botanical classification nomenclature was based on the occurrence the two characteristic species (Strohbach et al, 2004) and the related environmental setting of each vegetation type (e.g. *Enneapogon desvauxii - Eriocephalus luederitzi* short shrublands on calcareous omiramba & pans). In respect of the classification scheme the essential earth observation variables to classify different land-cover types in the Kalahari can be grouped in observations related to vegetation structure (e.g. life form and coverage) and environmental aspects. The relevant variables from the satellite earth observations representing phenological information and the environmental settings are shown in Table 4-1. The Kalahari vegetation types can be distinguished by distinct temporal characteristics of the vegetation activity during the growing season, as displayed in Figure 4-1.

4.2.3 Satellite data and preprocessing

Five Landsat-7 ETM+ scenes with six channel reflectance bands (bands 1-5, 7, 0.45 - 2.35 μm spectral range, Path/Row/date: 176/074/04-06-2000, 176/075/04-06-2000, 177/074/24-04-2000; 177/075, 24/04/2000, 178/074/ 17-05-2000, GLCF 2007) were pre-processed in terms of geolocation and map projection. 16-day composite images of MODIS collection 5 product MOD13Q1 in sinusoidal projection, surface reflectance bands and the enhanced vegetation index (EVI) in the original 232 m resolution were used for generating features for time series image classification. Pre-processing included subsetting and quality analysis using the Time Series Generator (TiSeG) software (Colditz *et al.*, 2008). Low quality data were identified based on the MODIS Quality Assessment Science Data Sets and replaced by linear interpolation. Annual time series were calculated for EVI and the Blue, Red, Nir, Mir spectral bands. A temporal segmentation was conducted on the annual time series for the vegetation period of 2004/2005. Four temporal segments were derived following the main seasonal characteristics of vegetation activity during the growing season in the NE-Namibia. The first temporal segment (Sep.'04-Dec.'04) features the dry winter season and rainy season transition. Segment two (Jan.'05-Apr.'05) features the main growing period during the rainy season, segment three (May'05-Jul.'05) the transition from rainy to dry season, and segment four (Aug.'05-Sep.'05) the following winter season. Statistical metrics were calculated from the temporal segmentation to capture the temporal variability for each temporal segment from the spectral bands, as shown in Figure 4-2.

4.2.4 Separability analysis and classification of MODIS time series

A separability analysis was performed on vegetation type classes and LCCS-based structural groups. A number of 100 intra-annual MODIS time series metrics computed on the seasonal segments were used as input variables for the Bhattacharyya distance measurement using the Multispec software (Landgrebe & Biehl, 2001). The distance measure indicates the suitability of satellite time series metrics in terms of signal variability to statistically discriminate vegetation community classes in a temporal and spectral feature space in a valid data range from 0 to 30,000.
The MODIS seasonal segment features were used as input variables for the land cover classification. The classification was performed using a tree-based ensemble Random Forest classifier (Breiman, 2001) using the vegetation type training database. The Random Forest package implemented in the R statistic software was used for the data processing. Graminoid crops and bare areas and pans were added to the legend to properly represent all land cover types. The classification accuracy assessment, discussed in section 4.5.2 is based on the out-of-bag (OOB) bootstrap sampling procedure, as integrated in the Random Forest classification procedure. The OOB classification error is a more conservative estimation of the classification accuracy compared to an independent test set, as shown in various studies and was used here for the accuracy assessment. Further details on the classification procedure are described in Hüttich *et al.* (2009).

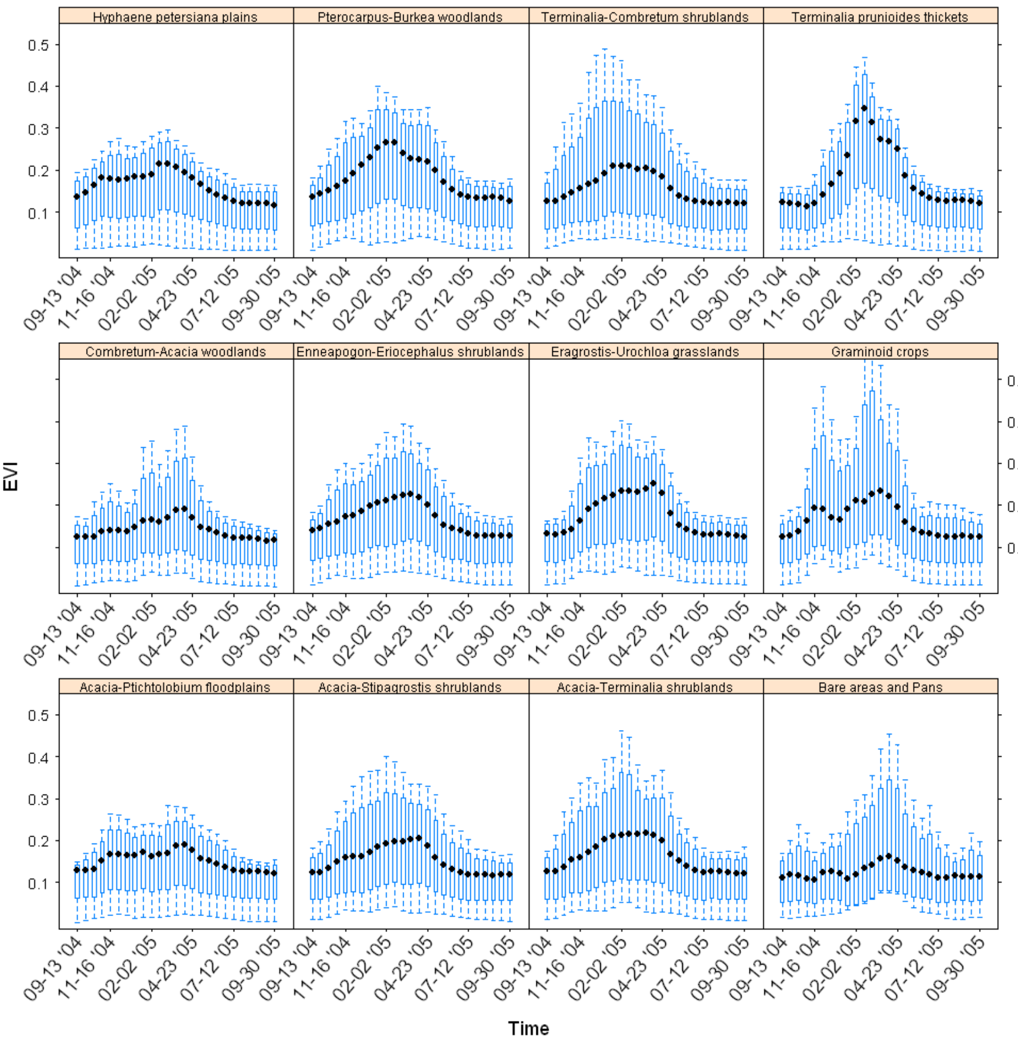


Figure 4-1 Phenological characteristics of the 10 vegetation type classes derived from the botanical field survey indicated by the 16-day MODIS time series of the enhanced vegetation index (EVI) from mid-September 2004 to end of September 2005. The classes bare areas / pans and graminoid crops were added to the classification scheme to properly represent the major land-cover types.

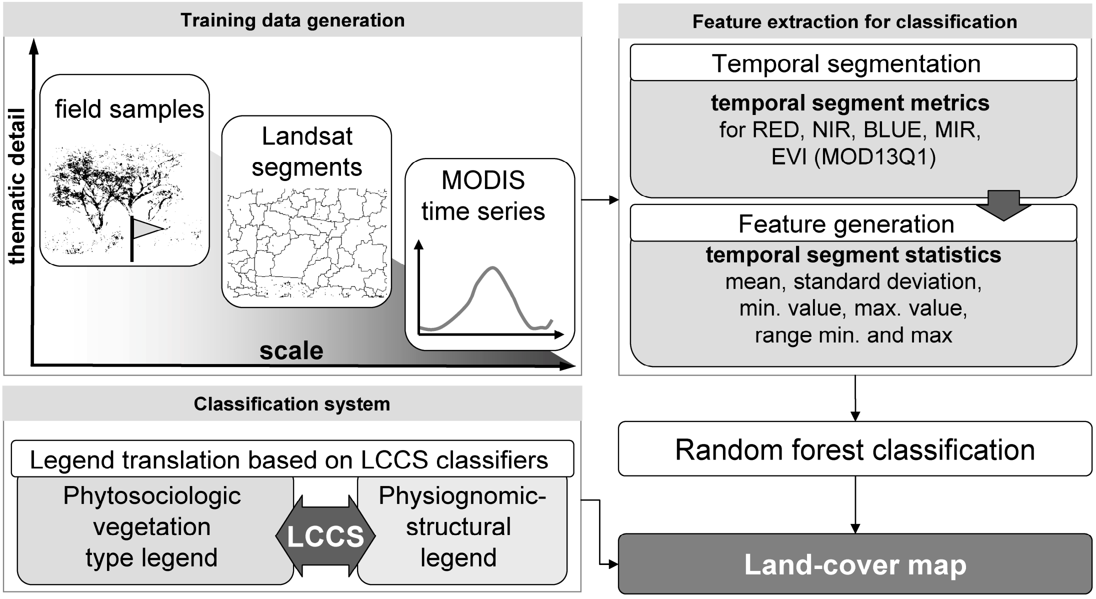


Figure 4-2 Flowchart showing the process the training data generation, feature extraction and classification and the classification scheme generation. The segmentation derived from Landsat imagery was used for the spatial up-scaling from *in-situ* to MODIS scale. The botanical field plots were intersected with the surrounded Landsat segments. The segments assigned with a training class label were used as training database for MODIS time series metrics. The legend translation from the floristic vegetation type legend to a physiognomic-structural legend was based on the specification of LCCS classifiers and modifiers. The classification process including the feature extraction and random forest classification resulted in the final LCCS-based land-cover map.

4.3 Land cover classification system

Global classifications were developed based on the biome concept, including floristic and zoogeographic provinces, and global maps of vegetation types for terrestrial bioregions (Olson *et al.*, 2001) and as baseline legends for global land-cover database initiatives, e.g. the IGBP DISCover project of the International Geosphere-Biosphere Program (IGBP, Loveland *et al.*, 2000). Remote-sensing based classification schemes are based on the structural components of the plant canopy, such as aboveground live biomass, leaf longevity and leaf type (Running *et al.*, 1995). Generic classification schemes were adapted to the African savanna ecosystems integrating broad growth-form and structural classes (Edwards, 1983; Thompson, 1996).

To meet the requirements of unambiguousness and transferability in defining entities of land-cover characteristics without class boundary overlaps, the FAO LCCS was designed as an *a priori* classification system. The main advantages of LCCS are its independence of map scale, data source, geographic location, and field of application (Loveland, 2008). The class definition process is structured in an initial dichotomous and a subsequent modular-hierarchical phase. After the definition of the broad land-cover type in the dichotomous phase, land-cover classes were defined by a set of predefined pure classifiers that can be combined with environmental and specific technical attributes. The class definition can be refined by using modifiers (linked to each land-cover classifier). Land-cover classes were assigned to a Boolean formula (combination of selected classifiers), a unique numerical LCCS code, and a standard class name (Di Gregorio, 2005).

This study focussed on the assessment of the land-cover type *Natural and semi-natural terrestrial vegetation* (LCCS level A12). Classification rules of that module are defined according to physiognomic-structural aspects, such as (a) physiognomy; (b) vertical and horizontal arrangement; (c) leaf type; and (d) leaf phenology. The classifiers related to the vegetation distribution are *Life Form*, *Height*, and *Cover*. A complete list of classifiers used is visible in Table 4-2**.**

Table 4-2 List of LCCS levels, classifiers for the legend definition of Kalahari vegetation types. Modifiers are written in italic letters.

LCCS level	*LCCS Classifier*	*LCCS code*
Bare Areas	Hardpans	*B16*
Cultivated and Managed Terrestrial Areas	Graminoid Crops	A11
Life Form (Level 1)	Trees (Main Layer)	A3
	Shrubs (Main Layer)	A4
	Graminoid (Main layer)	A6
Coverage (Level 1)	Closed > (70-60)% (Main layer)	A10
	Open General (70-60)-(20-10)% (Main Layer)	A11
	Open (70-60) – 40% (Main Layer)	*A12*
	Very Open 40 – (10-20)% (Main Layer)	*A13*
	Sparse (20-10)-1% (Main Layer)	A14
	Scattered 4 - 1%	*A16*
	30-3m (Trees Height Main Layer)	B2
	5-0.3m (Shrubs Height Main Layer)	B3
	Medium High 3 – 0.5m (Shrubs Height Main Layer)	*B9*
	Dwarf - < 0.5m (Shrubs Height Main Layer)	*B10*
	Medium to High 5 – 0.5m (Shrubs Height Main Layer)	*B14*
	3- 0.03m (herbaceous height mean layer)	B4
	Medium Tall	*B12*
Leaf Type / Life Cycle (Level 2)	Broadleaved	D1
	Deciduous	E2
	Mixed	*E3*
Stratification (Level 3)	Single Layer	F1
	Second and/or Third layer present	F2
	Herbaceous Vegetation (Second or Third Layer)	F4
	Trees (Second or Third Layer)	F5
	Shrubs (Second or Third layer)	F6
	Closed (> 70-60%) to Open (70-60)-(20-10)% (Second or Third Layer)	F7 *F9*
	Open (70 - 60) – (20 - 10)% (Second or Third Layer)	F10
	Sparse (20-10)- 5%	G2
	> 30-3m (Tree Height Second or Third Layer)	G3
	5 - 0.3m (Shrubs Height Second or Third Layer)	*G9*
	Medium High	G4
	3-0.03m (Herbaceous Height Second or Third Layer)	*G6*
	Medium High 14 – 7m (Tree Height Second or Third Layer)	*G7*
	Low 3 - 7m (Tree Height Second or Third Layer)	*G11*
	Medium to Tall 3 – 0.3 (Herbaceous Height Second or Third Layer)	*G12*
	Short 0.3 – 0.03m(Herbaceous Height Second or Third Layer)	
Landform, Lithology/ Soils (Level 4)	Flat to almost Flat Terrain	L5
	Gently Undulating to Undulating Terrain	L6
	Hilly Terrain, Plain, Depression, Valley Floor	L8/11/13/15
	Wadi, Ridges	L16/24
	Clay, Calcareous rock	M211/230
	Soil Surface, Loose and Shifting Sands	N2/3
	Stony (5-40%), With Dunes, Hardpans, Petrocalcic	N4/6/7/9
	Arenosols, Calcisols, Fluvisols	N1105/1106/1110
climate, altitude erosion (Level 5)	Subtropics - Summer Rainfall, Dry Semi-Arid	O2/11
	1000-1500m, Water Erosion, Wind Erosion	P10, Q3/4
	Sheet Erosion	Q6
Floristic Aspect (Level 6)	*Pterocarpus angolensis - Burkea africana woodlands*	ZT01
	Terminalia sericea Combretum collinum shrub- and bushlands	ZT02
	Acacia erioloba - Terminalia sericia bushlands	ZT03
	Eragrostis rigidior - Urochloa brachyura grasslands	ZT04
	Acacia mellifera -Stipagrostis uniplumis shrublands, Typicum	ZT05
	Enneapogon desvauxii - Eriocephalus luederitzi short shrublands on calcareous omiramba & pans	ZT11
	Acacia luederitzii - Ptichtolobium biflorum floodplains of the Omatako Omuramba	ZT010
	Hyphaene petersiana plains	ZT012
	Terminalia prunioides thickets	ZT013
	Combretum imberbe - Acacia tortilis woodlands	ZT014

The specific technical attribute *Floristic Aspect* can be defined for each land-cover class, separated in single plant species (classifiers: *Dominant Species* and *Most Frequent Species*) and plant groups. Groups of plants can be distinguished between statistically derived plant groups (e.g Braun-Blanquet) or groups classified without statistical methods. If the name of the plant group is added through the user-defined attribute option the specific name is added in the user legend. Vegetation classes can also be combined with *Environmental Attributes*, e.g. *Climate*, *Landform*, *Altitude*, and *Soils* (Di Gregorio, 2005).

4.4 Scaling of land cover data

A spatial up-scaling procedure was performed to combine the *in-situ* point locations with a coarse 232-m pixel resolution of the MODIS time series metrics. The database of 422 relevés was intersected with homogeneous polygons derived from an image segmentation of Landsat data. The medium resolution of 30 m was used as an interim stage in the up-scaling process. Only MODIS pixels inside of the intersected segments were used as a training database for the image classification. The shape characteristics of the vegetation patterns in the Landsat imagery enhance the distinction of vegetation boundaries within the region-growing process.

The thematic scaling procedure comprises the translation of the local site information of the vegetation survey towards a legend according to LCCS. Based on the result of the phytosociological analysis from the vegetation and soil survey, ten synoptic vegetation type classes and two additional land-cover classes were defined as the user legend for the classification. The legend was generated following the hierarchical concept of LCCS. For each class the *in-situ* land-cover information were assigned to the corresponding classifier and added to the final legend. As displayed in Table 4-3 (showing the LCCS codes for the vegetation type classes) the class hierarchy starts with the definition of the main-, second-, and third layer's growth form (if stratification of different growth forms is present), height and coverage, followed by leaf type/ live cycle, and phenology. The related land-cover aspects of the pedological and geological context, as well as the topographical characterisation were translated from the soil survey (Strohbach *et al.*, 2004) into the Landform, Lithology/ Soils classifiers. Climate, altitude and erosion were defined in the fifth level of the classification hierarchy. The corresponding classifier for the creation of a vegetation type class is *Floristic aspect*. Since the vegetation survey is based on the Braun-Blanquet approach, the option of *Statistically derived plant groups* was chosen for the finalizing definition of the ten vegetation association class labels.

4.5 Results and discussion

4.5.1 Using LCCS classifiers: a flexible thematic legend

The translation of the vegetation type classes based on the phytosociologic plant community associations and plot descriptions resulted in a legend of multiple levels of the thematical detail. The standardised LCCS label includes the three uppermost LCCS levels (life form, coverage, leaf type/life cycle, and stratification, see Table 4-2). The predominant structural vegetation type of the Kalahari was identified as broadleaved deciduous shrubland with herbaceous understory within the main land-cover type *cultivated and managed terrestrial areas*. Coverage and height of the main layer and the subsequent layers are varying among the plant community classes.

The LCCS code (Table 4-3) reflects the typical coexistence of trees, shrubs, and grasses for an open savanna ecosystem among the ordinance of the LCCS classifiers of the vegetation type classification. The presence of a second or third layer is indicated by the F2 classifier. The growth form of that layer is defined by the F4 and F6 classifiers *Herbaceous Vegetation (Second or third Layer)* and *Shrubs (Second or third Layer)*. The resultant vegetation type classes are indicated in the specific

Table 4-3 Legend of the Kalahari vegetation types. Synoptic vegetation type classes are linked to the LCCS legend based on the LCCS code. The LCCS level indicates the classifiers used for the class definition. Note the varying s between LCCS and the classification scheme after Edwards (1983).

Synoptic Vegetation Type Legend	***LCCS Legend***	***LCCS Level***	***LCCS Code***	***Veg. classification after Edwards***
Pterocarpus angolensis - Burkea africana Savanna	Broadleaved Deciduous Woodland With Shrubs And Herbaceous Layer	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F2.F5.F10. G2-A13.B9.F9.G11.G6- L11.L5.N3.N6.N1105.O2.O11.P10.**Zt01**	21104-32174-L11L5N3N6N 1105O2O11P10Zt01	Tall moderately closed bushland
Combretum imberbe - Acacia tortilis Savanna	Broadleaved Deciduous (40 - (20 -10) %) Medium High Shrubland with Open Medium to Tall Herbaceous and Medium High Emergents	A3.A11.B2.XX.D1.E2.F2.F6.F7.G3.F2.F4.F7.G 4-A13.B7.G9.F9.G12- L11.L5.N1106.O2.O11.P10.**Zt14**	22687-L11L5N1106O2O11P10Zt14	Tall semi-open woodlands
Terminalia sericea Combretum collinum Savanna	Broadleaved Deciduous (40 - (20 -10) %) Medium To High Shrubland with Open Medium to Tall Herbaceous and Low Emergents	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F2.F5.F10. G2-A13.B14.F9.G11.G7- L11.L5.N3.N6.N1105.O2.O11. P10.Q4.**Zt02**	21104-18308- L11L5N3N6N1105O2O11P10Q4Zt02	Short moderately closed bushland
Acacia erioloba - Terminalia sericia Savanna		A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F2.F5.F10. G2-A13.B14.F9.G11.G7- L13.L5.N1106.O2.O11.P10.**Zt03**	21104-18308- L13L5N1106O2O11P10Zt03	Tall moderately closed bushland
Hyphaene petersiana Savanna	Mixed ((70 - 60) - 40%) Medium To High Shrubland with Open Short Herbaceous	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F1- A12B14E3F9G12-L11.L6.M230. N2.N4.N1106.O2.O11.P10.Q6.**Zt12**	21103-595-L11L6M230N 2N4N1106O2O11P10Q6Zt12	Short moderately closed bushlands
Acacia mellifera -Stipagrostis uniplumis Savanna	Broadleaved Deciduous ((70 - 60) – 40 %) Medium High Shrubland with Open Short Herbaceous	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F1- A13.B14.F9.G12-L11.L6.N3. N1105.O2.O11.P10.**Zt05**	21103-3712- L11L6N3N1105O2O11P10Zt05	High moderately closed shrublands
Enneapogon desvauxii - Eriocephalus luederitzi Savanna on calcareous omiramba & pans	Broadleaved Deciduous (40 - (20 -10) %) Medium High Shrubland with Open Short Herbaceous	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F1- A13.B9.F9.G12-L16.L6.M230. N2.N4.N1106.O2.O11.P10.Q3.**Zt11**	21103-4506-L16L6M230 N2N4N1106O2O11P10Q3Zt09	Tall moderately closed shrubland
Acacia luederitzii - Ptichtolobium biflorum Savanna of the Omatako Omuramba floodplains	Broadleaved Deciduous (40 - (20 -10) %) Medium High Shrubland with Open Short Herbaceous	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F1- A13.B9.F9.G12-L11.L6.N7.N9. N1106.O2.O11.P10.Q6.**Zt10**	21103-4506- L11L6N7N9N1106O2O11P10Q6Zt10	Short moderately closed bushland
Terminalia prunioides thicket Savanna	Broadleaved Deciduous (40 - (20 -10) %) Medium To High Shrubland with Open Short Herbaceous and Low Emergents	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F2.F5.F10. G2-A13.B14.F9.G12.G7- L24.L8.M211.N3.N1105. O2.O11.P10.**Zt13**	21104-18309- L24L8M211N3N1105O2O11P10Zt13	Tall moderately closed thickets
Eragrostis rigidior - Urochloa brachyura Savanna	Medium Tall Grassland with Medium High Shrubs	A6.A10.B4.XX.XX.XX.F2.F6.F10.G3-B12.G9- L15.L6.N3.N1106.O2. O11.P10.Q4.**Zt04**	20443-12292-L15L6N3N11 06O2O11P10Q4Zt04	Tall semi-open Shrubland
Graminoid Crops	Graminoid Crops	A11	10037	-
Bare Areas, Hardpans	Bare Areas, Hardpans	B16	6003	-

technical attribute *Floristic Aspect*, indicated in the final classifiers (ZT01 – ZT14). A comparative assessment of the type of the second layer shows that the majority of classes have a main shrub layer and a subsequent herbaceous layer. The remaining classes are classified as woodlands and grasslands.

The generally open shrublands (10 – 40 % coverage as defined in the A11 classifier) include the *Combretum imberbe – Acacia tortilis*, *Terminalia sericea – Combretum collinum*, *Enneapogon desvauxii – Eriocephalus luederitzii*, *Acacia luederitzii – Ptichtolobium biflorum*, and *Terminalia pruniodes vegetation associations*. Within this group the characteristics of the herbaceous layer are slightly varying and are defined in the modifier concerning the vegetation coverage of the second or third layer (F7) and the height (G2, G3, G4). The *Acacia mellifera – Stipagrostis uniplumis* savannas differ from these vegetation types with a higher coverage of the main shrub laver (40 – 70 %). The occurrence of evergreen and deciduous plants in the *Hyphaene petersiana* savanna caused a different labelling (Mixed ((70 - 60) – 40 %) Medium to High Shrubland with Open Short Herbaceous). The vegetation classification comprises one woodland class (*Pterocarpus angolensis – Burkea africana*) where trees are the dominant life form, as defined with the A4 classifier in the first level. Due to a dominating coverage of grasses compared to shrubs, the *Eragrostis ridigor – Urochloa brachyura* savannas were classified as medium to tall grasslands with medium to high shrubs.

Besides the classifiers related to the dominant and subsequent vegetation layers, a further specification of additional environmental attributes was conducted. Information of landform, lithology, and soils are key parameters for the occurrence of patterns of plant associations and can be useful information for the spectral analyses in the remote sensing imagery, as stated in section 4.5.2. Figure 4-3 shows the spatial distribution of Kalahari land-cover classes classified in a two-level scheme. Map (A) shows the main growth-forms and vegetation coverage. The subsequent vegetation type classification is shown in Map (B).

4.5.2 Spectral dependencies for class separability and mapping accuracy

The separability analysis of the time series metrics was performed to highlight the classification capabilities based on temporal and spectral feature sets for two thematic levels. Table 4-4 shows the inter-class Bhattacharyya distance (B-distance) measures between the aggregated LCCS classes (representing the main growth form and vegetation cover) and the synoptic vegetation type classes. Bare areas and pans show the best inter-class separabilities for the aggregated LCCS classes. The spectral separability increases with increasing vegetation coverage and growth form for the compared classes. Contrary values become apparent for the woodlands, where the best B-distance values were computed for differing life form classes, e.g. grasslands (37.4). The spectral separability decreases for classes with similar vegetation physiognomy and vegetation coverage. The shrubland classes with 10 to 40 % and 40 to 70 % vegetation cover achieved the lowest B-distance values (1.88). A more differentiated result is visible for the vegetation type classes. Compared to the broad structural vegetation classes the range of B-distance values is more distinct. The *Combretum imberbe - Acacia tortilis* Savanna achived the highest B-distance values with all competitive classes. The best separability was achieved against *Terminalia prunioides* Savanna (118) and *Hyphaene petersiana* plains (156). Besides contrary phenological characteristics of these classes, such exposed B-distance values can be explained by the environmental context. However, as the vegetation physiognomy of these classes is slightly differing in shrub cover, different environmental settings were defined in the LCCS class description. The *Terminalia prunioides* Savanna is occurring on hilly terrain with ridges and clayey arenosols (classifiers: L8, L24, M211, N1105). *Hyphaene petersiana* plains are characterised as plains and gently undulating to undulating terrain with the occurrence of calcareous rocks and clacisols (classifiers: L11, L6, M230, N1106). The LCCS classifier settings of the environmental attributes show that soil type and structure are differing between those classes. The pedologic

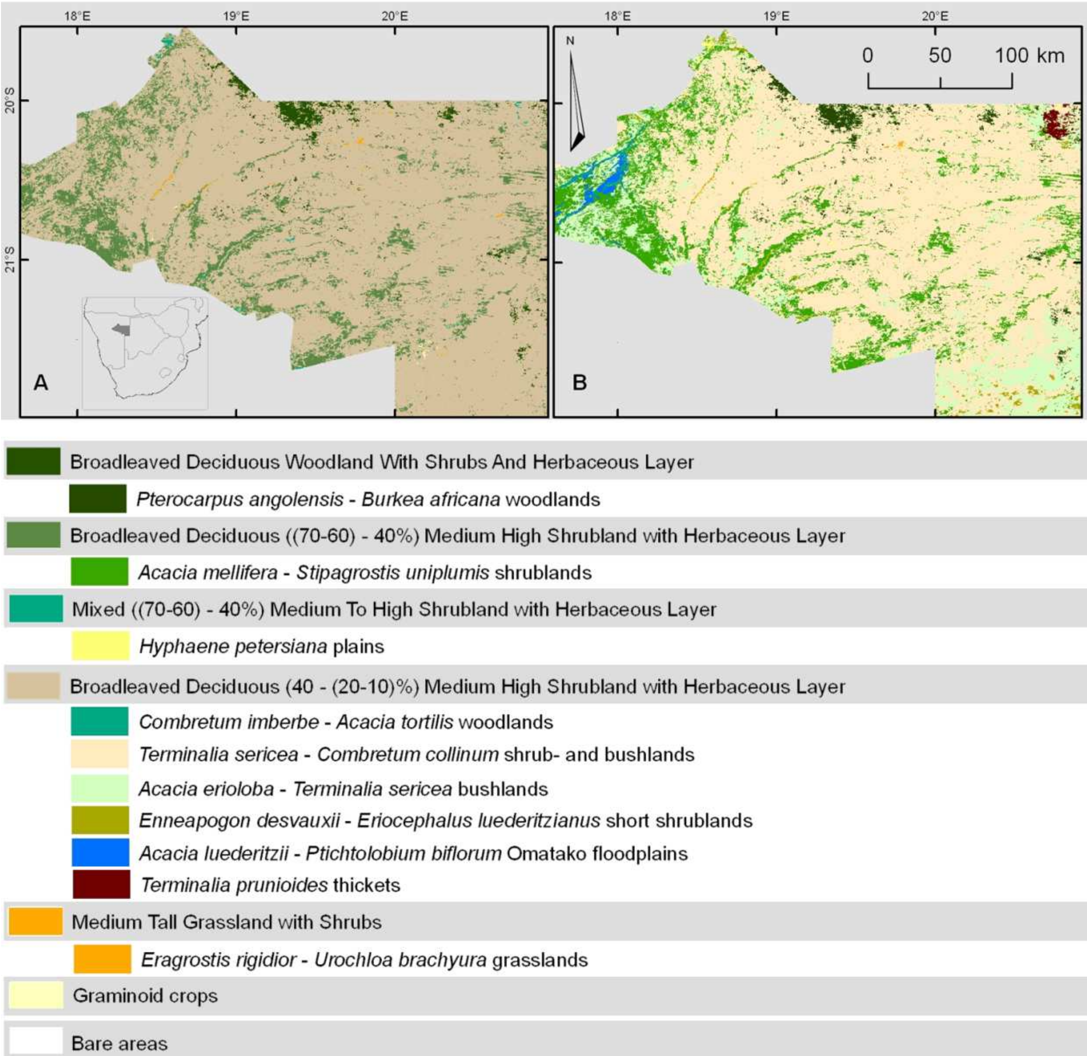


Figure 4-3 Classification of Kalahari savanna vegetation types using MODIS time series metrics of the year 2004 with a LCCS-based land-cover legend. The flexible legend comprises a broad classification of the main physiognomic-structural aspects (map A). The class label is based on the class-specific LCCS code and consists of the description of the leaf type and phenology of the main layer with the percent coverage followed by the cover and life form of the subsequent vegetation layers. The vegetation type legend consists of the phyto-sociologic plant association label as defined in the LCCS classifier *Floristic Aspect*.

situation can be a key variable for a spectral separation of land-cover types in sparsely vegetated savanna environments. The classification error assessment reflects the dependent relationship between growth form and coverage within the inter-class B-distance values. An overall mapping accuracy of 92.48 % was achieved for the land-cover map based on annual time series covering the growing season of 2004/2005, as shown in Table 4-5. Beside the class of bare areas and hardpans the best mapping accuracies achieved land-cover types with specific soil and terrain specifications in their environmental aspects, e.g. *Terminalia prunioides* thickets (95.03 %), *Combretum imberbe - Acacia tortilis* woodlands (94 %), and *Acacia luederitzii - Ptichtolobium biflorum* floodplains of the Omatako Omuramba (93.85 %). Land-cover types with an open shrub cover on deep Kalahari sands achieve slightly decreased accuracies between 83 % and 94 %. Land-cover classes with a high fractional cover of woody vegetation reached increased scores for both producer's and user's accuracy. The user's accuracy decreases for classes with a significant herbaceous layers (e.g. *Acacia mellifera - Stipagrostis uniplumis* shrublands) or main grassland classes like *Eragrostis rigidior - Urochloa brachyura* grasslands (75.66 % user's) and graminoid crops (48.46 % user's). The lowest distance values (0.88) were computed between the vegetation types, where the dominant species are slightly varying (*Terminalia sericea - Combretum collinum* shrub- and bushlands, *Acacia erioloba - Terminalia sericia* bushlands). The similar vegetation structure and environmental settings were indicated by the common set of classifiers for these classes.

Table 4-4 Confusion matrix showing average Bhattacharyya distance measures between vegetation type classes (bottom left) and the aggregated LCCS classes (upper right) calculated on MODIS intra-annual segment metrics. See Table 2 for the detailed labels of the vegetation type classes in level six of the classification hierarchy. Grey scale codes of the aggregated LCCS classes: Grasslands (light grey), Broadleaved deciduous (40 - (20 - 10) %) medium to high shrubland with herbaceous layer (medium light grey), Broadleaved Deciduous ((70 - 60) - 40%) medium to high shrubland with herbaceous layer (medium dark grey), Woodlands (dark grey), Bare Areas and Hardpans (grey lines).

		Aggregated LCCS Legend											
		Crops	ZT04	ZT03	ZT10	TZ14	ZT11	ZT13	ZT02	ZT05	ZT12	ZT01	Pans
Vegetation Type Legend	Crops		-	15.3	15.3	15.3	15.3	15.3	15.3	23.5	23.5	37.4	56.1
	ZT04	52.7		15.3	15.3	15.3	15.3	15.3	15.3	23.5	23.5	37.4	56.1
	ZT03	12.4	53		-	-	-	-	-	1.88	1.88	9.77	71.1
	ZT10	63.2	73.8	27.5		-	-	-	-	1.88	1.88	9.77	71.1
	ZT14	96.2	64.9	63.7	34.1		-	-	-	1.88	1.88	9.77	71.1
	ZT11	24.1	58.2	4.21	27.2	72.1		-	-	1.88	1.88	9.77	71.1
	ZT13	28.5	49.5	26.1	91.6	**118**	39.5		-	1.88	1.88	9.77	71.1
	ZT02	13.7	50.7	**0.88**	27.1	60.5	4.65	34		1.88	1.88	9.77	71.1
	ZT05	11.8	52.8	**1.52**	21.6	60.5	4.96	32.9	2.25		-	11.4	76.2
	ZT12	39.3	64.4	26.1	87.2	**156**	27.9	74.7	26.6	24.5		11.4	76.2
	ZT01	25	56.8	9.09	56.7	102	13.2	36.2	8.83	12.6	26.7		82.5
	Pans	29.7	44	17.8	37.7	39.2	24.8	50.3	17.8	21.9	74	37.4	

Table 4-5 Confusion matrix showing the producer's, user's, and overall accuracies of the Random Forest classification for the Kalahari vegetation type classes (see Table 2 for the detailed of the class labels).

		Out-of-bag validation samples															
		crops	ZT03	ZT10	ZT05	ZT14	ZT11	ZT12	ZT01	ZT13	ZT02	ZT04	Bare	Total	Prod's (%)	User's (%)	Overall (%)
Classification	crops	**63**	13	0	23	0	0	0	1	0	30	0	0	130	98.44	48.46	48.09
	ZT03	1	**3293**	3	85	1	3	0	2	0	92	0	0	3480	91.14	94.63	86.66
	ZT10	0	7	**748**	8	4	0	0	0	0	3	0	0	770	96.52	97.14	93.85
	ZT05	0	123	7	**3563**	5	2	3	1	1	129	0	0	3834	91.95	92.93	85.94
	ZT14	0	1	8	6	**564**	1	0	0	0	2	0	0	582	96.91	96.91	94.00
	ZT11	0	20	2	56	3	**442**	11	1	0	33	0	0	568	96.72	77.82	75.81
	ZT12	0	4	0	5	0	8	**296**	0	0	3	0	0	316	95.48	93.67	89.70
	ZT01	0	7	0	4	0	0	0	**728**	0	19	0	0	758	97.33	96.04	93.57
	ZT13	0	2	0	0	0	0	0	0	**212**	6	0	0	220	99.53	96.36	95.93
	ZT02	0	130	3	119	3	1	0	14	0	**3228**	2	0	3500	89.72	92.23	83.41
	ZT04	0	13	0	6	2	0	0	0	0	53	**230**	0	304	99.14	75.66	75.16
	Bare	0	0	0	0	0	0	0	1	0	0	0	**119**	120	100.00	99.17	99.17
	Total	64	3613	771	3875	582	457	310	748	213	3598	232	119	**14582**	**92.57**	**91.98**	**92.48**

4.5.3 Experiences using LCCS in Southern African savannas

LCCS has been successfully implemented in global land-cover mapping initiatives and provides a land-cover standard, independent of geographic scale. In this study the suitability of LCCS-based vegetation mapping in a dry semi-arid environment was tested. Regarding the *a priori* legend definition, LCCS proved to be a useful tool set to define and highlight different thematic levels of detail of the ecosystem. The classifiers approach is capable of identifying woodlands, shrublands, and grasslands as the main structural land cover types. However, the comparison of the growth-form with the classification system after Edwards (1983) shows considerable differences between the class-specific labels. As shown in Table 4-3, *Pterocarpus angolensis - Burkea africana* Savannas were assigned to the LCCS class broadleaved deciduous woodland with shrubs and herbaceous layer. After Edwards this class is labelled as tall moderately closed bushland, whereas the grasslands were classified as tall semi-open shrubland (after the Edward's scheme). Even since differences in different classification schemes are apparent, LCCS considers the layering characteristics of woody and herbaceous life forms in savanna ecosystems. The final user label is based on the combination of classifiers for a specific class. Since most of the existing LCCS applications are based on broad top-down approaches, the presented legend confirms the suitability of LCCS for standardised land-cover description in very heterogeneous landscapes.

Beyond the physiognomic characterisation of vegetation, the definition of environmental and specific technical attributes proved to be essential for the detailed characterisation of semi-arid vegetation types. The analysis of inter-class B-distance measurements indicated a relationship between life-form composition and spectral separability. For semi-arid environments the pedologic and lithologic context is a key parameter influencing the phenology and the spectral separation of plant community patterns in satellite imagery (Childs, 1989; Archibald & Scholes, 2007). The integration of the environmental classifier's settings in a spectral analysis can help to deepen the understanding of effects of surface reflectance characteristics to mapping dependencies.

The major effort of a standardised classification system is to eliminate inconsistent classification rules and to reduce the number of impractical combinations of classifiers (Di Gregorio, 2005). In order to optimise the classification system for semi-arid environments we recommend the implementation of a wider range of environmental attributes, e.g. soil texture and soil colour. Multiple selections of classifiers should be allowed for the environmental aspect section. The class boundary definition of up to three life form levels proved to be applicable. Some restrictions were recognised for the definition of subsequent layers. A more unrestrictive and thus flexible rule set could increase the classification conciseness and usefulness of the land-cover data for different user communities. Further limitations became apparent for the set of leaf type classifiers in the second modular level. The plant group of *Acaciatea* is characterised by a fine leafed physiognomy. An appropriate classifier should be added to the leaf type section.

4.6 Conclusions

In this study *in-situ* botanical field samples were integrated in a vegetation type classification on regional scales based on MODIS time series metrics. The mapping legend was defined using the FAO Land Cover Classification System. The mapping framework comprised the assessment of land-cover data on different thematic scales. MODIS time series metrics were linked with the ecological specifications of ten phytosociologic vegetation types. Capabilities and limitations of LCCS as a common land cover language were analysed for a dry semi-arid ecosystem. The results could point out distinct benefits for the remote sensing community related to large-area applications, but also for research communities, working on a local plot scale, related to biodiversity and landscape ecology. Detailed plot descriptions could be extrapolated on a regional scale by integrating coarse earth observation data.

By using a flexible classification scheme, the mapping capabilities could be analysed on different thematic information levels. The modular and hierarchical structure of LCCS therefore allowed an objective assessment of phenological and spectral inter-class relationships under consideration of additional environmental features influencing the land-cover. Integrated assessments of ecological data are one of the key challenges in interdisciplinary research to deepen the understanding of ecosystems functioning.

The application of LCCS as a common land-cover language in ecosystem assessment studies can be an important step towards an effective communication between different scientific communities (e.g. remote sensing, biodiversity, and landscape ecology) and space agencies. Land-cover datasets based on a standardised classification scheme are baseline information to set up a land-cover and biodiversity monitoring structure and to assess the state of ecosystems on a global level. In conclusion, LCCS can play a key role as a land cover language to set up interdisciplinary research structures in terms of standardisation and communication in the framework of CBD and GEO-BON.

The results on the general applicability of the whole set of classifiers for the major land-cover type *Natural and semi-natural vegetation* highlighted the existing limitations of LCCS version two for Southern African savanna ecosystems. A detailed bottom-up assessment of the classifier rule set for the world's major bioregions is highly recommended for the major land-cover types. Especially LCCS classifiers for fine scattered or dynamic landscapes should be focussed for validation initiatives based on *in-situ* measurements regarding the international standardisation process through the ISO TC211.

The legend definition indicated a predominance of open shrublands in the Kalahari of north-eastern Namibia. Several implications arise for existing global land cover maps where open grasslands are the main land cover classes detected in that region. Further research has to be conducted to condense the validation database and calibration process and to improve classifications the for the savanna biomes in the upcoming global mapping initiatives.

5 Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems

Implications for Global Monitoring[3]

Abstract – Land-cover is an important parameter in analysing the state and dynamics of natural and anthropogenic terrestrial ecosystems. Land-cover classes related to semi-arid savannas currently exhibit among the greatest uncertainties in available global land cover datasets. This study focuses on the Kalahari in northeastern Namibia and compares the effects of different composite lengths and observation periods with class-wise mapping accuracies derived from multi-temporal MODIS time series classifications to better understand and overcome quality gaps in mapping semi-arid land-cover types. We further assess the effects of precipitation patterns on mapping accuracy using Tropical Rainfall Measuring Mission (TRMM) observation data. Botanical field samples, translated into the UN Land Cover Classification System (LCCS), were used for training and validation. Different sets of composites (16-day to three-monthly) were generated from MODIS (MOD13Q1) data covering the sample period from 2004 to 2007. Land-cover classifications were performed cumulatively based on annual and inter-annual feature sets with the use of random forests. Woody vegetation proved to be more stable in terms of omission and commission errors compared to herbaceous vegetation types. Generally, mapping accuracy increases with increasing length of the observation period. Analyses of variance (ANOVA) verified that inter-annual classifications significantly improved class-wise mapping accuracies, and confirmed that monthly composites achieved the best accuracy scores for both annual and inter-annual classifications. Correlation analyses using piecewise linear models affirmed positive correlations between cumulative mapping accuracy and rainfall and indicated an influence of seasonality and environmental cues on the mapping accuracies. The consideration of the inter-seasonal variability of vegetation activity and phenology cycling in the classification process further increases the overall classification performance of savanna classes in large-area land-cover datasets. Implications for global monitoring frameworks are discussed based on a conceptual model of the relationship between observation period and accuracy.

[3] Submitted as: Hüttich, C., Herold, M., Wegmann, M., Cord, A., Strohbach, B., Schmullius, C., & S., Dech (n.d.): Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for Global Monitoring. *Remote Sensing of Environment*. (in review).

5.1 Introduction

Land-cover is one of the most important environmental parameters. Observations on the state and the dynamics of land cover are basic indicators in assessing land degradation, biodiversity, nature conservation, and food security. Vegetation maps offer basic information for environmental management and modelling (Dougill & Trodd, 1999) and land planning (Herold, 2009). Global characterisations of the main vegetation types have been achieved within different remote sensing based mapping initiatives, such as the 1-km datasets IGBP DISCover (Loveland *et al.*, 2000), the MODIS land cover product (Friedl *et al.*, 2002), the UMD land cover product (Hansen *et al.*, 2000), and GLC2000 (Mayaux *et al.*, 2004). Improvements towards a better spatial resolution were delivered by the GlobCover initiative with a global 300-m land cover product for the years 2005-2006 (Defourny *et al.*, 2009) and the MODIS land cover product with a spatial resolution of 500-m, delivered by the MODIS land cover team (Friedl *et al.*, 2010) on an annual basis. Different project histories and technical restrictions led to major restrictions regarding the input datasets, different levels of thematic information and different classification methods (bottom-up vs. top-down approaches). However, recent initiatives on the harmonisation and standardisation of existing global land cover datasets fostered an increased interoperability of land cover information among the 'global' research communities, such as remote sensing and carbon modelling but also emphasise inherent strengths and weaknesses among the coarse scale land-cover datasets (Herold *et al.*, 2008; Jung *et al.*, 2006). International initiatives such as the Global Earth Observation System of Systems (GEOSS) established by the Group of Earth Observations (GEO, Muchoney, 2008; Scholes *et al.*, 2008) aim at supporting earth observations and scientific exchange for assessing biodiversity and the status of ecosystems on various scales. The standardisation of the scientific exchange of environmental data accrues several advantages for the user community. We present here an example by linking *in-situ* land-cover data with coarse scale satellite time series to assess the capabilities of land-cover mapping in semi-arid environments in the framework of global land-cover mapping approaches.

Harmonisation and standardisation initiatives have recently advanced the applicability of land-cover data. The Land Cover Classification Scheme (LCCS, Di Gregorio, 2005) of the Food and Agriculture Organisation (FAO) emerged as the most common accepted standard for land-cover (Herold *et al.* 2009) and has been submitted to the International Organisation for Standardisation (ISO) to evolve better international standardisation (TC211, GLCN, 2009). The independent diagnostic criteria of LCCS characterizing vegetation physiognomy and land-cover (classifiers) help to overcome existing limitations of class definitions and related class boundary overlaps (Jansen & Di Gregorio, 2002), as analysed by Thompson (1996) for Southern African vegetation types. Despite inter-comparisons of existing land-cover data, further research has to be addressed to link *in-situ*- and satellite earth observation data. The knowledge of environmental cues and internal technical limitations (e.g. composite method and length) of coarse scale land-cover mapping approaches in semi-arid savanna ecosystems is still insufficient. A number of research initiatives focussed on characterizing the Southern African savannas by integrating multi-scale environmental data, e.g. for ground vegetation surveys (Strohbach 2001) and the integration of *in-situ* data for vegetation mapping using remote sensing techniques in the framework of the ACACIA and BIOTA projects in Namibia and South Africa (Burke & Strohbach, 2000; Hüttich *et al.*, 2009; Keil *et al.*, 2010), or calibration and validation initiatives of remotely sensed biophysical parameters in the framework of SAFARI 2000 (Scholes *et al.*, 2004; Privette *et al.*, 2004; Scholes *et al.*, 2002).

5.1.1 Compositing approaches and periods in global land-cover datasets

Compositing has been widely used in order to generate cloud-free spatially consistent images from satellite time series and to correct negative effects of differing viewing zenith angles on reflectance

(Cihlar *et al.*, 1994; Roujean *et al.*, 1992). Differences are evident in terms of the compositing method and compositing period which is strongly related to the technical design of the sensor system. Three main compositing methods were applied on satellite time series data. The most common method is the maximum value composite (MVC) approach (Holben, 1986), where the maximum value of the normalised difference vegetation index (NDVI) is selected over a particular composite period to minimise negatively biased atmospheric effects. However, this method is susceptible to anisotropic vegetation structure and unstable atmospheric conditions (van Leeuwen, 1999). A more advanced approach is the best index slope extraction (Viovy *et al.*, 1992) which accounts for a realistic representation of the time series by retaining more valuable time steps than MVC. Data averaging on a per-pixel basis in each spectral band is used in the mean compositing method and used to be the best method for synthesizing the ENVISAT MERIS time series in the framework of the GlobCover project (Vancutsem *et al.*, 2007a; Vancutsem *et al.*, 2007b). The Hagolle algorithm is based on the bidirectional compositing algorithm (Duchemin *et al.*, 2002) and fits a BRDF model to the reflectance data (Hagolle *et al.*, 2005).

Table 5-1 Summary of some global land-cover products and the related composite periods of the input features for land-cover classification compared to the classification accuracies of semi-arid land-cover classes.

Dataset description	***Composite period***	***Accuracy of classes related to semi-arid ecosystems***	
		Label	***Acc. (%)***
IGBP DISCover, 1 km, AVHRR, 1992-1993 (T. R. Loveland et al., 2000)	Maximum value compositing of 10-day composites into monthly composites	Closed shrublands	65.0*
		Open shrublands	78.0*
		Woody savannas	58.0*
		Savannas	42.0*
		Grasslands	58.0*
		Overall accuracy, (Scepan, 1999)	66.9
MODIS land cover product,1 km, MODIS, 2000 (M. A. Friedl et al., 2002)	Nadir BRDF-adjusted 16-day composites	Closed Shrubland	46.2*
		Open Shrubland	46.8*
		Woody savanna	37.5*
		Savanna	39.1*
		Grasslands	55.3*
		Overall accuracy,(Hodges, 2002)	**71.6**
MODIS land cover product collection 5 (MCD12Q1), 500 m, MODIS, 2001-2008 (Friedl *et al.*, 2010)	Monthly Averaged composites from 8-day composites	Open shrublands	47.0*
		Closed shrublands	74.1*
		Woody savannas	34.0*
		Savannas	39.0*
		Grasslands	55.9*
		Overall accuracy (2005 classification, (Friedl *et al.*, 2010))	**75.8**
UMD land cover product, 1 km, AVHRR, 1992-1993, (Hansen *et al.*, 2000)	Maximum value compositing of 10-day composites into monthly composites	No validation information available	
GLC2000, 1 km, 2000, SPOT VEGETATION (Bartholomé & Belward, 2005)	Monthly to three-monthly composites based on statistical averaging of daily mosaics	Shrub Cover, closed-open	47.3*
		Sparse Herbaceous or sparse shrub cover	50.5*
		Herbaceous Cover, closed-open	33.4*
		Overall accuracy, (Mayaux *et al.*, 2006)	**68.6**
GlobCover version 2, 300 m, ENVISAT MERIS, 2005-2006 (Vancutsem *et al.*, 2007; Defourny *et al.*, 2009)	Bi-monthly composites based on mean compositing of 10-day composites	Mosaic forest or shrubland (50-70%) / grassland (20-50%)	51.0*
		Mosaic grassland (50-70%) / forest or shrubland(20-50%)	65.2*
		Closed to open (>15%) (broadleaved or needle leaved, evergreen or deciduous) shrubland (<5m)	40.6*
		Closed to open (>15%) herbaceous vegetation	16.4*
		Sparse (<15%) vegetation	10.9*
		Overall accuracy, (Bicheron et al., 2008)	**67.1**

*User's accuracy

The composite periods and methods differ between the global land cover mapping approaches, as they are imbedded in different project structures with differing thematic issues and technical constraints of the satellite sensor. As listed in Table 5-1 the compositing periods range from 16-day composites to three-monthly composites. Most of the global maps are based on monthly composites, such as DISCover dataset (Loveland *et al.*, 2000), the University of Maryland land-cover map (UMD, Hansen *et al.* 2000), and the MODIS land cover product collection 5 (Friedl *et al.*, 2010).

The original 10-day composites (AVHRR) and 8-day composites (MODIS) were re-composited to monthly composites to reduce data volume while keeping the phenological information, as shown in earlier studies (Townshend & Justice, 1986; Tucker *et al.*, 1985). A regional adaption of the composite period (monthly to three-monthly composites) was done in producing the GLC2000 dataset (Bartholomé & Belward, 2005). For the African subset of GLC2000, monthly composites were generated based on a physical and a statistical approach (Fritz *et al.*, 2004; Vancutsem *et al.*, 2007b). Besides averaging the most applied method of compositing is MVC, as shown in Table 5-1.

5.1.2 Spatiotemporal variability in mapping savannas at global scales

Savanna ecosystems are one of the most spatiotemporally heterogeneous biomes. They are characterised as landscapes with co-existing vegetation layers of herbaceous vegetation and woody plant communities (trees and shrubs), determined by precipitation rates, fire, and grazing processes (Frost, 1996; Sankaran *et al.*, 2005; Privette *et al.*, 2004; Scholes *et al.*, 2002). This might be a key factor why land-cover classes related to savanna ecosystems have moderate map accuracies (Herold *et al.*, 2008; Jung *et al.*, 2006) and thus show increased inconsistencies when comparing the global land-cover maps. Besides differences in the classification schemes, the comparison of the Global Land Cover 2000 (GLC2000) and MODIS land cover datasets (Giri *et al.*, 2005) show increased areal disagreements for *woody savannas* (141.16 %), *open shrublands* (73.76 %), and *savannas* (13.95 %). Further, very high areal disagreements were revealed for *grasslands* and *shrubs* by comparing the IGBP DISCover and UMD products (Hansen & Reed, 2000). Furthermore, the comparison of present inter-annual land-cover databases, such as the MODIS land-cover product collection 5 (MCD12Q1, Friedl *et al.*, 2010) highlights inconsistencies in terms of the spatial distribution of the class assignments between different years. Figure 5-1 shows the spatial variability of land-cover types mapped for three years over the *Kalahari* and *Thornbush-* and *Mountain savanna* transition in NE Namibia.

This example of the spatiotemporal variability in mapping plant functional types (grasses and shrubs) clearly indicates the impacts of certain environmental factors on the resulting land-cover map. Especially the inter-annual variability of precipitation rates (shown with the areal statistics of annual precipitation sums for the presented area in Figure 5-1d) might be a key factor for such differing mapping results. In fact, the temporal dependencies in mapping semi-arid ecosystems and the implications for global scale land-cover mapping approaches are still insufficiently analysed. This case study analyses the effects of the factors (1) composite period and (2) observation period on the mapping accuracy. Observation period is defined as the period in which MODIS time series features are included in the classification.

5.1.3 Aims and objectives

Semi-arid land-cover classes at global scales are still insufficiently mapped and their accuracies differ strongly between the datasets (Table 5-1). The availability of *in-situ* observations plays a critical role for satellite-based land-cover mapping as they provide reliable land cover information (e.g. land-cover type and characteristics, Herold, 2009). The integration of ground-based measurements is a forthcoming task for future global land-cover monitoring initiatives, e.g. in the framework of

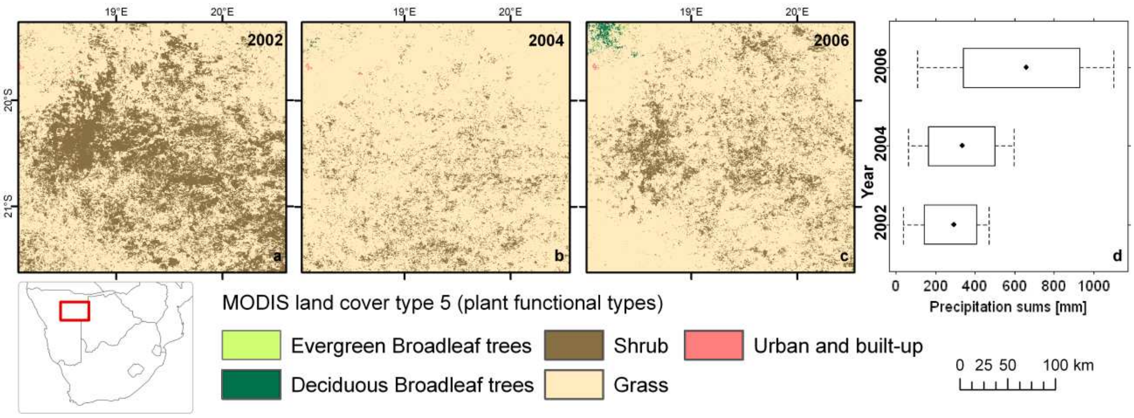


Figure 5-1 Comparison of the plant functional type land-cover maps (a-c) derived from the MODIS MOD12Q1 global land-cover database (Friedl *et al.*, 2010) for the years 2002, 2004, and 2006. The areal precipitation sums (d) of the presented spatial extend, derived from data of the Tropical Rainfall Measuring Mission Satellite (TRMM, Simpson *et al.*, 1996), indicate the linkage of the inter-annual variability of the rainfall rates and spatial variations of class assignments in the Southern African savannas.

GEOSS and the Committee of Earth Observation Satellite's (CEOS) Land Product Validation Subgroup (LPC). During the field campaigns in the framework of the BIOTA-Southern Africa project (Strohbach & Jürgens, 2010; Keil *et al.*, 2010) a number of standardised ground measurements was collected. Up-scaling analyses of *in-situ* vegetation data form plot-level to a 'coarse' MODIS scale were performed for large-area mapping of Kalahari vegetation types (Hüttich *et al.*, 2009) and fractional vegetation cover (Gessner *et al.*, 2009). Studying spatiotemporal patterns of land cover types in Southern African savannas across scales affected positive feedbacks for the involved research communities in order to:

a) encourage and establish more standardised monitoring schemes based on LCCS, since definitions of savanna vegetation is scale-dependent and still differing among the research communities (data harmonisation).

b) improve large-area mapping through spatial up-scaling (*in-situ* to landscape level). The integration of in-situ data in large-area mapping initiatives affected an enhanced thematic resolution.

c) enhance plant community characterisations through temporal up-scaling. Linking satellite time series to periodically acquired vegetation survey data enhanced the characterisation of vegetation dynamics and phenological patterns.

The study here aims to integrate *in-situ* land-cover data from the BIOTA network within a coarse scale land-cover mapping framework, based on different temporal settings of satellite time series sets. The following research questions are addressed in order to provide a more comprehensive understanding of the temporal effects for mapping semi-arid vegetation cover types based on multi-temporal remote sensing information:

- Which effect do temporal compositing and variations of the observation period have on the land cover classification accuracy in semi-arid areas?

- What are the effects of inter-annual rainfall variability on the land-cover classification accuracies in rainfall-driven savannas?

- How can the remotely sensed feature assembly be adapted to consider inter- and intra-annual variations, affecting spatiotemporal patterns of the vegetation activity and phenological cycling?

In order to understand and overcome the existing quality gaps in semi-arid land-cover maps derived from satellite time series observations, this study is methodologically aiming to:

- Analyse effects of temporal compositing on classification accuracies indicating (inter-) annual temporal requirements and environmental constraints in mapping semi-arid cover types.

- Analyse the effect of varying observation periods on land-cover classification accuracies considering the inter-annual rainfall variability.

- Evaluate the implications for global monitoring initiatives in order reveal options for improving and adapting the mapping design on the actual ecosystem-specific requirements.

5.2 Data and methods

5.2.1 Study area

The study area comprises the eastern communal areas in the Namibian eastern Kalahari with a geographic extent from 17°30'E to 21°00'E and 19°45'S to 21°45'S (Figure 5-2) covering different vegetation types after Giess (1971), such as the *Kalahari*, *Karstveld*, and the *Thornbush Shrubland*. Most parts of the investigation area are covered by deep Kalahari sands. Typical for the Kalahari sand plateau, the transition between the Central Namibian Highlands and the Kalahari Basin is characterised by the SW-NE oriented shallow water courses with no visible gradients or visible water course (so-called *Omirimba*, Figure 5-2b, King, 1963). The climate is sub-continental, dominated by highly variable summer rainfall and has an annual precipitation of 350 - 450 mm in the long-term mean (Figure 5-2c).

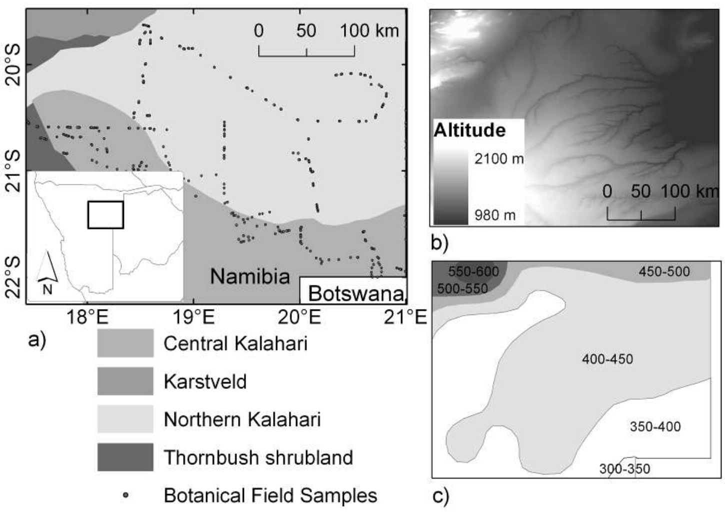


Figure 5-2 Overview of the study region showing (a) the main vegetation types after Giess (1971), overlain with the distribution of *in-situ* data, assessed in the framework of the Desert Margin Project (Strohbach *et al.*, 2004), (b) the topography (SRTM, Farr *et al.*, 2007) and (c) the spatial distribution of the mean annual precipitation (Mendelsohn *et al.*, 2002).

5.2.2 Field data processing

A reconnaissance survey of the landscapes, soils, and vegetation has been carried out during the growing season of the year 2004, providing a detailed description of land types, vegetation composition and physiognomy and habitat settings (Strohbach *et al.*, 2004). A stratified sampling with a standardised plot size of 20 x 50 m following the Braun-Blanquet approach was applied, as described in Strohbach (2001). A phythosociologic analysis was performed resulting in 12 main vegetation types according to the dominant species occurrence. This synoptic vegetation legend was translated into a classification nomenclature according to LCCS (Table 5-2). The result was a classification scheme based on the physiognomic-structural aspects, such as (a) physiognomy; (b) vertical and horizontal arrangement; (c) leaf type; and (d) leaf phenology (Di Gregorio, 2005). Five land-cover classes were derived within the LCCS category of major land-cover types *Natural and semi-natural vegetation* (LCCS level A12). Beside the *medium tall grassland with medium high shrubs* and *broadleaved deciduous woodland with shrubs and herbaceous layer*, the *broadleaved deciduous medium high shrublands with an open short herbaceous layer* build the largest cover type, grouped into two coverage types (open and very open, see Table 5-2). As the legend definition in LCCS uses independent diagnostic criteria, the layering system of the vegetation structure is taken into account, and so, the highest degree of thematic information of the vegetation structure is included in a single legend. Beside the general advantages of flexibility and transparency, the main advantage of the layering concept becomes important if analysing highly heterogenic landscapes. Another main advantage is the (most possible) independence of geographic scale. The floristic legend of the *in-situ* database used is translated on a broad physiognomic thematic level that considers the coarse spatial resolution of the satellite time series data. Doing so, the final land-cover types are classified in layered vegetation classes, including at least two life forms, described in the layer specifications in Table 5-2. The multi-layer concept of LCCS therefore accounts for the representation of the typical coexistence of trees, shrubs, and grasses in savanna ecosystems. A detailed description of the legend translation and vegetation distribution of the study area is given in Strohbach *et al.* (2004) and Hüttich *et al.* (2009, 2010).

5.2.3 Satellite data

MODIS data - Time series of the MODIS MOD13Q1 collection five product at the original 232 m resolution (250 m is used by convention) in sinusoidal projection were utilised as input for the land-cover classification. The original 16-day composite images of the blue (459-479 nm), red (620-670 nm), near infrared (NIR, 841-876 nm) and middle infrared (MIR, 1,230-1,250 nm) reflectance bands, as well as the enhanced vegetation index (EVI, Huete *et al.*, 2002) were analysed for three growing seasons from September 2004 to October 2007. Due to cloud effects and differing sun zenith angles the pre-processing included the identification and removal of low quality pixels based on the MODIS quality assessment science datasets (Justice *et al.*, 2002). Using the Time Series Generator software (TiSeG, Colditz *et al.*, 2008) data gaps were filled by linear interpolation. As further refinement an adaptive Savitzky-Golay filter was applied on each band of time series observations using the TIMESAT software package (Jönsson & Eklundh, 2004) to generate meaningful time series for a most realistic representation of the vegetation phenology. The main advantage becomes visible in Figure 5-3 by comparing the original EVI time series with the Savitzky-Golay filterd data. The preservation of the mean position of the seasonal peak and the adaption to the upper envelope of the data is particularly useful in semi-arid and rainfall-driven ecosystems with high temporal variations of the phenology patterns.

TRMM data - Time series from the Tropical Rainfall Measuring Mission (TRMM, Simpson *et al.*, 1996) were used as auxiliary datasets for the correlation analyses of the temporal dynamics of annual and inter-annual rainfall patterns with cumulative mapping accuracies. Monthly aggregated TRMM observations with an original resolution of $0.25° \times 0.25°$ were resampled to the 250 m

MODIS resolution. Monthly precipitation sums covering the sample period (Sep. 2004 – Oct. 2007) were used to compute spatial statistics of the study region (Figure 5-4) indicating the spatial mean, minimum, maximum and range of monthly precipitation rates over the study area.

5.2.4 Composite period sets

Four sets of varying composite periods were generated to explore the dependency of the length of the composite period on the classification performance. Following the common composite lengths used for the main global land-cover classifications (Table 5-1), the classification sets were subdivided in 16-day, monthly, bi-monthly, and three-monthly composites. Averaging and MVC are the common compositing approaches in the global community. This study followed the statistic approach having referring to the MVC result. The aggregated value for each composite period was calculated per-pixel for the blue, red, NIR, MIR, and EVI bands by choosing the maximum value of the given composite period. This reduced the number of input features for the classifier with increasing composite period (Table 5-3).

Table 5-2 Classification scheme according to LCCS, based on a legend translation of floristic vegetation types (Strohbach *et al.*, 2004). Note the vegetation structure description for each vegetation layer derived from the LCCS classifier and modifier settings, shown in the LCCS level (A6=graminoids, A4=shrubs, A3=trees, see Di Gregorio (2005) for a full description of classifiers).

LCCS class name	*Vegetation structure*	*LCCS level*	*Sample size [pixel]*
Medium Tall Grassland with Medium High Shrubs	Main layer specifications: ▪ life form: graminoids ▪ coverage: closed > (70-60)% ▪ height: medium tall (3 - 0.03m) Second layer specifications: ▪ life form: shrubs ▪ coverage: sparse (20-10) - 5% ▪ height: medium high (5 - 0.3m)	A6.A10.B4.XX.XX.XX.F2.F6.F10.G3-B12.G9	434
Broadleaved Deciduous ((70-60) - 40%) Medium High Shrubland with Open Short Herbaceous	Main layer specifications: ▪ life form: shrubs ▪ coverage: open (70-60) - (20-10)% ▪ height: medium high (5 - 0.3m) Second layer specifications: ▪ life form: herbaceous vegetation ▪ coverage: closed (> 70-60%) to open (70-60) - (20-10)% ▪ height: short (3 - 0.03m)	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F1-A13.B14.F9.G12	2316
Broadleaved Deciduous (40 - (20-10)%) Medium To High Shrubland with Open Short Herbaceous	Main layer specifications: ▪ life form: shrubs ▪ coverage: very open (40 - (20-10)% ▪ height: medium to high (5 - 0.3m) Second layer specifications: ▪ life form: herbaceous vegetation ▪ coverage: closed (> 70-60%) to open (70-60) - (20-10)% ▪ height: short (3 - 0.03m)	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F2.F5.F10.G2-A13.B14.F9.G11.G7	6140
Broadleaved Deciduous Woodland With Shrubs And Herbaceous Layer	Main layer specifications: ▪ life form: trees ▪ coverage: open general (70-60) - (20-10)% ▪ height: 30 - 3m Second layer specifications: ▪ life form: shrubs ▪ coverage: closed (> 70-60%) to open (70-60) - (20-10)% ▪ height: 5 - 0.3m Third layer specifications: ▪ life form: herbaceous vegetation ▪ coverage: closed (> 70-60%) to open (70-60) - (20-10)% ▪ height: 3 - 0.03m	A4.A11.B3.XX.D1.E2.F2.F4.F7.G4.F2.F5.F10.G2-A13.B9.F9.G11.G6	758
Bare areas and pans	Main layer specifications: bare areas and hardpans	B16	120

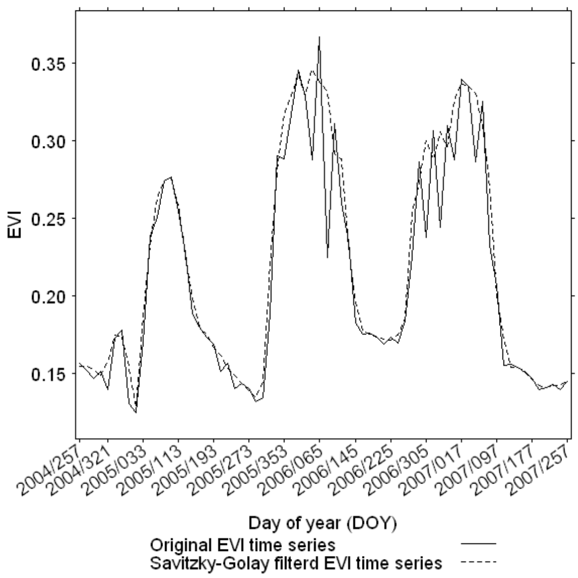


Figure 5-3 Exemplary plot of three growing seasons of EVI time series (Sep. 2004 – Oct. 2007) showing the original EVI values compared to the enhanced Savitzky-Golay filtered EVI time series.

Table 5-3 Number of composite image features in the different feature sets, annual and inter-annual for the sample period 2004-2007.

Feature set	*Number of features*
Annual 16-day	115-120
Annual Monthly	60
Annual Bi-monthly	30
Annual Three-monthly	20
Inter-annual 16-day	350
Inter-annual monthly	175
Inter-annual bi-monthly	90
Inter-annual three-monthly	60

5.2.5 Land-cover classification

The land-cover classifications were performed on an annual basis covering the growing seasons from 2004 – 2007 to analyse the effects of the inter-annual variability on the mapping accuracy. The growing season period was defined to last from September to October to consider the typical summer rainfall regime. Furthermore, inter-annual classifications were performed covering the composite images over the entire sample period. Within the grouping of annual and inter-annual feature sets, the classifications were further subdivided into different groups following the four groups from 16-day to three-monthly composite periods. The input feature sets comprised the temporal composites of the smoothed MOD13Q1 bands. The annual and inter-annual classifications were cumulatively classified which enables further analyses of the temporal response of classification accuracies compared to precipitation patterns and the dependencies of differing observation periods.

Non-parametric tree-based classification algorithms have proven to be effective for land-cover classifications using high dimensional satellite imagery, such as for mapping cover types globally based on coarse scale satellite time series (Brodley & Friedl, 1997; DeFries *et al.*, 1995; Hansen *et al.*, 2000) or for local scale ecotype mapping purposes using hyperspectral imagery (Chan & Paelinckx, 2008). The overall performance of these decision tree (DT) based approaches was optimised in terms of robustness using ensembles of DT in combination with bagging, e.g. Random Forests (Breiman, 2001) and boosting (Pal, 2005; Chan & Paelinckx, 2008; Watts & Lawrence, 2008). In this study the random forest classifier was applied for the classification of the time series feature sets using the random forest package implemented in the R statistics language (Liaw & Wiener, 2002).

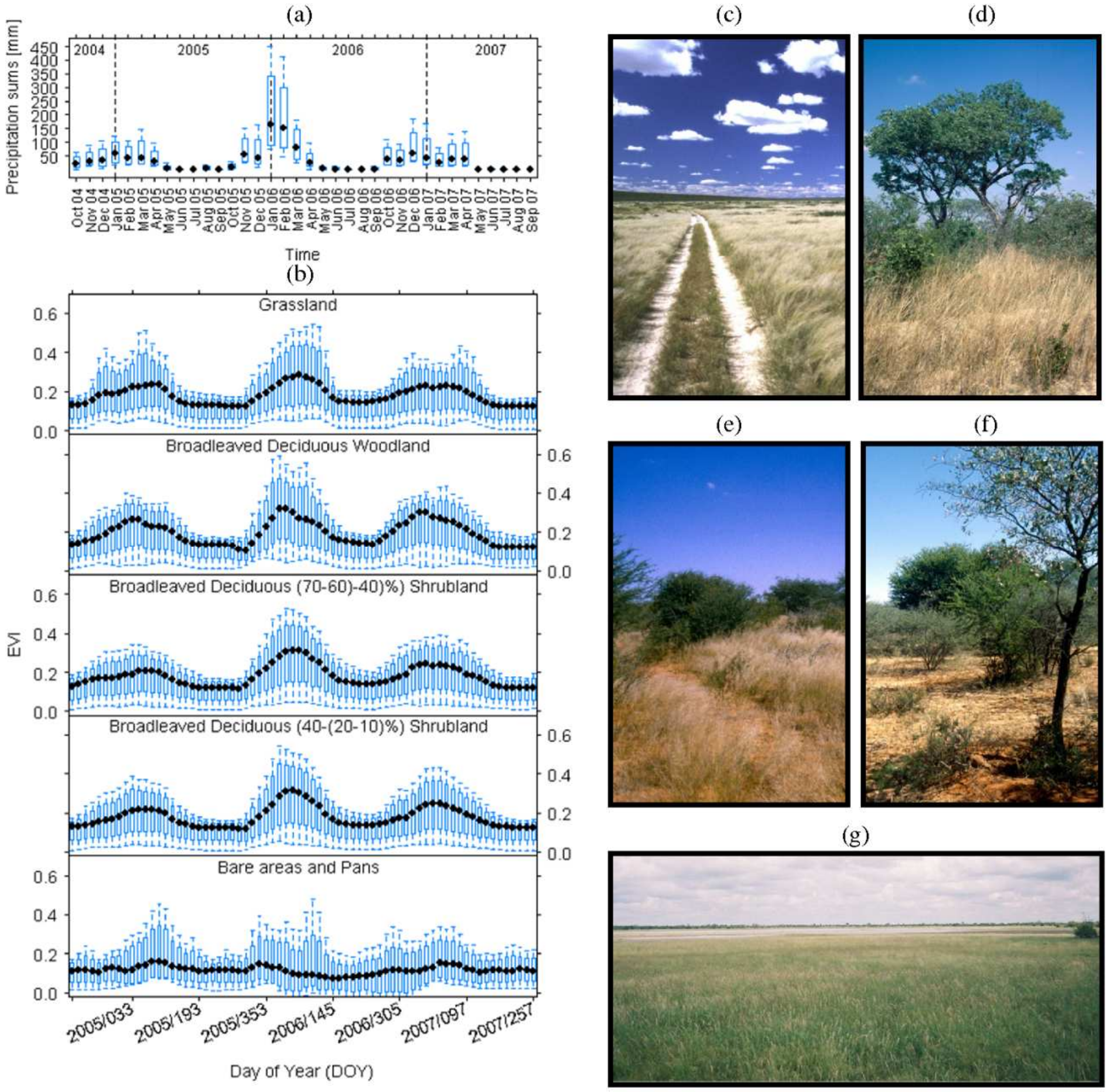


Figure 5-4 Distinct spatiotemporal variations of the rainfall patterns in the study region for the observation period, indicated by (a) monthly TRMM observations covering the growing seasons from 2004/05 to 2006/07. Similar pattern are visible in (b) the training site statistics of the EVI time series of the land cover classes grassland (c), woodland (d), open shrubland (e), very open shrubland (f), and bare areas and pans (g).

Within the methodological framework of random forest multiple trees were produced and combined based on majority voting. For the training of each particular tree, one third of the training data, the so-called out-of-bag (OOB) sample, is excluded and used for validation. An overall OOB prediction error is given using the majority vote of all classifications within the tree ensemble. Several studies showed that using the OOB prediction error provides a more conservative classification error assessment than using an independent test set for the validation

(Breiman, 2001; Gislason *et al.*, 2006). The OOB error is used here for the accuracy analyses. A detailed description of random forests is given in Breiman (2001) and Hastie *et al.* (2009). Kappa statistics and the class-wise mapping accuracy were computed using the OOB error confusion matrixes.

Mapping accuracy represents the number of correctly identified pixels of the mapping area in relation to the omission and commission errors, as shown in the equation as follows. *MA* is the class-wise mapping accuracy derived from the OOB error matrix (*j* = columns, *i* = rows), where n_{ij} is the number of correctly classified pixels, n_{i+} the number of samples classified into category *i* in the classification, n_{+j} the number of samples classified into category *j* in the reference data set (Congalton & Green, 2009).

$$MA = \frac{n_{ij}}{n_{ij} + n_{i+} + n_{+j}}$$

5.2.6 Statistical framework

Annual and inter-annual time series of the cumulatively generated classification accuracies were tested to prove the significance of existing differences among different groups in the cumulative classification runs, such as (a) observation period, (b) composite periods, and (c) land-cover classes. A stratified random sampling (N=500 per land-cover class) was performed for all land-cover classes on each single cumulative classification result and the monthly cumulative TRMM datasets. Different runs of analyses of variance (One-way-ANOVA) were applied on the class-specific cumulative mapping accuracy outputs of the experimental design. Following the ANOVA framework, the F statistic was used as the test of significance and to prove if the null hypothesis (H_0) can be rejected for the following cases:

a) H_0 = no difference between annual or inter-annual classifications, no effect of the length of the observation period on the classification accuracy.

b) H_0 = no effect of composite periods on the classification accuracy, no measured difference among the classification accuracies if 16-day, monthly, bi-monthly, or three-monthly composites are used for annual and inter-annual classifications.

c) H_0 = no effect of observation period and composite length on class-specific accuracies, no difference among the class-specific accuracies comparing 16-day to three-monthly composites for annual and inter-annual classifications.

Based on the ANOVA results, multiple mean comparisons were applied to identify the magnitude, direction, and significance of the differences among the groups. The T statistic was used to test the significance if the group means significantly differed compared to a reference group which had been defined for each ANOVA run.

The functional responses of cumulative mapping accuracies to rainfall were analysed using linear regression models. A piecewise linear model was applied to correlate cumulative mapping accuracies with cumulative monthly rainfall. Such 'broken-stick' models are useful in modelling abrupt ecological and climatic thresholds, defined as a substantial change in a response variable. Thresholds are indicated at unknown points, so-called 'breakpoints' that join the two straight lines *sharply* in the model (Toms & Lesperance, 2003). Breakpoints were derived to assess the functional responses of classes-specific accuracies to cumulative precipitation. The sharp piecewise linear regression was applied on a 95% confidence level using the R package *Significant Zero Crossings* (SiZer, Sonderegger *et al.*, 2009).

5.3 Results

5.3.1 Effects of observation period classification accuracy

The classification accuracy increases with increasing length of the observation period, as shown by the distribution of the Kappa coefficients (Figure 5-5) of the cumulative multi-annual classifications. Except for the land-cover types related to *bare areas and pans*, the vegetated classes *grasslands*, *shrublands* and *woodlands* show increasing mapping accuracies with increasing observation period. The cumulative accuracy changes can be subdivided into three dynamic phases: an initial increasing phase, a phase of levelling where the accuracy scores remain stable, and a fluctuating phase. As shown in the graphs of the cumulative annual classifications (Figure 5-6), a gain of accuracy is evident in the beginning observation period between September and December. The functional response of increasing mapping accuracies depends further on land-cover type and precipitation amount during the corresponding year. Class-specific differences become apparent in the starting level and slope, fluctuation of the accuracy gain, and the final accuracy score. Land-cover types with a higher fraction of woody biomass (*woodlands*, *shrublands*) are preferred by the levelling phase. *Grasslands* show increased fluctuations in the end of the rainy season and the beginning dry season.

By comparing the final accuracy measures it becomes apparent that the annual classifications show distinct differences of the class-wise accuracies between the years. Inter-annual classifications are characterised by a more distinct levelling, resulting in higher final accuracies compared to the annual classifications. As an example, the *grassland* class benefits most when inter-annual time series are used for the classification. Annual accuracies using bi-monthly composites are varying between 57 %, 48 %, and 49 % between 2004 and 2007, whereas 58 % could be measured using inter-annual time series metrics.

The ANOVA comparing the two groups of annual and inter-annual classification accuracies resulted into a rejection of null hypothesis (a). As listed in Table 5-4, classifications performed on an annual basis produce significantly different results within different years (F value = 16.75). Annual multiple mean comparisons could not prove a relationship to annual rainfall patterns, since the comparatively wet rainy season of 2006/05 did not achieved significantly different accuracies. The ANOVA of the inter-annual classifications show, compared to the annual, an increased F value (187.14) and significantly higher accuracies with each year that was added to the cumulative classification. 14.07 % and 18.07 % higher mean accuracies were achieved when including the growing seasons of 2005/06 and 2006/07, respectively, compared to the reference year of 2004/05. Thus, classification accuracies increase with the length observation period.

5.3.2 Effects of composite length on classification accuracy

Composite length is a critical parameter in highly spatiotemporally dynamic savannas. Mapping accuracy (Figure 5-6) shows that composite length has an increased impact on dynamic classes such as *bare areas and pans* and *grasslands*. This is indicated by the high differences of the final accuracies of the annual classifications. For example, *grasslands* in 2005/06 achieve the best accuracy scores when 16-day composites are used (56 %) compared to three-monthly composites (38 %). Similar patterns can be observed for *shrublands* and *woodlands* where smaller composite lengths result into higher accuracy scores in the annual classification runs. For the inter-annual classification the final accuracy scores are increased and the ranking is not as pronounced compared to the annual classifications.

The ANOVA runs using the composite length as the grouping criterion reject the null hypothesis (b) and show significant differences between the composite periods, both for annual (F value =

57.47) and inter-annual classifications (F value = 32.09). The results of the multiple mean comparisons (using monthly composites as reference group) prove the hypothesis that smaller temporal composites lead to increased classification accuracies. For annual classifications the mean accuracy values of the 16-day composites are with 6.39 % higher than the monthly composite classification runs (64.07 %). Bi- and three-monthly composites achieve with -6.63 % and -9.2 % lower accuracies than 16-day and monthly composites.

Table 5-4 Results of the one-way ANOVA and multiple mean value comparisons of annual and inter-annual cumulative classification accuracies compared with composite length, land-cover class, and observation time.

ANOVA Run	*DF*	*Sum Sq*	*Mean Sq*	*F Value*	*Pr (>F)*	*Mean Pr(>|t|)*	*Mean Pr(>|t|)*	*Mean Pr(>|t|)*	*Mean Pr(>|t|)*	*Mean Pr(>|t|)*
Composite						1-m	16-d	2-m	3-m	-
annual	3	52919	17639.6	57.47	< 2.2e-16 ***	64.07 (ref) < 2e-16 ***	6.39 1.10e-06 ***	-6.63 4.38e-07 ***	-9.2 3.01e-12 ***	-
Inter-annual	3	25632	8544.2	32.09	< 2.2e-16 ***	76.44 (ref) < 2e-16 ***	1.58 0.201	-6.8 4.19e-08 ***	-8.43 1.21e-11 ***	-
Class						Vos	P	G	Os	W
Annual (total)	4	281589	70397	476.41	< 2.2e-16 ***	63.11 (ref) < 2e-16 ***	14.68 < 2e-16 ***	-26.82 < 2e-16 ***	-1.82 0.0721 .	7.02 6.15e-12 ***
Inter-annual (total)	4	219930	54983	432.41	< 2.2e-16 ***	74.89 (ref) < 2e-16 ***	10.86 < 2e-16 ***	-25.05 < 2e-16 ***	-2.33 0.0147 *	7.21 7.23e-14 ***
Annual (16-day)	39	2794144	71645	692.53	< 2.2e-16 ***	72.4 (ref) < 2e-16 ***	79.92 (ref) < 2e-16 ***	48.95 (ref) < 2e-16 ***	71.22 (ref) < 2e-16 ***	79.83 (ref) < 2e-16 ***
Annual (1-mth)	39	2794144	71645	692.53	< 2.2e-16 ***	-6.5 2.99e-14 ***	-1.35 0.110625	-10.34 < 2e-16 ***	-6.98 < 2e-16 ***	-6.83 8.35e-16 ***
Annual (2-mth)	39	2794144	71645	692.53	< 2.2e-16 ***	-13.45 < 2e-16 ***	-5.12 3.59e-10 ***	-17.46 < 2e-16 ***	-14.51 < 2e-16 ***	-14.34 < 2e-16 ***
Annual (3-mth)	39	2794144	71645	692.53	< 2.2e-16 ***	-17.26 < 2e-16 ***	-1.89 0.025503 *	-22.88 < 2e-16 ***	-18.26 < 2e-16 ***	-17.61 < 2e-16 ***
Inter-ann. (16-day)	39	2794144	71645	692.53	< 2.2e-16 ***	7.9 < 2e-16 ***	6.1 6.29e-13 ***	11 < 2e-16 ***	8.11 < 2e-16 ***	5.64 2.92e-11 ***
Inter-ann. (1-mth)	39	2794144	71645	692.53	< 2.2e-16 ***	6.73 2.11e-15 ***	7.14 < 2e-16 ***	6.05 9.90e-13 ***	5.55 6.26e-11 ***	5.6 4.24e-11 ***
Inter-ann. (2-mth)	39	2794144	71645	692.53	< 2.2e-16 ***	0.05 0.956277	4.92 6.45e-09 ***	-5.44 1.48e-10 ***	-1.99 0.01897 *	-0.06 0.945036
Inter-ann. (3-mth)	39	2794144	71645	692.53	< 2.2e-16 ***	-3.7 1.31e-05 ***	5.54 6.35e-11 ***	-6.69 3.19e-15 ***	-5.29 4.55e-10 ***	-0.95 0.262745
Time						2004/05	2005/06	2006/07	-	-
Annual	2	11128	5564.2	16.57	7.672e-08 ***	62.59 (ref) < 2e-16 ***	2.01 0.09 .	-4.63 9.42e-05 ***	-	-
Inter-annual (cumulative)	2	83954	41977	187.14	< 2.2e-16 ***	62.21 (ref) <2e-16 ***	14.07 <2e-16 ***	18.07 <2e-16 ***	-	-

Significance levels: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '; class labels: Vos…Shrubland (very open), P…Bare areas and Pans, G…Grassland, Os…Shrubland (open), W…Woodland. (ref) indicates the control group for the multiple mean comparisons.

The inter-annual monthly mean accuracy values are with 76.44 % significantly higher compared to the annual results. The mean comparisons of the monthly and 16-day composites could not achieve significant differences. Bi- and tree-monthly composites show similar patterns compared to the annual classifications. The temporal profile of the Kappa statistic (Figure 5-5) also indicates an increased relevance of smaller composite lengths. However, using smaller temporal composites leads to increased variations due to rainfall and vegetation dynamics, as given in the growing season of 2005/06 for 16-day and monthly composites.

5.3.3 Effects of observation period and composite length on class-specific accuracies

Class-specific differences of observation period and composite length were assessed with the ANOVA run based on a combined grouping of observation period and composite length with 39 degrees of freedom (Table 5-4). Annually classified 16-day composites were used as reference group for the multiple mean comparisons. The F value of 692.53 proves that significant differences are existent among the groups, and thus, rejects the null hypothesis (c). As mentioned above, observation period has the greatest impact on *grasslands*. Inter-annual classifications increase the mean accuracies of 16-day and monthly composites by 11 % and 6.05 % and decrease the difference to classifications based on the bi- and three-monthly composites (-5.44 and -6.69). Similar patterns can be observed when different composite lengths are compared. The mapping accuracy of the *grassland* class increases with lower composite length and longer observation period. *Shrublands* and *woodlands* achieved similar results with lower deviations of the mean values. For those classes the use of inter-annual time series metrics could not enhance the accuracy when bi-monthly and three-monthly composites were used (Pr(>|t|): Vos = 0.96, Os = 0.02, W = 0.95 and 0.26). For annual classifications the comparison of the group means for *bare areas and pans* could not detect significant differences when 16-day and monthly composites were used. This class benefits most from an increasing observation period since all group means of the inter-annual classifications are significantly higher than the annual classifications.

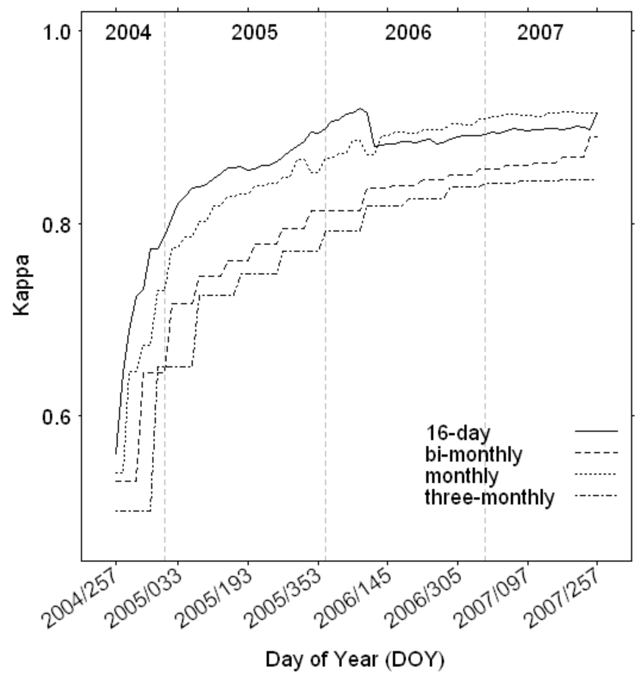


Figure 5-5 Kappa coefficients of cumulative multi-annual classifications using sets of different composite lengths.

5.3.4 Dependencies of rainfall variability and observation period

The comparisons of cumulative monthly precipitation with cumulative mean accuracies using a piecewise linear regression technique resulted for all classes in positive correlations ($p < 0.05$). For annual as well as for inter-annual classifications a deterministic relationship between classification accuracy and cumulative rainfall (and thus observation period) could be proved. A stronger linear relationship was found for grasslands for all observation periods ($R^2 = 0.79$ to 0.85) and vegetation types dominated by woody vegetation, such as *open shrublands* (R^2 = 0.75 to 0.84), *very open shrublands* (R^2 = 0.69 to 0.85), and *woodlands* (R^2 = 0.68 to 0.84). Accuracies of *bare areas and pans* are less coupled to rainfall, which is apparent in the inter-annual classification results (R^2 = 0.36) and the increased range in the annual correlations (R^2 = 0.36 to 0.79).

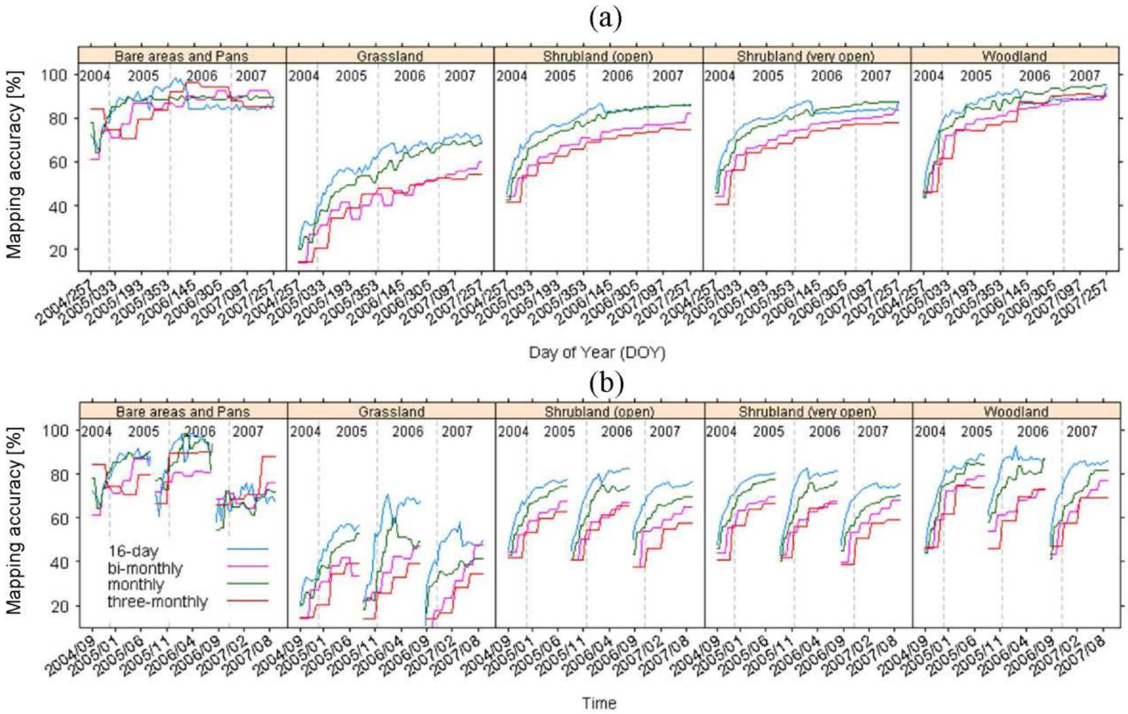


Figure 5-6 Class-wise mapping accuracies of cumulative (a) annual and (b) inter-annual classifications visualise the class-specific improvement of the classification performance with increasing observation period.

As shown in Figure 5-7, a cumulative precipitation amount of 1800 mm could be measured by the TRMM observations for the overall observation period (September 2004 – October 2007). Effects of inter-annual variations in rainfall indicate the annual correlations, e.g. the comparatively wet growing season (September 2005 to October 2006) with approximately 1000 mm precipitation. Except of *bare areas and pans* and *grasslands*, the distribution of precipitation versus accuracy follows a Michaelis-Menton distribution, as described in Bucini *et al.* (2007). This is more pronounced at the inter-annual classifications. The shape of the distribution is going to saturate with high precipitation sums, where the accuracy levels out. Pointing on the annual classifications, the shape of the distribution deviates from Michaelis-Menton and the accuracy values decrease after a critical rainfall amount has reached, following a 'broken-stick' distribution.

Using a piecewise regression model the threshold values, both accuracy level and cumulative precipitation, could be detected at which the accuracy moves toward saturation or changes to decreasing values. Moreover, the slopes of the two regression lines indicate the strength of the relationship between mapping performance and rainfall amount. By analysing the breakpoints at the inter-annual classifications (Figure 5-7), it is clear that the accuracy levels are higher than the annual results, ranging from 75 % to 90 % except for the *grassland* class.

Breakpoints at the classification example of 2005/06 are found at a lower accuracy level. The thresholds depend on the respective observation period. Different thresholds were detected for annual and inter-annual classifications, ranging from 168 mm to 464 mm (inter-annual) and 108 mm to 594 mm (annual). However, the trigger of the threshold can be seen in the cumulative precipitation. For all classes a rapid improvement of the accuracy was found at the initial observation period, which is shown by increased regression slopes below the breakpoint. By comparing the slopes of the first regression line (Table 5-5), it is obvious that *woodlands* (maximum initial slope 0.2333 in 2004/05) and *shrublands* (e.g. for *open shrubs* 0.1303 in 2005/06, *very open shrubs* 0.1678) benefit most from the earlier observations. The second slope indicates how effective an increasing observation period is in order to improve the mapping accuracy. Different patterns are obvious for annual and inter-annual classifications.

The inter-annually derived accuracies show slightly positive slopes for all classes after the change point, whereas the temporally dynamic classes show slightly negative slopes after a critical rainfall amount between 250 mm and 500 mm. Negative second slopes were derived *for bare areas and pans* for all annual correlations (-0.0016 to -0.01) and for *grasslands* in the wet rainy season of 2005/06 (-0.01). Compared to the annually derived classifications, the second slopes are slightly increased and lead to higher accuracies when longer time series are used. This implies that the extended temporal rainfall profile in the feature set leads to stable mapping results.

The breakpoints range from 300 mm in regular rainy seasons and 500 mm in the extreme rainy season of 2005/06. The mean breakpoint was found at 256.7 mm ± 146.9 mm, which marks the range of long-term-mean MAP (Figure 5-2c). These cumulative rainfall amounts are reached in the first quarter of the rainy season (compare Figure 5-3a).

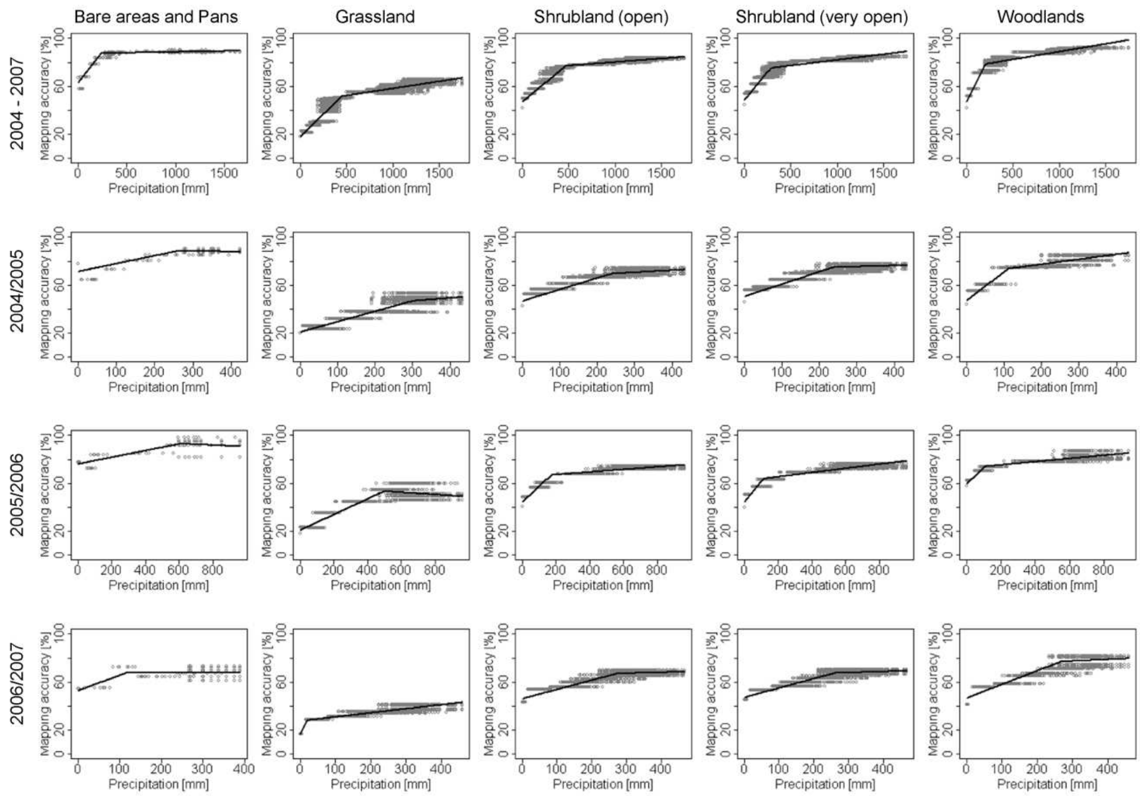


Figure 5-7 Piecewise linear regression between cumulative precipitation and class-specific cumulative mapping accuracies. The temporal relationship between accuracy and precipitation is compared between inter-annual (2004 - 2007) and annual classification of the growing seasons 2004/2005, 2005/06). Accuracies are based on classifications using monthly composites. Regressions were performed on a 95% confidence level.

Table 5-5 Correlation of cumulative precipitation with cumulative annual and inter-annual mapping accuracies using piecewise linear regression.

Class	*Observation period*	R^2	*Breakpoint (mm)*	*Initial slope*	*Slope change*	*Second slope*
Bare areas and pans	2004 - 2007	0.36	241	0.0994	-0.0996	0.0014
	2004/2005	0.79	256.3	0.0638	-0.0713	-0.0039
	2005/2006	0.7	594	0.0275	-0.0383	-0.0102
	2006/2007	0.4	119	0.1249	-0.1301	-0.0016
Grassland	2004 - 2007	0.8	439	0.0753	-0.0642	0.0117
	2004/2005	0.85	298.26	0.0864	-0.0648	0.0220
	2005/2006	0.8	496.53	0.0652	-0.0758	-0.0104
	2006/2007	0.79	19.6	0.5720	-0.5511	0.0338
Shrubland (very open)	2004 - 2007	0.69	280.77	0.0925	-0.0845	0.0093
	2004/2005	0.85	236.76	0.1021	-0.0938	0.0085
	2005/2006	0.81	112.64	0.1678	-0.1524	0.0175
	2006/2007	0.84	259	0.0775	-0.0706	0.0076
Shrubland (open)	2004 - 2007	0.75	463.84	0.0637	-0.0585	0.0058
	2004/2005	0.84	237	0.0956	-0.0795	0.0168
	2005/2006	0.82	167.52	0.1303	-0.1234	0.0105
	2006/2007	0.84	269.9	0.0767	-0.0708	0.0067
Woodland	2004 - 2007	0.68	172	0.1487	-0.1696	0.0129
	2004/2005	0.84	108.26	0.2332	-0.2022	0.0408
	2005/2006	0.77	96.31	0.1277	-0.1292	0.0128
	2006/2007	0.81	266.1	0.1120	-0.1043	0.0109

5.4 Discussion

5.4.1 Effects of environmental cues

The distribution of different physiognomic vegetation classes can be very scattered, mainly caused by fire, grazing effects, and different land-use practices (Archibald & Scholes, 2007; Scanlon *et al.*, 2007; Thomas & Twyman, 2004). As analysed by Archibald & Scholes (2007) and according to the class-specific accuracies in Figure 5-6, woody vegetation shows a less sensitive inter-annual variability than herbaceous life-forms. Mean annual precipitation is a key variable determining the occurrence of woody and herbaceous life-forms (Bucini *et al.*, 2007; Sankaran *et al.*, 2005). Landscape heterogeneity affects the coexistence of woody and herbaceous life-form classes using a classification scheme adapted to 'coarse' multi-temporal imagery, as indicated by the specific layer descriptions in Table 5-2. Beside the above-mentioned factors that influence the occurrence of vegetation patterns, environmental cues are key regulators of vegetation activity. Those differences might be explained by memory effects (mainly for woody species), and differing magnitudes of environmental cues related to water availability and temperature influencing plant phenology (Childs, 1989; Goward & Prince, 1995). Various studies found that soil moisture is the most important trigger for the initiation of herbaceous biomass production, whereas the initiation of leafs for woody species is mainly triggered by temperature (Shackleton, 1999; 2002). Beside leaf flush in accordance to initial rainfall events, memory effects causing pre-rain flushing was observed for some woody species in Sub-Saharan savannas (Childs, 1989).

5.4.2 Determinants of mapping accuracy

As demonstrated, the rainfall patterns can underlie remarkable spatial variations combined with a pronounced inter-annual variability (Figure 5-4 – Figure 5-7). The results prove the hypothesis that rainfall is a key parameter influencing the mapping performance when satellite time series are used. Previous analyses of station data covering mean annual precipitation (MAP), temperature, and soil characteristics coupled with eight years of AVHRR observations showed that woody cover is mainly a function of MAP in arid and semi-arid savannas. Based on piecewise linear regressions, (Sankaran et al., 2005) estimated for the African continent a MAP of 650 ± 134 mm at which the maximum tree cover is attained. A minimum of 101 mm MAP is required for the occurrence of trees. Similar results were obtained by Bucini & Hanan (2007) when the MODIS 500 m tree cover product (Hansen *et al.*, 2003) was compared with MAP. Compared to the quintile regression line representing the upper boundary of woody cover as presented by Sankaran *et al.* (2005), the comparison of MODIS with MAP achieved the best functional responses with a sigmoid relationship. The relationship of mapping accuracy and cumulative precipitation was found to be comparable to results of Sankaran *et al.* (2005), where the occurrence of woody cover increases with increasing MAP, and levels between 516 and 784mm. In fact, the breakpoint at which the maximum tree cover is attained is over the range at which the maximum accuracy (19.6 – 496.53) is attained or levels out. Thus, the functional response of woody cover occurrence and mapping accuracy to rainfall seems to be comparable.

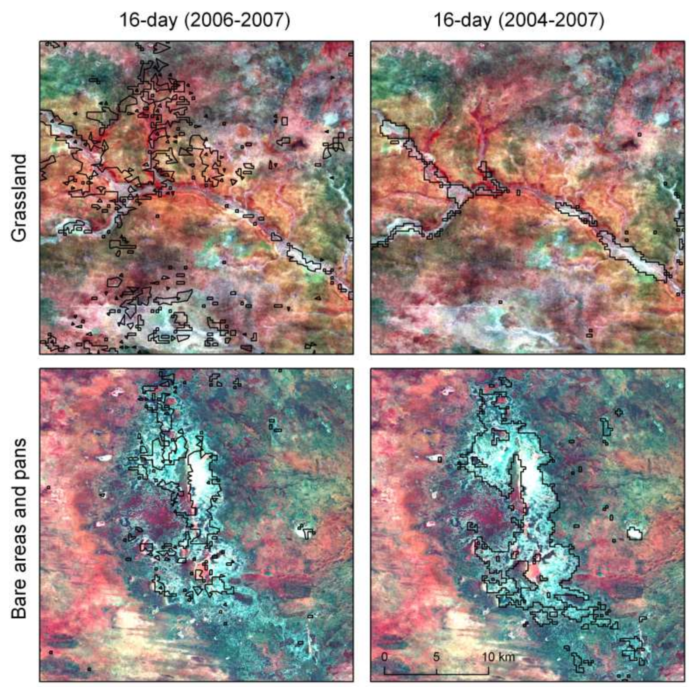


Figure 5-8 Two examples of MODIS time series classifications showing differing spatial patterns of class distributions depending on the observation periods of 16-day composites compared with a Landsat RGB display of the Central Kalahari drainage system (upper: Landsat TM RGB: 4-3-2, acquisition: 04-27-2004, path 176, row 74) and the *Naye-Naye* salt pans in the Kalahari (lower: Landsat TM RGB: 4-3-2, acquisition: 04-20-2004, path 177, row 74). The black-lined polygons indicate the cartographic delineation of grasslands (upper row) and *bare areas and pans* (lower row).

Further, it becomes obvious that the degree of accuracy improvement (accuracy value and slope at which the breakpoints are derived) is a function of how variable any given land-cover type is. Mean EVI values (Figure 5-3b) and cumulative (annual and inter-annual) mapping accuracies (Figure 5-6b) indicate that very dynamic land-cover classes like *grasslands* (high inter-annual variations in photosynthetic activity) and *bare areas and pans* (alternating occurrence of herbaceous cover and flooding situations) benefit most in terms of accuracy when longer time series are used for classification. Examples are given in Figure 5-8 comparing annual and inter-annual classifications of *grasslands* and *bare areas and pans*. A remarkable reduction of omission and commission errors is obvious using multi-seasonal image features from 2004 to 2007. The delineation of the *grassland* class is limited to the incised river bed. The example of the *Naye-Naye* salt pan system shows that a realistic spatial delineation of *bare areas and pans* is obtained by integrating multi-seasonal land surface dynamics in the feature set.

Cover-types with a certain fraction of woody life forms (*woodlands* with open to closed tree and shrub layer, *shrublands* with a very open to open shrub layer, compare Table 5-2) have a rapid increase and stable accuracies over time in common. In respect of the temporal change patterns of the mapping accuracy and the breakpoint and slopes from the piecewise linear model, we propose a conceptual model of the relationship between mapping accuracy and observation period as a function of precipitation amount and magnitude of change between certain land-cover stages. As shown in Figure 5-9, the accuracy increases with increasing observation period (and thus precipitation input). The upper and lower accuracy limits are shown to be a function of the variability and changeability of a given land-cover type. Variability of land-cover can be defined as the number and magnitude of land-cover stages that are passed over time for certain land-cover types. For example*, grasslands* can pass different land-cover stages over the growing season, such as almost barren ground in the dry season (due to grazing or fire) or a consistent coverage of photosynthetic active vegetation in the rainy season. *Grasslands* have further a rapid response of biomass production to rainfall characterizing them to have a higher magnitude and variability between land-cover stages. *Woodlands* are less coupled to soil moisture (due to a deeper root system and pre-rain flush effects), have lower variability in phenological cycling, and can therefore be classified as less variable land-cover types.

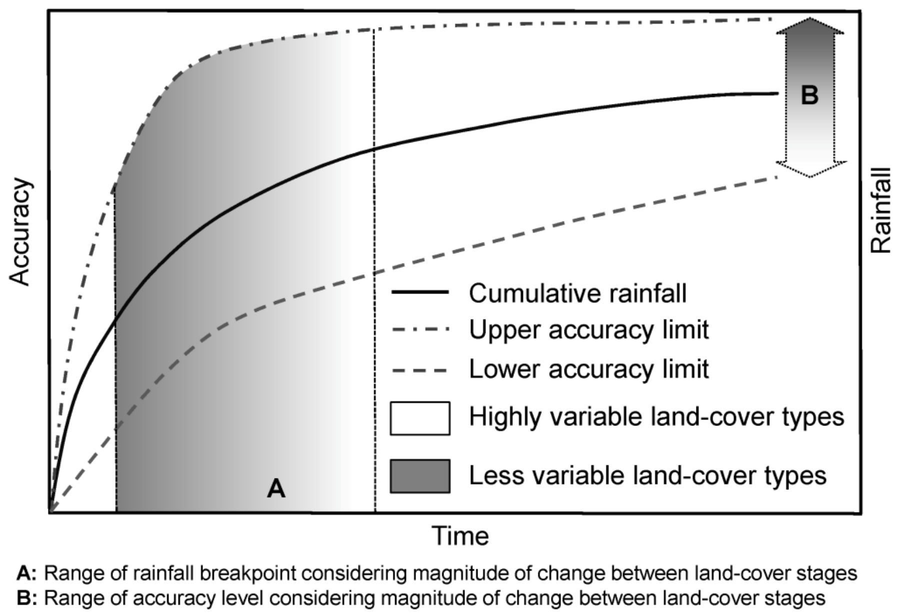


Figure 5-9 Conceptual model of the relationship between mapping accuracy and observation period as a function of precipitation input and magnitude of change between land-cover stages.

Degree and changeability of land-cover stages determine the rainfall threshold (breakpoint at which the accuracy levels out, Figure 5-9a) and the accuracy level (upper or lower accuracy limit, Figure 5-9b). In terms of mapping accuracy, highly variable land-cover types benefit most from multi-seasonal time series compared to annual classifications.

5.5 Summary and conclusions

In this study, we compared the effects of different composite lengths and observation periods with class-wise mapping accuracies derived from multi-temporal MODIS time series classifications. Further on, we assessed the effect of precipitation patterns using TRMM data on mapping accuracy in a dry semi-arid savanna environment. The commonality of the recent globally available land-cover datasets are comparatively decreased accuracies for classes related to the savanna biome. This study points at some aspects regarding the conceptual design of large-area and global land-cover mapping frameworks.

- *Which effect does temporal compositing and variations of the observation period have on the land cover classification accuracy in semi-arid areas? How can the remotely sensed feature assembly be adapted to consider inter- and intra-annual variations, affecting spatiotemporal patterns of the vegetation activity and phenological cycling?*

Compositing and temporal sampling of composite periods are crucial methods for data compression, increase the number of calibrated and quality observations and for reduction of radiometric distortions (Vancutsem *et al.*, 2007a) and, thus, the removal of 'noise' causing decreased classification results (Maxwell *et al.*, 2002). In particular, within the framework of global and large-area regional land-cover mapping initiatives, noise reduction and the generation of meaningful input features is necessary for the optimisation of the final classification result.

Several implications arise concerning the existing methodological setups of global land-cover mapping initiatives respective to semi-arid savanna environments. Most of the global land-cover datasets (IGBP DISCover, MODIS land cover product collection 5, UMD, and GLC2000, see Table 5-1) were based on monthly aggregated composites, which proved to be advantageous regarding high accuracy scores for all cover types in the inter-annual classifications and the comparative robustness against inter-seasonal rainfall variability. The longer composite periods used for the generation of the GlobCover (bi-monthly) and GLC2000 datasets (monthly to three-monthly) obviously expose most limitations for mapping very dynamic land-cover types.

In regard to the effect of the observation period on mapping accuracy it was found that multi-annual classifications caused reliable results. Annual classifications led to a considerable inter-annual variability of accuracy dependent on the given rainfall season. 'Traditional' global land-cover mapping initiatives use annual assemblies of satellite time series covering at least one growing season. Global datasets serve the purpose of an objective and spatiotemporally consistent monitoring of the environment, and will thus be periodically updated.

In order to increase the inter-annual comparability of those geo-information products in particular in savanna environments, we suggest using multi-annual time series feature sets for the classification covering at least two growing seasons. It is advised to use smaller composite lengths if classifications are performed on small (intra-seasonal to annual) observation periods (Figure 5-5).

- *What are the effects of inter-annual rainfall variability on the land-cover classification accuracies in rainfall-driven savannas?*

A distinct intra- and inter-annual variability of the classification performance was observed, where woody-dominated classes proved to be more stable than *grasslands* or classes with an herbaceous vegetation layer (Figure 5-6b). Furthermore, an apparent optimisation of class-specific mapping accuracies was achieved using input feature sets of MODIS time series covering multiple vegetation periods. The results on the effect of the observation period exposed specific temporal requirements and the influence of environmental cues on the classification performance. We showed that mapping accuracy is correlated with rainfall. This explains why the occurrence of herbaceous and woody land-cover types spatially varies between the years (Figure 5-1). Increased rainfall can have a negative impact on accuracy for variable land-cover types as shown for the annual classifications (breakpoints at a maximum of 500 mm).

A general question arising is: How can we adapt the feature space for a most accurate classification of a specific land-cover type? Such kinds of 'customised' assemblies of feature sets can be useful towards a reliable spatial representation of land-cover types with a distinct dynamic component. In respect of dry savanna ecosystems the inclusion of the multi-annual phenological variability in the classification process led to an improvement of the final mapping result. Kaptué *et al.* (2010) have recently mapped ecosystems on continental scale in Africa using eight years of AVHRR time series which proved to be promising for future large-area ecosystem assessments.

This study emphasises the general need to analyse the determinants and key factors for land-cover mapping in very dynamic landscapes, including the spatiotemporal heterogeneity of cover types (e.g. wetlands, tundra, savannas, and transition zones in general). Regional stratification strategies, as done for the GLC2000 and GlobCover initiatives, helped so far to include the climatic variability and ecosystem-dependent vegetation dynamics for the generation of global land-cover databases. From a data-perspective the remote sensing community lives in a data-rich world (regarding existing satellite time series databases, data access, database management, computing power, and classification techniques), which signifies that dynamic processes (defining a specific land-cover type) were covered in existing hyper-temporal satellite databases.

The demonstrated spatiotemporal variations of large-area land-cover products and spatial heterogeneity of different life form groups in savannas face the following challenging tasks:

(a) *The accurate mapping of vegetation types comprising specific phenological characteristics of multiple life-form layers.* A dominant photosynthetic activity of a single vegetation layer (e.g. grasses vs. woody vegetation) can affect incorrect class assignments. The inclusion of multiple growing seasons and the adaption of the feature setup on the vegetation phenology can lead more stable vegetation classifications.

(b) *The accurate thematic representation of a complex vegetation structure within a categorical classification scheme using coarse scale time series.* A classified pixel represents in many remote sensing datasets a compound of different life-forms in this region. The use of high resolution time series data that are becoming more available (e.g. RapidEye, Landsat Data Continuity Mission, and Sentinel-II) will enhance the accurate mapping of semi-arid land-cover types, characterised by a very high patchiness of the vegetation structure. The application of a standardised and flexible classification system, as defined in the FAO Land Cover Meta Language (LCML, LCCS-3), further improves the thematic representation of heterogeneous landscapes in coarse scale land-cover products.

(c) *The generation, provision, and standardisation of in-situ reference databases.* Southern African savannas are poorly populated, with a limited infrastructure, and poorly analysed concerning vegetation composition, distribution, and phenology. The inclusion of *in-situ* reference data in the framework of large-area land-cover mapping initiatives is limited. The collection and translation of available botanical databases in LCCS/LCML cause the most effective transfer of land-cover information from local to regional and global scales.

Particularly the savanna biome includes a number of so-called “dynamic classes” or classes with a very high spatiotemporal variability of certain land-cover stages. While mapping such a dynamic landscape in terms of vegetation structure and dynamics with time series data, the final map often represents simplified thematic information (compared to homogeneous cover types, e.g. tropical rain forests and boreal evergreen forests). Further interdisciplinary research has to be conducted to reveal the influence of ecosystem-specific dynamics in the framework of large-area mapping through ecosystem-oriented case studies to capture the dynamic characteristics of cover types within the setup of feature sets for the classification of satellite time series.

6 The Potential of MODIS Time Series Metrics for bottom-up Vegetation Mapping in a semi-arid Savanna Ecosystem in Namibia[4]

Abstract - Reliable land cover and vegetation maps are key information for land planning and resource management. Less than 40 % of Namibia's semi-natural environments have been mapped in terms of vegetation distribution. This study shows the potential of time series metrics from the Moderate Resolution Imaging Spectroradiometer (MODIS) for large-area vegetation mapping in Namibia. Annual and inter-annual MODIS time series metrics were linked to *in-situ* data and used as features for land-cover and vegetation mapping based on Random Forests. The results showed that land-cover classifications based on inter-annual time series metrics achieved increased mapping accuracies, compared to classifications using annual time series. A distinct seasonal variation of the classification accuracies was evident. Integrated analyses of *in-situ* and coarse EO time series contribute for an extended ecosystem understanding in seasonal savanna ecosystems.

[4] Published as: Hüttich, C., Strohbach, B., Herold, M., Keil, M., & Dech, S. (2010): The Potential of MODIS Time Series Metrics for bottom-up Vegetation Mapping in a semi-arid Savanna Ecosystem in Namibia. *Proceedings of the ESA Living Planet Symposium, 28. June - 2. July 2010, Bergen, Norway* (pp. 1-8).

6.1 Introduction

Geo-information data on land-cover are important information sources for the management of urban and natural environments. Land-cover has recently evolved as a key variable in the framework of the assessment of the status of the development of the standards for the terrestrial essential climate variables (Herold, 2009). Despite existing operational land-cover monitoring initiatives at global scales (Loveland *et al.*, 2000; Defourny *et al.*, 2009; Friedl *et al.*, 2010), there are major limitations evident in observing and monitoring highly dynamic ecosystems, e.g. arid and semi-arid ecosystems and wetlands (Giri *et al.*, 2005; Herold *et al.*, 2008). Advancing progress in the standardisation of the observation and characterisation of terrestrial ecosystems is provided by the UN-FAO Land Cover Classification System (LCCS, Di Gregorio, 2005). LCCS was used in several studies as a "common land-cover language" and bridges the gap between the common observational scales, such as *in-situ*, local to regional and global scales (Neumann *et al.*, 2007; Hüttich *et al.*, 2010). A number of studies conducted phenological characterisations of southern African Savannas (Sekhwela & Yates, 2007; Archibald & Scholes, 2007; Shackleton, 1999; Childs, 1989). Phenological metrics derived from 250 m to 1000 m MODIS time series data were used for the characterisation and monitoring of arid and semi-arid vegetation types (Sedano *et al.*, 2005), and were combined with decision tree-based machine learning algorithms (Hansen *et al.*, 2000; Steenkamp *et al.*, 2008). The advantages and limitations of different semi-arid land-cover monitoring concepts were employed in the BIOTA Southern Africa project (Krug *et al.*, 2006). Distinct inaccuracies and challenges in characterizing land-cover types related to the savanna biome were identified, such as (a) the limited availability of up-to-date vegetation maps of Namibia on a national scale, (b) an distinct negative influence of the inter- and intra-annual phenological variability on the accuracies of land-cover maps, and (c) significant differences between land cover classification systems and products (Hüttich *et al.*, 2009). In order to improve integrated characterisations and monitoring mechanisms in arid and semi-arid environments, this paper aims to:

- present an integrated concept for a bottom-up land-cover assessment and fuzzy vegetation type mapping in a dry semi-arid savanna ecosystem based on local scale *in-situ* botanical survey data with coarse scale (MODIS) satellite time series data.
- analyse the temporal requirements of time series metrics derived from MODIS for single annual and long-term inter-annual image classifications from 2001 to 2007.
- show the potential of fuzzy mapping of semi-arid vegetation types by integrating *in-situ* vegetation databases as presence-absence data.

6.2 Data and methods

6.2.1 Study area

Two methodological approaches of the fuzzy vegetation type detection and a general land-cover mapping based on LCCS were conducted on two different spatial extends. The study area comprises the northern part of Namibia (fuzzy vegetation type mapping) and the communal areas in the eastern Kalahari of Namibia (land-cover assessment). The area is characterised by a sub-continental climate with a long term annual average summer rain period from 300-500 mm and often erratic rainfall events (Mendelsohn *et al.*, 2002).

6.2.2 Field data

The *in-situ* reference database, including 420 field plots for the land-cover assessment and over 10.000 plots for the fuzzy vegetation type mapping was generated based on various field surveys in the framework of the BIOTA Southern Africa project and the Desert Margins Project. The Braun-Blanquet approach was used for the field surveys. A phytosociological analysis was conducted resulting in a synoptic vegetation type legend. The botanical classification nomenclature was based on the occurrence of the dominant species (Strohbach, 2001; Strohbach *et al.*, 2004). The phytosociologic vegetation type legend including ten vegetation type classes was translated to a physiognomic-structural LCCS-based legend using the LCCS classifiers set as translation entities. A spatial up-scaling procedure was performed to combine the in-situ point locations with a coarse 232 m pixel resolution of the MODIS time series metrics. The database was intersected with homogeneous polygons derived from an image segmentation of Landsat data to increase the training sample size for classification.

6.2.3 Satellite data and preprocessing

Five Landsat-7 scenes with six channel reflectance bands were pre-processed in terms of geolocation and map projection. 16-day composite images of MODIS collection 5 product MOD13Q1 in sinusoidal projection, surface reflectance bands and the enhanced vegetation index (EVI) on the original 232 m resolution were used for generating time series image classification features. Pre-processing included subsetting and quality analysis using the Time Series Generator (TiSeG) software (Colditz *et al.*, 2008). Low quality data were identified based on the MODIS Quality Assessment Science Data Sets and excluded by linear interpolation. Inter-annual time series (2001-2007) were calculated for EVI and the BLUE, RED, NIR, MIR spectral bands. Four temporal segments per year were derived from the annual time series following the main seasonal characteristics of vegetation activity in Namibia. The first segment (Jan. - Mar.) features the main rainfall period, segment two (Apr - Jun.) the fall and winter season, and segment three (Jul. - Sep.) the dry spring. The fourth segment (Oct. – Dez.) features the early summer dry-rainy season transition. Statistical metrics were calculated after Conrad *et al.* (2007) from temporal segments indicating temporal mean, standard deviation, minimum and maximum value, and range (between minimum and maximum). Inter-annual statistics were calculated from the annual segment statistics.

6.2.4 Random Forest framework

Remote Sensing studies related to time series analyses yield data sets with a large number of predictors (e.g. acquisition times, temporal segment statistics, or phenological metrics) and often a small sample size of the target classes. Beside established machine learning algorithms, such as support vector machines or neural networks, random forests (RF) recently became popular for classification applications.

The RF classification and regression algorithm builds upon bagging classification trees (Breiman, 2001). A random feature selection procedure was implemented, so that the single trees are grown on different bootstrapped samples from the training sample. The random feature selection at each node leads to decreased correlations between the trees. With each grown tree the forest error rate decreases. The classification accuracy is computed based on the out-of-bag (OOB) bootstrap samples. The OOB observations are used to generate an importance measurement for each candidate predictor. A detailed description of RF is given by Archer & Kimes (2008) and Breiman (2001).

Within this study RF were used for the importance measurement for each candidate predictor. RF classification was used for the land-cover mapping. RF regression was used for the fuzzy mapping of specific vegetation types. OOB error estimations were used for the accuracy assessment.

6.3 Results and discussion

6.3.1 A flexible thematic legend: experiences using LCCS classifiers in southern African savannas

The application of the vegetation type legend based on the phytosociologic plant community associations resulted in a legend of multiple levels of thematical detail, a physiognomic-structural and a floristic vegetation type legend (Figure 6-1). The predominant structural vegetation type of the Kalahari was identified as broadleaved deciduous shrubland with herbaceous understory, which is included in the main land-cover type cultivated and managed terrestrial areas.

The comparison of the OOB errors classifications per growing season with the inter-annual classification (Figure 6-2) indicates a higher stability of the classification performance by using inter-annual phenological and spectral time series features.

Former studies showed that woody vegetation is less sensitive to inter-annual variability than grasses (Archibald & Scholes, 2007). An inclusion of features in the classification process containing information on the inter-annual variability of the photosynthetic activity during the growing seasons can help to separate different life-form classes in savannas, which is a response of spatially and temporally variable rainfall events.

Besides the LCCS classifiers characterizing the vegetation physiognomy (life, form, coverage, height) of the dominant and subsequent vegetation layers, a further specification of additional environmental attributes was conducted. Information of landform, lithology, and soils are key parameters for the occurrence of patterns of plant associations and can be crucial information for the spectral analyses in the remote sensing imagery. The definition of environmental and specific technical attributes proved to be essential for the detailed characterisation of semi-arid vegetation types. The major advantage of using a standardised classification system is to eliminate inconsistent classification rules and to reduce the number of impractical combinations of classifiers (Di Gregorio, 2005). In order to optimise the classification system for semi-arid environments, the implementation of a wider range of environmental attributes (e.g. soil texture and soil colour) is recommended. Limitations became apparent for the set of leaf type classifiers in the second modular level. The set of leaf type classifiers did not cover the whole diversity of leaf types. Since the plant group of *Acaciatea* is characterised by a fine leafed physiognomy, the inclusion of an appropriate classifier to the leaf type section of the modular phase in LCCS is suggested.

6.3.2 Temporal requirements in mapping seasonal semi-arid environments

Southern African savannas are characterised by distinct dry and rainy seasons with a high spatiotemporal variability of precipitation rates. The effect of the acquisition date on the overall classification accuracy was tested for singe classifications of different composite periods (16-days, monthly, bi-monthly, and three-monthly) for the growing seasons from 2004 to 2007. The kappa statistics, visualised in Figure 6-3, indicate a clear seasonality following the precipitation patterns during the growing season. In general, the classification accuracy increases prior to the start of the rainy season and decreases equal to the transition from rainy to dry season. The comparison of the feature sets based on maximum value composites and the set of temporal segment statistics (min., max. mean, range) indicate an increased stability of the classification performance due to precipitation patterns. In contrast to the maximum value feature set, the extended set shows a distinct temporal stability and increased kappa values due to seasonal variations with an increasing composite period. The most robust classification performance was achieved using three-monthly composites of the time series segment statistics set.

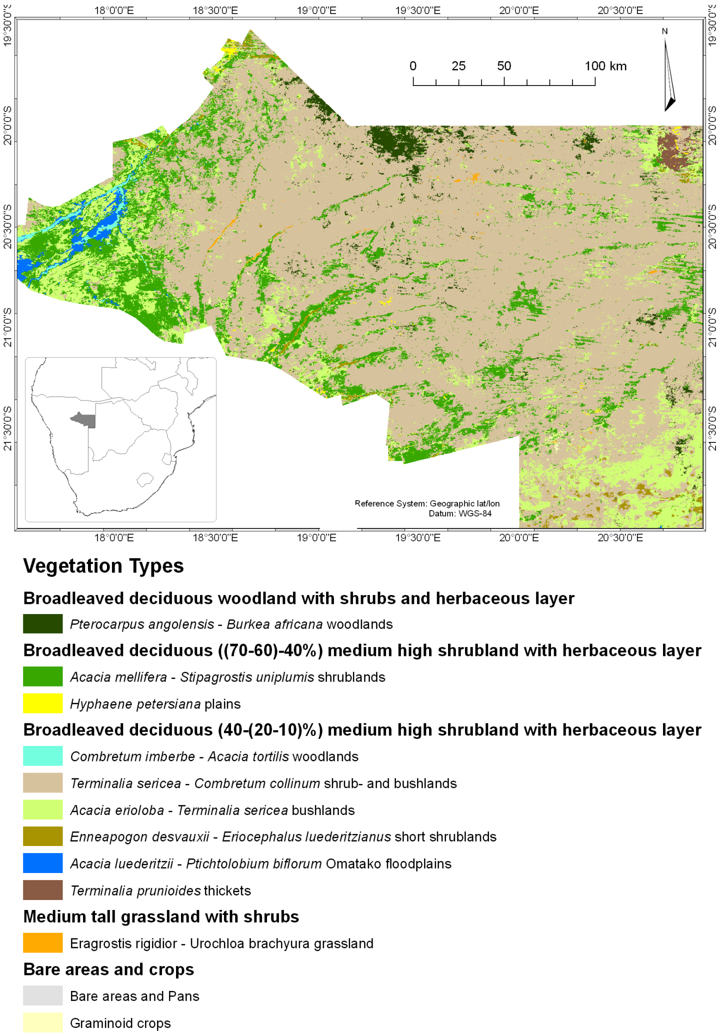


Figure 6-1 Classification of Kalahari vegetation types using inter-annual MODIS time series metrics (2001-2007) with a LCCS-based land-cover legend.

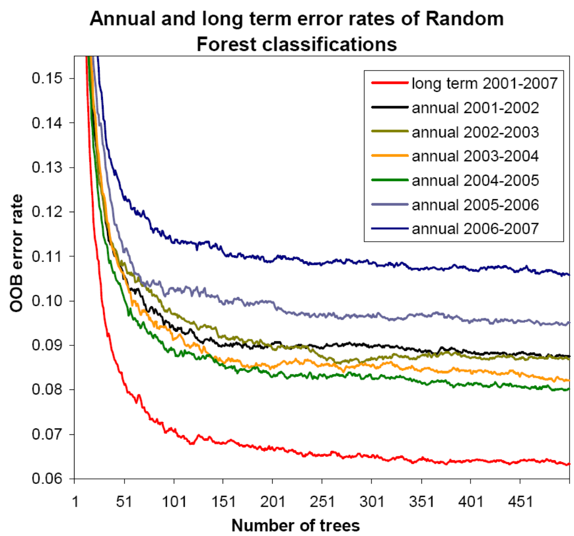


Figure 6-2 Out-of-bag errors for annual classifications per growing season compared to the classification error using inter-annual time series metrics.

6.3.3 Implications for global monitoring

Savannas, defined as landscapes with coexisting vegetation of woody and herbaceous life-forms, are characterised by a very high spatiotemporal variability of vegetation activity. As visualised for three growing seasons in Figure 6-3d, the precipitation patterns in the southern African subtropical savannas show a distinct seasonality and inter-annual variability. Considering this, mapping arid and semi-arid vegetation structure at global scales using coarse scale satellite data is challenging. Thus, the mapping accuracies of savannas are under-average compared to other biomes.

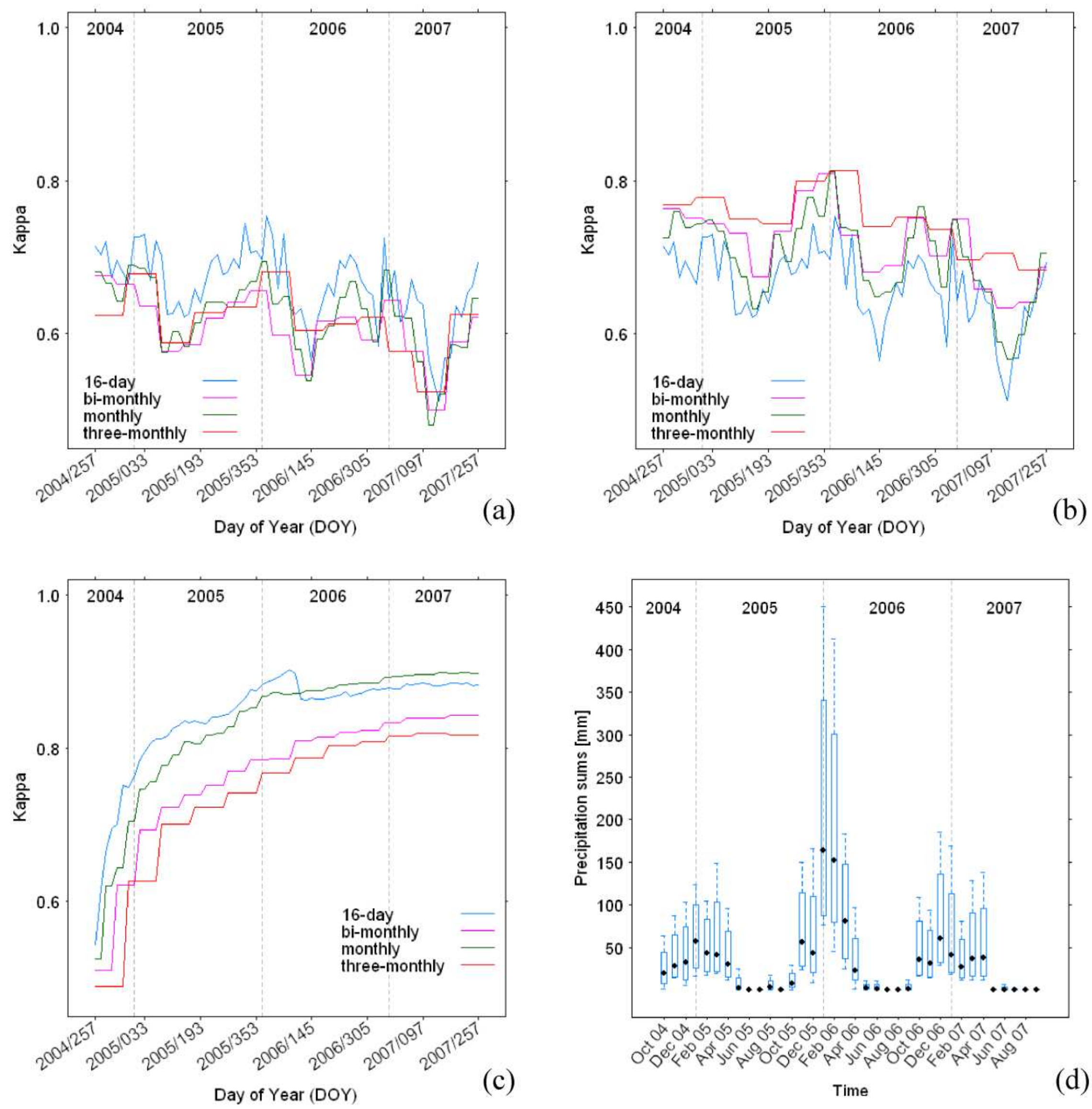


Figure 6-3 Kappa coefficients of different sets of separately classified composite periods indicate seasonal effects and effects of the length of the composite periods on the overall classification accuracy. Different seasonal patterns become apparent by comparing the classification results using maximum value composites (a) with kappa statistics from input features covering max., min, range, and mean segment statistics (b). The best results achieve classifications using inter-annual feature sets, shown for cumulative classifications using time series features covering the whole sample period from 2004 to 2007 (c). The kappa statistics are compared to the spatial statistics of monthly precipitation sums (d) over the study region derived from Tropical Rainfall Monitoring Mission (TRMM) data.

In this study, the implications of large-area land-cover mapping were simulated by classifying maximum value composites with different composite periods, as used in the global initiatives (IGBP DISCover: monthly, MODIS Land Cover: 16-day to monthly, GLC2000: monthly to three-monthly, GlobCover: bi-monthly). The classifications of single composites (Figure 6-3a) indicate that best accuracies were achieved when 16-day composites are used. In general, the mapping performance decreases with increasing composite period length.

For mapping in arid and semi-arid savannas several implications are arising. First, the bi-monthly composite period as used for the production of the GlobCover land-cover map appears to be inappropriate to consider the temporal variability in savanna biomes. Small sample period lengths (e.g. as used for MODIS LC) account for short-term phenological responses of woody and herbaceous life forms due to erratic precipitation patterns. Second, the overall sample period plays a critical role in order to produce meaningful mapping accuracies. As visualised in Figure 6-3c, the processing of a minimum of one growing season is required to consider the phenological cycle of dry and rainy season vegetation states. The smaller the overall sample period of time series features is, the smaller should be the composite period length to increase the mapping performance.

6.3.4 Fuzzy classification of arid and semi-arid vegetation types using random forest regression

Vegetation mapping can be very problematic in very heterogeneous landscapes where single vegetation types are poorly represented by coarse pixels. Fuzzy logic offers an alternative to 'hard' classifications based by evaluating the membership of a specific class to each pixel. Fuzzy membership is based on fuzzy set theory by assuming that membership to a given category will range from 100 % membership to non-membership (Gopal & Woodcock, 1994). A supervised RF regression technique is used to estimate the per-pixel membership for selected vegetation types in northern Namibia. A field database including an inventory of over 10.000 field plots was used to define presence and absence locations.

The proportional vegetation maps (Figure 6-4) display the probability of occurrence for the classes wetlands and hardpans (Figure 6-4b), *Terminalia sericea – Combretum collinum* shrub and bushlands (Figure 6-4c), and *Acacia erioloba – Terminalia sericea* shrublands (Figure 6-4d). The fuzzy vegetation type maps are (a) compared to the class boundaries of the official vegetation type map (Figure 6-4a) after Giess (1971), and (b) overlain with the field plots locations from the vegetation database of the National Botanical Research Institute of Namibia (NBRI) and the field plots of BIOTA network (www.biota-africa.org).

The examples shown in Figure 6-4 demonstrate the improvements for mapping the vegetation distribution by combining *in-situ* databases with coarse earth observation time series. The continuous classification result of percentage class membership account for the spatiotemporal heterogeneity of savanna landscapes. An example is given for the fuzzy mapping of wetlands. Figure 6-4b shows the Etosha pan, associated as an important water resource and wetland habitat in northern Namibia. The map shows areas where different wetland dynamics are occurring. Areas with high class membership values indicate the core pan with the longest flooding periods. The class membership decreases at the borders of the pan due to the increasing occurrence of herbaceous vegetation. The influence of the northern Cuveley drainage is indicated by moderate class membership values for wetlands due to sub-pixel occurrences of fine scattered wetlands. The spatial representation of characteristic vegetation types and transitional zones is shown for *Terminalia - Combretum* savannas (Figure 6-4c). Transitional zones to incised river beds and the bordering vegetation types are well represented by lower class membership values. The probability of occurrence of the *Acacia – Terminalia* savannas (Figure 6-4d) marks the transitional zone from deep Kalahari sands to more loamy and calcareous soil conditions, to be found in central Namibia and the Cuveley drainage system. The fuzzy per-pixel representation supports the spatial distribution modeling of fine scattered semi-arid vegetation types. The moderate and lower class memberships indicate the sub-pixel proportions of a specific vegetation type.

The validation was conducted using the out-of-bag samples of the random forest regression. In general, decreasing RMSE can be observed with increasing number of regression trees in the ensemble (Figure 6-5). For the probability maps shown in Figure 6-4, the accuracy stabilises after 50 trees were grown. The best classification performance achieved hardpans (RMSE=2.4) followed by *Acacia-Terminalia* shrublands (RMSE=13.3) and *Terminalia-Combretum* shrublands (RMSE=13.7). The perturbation process of the random forest framework affects an increase of accuracy when classifying very similar vegetation types, although a large number of similar features (so-called weak predictors) are found. The advantage of using out-of-bag data for the error assessment is that the training database required no reductions in terms of sample size which is particularly important to properly represent small classes or classes with lower plot numbers due to limited accessibility.

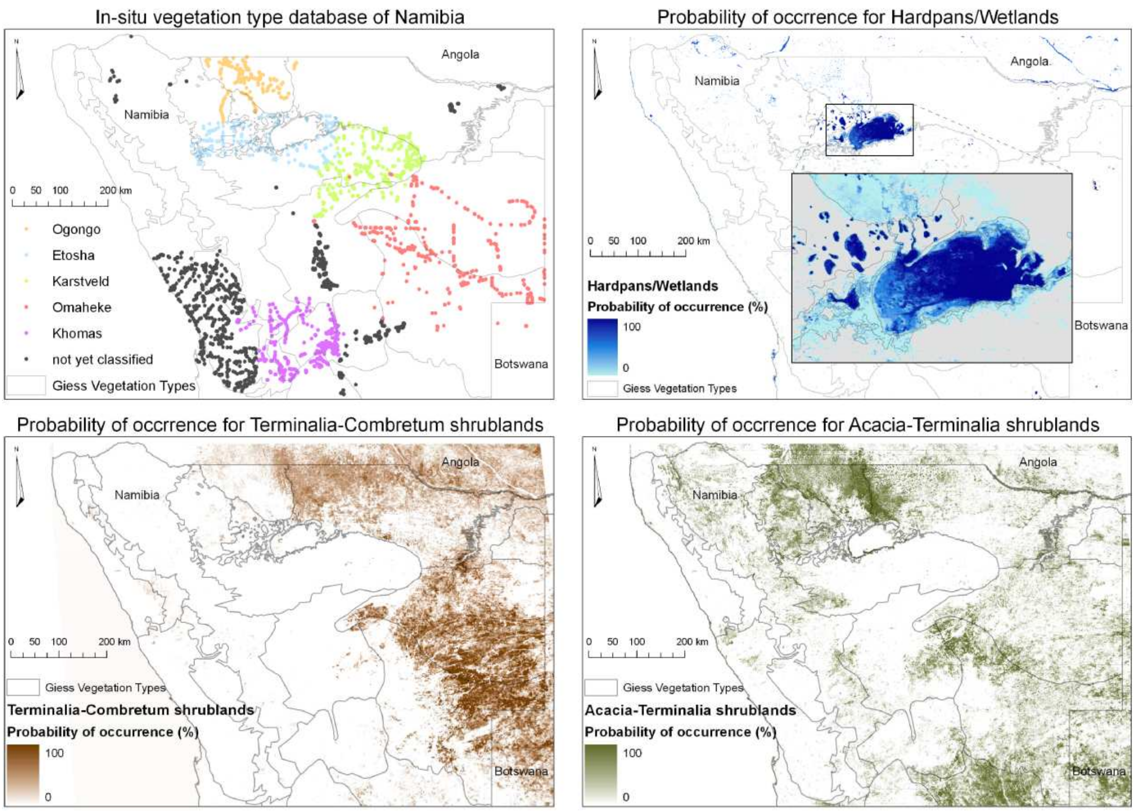


Figure 6-4 Fuzzy vegetation type maps through the integration of existing *in-situ* vegetation type databases in Namibia collected during various field surveys. Field data used as training samples from different field campaigns are displayed in (a). The probability of occurrence for hardpans and wetlands are shown in (b). The improvement of the vegetation type map after Giess (1971) are exemplarily illustrated for the vegetation types of *Terminalia-Combretum* shrublands (c) and *Acacia-Terminalia* shrublands (d).

6.3.5 Assessment of variable importance

A key advantage of the random forest framework is the implementation of an internal computation of the variable importance for each candidate predictor based on the OOB observations as mean decrease of accuracy. The RF variable importance measures provide insights on what covariates are most useful in respect to the predictive model derived. As an example, the variable importance ranking for hardpans (Figure 6-6a) and *Terminalia – Combretum* savannas are analysed. It is obvious that the high-ranked MODIS time series features indicate specific ecologic processes and dynamics that characterise and define the ecotype being mapped.

For example, for mapping seasonal flooded wetlands (hardpans), within the 30 high-ranked inter-annual statistics (derived from temporally segmented time series features), the most counted features are the middle infrared (MIR) and blue bands with 12 and seven counts. The dominance of the MIR and BLUE reflectance bands suggests that reflectance characteristics of bright saline soils and water absorption are the main features to separate this class. Temporal dependencies can be exposed by analysing the segment statistics. For example, the mean range of the blue band in segment one (Jan. – Mar, BLUE_range_1_mean) features the flooding process of the hardpans in the peak rainy season.

A different feature ranking is apparent for the *Terminalia-Combretum* savannas (Figure 6-6b). The enhanced vegetation index (EVI) is with 18 counts of the 30 top-ranked features the most important candidate predictor variable. As this class is characterised after LCCS as *broadleaved (40 - (20 - 10) %) medium high shrubland with an herbaceous layer*, the inclusion of the differences in the phenological cycle of different vegetation types seems to be crucial for the classification performance.

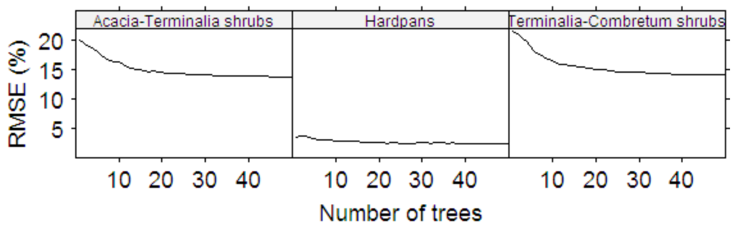


Figure 6-5 RMSE of random forest regression results based on out-of-bag estimates.

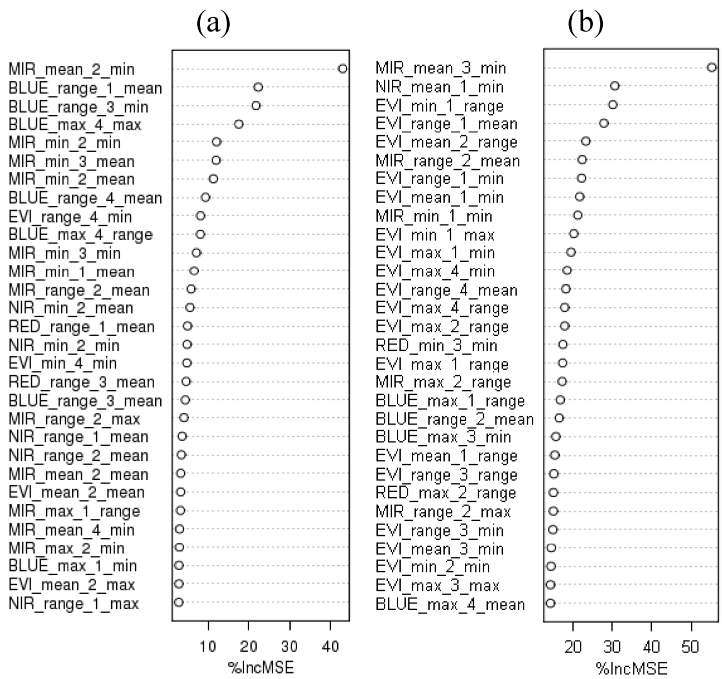


Figure 6-6 Exemplary variable importance ranking for hardpans (a) and *Terminalia-Combretum* savannas (b) of the mean increase of mean square error (MSE) or mean decrease of accuracy for the 30 high-ranked predictor variables.

6.4 Conclusions

The integration of available environmental data sources on different scales in a comparable format is a key task for the future development of earth observation initiative in less investigated and remote regions. The integrated use *in-situ* databases and coarse earth observation data used for mapping land-cover and vegetation types in Namibia, affected positive feedbacks for the local scale and wall-to-wall perspectives.

On the one hand, for local scale applications, the thematic information and accuracy of geo-data on vegetation cover and distribution could be increased by using inter-annual satellite-based phenological observations. On the other hand, the *in-situ* reference database used could possibly refine and correct the results on large-area land-cover applications in southern African savanna ecosystems. Strengths and weaknesses of the FAO LCCS for regional land-cover studies could be particularly highlighted. So, the usually given physiognomic-structural class assignments of Namibian savanna types in the main global land-cover datasets (open grasslands) need to be revised with regard to class assignments with higher proportions of woody vegetation.

Mapping in arid and semi-arid environments is closely linked with increased spatiotemporal variations of the classification result. The results on the effect of different composite and sample periods, usually applied in global mapping initiatives, showed that short composite period lengths take the short-term variations of subtropical climate conditions into account. Further on, it has been demonstrated that image classification feature sets capturing one single growing season are evidently insufficient to produce reliable land-cover classifications. The implementation of inter-annual time series features in the classification procedure should be focused in future regional mapping initiatives in savanna ecosystems. The analysis of more than one growing season could help to capture land-cover types with a high spatiotemporal variability in order to improve the accuracy of savanna-related classes in the global land-cover datasets.

The analyses on the variable importance measurements proved to be helpful in order to understand the predictive value of temporal features in the classification process. Further on, importance measures support an optimisation of the classification performance. The classification of time series often deals with a high number of features. The reduction of the feature space on the most relevant features can enhance the classification performance in terms of computation time.

The potential of integrating *in-situ* vegetation databases with MODIS time series analyses for the improvement and update of vegetation maps in Namibia was demonstrated. The use of probability maps of vegetation distribution accounts for a realistic spatial representation of the heterogeneous semi-arid vegetation structure and an appropriate cartographic representation of seasonal land cover types. Further research has to be conducted in order to analyse the potential of general land-cover maps and probability-based vegetation maps as management tool for an (a) operational land-use monitoring, (b) the evaluation of ecosystem services and biodiversity assessments, (c) and the update, refinement, and standardisation of existing vegetation distribution databases.

7 Synthesis: Concepts for integrated land-cover assessments

Abstract - The literature review (section 2.1.2) showed that ecosystem cycling and functioning in savannas is not fully understood. In particular the effects of environmental cues like fire, herbivory and rainfall on vegetation distribution patterns and dynamics are insufficiently analysed. However, one of the determinants of woody vegetation cover was proved to be annual precipitation (Sankaran *et al.*, 2005). The factors influencing mapping performance have poorly been analysed while using multi-temporal (annual and inter-annual) satellite time series. Moreover, no alternatives were provided by the remote sensing community to improve the mapping performance in semi-arid regions, either in regional studies, or in global mapping initiatives. Data standardisation and integration in mapping applications play a key role in the process of removing uncertainties in land-cover datasets in semi-arid savannas.

In respect of the main research topics of ***integrated vegetation type mapping*** (data mining), ***standardisation of earth observation data*** (standardisation), and ***implications for improved large-area monitoring*** (improvement) of land-cover mapping methodologies (as introduced in chapter one), here, a synthesis summary of research is given. The main conclusions of this dissertation are presented resulting in a number of arisen tasks and questions for future research. As a result of the integrated use of biological, biogeographic, and remote sensing methodologies and paradigms, conclusions and some final remarks are proposed in order to make a contribution towards the future development of integrated land-cover assessments in semi-arid ecosystems. Further research needs are defined aiming at the provision of improvements and alternatives for the large-area land-cover assessment and monitoring using multi-temporal satellite imagery.

7.1 Summary

Integrated vegetation type mapping – Even though a number of vegetation surveys have been conducted in Namibia (Burke & Strohbach, 2000) with the integrated use of *in-situ* and remotely sensed data (Strohbach & Jürgens, 2010), there is a lack of consistent and area-wide geo-information on vegetation cover and vegetation type distribution.

Given the first task of developing an integrated vegetation type mapping framework, an *in-situ* vegetation database was integrated in a multi-stage image classification framework. The up-scaling from point to 250 m MODIS scale was performed using an intermediate segmentation of Landsat imagery. Including data with an intermediate resolution enabled (a) the reference of the training data to spectrally homogeneous landscape features and (b) the extension of the training database.

The use of temporally segmented time series features proved to be advantageous for mapping in seasonal environments with distinct dry and rainy seasons. Temporally segmented metrics resulted in a reduction of data noise such as artefacts due to cloud cover and missing values. The classification of annual time series features resulted in accuracies ranging from 0.87 to 0.91 (kappa coefficients from 2001 to 2007), while including inter-annual time series features (2001 to 2007) improved the accuracy significantly (user's = 94.86 %, kappa = 0.93).

It could be shown that, even if 'coarse' satellite data (250 m × 250 m) were used, very similar vegetation types (in terms of plant species composition) could be distinguished. Further insight into the predictive strength of different derivates of time series features was provided using variable importance measurements which have been implemented in the random forest classification framework. The comparison of phenological (date-related) features with spectral (intensity-related) features resulted in a general preference of intensity-related features. Analyses of variable importance ranking position of the MODIS spectral bands indicated an increased significance of surface reflectance properties (e.g. soil properties section 3.4.2).

In general, the integrated use and assessment of machine-learning techniques (Random Forests) and remotely sensed image features holds significant potential for multi-dimensional data analyses. In particular, an increase in knowledge can be expected for the development of an ecologically-based understanding of remote sensing parameters. Using variable importance functionalities while classifying satellite time series, would lead to an increased understanding of the bio-physical relevance of each feature in the set. This would result in (a) the improvement of the classification accuracy by choosing appropriate features, (b) the reduction of the feature space and thus less computational costs, and finally (c) the possibility to assess environmental cues for the classification of semi-natural semi-arid vegetation types.

Spatial-explicit vegetation type mapping can be very problematic in heterogeneous landscapes, because only few pixels contain pure information on a single class (e.g. life forms, physiognomic-structural classes, or vegetation communities). Fuzzy logic can thus be an alternative to 'hard' classifications. The fuzzy mapping of vegetation types (section 6.3.4) demonstrated significant improvements of the existing national vegetation map developed by Giess (1971). Improvements can be stated as (a) a realistic spatial distribution of fine scattered semi-arid vegetation types, (b) the detection of transition zones between different vegetation types, and (c) the possibility to detect dynamic land-cover classes and different land-cover stages (e.g. flooding and dry situations in wetlands and salt pans, or photo-synthetically active or dry grasslands).

Standardisation of earth observation data – About 20 % of the earth's surface is covered by the savanna biome. Currently, there are several uncertainties in the assessment and spatial characterisation of these semi-arid vegetation types. The main occurring problems can be summarised as (a) the application of different nomenclatures and classification schemes in mapping initiatives and thus, the non-compatibility of land-cover datasets, (b) the (existence of) non-transparent and inflexible reference databases (e.g. vegetation and land-cover descriptions from various field campaigns), and little experiences among the scientific remote sensing and land-cover mapping community with the integration of mapping standards in remote sensing applications.

Regarding the objective to assess the capabilities of the standardisation of earth observation data, a translation of *in-situ* data with a phytosociologic nomenclature into the FAO land cover classification system was conducted. The result was a flexible thematic legend including multiple levels of thematic detail, such as the phytosociologic vegetation type legend and the physiognomic-structural legend. The purpose of LCCS could be confirmed to objectively describe all land-cover components and include them in the class description. The usefulness of LCCS for the application on savanna vegetation structure data was proved, since its typical coexistence of trees, shrubs, and herbaceous life forms as well as the floristic composition is implementable in a single legend (section 4.5.1). Thus, LCCS considers the layering characteristics of savannas. Comparisons of the land-cover classes in terms of spectral separability (B-distance) showed that the separability increases with increasing vegetation coverage and vegetation height. It was found, that the traceability a map legend by using LCCS classifiers is useful for the interpretation of distinct differences of separability distance values between specific classes, and for the understanding the environmental context (e.g. soil properties and phenological cycling) and its impact on the sensor measurement (section 4.5.2).

Comparisons of LCCS with the classification scheme after Edwards (1983) pointed on still existing differences between land-cover nomenclatures. However, including environmental attributes in the legend definition improved the understanding of the satellite measurement. Some limitations became apparent concerning the class definition boundaries, e.g. restrictions in the definition of subsequent layers and the ranges of the categories *cover* and *height* (section 4.5.3).

Uncertainties of using LCCS for regional land-cover studies could be highlighted. ‘Open grasslands’ for example, is the usually given physiognomic-structural class assigned of southern African savanna types in the main global land-cover datasets. Regarding the legend translation in section 4.5.1, the physiognomic-structural class labelling in the global datasets needs to be revised with regard to class assignments with higher proportions of woody vegetation.

In summary, LCCS can be understood as a ‘common land-cover language’ and can assist to improve the communication between different scientific disciplines. For instance in semi-arid regions, a high degree of inter-disciplinary communication is required. Standardisation is an essential step to increase the applicability of land-cover data.

Implications for improved large-area monitoring – Savannas exhibit the highest uncertainties in global mapping initiatives. The determinants of comparatively low classification accuracies have not been quantitatively analysed yet. Possible solutions and adaptations for technical design of future land-cover mapping and monitoring initiatives have not been discussed yet, in order to overcome the cartographic uncertainties and increase the comparability of multi-temporal map products in the future. Solutions for the improvement of large-area monitoring of savannas were evaluated. In this context the effects of temporal compositing and varying observation periods for large-area land-cover mapping were analysed in the test site of the north-eastern Kalahari in Namibia.

The comparison of mapping accuracies from multi-temporal time series classifications of different observation periods revealed a significant inter-seasonal variability in the overall accuracy level. Using inter-annual assemblies of MODIS time series features resulted in increasing mapping accuracies for all land-cover classes. Different ANOVA runs that are comparing the effect of observation period and composite length on the class-specific accuracy (section 5.3.1 and 5.3.2), as well as piecewise linear regressions between mapping accuracies and precipitation amounts derived from TRMM observations (section 5.3.4) proved the following statements:

i. Mapping accuracy increases with increasing observation period.
ii. Using small composite period lengths leads to increased mapping accuracies.
iii. The relationship between mapping accuracy and observation period was found to be a function of precipitation input and the magnitude of change between land-cover stages.

Special attention is given to the last statement since it implies a dependency of the accuracy of a land-cover map from (a) the observation period and (b) the variability of the condition of a certain land-cover type.

Precipitation is the controlling parameter. Breakpoints derived from piecewise linear regressions between cumulative precipitation and cumulative mapping accuracies cause the occurrence of a 'broken-stick' distribution. This means that the mapping accuracy increases until the peak of the rainy season is reached (256.7 mm ± 146.9 mm) and (a) levels out with slightly decreased second slopes if inter-annual time series are used for classification or (b) either levels out or even decreases in the second half of the growing period (section 5.3.4).

Several implications arise for global monitoring mechanisms if these findings are transferred to the setup of 'traditional' land-cover mapping designs which were (a) usually designed for an annual update of the global land-cover, and (b) differ in the compositing technique applied (e.g. GlobCover with bi-monthly composites and MODIS land cover collection 5 with monthly composites). Future mapping improvements can be obtained by using longer observation periods in combination with small composite lengths in savannas.

The variations in class-specific accuracies of annually produced land-cover maps seem to be too high in semi-arid ecosystems so that further land change analyses obviously cannot be based on such maps. The use of multi-annual feature sets with the smallest possible composite length is suggested for classification, to increase the comparability and integrity of semi-arid land-cover maps (section 5.4.2).

Improvement of the regionalisation of plant communities in Namibia –In fact, the long-term goal of a national update and thematic improvement of the current vegetation databases in Namibia has not been accomplished yet, e.g. by the scientific network within the BIOTA Southern Africa project within the last ten years. Within this thesis the potential of inter-annual MODIS time series was assessed in regard to the improvement of the regionalisation of plant communities in Namibia.

Vegetation type maps were derived on two scales. First, a 'hard' classification of vegetation types was conducted on an administrative level covering the communal areas of Omaheke in north-eastern Namibia. Second, fuzzy vegetation mapping for selected vegetation types was conducted on regional scale covering the northern part of Namibia. For both applications, existing *in-situ* databases were integrated in the training process to apply a random forest classification (NE Namibia, section 6.3.1) and regression (northern Namibia, section 6.3.4) techniques on multi-annual MODIS time series.

The integration of *in-situ* data to a multi-scale framework is one innovation that is leading to improved knowledge of the regionalisation of Namibian vegetation types. Former vegetation mapping projects in Namibia were conducted by Strohbach *et al.* (2004) using *in-situ* data and Landsat imagery. Due to different acquisition dates, and thus, different land-cover stages observed in the imagery (e.g. different atmospheric, phenological, or soil moisture conditions) the resulting vegetation map had major cartographic errors (linear features due to image overlap) and low thematic accuracies. Colditz *et al.* (2007) followed a top-down approach, while using the national vegetation type map after Giess (1971) as reference to train a tree-based classifier applied on annual

MODIS time series data. However, the MODIS-derived time series metrics could improve the spatial delineation of the broad vegetation type classes, however, thematic information could not be spatially improved.

Within this dissertation a bottom-up image processing framework was developed. Starting from the *in-situ* level, point reference data was spatially up-scaled to the 250 m MODIS resolution via an intermediate Landsat scale. Based on this, MODIS time series could be trained with a thematically improved reference database (section 3.2.4). Using the multi-scale framework, a thematically detailed reference database could be combined with multi-temporal imagery, capturing important phenological properties of plant communities in the Kalahari.

More than 10,000 field observations on land-cover and vegetation compositions were collected during field surveys of the last decades (Strohbach, pers. comm.). However, due to lacking and incompletely pre-processed *in-situ* reference data on typical and dominant plant communities or major land-cover classes, only a limited number of these data can be used for supervised vegetation mapping in a remote sensing framework. Fuzzy classification approaches, as exemplary shown in section 6.3.4, can be applied on incompletely classified reference data (Figure 6-4a) if applied in a presence-absence framework. These class-wise probability maps (Figure 6-4b-c) consider (a) the realistic spatial representation of transition zones between semi-natural vegetation types, which cannot be detected by discrete classification approaches, (b) the coarse spatial resolution of the MODIS time series data through a sub-pixel representation of the target class, and (c) the possibility to keep insufficiently pre-processed reference data in the training process of the regression model.

This is particularly of interest, since the majority of the available reference data are not aggregated to a vegetation type nomenclature. The classification of species data on a community level is a precondition for their integration in a coarse scale remote sensing framework. Using the currently available reference database in a presence-absence framework implicates (a) a compromise between regionalisation and incomplete ground information and (b) a speeding of the geographic regionalisation process of vegetation types on a regional to national scale.

The results of the fuzzy mapping approach of the vegetation types in the Kalahari follow the coarse delineation of the national vegetation map after Giess (1971). They allow for a more detailed phytosociological differentiation of the vegetation types defined by Giess.

7.2 Conclusions

The semi-natural savannas in Namibia are important natural resources for agricultural planning, rangeland management and tourism. Their existence and integrity is a precondition for the comparatively high biodiversity rates in southern Africa. The different types of savannas in Namibia hold a number of important ecosystem services such as water provision, eco-tourism, and food production through large-scale livestock farming.

The key issue of this dissertation was to improve vegetation mapping applications with the integrated use of earth observation data. The pre-conditions and methodological issues are described in the following. The spatial distribution of savanna vegetation types in Namibia is well documented at local scales and is still an important long-term research issue. Remote sensing techniques play an important role for the basic mapping and monitoring of vegetation types. However, the operational monitoring of land-cover in the arid and semi-arid savannas is challenging. The main reasons can be seen in (a) the complex configuration of the vegetation structure and thus a high degree of landscape heterogeneity, (b) the temporal variability of vegetation activity and phenological cycling due to variable climate conditions, and (c) the influence of other non-linear processes such as grazing, fire, and population dynamics. The effects of these processes and interactions between them have not been understood yet. Research has to be addressed to (a) evaluate the influence of such processes on land-cover information products and (b) find ways how to treat them in remote sensing based land-cover mapping frameworks.

Following the research framework and assumptions stated in section 1.2, the main objectives of this dissertation were methodically to conduct an ***integrated vegetation type mapping*** framework and the ***standardisation of earth observation data*** which have been collected in the interdisciplinary research initiative of the BIOTA Southern Africa project. The results were further used to discuss the ***implications for improved global monitoring***. Based on the main research questions of this dissertation, some concluding remarks and suggestions towards ***concepts for integrated land-cover assessments*** are summarised in the following.

7.2.1 How can satellite data classification techniques be adapted to semi-arid environments in order to account for the temporal requirements in mapping seasonal semi-arid environments?

The analysis of single acquisition dates proved to be disadvantageous for mapping semi-arid environments with alternating wet and dry seasons. For example, Strohbach *et al.* (2004) evaluated the limitations of the use of single acquisition dates of Landsat imagery for mapping Kalahari vegetation types. Most limitations emerged due to the lack of temporal information. The integration of multi-temporal features is a precondition for a reliable statistical separation of deciduous vegetation types.

The adaption of classification techniques to semi-arid environments thus comprises (a) the **adjustment of the feature space** to the temporal properties of the ecosystem and (b) the **choice of an adjusted classification method** which accounts for a multi-temporal feature space comprising a large number of predictor variables. Acceptable accuracies were achieved when analysing annual MODIS time series. It could be proved that a temporal resolution of 16 days achieved increased mapping accuracies compared to feature sets with longer composite lengths. Thus, the temporal resolution is an important property of the satellite data for large-area mapping purposes in semi-arid environments.

Precipitation is the controlling variable on vegetation activity in southern African savannas (Sankaran *et al.*, 2005). Regressions between mapping accuracy and precipitation showed a positive correlation for all land-cover classes. The spatial distribution of arid and semi-arid land-cover classes in the land-cover classification will thus always depend on the intensity of the previous and corresponding rainy season. A general question is how to avoid the mapping of different land-cover stages instead of land-cover types. One solution can be the inclusion of the class-specific inter-annual characteristics in the classification process. This is achieved by integrating multi-annual time series in the feature space.

An important precondition for multi-annual land-cover classifications is a certain degree of stability for the observed land-cover due to modifications and transformation processes. Fire and grazing are apparent processes but do not change the general vegetation type occurrence in the short-term. The integration of (long-term) multi-annual satellite time series in the feature set enables the integration of inter-annual characteristics (of specific vegetation types) in the classification process. Inter-annual feature sets avoid classification errors due to temporal shifts of the rainy season and differing magnitudes of the rainfall amount.

The accuracy of land-cover maps with a structural-physiognomic legend also benefits from longer than annual observation periods. Inter-annual variations of class assignments can be reduced if the characteristics of different life forms (grasses, shrubs, and trees) are considered in the classification process. Image classification feature sets have thus to be adjusted to the temporal patterns of vegetation dynamic characteristics of the target classes. For the typical vegetation structure in savannas it can be summarised as follows:

Herbaceous vegetation: The occurrence of herbaceous life forms such as grasses and forbs is a key property of a savanna landscape. The biomass production of annual and perennial grass and forbs species is characterised by a rapid response to environmental cues. Most importantly, the photosynthetic activity is controlled by soil moisture and thus rainfall. Regarding the high spatiotemporal variability of the precipitation amounts in southern African savannas, the delineation of herbaceous classes in a land-cover map is characterised by strong inter-annual variations between different growing seasons. For example, grasslands can be mapped as barren ground in a poor rainfall year if no vegetation activity is apparent. Using time series data, the characteristics of high inter-annual ranges of the vegetation activity (expressed by high ranges of the vegetation indices of NDVI and EVI) can be a crucial feature for the statistical separation of the grassland class. Comparisons of classification results using annual and inter-annual feature sets showed that herbaceous land-cover classes benefit most from the lengthening of the observation period. This implies that the inclusion of multiple land-cover stages of a certain land-cover class (e.g. for grasslands: barren ground in the dry season and extensive coverage of green vegetation in the rainy season) reduces the uncertainties in the classification process.

Woody vegetation: Shrubs are the dominant life form in the central Kalahari in north-eastern Namibia, covering most parts of the area. The dominance of trees increases northwards following the NW-SE rainfall gradient. Compared to herbaceous life forms, the phenological cycling of woody vegetation is less sensitive to precipitation. Moreover, an important trigger of biomass production not to be underestimated is temperature. Phenomenon of pre-rain flush was observed for some characteristic vegetation types (e.g. *Acaciatea* savannas) and for some tree species (e.g. *Pterocarpus angolensis*, Shackleton, 2002). This implies for remote sensing mapping applications, that reliable results can be achieved for woody land-cover classes, when at least annual time series metrics are used. The comparative assessments of the effects of composite length on mapping accuracy showed little variations in the shrubland and woodland classes. This means that the currently used feature assemblies (annual time series with 16-day to three-monthly composite lengths) in the global mapping initiatives fulfil the requirements of mapping woody vegetation in southern African savannas.

The occurrence of a 'pure pixel' of a single life form is unlikely in savannas when MODIS data of a spatial resolution of 250 m are used. Most of the globally available land-cover datasets were developed using a 'hard' classification scheme. Climatic gradients of the occurrence and mixture of life forms can be well represented if continuous nomenclatures are used. The applicability and potential of fractional cover mapping of non-vegetated areas, woody and herbaceous vegetation was demonstrated by Gessner *et al.* (2009). However, there is a lack of continuous land-cover datasets at global scales. The MODIS derived vegetation continuous field product (VCF, Hansen *et al.*, 2003) is the most recent available dataset, but a frequent update is missing.

Time series data are an indispensible component for mapping seasonal environments. With regard to the temporal requirements of mapping savanna vegetation types, the technical and methodological needs for mapping initiatives are increasing. Examples are given by the increasing number of (freely available) satellite time series data (e.g. ten years of MODIS data, MERIS, and the future Sentinel satellites) that will lead to increasing challenges to the image classification methods used. The consequence is an increasing number of predictor variables. The reduction of the feature space on the 'most important and effective' features is highly recommended.

Different implications arise for supervised and unsupervised classification approaches. The 300 m GlobCover product is based on an unsupervised clustering, whereas the 500 m MODIS product is derived from a tree-based ensemble classifier. Unsupervised methods have distinct limitations in terms of feature space adjustment, since the classification process does not allow any interaction with the input feature set. Supervised methods allow the adaption of the feature set on the target classes. More specifically, tree-based ensemble bagging and boosting methods such as Random Forest and Adaboost allow an assessment of the predictive effectiveness through specific variable importance functions (e.g. decrease of accuracy and decrease of Gini index in Random Forest).

In particular within the *R* community, the number of statistical packages comprising predictive model adaption and variable importance assessments is increasing. Examples are given with the *caret* package (Kuhn, 2008). Within this dissertation feature sets with more than 400 MODIS metrics were used for classification. In order to (a) enhance the classification performance in terms of computation time and (b) reduce the feature set on the most important predictive variables, a variable importance analysis before the classification process is highly recommended.

Regarding the choice of the classification method, the application of ensemble methods proved to be advantageous for mapping vegetation types with very similar phenological characteristics. As stated in the introducing hypotheses, satellite time series metrics are 'weak' predictor variables for the classification of savanna vegetation types. The iterative classification process of Random Forests coupled with the randomisation of features and training data (a) allows the evaluation of the predictive value of a high number of variables, (b) reduces the probability of over-fitting in the classification tree ensemble (and is thus transferrable to a large area), and is (c) applicable on high dimensional feature sets. In summary, non-parametric tree-based ensemble models are suggested for mapping dynamic landscapes with satellite time series metrics. Other models (e.g. Support Vector Machines or Adaptive Boosting) have not been compared in this dissertation but bear a very high potential for further research activities.

7.2.2 How can earth observation information from *in-situ* to regional scales be harmonised into a standardised framework for the regionalisation of ecosystem diversity assessments?

The current and future mapping and monitoring initiatives do not have to fear a lack of earth observation data. On the one hand, a number of more than 10,000 *in-situ* samples were collected during the last decades in the case of Namibia. On the other hand, the data policies of international space agencies are moving towards a free access of satellite databases (e.g. the USGS Landsat database covering 35 years of Landsat acquisitions). This bears a very high potential but also challenges towards a future regionalisation of ecosystem diversity assessments.

Harmonisation and standardisation is the most important task in this context and has to be done prior to data analyses. The results of this dissertation showed, that the integrated use of harmonised data achieved benefits at both *in-situ* and regional scales. On the one hand, satellite time series data can measure slight phenological differences over multiple years. The precondition is a good temporal resolution. On the other hand, these hyper-temporal dynamics including phenological, land-use, and other ecosystem-specific processes can only be evaluated with the use of *in-situ* knowledge.

A key characteristic of multi-scale applications is that multiple research communities are included in the data analyses process. For example, the regionalisation of thematic biodiversity or ecosystem diversity information (such as the spatial distribution of vegetation types) from *in-situ* to a national scale is a long-term process in Namibia and includes research activities from various disciplines, such as vegetation and population ecology (field surveys and phytosociological classifications), geo-informatics (data management), and remote sensing (satellite data acquisition pre-processing, and analyses). In this context, the standardisation of data processing and analysis techniques and the harmonisation of existing data is an essential step and cause positive feedbacks in multiple directions (sciences related to botany, biogeography, and remote sensing).

Standards have to be developed and respected on the thematic and technical side of the regionalisation framework. The thematic standardisation includes the definition of rules for up- and downscaling of thematic information (e.g. class definitions and classification schemes). Existing databases and processing workflows have to be translated and harmonised according to these standards. Table 7-1 summarises the methods applied in this dissertation based on (a) the data which have been adapted from other research projects and (b) the methods used for the integration in a multi-scale land-cover assessment framework. The methodological overview may be seen as a suggestion for the further refinement of technical land-cover standards from *in-situ* to global scales.

Table 7-1 Overview of integration methods for *in-situ* and remote sensing data in the land-cover and vegetation type mapping framework.

Analysis step	*Data and Products*	*Method*	*Description*
Field survey	Vegetation structure: species abundance and richness, height, growth form, cover abundance (cover in %) of every height class	Braun-Blanquet (Braun-Blanquet, 1964; Van Der Maarel, 1975)	Vegetation survey follows a systematic description and definition of plant communities, sampling requires predefined homogeneous plant distribution patterns, e.g. for this study 20 m × 50 m
	Environmental attributes: slope, terrain type, aspect, stone, cover estimation, lithology, (parent material), erosion, severity, surface sealing/crusting, disturbances	Braun-Blanquet	
Phytosociologic classification	Phytosociologic vegetation type legend (vegetation types legend)	TWINSPAN (Hill, 1979) and PHYTOTAB (Westfall *et al.*, 1997)	vegetation associations were clustered after characteristic, differentiating and typical species
Translation of the vegetation type legend into LCCS	Physiognomic-structural legend	LCCS-2 (Di Gregorio, 2005)	Translation of the vegetation type legend (vegetation structure and environmental attributes) into the standard LCCS legend using LCCS classifiers and modifiers
Up-scaling from *in-situ* to MODIS scale	Training database	Image segmentation (Definiens, 2009)	Object-oriented segmentation performed on Landsat data to retrieve homogeneous objects based on similar reflectance settings to regionalise in-situ information of the training points on a coarse MODIS pixel size.
Feature space generation	MODIS time series metrics	Temporal segmentation; TIMESAT (Jönsson & Eklundh, 2004)	Derivation of predictor variables for land-cover classification based on (a) temporal segmentation and (b) derivation of phenological metrics
Classification and regression techniques applied on MODIS time series data	Land-cover map based on a flexible legend (vegetation type legend and physiognomic-structural legend), probability-based legend	Tree-based ensemble methods (Random Forest, Liaw & Wiener, 2002)	Supervised classification of MODIS time series metrics by applying the Random Forest framework

Note that the analyses steps of the field survey and phytosociologic classification were conducted by the National botanical research Institute of Namibia (NBRI).

Thematic standards have to be defined on multiple scales and should ideally start at the *in-situ* level. For vegetation type mapping purposes the use of the Braun-Blanquet approach proved to be applicable since the method is flexible and can be adapted to any vegetative land-cover type. Regarding the minimum mapping unit of the satellite imagery, special attention has to be given to the definition of homogeneous areas. Moreover, the method is implemented in the FAO-LCCS. Therewith it is possible to apply a global scale land-cover standard to regional and local mapping initiatives which enables the comparability and usefulness of regional to national vegetation maps to a broader user community. LCCS is the most accepted land-cover standard. This dissertation proved that results from the Braun-Blanquet approach can be integrated in LCCS since most of the LCCS classifiers are considered in the system.

The phytosociologic classification, as used in this dissertation, proved to be useful for the translation of a 'fine' floristic legend into a 'coarse' land-cover legend. The application of specific tools for the clustering of vegetation databases after characteristic, differentiating and typical species is suggested, rather than a subjective translation of land-cover classes. An integration of phytosociologic methods into LCCS-2 has not been finalised yet. However, the development of the Land Cover Classification System is an ongoing process. So, a new version has been released (LCCS-3) called Land Cover Meta Language (LCML). Further research has to be addressed to assess and report the integrity of phytosociological applications into LCCS-3.

This dissertation proved the applicability of LCCS-2 for vegetation type mapping in the Kalahari. Moreover, the application of LCCS is highly recommended of any kind of biogeographic mapping since it (a) fosters the thematic improvement of a 'common land-cover language' and (b) fosters the establishment and perception of LCCS among a wider scientific community. Most importantly, LCCS-based geo-information products include the necessary metadata (such as the list of *classifiers* and *modifiers*) for an objective comparison and synergistic use of different map products.

Defining **technical standards** for the regionalisation of ecosystem diversity assessments is challenging, since data processing has to be adapted to the geographic scale and the research aims. At this stage, some general recommendations can be provided for mapping in southern African savannas.

Vegetation types and general land-cover classes can often be distinguished reliably, by using the temporal patterns of land-cover classes. Therefore, the processing of satellite imagery using the best possible temporal resolution is recommended. The analysis of single-date imagery (except of VHR data) involves the risk of mapping land-cover stages instead of land-cover classes. Multi-temporal imagery is often available on coarse spatial resolutions, e.g. derived from the sensors of MODIS, SPOT-VEGETATION, and MERIS. However, the freely accessible high resolution data archives (Landsat) and upcoming missions such as the Sentinel-II and the Landsat Data Continuity Mission (LDCM) have the potential to provide high resolution time series imagery. Satellite time series with high and very high resolutions will significantly increase land-cover mapping in arid and semi-arid ecosystems.

As mentioned in section 5.5, the following issues have to be regarded for a future setup of technical standards for mapping in highly dynamic savanna environments:

i. *The consideration of specific phenological characteristics of multiple life-form layers.* A dominant photosynthetic activity of a single vegetation layer (e.g. grasses vs. woody vegetation) can cause incorrect class assignments. The inclusion of multiple growing seasons and the adaption of the feature setup on the vegetation phenology can result in more stable vegetation classifications.

ii. *The consideration of a complex vegetation structure within a categorical classification scheme using coarse scale time series.* In many remote sensing datasets a classified pixel represents a mixture of different life forms in this region. While high resolution time series data is getting increasingly available, it will enhance the accurate mapping of semi-arid land-cover types, characterised by a very high degree of patchiness of the vegetation structure. The application of a standardised and flexible classification system, as defined in the FAO Land Cover Meta Language (LCML, LCCS-3), can further improve the thematic representation of heterogeneous landscapes in coarse scale land-cover products.

iii. *The generation, provision, and standardisation of in-situ reference databases.* Southern African savannas are poorly populated, with a limited infrastructure, and they are poorly analysed concerning vegetation composition, distribution, and phenology. Including *in-situ* reference data in the framework of large-area land-cover mapping initiatives is limited so far. The collection and translation of available botanical databases in LCCS/ LCML cause the most effective transfer of land-cover information from local to regional and global scales.

Savannas are ecosystems with a high magnitude of **biocomplexity**. Following the concept of biocomplexity according to Cadenasso *et al.* (2006), the complicatedness can be stratified in the scale of measurement (space and time) and the component specification (patterns and processes). Remote sensing land-cover initiatives have a long experience in integrative data assimilations such as multi-temporal and multi-resolution classification concepts. The patterns and processes causing the spatial distribution and dynamics of certain land-cover types have not been considered in the 'classical' sense of a land-cover mapping framework.

For example, the results in sections 3.4.2 and 5.3.4 point on the precipitation pattern to be a key parameter controlling the condition of land cover, and thus, the final mapping result. The lengthening of the observation period can adapt the properties of remote sensing data on the ecosystem-specific behaviour. Such kind of 'customised' remote sensing can help to integrate the temporal component of structural complexity of an ecosystem in the technical design of a land-cover mapping framework.

7.2.3 What are the strengths and weaknesses of the FAO and UNEP Land Cover Classification System (LCCS), if applied in a standardised bottom-up mapping framework in a semi-natural semi-arid environment?

The FAO and UNEP LCCS provides a scale independent system for the classification of land-cover. LCCS is the international most accepted classification system and has been submitted to become an international standard through the Technical Committee of the International Organisation for Standardisation (ISO TC 211).

In this dissertation LCCS was applied in a semi-arid ecosystem of the Kalahari in north-eastern Namibia. A phytosociologic legend of vegetation types was translated into LCCS, based on the implemented classifiers and modifiers. In this process, the highest thematic degree of detail was used for the derivation of the classification system of Kalahari vegetation types showing the capabilities and limitations of LCCS.

In general, LCCS proved to be applicable in semi-arid environments. One of the strengths is the layering system which allows the definition of different strata of life-forms. The hierarchical system of LCCS-2 allows for the detailed description of the vegetation structure of the tree, shrub, and herbaceous layers, which accounts for the typical savanna vegetation structure, the coexistence of trees, shrubs, and grasses. Some limitations became apparent due to the static definition of class boundaries, such as the classifiers *height* and *coverage*. Further problems raised in the definition of a third vegetation layer. In particular, it is not possible to select a second shrub layer with all combinations of the first and second vegetation layer. Those limitations may be solved with the release of LCCS-3. Further research has to be conducted to transfer the vegetation type legend into the Land Cover Meta Language (LCCS-3).

Beside the definition of classifiers related to vegetation structure, the definition of environmental attributes such as soil type and lithology proved to be advantageous for mapping Kalahari vegetation structure. The refinement of the attribute list of each vegetation type class helps the remote sensing analyst to 'understand' the remote sensing measurement and the class signature. A simple example is given with soil type, which is a determining geo-factor for the occurrence of distinct vegetation types. In this context, the future implementation of the attribute soil colour would be helpful since it is a proxy for certain soil properties (e.g. white sands with lower loam and silt content than red sands) and a dominant feature in a remote sensing image covering open savanna landscapes.

A more unrestrictive and thus flexible rule set could increase the classification conciseness and usefulness of the land-cover data for different user communities. Further limitations became apparent for the set of leaf type classifiers in the second modular level. The plant group of *Acaciatea* is characterised by a fine leafed physiognomy. An appropriate classifier should be added to the leaf type section.

The definition of environmental and specific technical attributes proved to be essential for the detailed characterisation of semi-arid vegetation types. The possibility of defining the specific technical attribute *Floristic Aspect* addresses an important interface to include botanical and biodiversity databases. However, there is potential for a further refinement of the floristic attributes which can fasten the acceptance of LCCS among the botanical sciences. On the one hand, the application of LCCS on the *in-situ* level enables the integration of (existing) extensive *in-situ* reference databases in regional to national land-cover mapping initiatives. On the other hand, the visibility, transparency, usefulness, and applicability of botanic vegetation maps can be increased by integrating a floristic classification scheme in LCCS.

The classifiers approach is capable of identifying woodlands, shrublands, and grasslands as the main structural land cover types. Comparisons of the class boundaries of the main vegetation layers (section 2.2.2.2) showed deviations of the class boundaries in open land-cover types (grasslands) and transitional zones (open woodlands). However, the comparability of historic maps or geo-information products applied on other nomenclatures with LCCS-based land-cover maps may be limited in the mentioned open landscapes and addresses need for future research on legend translation and harmonisation in semi-arid environments.

7.2.4 How can the global remote sensing community learn from regional biodiversity and vegetation type assessments and *vice versa*?

As mentioned above the application of LCCS enables the almost complete integration of class descriptions in a flexible legend system. If LCCS-based mapping is applied in a bottom-up framework, several opportunities will arise for the global remote sensing community. A precondition is that LCCS acts as ‘common land-cover ontology’ between scientific disciplines.

First, international panels on monitoring and system analyses of the earth demand the integration of *in-situ* databases in global earth observation mechanisms. The integrated use of *in- situ*-, Landsat-, and MODIS time series data in the Kalahari study area could successfully up-scale plant community diversity data on regional scales. In this context, the transformation and integration of an existing biological database in coarse scale mapping framework highlights the potential for future operational and standardised monitoring mechanisms. Land-cover is an important parameter for the nine social benefit areas of GEOSS. The future use of ‘flexible’ land-cover observations will adapt land-cover observations to the requirements of the implementation plan. This will also help to support the reporting of the CBD focal area *Indicators for immediate testing*, as stated by Muchoney (2008):

- *“Status and trends of the components of biological diversity”*
- *“Trends in extent of selected biomes, ecosystems, and habitats”*
- *“Trends in abundance and distribution of selected species”*
- *“Sustainable use of forest area, agricultural, and aquaculture ecosystems under sustainable management numbers and cost of alien invasions”*

- *"Ecosystem integrity and ecosystem goods and services"*
- *"Connectivity/fragmentation of ecosystems"*
- *"Incidence of human-induced ecosystem failure"*
- *"Health and well-being of people living in biodiversity"*

Regional biodiversity and vegetation type assessments can directly contribute to the demands of the Group on Earth Observation's Biodiversity Observation Network (GEO-BON). The up-scaling and mapping of biodiversity indicators and regionalisation of in-situ measurements are core tasks of GEO-BON. Standardised monitoring mechanisms allow for an inter-comparison of maps on biodiversity and ecosystem diversity. The attribution of currently existing land-cover datasets is missing. However, the collection, adaption, and transfer of regional biodiversity assessments can assist the transfer of biodiversity observations (e.g. species and remotely sensed images) to observation products (tabulations, models, GIS databases, and attribute maps), which are claimed as the key input data for the GEO-BON user interface (section 2.1.1.1).

Second, the global remote sensing community can extend its reference databases by recognising and integrating regional standardised biodiversity and ecotype assessments in calibration and validation activities.

This dissertation points on the uncertainties of mapping savannas at global scales and suggests possible solutions for improvements by adapting the remotely sensed feature sets and classification methods. However, a certain degree of uncertainty and classification errors (not to be underestimated) may be due to the selection of incorrect training data or error propagation in the training process. Often, training data of time series data for global map productions are based on secondary high resolution data. Reference data are often derived from existing vegetation databases. Training data may therefore include the 'traditional' or 'common' class descriptions according to the 'typical occurrence' and perception of the savanna biome. Interestingly, major parts of savannas in the Kalahari basin (including Namibia, South Africa, Botswana, Angola, and Zambia) are classified as *open grasslands* in the global land-cover datasets.

The reality is divergent. Major parts of the vegetation structure in the Kalahari do not agree to the common definition of savanna biome, e.g. as defined in Eiten (1992): Lands *"where the presence of a dominating herbaceous ground layer (or forbs, or dwarf shrubs) is required"*. Long-term anthropogenic land use caused a creeping bush encroachment process with the result that dense shrubs became the dominating vegetation layer. This situation is frequently being recorded and reported during numerous vegetation surveys and may be an important information source for calibration and validation activities for the global remote sensing community.

7.3 Future research needs

This dissertation faces various issues on conceptional data integration, methodological concepts in the field of time series analyses, and applied mapping purposes. Even though the major aims of research were achieved, a number of further research questions rose during the studies. Some important future research needs are listed in the following.

New vegetation map of Namibia – One of the main long-term research goals of the BIOTA Southern Africa network and in particular the National Botanical Research Institute of Namibia is an update of the official vegetation map of Namibia after Giess (1971) which is based on extensive field survey data. Remote sensing will be an essential tool for a national update of the current vegetation type distribution. In particular satellite time series capture the phenological characteristics of certain vegetation types. The potential of eight years of MODIS data was demonstrated in the Hereroland test site. Future research addresses the extension of the mapping area onto the national level. As mentioned before, the main workload will be the pre-processing of more than 10,000 available field samples, which comprises the integration in a common database, the classification on an aggregated vegetation type classification scheme, and the integration in a GIS environment.

Harmonisation of national vegetation databases – National vegetation maps in southern Africa differ considerably in terms of thematic detail, minimum mapping unit, input data, and the classification system applied. Examples are Namibia and South Africa, where different vegetation type units were assigned across the border but in the same landscape. Remote sensing data allow for an objective and transboundary monitoring of land. The potential of integrating satellite time series data in recent mapping initiatives has not been regarded yet to come up with border-line effects in regional to sub-continental biogeographic geo-information products.

Phenology of semi-arid vegetation types – The semi-natural savannas in Namibia are of important economic value since they provide the source of the main ecosystem services, e.g. food production by large-scale life stock farming, water supply, and eco-tourism. The monitoring of the phenological cycling of the savanna types in Namibia has not been conducted yet on regional scale. The results of the time series analyses in this dissertation showed that phenological patterns are the key properties for the delineation of vegetation types with sometimes slightly differing species compositions. However, a systematic assessment and monitoring of vegetation phenology in Namibia is lacking. Future research has to address the setup of *in-situ* phenology observatories since currently existing phenology information are based on a few studies. Moreover, the potential of the satellite-based derivation of phenological metrics has not been assessed on vegetation type level. A few studies were carried out in southern Africa using AVHRR and MODIS time series (Steenkamp *et al.*, 2008). However these phenological observations are based on coarse biome classes. The assessment of phenological cycling for plant communities will (a) deepen the knowledge and understanding of vegetation ecology in semi-arid systems and (b) provide important information on the state of the vegetation which can be used for environmental planning on national level.

Variable importance in remote sensing – The application of machine learning techniques in remote sensing has been increasing over the last decades. The free access of statistical packages (e.g. through the *R* community) fostered the application of multiple statistical learning algorithms rather than using just one method of classification purposes (e.g. tree-based learning algorithms like Random Forests and Adaboost, Support Vector Machines, or Neural Networks). This bears a high potential for (a) the combination of multiple models, (b) further inter-comparisons of their performance on remote sensing data. Especially the last statement has been poorly assessed in the field of remote sensing. Variable importance measurements are widely used in other scientific disciplines, such as genetics and molecular biology working with multivariate data. With regard to increasing satellite data archives or synergistic multi-sensor applications the assessment of existing variable importance measurement techniques can help to increase the understanding of the predictive value of satellite measurements and can help to optimise classification workflows in terms of accuracy and computation time.

LCCS experience – The FAO and UNEP Land Cover Classification System is going to become the global standard for land-cover and has its own requirement to classify any land-cover type on the earth. However, the suitability of LCCS for bottom-up vegetation type assessments could be demonstrated within this dissertation. Further research has to be carried out to (a) evaluate the classifier and modifier settings in other dynamic landscapes, and to report the applicability and general experience of the Land Cover Meta Language (LCCS-3) among the remote sensing community.

Global land-cover mapping – Great progress has been made in the harmonisation and standardisation of land-cover monitoring mechanisms, including calibration and validation, accuracy assessment, legends, and classification techniques. Within this process, the highest uncertainties were observed in mapping heterogenic and dynamic landscapes (e.g. savannas, wetlands, and agricultural lands). Within this dissertation some suggestions and limitations for improvements were provided for mapping in semi-arid landscapes. Future research has to address the integration of dynamic processes of ecosystems in the mapping framework, e.g. phenology, changing hydrological situations, and human interactions. Further conceptional research has to be conducted to define standards for the definition and integration of land-cover states in global land-cover projects.

References

Acocks, J. P. H. (1988): Veld types of South Africa, Memoirs of the Botanical Survey of South Africa, Government Printer.

Ahlqvist, O. (2007): 'In search of classification that supports the dynamics of science: the FAO Land Cover Classification System and proposed modifications', Environment and Planning B: Planning and Design, 35(1), pp. 169-186.

Allsopp, N., Laurent, C., Debeaudoin, L. M. C. and Igshaan Samuels, M. (2007): 'Environmental perceptions and practices of livestock keepers on the Namaqualand Commons challenge conventional rangeland management', Journal of Arid Environments, 70(4), pp. 740-754.

Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E. (1976): 'A Land Use And Land Cover Classification System For Use With Remote Sensor Data', Geological Survey Professional Paper 964, United States Department of the Interior and USGS, 964, p. 28.

Archer, K. and Kimes, R. (2008): 'Empirical characterisation of random forest variable importance measures', Computational Statistics & Data Analysis, 52(4), pp. 2249-2260.

Archer, S. (1995): 'Tree-Grass Dynamics in a Prosopis-Thornscrub Savanna Parkland - Reconstructing the Past and Predicting the Future', Ecoscience, 2(1), pp. 83-99.

Archibald, S. and Scholes, R. J. (2007): 'Leaf green-up in a semi-arid African savanna – separating tree and grass responses to environmental cues', Journal of Vegetation Science, pp. 583-594.

Arndt, D. S., Baringer, M. O. and Johnson, M. R. (2010): 'State of the Climate in 2009', Bulletin of the American Meteorological Society, 91(6), pp. 1-244.

Asner, G. P., Levick, S. R., Kennedy-Bowdoin, T. Y., Knapp, D. E., et al. (2009): 'Large-scale impacts of herbivores on the structural diversity of African savannas.', Proceedings of the National Academy of Sciences of the United States of America, 106(12), pp. 4947-52.

Augustine, D. J. (2003): 'Spatial heterogeneity in the herbaceous layer of a semi-arid savanna ecosystem', pp. 319-332.

Bankhofer, U. and Vogel, J. (2008): Datenanalyse und Statistik, Wiesbaden, Gabler.

Bartholomé, E. and Belward, A. S. (2005): 'GLC2000: a new approach to global land cover mapping from Earth observation data', International Journal of Remote Sensing, 26(9), pp. 1959-1977.

Beckage, B., Platt, W. J. and Gross, L. J. (2009): 'Vegetation, fire, and feedbacks: a disturbance-mediated model of savannas.', The American naturalist, 174(6), pp. 805-818.

Beerling, D. J. and Osborne, C. P. (2006): 'The origin of the savanna biome', Global Change Biology, 12(11), pp. 2023-2031.

Begon, M. E., Harper, J. L. and Townsend, C. R. (1998): Ökologie, Sauer, K. (ed.), Heidelberg, Spektum Verlag.

Bertin, R. I. (2008): 'Plant Phenology And Distribution In Relation To Recent Climate Change', The Journal of the Torrey Botanical Society, 135(1), pp. 126-146.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., et al. (2008): GLOBCOVER: Products Description and Validation Report.

Bonan, G. B., Levis, S., Kergoat, L. and Oleson, K. W. (2002): 'Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models', Global Biogeochemical Cycles, 16(2).

Bossard, M., Feranec, J. and Otahel, J. (2000): CORINE land cover technical guide – Addendum 2000, European Environment Agency.

Botha, L. (1999): 'History of drought in Namibia', Agricola, 10, pp. 5-12.

Braun-Blanquet, J. (1964): Pflanzensoziologie - Grundzüge der Vegetationskunde, Springer.

Breiman, L. (2001): 'Random Forests', Machine Learning, 45, pp. 5-32.

Breiman, L. (1999): Random Forests - Random Features. Technical Report 567, Berklay.

Breiman, L., Cutler, A., Liaw, A. and Wiener, M. (2006): 'randomForest: Breiman and Cutler's Random Forests for Classification and Regression', [online] Available from: http://stat-www.berkeley.edu/users/breiman/RandomForests, (Date Accessed: 22-07-2010).

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984): Classification and Regression Trees, Wadsworth, Wadsworth Pacific Grove CA.

Brodley, C. E. and Friedl, M. A. (1997): 'Decision Tree Classification of Land Cover from Remotely Sensed Data', Remote Sensing of Environment, 61, pp. 399-409.

Brown, G. (2004): 'Diversity in Neural Network Ensembles', IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Computer Society.

Bucini, G. and Hanan, N. P. (2007): 'A continental-scale analysis of tree cover in African savannas', Global Ecology and Biogeography, 16(5), pp. 593-605.

Burke, A. and Strohbach, B. J. (2000): 'Review: Vegetation Studies in Namibia', Dinteria, (26), pp. 1-24.

Burke, A. and Strohbach, B. J. (2000): 'Review: Vegetation Studies in Namibia', (26), pp. 1-24.

Burnett, C. (2003): 'A multi-scale segmentation/object relationship modelling methodology for landscape analysis', Ecological Modelling, 168(3), pp. 233-249.

Cadenasso, M., Pickett, S. and Grove, J. (2006): 'Dimensions of ecosystem complexity: Heterogeneity, connectivity, and history', Ecological Complexity, 3(1), pp. 1-12.

Callicott, J. B., Rozzi, R., Delgado, L., Monticino, M., et al. (2007): 'Biocomplexity and conservation of biodiversity hotspots: three case studies from the Americas.', Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 362(1478), pp. 321-33.

Camberlin, P., Martiny, N., Philippon, N. and Richard, Y. (2007): 'Determinants of the interannual relationships between remote sensed photosynthetic activity and rainfall in tropical Africa', Remote Sensing of Environment, 106, pp. 199 - 216.

CBD (2006): Global Biodiversity Outlook 2, Global Biodiversity, Montreal, [online] Available from: http://www.cbd.int/doc/gbo/gbo2/cbd-gbo2-en.pdf, (Date Accessed: 25-01-2010).

CBD (2005): Handbook of the Convention on Biological Diversity Including its Cartagena Protocol on Biosavety, 3rd edition, 3rd ed. Montreal, [online] Available from: http://www.cbd.int/handbook/, (Date Accessed: 12-08-2010).

Chan, J., Huang, C. and DeFries, R. (2001): 'Enahanced algorithm performance for land cover classification from remotely sensed data using bagging and boosting', IEEE Transactions on Geoscience and Remote Sensing, 39, pp. 693-695.

Chan, J. and Paelinckx, D. (2008): 'Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery', Remote Sensing of Environment, 112(6), pp. 2999-3011.

Chaudhuri, G., Borwankar, J. D. and Rao, P. (1991): 'Bhattacharyya distance-based linear discriminant function for stationary time series', Communications in Statistics, 20, pp. 2195-2205.

Childes, S. L. (1989): 'Phenology of nine common woody species in semi-arid, deciduous Kalahari Sand vegetation', Vegetatio, 79, pp. 151-163.

Cihlar, J., Manak, D. and Dapos, I. (1994): 'Evaluation of compositing algorithms for AVHRR data over land', IEEE Transactions on Geoscience and Remote Sensing, 32(2), pp. 427-437.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A. and Schwartz, M. D. (2007): 'Shifting plant phenology in response to global change', Trends in Ecology and Evolution, 22(7).

Colditz, R. R., Conrad, C., Wehrmann, T., Schmidt, M. and Dech, S. (2008): 'TiSeG: Flexible software tool for time-series generation of MODIS data utilizing the quality assessment science data set', IEEE Transactions on Geoscience and Remote Sensing, 46, p. 3296–3308.

Colditz, R. R., Keil, M., Strohbach, B. J., Geßner, U., et al. (2007): 'Vegetation structure mapping with remote sensing time series: Capabilities and improvements', In 32nd International Symposium on Remote Sensing of Environment, June 25 - 29 2007, Costa Rica, ISRSE, pp. 1-4.

Colwell, R. (1998): 'Balancing the biocomplexity of the planet's living systems: a 21st century task for science', BioScience, 48, pp. 786-787.

Colwell, R. (1999): 'Complexity and connectivity: a new cartography for science and engineering', In AGU Fall Meeting, San Francisco, American Geoscience Union.

Comber, A. J. (2008): 'The separation of land cover from land use using data primitives', Journal of Land Use Science, 3(4), pp. 215-229.

Comber, A. J., Fisher, P. F. and Wadsworth, R. A. (2005): 'What is land cover?', Environment and Planning B Planning and Design, Pion, 32, p. 199–209.

Congalton, R. G. and Green, K. (2009): Assessing the accuracy of remotely sensed data: principles and practices, 2nd ed. Boca Raton, CRC Press.

Conrad, C., Colditz, R, Dech, S. and Vlek, P. L. G. (2007): 'Combining temporal segmentation and classification trees for time series classification', In Fourth International Workshop on the Analysis of Multitemporal Remote Sensing Images, Multitemp, 18–20 July 2007, Leuven.

Cord, A., Conrad, C., Schmidt, M. and Dech, S. (2010): 'Standardized FAO-LCCS land cover mapping in heterogeneous tree savannas of West Africa', Journal of Arid Environments, 74(9), pp. 1083-1091.

De Pauw, E. (1996): Agroecological zones of Namibia. First appoximation. Technical Cooperation Project TCP/Nam/6611(A) "Initiation of National Agroecological Zoning Procedures". Technical report No 1, Rome.

De Pauw, E. and Coetzee, M. E. (1999): 'Production of an Agro-Ecological Zones Map of Namibia (first approximation). Part I: Condensed Methodology', Agricola, 10, pp. 27-31.

De Pauw, E., Coetzee, M. E., Calitz, A. J., Beukes, H. and Vits, C. (1999): 'Production of an Agro-Ecological Zones Map of Namibia (first approximation). Part II: Results', Agricola, 10, pp. 33-43.

Definiens (2009): Definiens Developer 7. Reference Book,.

Defourny, P., Bicheron, P., Brockman, C., Bontemps, S., et al. (2009): 'The first 300 m global land cover map for 2005 using ENVISAT MERIS time series: a product of the GlobCover system', In Proceedings of the 33th International Symposium of Remote Sensing of Environment, May 04-08, Stresa, pp. 1-4.

DeFries, R. S., Hansen, M. C., Townsend, J. R. G., Janetos, A. C. and Loveland, T. R. (2000): 'A new global 1-km dataset of percentage tree cover derived from remote sensing', Global Change Biology, 6, pp. 247-254.

DeFries, R., Hansen, M. and Townshend, J. R. G. (1995): 'Global discrimination of land cover types from metrics derived from AVHRR pathfinder data', Remote Sensing of Environment, 54(3), pp. 209-222.

DeFries, R., Hansen, M., Townshend, J. and Sohlberg, R. (1998): 'Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers', International Journal of Remote Sensing, 19(3141-3168).

Delbart, N., Letoan, T., Kergoat, L. and Fedotova, V. (2006): 'Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004)', Remote Sensing of Environment, 101(1), pp. 52-62.

Di Gregorio, A. (2005): Land Cover Classification System. Classification concepts and user manual. Software version 2, 2nd ed. Rome, FAO, Environment and Natural Resources Series number 8.

Di Gregorio, A. and Jansen, L. J. M. (2000): Land Cover Classification System (LCCS): Classification Concepts and User Manual:LCCS version 1.0, System, Environment and Natural Resources Service,GCP/RAF/287/ITA Africover - East Africa Project and Soil Resources, Management and Conservation Service,Food and Agriculture Organization of the United Nations.

Dierssen, K. (1990): 'Einführung in die Pflanzensoziologie (Vegetationskunde)', Darmstadt, Wissenschaftliche Buchgesellschaft.

Dougill, A. J. and Thomas, A. D. (2004): 'Kalahari sand soils: spatial heterogeneity, biological soil crusts and land degradation', Land Degradation Development, 15(3), pp. 233-242.

Dougill, A. and Trodd, N. (1999): 'Monitoring and modelling open savannas using multisource information: analyses of Kalahari studies', Global Ecology and Biogeography, pp. 211-221.

Du Plessis, W. P. (1999): 'Linear regression relationships between NDVI, vegetation and rainfall in Etosha National Park, Namibia', Journal of Arid Environments, pp. 235-260.

Du Plessis, W. P., Breedenkamp, G. J. and Trollope, W. S. W. (1998): 'Response of herbaceous species to a degradation gradient in the western region of Etosha National Park, Namibia', Koedoe, 41, pp. 19-28.

Duchemin, B., Berthelot, B., Dedieu, G., Leroy, M. and Maisongrande, P. (2002): 'Normalisation of directional effects in 10-day global syntheses derived from VEGETATION/SPOT: II. Validation of an operational method on actual data sets', Remote Sensing of Environment, 81(1), pp. 101-113.

D'Odorico, P., Caylor, K., Okin, G. S. and Scanlon, T. M. (2007): 'On soil moisture–vegetation feedbacks and their possible effects on the dynamics of dryland ecosystems', Journal of Geophysical Research, 112(4), pp. 1-10.

D'Odorico, P., Laio, F. and Ridolfi, L. (2006): 'A probabilistic analysis of fire-induced tree-grass coexistence in savannas.', The American naturalist, UChicago Press, 167(3), p. E79-E87.

Eamus, D., Myers, B., Duff, G. and Williams, D. (1999): 'Seasonal changes in photosynthesis of eight savanna tree species.', Tree physiology, 19(10), pp. 665-671.

Edwards, D. (1983): 'A broad-scale structural classification of vegetation for practical purposes', Bothalia, 14, pp. 705-712.

Ehleringer, J. and Bjorkman, O. (1977): 'Quantum Yields for CO2 Uptake in C3 and C4 Plants: Dependence on Temperature, CO2, and O2 Concentration', Plant Physiology, 59(1), pp. 86-90.

Eiten, G. (1992): 'How names are used for vegetation', Journal of Vegetation Science, 3(3), pp. 419-424.

Eiten, G. (1968): 'Vegetation Forms - A classification of stands of vegetation based on structure, growth form of the components, and vegetative peridicity', Boletim do Instituto de Botânica, 4(1), pp. 1-88.

Ewald, J. (2003): 'A critique for phytosociology', Journal of Vegetation Science, John Wiley & Sons, 14(2), p. 291–296.

FAO (1993): Global and National Soils and Terrain Digital Databases (SOTER), Rome.

FAO (2009): 'Global Land Cover Network. Global Alliance for Standard Multi-Purpose Land Cover Data Production', Land and Water Division of the Food and Agriculture Organization of the United Nations.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., et al. (2007): 'The Shuttle Radar Topography Mission', Reviews of Geophysics, 45.

Fisher, J. and Mustard, J. (2007): 'Cross-scalar satellite phenology from ground, Landsat, and MODIS data', Remote Sensing of Environment, 109(3), pp. 261-273.

Friedl, M. A. and Brodley, C. E. (1997): 'Decision tree classification of land cover from remotely sensed data', In: Remote Sensing Of Environment, Vol., 61(3), pp. pp399-409.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., et al. (2002): 'Global land cover mapping from MODIS: algorithms and early results', Remote Sensing of Environment, 83, pp. 287 - 302.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., et al. (2010): 'MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets', Remote Sensing of Environment, 114(1), pp. 168-182.

Frost, P. (1996): The ecology of miombo woodlands (Africa), Jakarta, CIFOR.

Funk, C. and Budde, M. (2009): 'Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe', Remote Sensing of Environment, 113(1), pp. 115-125.

Gessner, U., Conrad, C., Hüttich, C., Keil, M., et al. (2008): 'A multi-scale approach for retreiving proportional cover of life forms', In Proceedings of the IEEE International Geoscience and RemoteSensing Symposium, Boston USA, July 6 - 8, pp. 3-4.

Gessner, U., Klein, D., Conrad, C., Schmidt, M. and Dech, S. (2009): 'Towards an automated estimation of vegetation cover fractions on multiple scales: Examples of Eastern and Southern Africa', In Proceedings of the 33th International Symposium of Remote Sensing of Environment, May 04-08, Stresa, pp. 1-4.

Giess, W. (1971): 'A preliminary vegetation map of South West Africa', Dinteria, 4, pp. 1-114.

Giglio, L., Descloitres, J., Justice, C and Kaufman, Y. J. (2003): 'An Enhanced Contextual Fire Detection Algorithm for MODIS', Remote Sensing of Environment, 87(2-3), pp. 273-282.

Gillespie, T. W., Foody, G. M., Rocchini, D., Giorgi, A. P. and Saatchi, S. (2008): 'Measuring and modelling biodiversity from space', Progress in Physical Geography, 32(2), pp. 203-221.

Giri, C., Zhu, Z. and Reed, B. (2005): 'A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets', Remote Sensing of Environment, 94(1), pp. 123-132.

Gislason, P., Benediktsson, J. and Sveinsson, J. (2006): 'Random Forests for land cover classification', Pattern Recognition Letters, 27(4), pp. 294-300.

GLCN (2009): 'Land Cover Ontology. LCCS 2', [online] Available from: http://www.glcn.org/ont_1_en.jsp, (Date Accessed: 25-01-2010).

Gopal, S. and Woodcock, C. (1994): 'Theory and Methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets', Photogrammetric Engeneering and Remote Sensing, 60(2), pp. 181-188.

Goward, S. N. and Prince, S. D. (1995): 'Transient effects of climate on vegetation dynamics: satellite observations', Journal of Biogeography, 22, pp. 549-563.

Green, J. L., Hastings, A., Arzberger, P, Ayala, F. J., et al. (2005): 'Complexity in Ecology and Conservation: Mathematical, Statistical, and Computational Challenges', BioScience, 55(6).

Group on Earth Observation (GEO) (2005): 'GEOSS 10-year implementation plan reference document', [online] Available from: http://www.earthobservations.org/documents/10-Year Plan Reference Document.pdf, (Date Accessed 8-03-2011).

Guerschman, J. P., Hill, MJ, Renzullo, L. J., Barrett, D. J., et al. (2009): 'Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors', Remote Sensing of Environment, 113(5), pp. 928-945.

Hagolle, O., Lobo, A., Maisongrande, P., Cabot, F., et al. (2005): 'Quality assessment and improvement of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images', Remote Sensing of Environment, 94(2), pp. 172-186.

Hall, L. S., Krausman, P. R. and Morrison, M. L. (1997): 'The habitat concept and a plea for standard terminology', Wildlife Society Bulletin, 25(1), pp. 173-182.

Hanan, N. P., Sankaran, M., Ratnam, J., Dangelmayr, G., et al. (2006): Biocomplexity in African Savannas. Scientific Reports on Research Projects undertaken in the Kruger National Park during 2005, Pretoria.

Hansen, L. K. and Salamon, P. (1990): 'Neural network ensembles', IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(10), pp. 993-1001.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., Carroll, M., et al. (2003): 'Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm', Earth Interactions, 7(10).

Hansen, M. C., DeFries, R. S., Townshend, J. R. G. and Sohlberg, R. (2000): 'Global land cover classification at 1 km spatial resolution using a classification tree approach', International Journal of Remote Sensing, 21(6 & 7), pp. 1331-1364.

Hansen, M. C., DeFries, R., Townshend, J. R. G., Sohlberg, R., et al. (2002): 'Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data', Remote Sensing of Environment, 83, pp. 303 - 319.

Hansen, M. C. and Reed, B. (2000): 'A comparison of the IGBP DISCover and University of Maryland 1 km global land cover products', International Journal of Remote Sensing, 21(6 & 7), pp. 1365-1373.

Hastie, T., Tibshhirani, R. and Friedmann, J. (2009): The Elements of Statistical Learning-Data Mining, Inference, and Prediction, Hastie, T., Tibshhirani, R., and Friedmann, J. (eds.), New York, Springer.

Hennekens, S. (1996): 'TURBO(VEG) Software Package for Input, Processing, and Presentation of Phytosociological Data. User's Guide.', Wageningen.

Herold, M. (2009a): Assessment of the Status of the Development of the Standards for the Terrestrial Climate Variables. Land Cover. T9 Report, Rome, [online] Available from: http://www.fao.org/gtos/doc/ECVs/T09/T09.pdf.

Herold, M. (2009b): 'Building saliency, legitimacy, and credibility towards operational global and regional land cover observations and assessments', In Proceedings of the International Symposium of Remote Sensing of the Environment, 4 -8 May, Stresa, Italy.

Herold, M., Latham, J. S., Di Gregorio, A. and Schmullius, C. C. (2006): 'Evolving standards in land cover characterization', Journal of Land Use Science, 1(2), pp. 157-168.

Herold, M., Liu, X and Clarke, K. C. (2003): 'Spatial Metrics and Image Texture for Mapping Urban Land Use', Photogrammetric Engineering & Remote Sensing, 69(9), pp. 991-1001.

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A. and Schmullius, C. (2008): 'Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets', Remote Sensing of Environment, 112(5), pp. 2538-2556.

Herold, M., Schmullius, C. and Arino, O. (2008): 'Building Saliency, Legitimacy, and Credibility towards operational global Land Cover Observations', In Proceedings of the 2nd MERIS (A)ATSR User Workshop 22-26 September, Frascati, Italy, ESA SP-666, pp. 1-5.

Herold, M., Woodcock, C. E., Di Gregorio, A., Belward, A. S., et al. (2006): 'A joint initiative for harmonization and validation of land cover datasets', IEEE Transactions on Geoscience and Remote Sensing, 44, pp. 1719-1727.

Herold, M., Woodcock, C. E., Loveland, T.R., Townshend, J., et al. (2008): 'Land-Cover Observations as Part of a Global Earth Observation System of Systems (GEOSS): Progress, Activities, and Prospects', IEEE Systems Journal, 2(3), pp. 414-423.

Herold, M., Woodcock, C., Stehman, S., Baret, F., et al. (2009): 'The GOFC-GOLD/CEOS land cover harmonization and validation initiative: Technical design and implementation framework', In Proceedings of the 33th International Symposium of Remote Sensing of Environment, May 04-08, Stresa.

Higgins, S. I., Bond, W. J. and Trollope, W. S. W. (2000): 'Fire, resprouting and variability: a recipe for grass-tree coexistence in savanna', Journal of Ecology, British Ecological Society, 88(2), pp. 213-229.

Hill, M. O. (1979a): Decorana — a Fortran program for detrended correspondence analysis and reciprocal averaging., Systematics, E. A. (ed.), Cornel University Ithaca.

Hill, M. O. (1979b): TWINSPAN: a FORTRAN Program for Arranging Multivariate Data on an Ordered Two-Way Table by classification of the Individuals and Attributes., Ecology And Systematics, Cornell University.

Hird, J. and Mcdermid, G. (2009): 'Noise reduction of NDVI time series: An empirical comparison of selected techniques', Remote Sensing of Environment, 113(1), pp. 248-258.

Ho, T. K. (1995): 'Random decision forests', In Third International Conference on Document Analysis and Recognition ICDAR95, IEEE Computer Society, pp. 278-282.

Hodges, J. (2002): 'MODIS MOD12 Land Cover and Land Cover Dynamics Products User Guide. Validation of the consistent-year V003 MODIS land cover product', [online] Available from: http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD12, (Date Accessed 8-03-2010).

Holben, B. N. (1986): 'Characteristics of maximum-value composite images from temporal AVHRR data', International Journal of Remote Sensing, 7(11), pp. 1417-1434.

Hollingsworth, A., Uppala, S., Klinker, E., Burridge, D., et al. (2005): 'The transformation of earth-system observations into information of socio-economic value in GEOSS', Quarterly Journal of the Royal Meteorological Society, 131(613), pp. 3493-3512.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., et al. (2002): 'Overview of the radiometric and biophysical performance of the MODIS vegetation indices', Remote Sensing of Environment, 83, pp. 195 - 213.

Huete, A., Didan, K., Miura, T., Rodriguez, P., et al. (2002): 'Overview of the radiometric and biophysical performance of the MODIS vegetation indices', Remote Sensing of Environment, 83, pp. 195 - 213.

Hüttich, C., Herold, M., Strohbach, B. J. and Dech, S. (2010): 'Integrating in-situ-, Landsat-, and MODIS data for mapping in Southern African savannas: Experiences of LCCS-based land-cover mapping in the Kalahari in Namibia', Environmental monitoring and assessment, E-pub, [online] Available from: http://www.springerlink.com/content/8274n71420g2t2ww/.

Hüttich, C., Gessner, U., Herold, M., Strohbach, B. J., et al. (2009): 'On the Suitability of MODIS Time Series Metrics to Map Vegetation Types in Dry Savanna Ecosystems: A Case Study in the Kalahari of NE Namibia', Remote Sensing, 1(4), pp. 620-643.

Jankowitz, W. and Venter, H. (1987): 'Die plantgemeenskappe van die Waterberg-platopark', Madoqua, 15, pp. 97-146.

Jansen, L. J. M. and Di Gregorio, A. (2002): 'Parametric land cover and land-use classi cations as tools for environmental change detection', Agriculture, Ecosystems and Environment, 91, pp. 89-100.

Jeltsch, F., Milton, S. J., Dean, W. R. J., Rooyen, N. V. and Moloney, K. A. (1998): 'Modelling the impact of small-scale heterogeneities on tree-grass coexistence semi-arid savannas', Journal of Ecology, British Ecological Society, 86(5), pp. 780-793.

Jeltsch, F., Weber, G. E. and Grimm, V. (2000): 'Ecological buffering mechanisms in savannas: A unifying theory of long-term tree-grass coexistence', Plant Ecology, Springer, 150(1), pp. 161-171.

Jolly, W. M. and Running, S. W. (2004): 'Effects of precipitation and soil water potential on drought deciduous phenology in the Kalahari', Global Change Biology, 10, pp. 303-308.

Joubert, D. F., Rothauge, A. and Smit, G. N. (2008): 'A conceptual model of vegetation dynamics in the semiarid Highland savanna of Namibia, with particular reference to bush thickening by Acacia mellifera', Journal of Arid Environments, 72(12), pp. 2201-2210.

Jung, M., Henkel, K., Herold, M. and Churkina, G. (2006): 'Exploiting synergies of global land cover products for carbon cycle modeling', Remote Sensing of Environment, 101(4), pp. 534-553.

Justice, C. O., Vermote, E. F., Masuoka, E., Wolfe, R. E., et al. (2002): 'An overview of MODIS Land data processing and product status', Science, 83, pp. 3 - 15.

Justice, C. O., Vermote, E., Townshend, J. R. G., DeFries, R., et al. (1998): 'The Moderate Resolution Imaging Spectroradiometer (MODIS): Land Remote Sensing for Global Change Research', IEEE Transactions on Geoscience and Remote Sensing, 36(4), pp. 1228-1249.

Justice, C. O., Townshend, J. R. G., Vermote, F., Masuoka, E., et al. (2002): 'An overview of MODIS Land data processing and product status', Science, 83, pp. 3 - 15.

Jönsson, P. and Eklundh, L. (2004): 'TIMESAT - a program for analyzing time-series of satellite sensor data', Computers & Geosciences, 30(8), pp. 833-845.

Jürgens, N., Burke, A., Sally, M. and Jacobson, K (1997): 'Desert', In Vegetation of Southern Africa.

Kaptué Tchuenté, A. T., De Jong, S. M., Roujean, J-L, Favier, C. and Mering, C. (2010): 'Ecosystem mapping at the African continent scale using a hybrid clustering approach based on 1-km resolution multi-annual data from SPOT/VEGETATION', Remote Sensing of Environment, Elsevier Inc., 115(2), pp. 452-464.

Keil, M., Gessner, U., Hüttich, C. and Colditz, R. (2010): 'Large-scale vegetation assessments in southern Africa: concepts and applications using multi-source remote sensing data', In Schmiedel, U. and Jürgens, N. (eds.), Biodiversity in southern Africa. Volume 2: Patterns and processes at regional scale, Göttingen, Windhoek, Klaus Hess Publishers, pp. 32-45.

Kent, M. (2006): 'Numerical classification and ordination methods in biogeography', Progress in Physical Geography, 30(3), pp. 399-408.

Kent, M., Gill, W. J., Weaver, R. E. and Armitage, R. P. (1997): 'Landscape and plant community boundaries in biogeography', Progress in Physical Geography, 21(3), pp. 315-353.

King, L. (1963): South African Scenery, New York, Hafner Publishing Company.

Klein, D. and Roehrig, J. (2006): 'How does vegetation respond to rainfall variability in a semi-humid west african in comparison to a semi-arid east african environment?', Proceedings of the 2nd Workshop of the EARSeL Special Interest Group on Land Use and Land Cover, Bonn, Germany, September 28–30, 2006; Volume 1, 149–156.

Krug, C. B., Elser, K. J., Hoffman, M. T., Henschel, J., et al. (2006): 'North–South cooperation through BIOTA: an interdisciplinary monitoring programme in arid and semi-arid southern Africa', South African Journal Of Science, 102, pp. 1-4.

Kuhn, M. (2008): 'caret Package', Journal Of Statistical Software, 28(5).

Lambin, E. (1999): 'Monitoring forest degradation in tropical regions by remote sensing: some methodological issues', Global Ecology and Biogeography, 8, pp. 191-198.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., et al. (2001): 'The causes of land-use and land-cover change: moving beyond the myths', Global Environmental Change, 11, pp. 261-269.

Landgrebe, D. and Biehl, L. (2001): An Introduction to MultiSpec, West Lafayette.

Laris, P. (2005): 'Spatiotemporal problems with detecting and mapping mosaic fire regimes with coarse-resolution satellite data in savanna environments', Remote Sensing of Environment, 99(4), pp. 412-424.

Latifovic, R., Zhu, Z.-L., Chihar, J., Giri, C. and Olthof, I. (2004): 'Land cover mapping of North and Central America—Global Land Cover 2000', Remote Sensing of Environment, 89(1), pp. 116-127.

Le Roux, C. L. G., Grunow, J. O., Morris, J. W., Bredenkamp, G. J. and Scheepers, J. C. (1988): 'A classification of the vegetation of the Etosha National Park', South African Journal Of Botany, 54(1-10).

Leeuwen, W. van (1999): 'MODIS Vegetation Index Compositing Approach A Prototype with AVHRR Data', Remote Sensing of Environment, 69(3), pp. 264-280.

Leimgruber, P., Christen, C. A. and Laborderie, A. (2005): 'The Impact of Landsat Satellite Monitoring on Conservation Biology', Environmental Monitoring and Assessment, 106(1-3), pp. 81-101.

Leser, H. (1972): 'Geoökologische Verhältnisse der Pflanzengesellschaften in den Savannen des Sandveldes um den Schwarzen Nossob und um Epukiro', Dinteria, 6, pp. 1-43.

Liaw, A. and Wiener, M. (2002): 'Classification and regression by random forest', R News, 2, pp. 18-22.

Loveland, T. (2008): North America Land Cover Summit, Campbell, J., Jones, K., Smith, J., and Koeppe, M. (eds.), Washington, Association of American Geographers.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., et al. (2000): 'Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data', International Journal of Remote Sensing, 21(6 & 7), pp. 1303-1330.

Ludwig, F., Kroon, H., Prins, H. H. T. and Berendse, F. (2001): 'Effects of nutrients and shade on tree-grass interactions in an East African savanna', Journal of Vegetation Science, 12(4), pp. 579-588.

Martiny, N., Richard, Y. and Camberlin, P. (2005): 'Interannual persistence effects in vegetation dynamics of semi-arid Africa', Geophysical Research Letters, American Geophysical Union, 32(24), pp. 24403-24420.

Maxwell, S. K., Hoffer, R. M. and Chapman, P. L. (2002): 'AVHRR composite period selection for land cover classification', International Journal of Remote Sensing, 23(23), pp. 5043-5059.

Mayaux, P., Bartholome, E., Belward, A and Fritz, S. (2004): 'A new land-cover map of Africa for the year 2000', Journal of Biogeography, 31, pp. 861-877.

Mayaux, P., Eva, H., Gallego, J., Strahler, A. H., et al. (2006): 'Validation of the global land cover 2000 map', IEEE Transactions on Geoscience and Remote Sensing, 44(7), pp. 1728-1739.

Mayaux, P., Massart, M., Vancutsem, C., Cabral, A., et al. (2003): 'A land cover map of africa carte de l'occupation du sol de l'afrique', [online] Available from: http://publications.jrc.ec.europa.eu/repository/handle/111111111/10815, (Date Assessed 10-03-2011).

McDermid, G. J., Franklin, S. E. and LeDrew, E. F. (2005): 'Remote sensing for large-area habitat mapping', Progress in Physical Geography, 29(4), pp. 449-474.

Mendelsohn, J., Jarvis, A., Roberts, C. S. and Robertson, T. (2002): Atlas of Namibia : A portrait of the Land and it's People, Cape Town, David Philip Publishers.

Mendelsohn, J. and Obeid, S. (2002): The Communal Lands in Eastern Namibia, Windhoek, Raison.

Michener, W. K., Baerwald, T. J., Firth, P., Palmer, M. A., et al. (2001): 'Defining and Unraveling Biocomplexity', BioScience, 51(12), pp. 1018-1023.

Muchoney, D. (2008): 'Earth observations for terrestrial biodiversity and ecosystems', Remote Sensing of Environment, 112(5), pp. 1909-1911.

Na, X., Zhang, S., Li, X., Yu, H. and Liu, C. (2010): 'Improved Land Cover Mapping using Random Forests Combined with Landsat Thematic Mapper Imagery and Ancillary Geographic Data', Photogrammetric Engeneering & Remote Sensing, 76(7), pp. 833-840.

Nagendra, H. (2001): 'Using remote sensing to assess biodiversity', International Journal of Remote Sensing, 22(12), pp. 2377-2400.

Neumann, K.., Herold, M., Hartley, A. and Schmullius, C. (2007): 'Comparative assessment of CORINE2000 and GLC2000: Spatial analysis of land cover data for Europe', International Journal of Applied Earth Observation and Geoinformation, 9(4), pp. 425-437.

Oldeland, J., Dorigo, W., Lieckfeld, L., Lucieer, A. and Jürgens, N. (2010a): 'Combining vegetation indices, constrained ordination and fuzzy classification for mapping semi-natural vegetation units from hyperspectral imagery', Remote Sensing of Environment, 114(6), pp. 1155-1166.

Oldeland, J., Dorigo, W., Wesuls, D. and Jürgens, N. (2010b): 'Mapping Bush Encroaching Species by Seasonal Differences in Hyperspectral Imagery', Remote Sensing, 2, pp. 1416-1438.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., et al. (2001): 'Terrestrial ecoregions of the world: A new map of life on earth', BioScience, 51, p. 933−938.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., et al. (2001): 'Terrestrial Ecoregions of the World: A New Map of Life on Earth', BioScience, 51(11), pp. 933-938.

Pal, M. (2005): 'Random forest classifier for remote sensing classification', International Journal of Remote Sensing, 26(1), pp. 217-222.

Pereira, H. M. and Cooper, H. D. (2006): 'Towards the global monitoring of biodiversity change.', Trends in ecology & evolution (Personal edition), 21(3), pp. 123-9.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., et al. (2005): 'Using the satellite-derived NDVI to assess ecological responses to environmental change.', Trends in ecology & evolution, 20(9), pp. 503-10.

Prasad, A. M., Iverson, L. R. and Liaw, A. (2006): 'Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction', Ecosystems, 9(2), pp. 181-199.

Privette, J. L., Tian, Y., Roberts, G., Scholes, R. J., et al. (2004): 'Vegetation structure characteristics and relationships of Kalahari woodlands and savannas', Global Change Biology, 10, pp. 281-291.

Quinlan, J. R. (1993): C4.5: Programs for Machine Learning, C45 Programs for Machine Learning, Morgan Kaufmann.

Reed, M. S., Dougill, A. J. and Taylor, M. J. (2007): 'Integrating local and scientific knowledge for adaption to land degradation: Kalahari management options', Land Degradation & Development, 18(3), pp. 237-353.

Richter, C. G. F. and Rethman, N. F. G. (1999): 'The influence of tree density on the grass layer of three semi-arid savanna types of southern Africa', Afr J Range Forage Sci, 18(1), pp. 9-18.

Robinson, E. (1976): 'Phytosociology of the Namib Desert Park, South West Africa'.

Roujean, J. L., Leroy, M. and Deschamps, P. Y. (1992): 'A Bidirectional Reflectance Model of the Earth□s Surface for the correction of Remote Sensing Data', Journal of Geophysical Research, 97(18), pp. 20,455-20,468.

Roy, D. P., Lewis, P. E. and Justice, C. O. (2002): 'Burned area mapping using multi-temporal moderate spatial resolution data—a bi-directional reflectance model-based expectation approach', Remote Sensing of Environment, 83, pp. 263 - 286.

Roy, D., Boschetti, L., Justice, C and Ju, J. (2008): 'The collection 5 MODIS burned area product - global evaluation by comparison with the MODIS active fire product', Remote Sensing of Environment, 112, pp. 3690-3707.

Running, S. W., Loveland, T. R., Pierce, L. L., Nemany, R. R. and Hunt, E. R. (1995): 'A Remote Sensing Based Vegetation Classification Logic for Global Land Cover Analysis', Remote Sensing of Environment, 51, pp. 39-48.

Sankaran, M., Hanan, N. P., Scholes, R. J., Ratnam, J., et al. (2005): 'Determinants of woody cover in African savannas.', Nature, 438(7069), pp. 846-9.

Sannier, C. A. D., Taylor, J. C., Du Plessis, W. and Campbell, K. (1998): 'Real-time vegetation monitoring with NOAA-AVHRR in Southern Africa for wildlife management and food security assessment', International Journal of Remote Sensing, 19, pp. 621-639.

Satake, M., Oshimura, K., Ishido, Y., Kawase, S. and Kozu, T. (1995): 'TRMM PR data processing and calibration to be performed by NASDA', In IEEE Conference Proceedings, Geoscience and Remote Sensing Symposium, IGARSS '95, Quantitative Remote Sensing for Science and ApplicationsApplications, Firenze, Italy, July 10–14.

Scanlon, T. M., Albertson, J. D., Caylor, K. K. and Williams, C. A. (2002): 'Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem', Remote Sensing of Environment, 82, pp. 376 - 388.

Scanlon, T. M., Caylor, K. K., Levin, S. A. and Rodriguez-Iturbe, I. (2007): 'Positive feedbacks promote power-law clustering of Kalahari vegetation.', Nature, 449(7159), pp. 209-12.

Scepan, J. (1999): 'Thematic Validation of High-Resolution Globla Land-Cover Data Sets', Photogrammetric Engeneering and Remote Sensing, 65(9), pp. 1051-1060.

Scholes, R. J. (1990): 'The Influence of Soil Fertility on the Ecology of Southern African Dry Savannas', Journal of Biogeography, Blackwell Publishing, 17(4/5), pp. 415-419.

Scholes, R. J. and Archer, S. R. (1997): 'Tree-grass interactions in savannas', Annual Review of Ecology and Systematics, Annual Reviews, 28(1), pp. 517-544.

Scholes, R. J. and Biggs, R. (2005): 'A biodiversity intactness index', Nature, 434, pp. 45-49.

Scholes, R. J., Dowty, R. P., Caylor, K, Parsons, D. A. B., et al. (2002): 'Trends in savanna structure and composition along an aridity gradient in the Kalahari', Journal of Vegetation Science, 13, pp. 419-428.

Scholes, R. J., Frost, P. G. H. and Tian, Y. (2004): 'Canopy structure in savannas along a moisture gradient on Kalahari sands', Global Change Biology, 10, pp. 292-302.

Scholes, R. J., Mace, G. M., Turner, W, Geller, G. N., et al. (2008): 'Toward a global biodiversity observing system', Science (New York, N.Y.), 321(5892), pp. 1044-1055.

Sedano, F., Gong, P. and Ferrao, M. (2005): 'Land cover assessment with MODIS imagery in southern African Miombo ecosystems', Remote Sensing of Environment, 98(4), pp. 429-441.

Sekhwela, M. and Yates, D. (2007): 'A phenological study of dominant acacia tree species in areas with different rainfall regimes in the Kalahari of Botswana', Journal of Arid Environments, 70(1), pp. 1-17.

Seni, G. and Elder, J. (2010): Ensemble methods in data mining. Improving accuracy through combining predictions, Statistics, Morgan & Claypool.

Shackleton, C. M. (2002): 'Growth patterns of Pterocarpus angolensis in savannas of the South African lowveld', Forest Ecology amd Management, 166, pp. 85-97.

Shackleton, C. M. (1999): 'Rainfall and topo-edaphic in uences on woody community phenology in South African savannas', Global Ecology and Biogeography, 8, pp. 125-136.

Sheuyange, A., Oba, G. and Weladji, R. B. (2005): 'Effects of anthropogenic fire history on savanna vegetation in northeastern Namibia.', Journal of environmental management, 75(3), pp. 189-98.

Shugart, H. H., Macko, S. A., Lesolle, P., Szuba, T. A., et al. (2004): 'The SAFARI 2000-Kalahari Transect Wet Season Campaign of year 2000', Global Change Biology, 10(3), pp. 273-280.

Simpson, J., Kummerow, C. and Adler, R. F. (1996): 'On the Tropical Rainfall Measuring Mission (TRMM)', Meteorology and Atmospheric Physics, 60(1), pp. 19-36.

Skidmore, A. K., Oindo, B. O. and Said, M. Y. (2009): 'Biodiversity Assessment by Remote Sensing', In Proceedings of the International Symposium of Remote Sensing for the Environment, 3-8 May, Stresa, Italy.

Sonderegger, D. L., Wang, H, Clements, W. H. and Noon, B. R. (2009): 'Using SiZer to detect thresholds in ecological data', Frontiers in Ecology and the Environment, Ecological Society of America, 7(4), pp. 190-195.

Steenkamp, K., Wessels, K., Archibald, S. and Maltitz, G. von (2009): 'Satellite derived phenology of southern Africa for 1985-2000 and functional classification of vegetation based on phenometrics', In Proceedings of the International Symposium of Remote Sensing for the Environment, 3-8 May, Stresa, Italy, pp. 10-14.

Strahler, A. H., Woodcock, C. E. and Smith, J. A. (1986): 'On the nature of models in remote sensing', Remote Sensing of Environment, 20(2), pp. 121-139.

Strobl, C., Boulesteix, A.-L., Zeileis, A. and Hothorn, T. (2007): 'Bias in random forest variable importance measures: illustrations, sources and a solution.', BMC bioinformatics, 8, p. 25.

Strohbach, B., Calitz, A. and Coetzee, M. (1996): 'Erosion Hazard Mapping: Modelling vegetative cover', Agricola, 8, pp. 53-59.

Strohbach, B. J. (2001): 'Vegetation Survey of Namibia', Namibia Scientific Society, 49.

Strohbach, B. J. and Jürgens, N. (2010): 'Towards a user-friendly vegetation mapping of Namibia: ground trouthing approach to vegetation mapping', In Schmiedel, U. and Jürgens, N. (eds.), Biodiversity in southern Africa 2: Patterns and processes at regional scale, Göttingen, Windhoek, Klaus Hess Publisher, pp. 46-56.

Strohbach, B. J. and Petersen, A. (2007): 'Vegetation of the central Kavango woodlands in Namibia: An example from the Mile 46 Livestock Development Centre', South African Journal of Botany, 73(3), pp. 391-401.

Strohbach, B.J., Strohbach, M., Katuahuripa, J.T and Mouton, H. D. (2004): A Reconnaissance Survey of the Landscapes, Soils and Vegetation of the Eastern Communal Areas (Otjiozondjupa and Omaheke Regions), Namibia, Windhoek.

Strohbach, J. and Petersen, A. (2007): 'Vegetation of the central Kavango woodlands in Namibia: An example from the Mile 46 Livestock Development Centre', South African Journal Of Botany, 73, pp. 391 - 401.

Sullivan, S. (1999): 'The impacts of people and livestock on topographically diverse open wood- and shrub-lands in arid north-west Namibia', Global Ecology and Biogeography, 8, pp. 257-277.

Swap, R. J., Annegarn, H. J. and Otter, L. (2002): 'Southern African Regional Science Initiative (SAFARI 2000): summary of science plan', South African Journal Of Science, (April), pp. 119-124.

Tateishi, R. and Ebata, M. (2004): 'Analysis of phenological change patterns using 1982-2000 Advanced Very High Resolution Radiometer (AVHRR) data', International Journal of Remote Sensing, 25(12), pp. 2287-2300.

Thomas, D. S. G. and Shaw, P. A. (1991): Kalahari Environment, Cambridge, Cambridge University Press.

Thomas, D. S. G. and Twyman, C. (2004): 'Good or bad rangeland? Hybrid knowledge, science, and local understandings of vegetation dynamics in the Kalahari', Land Degradation & Development, 15(3), pp. 215-231.

Thompson, M. (1996) 'A standard land-cover classification scheme for remote sensing applications in South Africa', South African Journal Of Science, 92, pp. 34-42.

Toms, J. D. and Lesperance, M. L. (2003): 'Piecewise regression: A tool for identifying ecological thresholds', Ecology, Eco Soc America, 84(8), pp. 2034-2041.

Townshend, J. R. G. and Justice, C. O. (1986): 'Analysis of the dynamics of African vegetation using the normalized difference vegetation index', International Journal of Remote Sensing, 7(11), pp. 1435-1445.

Townshend, J. R., Latham, J., Arino, O., Balstad, R., et al. (2008): Integrated Global Observations of the Land: an IGOS-P Theme. IGOL Report No. 8, GTOS 54, Rome.

Townshend, J. R., Latham, J., Arino, O., Balstad, R., et al. (2007): 'Integrated Global Observation of Land: an IGOS-P Theme', [online] Available from: www.fao.org/gtos.

Trodd, N. M. and Dougill, A. J. (1998): 'Monitoring vegetation dynamics in semi-arid African rangelands', Applied Geograpghy, 18(4), pp. 315-330.

Trollope, W. S. W. (1998): 'Effect and Use of Fire in the Savanna Areas of Southern Africa',.

Tucker, C. T., Townshend, J. R. G. and Goff, T. E. (1985): 'African Land-Cover Classification Using Satellite Data', Science, 227, pp. 369-375.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., et al. (2003): 'Remote sensing for biodiversity science and conservation', Trends in Ecology & Evolution, 18(6), pp. 306-314.

Van Der Maarel, E. (1975): 'The Braun-Blanquet approach in perspective', Vegetatio, 30(3), pp. 213-219.

Vanacker, V., Linderman, M., Lupo, F., Flasse, S. and Lambin, E. (2005): 'Impact of short-term rainfall fluctuation on interannual land cover change in sub-Saharan Africa', Global Ecology and Biogeography, pp. 123-135.

Vancutsem, C., Bicheron, P., Cayrol, P. and Defourny, P. (2007): 'An assessment of three candidate compositing methods for global MERIS time series', Canadian Journal of Remote Sensing, 33(6), pp. 492-502.

Vancutsem, C., Pekel, J. F., Bogaert, P. and Defourny, P. (2007): 'Mean Compositing, an alternative strategy for producing temporal synthesis. Concepts and performance assessment for SPOT VEGETATION time series', International Journal of Remote Sensing, 28(22), pp. 5123-5141.

Verburg, P. H., De Steeg, J. van, Veldkamp, A and Willemen, L. (2009): 'From land cover change to land function dynamics: a major challenge to improve land characterization.', Journal of environmental management, 90(3), pp. 1327-1335.

Verlinden, A. and Laamanen, R. (2006): 'Long term fire scar monitoring with remote sensing in Northern Namibia: relations between fire frequency, rainfall, land cover, fire management and trees.', Environmental monitoring and assessment, 112(1-3), pp. 231-53.

Verlinden, A. and Laamanen, R. (2001): 'Modeling woody vegetation resources using Landsat TM imagery in northern Namibia', Africa, 2(1999).

Verlinden, A. and Laamanen, R. (2001): 'Modeling woody vegetation resources using Landsat TM imagery in northern Namibia', Southern African Forestry Journal, 2(1999).

Viovy, N., Arino, O. and Belward, A. S. (1992): 'The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series', International Journal of Remote Sensing, 13(8), pp. 1585-1590.

Vogel, M. (2005): 'Erfassung von Vegetationsveränderungen in Namibia mit Hilfe von Fernerkundungs-Change-Detection-Verfahren und unter Beruecksichtigung rezenter Niederschlagsereignisse', Dissertation, University of Wuerzburg.

Volk, O. (1966): 'Die Florengebiete von Südwestafrika', Journal of the South Western African Scientific Society, 20, pp. 25-58.

Volk, O. and Leippert, H. (1971): 'Vegetationsverhältnisse im Windhoeker Bergland, Südwestafrika', Journal of the South West African Scientific Society, 30, pp. 5-44.

Wagenseil, H. and Samimi, C. (2007): 'Woody vegetation cover in Namibian savannahs: a modelling approach based on remote sensing', Erdkunde, 61(4), pp. 325-334.

Walters, S. (2007): 'Modeling scale-dependent landscape pattern, dispersal, and connectivity from the perspective of the organism', Landscape Ecology, 22(6), pp. 867-881.

Ward, J. H. (1963): 'Hierarchical grouping to optimize an objective function', Journal of the American Statistical Association, American Statistical Association, 58(301), pp. 236-244.

Waske, B. and Braun, M. (2009): 'Classifier ensembles for land cover mapping using multitemporal SAR imagery', ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier B.V., 64(5), pp. 450-457.

Watts, J. D. and Lawrence, R. L. (2008): 'Merging random forest classification with an object-oriented approach for analysis of agricultural lands', In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII, Beijing, pp. 579-582.

Wessels, K. J., Prince, S. D., Frost, P. E. and Zyl, D. Van (2004): 'Assessing the effects of human-induced land degradation in the former homelands of northern South Africa with a 1 km AVHRR NDVI time-series', Remote Sensing of Environment, 91, pp. 47 - 67.

Wessels, K. J., Prince, S. D., Malherbe, J., Small, J., et al. (2007): 'Can human-induced land degradation be distinguished from the effects of rainfall variability? A case study in South Africa', Journal of Arid Environments, 68, pp. 271-297.

Westfall, R., Theron, G. and Rooyen, N.van (1997): 'Objective classification and analysis of vegetation data', Plant Ecology, 132, pp. 137-154.

White, J. C., Wulder, M., Goward, S. N., Masek, J. G., et al. (2008): 'Landsat continuity: Issues and opportunities for land cover monitoring', Remote Sensing of Environment, 112(3), pp. 955-969.

White, M. A., Beurs, K. M. de, Didan, K., Inouye, D. W., et al. (2009): 'Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006', Global Change Biology, 15(10), pp. 2335-2359.

Wilson, E. O. (1988): Biodiversity, Washington, National Academy Press.

Woodcock, C. and Harward, V. J. (1992): 'Nested-hierarchical scene models and image segmentation', International Journal of Remote Sensing, TAYLOR & FRANCIS LTD, 13(16), pp. 3167-3187.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H. and Zhong, L. (2005): 'Monitoring the response of vegetation phenology to precipitation in Africa by coupling MODIS and TRMM instruments', Journal of Geophysical Research, 110(D12), pp. 1-14.

Zhang, X., Friedl, M. A. and Schaaf, C. B. (2006): 'Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements', Journal of Geophysical Research, 111(G4), pp. 1-14.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., et al. (2003): 'Monitoring vegetation phenology using MODIS', Remote Sensing of Environment, 84, pp. 471 - 475.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H. and Zhong, L. (2005): 'Monitoring the response of vegetation phenology to precipitation in Africa by coupling MODIS and TRMM instruments', Journal of Geophysical Research, 10, D12103., doi, p. 101029/.

Zhang, X., Hodges, J. C. F., Schaaf, C. B., Friedl, M. A., et al. (2001): 'Global vegetation phenology from AVHRR and MODIS data', In International Geoscience and Remote Sensing Symposium, p. 2262–2264.

Zhang, X., Friedl, M. and Schaaf, C. (2009): 'Sensitivity of vegetation phenology detection to the temporal resolution of satellite data', International Journal of Remote Sensing, 30(8), pp. 2061-2074.

Appendix

Appendix A Overview of the publications that contribute to this dissertation

Authors	*Title*	*Publication*	*Date submitted*	*Status*	*Author's contribution*	*Signature of supervisor*
C. Hüttich, U. Gessner, M. Herold, B. Strohbach, M. Schmidt, M. Keil, S. Dech	On the Suitability of MODIS Time Series Metrics to Map Vegetation Types in Dry Savanna Ecosystems: A Case Study in the Kalahari of NE Namibia	Remote Sensing	3 July 2009	Accepted: 24 Sep. 2009, Published 30 Sep. 2009	Development of experimental design, data processing, presentation and discussion of results, manuscript editing	…………….....
C. Hüttich, M. Herold, B. Strohbach, S. Dech	Integrating in-situ, Landsat, and MODIS data for mapping in Southern African savannas: experiences of LCCS-based land-cover mapping in the Kalahari in Namibia	Environmental Monitoring and Assessment	31 Dec. 2009	Accepted: 27 June 2010, E-Pub ahead of print	Development of experimental design, data processing, presentation and discussion of results, manuscript editing	…………….....
C. Hüttich, M. Herold, M. Wegmann, A. Cord, B. Strohbach, C. Schmullius, S. Dech	Effects of temporal compositing for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring	Remote Sensing of Environment	30 March 2010	First revision in review	Development of experimental design, data processing, presentation and discussion of results, manuscript editing	…………….....
C. Hüttich, B. Strohbach, M. Herold, M. Keil, S. Dech	The Potential of MODIS Time Series Metrics for bottom-up Vegetation Mapping in a semi-arid Savanna Ecosystem in Namibia	Proceedings of the ESA Living Planet Symposium, 28 June -2 July, Bergen, Norway, ESA Special Publication (SP-686)	2 July 2010	ESA Publication SP-686 (CD ROM)	Development of experimental design, data processing, presentation and discussion of results, manuscript editing	…………….....

Appendix B Synoptic table of the phytosociological analysis (Strohbach *et al*., 2004)

Typecode	10	21	22	23	24	25	110	111	120	131	132	133	134	140	150	160	170	180	190
Order																			
Association	Pterocarpus angolensis -Burkea Africana woodlands	Terminalia sericea Combretum collinum shrub- and bushlands					Acacia erioloba - Terminalia sericia bushlands		Eragrostis rigidior - Urochloa brachyuran shrublands	Acacia mellifera - Stipagrostis uniplumis shrublands				Enneapogon desvauxii -Eriocephalus luederitzi short shrublands on calc. omiramba and pans	Acacia luederitzii -Ptichtolobium biflorum floodplains of the Omatako Omuramba	Enneapogon desvauxii - Acacia hebeclada wooded omiramba	Hyphaene petersiana plains	Terminalia pruniodes thickets	Combretum imberbe -Acacia tortilis woodlands
Subassociation/Variation		Schinziophyton rautanenii high dunes	Terminalia sericea - Combretum collinum shrub-and bushlands	Commiphora angolensis variant	Grewia flava variant	Acacia luederitzii subassociation	Typicum	Grewia flava - Stipagrostis uniplumis variant		Rhigozum brevispinosum subassociation Typicum	Acacia mellifera - Stipagrosti uniplumis shrublands	Catophractes alexandri subassociation	Ziziphus mucronatha -Acacia hebeclada subass.						
	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.	Avg.
Number of relevés	17	4	42	100	21	45	46	13	12	20	21	14	7	9	6	5	8	3	3
Number of species	36,5	31,3	33,1	39,7	39,4	41,6	39,6	38,6	39,1	37,7	40,2	37,4	41,7	31,7	48,0	32,4	44,6	27,0	23,3
Cover total (%)	83,5	76,3	73,6	77,2	75,6	65,8	69,6	80,4	81,7	67,3	70,5	72,9	65,0	66,1	55,0	75,0	73,1	80,0	45,0

Cover tree layer (%)	11,1	10,5	3,3	2,4	3,3	5,7	4,6	3,1	0,4	2,9	1,7	0,1	2,3	2,2	10,3	3,0	4,8	18,3	12,3
Cover shrub layer (%)	40	38,8	40,0	45,9	46,2	40,4	36,7	28,5	16,1	39,8	40,0	47,9	35,0	23,6	31,7	48,0	41,9	38,3	6,7
Cover herb layer (%)	40	35,0	34,6	32,7	31,2	21,8	30,8	52,5	72,9	27,9	33,7	26,1	27,7	44,7	21,7	28,0	27,5	28,3	26,7
Annual grass cover	29,1	30,3	18,5	16,7	12,0	10,2	12,0	12,5	25,2	11,3	15,4	13,2	7,4	26,9	11,0	19,0	21,1	28,5	18,8
Perennial grass cover	7,7	3,9	11,5	9,2	9,9	5,9	12,0	25,4	30,6	10,0	15,5	12,2	10,9	16,7	11,1	4,1	5,9	0,4	7,7
Height (highest) trees (m)	11,3	11,0	5,7	5,8	6,7	8,0	6,3	5,0		5,8	7,6		5,0	9,0	7,5		9,3	10,0	13,0
Height lowest trees (m)	6	6,0	4,4	4,2	4,7	5,8	4,5	4,0		4,5	5,0		4,0	4,5	4,8		5,5	5,5	6,7
Height (highest) shrubs (m)	3	4,3	3,8	3,6	3,4	4,2	3,9	2,0	2,1	3,4	3,8	2,6	2,5	2,5	3,3		4,3	3,5	2,3
Height lowest shrubs (m)	0,9	1,3	0,8	0,8	1,1	0,6	0,9	0,9	0,6	1,2	1,0	0,9	0,7	0,6	1,1		1,5	1,2	0,9
Aver. height (high) herbs (cm)	107,5	66,7	91,7	69,0	57,5	55,0	48,4	73,3	58,6	56,7	48,2	37,5	30,0	30,0	31,7		60,0	40,0	30,0
Aver. height lowest herbs (cm)	4,3	5,0	5,0	4,0	4,1	2,5	4,2	3,0	3,6	4,5	3,0	3,5	2,0	3,7	2,8		1,8	5,0	1,7
Structure	Bmct	Bmct	Bmcs	Smch	Bmcs	Bmct	Bmct	Bsot	Ssot	Smct	Smch	Smct	Smct	Ssot	Bmcs	Smct	Bmcs	Tmct	Wsot
Growing period zone	4	4	4	3 - 4	4	4	3 - 4	4	4	4	3 - 4	3 - 4	4	3 - 4	4	4	3	4	4
Stratigraphy	Kalah	Kalah	Kala	Kala Nos. Gr.	Kala, Noss. Group	Kala, Noss. Group	Kala Nos Gr.	Kalah	Kalah	Kalah	Kalah	Kalah	Noss. Gr.,	Kala, Noss. Gr.	Kalah	Kala, Noss. Group	Kalah		Kalah
Lithology	UE	UE													UF		UF / Uca		
Cover: Gravel	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,3	0,6	0,4	0,0	1,4	0,3	0,0	0,0
Cover: Small stones	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,4	0,0	1,2	0,3	0,0	0,0
Cover: Medium stones	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,2	0,1	0,8	0,0	1,6	0,6	0,0	0,0
Cover: Large stones	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,7	0,0	1,2	0,6	0,3	0,0
Rock	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,2	0,0	2,2	0,1	0,0	0,0
Aspect (degrees)																			
Altitude (m)	1245,4	1201,0	1291,3	1271,1	1370,9	1354,7	1323,6	1277,2	1228,2	1320,8	1331,4	1348,9	1282,0	1236,6	1308,5	1222,0	1238,4	1117,7	1300,0
Soil depth	150	150,0	147,4	143,4	123,8	133,8	108,7	103,8	60,0	120,5	98,1	77,9	75,7	58,9	15,0	36,0	10,0	33,3	0,0
Species group A	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy	Constancy
Pterocarpus angolensis	65%																		
Crotalaria flavicarinata	35%																		
Clerodendrum dekindtii	29%		2%				2%												
Combretum zeyheri	24%																		
Strychnos cocculoides	18%			1%															
Species group B																			
Schinziophyton rautanenii		100%																	
Strychnos pungens		50%																	
Species group C																			
Burkea africana	100%	50%	52%	8%															
Hermannia eenii	76%		26%	8%		2%	7%	23%			10%								
Combretum engleri	35%	75%	31%	15%		4%	2%												
Panicum kalaharense	65%	75%	29%	8%	5%				8%										
Oxygonum delagoense	35%		21%	18%	5%	4%													

Salacia luebbertii	41%	25%	12%	8%															
Mariscus confusus	18%		14%	10%		2%					5%								
Baissea wulfhorstii	18%	25%	2%	6%															
Tephrosia cephalantha v. decumbens	24%		5%	2%															
Species group D																			
Combretum collinum	94%	100%	98%	75%	71%	71%	4%		8%		5%	7%							
Ochna pulchra	100%	100%	88%	44%	14%	29%													
Eragrostis pallens	47%		81%	49%	62%	29%	4%												
Acacia ataxacantha	59%	75%	43%	60%	24%	31%	11%		8%	5%									
Grewia avellana	41%		43%	44%	33%	31%	11%			5%		7%				20%	13%	67%	
Aristida stipitata	47%	25%	43%	34%	38%	38%	26%						14%						
Combretum psidioides	82%	50%	95%	22%	5%	7%													
Pogonarthria squarrosa	18%		17%	31%	52%	18%	4%			5%		7%							
Dichapetalum cymosum	41%	25%	31%	19%	29%	4%	2%		8%										
Phyllanthus omahekensis	35%		31%	24%	5%	7%	4%												
Species group E																			
Aristida meridionalis			5%	8%	5%	2%	9%	38%	8%										
Species group F																			
Megaloprotachne albescens	100%	100%	88%	78%	43%	42%	15%				5%								
Croton gratissimus	24%	100%	33%	31%	5%	9%	4%			5%	14%						25%	67%	
Mundulea sericea	29%	50%	24%	18%	10%	22%	11%			15%	14%					20%	13%	67%	
Hemizygia bracteosa	12%		21%	21%	5%	7%	4%			5%									
Sida ovata	18%		10%	13%		22%	9%		8%	5%									
Species group G																			
Terminalia sericea	100%	100%	98%	98%	95%	96%	78%	46%	33%	5%									
Xenostegia tridentata	94%	100%	83%	80%	48%	44%	35%	23%	33%	5%		7%					13%		
Digitaria seriata	94%	100%	93%	51%	57%	22%	20%	15%	33%	5%	5%		14%					33%	
Tephrosia lupinifolia	88%	75%	40%	51%	38%	40%	30%	15%	17%				14%				13%		
Chamaecrista falcinella v. parviflora	59%	25%	14%	30%	43%	36%	26%		17%										
Hibiscus meeusei	65%		50%	33%	5%	4%	9%		25%	5%					17%				
Basananthe pedata	18%		17%	22%	14%	22%	7%	15%	42%	5%			14%						
Melinis repens s. repens	35%		31%	26%	10%	4%	4%	23%	17%			7%							
Brachiaria nigropedata	29%		17%	15%	14%	9%	2%	15%	42%	5%		7%	14%						
Species group H																			
Indigofera bainesii			5%	30%	10%	36%	4%	8%	33%	10%			14%						
Asparagus suaveolens			5%	13%	5%	36%	13%	23%		20%	5%	7%							
Oligomeris linifolia			7%	12%	14%	11%	13%			10%	5%								
Ipomoea verbascoidea			10%	10%		7%	4%			10%		7%							
Species group I																			
Limeum fenestratum	65%	100%	38%	56%	48%	47%	46%	69%	8%	10%	24%	7%	14%						

Eragrostis lehmanniana	12%		21%	41%	57%	40%	28%	8%		35%	10%	7%		11%					
Ipomoea magnusiana	35%	50%	26%	41%	14%	33%	15%		8%	10%	14%			11%			13%		
Ipomoea chloroneura	35%		24%	32%	14%	24%	15%	8%	67%	20%	14%		14%						
Clerodendrum uncinatum	18%		2%	21%	24%	27%	11%				10%								
Cyperus margaritaceus	12%		17%	15%	24%	18%	7%	8%		5%	10%								
Species group J																			
Rhigozum brevispinosum				38%	48%	62%	41%	38%	33%	100%	38%	79%	14%						
Tylosema esculentum				15%	33%	33%	39%	92%		30%	62%	50%							
Neorautanenia amboensis				10%	5%	4%	7%	31%	50%	10%	14%	14%				20%			
Species group K																			
Merremia verecunda	6%	25%	24%	34%	38%	47%	37%	69%	42%	30%	33%	14%	14%						
Rhynchosia totta	6%		29%	25%	29%	36%	20%	15%	8%	15%	14%	21%	14%				13%		
Ozoroa paniculosa			12%	11%	38%	38%	22%	62%	8%	5%	19%	14%							
Crotalaria sphaerocarpa	6%	25%	17%	14%	14%	4%	22%	8%	17%	20%	19%	14%					13%	33%	
Tragia dioica			5%	7%	5%	13%	24%	8%	8%	10%	10%	29%	14%	11%		20%			
Species group L																			
Vernonia poskeana	71%	25%	67%	75%	76%	73%	61%	100%	83%	40%	19%	36%					13%		
Bauhinia petersiana s. macrantha	100%	75%	79%	81%	76%	62%	59%	8%		25%		14%					13%		
Acanthosicyos naudinianus	65%	25%	43%	50%	71%	27%	22%	62%	58%	25%	24%	14%							
Indigofera daleoides v. daleoides	18%		10%	20%	19%	16%	37%	54%	25%	25%	10%	29%		11%		20%	13%		
Species group M																			
Otoptera burchellii				1%	5%	2%	15%	31%	50%	25%	48%	64%	71%	11%		20%			
Ipomoea oblongata				1%	19%	4%	13%	38%		5%	14%	14%	29%						
Species group N																			
Acacia erioloba				10%	38%	16%	83%	77%	8%	15%	43%	43%	71%	11%	17%	20%			
Requienia sphaerosperma			2%	12%	62%	40%	50%	15%	17%		5%	7%	29%						
Indigofera vicioides				2%	5%	11%	4%		25%	20%	14%	21%	43%	11%	17%				
Eragrostis omahekensis			2%	10%	19%	31%	11%	8%	8%	5%		14%	57%	11%		20%			33%
Elephantorrhiza elephantina			2%	2%	57%	22%	20%	23%			5%		43%						
Species group O																			
Talinum crispatulum				10%	19%	22%	35%	54%	50%	25%	57%	14%	71%	11%					
Ipomoea bolusiana				10%	14%		22%	38%	33%	30%	19%	21%	29%				13%		
Kohautia caespitosa s. brachyloba				4%	10%	4%	11%	38%	8%	30%	29%	29%	29%	11%	17%	20%			
Species group P																			
Eragrostis rigidior			5%	30%	43%	36%	30%	92%	100%	35%	48%	43%	57%	11%			13%		
Dicoma schinzii			5%	19%	52%	49%	63%	46%		20%	14%	21%	71%	11%					
Rhus tenuinervis	18%		5%	23%	14%	27%	20%	38%	8%	45%	14%		43%			40%	13%	67%	
Hermannia tomentosa		25%	5%	18%	19%	20%	11%	8%	8%				29%						33%
Species group Q																			
Tephrosia burchellii	65%	75%	69%	78%	71%	93%	67%	38%	42%	70%	71%	36%	86%		17%		13%		
Lonchocarpus nelsii	88%	75%	81%	72%	52%	64%	28%	62%	50%	15%	19%	21%	29%						

Sesamum capense	29%	50%	40%	37%	14%	16%	22%	62%	58%	45%	19%	21%	43%						
Waltheria indica	18%	75%	19%	15%	24%	4%	26%	23%	17%		5%		43%				13%		
Species group R																			
Eragrostis echinochloidea												7%		89%					
Eriocephalus luederitzianus														44%					
Cynodon dactylon														44%					
Salsola tuberculata														33%					
Panicum coloratum														33%					
Heliotropium ovalifolium												7%		22%					
Eragrostis pilgeriana														22%		20%			
Eragrostis truncata														22%					
Peliostomum leucorrhizum														22%		20%			
Tragia okanyua														22%					
Aptosimum albomarginatum												7%		22%					
Species group S																			
Schmidtia pappophoroides	6%	25%	2%	24%	62%	29%	13%	77%	58%	20%	5%	7%	29%	33%			13%		
Hirpicium gazanioides					5%		17%			30%	24%		43%	22%		20%	13%		
Tarchonanthus camphoratus				3%	14%	18%	15%	46%	17%	15%	29%	43%	43%	22%					
Walleria nutans			2%	3%	24%	18%	15%	8%	8%	25%	10%	21%		22%					
Species group T																			
Commiphora angolensis	24%	25%	43%	82%	19%	42%	48%		25%	40%	33%	57%		22%			13%		
Raphionacme velutina	35%		33%	14%	43%	20%	15%	8%	25%	30%	24%	14%	29%	22%					
Species group U																			
Felicia clavipilosa s. clavipilosa				2%			4%			20%					100%				33%
Xerophyta humilis															50%				
Barleria lancifolia												7%			50%				
Species group V																			
Ptycholobium biflorum s. angolensis										10%	33%	7%		11%	67%				
Species group W																			
Kleinia longiflora				4%	29%	31%	35%	15%		55%	62%	64%	14%	11%	83%	20%	13%		33%
Eragrostis biflora				3%	10%	18%	13%	8%		35%	14%	14%	14%		33%	20%		33%	
Blepharis integrifolia				1%	5%	2%	9%	15%	8%	15%	5%	14%	29%	11%	33%				
Species group X																			
Eragrostis dinteri			24%	26%	10%	9%	63%	38%		25%	57%	29%			83%		13%		
Kyphocarpa angustifolia			5%	10%	19%	18%	39%	15%	50%	30%	43%	29%	14%	11%	100%		13%		
Pollichia campestris			5%	4%	14%	36%	30%	8%	17%	45%	24%	21%	43%		33%	20%	13%		
Species group Y																			
Evolvulus alsinoides	35%	25%	31%	40%	33%	38%	39%	23%	8%	40%	52%	43%	57%		50%		13%		
Commelina africana v. krebsiana	71%		33%	44%	19%	18%	4%	15%	33%	35%	29%				67%				
Ipomoea obscura	35%		38%	36%	29%	24%	11%	23%	25%	10%	5%	29%	14%		33%				

Acrotome angustifolia	24%	50%	29%	32%	19%	16%	35%	15%		10%	5%	7%	14%	11%	33%				
Species group Z																			
Catophractes alexandri							4%	8%	8%	15%	24%	100%	14%	56%	33%	40%	13%		
Phyllanthus maderaspatensis						2%	2%		8%	10%	14%	29%	71%	11%	50%	40%			
Species group AA																			
Lantana angolensis				6%	10%	31%	24%	15%		30%	24%	57%	43%	44%	17%	40%			
Solanum burchellii				6%		22%	22%		25%	15%	19%	29%	14%			40%			
Species group AB																			
Hyphaene petersiana														11%			75%		
Enneapogon scoparius												7%		11%			50%		
Sporobolus ioclados																	38%		
Cucumis meeusei																	25%		
Elaeodendron transvaalense																	25%		
Eragrostis aspera																	25%		
Cardiospermum halicacabum																	25%		
Species group AC																			
Enneapogon desvauxii											5%	21%		89%		100%	75%		
Species group AD																			
Triraphis purpurea				1%			7%			10%	52%	29%	14%	22%			63%		
Seddera suffruticosa							4%	8%	8%	5%	29%	36%	14%	44%		40%	63%	33%	
Brachiaria deflexa							4%			5%	19%	7%		33%	83%		100%		
Geigeria ornativa				1%			4%		17%	5%	24%	7%		33%	67%	40%	38%		
Tragus racemosus				1%						5%	10%	7%	14%	67%		20%	75%		
Aristida rhiniochloa						2%	2%	8%		5%	24%	14%			50%		63%		
Cenchrus ciliaris											10%	21%	14%	56%		40%	25%		
Species group AE																			
Acacia mellifera s. detinens				13%	48%	78%	78%	38%	42%	100%	100%	93%	86%	33%	100%	100%	100%	33%	33%
Clerodendrum ternatum			2%	29%	57%	33%	30%	54%	58%	65%	57%	86%		22%	33%	20%	38%	33%	
Pogonarthria fleckii			2%	18%	24%	13%	39%	46%	50%	50%	71%	50%	57%	11%	83%	20%	75%		
Acacia luederitzii				7%	14%	47%	30%	31%		50%	33%	21%	29%	33%	100%		88%	33%	
Ziziphus mucronata				1%	24%	13%	35%	38%	33%	25%	33%	14%	100%	22%	83%	80%	50%		33%
Aristida adscensionis				3%	5%		20%	8%	8%	20%	29%	36%	14%	44%	83%	20%	75%		
Heliotropium steudneri		25%	2%	9%	10%	9%	9%	69%	17%	30%	14%	14%	14%	22%	17%	20%	25%		
Ipomoea sinensis				6%	10%	9%	9%	8%	67%	20%	29%	7%		11%	17%		25%		
Commiphora pyracanthoides				2%			7%	8%	25%	25%	29%	7%	14%		33%	20%	75%		
Sansevieria aethiopica				11%	5%	11%	7%			30%	10%	14%				20%	25%		
Spermacoce senensis			2%	6%		2%	26%	8%		15%	14%				50%		50%		
Zornia glochidiata				2%		2%	22%	23%	42%	5%	10%	29%	29%				38%		
Limeum sulcatum				3%	5%	2%	4%		8%		38%	21%		22%	50%		50%		
Species group AF																			
Stipagrostis uniplumis v. uniplumis	82%	75%	93%	87%	90%	78%	74%	100%	92%	70%	95%	71%	86%	67%	33%	60%	63%	33%	

Melinis repens s. grandiflora	29%	25%	67%	85%	76%	73%	83%	69%	75%	45%	62%	64%	43%	33%	50%	40%	25%		
Gisekia africana	88%	100%	60%	65%	52%	58%	54%	85%	42%	55%	62%	50%	57%	22%	83%	80%	63%		
Grewia retinervis	82%	50%	71%	86%	62%	91%	70%	23%	8%	65%	57%	29%	29%	11%	67%		75%		
Oxygonum alatum	47%	50%	43%	52%	86%	40%	57%	77%	75%	30%	33%	21%	29%	11%			50%		
Bulbostylis hispidula	71%	100%	67%	74%	81%	71%	39%	69%	83%	55%	33%	43%	29%	11%	33%		25%		
Phyllanthus pentandrus	65%	75%	81%	63%	43%	60%	52%			30%	24%	36%	29%	33%	67%	20%	63%	33%	
Cleome rubella			31%	28%	19%	16%	37%	31%	58%	30%	33%	36%	29%		67%		38%		
Asparagus nelsii	18%		29%	49%	10%	42%	33%	23%		20%	24%	29%	29%	11%		20%	38%	33%	
Ipomoea hackeliana	53%	25%	40%	33%	48%	62%	43%	15%	42%	30%	10%	21%	14%		17%		38%		
Acrotome inflata	29%	25%	26%	43%	14%	31%	30%	23%	8%	15%	29%		14%	11%	33%		50%		33%
Indigofera charlieriana v. charlieriana	6%	25%	10%	21%	14%	20%	11%	38%		20%	19%	29%		22%	33%		38%	33%	33%
Momordica balsamina	29%	100%	36%	27%	10%	36%	17%	8%	42%	5%	14%			22%			38%		
Commiphora africana	29%		12%	15%	24%	13%	13%		17%	30%	14%	14%			17%	20%	38%		
Indigofera flavicans	6%		7%	13%	19%	22%	17%	15%		15%	10%	7%	14%				38%		
Species group AG																			
Terminalia prunioides											5%						75%	100%	
Grewia bicolor			2%				2%	8%			5%					20%	25%	100%	
Species group AH																			
Grewia flava				42%	76%	80%	72%	92%	75%	90%	86%	86%	86%	67%	83%	80%	50%	100%	
Species group AI																			
Acacia fleckii	6%		12%	68%	76%	71%	93%	54%	33%		19%	29%	29%	11%			13%	100%	
Species group AJ																			
Urochloa brachyura	6%		21%	74%	90%	87%	87%	85%	100%	85%	57%	71%	43%	22%	17%	20%	63%	100%	
Ehretia rigida			5%	12%	10%	38%	28%	31%	17%	55%	48%	43%	43%	22%	50%	60%	25%	67%	
Species group AK																			
Dichrostachys cinerea	18%	50%	29%	59%	67%	82%	89%	46%	75%	55%	71%	57%	43%	11%		20%	50%	100%	
Species group AL																			
Boscia albitrunca	18%	25%	14%	45%	38%	56%	37%	46%	58%	100%	62%	57%	57%	22%	83%	40%	13%	100%	
Asparagus cooperi	53%		26%	29%	19%	40%	15%	46%	17%	40%	43%	36%	57%	67%	67%	60%	38%	67%	
Species group AM																			
Mollugo nudicaulis																			100%
Lotononis listii																			100%
Species group AN																			
Combretum imberbe																20%	63%		100%
Species group AO																			
Dactyloctenium aegyptium						2%		8%		10%	10%	7%	14%		67%		13%	33%	100%
Chloris virgata									8%	5%	5%	14%	14%	11%	83%		13%		100%
Acacia tortilis s. heteracantha							11%								50%	20%			100%
Species group AP																			
Eragrostis rotifer																			67%
Schoenoplectus senegalensis																			33%
Species group AQ																			

Panicum gilvum																			
Diandrochloa pusilla																			
Pycreus macrostachyos																			
Cyperus species																			
Kohautia virgata														11%					
Vahlia capensis																			
Marsilea nubica v. gymnocarpa																			
Schoenoplectus muricinux																			
Lindernia parviflora																			
Species group AR																			
Acacia hebeclada s. hebeclada						4%	15%	31%	50%	5%	29%	21%	86%	44%	50%	60%			67%
Species group AS																			
Aristida congesta s. congesta				20%	48%	56%	41%	46%	75%	50%	38%	50%	86%	33%	67%	20%	13%		67%
Eragrostis porosa			2%	3%	5%	11%	37%	15%	42%	45%	76%	79%	71%	100%	100%	80%	75%	67%	100%
Tragus berteronianus				4%	10%	20%	28%	8%	25%	55%	71%	64%	71%	44%	100%	40%	25%		67%
Eragrostis trichophora				7%	5%	2%	39%	15%	50%	25%	38%	36%	29%	22%	100%		75%		67%
Chamaesyce inaequilatera			2%	2%	19%	16%	15%	38%		20%	10%	29%	43%	44%	17%	20%	50%		67%
Sida cordifolia				11%	5%	20%	30%			25%	10%	14%	14%	11%	17%	20%			67%
Species group AT																			
Indigofera filipes		50%	36%	40%	29%	51%	13%			5%									
Pergularia daemia			7%	23%	14%	22%	22%		8%	15%	10%			11%		20%	13%		
Blepharis obmitrata				22%	14%	27%	30%			15%					17%				
Pupalia lappacea						20%	9%			30%	33%	21%	14%	44%	67%	80%	13%	67%	
Citrullus lanatus	6%	75%	17%	7%		4%	11%	15%	50%	10%		7%			33%		13%		
Striga gesnerioides			5%	11%	14%	18%	13%	23%	25%			7%				20%			
Hermannia modesta						2%	2%	23%	17%	15%	24%	43%	43%	89%	33%	80%	25%		
Enneapogon cenchroides			2%	4%	5%	4%	7%	8%		15%	33%	36%		44%	50%	20%		67%	
Corchorus tridens							2%	8%	42%	30%	19%	29%			50%		50%		100%
Chamaecrista absus			2%	4%	5%		9%		75%	20%	5%	7%					13%		
Indigofera baumiana	6%	25%	12%	15%	14%	9%	2%	8%		5%									
Dicoma tomentosa				1%		2%			17%	25%	43%	14%	29%		83%		50%		
Talinum caffrum				1%		7%	7%		25%	20%	24%	21%		22%	50%		38%		
Eragrostis jeffreysii	6%		10%	11%	5%	18%	2%				5%								
Sarcostemma viminale	6%		2%	9%		2%	11%			5%	10%	7%		11%					
Achyranthes aspera v. sicula					10%	13%	4%				10%		14%	22%	17%	20%	25%	100%	33%
Anthephora pubescens			5%	9%	5%	4%	4%	15%	17%			7%							
Tribulus terrestris			2%			4%	2%		8%	15%	5%	14%	14%	11%	50%	60%	25%		33%
Hoffmannseggia burchellii			2%	7%	10%	16%	9%	8%		5%									
Barleria macrostegia				1%	19%	2%	11%	31%	8%	10%	10%		14%					33%	
Combretum hereroense	6%						7%	31%			14%	7%	14%	22%		40%	75%		
Aristida effusa				2%	5%	2%	9%				24%	14%	14%	11%		80%		33%	

Maytenus senegalensis				2%		2%	11%	8%	8%	10%	5%	7%	29%			20%	25%		
Tribulus zeyheri						11%	9%	8%		5%	10%			11%	17%	40%	25%		
Lapeirousia bainesii			2%	5%	5%	4%	7%	8%	17%		5%	7%							
Melhania acuminata				4%	5%	4%	15%			10%	5%	7%	14%						
Ipomoea coptica				7%	5%		2%			10%			14%		50%		25%		
Senna italica					10%	4%	7%	8%	33%	5%	10%		29%	22%					
Pavonia clathrata	12%			13%	14%	4%													
Crotalaria podocarpa				3%	5%		4%		42%	15%		7%		11%	17%	20%			
Aspect (degrees)				1%	5%	7%	9%		17%	5%	5%		14%						
Setaria verticillata						2%	4%	8%			14%		14%	56%	17%	20%	13%	67%	33%
Hermbstaedtia odorata				1%			11%		8%		14%		14%	33%				67%	
Helinus spartioides			2%	2%	5%		2%	8%			24%	36%							
Eragrostis nindensis				2%		2%		15%	25%		14%		29%	22%			25%		
Monechma divaricatum						7%	7%			15%	14%	14%	14%	22%		20%			
Gomphocarpus tomentosus		25%	2%	7%	5%	9%	4%									20%			
Albizia anthelmintica						2%	2%	8%			33%	7%		11%			25%	67%	
Ophioglossum polyphyllum				3%			2%	15%	8%		5%	7%	57%	11%			25%		
Jatropha erythropoda			5%	4%	5%	4%	4%	8%	8%		5%	14%							
Triraphis schinzii			12%	7%	10%		2%	8%											
Aristida stipoides				1%			7%			5%				11%	50%		88%		
Cephalocroton mollis					5%		4%	8%		5%	19%	21%					13%	67%	
Commelina benghalensis				1%		4%	4%			5%	19%	7%			17%			33%	
Kalanchoe lanceolata				2%		13%	2%			10%		14%	14%				13%		
Limeum arenicolum			2%	1%		2%					29%		14%	22%			38%		
Monechma spartioides				1%	5%		2%			10%	14%		14%	11%	50%		13%		
Pavonia burchellii				2%		11%	2%	8%		5%	10%		14%		33%				
Vigna unguiculata	6%	25%	17%	4%					8%										
Euphorbia forskalii	12%		7%	4%	10%	2%					5%								
Eragrostis viscosa				1%			2%		17%	10%			14%		33%		25%		100%
Rhynchosia resinosa	6%		12%			2%	2%	15%		5%	5%					20%			
Tricholaena monachne				9%	5%		2%	8%		5%									
Solanum delagoense									8%	10%	19%	7%			33%		25%		
Hibiscus micranthus										15%	5%				83%			33%	
Cucumis africanus				2%		2%	2%		8%		10%		14%	22%					
Ocimum americanum v. americanum				1%				8%	8%	5%	10%	14%	14%		33%		13%		
Dactyloctenium giganteum						7%	4%				5%								
Talinum tenuissimum			2%	2%	10%	2%	2%		8%								25%		
Cyamopsis serrata			2%			2%	9%		17%	5%	10%								
Lotononis platycarpa						2%			17%	5%	5%	7%	14%	11%	17%	20%	13%		
Lycium bosciifolium							2%			10%		7%			83%				33%

Portulaca kermesina				2%		2%	4%			5%		14%							
Nolletia tenuifolia				6%	5%	4%				5%									
Tricliceras schinzii s. schinzii v. jutt				1%			17%												
Dipcadi marlothii			5%	3%	5%	2%				5%	5%								
Limeum myosotis				1%			7%	8%			10%	7%				20%	13%		
Panicum maximum						2%				5%		7%	14%		17%		25%		33%
Zehneria marlothii			2%	1%	5%					10%	10%		14%	11%			13%		
Sericorema sericea								8%		5%		14%				60%	25%		
Pegolettia senegalensis				1%			7%	8%			10%						38%		
Monsonia angustifolia				1%	5%	4%	4%				5%	7%	14%						
Acalypha segetalis				2%							10%			11%		20%	25%	67%	
Gnidia polycephala							7%	8%		5%	5%		29%			20%			
Ipomoea hochstetteri			2%			4%	2%	8%	8%	5%			14%						
Aristida pilgeri							11%	8%	17%		5%								
Crotalaria steudneri									25%		5%			11%	33%		13%		33%
Perotis patens			2%	6%		2%													
Crotalaria pisicarpa									33%	5%									
Portulaca hereroensis													14%	11%	33%	20%	38%		
Melhania virescens							2%				5%			33%		40%			
Kyllinga alba			2%	3%		2%	4%						14%						
Amaranthus thunbergii				1%		2%				5%	5%		14%		33%			33%	
Acanthospermum hispidum						2%	2%							11%	17%	40%			67%
Trochomeria debilis				1%		2%					5%		14%						
Kalanchoe rotundifolia										15%	5%		14%		17%		13%		
Alectra orobanchoides						7%			8%	5%									
Polygala schinziana				4%	5%	2%	2%												
Ipomoea welwitschii				3%	10%	2%					5%								
Peltophorum africanum				1%			2%				5%						50%		
Cyperus fulgens						2%			17%	10%		7%							33%
Maerua schinzii								8%			5%					20%	38%	33%	
Entada arenaria	18%		5%	1%	5%														
Sporobolus fimbriatus							4%	8%	8%			7%			33%				
Fockea angustifolia				2%	5%	4%	4%												
Antizoma angustifolia					5%			8%		15%		14%							
Aptosimum arenarium							2%	8%	17%			7%	14%						
Chascanum pinnatifidum									17%	5%		21%	14%						
Calostephane divaricata										10%					50%				33%
Cyperus amabilis				4%	5%				8%										
Diospyros lycioides							2%	8%	25%							20%			
Tapinanthus oleifolius							4%			5%			14%						
Lindneria clavata				2%				8%				7%		11%					

Monechma debile						2%				5%					17%			33%	67%
Helinus integrifolius										5%	5%	14%					13%	33%	
Sericorema remotiflora				1%		2%		23%						11%					
Vernonia fastigiata							2%	8%					14%				38%		
Aptosimum angustifolium						2%	9%												
Acacia nebrownii													29%	22%		20%			
Oropetium capense														33%		40%			
Tephrosia dregeana						2%								11%		60%			
Blepharis maderaspaten				2%	5%	2%	2%												
Brachiaria marlothii														22%	17%				67%
Hibiscus elliottiae											10%							33%	
Cyphostemma congestum				1%			2%				10%	7%							
Hermannia guerkeana						2%	2%						14%						
Digitaria velutina				1%							5%				17%		13%	33%	
Ornithoglossum vulgare						2%			8%		5%			22%					
Ximenia americana							2%				5%						38%		
Leucosphaera bainesii											10%				17%	20%		33%	
Rhynchosia sublobata				1%						10%							25%		
Becium filamentosum							2%						14%	11%					
Mollugo cerviana													14%		50%		13%		
Bidens schimperi			2%	3%															
Barleria senensis							2%			5%					33%				
Acrosanthes angustifolia						2%					10%					20%			
Barleria lanceolata							2%				5%				17%				
Alectra sessiliflora				1%			2%					7%							
Hibiscus calyphyllus										5%		14%						33%	
Polygala leptophylla					5%		2%				5%								
Leucas pechuelii											10%			22%					
Guilleminea densa													14%		17%				67%
Hibiscus vitifolius					5%										33%				
Cymbopogon plurinodis									8%			7%	14%	11%					
Cucumis kalahariensis						2%		8%	8%										
Euphorbia spartaria				1%		4%				5%									
Corallocarpus triangularis			2%										14%					33%	
Ceropegia lugardae				2%	5%		2%												
Maerua juncea				1%						5%			14%				13%		
Aptosimum decumbens						2%			8%										
Limeum pterocarpum											10%				17%				
Alternanthera pungens													29%			20%			33%
Aptosimum glandulosum													14%	11%		20%			33%
Indigofera			7%																

arenophila																			
Cynanchum orangeanum									17%				14%						
Aerva leucura																60%			
Chenopodium pumilio								8%					14%	11%					
Neorautanenia mitis					5%							7%							
Tavaresia barklyi				1%		2%				5%									
Solanum supinum				1%			2%					7%							
Chenopodium petiolariforme						2%								11%		20%			
Nolletia gariepina				2%															
Boscia foetida											10%								
Hibiscus caesius												14%							
Justicia odora													29%						
Marsdenia sylvestris																		67%	
Heteropogon contortus									8%				14%						
Jamesbrittenia atropurpurea s. atropurpu														11%		20%			
Helichrysum argyrosphaerum											5%				17%				
Marsilea macrocarpa																			33%
Kohautia azurea										5%			14%						
Euphorbia crotonoides			2%		5%														
Vigna lobatifolia				1%		2%													
Commelina livingstonii				1%			2%												
Hypoestes forskaolii						2%										20%			
Erlangea misera				1%													13%		
Trichoneura grandiglumis									8%			7%							
Opuntia species							2%			5%									
Vangueria infausta	6%			1%															
Limeum viscosum						2%											13%		
Schmidtia kalihariensis							2%						14%						
Indigofera hochstetter										5%							13%		
Kohautia cynanchica												7%				20%			
Jacquemontia tamnifolia				1%							5%								
Eragrostis annulata											5%						13%		
Ipomoea crassipes							2%				5%								
Corallocarpus welwitschii			2%	1%															
Solanum kwebense				1%														33%	
Abutilon austro-africanum										5%		7%							
Securidaca longepedunculata	6%			1%															
Cymbopogon excavatus													14%	11%					
Setaria pumila																	13%		33%
Boophane disticha						2%											13%		

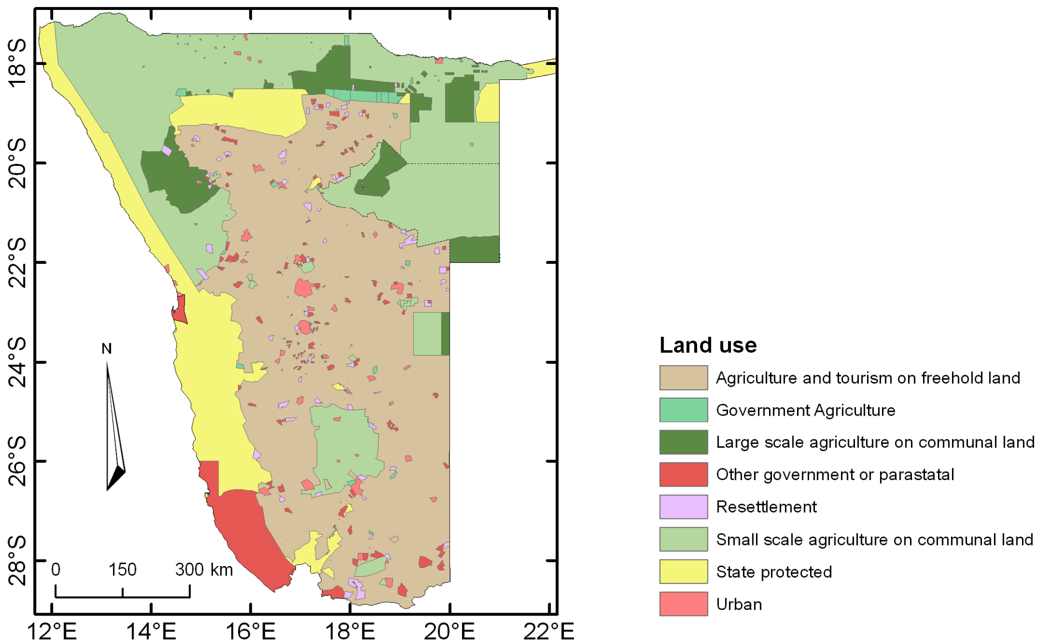
Panicum lanipes													14%	11%					
Gloriosa superba				1%							5%								
Aloe hereroensis																20%			
Ipomoea adenioides																20%			
Dactyliandra welwitschii												7%							
Nerine laticoma														11%					
Ruellia species																	13%		
Ammocharis coranica									8%										
Eriospermum species											5%								
Eragrostis cylindriflora																			
Petalidium engleranum																20%			
Plinthus sericeus														11%					
Harpagophytum zeyheri																			
Elionurus tripsacoides	6%																		
Indigastrum parviflorm s. occidentalis														11%					
Aizoon virgatum																20%			
Crossopteryx febrifuga				1%															
Commelina forskaolii																			
Acacia karroo																20%			
Datura stramonium																20%			
Bidens pilosa																	13%		
Sansevieria pearsonii										5%									
Stipagrostis hirtigluma														11%					
Trachyandra laxa																			
Endostemon tereticaulis																	13%		
Hibiscus sidiformis													14%						
Lycium eenii											5%								
Polycarpaea corymbosa							2%												
Hermbstaedtia fleckii						2%													
Heliotropium marifolium										5%									
Microchloa caffra													14%						
Bidens biternata																20%			
Solanum incanum															17%				
Flaveria bidentis																20%			
Psydrax livida	6%																		
Crotalaria species	6%																		
Peristrophe hereroensis															17%				
Lapeirousia odoratissima			2%																
Trochomeria macrocarpa s. vitifolia													14%						

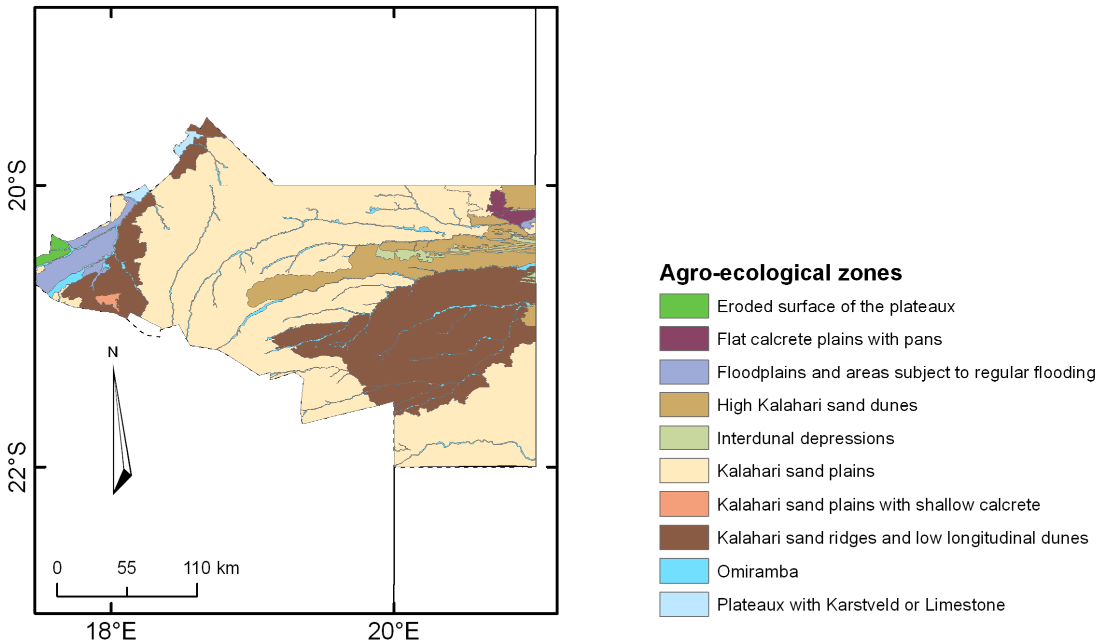
Helichrysum candolleanum				1%															
Ipomoea coscinosperma														11%					
Adenia repanda																			
Corchorus schimperi																		33%	
Striga asiatica			2%																
Hermannia bicolor										5%									
Tagetes minuta																20%			
Ximenia caffra			2%																
Sebaea grandis							2%												
Barleria albi-pilosa										5%									

Appendix C Agro-ecological zones after De Pauw & Coetzee (1999) and Strohbach *et al.* (2004, modified).

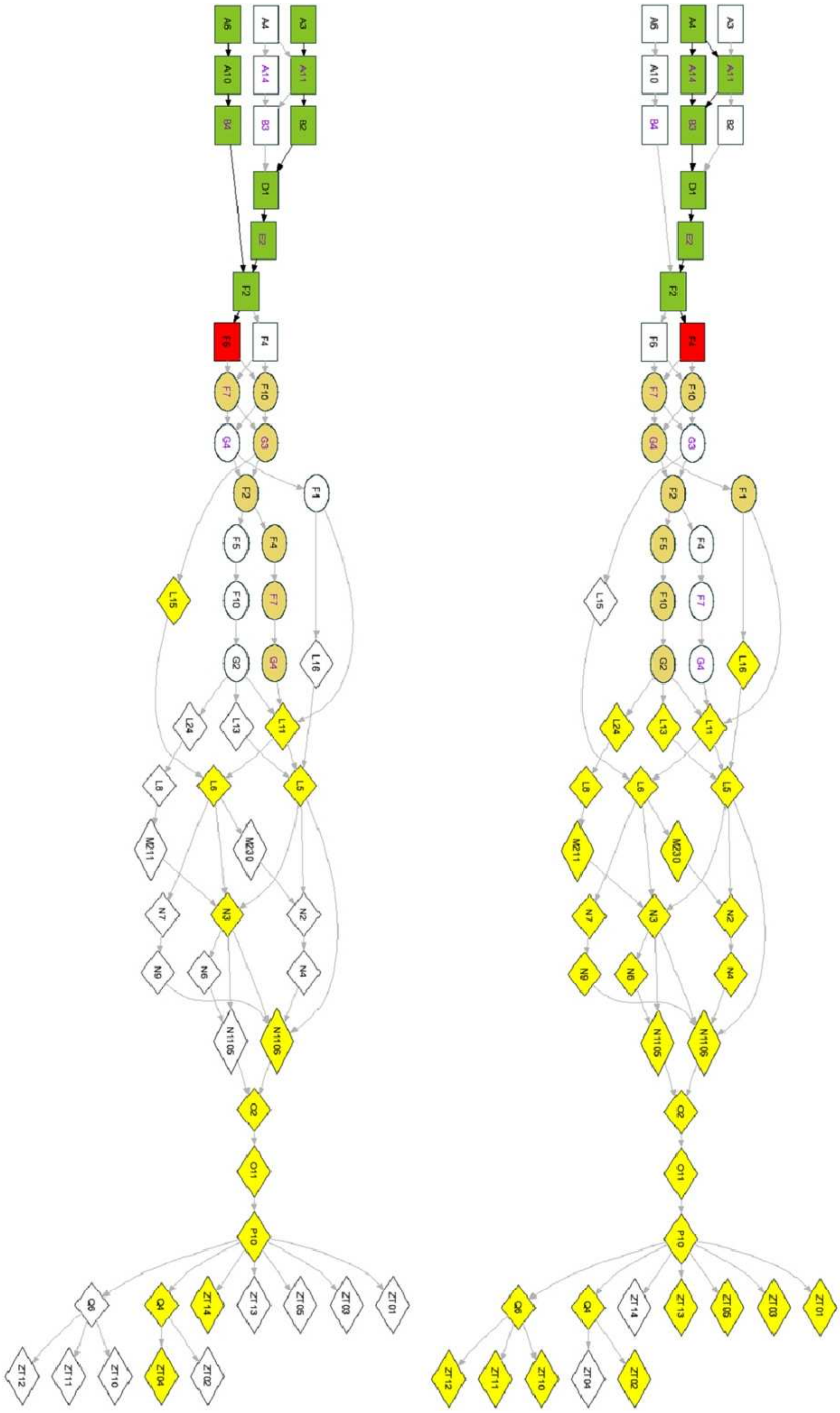
Main Zone	***AEZ description***	***Eval. crop production***	***Agric. potential***
Central Plateau	CPL1: Southern Omatako plains	Not suitable for crop production due to short dependable growing period, combined low water holding capacity and low nutrient status	Large livestock production on extensive grazing.
	CPL2: Fringe plains	Not suitable for crop production due to low water holding capacity and low nutrient status	Large livestock production on extensive grazing, limited cropping
Kalkfeld	KALK3: Kalkfeld	Unsuitable for crops due to shallow soils, on deeper soils however good for cropping.	Large livestock production on extensive grazing, limited cropping
Kalahari Sands Plateau	KAL1: stabilised WE dunes with few pans	Unsuitable for crops due to low dependable growing period and sandy soils.	Large livestock production on extensive grazing.
	KAL10: Tsumkwe panveld	Unsuitable for crops due to low dependable growing period and sandy soils. Small-scale irrigation (garden-scale) for food production possible.	Large livestock production on extensive grazing.
	KAL3-4: stabilised sand drift with few pans	Unsuitable for crops due to low dependable growing period and sandy soils.	Large livestock production on extensive grazing.
	KAL5: Incised river valleys	Unsuitable for crops due to low dependable growing period and sandy soils.	Large livestock production on extensive grazing.

Appendix D Overview of land use in Namibia (Mendelsohn *et al.*, 2002)and the agro-ecological zones in the study area after De Pauw & Coetzee (1999).





Appendix E Flowchart of LCCS classifiers used for the classification of Kalahari vegetation types. The nomenclature of the classifiers and a description is listed in table 2. The user label is specified in the *Floristic Aspect* classifier (e.g. ZT04 = *Eragrostis rigidior - Urochloa brachyura grasslands*). The structure of the classifier conjunctions visualises the characteristic layered vegetation structure of coexisting herbaceous and woody life-forms of open savanna environments, as defined in the *Main Layer* and Second and *Third Layer* classifiers. The upper flowchart visualises the classifiers used for the definition of vegetation types with a main shrub layer (F4). The lower chart shows the classifiers for a grassland class (F6). Using a set of standard class descriptions enables the translation of a floristic vegetation type class to the broader level describing physiognomic-structural level.



Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und unter Verwendung der angegebenen Hilfsmittel, persönlichen Mitteilungen und Quellen angefertigt habe.

Jena, April 2011 (Christian Hüttich)

Tabellarischer Lebenslauf

1. Berufliche Tätigkeiten

Seit 05/2010	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Fernerkundung des Instituts für Geographie und Geologie der Julius-Maximilians-Universität Würzburg - Mitarbeit im Projekt „GlobWetland-II", gefördert durch ESA
04/2007-04/2010	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Fernerkundung des Instituts für Geographie und Geologie der Julius-Maximilians-Universität Würzburg - Mitarbeit und Promotion im Projekt „BIOTA-Süd", gefördert durch BMBF
01/2007-04/2007	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Erdbeobachtung des Instituts für Geographie der FSU Jena - Erfassung von „Disturbances" mittels Landsat und MERIS Satellitendaten auf Testflächen in Sibirien zur Modellierung von Degradationserscheinungen im Permafrost
05/2005-09/2005	Studentische Hilfskraft am Lehrstuhl für Erdbeobachtung des Instituts für Geographie der FSU Jena
12/2004 - 01/2005	Hochschulpraktikum bei KARTA.GO – Gesellschaft für raumbezogenes Informationsmanagement mbH in Bonn
08/2004 - 10/2004	Hochschulpraktikum am Deutschen Fernerkundungs-Datenzentrum (DFD) des Deutschen Zentrums für Luft- und Raumfahrt (DLR) in Oberpfaffenhofen
07/2003 - 09/2003	Freie Mitarbeit im Ingenieurbüro JENA-GEOS in Jena
03/2002 – 10/2003	Studentische Hilfskraft am Lehrstuhl für Geoinformatik des Instituts für Geographie der FSU Jena

2. Ausbildung

WS 1998/99-SS 2006	Studium der Geographie an der Friedrich-Schiller-Universität in Jena, Abschluss: Diplom Thema der Diplomarbeit: „Assessing Land Cover Change Processes in Northern Eurasia based on SPOT-VEGETATION Time Series Analysis" Fachgebiete: Geoinformatik/Fernerkundung, Sozialgeographie Physische Geographie
09/1991 – 09/1997	Adolf-Reichwein Gymnasium, Jena Abschluss: Abitur

3. Publikationen

Artikel in wiss. Journalen (peer-reviewed)

Hüttich, C., Herold, M., Strohbach, B., & Dech, S. (2010). Integrating in-situ-, Landsat-, and MODIS data for mapping in Southern African savannas: Experiences of LCCS-based land-cover mapping in the Kalahari in Namibia. *Environmental monitoring and assessment*, *E-pub*. doi: 10.1007/s10661-010-1602-5.

Hüttich, C., Gessner, U., Herold, M., Strohbach, B. J., Schmidt, M., Keil, M., et al. (2009). On the Suitability of MODIS Time Series Metrics to Map Vegetation Types in Dry Savanna Ecosystems: A Case Study in the Kalahari of NE Namibia. *Remote Sensing*, *1*(4), 620-643. doi: 10.3390/rs1040620.

Hüttich, C., Herold, M., Schmullius, C., Egorov, V., & Bartalev, S. A. (2007). Indicators of Northern Eurasia's land-cover change trends from SPOT-VEGETATION time-series analysis 1998-2005. *International Journal of Remote Sensing*, *28*(18), 4199-4206. doi: 10.1080/01431160701442054.

Buchkapitel (peer-reviewed)

Keil, M., Gessner, U., & Hüttich, C. (2010). Using the MODIS Enhanced Vegetation Index (EVI) for BIOTA transect analyses. In Schmiedel & Jürgens (Eds.), *Biodiversity in southern Africa. Volume 2: Patterns and processes at regional scale* (Vol. 2, pp. 145-148). Göttingen, Windhoek: Klaus Hess Publishers.

Keil, M., Gessner, U., & Hüttich, C. (2010). Elevation profile along transects. In Schmiedel & Jürgens (Eds.), *Biodiversity in southern Africa. Volume 2: Patterns and processes at regional scale* (Vol. 2, pp. 81-83). Göttingen, Windhoek: Klaus Hess Publishers.

Keil, M., Gessner, U., Hüttich, C., & Colditz, R. (2010). Large-scale vegetation assessments in southern Africa: concepts and applications using multi-source remote sensing data. In Schmiedel & Jürgens (Eds.), *Biodiversity in southern Africa. Volume 2: Patterns and processes at regional scale* (Vol. 2, pp. 32-45). Göttingen, Windhoek: Klaus Hess Publishers.

Keil, M., Gessner, U., & Hüttich, C. (2010). Spatial patterns of topography in southern Africa. In U. Schmiedel & N. Jürgens (Eds.), *Biodiversity in southern Africa. Volume 2* (Vol. 2, pp. 29-31). Göttingen, Windhoek: Klaus Hess Publishers.

Vorträge

Hüttich, C., Strohbach, B., Herold, M., Keil, M., & Dech, S. (2010). The Potential of MODIS Time Series Metrics for bottom-up Vegetation Mapping in a semi-arid Savanna Ecosystem in Namibia. *Proceedings of the ESA Living Planet Symposium, 28. June - 2. July 2010, Bergen, Norway* (pp. 1-8).

Hüttich, C., Dech, S., Gessner, U., Keil, M., Schmidt, M., & Fox, T. (2009). Remote Sensing for Mapping Vegetation Types and Dynamics in Savanna Ecosystems of Namibia: Concepts for Integrated Vegetation Diversity Assessments. *2nd DIVERSITAS Open Science Conference, 13*

– 16 October 2009, Cape Town.

Hüttich, C., Herold, M., & Schmullius, C. (2007). Observing human activities in Northern Eurasia from DMSP nighttimes lights time series data. *58. Berg- und Hüttenmännischer Tag - Freiberger Forschungsforum 2007 Freiberg, Germany, June 13th - 15th 2007.*

Hüttich, C., Herold, M., Schmullius, C., Egorov, V., & Bartalev, S. A. (2006). Trends in Northern Eurasia's land cover derived on SPOT-VGT time series analysis 1998-2005. *2nd Göttingen GIS & Remote Sensing Days. Global Change Issues in Developing and Emerging Countries, 4-6 October 2006.*

Poster

Hüttich, C., Gessner, U., Schmidt, M., Keil, M., Strohbach, B., Herold, M., et al. (2009). Ecosystem Assessment in Savanna Ecosystems in Namibia based on multi-scale Earth Observation data. *Proceedings of the International Symposium of Remote Sensing for the Environment, 3-8 May.* Stresa, Italy.

Gessner, U., Conrad, C., Hüttich, C., Keil, M., Schmidt, M., Schramm, M., et al. (2008). A multi-scale approach for retreiving proportional cover of life forms. *Proceedings of the IEEE International Geoscience and RemoteSensing Symposium, Boston USA, July 6 - 8* (pp. 3-4).

Hüttich, C., Gessner, U., Keil, M., Schmidt, M., & Strohbach, B. J. (2008). Vegetation Type Mapping in NE-Namibia based on Botanical Assessments and multi-temporal Satellite Observations. *International Congress ´Biodiversity in Africa – Observation and Sustainable Management for our Future!´, 29 Sep-3 Oct 2008, Spier, South Africa.*

Hüttich, C., Gessner, U., Schmidt, M., Keil, M., & Dech, S. (2008). Vegetation structure and land dynamics in Southern Africa' s Savannas: Multi-scale earth observation data for ecosystem and biodiversity monitoring. *Geophysical Research Abstracts, Vol. 10, European Geosciences Union General Assembly 2008, 13-18th April, 10.*

Hüttich, C., Herold, M., & Schmullius, C. (2007). Land Cover Change in Northern Eurasia: Synergy indications from various coarse resolution satellite observations. *Proceedings of the Tri-National Conference 2007 of the Swiss, Austrian and German Society for Photogrammetry and Remote Sensing, 19.-21.06.2007, Basel, Switzerland.*

Hüttich, C., Herold, M., Schmullius, C., Egorov, V., & Bartalev, S. A. (2007). Indicators of Northern Eurasia's land-cover change trends from SPOT-VEGETATION time-series analysis 1998-2005. *Geophysical Research Abstracts, Vol. 9, European Geosciences Union General Assembly 2007, 15-20th April.*

Hüttich, C., Herold, M., Schmullius, C., Egorov, V., & Bartalev, S. A. (2005). Spot-VGT NDVI and NDWI Trends 1998-2005 as Indicators of recent Land Cover Change Processes in Northern Eurasia. *Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover* (pp. 336-344). Bonn.

Jena, April 2011 (Christian Hüttich)