Vegetation Dynamics in West Africa

Spatio-temporal Data Fusion for the Monitoring of Agricultural Expansion

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Julius-Maximilians-Universität Würzburg

vorgelegt von Kim Knauer
September 2017

(Titelbild: Mittlerer NDVI 2014 für Burkina Faso, abgeleitet aus Zeitreihen, die mittels des ESTARFM-Frameworks erstellt wurden)

Abstract

West Africa is one of the fastest growing regions in the world with annual population growth rates of more than three percent for several countries. Since the 1950s, West Africa experienced a fivefold increase of inhabitants, from 71 to 353 million people in 2015 and it is expected that the region's population will continue to grow to almost 800 million people by the year 2050. This strong trend has and will have serious consequences for food security since agricultural productivity is still on a comparatively low level in most countries of West Africa. In order to compensate for this low productivity, an expansion of agricultural areas is rapidly progressing. The mapping and monitoring of agricultural areas in West Africa is a difficult task even on the basis of remote sensing. The small scale extensive farming practices with a low level of agricultural inputs and mechanization make the delineation of cultivated land from other land cover and land use (LULC) types highly challenging. In addition, the frequent cloud coverage in the region considerably decreases the availability of earth observation datasets. For the accurate mapping of agricultural area in West Africa, high temporal as well as spatial resolution is necessary to delineate the small-sized fields and to obtain data from periods where different LULC types are distinguishable. However, such consistent time series are currently not available for West Africa. Thus, a spatio-temporal data fusion framework was developed in this thesis for the generation of high spatial and temporal resolution time series.

Data fusion algorithms such as the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) enjoyed increasing popularity during recent years but they have hardly been used for the application on larger scales. In order to make it applicable for this purpose and to increase the input data availability, especially in cloud-prone areas such as West Africa, the ESTARFM framework was developed in this thesis introducing several enhancements. An automatic filling of cloud gaps was included in the framework in order to use even partly cloud-covered Landsat images for the fusion without producing gaps on the output images. In addition, the ESTARFM algorithm was improved to automatically account for regional differences in the heterogeneity of the study region. Further improvements comprise the automation of the time series generation as well as the significant acceleration of the processing speed through parallelization. The performance of the developed ESTARFM framework was tested by fusing an 8-day NDVI time series from Landsat and MODIS data for a focus area of 98,000 km² in the border region between Burkina Faso and Ghana. The results of this test show the capability of the ESTARFM framework to accurately produce high temporal

resolution time series while maintaining the spatial detail, even in such a heterogeneous and cloud-prone region.

The successfully tested framework was subsequently applied to generate consistent time series as the basis for the mapping of agricultural area in Burkina Faso for the years 2001, 2007, and 2014. In a first step, high temporal (8-day) and high spatial (30 m) resolution NDVI time series for the entire country and the three years were derived with the ESTARFM framework. More than 500 Landsat scenes and 3000 MODIS scenes were automatically processed for this purpose. From the fused ESTARFM NDVI time series, phenological metrics were extracted and together with the single time steps of NDVI served as input for the delineation of rainfed agricultural areas, irrigated agricultural areas and plantations. The classification was conducted with the random forest algorithm at a 30 m spatial resolution for entire Burkina Faso and the three years 2001, 2007, and 2014. For the training and validation of the classifier, a randomly sampled reference dataset was generated from Google Earth images based on expert knowledge of the region. The overall classification accuracies of 92% (2001), 91% (2007), and 91% (2014) indicate the well-functioning of the developed methodology. The resulting maps show an expansion of agricultural area of 91% from about 61,000 km² in 2001 to 116,900 km² in 2014. While rainfed agricultural areas account for the major part of this increase, irrigated areas and plantations also spread considerably. Especially the expansion of irrigation systems and plantation area can be explained by the promotion through various national and international development projects. The increase of agricultural areas goes in line with the rapid population growth in most of Burkina Faso's provinces which still had available land resources for an expansion of agricultural area. An analysis of the development of agricultural areas in the vicinity of protected areas highlighted the increased human pressure on these reserves. The protection of the remnant habitats for flora and fauna while at the same time improving food security for a rapidly growing population, are the major challenges for the region in the future.

The developed ESTARFM framework showed great potential beyond its utilization for the mapping of agricultural area. Other large-scale research that requires a sufficiently high temporal and spatial resolution such as the monitoring of land degradation or the investigation of land surface phenology could greatly benefit from the application of this framework.

Kurzfassung

Westafrika ist eine der am schnellsten wachsenden Regionen der Welt. Das jährliche Bevölkerungswachstum liegt in mehreren Ländern der Region über drei Prozent. Seit den 1950er Jahren erlebte Westafrika eine fünffache Zunahme der Einwohner, von 71 auf 353 Millionen Menschen im Jahr 2015 und es wird erwartet, dass die Bevölkerung der Region bis zum Jahr 2050 auf fast 800 Millionen Menschen steigen wird. Dieser Trend hat ernsthafte Konsequenzen für die Nahrungsmittelsicherheit, da die landwirtschaftliche Produktivität in den meisten Ländern Westafrikas immer noch auf einem vergleichsweise niedrigen Niveau liegt. Um diese geringe Produktivität zu kompensieren, schreitet die Ausweitung der landwirtschaftlichen Flächen rasch voran. Die Kartierung und Überwachung der landwirtschaftlichen Gebiete in Westafrika ist eine schwierige Aufgabe, auch auf der Basis der Fernerkundung. Die kleinräumigen, extensiven Anbaumethoden mit geringem landwirtschaftlichen Input und mangelnder Mechanisierung stellen für die Abgrenzung der Anbauflächen von anderen Landbedeckungs- und Landnutzungstypen (LULC) eine große Herausforderung dar. Darüber hinaus wird die Verfügbarkeit von Erdbeobachtungsdaten durch häufige Wolkenbedeckung in der Region erheblich reduziert. Für die genaue Kartierung der landwirtschaftlichen Fläche in Westafrika ist eine hohe zeitliche sowie räumliche Auflösung notwendig, um die meist kleinen Felder abgrenzen zu können und Daten aus Zeiträumen zu erhalten, in denen unterschiedliche LULC-Typen unterscheidbar sind. Allerdings sind solch konsistente Zeitreihen für Westafrika derzeit nicht verfügbar. Daher wurde in dieser Arbeit ein raum-zeitliches Datenfusions-Framework entwickelt, um räumlich und zeitlich hoch aufgelöste Zeitreihen zu erzeugen.

Daten-Fusions-Algorithmen wie das Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) erfreuten sich in den letzten Jahren zunehmender Beliebtheit, wurden aber kaum auf größeren Skalen verwendet. Um die Nutzbarkeit für diesen Zweck zu verbessern und die Verfügbarkeit der Eingangsdaten vor allem in wolkenreichen Gebieten wie Westafrika zu erhöhen, wurde das ESTARFM-Framework in dieser Arbeit entwickelt. In diesem Framework wurden mehrere Verbesserungen eingebaut. Ein automatisches Füllen von Wolkenlücken wurde eingefügt, um auch teilweise wolkenbedeckte Landsat-Bilder für die Fusion zu verwenden, ohne Lücken auf den Ausgabebildern zu erzeugen. Darüber hinaus wurde der ESTARFM-Algorithmus verbessert, um regionale Unterschiede in der Heterogenität der Studienregion automatisch zu berücksichtigen. Weitere Verbesserungen umfassen die Automatisierung der Zeitreihenerzeugung sowie die signifikante Beschleunigung der

Prozessierungsgeschwindigkeit mittels Parallelisierung. Die Leistungsfähigkeit des entwickelten ESTARFM-Frameworks wurde durch die Fusion einer 8-tägigen NDVI-Zeitreihe aus Landsat- und MODIS-Daten für ein Fokusgebiet von 98.000 km² im Grenzgebiet zwischen Burkina Faso und Ghana getestet. Die Ergebnisse dieses Tests demonstrieren die Fähigkeit des ESTARFM-Frameworks, genaue, zeitlich hoch aufgelöste Zeitreihen zu erzeugen, die die räumlichen Details auch in einer solch heterogenen und wolkenreichen Region bewahren können.

Das erfolgreich getestete Framework wurde anschließend angewendet um konsistente Zeitreihen zu erzeugen, die als Grundlage für die Kartierung der landwirtschaftlichen Fläche in Burkina Faso für die Jahre 2001, 2007 und 2014 dienten. In einem ersten Schritt wurden mittels des ESTARFM-Frameworks zeitlich (8 Tage) und räumlich (30 m) hoch aufgelöste NDVI-Zeitreihen für das ganze Land und die drei Jahre abgeleitet. Mehr als 500 Landsat- und 3000 MODIS-Szenen wurden zu diesem Zweck automatisch verarbeitet. Aus den fusionierten ESTARFM NDVI-Zeitreihen wurden phänologische Metriken extrahiert und zusammen mit den einzelnen NDVI-Zeitschnitten als Basis für die Abgrenzung von Regenfeldbau, bewässerten landwirtschaftlichen Flächen und Plantagen genutzt. Die Klassifikation wurde mit dem Random-Forest-Algorithmus in einer räumlichen Auflösung von 30 m für ganz Burkina Faso und die drei Jahre 2001, 2007 und 2014 durchgeführt. Basierend auf Expertenkenntnissen der Region wurde ein zufällig beprobter Referenzdatensatz aus Google-Earth-Bildern generiert, der für die Kalibrierung und Validierung des Klassifikators diente. Die Klassifikationsgenauigkeiten von 92% (2001), 91% (2007) und 91% (2014) demonstrieren die Leistungsfähigkeit der entwickelten Methodik. Die daraus erzeugten Karten zeigen eine Ausweitung der landwirtschaftlichen Fläche um 91%, von etwa 61.000 km² im Jahr 2001 auf 116.900 km² im Jahr 2014. Während der Regenfeldbau den größten Teil dieser Zunahme ausmacht, haben sich auch bewässerte Flächen und Plantagen erheblich ausgebreitet. Insbesondere die Ausweitung der Bewässerungssysteme und der Plantagenfläche lässt sich mit der Förderung durch verschiedene nationale und internationale Entwicklungsprojekte erklären. Die Zunahme der landwirtschaftlichen Gebiete geht einher mit dem raschen Bevölkerungswachstum in den meisten Provinzen Burkina Fasos, welche noch über ungenutzte Landressourcen verfügen. Eine Analyse der Entwicklung der landwirtschaftlichen Fläche im Umland von Schutzgebieten verdeutlicht den erhöhten anthropogenen Druck auf diese Reservate. Der Schutz der verbliebenen Habitate für Flora und Fauna bei gleichzeitiger Verbesserung der Nahrungsmittelsicherheit für eine schnell wachsende Bevölkerung stellen die großen Herausforderungen für die Zukunft der Region dar.

Das entwickelte ESTARFM-Framework zeigt großes Potenzial für weitere Anwendungen, die über die Kartierung landwirtschaftlicher Fläche hinausgehen. Großflächige Forschung, die eine hinreichend hohe zeitliche und räumliche Auflösung erfordern, wie die Überwachung von

Bodendegradation oder die Untersuchung von Phänologie der Landoberfläche, könnten von der Anwendung des ESTARFM-Framework erheblich profitieren.

Acknowledgements

This work would not have been possible without the great support in all kinds of ways from various persons. First of all, I would like to thank Dr. Claudia Künzer for supervising my thesis, for very helpful discussions and for the guidance towards the submission of this work. Her supervision and commitment undoubtedly advanced me as a scientist. I am very thankful to Dr. Christopher Conrad for spontaneously taking over my second supervision and wasting his Christmas time in order to finish his report on this thesis. I also want to thank Prof. Dr. Rasmus Fensholt from the University of Copenhagen for his expert advice and feedback whenever I needed it and for welcoming me in Copenhagen. I am thankful for being given the opportunity to work and conduct my thesis at the German Remote Sensing Data Center (DFD) of the German Aerospace Center (DLR) and using the available resources necessary for the accomplishment of my work.

I was able to work in a great and highly motivating team that was always available for advice and discussion. Thanks go to Dr. Andreas Dietz, Dr. Corinne Frey, Juliane Huth, Dr. Christina Eisfelder, Verena Jaspersen, Malte Ahrens, Igor Klein and Dr. Patrick Leinenkugel for feedback, technical support and making this a good time. From this group, I have to specifically mention my mentor Dr. Ursula Geßner who was always there for helpful ideas, who lifted my spirits after setbacks or re-sharpened my focus when I was about to lose the bigger picture. I also very much appreciate the time with my PhD colleagues Christian Wohlfart, Kersten Clauss, Emmanuel Da Ponte and Marco Ottinger with whom I went through all the ups and downs that such a big work comes with. Special thanks also goes to Dr. Benjamin Mack for his great assistance in R project.

This thesis was conducted in the framework of the WASCAL project funded by the BMBF. I am thankful to all colleagues within the project who supported me during my field work, with administrative issues or with data and advice on my studies. I also want to thank Dr. Xiaolin Zhu for providing the code of the original ESTARFM and for some helpful comments. Further thanks go to Professor Crystal Schaaf and Dr. Qingsong Sun who kindly provided the new MODIS collection 6 data before the official distribution via the USGS MODIS websites.

What could I have accomplished without my friends and family who always backed me up and made my life a better one? I am very grateful to my parents Anita and Hartmut and my grandmother Lilo for their love and endless support. Last but not least, I want to thank Franziska for her love, for always being there for me, for never losing patience, for going without a lot of things, for always cheering me up…

Table of Contents

List of Figures

List of Tables

Glossary

ACi	African Cashew initiative
AGC	Apparent Green Cover
ANN	Artificial Neural Network
ASAR	Advanced Synthetic Aperture Radar
AVHRR	Advanced Very High Resolution Radiometer
BFA	Burkina Faso
BFAST	Breaks for Additive Seasonal and Trend
BMBF	Bundesministerium für Bildung und Forschung (German Federal Ministry of Education and Research)
BRDF	Bidirectional Reflectance Distribution Function
CERES	Crop Environment Resource Synthesis
CCI	Climate Change Initiative
CMAP	Climate Prediction Center Merged Analysis of Precipitation
CPU	Central Processing Unit
DFD	Deutsches Fernerkundungsdatenzentrum (German Remote Sensing Data Center)
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
DOY	Day Of the Year
ECOWAS	Economic Community of West African States
ENSO	El Niño-Southern Oscillation
EOS	Earth Observation System
EROS	Earth Resources Observation and Science Center
ESA	European Space Agency
ESTARFM	Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model
ETM+	Enhanced Thematic Mapper
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FEWS	Famine Early Warning System Networks
FIP	Forest Investment Program

Glossary

FPAR	Fraction of absorbed Photosynthetically Active Radiation
GDP	Gross Domestic Product
GFSAD30	Global Food Security-Support Analysis Data at 30 m
GIMMS	Global Inventory Modelling and Mapping Studies
GIZ	German Society for International Cooperation
GLC	Global Land Cover
GPCP	Global Precipitation Climatology Project
GPCC	Global Precipitation Climatology Centre
GPP	Gross Primary Production
GUF	Global Urban Footprint
GWI	Global Water Initiative
HDR	Human Development Index
HWSD	Harmonized World Soil Database
IDL	Interactive Data Language
IGB	Institut Géographique du Burkina Faso
IGS	International Ground Station
INSD	Institut National de la Statistique et de la Démographie
ISODATA	Iterative Self-Organizing Data Analysis
ITCZ	Intertropical Convergence Zone
JRC	Joint Research Centre
KTNP	Kaboré Tambi National Park
LAI	Leaf Area Index
LGAC	Landsat Global Archive Consolidation
LPJ-DGVM	Lund–Potsdam–Jena Dynamic Global Vegetation Model
LSP	Land Surface Phenology
LSM	Linear Spectral Mixture
LULC	Land Use and Land Cover
MAE	Mean Absolute Error
MERIS	Medium Resolution Imaging Spectrometer
MMT	Multisensor Multiresolution Technique
MODIS	Moderate-Resolution Imaging Spectroradiometer
MRT	MODIS Reprojection Tool
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NIR	Near InfraRed

NPP	Net Primary Productivity
OLI	Operational Land Imager
PNSR	National Programme for the Rural Sector
RLCM	Rapid Land Cover Mapper
RMSE	Root Mean Square Error
RUE	Rain-Use Efficiency
SAR	Synthetic Aperture Radar
SARRA-H	System for Regional Analysis of Agro-Climatic Risks
SDR	Rural Development Strategy
SLC	Scan Line Corrector
SPOT	Satellite Pour l'Observation de la Terre
SPSTFM	SParse-representation-based SpatioTemporal reflectance Fusion Model
SSC	Shoulder Scene Combination
SST	Sea Surface Temperature
STARFM	Spatial and Temporal Adaptive Reflectance Fusion Model
STRUM	Spatial and Temporal Reflectance Unmixing Model
SVM	Support Vector Machines
RF	Random Forest
TM	Thematic Mapper
TRMM	Tropical Rainfall Measuring Mission
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
VHR	Very High Resolution
VI	Vegetation Index
WASCAL	West African Science Service Center on Climate Change and Adapted Land Use

Definitions

Agricultural area, (rainfed) agricultural area is defined in this work as farmed land with partial tree cover (<10%, potentially fruit trees) comprising the predominant crop-fallow rotation system of annual crops and intensive pasture, i.e. shrub-free, partially fenced areas regularly used for grazing.

Agricultural expansion, temporary or permanent increase of land under agricultural use, most commonly at the expense of (semi-)natural areas.

Degradation, "land degradation means reduction or loss, in arid, semi-arid and dry sub-humid areas, of the biological or economic productivity and complexity of rainfed cropland, irrigated cropland, or range, pasture, forest and woodlands resulting from land uses or from a process or combination of processes, including processes arising from human activities and habitation patterns […]" (UNCCD, 2017).

Desertification, "desertification means land degradation in arid, semi-arid and dry sub-humid areas resulting from various factors, including climatic variations and human activities" (UNCCD, 2017).

Food security, "food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life" (World Food Summit, 1996).

Intensification of agriculture, intensification of agriculture can be defined as the increase in input (fertilizer, herbicides, improved seeds, mechanization etc.) to increase yield and productivity.

Sustainability of agriculture, "the goal of sustainable agriculture is to meet society's food and textile needs in the present without compromising the ability of future generations to meet their own needs. Practitioners of sustainable agriculture seek to integrate three main objectives into their work: a healthy environment, economic profitability, and social and economic equity" (UCDAVIS ASI, 2018).

Vegetation dynamics, vegetation dynamics is the quantitative and qualitative variation in plant communities over time. This can be the variation over the annual phenological cycle, the vegetation variability over several years, or a progressive change in vegetation over a longer period of time (Knauer et al., 2014).

1 Introduction

For the analysis of global prospects, West Africa is one of the most important regions in the world with great potentials but also considerable risks. Since the 1950s, West Africa experienced a fivefold increase of inhabitants, from 71 to 353 million people in 2015 (CILSS, 2016; United Nations, 2015). During the same time, the world population increased less than threefold. While the world population number is gradually stabilizing, the annual growth rates in West Africa are still enormous with more than 3% for some countries (OECD, 2009). It is one of the last regions in the world which is just at the beginning of demographic transition. Several West African countries are still experiencing constantly high birth rates while the death rates are already decreasing, resulting in a considerable population growth (OECD, 2009). Although predictions have to be taken with caution and depend on multiple factors in the region, it is expected that West Africa will grow to 516 million people in 2030 and to almost 800 million people by the year 2050 (United Nations, 2015). In the year 2063, the region could even reach a number of 1 billion people so that every tenth person living on earth would be West African. The enormous past and future population growth has serious implications for food security and environment (for definition of food security see page XX). In addition, major parts of the population are living below the poverty line despite an emerging economic growth in the region. Among the 14 least developed countries, eight are from West Africa with Niger having the lowest Human Development Index (HDR=0.348) of all countries globally (UNDP, 2015). The low development of most West African countries hinders the coping with the negative effects of population growth, especially in terms of food security.

Nonetheless, West Africa and especially its Sudanian Savanna zone is said to have a high agricultural potential (Callo-Concha et al., 2013; Vlek et al., 2010). If this potential could be exploited, the expected yields could stabilize food security in the region and West Africa could even become a major player in the global food market (NEPAD, 2013). However, so far, West Africa is only producing about 10% of its actual potential which is not enough for several countries to feed its own population (Brown et al., 2009). This situation could become even more critical in the future with rapidly increasing population. In order to compensate this trend and improve food security, yields would have to be increased considerably but during the last 50 years this could not be accomplished (FAO, 2016). The major reasons for this are the commonly inefficient farming practices in the region. A lack of own capital and access to credits results in little investment of the local farmers in their production. Thus, West African

agriculture is characterized by small-scale manual farming with a low level of agricultural inputs, mechanization, fertilizer and pesticide application as well as irrigation (FAO, 2014; Knauer et al., 2017). This low agricultural productivity is compensated in West Africa by a considerable expansion of cultivated land. However, this cannot be a sustainable solution as arable land is limited in West Africa. For example, it is expected that Burkina Faso will reach its maximum extent of cultivated land by the year 2030 (Knauer et al., 2017; Mathys and Gardner, 2009).

The expansion of cultivated land is seen as a main reason for land use and land cover (LULC) changes and a considerable threat to the environment, not only for West Africa but for the world in general (Ramankutty et al., 2008). It is estimated that about 40% of the earth's ice-free land surface is used for agriculture replacing the natural vegetation of savanna, grasslands and forests (Foley et al., 2005; Ramankutty et al., 2008). The clearing of tropical forests alone is responsible for about 12 to 16% of the annual anthropogenic CO_2 emissions to the atmosphere (Ramankutty et al., 2008). This deforestation again has a negative impact on regional and global climate and as such on regional farming. West Africa is already seen as one of the most vulnerable regions to climate change which could increase in the near future (Boko et al., 2007). The climate models consistently predict an increase in mean annual temperatures for the region of about 3 to 4°C until the end of the century (Boko et al., 2007). In addition, a higher variability of the rainy season and more extreme rainfall events are likely to occur. These changes will have serious consequences for human livelihood such as more frequent floods and droughts or a decrease in water availability. This will further affect agricultural production and food security by impacting yields and crop failure. To compensate these climatic effects on crop yields, a further expansion of agricultural land can be expected.

In order to break the cycle of agriculturally driven LULC conversion and climate change impacts on the region, suitable measures have to be implemented, to promote intensification on existing fields, to improve the adaptation capacity and implement a more a sustainable agriculture (for definitions of intensification and sustainability in agriculture see page XX). This can only be accomplished when the interactions between agricultural expansion, population growth, environment and climate change are fully understood (for definition of agricultural expansion see page XX). Thus, the starting point for such research should be the accurate mapping of agricultural area and its past development (for definition of agricultural area see page XX). Exact information on extent and location of agriculture over several years can serve as a basis for all subsequent analyses and modelling of regional climate change, future land use, food production and environmental consequences. For large regions on national or even supranational scale like West Africa, the only practical basis for the mapping and monitoring of agricultural area is remote sensing. During the last decades, several global or continental remote sensing based mapping efforts of land use and land cover have been undertaken that

also cover West Africa but their use on regional level was found to be limited (Gessner et al., 2012; Lambert et al., 2016).

1.1 The Need for Mapping of Agricultural Area in West Africa

So far, accurate information on the extent and development of agricultural area in West Africa is missing (section 3.3). The region is highly challenging for the derivation of such information based on remote sensing data due to its predominating agricultural practices, i.e. low degree of mechanisation and fertilizer input, small-sized fields with often partial tree cover. The frequent cloud cover during the growing season further hinders remote sensing approaches in West Africa. Consequently, available LULC classifications have their difficulties in accurately delineating agriculture in the region. However, there is a great demand for such information since it is required as a basis for a wide range of studies. Different regional climate or economic models that investigate the interplay between agricultural land use and environment or food markets can benefit from accurate maps of agricultural area (Forkuor, 2014). A proper prediction of expected national or provincial yields has to be based on information about the extent of cultivated land. If this input is already in a very coarse spatial resolution (Thornton et al., 1997) or of insufficient quality, the resulting yield estimations inherit considerable uncertainties.

Accurate and countrywide information on agricultural area can help to understand its spatio-temporal development. Such information can identify possible hotspots of expansion. Furthermore, these datasets can be coupled with population data in order to see where an increase in agricultural area can potentially improve food availability and where population growth cannot be compensated with an expansion of farmland. On the other hand, knowing the inventory and spatial distribution, available areas for a possible future expansion can be outlined. By monitoring the past development and annual growth of agricultural areas, it can also be estimated when a country will reach its limits of arable land. When the maximum extent of arable land is cultivated while the demand for food is still continuously increasing, it could have serious consequences for the food security of a country.

In addition to being a valuable source of information for food security studies, agricultural change maps also allow to follow and understand agricultural expansion from an environmental point of view. As indicated above, the ongoing conversion of natural vegetation into farmland is a serious threat to biodiversity, especially in the region of West Africa. The shrinking of forest and savanna areas has considerable consequences for their ecosystem services and functions. Increasing fragmentation of protected areas and a decreasing number and size of natural patches in between endangers genetic exchange among animal populations. These remnant reserves are threatened by an increasing population in their surroundings and their need for additional space. Insufficient protection and guarding can

thus result in uncontrolled exploitation of protected areas for firewood, meat or new fields. The monitoring of agricultural area in their vicinity can help to protect such reserves and control the compliance of their borders.

All in all, accurate information on the extent and development of agricultural area in West Africa can serve the sustainable management and policy planning on national level. On the basis of suitable maps, actions on regional development or nature conservation could be initiated and their effects could be monitored by a repeated classification of the country area.

In order to account for the comprehensive anthropogenic need of space for agricultural production, (rainfed) agricultural area is defined in this work as farmed land with partial tree cover (<10%, potentially fruit trees) comprising the predominant crop-fallow rotation system of annual crops and intensive pasture, i.e. shrub-free, partially fenced areas regularly used for grazing (Knauer et al., 2017).

1.2 Objectives, Structure and Context of the Thesis

The primary goal of this study is to map agricultural area in a high spatial resolution on a national scale for entire Burkina Faso and multiple years. Subsequently, changes during the period of investigation, reasons for different developments and relationships to population growth as well as implications for the environment shall be investigated. The challenging task can be divided into two main objectives:

I.   Development of a suitable earth observation based methodology for the derivation of agricultural area in small-scale farming systems with high temporal variability at high spatial resolution and national level.

As availability of earth observation data at high temporal and spatial resolution is limited for the study region, this objective shall be achieved by the development of a novel spatio-temporal data fusion framework in order to improve past data availability of remote sensors in the region. This framework shall be used to generate time series of high spatial and temporal resolution for entire Burkina Faso and the years 2001, 2007, and 2014. These synthetic time series are expected to improve the envisaged classification of agricultural areas in the country.

II. Mapping, quantification and analysis of agricultural expansion in Burkina Faso since the year 2001.

The spatio-temporal development of agriculture in Burkina Faso over 14 years shall be analysed and local differences and hotspots are identified. Furthermore, the relationship of

these developments to population growth will be investigated. Finally, implications of spreading agriculture for the status and conservation of protected areas will be examined.

The presented work was conducted in the context of the WASCAL (West African Science Service Center on Climate Change and Adapted Land Use) project (BMBF, 2017). The major goal of WASCAL is the development of science based services and effective adaptation and mitigation measures enhancing the resilience of anthropogenic and environmental systems to climate change and increased climate variability (BMBF, 2017). Researchers of different thematic fields such as climatology, agronomy or social sciences collaborate to improve the understanding of climate change and its impacts on the region and how to better adapt to it. Burkina Faso is in the centre of the WASCAL research activities and the derived maps can serve as input for climate, agronomic or economic models and are an important basis for the development of WASCAL's service provision. Furthermore, they can be applied to upscale local in-depth analyses on agricultural land use.

From the presented work, three SCI-listed journal publications and four contributions at national and international scientific conferences have already emerged and are listed below. It is one of the foremost principles of the WASCAL project to openly share information and results with the project partners, the scientific community and the affected people in order to increase knowledge about very recent developments and improve adaptation capacity. Withholding significant results would impede scientific discussion on the challenges of the region and hinder the development of possible solutions.

SCI-listed publications as first author:
- Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2014): Remote Sensing of Vegetation Dynamics in West Africa. *International Journal of Remote Sensing. 35 (17). 6357-6396*.
- Knauer, K.; Gessner, U.; Fensholt, R. and Kuenzer C. (2016): An ESTARFM Fusion Framework for the Generation of Large-Scale Time Series in Cloud-Prone and Heterogeneous Landscapes. *Remote Sensing. 8 (425).*
- Knauer, K.; Gessner, U.; Fensholt, R., Forkuor, G. and Kuenzer C. (2017): Monitoring Agricultural Expansion in Burkina Faso over 14 Years with 30 m Resolution Time Series: The Role of Population Growth and Implications for the Environment. *Remote Sensing. 9 (132).*

Conference contributions:
- Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2014): Challenges in the generation of LAI time series for West Africa: difficulties and potential improvements. Poster presentation at the GV2M, Avignon, 3rd-7th February 2014.

- Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2014): Remote sensing of vegetation dynamics in West Africa: improved satellite time series for phenological analyses. Poster presentation at the 5th ESA Advanced Training Course on Land Remote Sensing, Valencia, 8th-12th September 2014.

- Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2015): Remote sensing of vegetation dynamics in West Africa: improved satellite time series for phenological analyses. Poster presentation at the 36th ISRSE, Berlin, 11th-15th May 2015.

- Knauer, K.; Gessner, U.; Dech, S.; Fensholt, R. and Kuenzer, C. (2015): Fernerkundung der Vegetationsdynamik in Westafrika: Verbesserte Satelliten-Zeitreihe für landwirtschaftliche Analysen. Oral presentation at the Arbeitskreis Subsaharisches Afrika, Köln, 27th – 28th November 2015.

The presented work is subdivided into seven major chapters structuring the methodological procedure. First, the reader is introduced into the study area of West Africa with special focus on Burkina Faso, outlining physical as well as socio-economic characteristics (chapter 2). Then, the methodological background of the planned work is described and previous scientific studies on crucial aspects of this work are reviewed (chapter 3). A focus here lies on different approaches of agricultural area classification, currently available maps and challenges for the mapping in West Africa (section 3.3). This chapter also includes a pre-study for the identification of research gaps in the field of remotely sensed vegetation dynamics in West Africa. Hereafter, the data basis of this work is described including the used remote sensing datasets and their pre-processing, the generation of the reference database for the classification and the application of additional data and information (chapter 4).

The following two chapters are addressing the main objectives of this work as outlined above: in chapter 5, the development of a novel framework for spatio-temporal data fusion of dense, high spatial resolution time series is presented. Subsequently, the accuracy of this method is assessed and the suitability for the discrimination of different prevailing LULC classes is investigated in a focus area in southern Burkina Faso. In chapter 6, the developed methodology is applied to generate dense time series for entire Burkina Faso and three years (2001, 2007, and 2014) and the extent of agricultural area is derived from these datasets on the basis of land surface phenology (LSP) metrics. Subsequently, the agricultural expansion in Burkina Faso is investigated with respect to differences in the spatio-temporal development, its relationship to population growth and implications for protected areas.

Finally, the work is concluded with a summary of this work and its results (chapter 7). The suitability of the developed methodology for the accomplishment of the work's objectives is discussed and an outlook on further improvements and future research questions.

2 Study Area

The geographic focus of this work is on the continental part of West Africa and on the country of Burkina Faso in particular. According to the United Nations (United Nations Statistics Divison, 2013), continental West Africa consists of 15 countries (Figure 2.1): Benin, Burkina Faso, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone and Togo (excluding the islands of Saint Helena and Cabo Verde). The region extents from 18° West to 16° East and from 4° North to 27° North and covers an approximate area of 6.1 million km².
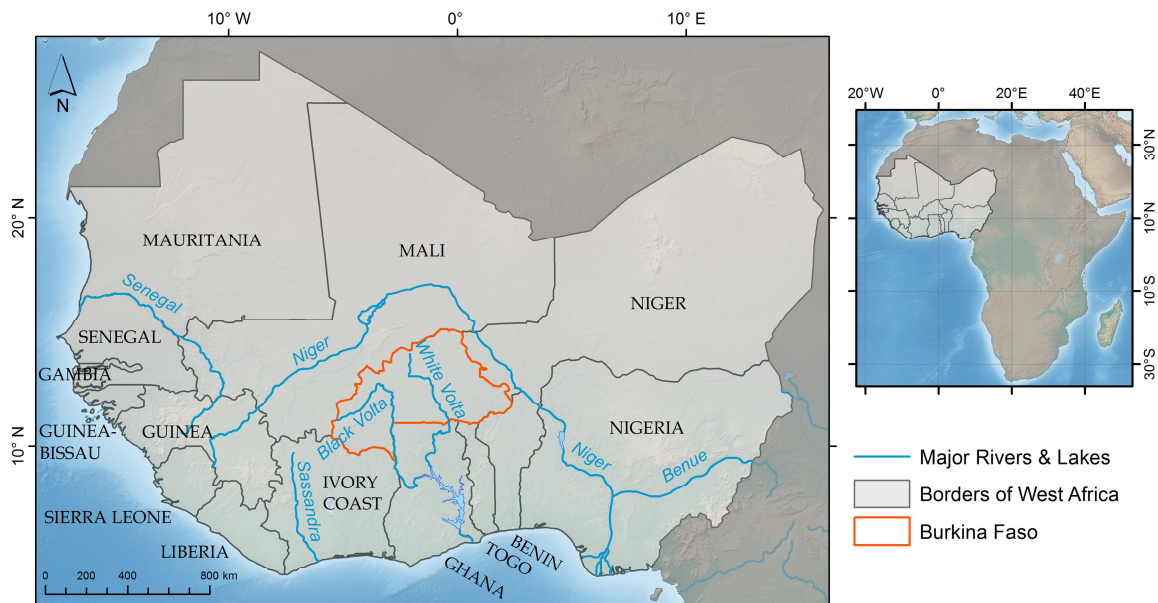
Figure 2.1: Countries and major rivers and lakes of West Africa.

Burkina Faso lies in the center of West Africa with an approximate size of 274,000 km². The country is in many ways representative for the region of West Africa: livelihood and vegetation are highly influenced by the prevailing climate, especially by the timing and length of the rainy season; as most countries in West Africa, Burkina Faso is characterized by small-scale rainfed subsistence farming and its population is experiencing a rapid growth which implicates considerable issues for food security. The country is located in the center of the West Sudanian and Sahelian Savanna, an area of high agricultural potential (Callo-Concha et al., 2013; Vlek et al., 2010). The exploitation of this potential could not only increase food

security within Burkina Faso but could also improve the situation in the greater region of West Africa. Thus, Burkina Faso is especially interesting as a focus country within the region for the investigation of agricultural development as conducted in this work.

In the following sections, the biophysical and socio-economic characteristics of West Africa and Burkina Faso in particular are presented. First, the regional climate, the topography and soils as well as the prevailing semi-natural vegetation are outlined. Hereafter, the regional population distribution and development are described and the anthropogenic land use is presented.

2.1 Climate

The climate of West Africa is primarily influenced by largescale processes of the atmosphere, i.e. the global circulatory system. Two large air masses meet over West Africa – the dry and hot winds of the Harmattan from the Sahara in the north and the humid, maritime winds of the Monsoon from the Atlantic Ocean in the south. The area where these two air masses meet is called the Inner Tropical Convergence Zone (ITCZ) and its annual north-south movement is the main influence factor on the formation of rainy seasons in the region. During summer of the northern hemisphere, the ITCZ is shifting north and the moist winds of the Monsoon are moving inland bringing rainfalls to the northern countries of West Africa. While the southern coastal areas of West Africa are experiencing two rainy seasons, only one rainy season occurs in the north. During the dry season, the ITCZ is shifting to the south and the Harmattan is bringing hot and dry air gathered over the Sahara.

As a result of the ITCZ movement, the rainfall distribution generally follows a north-south gradient in West Africa (Figure 2.2). While the northern regions bordering the Sahara experience mean annual rainfalls of less than 100 mm, the southern coastal regions receive up to 2000 mm and locally even more. Due to orographic effects and the specific movement of the Monsoon, the coastal regions of Sierra Leone and Guinea receive more than 4000 mm of mean annual rainfall. In Burkina Faso, the mean annual rainfall ranges from about 300 mm in the north to about 1100 mm in the south. However, the actual amount of rainfall in a single year and the timing and length of the rainy season are varying considerably in West Africa. In the north including Burkina Faso, the rainy season occurs between May and October and lasts for about three to four months with a general peak between July and August. In the southern regions, the two rainy seasons occur between March and July and between September and October (Knauer et al., 2014).

Contrasting the rainfalls, the mean annual temperatures increase from south to north. In the south, the temperatures vary between 24°C and 30°C with a daily amplitude of 3 to 5 °C (Hayward and Oguntoyinbo, 1987). In the northern regions, the mean annual temperatures lie

between 27 and 36°C with a considerably higher daily amplitude of 8 to 14 °C (Hayward and Oguntoyinbo, 1987).
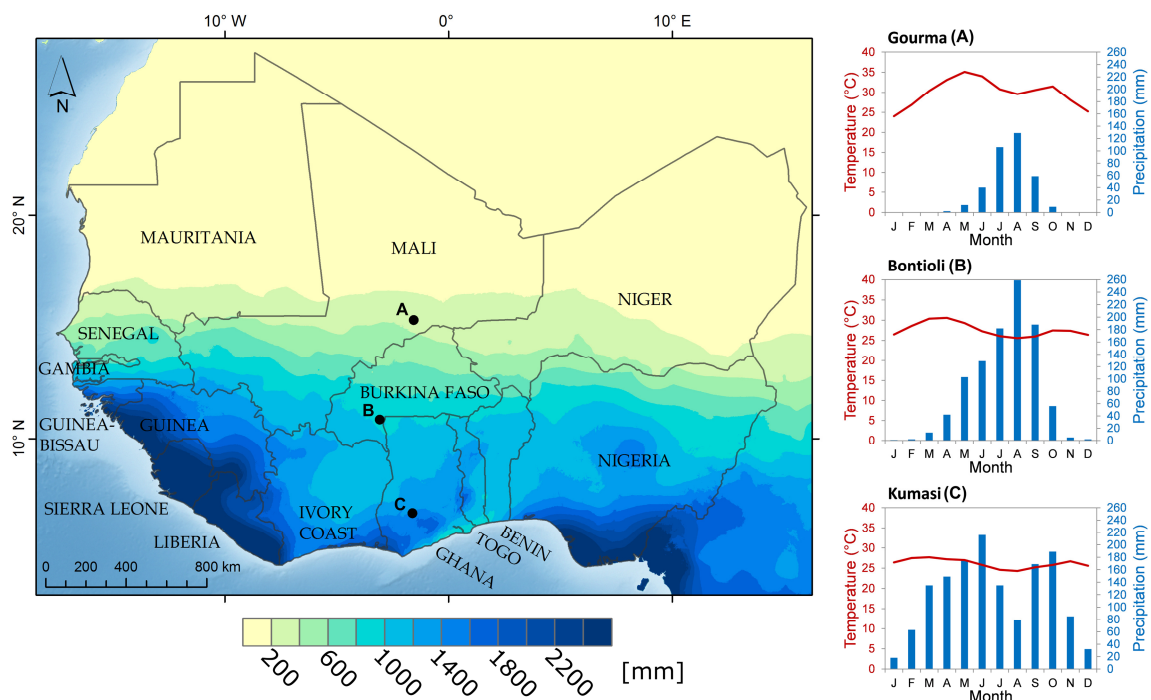
Figure 2.2: Long-time mean annual rainfall in mm and climate diagrams of three selected locations (Data Source: WorldClim, Hijmans et al. 2005).

Numerous studies investigated the influence of climate change on rainfall and temperature in West Africa (e.g. Neumann et al., 2007; Paeth et al., 2008; Paeth and Thamm, 2007). In general, climate change is predicted to increase the annual variability of the rainy season resulting in more serious floods and droughts (Boko et al., 2007; OECD, 2009). Several studies have detected a substantial decrease of rainfall in West Africa over the past 60 years (Nicholson et al., 2000; OECD, 2009). Especially in the northern regions of West Africa, i.e. the Sahel, there has been a continued reduction since the late 1960s with several periods below the long-term average (Nicholson et al., 2000; OECD, 2009). However, since the 1990s, an improvement of rainfall conditions could be detected from the records, especially in the Sahel region. This might be seen as a recovery of rainfalls from the previously dry years but its sustainability is unsure (OECD, 2009). During the past decades, a rise in the temperatures of about 1°C was detected for West Africa (Boko et al., 2007). This trend is estimated to continue in the future with a modelled increase of 3 to 4°C until the end of the century (Boko et al., 2007). The predicted trends in rainfall and temperature are expected to have serious consequences for the regional agricultural production and thus for food security in West Africa.

2.2 Geology, Topography and Hydrology

The geologic base of West Africa is dominated by the Precambrian West African Craton which has been folded and fractured since its formation (Büdel, 1981). In the Palaeozoic, several shallow seas covered West Africa, eroded and deposited considerable amounts of sediments which subsequently consolidated and today overlay large parts of the Precambrian rocks. In the following, large sandstone formations weathered during dry periods of the late Quaternary forming great sand sheets that further filled the relief of West Africa (CILSS, 2016). Although the general elevation of the region is rather low with an average of about 400 m above sea level, parts of the folded Precambrian Craton still emerge from the general level like the Aïr Mountains in northern Niger or the Fouta Djallon highlands in the centre of Guinea (CILSS, 2016). The latter is also called the 'water tower' of West Africa, since most of the great rivers of the regions spring from these highlands. The above mentioned orographic effects bring considerable rainfalls to the Fouta Djallon, which for example feed the Senegal River, the Gambia River or the Niger River (compare Figure 2.1). A further great river system of West Africa is the Volta River. It originates in Burkina Faso from its two main tributaries, the White Volta (Nakanbé) and the Black Volta (Mouhoun). Before it drains into the Gulf of Guinea it is dammed to form Lake Volta in Ghana, the biggest dammed lake in the world. In Burkina Faso, there are numerous small water reservoirs stemming from the damming of rivers but natural lakes are rare. The biggest reservoir of the country is formed at the Bagré dam retaining water from the White Volta.

2.3 Soils

The soils in West Africa are mainly influenced by the granitic bedrock material from the West African Craton and the past and present climatic conditions. Thus, the most northern part of West Africa bordering the Sahara Desert is covered with sand dunes (Figure 2.3). Adjacent to the south and approximately corresponding with the Sahel, a broad belt of Arenosols dominates the region. These comprise of deep sandy soils and developed in the region on sand dunes which were immobilized by vegetation cover. Due to their coarse texture and high permeability, they are characterized by a rather low water and nutrient storage capacity (IUSS Working Group WRB, 2015). Although these soils are not specifically fertile, dry farming is possible where annual rainfall exceeds 300 mm like in the north of Burkina Faso (IUSS Working Group WRB, 2015). However, these soils are highly sensitive to erosion which can result from overgrazing or cultivation without sufficient soil conservation measures (IUSS Working Group WRB, 2015).

Plinthosols and Lixisols cover great central parts of West Africa including southern Burkina Faso, southern Mali and Ghana. The humid climate of the past as well as the current sub-humid climate of the respective areas led to strong weathering of these soils (IUSS

Working Group WRB, 2015). Both soil types are characterized by a rather low fertility and require specially adapted tillage and erosion measures in order to conserve the soil for cultivation (IUSS Working Group WRB, 2015).

Covering the major part of the Ivory Coast, southwest Ghana and southeast Guinea, Acrisols are primarily found in the humid subtropical areas of West Africa and are generally acidic and highly weathered soils (Jalloh et al., 2011). Towards the south, in the tropical regions of West Africa, in Liberia, Sierra Leone and coastal Guinea, this weathering and leaching increases even more resulting in the typically yellow or reddish Ferralsols. Again, due to their rather low fertility, these soils are of limited suitability for agriculture and are ideally cultivated with fertilizer input.
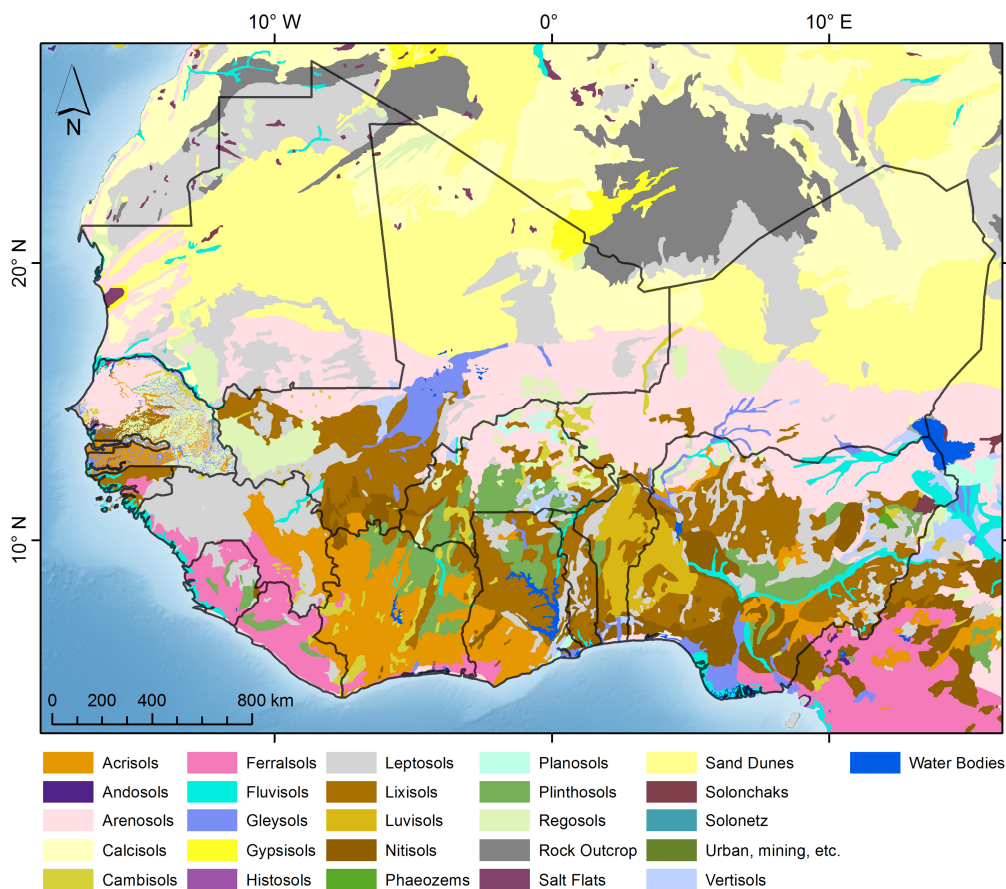
Figure 2.3: Soil types of West Africa according to the FAO HWSD classification scheme (Harmonized World Soil Database, FAO 2009).

2.4 Vegetation

The vegetation cover in West Africa generally follows the north-south gradient of the regional climate. From north to south, the density of vegetation commonly increases with sparse desert vegetation in the north to tropical rainforests in the coastal areas of West Africa (Figure 2.4). Besides the climatic factors, especially the anthropogenic influence has shaped

West Africa's vegetation through the high population density and considerable expansion of land use.

In the Sahara Desert, the northernmost ecoregion of West Africa, vegetation cover is sparse to absent, except for in some depressions with available water at or close to the surface (CILSS, 2016). Adjacent to the south, the Sahelian Acacia Savanna stretches from West to East Africa in a semiarid belt of 350 km average width including the northern parts of Burkina Faso (CILSS, 2016). It is characterized by 200 to 600 mm rainfall and a dry season of 8 to 9 months. Its vegetation is dominated by annual grasses of the genera *Aristida* and *Cenchrus* with a sparse (*Acacia*) tree and shrub layer (CILSS, 2016). The plant growth and also crop yields in this area are mainly determined by the precipitation resulting in a growing season which is primarily confined to the rainy season lasting for about three months (between June and September). During the dry season, the major part of the woody vegetation loses its leaves, the grasses become dry and fires are common (Knauer et al., 2014). The region has been shaped by a long history of land use, primarily pastoral nomadism with cattle as the primary livestock. During the last 30 years, the Sahel has received considerable attention with a controversial discussion about degradation or greening of the region which has been described and summarized in section 3.1.4 (for definition of degradation see page XX).

In the south, the Sahelian Acacia Savanna blends into the West Sudanian Savanna which covers the biggest part of Burkina Faso. This ecoregion is approximately defined by annual rainfalls between 600 mm and 1200 mm and a dry season between 5 and 7 months (CILSS, 2016). The woody component in this region is considerably higher than in the Sahel and further increase towards the south: its northern semi-natural areas are characterized by tree savanna changing into wooded savannas or open woodlands in its south. While the Sahel is characterized by short annual grasses, tall perennial grasses (mainly *Andropogon*) dominate in the Sudanian Savanna (CILSS, 2016). However, the West Sudanian Savanna is a region of intensive anthropogenic use and semi-natural areas become sparser. In general it is an area of high agricultural potential and rapidly growing population which resulted in a considerable expansion of agricultural area during recent decades (Callo-Concha et al., 2013; Knauer et al., 2017; Vlek et al., 2010).

At an annual rainfall of about 1200 mm, the West Sudanian Savanna turns into the Guinean Savanna or, as defined by Olson et al. (2001), the Guinean Forest – Savanna Mosaic (Figure 2.4). The semi-natural vegetation of this region is characterized by dense tree cover and ranges from deciduous woodland over closed deciduous forest to evergreen forest. The dry season prevails between 2.5 and 5 months and annual rainfalls can reach almost 2000 mm (Machwitz, 2010). In the coastal areas the forest – savanna mosaic of the Guinean region turns into moist forests which contain the last West African rainforests. However, despite the rather large expanse marked as Guinean Moist Forests by Olson et al. (2001), the rainforests only cover a small fraction of this area (about 15 %) and they are scattered over the ecoregion.

The whole Guinean region has been heavily altered and degraded by human practices, especially the slash-and-burn agriculture and the expansion of urban areas.
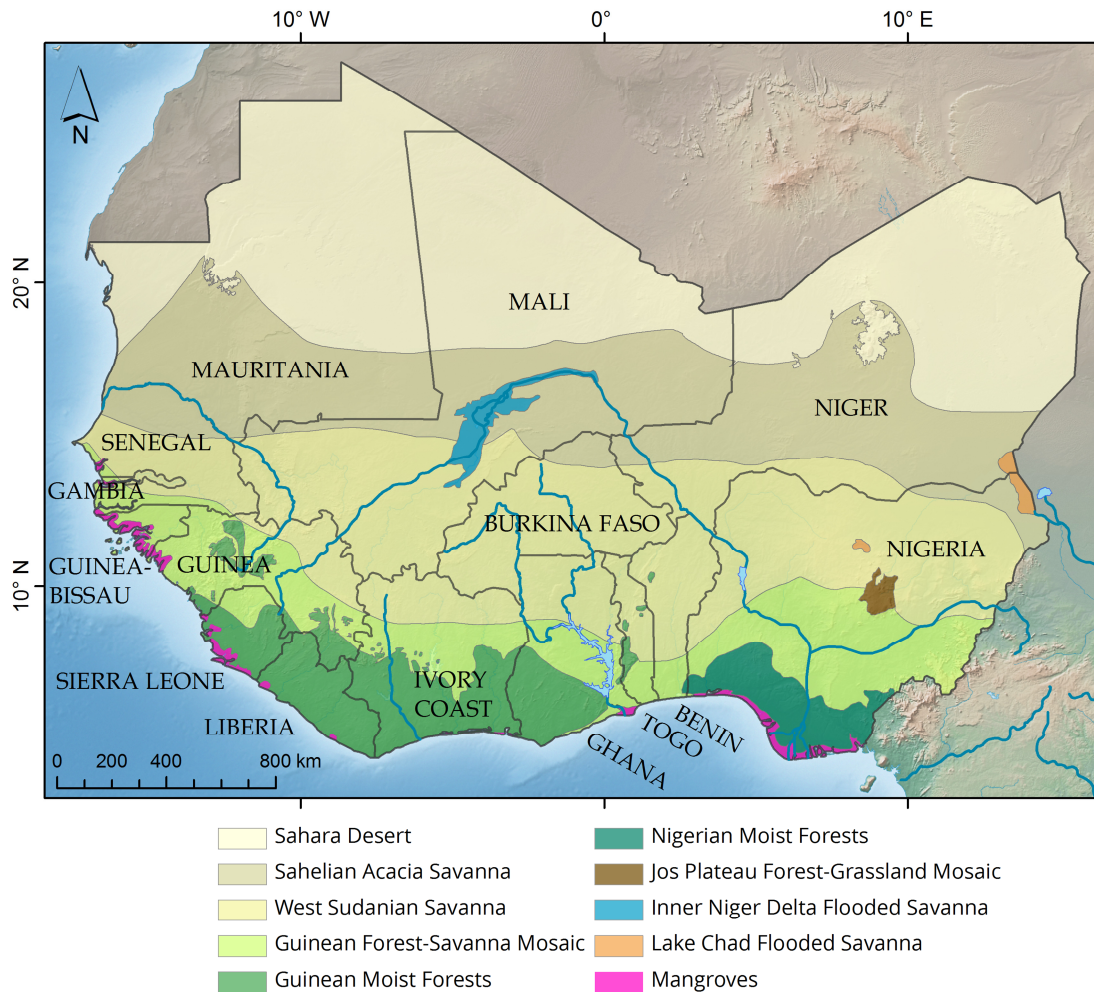
Figure 2.4: Major terrestrial ecoregions of West Africa according to Olson et al. (2001).

2.5 Population

West Africa is the fastest growing region in the world. With a growth from about 71 million people in 1950 to 353 million in 2015, its population experienced a fivefold increase while the world population only tripled (CILSS, 2016; United Nations, 2015). However, the population of West Africa is unevenly distributed which is a result of the history of human settlements and physical environment of the region (CILSS, 2016). In the northernmost areas of West Africa, the arid climate results in a very sparsely distributed population (Figure 2.5). In regions with more favourable conditions for agriculture, like the Senegalese Peanut Basin, the Niger-Nigeria border region or the central parts of Burkina Faso, population densities are considerably higher. Landlocked areas with constant access to water for agriculture and livelihood as along the Niger River in Mali also tend to have higher population numbers. In the coastal regions of West Africa, urban settlements and agglomerations considerably increase

population densities. Furthermore, in several countries such as Burkina Faso or Senegal, the respective capitals form extensive agglomerations resulting in high population numbers in their surroundings. The absolute national population numbers reveal the dominant role of Nigeria in West Africa: while Nigeria has an estimated population of 182 million people in 2015 (United Nations, 2015), the rest of the West African countries altogether account for 170 million people.

Though, the majority of West Africans still live in rural areas, the urban portion has increased from 8.3% in 1950 to 44% in 2015 (CILSS, 2016). While the rural population density is generally higher in the open savanna regions, e.g. of the Sudanian savanna, the originally extensively forested southern parts of West Africa experienced lower rural population numbers.

The predictions for West Africa continue the rapid population growth of the past: on the basis of a medium fertility model, the region will grow to an estimated 516 million people in 2030 and to almost 800 million people by the year 2050 (United Nations, 2015). This growth is sustained by the very young age structure of West Africa with almost half of the population being 15 years or younger (CILSS, 2016). An expected improvement of living conditions and thus an increase in life expectancy will boost this development in the near and medium future.
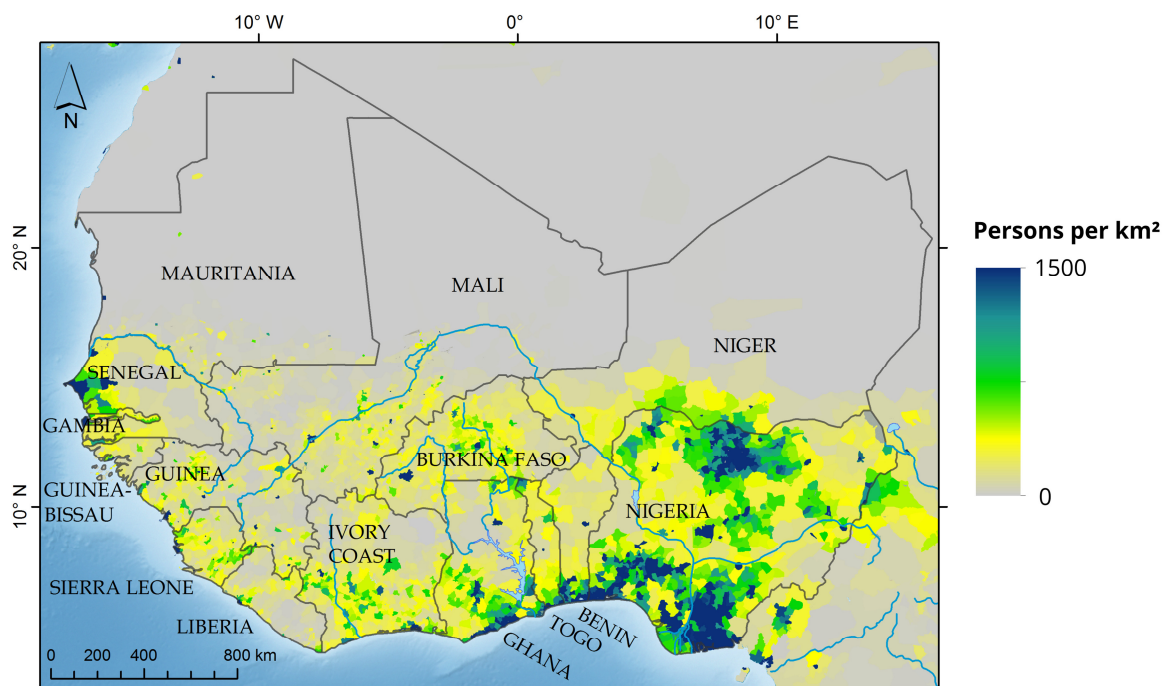
Figure 2.5: Population density per km² in West Africa for 2015 (CIESIN, 2016).

The above described population development and distribution in West Africa generally also apply for Burkina Faso. With an annual population growth rate of more than 3%, Burkina Faso ranks among the fastest growing countries in West Africa and the world in general. Only

since the year 2000, Burkina Faso's population grew by more than 50% from 11.6 million to 18.1 million people in 2015, and this growth is expected to continue (United Nations, 2015). In this agriculturally dominated country, the rural population still prevails but its regional distribution varies considerably with its physical environment (Figure 2.6). The major part of Burkinabe is concentrated in the central areas around the capital of Ouagadougou and around Bobo-Dioulasso. The northern Sahelian provinces as well as the more densely vegetated provinces in the southwest and southeast are rather sparsely populated. However, the development of rural population since 2001 shows that almost all provinces increased their numbers, even the more remote ones.
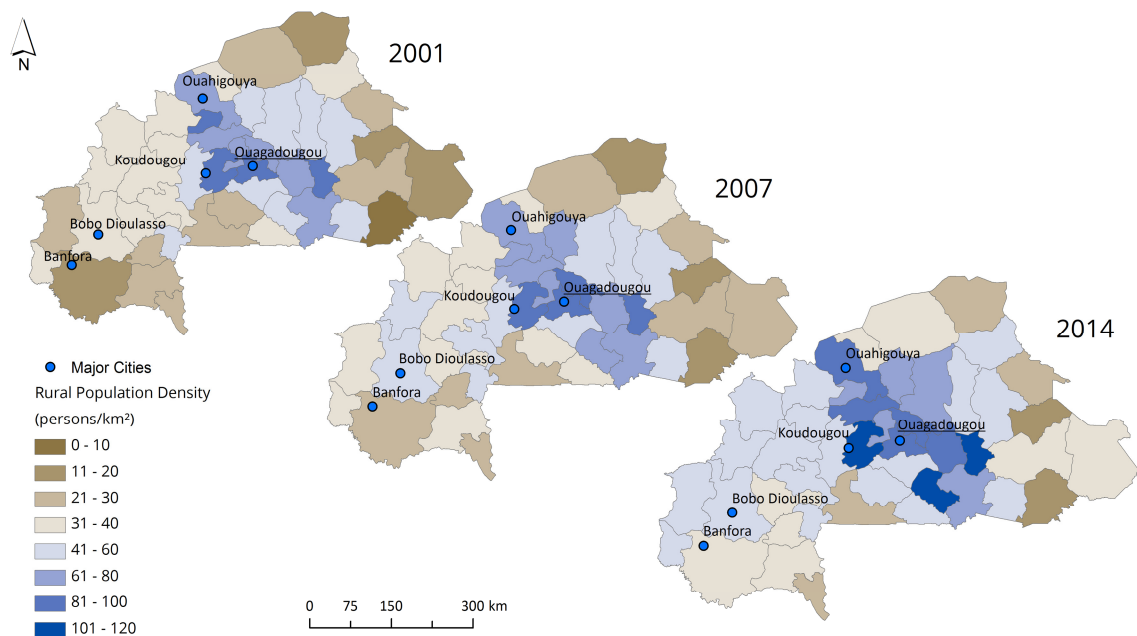
Figure 2.6: Rural population density for the provinces of Burkina Faso and years 2001, 2007 and 2014 (Source: Institut national de la statistique et de la démographie (INSD) 2014).

2.6 Economy and Land Use

Altogether, 15 West African countries are united in the Economic Community of West African States (ECOWAS) which was founded in 1975 (ECOWAS, 2017). All continental West African states are members of this union except for Mauretania which left in 2001 due to a planned stronger convergence to North Africa. The initial aim of this community was the self-subsistence and economic cooperation of the member states through the opening of a common domestic market and the foundation of a monetary and economic union. However, diverging interests between the francophone and anglophone countries so far hindered the introduction of a joint currency. In general, West Africa's economy is growing rapidly with GDP

(Gross Domestic Product) growth rates of up to 8.8% in 2015 (African Statistical Coordination Committee, 2016). Nonetheless, this growth so far is not reaching large parts of the population which is why several West African countries like Niger, Burkina Faso or Guinea still range among the least developed countries in the world (UNDP, 2015). West Africa's economy mainly relies on the export of agricultural products such as cotton, palm oil or cocoa and mineral resources like oil, gas or gold. Especially Nigeria, the biggest economy in West Africa, is exploiting considerable amounts of oil (853 million barrels in 2012) extracted offshore and onshore in the Niger Delta (Kuenzer et al., 2014).

The general economic situation of West Africa is also well reflected in Burkina Faso. Despite a sustained growth during the last decade with a GDP of 5.3% in 2015 (African Statistical Coordination Committee, 2016), Burkina Faso is one of the least developed countries in the world, positioned at rank 183 of 188 countries with a Human Development Index (HDR) of 0.402 (UNDP, 2015). Burkina Faso's economy depends heavily on agricultural production, livestock farming and the exploitation of mineral resources (FAO, 2014). In 2014, the most important export products were gold (2 billion US$), cotton (699 million US$) and refined petroleum (243 million US$) (OEC, 2017). Agriculture contributes to about one third to the national GDP with more than 90% of the countries workforce employed in the primary sector (FAO, 2014). Farming in Burkina Faso is generally characterized by small-scale rainfed subsistence cultivation with farm size below 5 ha. Due to a low level of mechanization and also fertilizer and pesticide input, the productivity of Burkina Faso's agriculture is still very low without a sufficient development during the last 50 years (FAO, 2016). Between 1961 and 2013, the cereal productivity of Burkina Faso increased from 0.41 t/ha/year to 1.16 t/ha/year (Figure 2.7). Although, the productivity almost tripled during this period, other developing or emerging countries like China or Brazil could increase their crop yields to almost 6 t/ha/year in the same time.

The most important staple crops in Burkina Faso are maize, millet and sorghum while the above-mentioned cotton is the most important crop in terms of export value. In irrigated areas around water reservoirs or along rivers, rice production is common. In the past, shifting cultivation with a long period of fallow and a relatively short period of cultivation was the major cropping system in West Africa (Forkuor, 2014). However, the rapid population growth forced farmers to switch to an annual cropping system without or with considerably shorter fallow periods. The primary type of cultivation is now intercropping where multiple different crops are cultivated on the same field (Forkuor, 2014). The implementation of this system is also a reaction to the increasing uncertainties in the patterns and length of the rainy season. On the basis of multiple crops with different water requirements on their fields, farmers are more resistant to crop failure (Forkuor, 2014).

The cultivation of tree crops such as mangos or cashew nuts is also common in Burkina Faso, so far mostly as part of the intercropping system which means that single trees are

spread over fields of other crop types like cereals. During recent years, also monocultures of tree plantations increase in number and size, although their total area is still low in comparison to other crops (section 6.4.2).
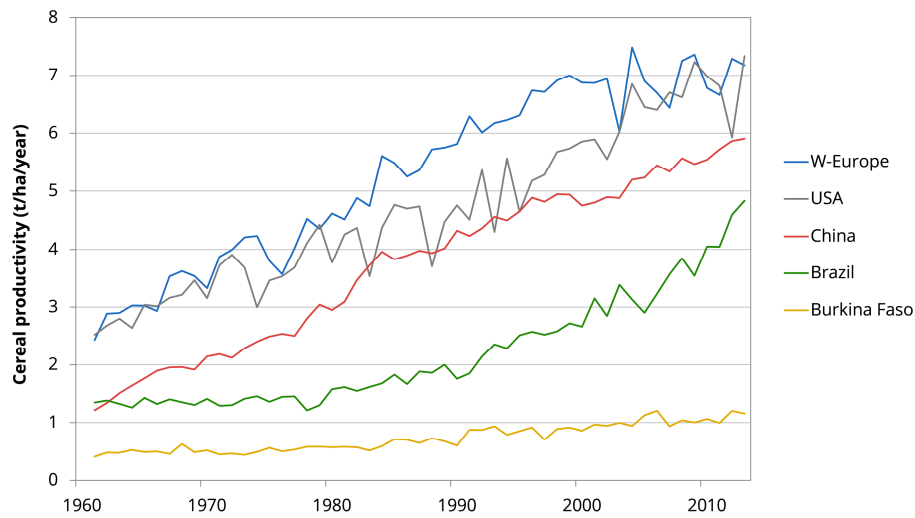
Figure 2.7: Cereal productivity in tons per hectare and year between 1961 and 2013 for Burkina Faso and other countries/regions for comparison (FAO, 2016).

3 Theoretical Background

The following chapter presents the theoretical background of this work giving an overview of existing datasets, methodologies and approaches that are suitable for the application in this study. Furthermore, this chapter describes why certain approaches were chosen over others and how, based on those, the methodology of this work was developed.

In a first part, selected parts of a conducted review study (Knauer et al., 2014) are presented identifying research gaps in the field of remotely sensed vegetation dynamics in West Africa (for definition of vegetation dynamics see page XX). Hereafter, a short overview of the most common satellite sensors with their technical specifications is given in order to show which remote sensing datasets are available for the mapping of agricultural area. Then, existing LULC classifications on different spatial scales are presented and their respective suitability for the monitoring of agricultural area in West Africa is outlined. Here, also the major challenges that impede such a monitoring in the region are discussed such as the lack of remote sensing data in a suitable temporal and at the same time spatial resolution. In a further section, the common spatio-temporal data fusion techniques that could be used to improve data availability, are summarized including their specific advantages and disadvantages that would affect an application in the region of West Africa. The last section presents the different classification methods varying from supervised to unsupervised and parametric to non-parametric types and shortly describes their general functioning and usability.

3.1 Remote Sensing of Vegetation Dynamics in West Africa

A comprehensive review study was conducted on prior research and on pressing, open research questions on vegetation dynamics in West Africa. This investigation and its results are summarized in this section and have been published in detail in Knauer et al. (2014).

3.1.1 Spatial Distribution of Reviewed Studies

Figure 3.1 gives an overview of the number of reviewed studies per country or region. Studies covering more than three countries were counted as an analysis of West Africa in this overview. All in all, 15 studies focused on vegetation dynamics in West Africa, 30 specifically investigated the Sahel region and 12 studies were reviewed that cover the entire continent or at least Sub-Saharan Africa. The coastal regions of West Africa only show low numbers of

studies with zero articles focusing on Gambia, Sierra Leone, Liberia, Togo or Guinea-Bissau. One reason for these trends is the dense cloud coverage prevailing in the coastal regions of West Africa which results in low availability of remote sensing data for the major part of the year. Another reason could also be the political instability in some of these countries which hampers at-ground research and field validation. In addition to the studies covering the entire Sahel region as mentioned above, the Sahelian countries Mali, Niger, Burkina Faso and Senegal exhibit the highest numbers of studies in West Africa. This can be explained by the high interest in the long-term vegetation development of the Sahel and the controversial discussion about degradation or re-greening in this region.
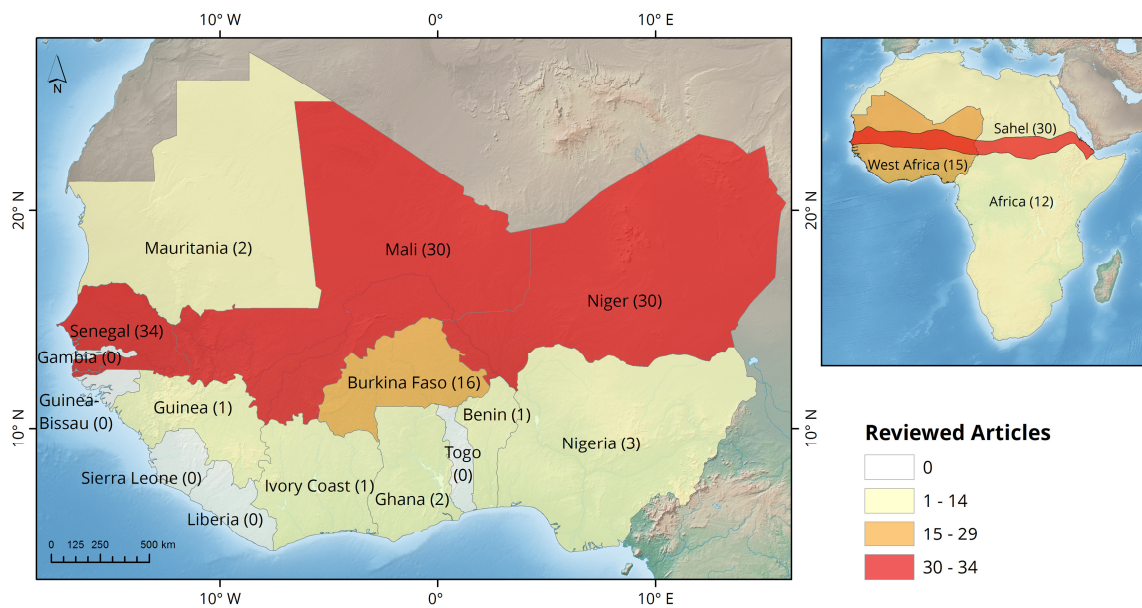
Figure 3.1: Spatial distribution of articles reviewed for this chapter. The number of studies per country or region included in brackets (modified after Knauer et al., 2014).

3.1.2 Assessment of Land Surface Phenology

Land surface phenology (LSP) is a consistent indicator of vegetation dynamics in a rapidly changing environment (Gonsamo et al., 2012). LSP is highly sensitive to climate change which is presumed to increase the annual variability of the growing season in West Africa (OECD, 2009). Thus, it is important to investigate and monitor the annual LSP in order to derive implications for humans and environment (Reed et al., 2009). This importance is reflected in the literature with rising numbers of studies during the last decade. LSP is defined after de Beurs and Henebry (2004, 2010) as the spatio-temporal development of vegetated land surfaces as detected from remote sensors. A focus in this field of research is on the timing and length of recurring season patterns (Gonsamo et al., 2012). While the phenology of vegetation is dealing with the seasonality of plants, LSP represents mixtures of land covers or plant communities because most remote sensors provide spatially aggregated information over areas of several square meters (Brown and de Beurs, 2008).

In one of the first analyses in this field of research, Reed et al. (1994) derived 12 LSP metrics like onset of greenness or date of maximum NDVI from AVHRR data for the USA. These metrics do not necessarily correspond to common phenological events like germination, tillering or flowering but may correlate with them.

One of the most commonly assessed LSP metrics, the start of season (SOS) or onset of greenness, has been estimated with a variety of different methods which consequently results in different dates of SOS for the same region. Thus, it is often unclear what SOS derived from remote sensing actually measures. It could be related to the first flush of greenness, the primary expansion of leaves or an early season growing peak (Reed et al., 2003). Nonetheless, the more or less arbitrarily set definitions of LSP metrics such as SOS can become meaningful and important when they are derived from a multi-annual time series for the analyses of changes in ecosystem characteristics (Reed et al., 1994). The most simple and at the same time most common way of defining SOS and the end of season (EOS) is based on threshold values applied to time series of vegetation indices. These thresholds are often defined retrieved relative to a baseline such as a mean value of NDVI over a part of the pre-growing season or between minimum and maximum of the NDVI curve.

Brown et al. (2010) investigated SOS derived from GIMMS (Global Inventory Modelling and Mapping Studies) NDVI datasets of entire Africa between 1981 and 2008 and the response of this metric to large-scale climate oscillations. In this study, they used the 50% point of green-up, which is also called mid-point NDVI. The dates of SOS and EOS were determined for every pixel as the point of time when the NDVI crosses the 50% level between minimum and maximum of NDVI in upward and downward direction respectively. Vrieling et al. (2013) applied the same method analysing the growing period over Africa based on GIMMS-3g NDVI (1981-2011). They justified these thresholds as the average point of time with the most rapid increase and decrease in greenness represented by NDVI. This extraction of phenological metrics with relative thresholds is included in the software TIMESAT which is frequently used in the literature (Eklundh and Jönsson, 2015). TIMESAT was developed for the analysis of time series of vegetation indices and includes time series filtering and the extraction of phenological metrics such as SOS or EOS. Heumann et al. (2007) used the software to analyse greening trends in the Sahel, Sudanian and Guinean region of Africa. For the extraction of NDVI, they used the 20% point of increase and decrease of the NDVI amplitude on the respective upward and downward limb of the curve. In addition, they analysed the length of the growing season, amplitude, and integrated NDVI (iNDVI). Time series with a too low seasonality for the extraction of phenological metrics defined by an average amplitude of less than 0.1 NDVI were excluded from the analysis. The results of this study are presented in section 3.1.4.

Ringelmann et al. (2004) and Kumar et al. (2002) used two different threshold methods for the extraction of phenological metrics which are not based on NDVI but focused on

backscatter and rainfall, respectively. Ringelmann et al. (2004) applied data from the SeaWinds scatterometer (2000-2003) and a simple backscatter threshold derived from the Ku-band to define SOS in Mali. They argued that this threshold can be used as a proxy for planting dates since they saw changes in backscatter mainly affected by changes in soil moisture (and growth of vegetation) which is the starting point of seeding for West African farmers. Kumar et al. (2002) assessed the beginning of the rainy season from AVHRR NDVI data to analyse subsequent growth rates of vegetation related to different soil types in Burkina Faso. They applied a threshold of 20 mm of cumulative rainfall which was reported to trigger plant growth of perennial vegetation. Their results indicate significantly higher growth rates of herbage vegetation on sandy soils than on clay soils, especially at the start of the rainy season.

The most complex method for the determination of phenological metrics is model fitting (Kirsten M. de Beurs and Henebry 2010). Because of the variable nature of VI curves, flexible models are developed that can be adapted to the underlying data (Kirsten M. de Beurs and Henebry 2010). Brown and De Beurs (2008) applied a quadratic regression model accounting for the precipitation gradient in the Sahel to identify the SOS. Relative humidity data accumulated over the year, was used as the independent variable (instead of simply the time of the year) in order to fit the model to AVHRR, SPOT and MODIS NDVI data. They developed an extensive searching algorithm testing variable seasonal window lengths and starting periods for each pixel to find the optimal model fit. Since the quadratic model fitted best for only the growing season taken into account, the SOS was determined as the first NDVI composite of the best fitting model. SOS estimations were validated with sowing dates derived through questioning farmers and the RMSE was found to be lowest for MODIS (RMSE = 12.2 days). It has to be considered that sowing can take one to several days depending on the method used, thus, at least 5 days of error are estimated to be inherent in this measurement by the authors (Brown and de Beurs 2008).

Butt et al. (2011) assessed the green-up and senescence of vegetation from MODIS NDVI data for a focus area in the bordering region between southern Mali and northern Burkina Faso. They fitted a double logistic regression to the NDVI time series and used the maximum of the second derivative as the date of green-up and 80% between minimum and maximum NDVI as the onset of senescence. The period between green-up and senescence was taken as length of growing season. The analysis of these metrics over eleven years (2000-2010) showed a weak phenological trend in green-up and senescence meaning a delay of 0.4 and 0.8 days/year, respectively. The regional analysis along a north-south gradient (11.5° - 17.5° N) showed a strong relation to the movement of the Intertropical Convergence Zone (ITCZ): at lower latitudes (11.5° - 13° N), a high green-up variability was found which is more likely at the pre-onset stage of the West African monsoon characterized by isolated rains. North of 13°, the green-up variability was detected to be lower which is rather influenced by rainfalls occurring

at the onset stage of monsoon when the ITCZ rapidly shifts northward bringing precipitation to a wide range of latitudes (Butt et al., 2011).

For entire Africa, Zhang et al. (2005) separately fitted sigmoidal models to the increasing and decreasing parts of the annual curves of MODIS Enhanced Vegetation Index (EVI). They subsequently derived phenological metrics such as 'onset of green-up' and 'dormancy onset' from these models, defined as extreme values in the change rate of the curvature. The response of these phenological metrics to rainfall was analysed on the basis of TRMM rainfall data. Zhang et al. (2005) determined distinct thresholds of cumulative rainfall for the green-up of vegetation in arid and semi-arid regions of Africa. Based on their methodology, Bobée et al. (2012) fitted a model to MODIS Leaf Area Index (LAI) time series for the derivation of SOS, EOS and length of growing season and for the analysis of the dependency of vegetation on rainfall.

Guan et al. (2014) fitted a double logistic curve to daily LAI data from the SEVIRI sensor for the African continent and subsequently derived several phenological metrics and so-called trajectory information like rates of emergence and senescence. They claimed that their algorithm is giving more flexibility in the definition of phenological parameters and the matching of them with in-situ observations. However, they did not include such a validation with at-ground data in their study.

Vintrou et al. (2014) assessed the quality of the MODIS Land Cover Dynamics product (MCD12Q2) including information about green-up, senescence, maturity and dormancy onset, by the comparison with the validated crop growth model SARRA-H (System for Regional Analysis of Agro-Climatic Risks). The MODIS product estimated SOS and EOS approximately 30 days before and 40 days after the index derived with SARRA-H. Since they found this bias to be regionally consistent, they drew a positive conclusion about the product and recommended it in combination with crop models for the application in food security systems (Vintrou et al., 2014).

An overall comparison of study results and estimated LSP metrics is difficult since the different methods for the generation of the same metric may relate to different phenological events. In addition, most of the times, it is not possible to validate the accuracy of the derived LSP metrics since accurate ground truth datasets are missing.

3.1.3 Assessment of Crop Yields

Except for one study focusing on the biomass of oil palm plantations in Benin (Thenkabail et al., 2004), all studies related to yield assessment were conducted on the Sahel and Sudanian region. The most common crops in the reviewed literature on yield assessment are millet and sorghum, which represent the major cultivated plants in the Sahel and Sudanian region. Moreover, the monitoring of rangelands in West Africa is a frequent topic in the literature. The majority of reviewed studies applied regression analyses between NDVI and final crop yield for the forecasting of yields. In addition to these empirical approaches, only

one study was found that used a more complex process-based crop simulation model for the assessment of yields in Burkina Faso (Thornton et al., 1997). In this section, the growth phases are always given with the actual time span of these phases as used in the respective study. This is necessary to avoid misinterpretations, since the growing period can be subdivided in either two (vegetative and reproductive period) or three growth phases (vegetative, reproductive and grain-filling period).

In several studies, single-date NDVI, mostly maximum value composites of medium spatial resolution time series, was related with yield measurements or biomass for the prediction of yields (e.g. Groten 1993; Fuller 1998). Maselli et al. (1993) applied standardized AVHRR NDVI values of the end of July, averaged over sub-departments of Niger for the forecasting of yields of millet and sorghum. This date was argued based on earlier studies to lie in a meaningful ecological period since it was identified as being highly responsive to absolute amounts of rainfall from the previous months and directly precedes the peak of vegetation activity (Maselli et al., 1992). This hypothesis was tested in later studies using the same approach (Bozzini and Maselli, 2002; Maselli et al., 2000) and after trying several acquisition dates, the AVHRR NDVI between end of August and middle of September (reproductive period) was found to be best suited for the forecasting of millet and sorghum yields.

Another frequently applied method is the integration of a vegetation index over a part of or the whole growing period and the subsequent correlation or regression with biomass or yield. Bartholome (1988) conducted one of the first relevant studies in Mali, deriving an integrated NDVI (iNDVI) from radiometric field measurements of millet and sorghum. They found the relationship between iNDVI and total biomass to be strongest for an integration of NDVI over the whole study period between mid-August and mid-October. Furthermore, the integration focusing on the reproductive period (here: maximum NDVI at beginning of September till beginning of October) delivered the best correlation with yields ($r \approx 0.9$). They found this integration method to yield more stable correlation coefficients than the use of single NDVI values ($r \approx 0.7$). However, due to an instrument failure they had to use the green band instead of the red band to compute the NDVI in their analysis.

An empirical millet yield forecasting model was developed by Rasmussen (1997) on the basis of AVHRR data applying it to two separate years (1990-1991) in a peanut-growing area in Senegal. Several integration periods of iNDVI were iteratively tested and the integrals over the reproductive period (45 days between end of August and mid-October) were found to provide the highest correlation coefficients with millet yield. In this study, a pre-crop iNDVI was determined iteratively as a reference level and subtracted from the iNDVI of the reproductive period. This was done because of a detected presence of green vegetation before seeding in the NDVI indicating a component of vegetation on the fields not related to crops. The final model of this study resulted in a correlation coefficient of 0.72 allowing the assessment of

yield data one month before harvest. In two additional studies by the same author (Rasmussen 1998a; Rasmussen 1998b) fixed periods were used, one over the grain-filling period (1st September to 10th October) and one over the month of May as the pre-season reference level. Furthermore, Rasmussen (1998a, 1998b) tested various climatic and environmental variables in single linear regressions as predictors for millet yield. He found the percentage of cultivated land and the Tropical Livestock Unit (TLU) density, measured in number of cattle per km², to significantly improve the accuracy of prediction with a maximum R^2 of 0.88. Based on Rasmussen's millet model, Knudby (2004) developed his groundnut yield model for Senegal also using an iNDVI over the reproductive period subtracted by an iNDVI from the pre-growing season. These periods were set relative to the estimated sowing date which was assessed similar to Groten (1993) as the first in a series of consistently increasing NDVI values resulting in a maximum R^2 of 0.64. However, this method also resulted in large local variations of sowing dates leading to unrealistically high local variations of estimated yields (Knudby, 2004). Töttrup and Rasmussen (2004) used a simplified approach based on an iNDVI of September in order to map long-term changes in crop productivity (predominantly millet and groundnut) of Senegal between 1981 and 2000. They detected a high spatial variability of trends with a generally negative tendency. This tendency could not exclusively be explained by rainfall and was attributed to changes in land use and land cover.

The process-based CERES-Millet (Crop Environment Resource Synthesis) crop simulation model was applied by Thornton et al. (1997) for the provinces of Burkina Faso. This model is using rainfall estimates from METEOSAT and a cropland use intensity map derived from Landsat imagery as well as additional information. The results generally indicate a good agreement between simulated and observed millet yields for two years (1986 and 1990) and in two selected provinces (for a comparative table see Thornton et al. (1997)).

Rojas et al. (2011) estimated drought probability for croplands in Africa on the basis of the NOAA Vegetation Health Index (VHI) product and GIMMS NDVI data for the period 1981 to 2006. The VHI is derived from NDVI and brightness temperatures and was averaged over the growing season in this study. This VHI integration period was derived from the GIMMS time series as the period between SOS (50% green-up) and the end of the grain-filling stage, which was defined as EOS (50% senescence) minus six weeks. These six weeks were assessed as ripening stage of crops and excluded from the integration since this period is less sensitive to water stress and thus not relevant for the drought assessment (Rojas et al., 2011). The aggregated results at sub-national level indicated several hot spots with a high drought probability in West Africa: a belt from the north of Senegal to the west of Burkina Faso and the region between southern Sierra Leone and northern Liberia.

Meroni et al. (2014) conducted a more recent approach on early warning of biomass production deficit of rangelands and croplands in the Sahel. They developed an algorithm based on cumulated fAPAR over the growing season derived from the Joint Research Centre

(JRC) FS-fAPAR product (SPOT VEGETATION, 1998-2012) and a historically defined deficit threshold. A deficit is detected at a certain point of time during the growing season by estimating the probability that the cumulated fAPAR will not be able to exceed the pre-defined threshold at the end of the growing season. This is method is conducted by comparing the current with past annual vegetation development from the 15 years of the time series. However, this rather low sample size, which is also used for the definition of SOS and EOS, seems statistically questionable.

In general, it is unknown if these models, which were designed for simplicity in order to guarantee local application, are really used today in the respective countries. However, it seemed as if up-to-date research on the topic of remote sensing based agricultural monitoring is rather sparse for West Africa, though this was not the major focus of the conducted review.

3.1.4 Vegetation Trends in West Africa

Especially in the Sahel and Sudanian region of West Africa, vegetation dynamics were controversially debated in the literature of the past 30 years. In their research note, Mbow et al. (2014) gave a brief overview of the progress in the monitoring of vegetation dynamics in West Africa. Land degradation and desertification on different scales and even an adverse greening trend have been detected and are still the subject of controversial discussions (Hein et al., 2011; Hein and De Ridder, 2006; Prince et al., 2007). For the proper analysis and interpretation of remote sensing data in this field of research, the meaning of these terms has to be fully understood. The United Nations Convention to Combat Desertification (UNCCD, 2017) has defined land degradation and desertification as follows:

"Land degradation" means reduction or loss, in arid, semi-arid and dry sub-humid areas, of the biological or economic productivity and complexity of rainfed cropland, irrigated cropland, or range, pasture, forest and woodlands resulting from land uses or from a process or combination of processes, including processes arising from human activities and habitation patterns [...]
"Desertification" means land degradation in arid, semi-arid and dry sub-humid areas resulting from various factors, including climatic variations and human activities.

Although, the definition of land degradation generally comprises several indicators of this process, most remote sensing based studies in West Africa only investigate one aspect such as the biological productivity. This may lead to flawed conclusions or further processes of land degradation might be missed in the analysis. The causes of land degradation are frequently discussed. They can generally be subdivided into directly human induced causes, for example due to agricultural overuse of land or they can be associated to climatic variations or climate change.

Greening or Degradation?

Starting with the severe droughts of the 1970s and 1980s, a southward expansion of the Sahara was assumed and the Sahel was said to be highly threatened by degradation (Tucker et al., 1991). On the basis of remote sensing, Tucker et al. (1991) were one of the first to reveal that vegetation was still able to recover from severe droughts like after 1984, one of the driest years of the century. Their results also highlighted the high inter-annual variability of vegetation in the Sahel (represented by NDVI) and its strong dependency on precipitation. The relatively simple methodology of this analysis was enhanced by Nicholson et al. (1998a) using a longer time series (1980-1995) and introducing several modifications (e.g. application of satellite-derived rainfall estimates and restricting the Sahel to its central and western part). The results of their study also confirm that there has not been a progressive shift of the desert boundary or a change of vegetation cover in the Sahel. In a later analysis, Eklundh and Olsson (2003) found a strong increase of seasonal iNDVI of the Sahel and Sudanian region derived from AVHRR between 1982 and 1999. This trend was interpreted as a recovery of vegetation from the droughts of the 1980s, although an increase of rainfall was also detected during this time. Based on an analysis of 23 years of GIMMS NDVI data (1981-2003), Anyamba and Tucker (2005) came to the same conclusions as Eklundh and Olsson (2003). They argued that the increase of rainfall was part of the long-term climatic history of the Sahel outlining that rainfall amounts were still considerably below the wetter conditions of 1930 to 1965. Seaquist et al. (2006) applied a simple light use efficiency model to convert AVHRR PAL NDVI (1982-1999) into growing season net primary productivity (NPP) and detected a similar greening trend for the Sahel and Sudanian region.

Heumann et al. (2007) conducted a phenological approach by applying a trend analysis of phenological metrics for the Sahel, Sudanian and Guinean region between 1982 and 2005. In general, they could confirm the previously detected greening trend and were able to separate it into two different types of greening: in the Sahel, the greening was associated to an increase of the seasonal NDVI amplitude, which they explained by the prevalence of annual grasses being more dependent on rainfall and filling patches where bare soil occurs during dry years. In the Sudanian region, the greening trend was related to an increase of the length of growing season. A better adaptation to drought conditions (especially of woody vegetation) was seen as the major reason, because tree cover is considerably denser and perennial grasses are more common in comparison to the Sahel. The production of tree leaves prior to the rainy season is expected to store the precipitation of the first rainfalls better, resulting in a prolonged growing season in the Sudanian region (Heumann et al., 2007).

All above-mentioned studies assume a greening of the Sahel and Sudanian region over the last 30 years, but the causes and interpretations of this finding derived from remote sensing are still controversial. From this greening, some studies concluded that degradation has not occurred in the Sahel region or at least not on a bigger scale (Nicholson et al., 1998b;

27

Tucker et al., 1991). As outlined above, rainfalls also increased during the last 30 years in the Sahel and Sudanian region which could possibly disguise human induced degradation (Fensholt et al., 2013). For this reason, it is essential to consider the role of climate and possible climatic changes for the interpretation of such trends in order to be able to distinguish human induced from climate induced vegetation trends.

Climate Influence

The droughts of the 1980s and the subsequent greening trend and increase of precipitation in the Sahel and Sudanian region lead to the question which climate variables are the main drivers affecting vegetation variability in West Africa (Brown et al., 2010; Jarlan et al., 2005; Kaspersen et al., 2011).

In section 3.1.2, Butt et al. (2011) described the influence of the ITCZ on the annual vegetation phenology. Philippon et al. (2007) applied GIMMS NDVI data covering 20 years (1982-2002), precipitation, outgoing long-wave radiation and sea surface temperatures (SST) as explanatory variables for the analysis of intra- and inter-seasonal vegetation variability in West Africa. Instead of the more common rotated principal component analysis, they used the independent component analysis for the extraction of two main modes of inter-annual variability. The first mode was called the 'Sahelian mode' and describes years with high photosynthetic activity, above-average rainfall and deep convection during summertime (August to October) over the Sahel and Sudanian region while at the same time the eastern tropical Pacific exhibits below-average SST (Philippon et al., 2007). The second mode was called the 'dipole mode' and showed opposing variations between the Sahel and Guinean region during summertime (July – September). During years with a high photosynthetic activity, the Sahel region received above average rainfalls while the Guinean region received below average rainfalls. In addition, the tropical Atlantic experienced a North-South dipole of SST during this mode. The second mode ('dipole mode') seems to characterize the variations connected to the latitudinal position of the ITCZ (Philippon et al., 2007). They furthermore analysed several phenological metrics like green-up and senescence for the intra- and inter-annual variability. An interesting result of this investigation is a year-to-year memory effect of vegetation meaning that the green-up (between June and July) is connected to the maximum NDVI of the following year.

Huber and Fensholt (2011) subdivided the Sahel from east to west into five regions in order to account for regional variability in their study. They applied a correlation analysis to investigate the relationship between NOAA SST, GIMMS NDVI and four climate indices namely the Indian Ocean Dipole, the Multivariate ENSO (El Niño-Southern Oscillation) Index, the North Atlantic Oscillation Index and the Pacific Decadal Oscillation. They found the NDVI of the western part of the Sahel to be associated to SST of large parts of the Atlantic, the Pacific and the Indian Ocean. An increased NDVI in Senegal and western Mauretania and Mali could be

related to above average SST in the eastern Atlantic and below average SST in the Pacific. Nonetheless, as noted by the authors, only linear relationships were investigated and non-linear features or interferences between the climate indices could also play a role (Huber and Fensholt, 2011).

The reviewed studies of this section outline the main drivers of regional precipitation and thus the annual vegetation cycle in West Africa. Because rainfall is the major limiting factor for vegetation growth in the Sahel and Sudanian region, the investigation of human impact cannot be done without taking precipitation into account. Suitable methods for this task are debated beyond the border of West Africa for various regional and global applications (e.g. Bai et al., 2008; Dent et al., 2009; Wessels, 2009).

Human Influence

Correlation and Regression Analyses. The most popular method for the detection of human-induced degradation is the correlation or regression between precipitation and NDVI. Areas showing low correlations are expected to be affected by human impact, although other variables like soil type should be considered as well. For the Sahel region (1982-2003), Herrmann et al. (2005) applied a relatively simple approach using GIMMS NDVI data and GPCP (Global Precipitation Climatology Project) and TRMM (Tropical Rainfall Measuring Mission) rainfall estimates. A linear regression between monthly NDVI and 3-month cumulative precipitation (the month of the NDVI record plus the two preceding months) was calculated and trends of NDVI residuals, i.e. the difference between observed and predicted NDVI, were derived. The cumulative rainfall of three months was used because it resulted in higher correlations with NDVI than for one or two months of cumulative rainfall. Herrmann et al. (2005) reasoned that the vegetation greenness in semi-arid environments is rather related to soil moisture than to immediate rainfall. Although they identified rainfall as the major factor for the increase in vegetation greenness in the Sahel region, their analysis of NDVI residuals indicated an additional, weaker factor interpreted as human impact.

The greening trend in the Sahel was also investigated by Olsson et al. (2005) who used precipitation data from 40 rainfall stations and TIMESAT-generated AVHRR NDVI integrals over July to September. They came to the conclusion that the greening in the Sahel cannot solely be explained by increasing rainfalls and discussed human influence like an improved land management with a higher agricultural productivity or land use changes resulting from migration (especially in Sudan). However, they could not confirm these hypotheses. An increase of agricultural production as influence on the observed greening trend was further analysed by Capecchi et al. (2008). They investigated the correlation between AVHRR NDVI, Global Precipitation Climatology Centre (GPCC) rainfall data and annual national millet yield (1986-2000). However, this analysis should rather be conducted on a per-pixel or at least local

analysis and not on countrywide averages but yield data on such a high spatial resolution are sparse.

M. Boschetti et al. (2013) used SPOT VEGETATION data and FEWS (Famine Early Warning System Networks) precipitation estimates to confirm the general greening trend in the Sahel of West Africa. In addition, they focused on the detection of local hot spots showing anomalous greening or anomalous degradation in the entire region of West Africa. A hotspot was defined as an area where a standardized iNDVI (July to October) was not correlated to annual rainfall over 13 years (1998-2010).

Identified hotspots were focused on the basis of land cover change detection from pairs of Landsat TM/ETM+ (Thematic Mapper, Enhanced Thematic Mapper Plus) images of the start and end of the investigation period. As the main reasons for anomalous greening, M. Boschetti et al. (2013) identified the abandonment of agro-pastoral land as a consequence of population displacement, an intensification of cropland and an expansion of vegetation cover (predominantly agriculture) on parts of former water bodies. Degradation was generally determined for regions of demographic pressure and increasing overuse of environmental resources.

Because such commonly applied linear regression models are not able to account for non-linear changes of vegetation, Jamali et al. (2014) used a polynomial fitting method for GIMMS NDVI data over the northern part of Africa between 1982 and 2006. They also predominantly found positive linear trends over the Sahel region and suggested climate variables to be the main drivers although they did not want to exclude anthropogenic drivers.

Non-linear trends were found to be much less frequent and were widely scattered clustering within or around areas of linear trends which emphasizes their importance for the determination of a more comprehensive change area.

Process-based Vegetation Models. Another approach for the identification of human-induced vegetation changes is the application of process-based vegetation models. With this method, vegetation development is simulated without anthropogenic influence and subsequently compared to vegetation index time series from remote sensing. Deviations between these datasets indicate a human influence. Hickler et al. (2005) and Seaquist et al. (2009) used the process-based ecosystem model LPJ-DGVM (Lund–Potsdam–Jena Dynamic Global Vegetation Model) for the Sahel. Hickler et al. (2005) compared simulated vegetation with GIMMS NDVI data for the period 1982 to 1998. The model did not include information on actual land use so only natural ecosystem dynamics were simulated. Since they found the LPJ-DGVM to closely reproduce the GIMMS NDVI time series, they attributed the Sahel wide greening trend to the change in rainfalls. In addition, these results were regarded as proof against a major human influence for degradation in the Sahel although the authors did not want to rule out anthropogenic influence on smaller scales. Seaquist et al. (2009) expanded

the LPJ-DGVM study to the period 1982 to 2002 and attempted to relate local differences between modelled and observed vegetation to human impact. This potential influence was represented in the model through datasets of land use and population density. Nonetheless, they were not able to explain the residuals between observed and predicted vegetation by human influence but rather found a slightly positive relationship between data-model agreement and grazing intensity which indicates that livestock grazing is not significantly affecting vegetation dynamics in the Sahel at larger scales (Seaquist et al., 2009).

Rain-Use Efficiency. A further common method for the identification of human induced vegetation trends is the rain-use efficiency (RUE). RUE is defined as ratio between above-ground NPP, mostly substituted by an integral of NDVI, and annual precipitation (Le Houérou, 1984). The idea behind this ratio is the normalization of the effect of precipitation change in vegetation trends. Thus, positive or negative trends of RUE should indicate an improvement or degradation, respectively. Anyway, its correct application and interpretation resulted in intense discussions in the literature (Hein et al., 2011; Hein and De Ridder, 2006; Prince et al., 2007, 1998). Although, the strong dependency of the vegetation in the Sahelian and Sudanian savannah on precipitation is widely accepted, the proportionality between the two variables and also between RUE and precipitation were debated. The discussion was mainly started by the studies of Hein (2006) and Hein and De Ridder (2006) claiming to have found a considerable flaw in the application of RUE in many remote sensing studies. For Niger and Senegal, they found a quadratic relationship between annual precipitation and RUE with a peak between 200 and 300 mm of cumulative rainfall and rejected the so far assumed linearity. Although their methodology contained some errors leading to direct replies in the literature (Prince et al., 2007; Retzer, 2006) and they subsequently re-analysed their datasets (Hein et al., 2011), the general challenge appeared to be justified. For years with low annual rainfall, (Hein and De Ridder, 2006) detected a deviation from linearity of RUE and rainfall towards lower values of RUE which they explained by greater losses of water through evaporation. For high annual rainfall, they found RUE to decrease and assumed other factors than water availability like nutrients to limit vegetation productivity. Nonetheless, they extracted and applied field data from other publications, in total six sites with varying numbers of annual observations (between 7 and 17 years). Thus, their conclusion of a quadratic behaviour of the RUE-rainfall relationship was only based on single observations for low and high rainfall years which resulted in wide confidence intervals for the respective fitted curves (Hein and De Ridder, 2006; Prince et al., 2007). Although this lower deviation of linearity could not be confirmed in other studies, the upper boundary, i.e. a saturation effect of vegetation on rainfall is under discussion. Different thresholds of saturation were found for areas of the Sahel and Sudanian region, e.g. 700-800 mm per year by Li et al. (2004), 800-1000 mm per year by Martiny et al. (2006) and 800 mm per year by Prince et al. (2007). The results

of these studies all indicate a general linearity between precipitation and NPP but the thresholds deviate in each study which indicates a biome-specific behaviour. Fensholt et al. (2013) tried to avoid a general assumption of linearity or proportionality and restricted their study to pixels with a proven linearity. They introduced two pre-conditions in their trend analysis of RUE in order to get substantial results: a significant positive linear relationship between annual precipitation (Climate Prediction Center Merged Analysis of Precipitation, CMAP) and iNDVI (AVHRR GIMMS) as well as no correlation between RUE and annual rainfall. The remaining RUE dataset did not indicate any widespread degradation as found by other studies. However, because of the applied pre-conditions, they excluded 63% of the Sahel for which no conclusion could be drawn. Fensholt et al. (2013) further investigated the impact of integration period of NDVI on the analysis of vegetation trends resulting in some possible differences between different studies. Their results indicate a higher suitability of the growing season integral in comparison to annually integrated NDVI leading to a considerably higher number of pixels in the Sahel with a significant trend over the period of investigation (75% versus 33.7%). This was explained by the impact of soil and dead annual vegetation on the reflectance signal of the dry season which introduces noise to the annual sums of NDVI (Fensholt et al., 2013).

The above described studies and the resulting discussion highlight the difficulties and also challenge the reliability of RUE for the detection of degradation in the Sahel and Sudanian region. The relationship between RUE, rainfall and degradation appears to require more research and a better data basis of at ground observations (Hein et al., 2011).

Local Analyses. The results presented by M. Boschetti et al. (2013) as well as the conclusions drawn in the other presented studies above indicate the need for a local analysis of the relationship between vegetation and rainfall in order to support large-scale studies. Bégué et al. (2011) conducted such a local study for the Bani catchment in Mali analysing the spatio-temporal development of GIMMS NDVI data (1982-2006) in combination with precipitation from 65 stations. Their results show a significant positive trend of NDVI over the entire study area but no clear trend of precipitation. They subsequently divided the study area into three eco-climatical regions and could explain the greening in the northern part by an expansion of cropland. The mixed results in the other two regions were explained by the natural vegetation dynamics relating to non-linear rainfall patterns (Bégué et al., 2011).

Rasmussen et al. (2014) conducted a local study in northern Burkina Faso analysing MODIS NDVI (MOD13Q1) trends (2000-2012). For one part of the study region, the observed trends were explained by soil erosion allocating material from plateaus to valleys and causing negative and positive NDVI trends respectively. For another part of the study region, negative trends were mainly attributed to an increased human activity.

For the Senegal, Li et al. (2004) analysed the NDVI-rainfall relationship (AVHRR GIMMS NDVI - rainfall stations) and found three areas of low correlation which were subsequently identified as stressed by urbanization or intensive grazing. A further study in Senegal (Ferlo region) was conducted by Martínez et al. (2011) investigating the relationship between Apparent Green Cover (AGC, derived from SPOT VEGETATION NDVI) and rainfall estimates for 2001 to 2009. They applied a linear regression between annual AGC and the logarithm of annual rainfall estimates in order to consider the non-linear behaviour of vegetation at high precipitation values. They found a general greening but could not be fully explain it by rainfalls. Focusing on areas with negative slope values in the regression confirmed several zones of degradation which were previously identified by Li et al. (2004).

Brandt et al. (2014a, 2014b) focused on local greening trends in two study areas in Senegal and Mali which could be found on several common time series products. A comprehensive investigation of high resolution imagery, field data and interviews with local people outlined several explanations for the greening: re-planting and protection, increased agro-forestry, an expansion of more robust species and a rise in annual precipitation (Brandt et al., 2014a). Nonetheless, an analysis of images from the Corona satellite from 1965 and the interviews could show that the condition of the ecosystem is considerably poorer than before the severe droughts of the 1970s and 1980s resulting in less natural vegetation, tree density and tree diversity (Brandt et al., 2014a).

For two study sites in Mali (Gourma region) and western Niger (Fakara region), Dardel et al. (2014) regressed GIMMS-3g NDVI over time (1981-2011) analysing regional trends which were subsequently compared to in-situ measurements of herbaceous biomass. The use of the very long time series of field measurements with 28 years for Gourma (1984-2011) and 17 years for Fakara (1994-2011) makes this study special. In a trend analysis for the entire Sahel, they could confirm the general re-greening hypothesis but also detected negative trends in the greater region of Fakara and in central and eastern Sudan. The focus analysis for the Gourma region showed significant positive trends over the whole time series of NDVI and in-situ measurements. Because the study period began in 1984, a year with a severe drought, the re-greening was seen as a sign of strong ecosystem resilience. Similar trends were found for the local analysis in the Fakara region until the mid-1990s. Afterwards, the local NDVI decreased, which was explained by increasing human land use like an expansion of croplands and increased grazing. Anyway, an additional stratification by soil type in the Gourma region showed that the greening is primarily related to sandy surfaces while shallow soils exhibit nearly no significant correlation. Furthermore, a large scale increase in surface run-off on shallow soils was determined in other studies of the Gourma region and further Sahel regions (Descroix et al., 2012, 2009) resulting in strong erosion for some areas. Dardel et al. (2014) concluded that local degradation may occur masked by a regional re-greening trend.

3.1.5 Discussion and Conclusion

In this section, the main findings of this review study and the resulting needs for future work on vegetation dynamics in West Africa are summarized and discussed. From these results, the topic of this work was deduced.

Need for Regionally Optimized Remote Sensing Products

Free global remote sensing products of e.g. NDVI, NPP or LAI from various sensors like MODIS, SPOT VEGETATION or AVHRR have greatly advanced research on vegetation dynamics. These products have been updated over several years and were improved for different areas of the world. Nonetheless, their regional applicability is often limited because they cover the whole globe and thus a variety of different ecosystems (Fensholt et al., 2006; Sjöström et al., 2013). Several issues arise for a diverse place like West Africa that emphasize the need for regionally optimized products. A common example is for areas with partially dense vegetation cover like evergreen forests where a saturation or underestimation can be observed in several global products like GIMMS-3g NDVI, CYCLOPES or GEOV1 LAI datasets (Dardel et al., 2014; Garrigues et al., 2008; Gessner et al., 2013). Therefore, a regional optimization might be necessary in order to achieve accurate results, especially for applications that specifically target such densely vegetated areas.

A further issue often is the spatial resolution of these products which is mostly of 1 km or less. In West Africa, land use and land cover and their changes are highly dynamic but happen on comparably small scales which might not be derived from global products. In addition, input datasets used for the generation of several global products (like climatic drivers or biome-specific parameters) are often on a rather coarse scale or information level. For example, Sjöström et al. (2013) emphasized the potential of the MODIS Gross Primary Production (GPP) dataset but suggested an enhancement of the parameters for the biome property look-up table as well as the climatic input data.

In areas with a high probability and coverage of clouds, a general problem of global products is the rather high percentage of data gaps. Data fusion approaches merging two sensors of different resolution, e.g. MODIS and Landsat could be explored for the reduction of data gaps while maintaining a high spatial resolution. A spatially and temporally consistent remote sensing dataset is important for various research topics, e.g. in phenological studies where long data gaps or strong temporal filtering and gap filling could mask annual variations like dry spells or fires in the phenological cycle.

Underrepresentation of the Southern Countries of West Africa

The map of study distribution in West Africa (Figure 3.1) clearly indicates an underrepresentation of southern countries in the literature. This trend is even more emphasized in contrast to the considerable research activities in the Sahel region focusing on

degradation or desertification. For some southern countries such as Liberia or Sierra Leone not even a single study was found specifically dealing with vegetation dynamics. One explanation could be the generally low availability of cloud-free remote sensing data in such regions which causes difficulties for the collection of consistent vegetation time series. A further reason for this underrepresentation might be the political instability of several of these countries in the past and present which makes field work and local research in general a difficult task. It could also be argued that research on phenology and vegetation dynamics is not as important in humid climates due to a high evergreen component of vegetation as it is in arid and semi-arid ecosystems. This might be part of the explanation, but the southern West African countries also experienced a serious land use intensification and an expansion of agricultural area during the past years which considerably affects the semi-natural vegetation. Studies on vegetation dynamics with a high temporal and spatial resolution are needed for these semi-natural areas in order to determine their condition and detect local degradation. This field of research becomes gradually more important due to its context in the REDD program. Since the condition of woodlands and forests considerably affects their capability to store carbon, not only their extent matters in this program but also potential degradation or improvement (Machwitz, 2010; Machwitz et al., 2011; Sasaki and Putz, 2009).

Challenge of Linking Land Surface Phenology with At-Ground Observations

One of the major current challenges in remote sensing is the linking of land surface phenology (LSP) with in-situ measurements of actual phenological events (Kirsten M. de Beurs and Henebry, 2010). This not only accounts for research in West Africa but also for other regions of the world where studies focusing on this topic are rather sparse. The major reason for this lack of studies obviously is the low data availability of at-ground phenological observations. In order to be useful for the analysis of vegetation dynamics, the in-situ observations need to be conducted regularly over a long period of time. In West Africa, phenological observation networks with a dense spatial coverage, such as used by White et al. (2009) for North America, do not exist. Brown and de Beurs (2008) for example had to use sowing dates resulting from the questioning of farmers to validate start of season estimations in the Sahel. However, this dataset can only be applied as a proxy for the start of the growing season or onset of greening.

Furthermore, scaling can be an issue and suitable methods need to be developed for a proper connection between the coarse resolution of commonly used sensors and point measurements of at-ground data. Nonetheless, such a linkage is essential for the understanding of how mixtures of LULC or plant communities as well as the different growth stages affect land surface phenology. This research could result in remote sensing based methods for the derivation of timing of actual phenological events or growth stages. Until now, metrics of LSP rather focus on the development of the NDVI curve than on the timing of

actual in-situ phenological events. Such approaches are suitable for relative analyses like change analyses of ecosystem characteristics but conclusions for the respective phenology of plants cannot be drawn (Reed et al., 1994). However, understanding the linkage between actual plant phenology and LSP as well as the subsequently derived metrics might improve the transferability of developed methods between study areas. Since vegetation phenology is a highly sensitive indicator for climate change, this field of research could give significant feedback to climate modellers and, especially in West Africa, might help to advance early warning systems for food security.

Need for Local and Comprehensive Analyses of Degradation

Several West African studies determined a greening trend in the Sahel and Sudanian region over the last 30 years, but also emphasised that degradation could still occur on a local scale (e.g. Fensholt et al., 2013; Seaquist et al., 2009). In this context, several studies investigating the NDVI-rainfall relationship found another influence factor on the NDVI which was presumed to be human impact (e.g. Herrmann et al., 2005). The general greening trend is well documented on a West African scale, but several areas were identified for possible degradation. Thus, local analyses at higher spatial resolution are needed to confirm these findings and study the causes for greening or degradation.

Herrmann and Tappan (2013) conducted an important field study in this context investigating the remotely sensed greening trend for the central Senegal. They focused on abundance and composition of woody vegetation in greening sites with botanical inventories and further field methods. Contrary to the general rise in precipitation detected for the Sahel after the droughts of the 1980s, the study area experienced comparably dry conditions until the mid-1990s. Herrmann and Tappan (2013) detected a rise of woody vegetation density, predominantly in the shrub layer, while the number of trees decreased. They determined a loss of species richness and a trend towards xeric species. While others interpreted this change as an indicator of degradation (Gonzalez, 2001; Gonzalez et al., 2004), Herrmann and Tappan (2013) emphasised that it could also be interpreted as an adaptation of the ecosystem to the arid climate conditions until the mid-1990s. They reasoned that the ecosystem could have found a new equilibrium state which so far was not reversed despite the increase in precipitation during recent years. The different possibilities of interpreting field study results point out the complexity of degradation and the difficulty of determining this process based on coarse remote sensing data. The studies conducted by M. Boschetti et al. (2013) and Dardel et al. (2014), presented in the section 3.1.4, further outlined possible causes of greening hiding local degradation. While Herrmann and Tappan (2013) stress the value of regional remote sensing studies as a starting point of investigation, they caution against premature conclusions.

A major issue can also be the definition of land degradation in remote sensing based analyses. Most studies focus on the biological productivity as a proxy and do not deal with the complexity, i.e. the biodiversity and species composition of the area under investigation. This might result in wrong conclusions when an increase in plant productivity is equalled with 'no degradation' or 'improvement' of vegetation conditions. Such conclusion can be erroneous when species compositions changes towards more productive but unpalatable herbs or bushes on over-grazed rangelands or in cases where alien species replace indigenous vegetation (Ward, 2005). For example Knauer (2011) showed for the Western Cape region in South Africa that the general remote sensing methods and the interpretations of results for land degradation have to be changed. Since the indigenous Fynbos vegetation is generally less productive than the invasive alien Acacia species, a rise in productivity could result from the displacement of the indigenous, diverse vegetation by invasive species. This is by definition considered as degradation concerning the biological complexity.

All in all, the presented examples demonstrate that the remotely sensed greening trend for the Sahel and Sudanian region is obviously caused by various factors that can be interpreted as an improvement but sometimes also as a degradation of ecosystems. In order to draw the right conclusions, multi-scale studies are recommended investigating local degradation and greening trends which were detected from coarse resolution datasets.

Conclusions for the Research Topic of this Work

The major identified research gaps derived from this pre-study and discussed in this section are summarized as follows:

- Even though the semi-arid regions of West Africa are a hot spot of food crises, studies on crop yield forecasting seem sparse and outdated in comparison to research in other regions of the world. Suitable operational methods could be used as an early indicator of food crises and could help in the regulation of prices and planning of necessary imports. However, for this purpose, accurate maps of agricultural area are an important prerequisite but as will be presented in section 3.3, are so far mostly missing for West Africa.
- Although, there are several global remote sensing products suitable for the analysis of vegetation dynamics, they have certain limitations for the use in West Africa. Thus, regionally optimized products seem to be needed in order to have a temporally and spatially consistent dataset necessary for several applications in the region.
- Studies related to vegetation dynamics and development are mostly missing for the southern countries of West Africa, which is most likely not an expression of low interest in these regions but rather an issue of low data availability due to frequent cloud coverage.

- Concerning the general monitoring of land surface phenology, a link to actual phenological stages of the vegetative cycle is mostly missing. This is also an issue of remote sensing data availability, especially in the region of West Africa, since consistent high spatial and temporal resolution time series are needed to link LSP with actual plant phenology.
- More comprehensive analyses covering the complexity of degradation seem to be needed in order to draw reliable conclusions on the vegetation trends in West Africa. For this purpose, local investigations of these trends could improve the general understanding. In order to draw representative conclusions from local analyses on regional scale, consistent high spatial resolution datasets are needed.

The identification of these research gaps finally lead to the definition of the above described objectives of this thesis. The last three points indicate a high need for consistent high spatial and high temporal resolution time series. Thus, in this work it was intended to develop a novel spatio-temporal data fusion framework in order to improve past data availability of remote sensors in the region (**Objective I**). The first two points indicate the need for regionally optimized information on agricultural area extent and changes with a high spatial resolution and accuracy. Thus, in this work, quantifying and analyzing the agricultural area's extent and changes in Burkina Faso throughout 14 years was defined as a final goal (**Objective II**).

3.2 Available Satellite Sensors

For the objectives of this study, remote sensing data is needed in a sufficiently high spatial and temporal resolution in order to delineate single fields and to distinguish agriculture from natural vegetation. In general, satellite sensors are divided into optical sensors and radar sensors. Here, it was decided to use optical sensors since the discrimination of agricultural area from other vegetation on the basis of radar data did not seem promising in West Africa. During the past decades, the number of earth observation satellite sensors has increased considerably, delivering a vast amount of multi-temporal and multi-spectral remote sensing data. Depending on the application and the budget, satellite datasets in varying spectral, temporal and spatial resolutions can be purchased or accessed free of charge (Table 3.1). However, till recently there was no single satellite sensor that can provide remote sensing datasets of (very) high spatial and temporal resolution at the same time. The MODIS sensor aboard the Terra and Aqua satellites for example, can provide a daily temporal coverage of the same region but its spatial resolution is of 250 m or lower. Commercial sensors such as the ones aboard the WorldView satellites have very high spatial resolutions of less than 3 m but because of their narrow swath width, the temporal and

spatial coverage is generally very low. With WorldView and some other commercial instruments, this revisit time can be increased by strategic pointing of the sensor, but the costs for such datasets are high.

Table 3.1: Overview of common optical satellite sensors; in case a satellite has more than one optical sensor, the sensor name is specified; Spatial resolutions: very high ***(<5 m), high ** (5-30 m), medium to low * (>30m); Repeat cycle: high *** (<5 days), medium ** (5-16 days), low * (> 16 days) (modified after B. Chen et al., 2015).

Satellite (Sensor)	Spatial Resolution	Repeat Cycle	Lifetime	Access
WorldView	***	*	2007 - present	Commercial
GeoEye-1	***	*	2008 - present	Commercial
QuickBird	***	*	2001 - present	Commercial
IKONOS	***	*	1999 - present	Commercial
SPOT HRV/HRVIR/HRG	**	*	1986 - present	Commercial
ALOS AVNIR-2	**	*	2006 - 2011	Commercial
RapidEye REIS	**	*	2008 - present	Commercial
Terra ASTER	**	*	1999 - present	Free
EO-1 Hyperion	**	*	2000 - present	Free
EO-1 ALI	**	*	2001 - present	Free
Landsat MSS/TM/ETM+/OLI	**	**	1972 - present	Free
Envisat MERIS	*	**	2002 - 2012	Free
Terra/Aqua MODIS	*	***	2000 - present	Free
SNPP VIIRS	*	***	2011 - present	Free
NOAA AVHRR	*	***	1982 - present	Free
SPOT-Végétation	*	***	1998 - present	Free
Proba-V	*	***	2013 - present	Free
Sentinel-3	*	***	2016 - present	Free
Sentinel-2A+B	**	***	2015/2017 - present	Free

Since several research applications require or would greatly benefit from high spatial and temporal resolution datasets, constellations of several satellites with high spatial resolution could be a solution to also achieve a sufficient temporal coverage. For this purpose, the European Space Agency (ESA) has started the Copernicus Sentinel program comprising different sensor platforms (Sentinel-1 to Sentinel-6) of which some shall be launched in a constellation of two identical satellites. Especially the Sentinel-2 series, which started in June 2015 with Sentinel-2A and was completed with Sentinel-2B in March 2017, will be of interest to gather data with a high spatial resolution (up to 10 m) and a frequent temporal coverage (up to 2 days in constellation). However, since no Sentinel datasets are available prior to 2015, other solutions are necessary to gather time series of high temporal and spatial resolution to accomplish the objectives of this work. Furthermore, in cloud-prone areas like West Africa, the availability of gap-free datasets from Sentinel-2 and other sensors is greatly reduced. In

addition, less datasets from Landsat sensors prior to Landsat-8 were recorded in regions such as West Africa further decreasing data availability.

3.3 Mapping of Agricultural Area in West Africa

In West Africa, the mapping and monitoring of agricultural area is crucial to understand implications for food security, regional development and for the protection of habitats and thus provide a reliable information basis for sustainable land use management. Statistical data on agriculture mostly exists on national or provincial level and is not outlining the exact coverage and spatial distribution of the areas and such datasets may comprehend considerable uncertainties (Knauer et al., 2017). Besides statistics, the extent of agricultural area can also be derived from remote sensing based land cover classifications. Such datasets are available for West Africa in varying spatial resolution and coverage and for different years. These products can be on a global scale, on a regional scale comprising several countries or focusing on a local scale covering small study areas on sub-national level.

3.3.1 Global Maps

An increasing number of global land cover datasets exists that also include different classes of agricultural area. Amongst the most common range the MODIS MCD12Q1 product (Friedl et al., 2010), the GlobCover classification (Bontemps et al., 2011), the Global Land Cover (GLC) 2000 (Bartholomé and Belward, 2005) or the ESA CCI (Climate Change Initiative) land cover product (ESA, 2016a).

The MCD12Q1 land cover product is a 500 m global, yearly classification of MODIS data from both Terra and Aqua satellites available for the years 2001 to 2013. It contains two classes of agricultural area: cropland and a mosaic of croplands and natural vegetation. Perennial woody crops are not included in the cropland class but assigned to the natural vegetation classes of forest or shrubland. Irrigated agricultural area is also not specified in a separate class in the MCD12Q1 land cover product. The ESA CCI land cover consists of three 300 m spatial resolution classifications for the years 2000, 2005 and 2010. Each of them represents the land cover of a 5-year period (1998-2002, 2003-2007, and 2008-2012) derived in a multi-year and multi-sensor strategy from MERIS and SPOT-Vegetation data (ESA, 2016a). The CCI land cover products contain the classes 'cropland, rainfed', 'cropland, irrigated or post-flooding' and two types of cropland-vegetation mosaics depending on the dominance of one of the classes. As for the MODIS product, plantations are not specified in this classification (ESA, 2016b). The GlobCover product is in a way the predecessor of the CCI products containing the same classes of croplands and global 300 m MERIS classifications for the years 2005 and 2009 (Bontemps et al., 2011). The last of the most common global land cover products is the GLC 2000 product, which was derived from the SPOT-Vegetation sensor in 1

km spatial resolution for the year 2000 and on two different thematic levels, a global and a regional level. The regional map for Africa contains four agricultural classes, a cropland class, cropland-vegetation mosaic, irrigated and tree crops.

Several studies comparing such global land cover products showed considerable discrepancies between them, especially for West Africa (Gessner et al., 2012; Lambert et al., 2016). Since these products were originally designed as input to climate change and modelling studies, exact regional extent of croplands is substantially variable, e.g. the GlobCover product estimates 20% more cropland globally than the MODIS product (Fritz et al., 2013). Focusing on West Africa, inconsistencies in cropland extent are often even higher than in other regions due to the special cropping systems and cloud coverage hampering input data availability for the classification (Gessner et al., 2012; Lambert et al., 2016).

During recent years, attempts have been made to increase the spatial resolution of global land cover datasets and several products in a 30 m spatial resolution such as the Chinese GLOBELAND30 or the FROM-GLC (Finer Resolution Observation and Monitoring of Global Land Cover) have been released (J. Chen et al., 2015; Yu et al., 2013). Both products are based on Landsat data and contain one cropland class (in the first version). While the GLOBELAND30 was derived for the year 2010, the FROM-GLC product used input datasets from the whole times series of Landsat TM and ETM+ and thus the result cannot be attributed to a single year. Though, the latter was seen as a first attempt to map global land cover at this scale, the use of this product especially in a fast changing landscape like West Africa is limited.

Another attempt to map global croplands at a 30 m resolution was started by USGS, NASA and others within the GFSAD30 (Global Food Security-Support Analysis Data at 30 m) project (USGS, 2017a). The aims are to derive global cropland extent from the Landsat archive for multiple years starting in 1990, for different crop types and intensities, and discriminating between rainfed and irrigated croplands. A first beta version of a global crop/non-crop map from this project can be inspected online. Specifically for Burkina Faso, this map so far shows substantial data gaps covering about one quarter of the country and considerable regional under- and overestimations.

3.3.2 Regional Maps

For West Africa, several studies highlighted the need for regionally optimized maps of land cover and agricultural area in particular (Forkuor, 2014; Fritz et al., 2010; Gessner et al., 2012; Lambert et al., 2016).

Tappan et al. (2016) generated three LULC classifications with 26 classes for the years 1975, 2000, and 2013 covering West Africa in 2 km spatial resolution. This was conducted on the basis of visual interpretation of Landsat images and the Rapid Land Cover Mapper (RLCM). The products were developed to reveal broad scale changes in West Africa's land resources over a long time span of 38 years. Although these are valuable information for climate

modelling and to show long-term, region wide developments, their suitability for detailed national analyses of agricultural area is limited.

Gessner et al. (2015b) applied a multi-sensor approach for the mapping of land cover in the core area of West Africa which corresponds to the focus region of the WASCAL project. They used MODIS data for the derivation of vegetation classes including agriculture, Envisat ASAR (Advanced Synthetic Aperture Radar) data for the delineation of water bodies and the Global Urban Footprint (GUF) based on TanDEM-X and TerraSAR-X data (Esch et al., 2013) for the classification of urban areas. The final product is in a 250 m spatial resolution for the year 2006 and contains 14 classes including one for agriculture. This land cover product gives a very good overview about the distribution of land cover, especially agriculture in West Africa. However, due to the spatial resolution of 250 m difficulties arise with the discrimination of natural vegetation classes and the typically small-scale agriculture of the region.

Lambert et al. (2016) generated a 100 m cropland map for the Sahel and Sudanian savanna of West Africa and the year 2014 from time series of the Proba-V satellite. As a basis for classification, they derived five temporal metrics from the time series: the minimum and maximum of NDVI, its rate of increase and decrease during the growing season as well as the maximum of the red band. A comparison of the results with several of the above mentioned global maps showed improved accuracies of crop delineation against most other products. Main sources of error were attributed to low data availability, especially during the peak growing season, and to small-scale cropping in heterogeneous areas despite the relatively high spatial resolution of 100 m (Lambert et al., 2016). Another general drawback of this approach is that it is not repeatable for preceding years in this spatial resolution due to the lifetime of the sensor and thus an analysis of past spatio-temporal changes is not possible.

3.3.3 Local Maps

With the opening of the Landsat archive and a growing number of very high spatial resolution sensors, the number of studies aiming at the classification of local, sub-national areas increased significantly in West Africa and in the world in general. Recent studies range from the mapping of local LULC changes (Zoungrana et al., 2015), over the investigation of drivers for such changes (Ouedraogo et al., 2015), the analysis of benefits of time series features for LULC mapping (Liu et al., 2016) to the discrimination of different crop types in West Africa (Forkuor et al., 2015, 2014). Zoungrana et al. (2015) for example derived three classifications (1999, 2006, and 2011) from Landsat data for a study area in the southwest of Burkina Faso. They found a considerable increase of agricultural area and bare surface over the period of investigation at the expense of semi-natural vegetation. In addition, they also investigated the added value of multi-temporal classification over mono-temporal classification and found significantly higher accuracies for the use of several input images. Forkuor et al. (2014) investigated the joint use of very high resolution optical data from

RapidEye with SAR data from the TerraSAR-X satellite for crop type mapping in northwestern Benin. They classified fields of five crop types and found the joint use of the two satellite systems to improve the accuracy for this purpose by 10% to 15% in comparison to the sole use of RapidEye data.

These studies are valuable in-depth analyses of the driving forces of LULC change in the region, improve the understanding of remote sensing features and signals in relation to at ground observations. However, the resulting maps of LULC or agricultural extent are not intended to give information on broader scales, i.e. on national or regional level since they are only covering small study areas.

3.3.4 Challenges for Mapping and Monitoring in West Africa

In the previous sections, multiple remote sensing based maps of LULC or agriculture in particular were presented. The difficulties that most of them face in the delineation of agricultural area in West Africa indicate that this is a highly challenging region. These problems mainly arise from the common West African cultivation practices as outlined in section 2.6. Since the level of mechanisation is low and farmers mostly cultivate less than 5 ha of land, agriculture in West Africa is conducted on a small scale. This results in highly heterogeneous landscapes with a mosaic of semi-natural vegetation, small-sized cropland and fallow. Several studies focusing on the regional optimization of LULC classifications have reported this landscape heterogeneity as a major source of error (Gessner et al., 2015b; Lambert et al., 2016). When focusing on a typical agricultural region in Burkina Faso comparing common field sizes with pixel sizes, this issue becomes apparent (Figure 3.2). Most global LULC maps such as the MODIS MCD12Q1 product do not exceed a spatial resolution of 500 m. This however results in a wide mixture of classes covered by a single pixel. While the average field size in West Africa is below 1 ha, a 500 m MODIS pixel covers about 25 ha of land. Even the 250 m resolution from MODIS as used in the regionally optimized product for West Africa by Gessner et al. (2015b) is way above the average field size covering about 6.25 ha per pixel. As indicated in Figure 3.2, the spatial resolution of Landsat (30 m) could be sufficient for an accurate delineation of agricultural area in West Africa and should result in an acceptable amount of mixed pixels in the classification output. The pixel coverage of about 0.09 ha is significantly lower than the average field size of about 1 ha.

However, not only the spatial resolution is important for a successful delineation of agricultural area, also a sufficient temporal resolution is crucial. Due to the often times rather inefficient cultivation practices with low input of fertilizer and pesticides, the spectral signal of an agricultural area can be difficult to distinguish from other classes, especially in a mono-temporal dataset. Agricultural area is often only distinguishable from semi-natural vegetation during salient periods of the cropping season, e.g. at its start or during its peak when the NDVI is highest. However, the timing of these dates is highly variable between different years

depending of the onset and length of the rainy season. Consequently, several studies showed the added value of multi-temporal input images or even the use of LSP features derived from time series for the classification of agricultural area in West Africa (Lambert et al., 2016; Liu et al., 2016; Zoungrana et al., 2015).
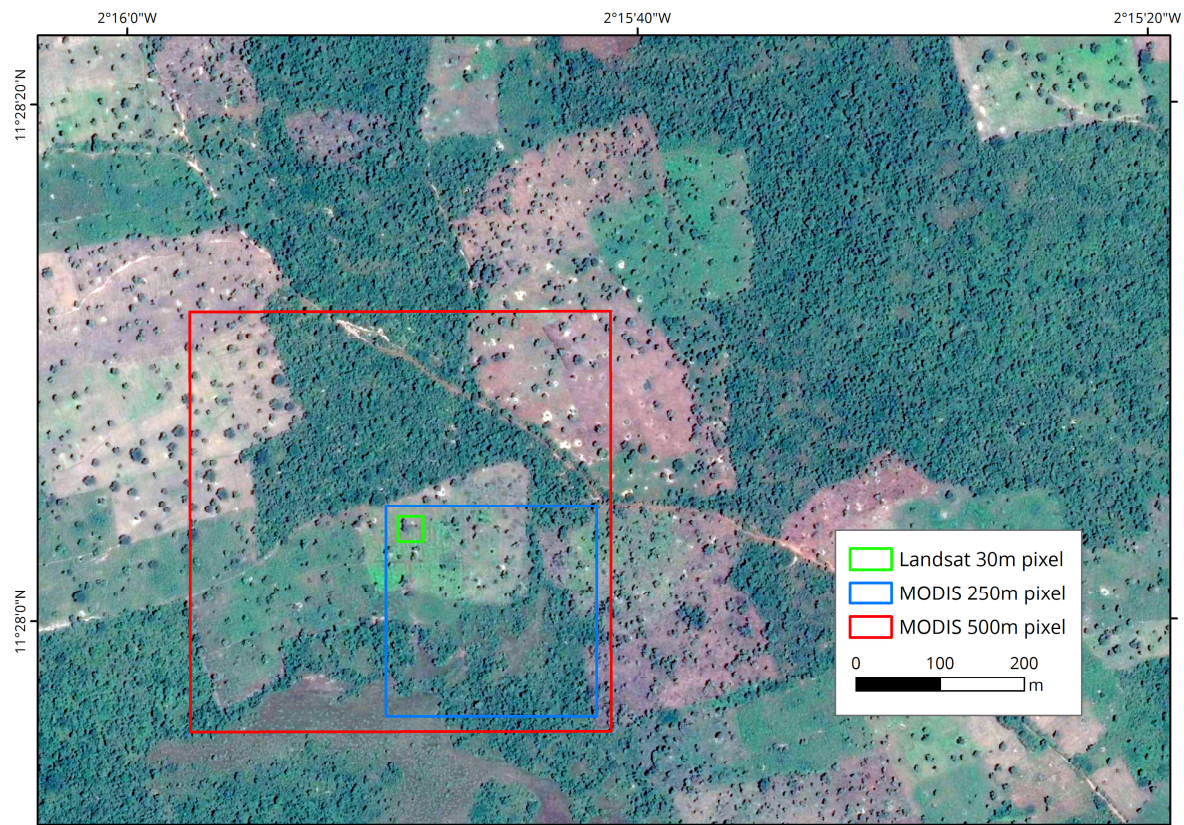
Figure 3.2: WorldView-2 image of a rural area in southern Burkina Faso with pixel extents of three exemplary spatial resolutions from MODIS (500 m and 250 m) and Landsat (30 m) in a rural area of Burkina Faso.

In addition, periods of high potential for class separability mainly occurring during the growing season are often affected by frequent cloud cover. This can be a serious issue for image acquisition in mono-temporal classification approaches but also for classifications based on time series when not enough datasets are available for the re-construction of the phenological cycle. Especially when working on larger scales covering several tiles of satellite data, it is almost impossible to obtain a consistent mono- or multi-temporal dataset. This is not only hindered by frequent cloud coverage but also by different acquisition dates of satellite paths which can result in different representations of phenological stages. Considerably varying data availability for multiple satellite tiles can result in spatial differences in the quality of the classification result and also in visible inconsistencies at tile borders as observed for example in the high resolution classification of GLOBELAND30 (J. Chen et al., 2015).

As will be shown in this work, not only rainfed agricultural area plays an important role in West Africa and Burkina Faso in particular but also irrigated cultivation and plantations. Irrigated agricultural area is often not classified in regional or global LULC maps or it is included in a single agriculture class. However, the monitoring of this cultivation practice in West Africa is important since irrigation can represent a way to cope with the above described varying timing and length of the rainy season and thus a factor of land use adaptation and food security. Furthermore, none of the above presented classifications includes plantations of e.g. fruit trees. In most maps, these areas fall within classes of dense woody vegetation. Nonetheless, plantations are of increasing importance as an alternative source of income for West African farmers. Thus, it would be of great interest to specifically map plantation area, but again this can be a difficult task due to the spectral similarity of plantations and forests.

All in all, West Africa is a highly challenging region for the mapping and monitoring of agricultural area. For a successful discrimination from semi-natural vegetation classes, sufficient spatial and temporal resolution is necessary in order to outline the comparably small-scale and unproductive agriculture of West Africa. Up to now, no single satellite sensor was able to provide a satisfactory combination of these two resolutions (compare section 3.2) for past and recent years which is indicated by the difficulties of existing global and regional maps in the delineation of agricultural area in West Africa.

3.4 Spatio-temporal Data Fusion

A powerful solution to overcome the trade-offs between pixel size and swath width of currently available satellite sensors is spatio-temporal data fusion. During recent years, various data fusion methods have been developed to blend data of different satellite sensors such as MODIS and Landsat to generate or as commonly used to 'predict' datasets of high temporal and spatial resolution (e.g. Gao et al., 2006; Hilker et al., 2009; Zhu et al., 2010). In general, all of these methods need at least one image pair at the same date consisting of a high and low spatial resolution scene and a low spatial resolution scene at the date of prediction as input for training. The result is a synthetic high spatial resolution image at the date of prediction. These spatio-temporal fusion methods commonly use the temporal information, e.g. the temporal change of reflectance between the pair date and the prediction date, from the low spatial resolution images of the input and the spatial detail from the high spatial resolution images (Zhu et al., 2016).

Following Zhu et al. (2016), current spatio-temporal data fusion techniques can be organized into three different groups that are based on either (1) unmixing, (2) dictionary-pair learning or (3) weighted functions.

3.4.1 Unmixing Based Data Fusion

The unmixing based spatio-temporal fusion techniques rely on the classic Linear Spectral Mixture Model (LSM) for the extraction of endmembers and their abundances on sub-pixel scale (Bioucas-Dias et al., 2012; Gevaert and García-Haro, 2015). For the purpose of remote sensing data fusion, the number of endmembers and abundances are derived from the high spatial resolution image and the spectral signatures of endmembers are unmixed from the low spatial resolution image (Gevaert and García-Haro, 2015). Generally, the unmixing based methods consist of four steps for the prediction of a synthetic high spatial resolution image:

1. Clustering or classification of the high spatial resolution image for the definition of endmembers.
2. Calculation of the fractions of endmembers within each low spatial resolution pixel.
3. Unmixing of the low spatial resolution pixels of the prediction date.
4. Assigning unmixed reflectance spectra to the synthetic high spatial resolution image at the prediction date (Gevaert and García-Haro, 2015; Zhu et al., 2016).

This method originates from the multi-sensor multi-resolution technique (MMT) developed at DLR by Zhukov et al. (1999). Thereafter, several studies and algorithms were built on this technique trying to improve drawbacks and raise the derived accuracies. Zurita-Milla et al. (2008) constrained the linear unmixing process to ensure that the resulting spectral reflectances are above zero and below a pre-defined upper limit. Gevaert and García-Haro (2015) however expected some unrealistic endmember spectra from this approach, since it is conducted separately for each band. In their unmixing based algorithm called Spatial and Temporal Reflectance Unmixing Model (STRUM), Gevaert and García-Haro (2015) directly used the spectral change between the low spatial resolution images to estimate the change of endmembers and applied a Bayesian approach incorporating prior spectral information to constrain the unmixing process. Amoros-Lopez et al. (2013) employed a cost function to regulate the unmixing process and to prevent the resulting endmember reflectances from differing too much from a pre-defined endmember spectrum for each class.

Studies on unmixing based data fusion so far often focus on the methodological development of algorithms and few apply the data fusion results to investigate specific thematic issues. The above mentioned Amoros-Lopez et al. (2013) applied their fused time series derived from Landsat-TM and MERIS to an agricultural area in Albacete, Spain for the purpose of crop monitoring. They highlighted the value of the fusion results for a better discrimination of small sized fields thanks to the high spatial resolution and a better separation of crop types due to the improved spectral and temporal resolution (Amoros-Lopez et al., 2013). Zurita-Milla et al. (2009) used their MERIS-derived fused time series for the monitoring of vegetation dynamics in the Netherlands with subsequent calculation of several

vegetation indices. In a later study (Zurita-Milla et al., 2011), they applied these vegetation index time series for land cover mapping and found that some of the classification results outperformed a corresponding land cover classification derived from Landsat-TM.

The major advantage of unmixing based methods in comparison to other spatio-temporal data fusion approaches is that they do not depend on matching spectral bands between low and high spatial resolution datasets (Gevaert and García-Haro, 2015). Thus, unmixing based methods can also be used to increase the spectral resolution of the synthetic high spatial resolution dataset and auxiliary land cover information can even completely replace high spatial resolution reflectances in the clustering process (Gevaert and García-Haro, 2015). However, the drawback of these spatio-temporal data fusion approaches is the pre-requirement that no land cover changes occurred between the input dates and the prediction date (Zhu et al., 2016). Furthermore, the quality of results of such methods is easily affected by co-linearity problems and noise from the input data since a moving window approach is used on the high spatial resolution reflectances for the unmixing process (Zhu et al., 2016).

3.4.2 Dictionary-Pair Learning Based Data Fusion

In the field of spatio-temporal data fusion, dictionary-pair learning based methods are rather new. This type of data fusion uses structural similarities between the high and low spatial resolution images to establish correspondences between them and capture the major features like land cover changes for the predictions (Huang and Song, 2012; Zhu et al., 2016). With the development of the SParse-representation-based SpatioTemporal reflectance Fusion Model (SPSTFM), Huang and Song (2012) were probably the first to transfer dictionary-pair learning from image super-resolution to spatio-temporal data fusion (Zhu et al., 2016). The SPSTFM directly establishes a relationship between the changes of two pairs of high and low spatial resolution images and uses this information for the training of two dictionaries. These are then applied to generate a synthetic high spatial resolution image at the date of prediction (Zhu et al., 2016). Since the need for two image pairs was seen as a drawback of this method, Song and Huang (2013) developed another technique that only uses one image pair of low and high spatial resolution data. The new algorithm first increases the spatial resolution of the low spatial resolution scenes at training and prediction date via sparse representation. Secondly, high-pass modulation is used for the fusion of these super-resolution images with the actual high spatial resolution image at the training date to generate a synthetic scene at the prediction date (Song and Huang, 2013). Although these dictionary-pair learning based methods are able to predict pixels with ongoing land cover changes, they only use statistical relationships between the image pairs instead of actual physical properties of the remote sensing signal (Zhu et al., 2016). Furthermore, when the scale difference between the low and high spatial resolution images is considerable, they cannot accurately reproduce the shape of objects and blur them in the output (Zhu et al., 2016). Since these techniques are relatively

new and are still in the stage of further development, they have not been applied extensively for the investigation of certain thematic issues.

3.4.3 Weighted Function Based Data Fusion

Probably the most frequently used type of spatio-temporal data fusion techniques so far are the ones based on weighted functions. These make use of similar surrounding high spatial resolution pixels and weighting functions to predict a pixel at the target date from the input images. The first method of this type was the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) developed by Gao et al. (2006). STARFM uses at least one pair of high and low spatial resolution images and a low spatial resolution image at the prediction date to generate a synthetic high spatial resolution scene. A general assumption of the algorithm is that changes in surface reflectance occurring between two dates are consistent and comparable at the different spatial resolutions if the low spatial resolution pixels are "pure" pixels, i.e. of only one land cover type (Gao et al., 2015; Zhu et al., 2016). This is determined by spectral similarity of high spatial resolution pixels within the same low spatial resolution pixel. In such a case, changes between the two dates of low spatial resolution pixels should be directly transferred to the high spatial resolution pixels to obtain a prediction:

$$H(x, y, t_p) = H(x, y, t_k) + (L(x, y, t_p) - L(x, y, t_k))$$

(3.1)

where (x,y) is the pixel location of the high spatial resolution image (H) and the one on the resampled low spatial resolution image (L), t_k is the acquisition date of these images and t_p the prediction date (Gao et al., 2015). However, since such an ideal situation cannot often be satisfied and most of these low spatial resolution pixels are rather mixed pixels, i.e. containing several land cover types, STARFM uses a weighted function including neighbouring spectrally similar pixels in the prediction (Gao et al., 2015). The whole pixel-wise prediction is conducted in a moving window approach as follows:

$$H\left(\frac{w}{2}, \frac{w}{2}, t_p\right) = \sum_{k=1}^{P} \sum_{i=1}^{N} W_{ik} \times (H(x_i, y_i, t_k) + (L(x_i, y_i, t_p) - L(x_i, y_i, t_k)))$$

(3.2)

Where w is the size of the moving window, ($w/2,w/2$) is the central pixel of the moving window, N is the number of spectrally similar pixels with the index i and P is the number of image pairs of high and low spatial resolution (Gao et al., 2015). The weight W_{ik} defines how much each of the neighbouring pixels contributes to the prediction of the central pixel of the moving window from the respective image pair k (Gao et al., 2015). This weight function is

determined by three measures: (1) the spectral difference between high and low spatial resolution pixels, (2) the temporal difference between the date of prediction and the image pair and (3) the geographic distance between the central pixel of the moving window and the candidate pixel (Gao et al., 2015). Pixels with low differences are assigned a higher weight.

The STARFM method also has some issues influencing the quality of prediction: as most weighted function based spatio-temporal data fusion methods, it cannot accurately predict abrupt disturbance events, especially when these disturbances are not visible in one of the high resolution images used as input. Furthermore, STARFM in particular is struggling with the prediction in heterogeneous landscapes due to its dependency on available pure, low spatial resolution pixels of homogeneous patches of land cover (Gao et al., 2006; Zhu et al., 2010).

The STARFM algorithm was modified and improved in subsequent years multiple times (e.g. Hilker et al., 2009; Wang et al., 2014; Zhu et al., 2010). One of these new methods based on STARFM is the Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) algorithm (Hilker et al., 2009). It was developed to improve the original algorithms capability to handle abrupt changes or disturbances between pair and prediction date. This is done by determining spatial changes from the high spatial resolution image and temporal changes from the low spatial resolution image allowing the algorithm to select optimal base date for the input image pairs (Hilker et al., 2009; Zhu et al., 2016). However, since the STAARCH algorithm subsequently uses the STARFM for the data fusion, it has the same limitation in heterogeneous landscapes.

The most frequently used modification of STARFM is probably the Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) developed by Zhu et al. (2010). It was mainly designed to improve the prediction quality of STARFM in heterogeneous landscapes (Zhu et al., 2010). In contrary to the original algorithm, which assumes the changes on the different spatial scales to be equivalent, ESTARFM assumes proportional changes between high and low spatial resolution pixels (Gao et al., 2015). In order to account for these proportional changes, a conversion coefficient is introduced in ESTARFM to translate changes from low spatial resolution to high spatial resolution for each pixel. For this purpose, ESTARFM always uses two image pairs of low and high spatial resolution, one from before and one after the date of prediction. The similar pixels for ESTARFM are also selected in a moving window approach. The exact derivation of similar pixels and conversion coefficients is described in chapter 5.2.

The final computation of a synthetic pixel in the center of a moving window with ESTARFM is conducted as follows:

$$H\left(\frac{w}{2},\frac{w}{2},t_p\right) = \sum_{k=1}^{2} T_k \times \left(H\left(\frac{w}{2},\frac{w}{2},t_k\right) + \sum_{i=1}^{N} W_{ik} \times v(x_i,y_i) \times \left(L(x_i,y_i,t_p) - L(x_i,y_i,t_k)\right)\right) \qquad (3.3)$$

where T_k is determining the temporal weight for each of the two image pairs, calculated based on the change magnitude between the low spatial resolution image at prediction date and the respective low spatial resolution images at the pair dates (Gao et al., 2015).

The major advantages of ESTARFM over the original algorithm are the modifications such as the conversion coefficient that target at the improvement of prediction in homogeneous areas. STARFM has difficulties in the prediction of objects that are significantly smaller than the low spatial resolution pixel if homogeneous pixels in this resolution are not found in the search window (Zhu et al., 2010). This issue results in blurry objects with indistinct boundaries. On the contrary, the modifications in ESTARFM facilitate an accurate delineation of such heterogeneous areas. However, it always needs two input image pairs instead of one for STARFM. This can be a problem in cloud-prone areas where gap-free high spatial resolution images can be difficult to find. In addition, ESTARFM still has issues with the prediction of abrupt changes or disturbances, especially if they are not detected on one of the input pair images. To mitigate this problem, it is recommended to use image pairs which are as close as possible to the prediction date but again this can be difficult in cloud-prone areas (Zhu et al., 2010).

Recent comparison studies of STARFM and ESTARFM highlight the respective advantages of these algorithms (Emelyanova et al., 2013; Jarihani et al., 2014). While ESTARFM is superior for regions with high spatial variance (heterogeneous areas), STARFM achieved higher accuracies for some cases with high temporal variance (rapid temporal changes). However, the overall accuracies for a whole time series or test set were found to be higher for ESTARFM (Emelyanova et al., 2013; Jarihani et al., 2014). In addition, Jarihani et al. (2014) demonstrated that the use of a vegetation index as input to the fusion results in higher accuracies than using reflectances and deriving the respective vegetation index from the output of the prediction since the error propagation is reduced this way.

The weighted function based methods described in this section have all proven to be highly suitable for the reconstruction of high spatial resolution phenology and have been applied in numerous studies since their development. Gao et al. (2013) generated STARFM NDVI time series from MODIS and Landsat for the purpose of crop condition monitoring on a focus site of about 225 km² in central Iowa, USA. From these time series, they derived land surface phenology (LSP) metrics with TIMESAT and found distinct features to distinguish different crop types (corn and soybean) from semi-natural vegetation types. Schmidt et al. (2015) demonstrated the potential of STARFM for the monitoring of forest disturbance and regrowth in Queensland, Australia. They fused a 12-year time series (2000-2011) from Landsat and MODIS for a study area of 2,960 km² and analyzed the results with the Breaks for Additive Seasonal and Trend (BFAST) algorithm (Verbesselt et al., 2010). This way, they identified events of forest clearing and were able to monitor the subsequent regrowth of vegetation.

On the basis of ESTARFM, Tewes et al. (2015) generated a time series from 250 m MODIS and 5 m RapidEye data for the monitoring of vegetation dynamics in a study area (1,434 km²) with semi-arid rangelands in South Africa. In this heterogeneous region, the ESTARFM time series was able to characterize the phenological development of different vegetation types and yielded more information than either of the two input time series (Tewes et al., 2015). In the Xinjiang Uyghur Autonomous Region of China, Gärtner et al. (2016) analyzed the benefit of ESTARFM time series derived from Landsat and RapidEye in comparison to the sole use of RapidEye data for monitoring of forest disturbances caused by insect defoliation (size of study area: 21 km²). They found the overall accuracies of disturbance detection to increase from 43% to about 64% using the fused time series.

The given examples for the use of STARFM and ESTARFM already indicate some general commonalities in the research with the weighted function based spatio-temporal data fusion. The majority of studies focus on the USA, China and Australia (Figure 3.3). Besides them, other countries occur much less in the weighted function based literature. The predominant countries USA and China are also the ones where the most algorithms were developed and subsequently tested. This results from the fact that a lot of these studies still focus on the methodological development itself and not on the investigation of specific geoscientific questions.
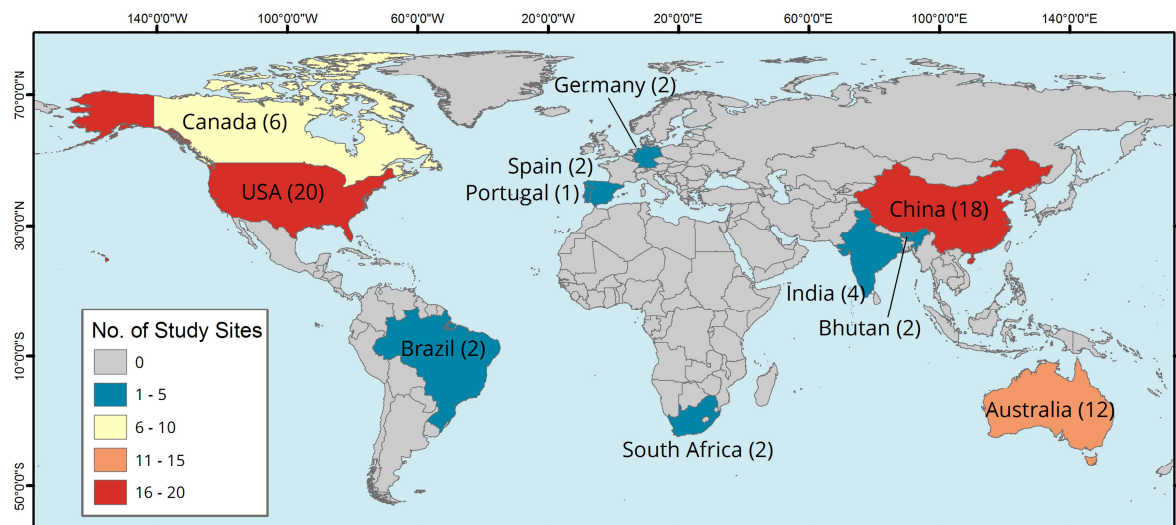
Figure 3.3: Number of study sites per country used in research with weighted function based spatio-temporal data fusion.

Whole continents like Africa or South America, where considerable research on land surface phenology is currently conducted are almost not investigated so far. However, LSP research in these areas could greatly benefit from the application of spatio-temporal data fusion since they also suffer from persistent cloud coverage reducing the amount of usable remote sensing datasets.

In addition, the study areas of the reviewed weighted function focused research are relatively small with an average size of 6,600 km² (Figure 3.4). These data fusion techniques have so far not been applied on national scale or for study areas greater than one Landsat scene (compare red line in Figure 3.4). Studies that investigated long and/or dense time series of several hundred synthetic images mostly focused on small-sized study areas below the general average value. One reason for the rather small investigated areas is that a lot of studies still focus on the methodological (further) development of algorithms and only need small study sites for performance tests. Another reason for the general trade-off between study area size and number of predictions is the considerable amount of computing time. Depending on the system specifications, the prediction of a single synthetic Landsat scene can take several hours. In addition, repeated manual input is necessary for the processing of a whole time series with most algorithms. The frequent use of such spatio-temporal data fusion approaches reflects the high interest in these techniques, but the distribution of studies and the sizes of study areas also point to some drawbacks which so far hindered broader geographic applications (Knauer et al., 2016).
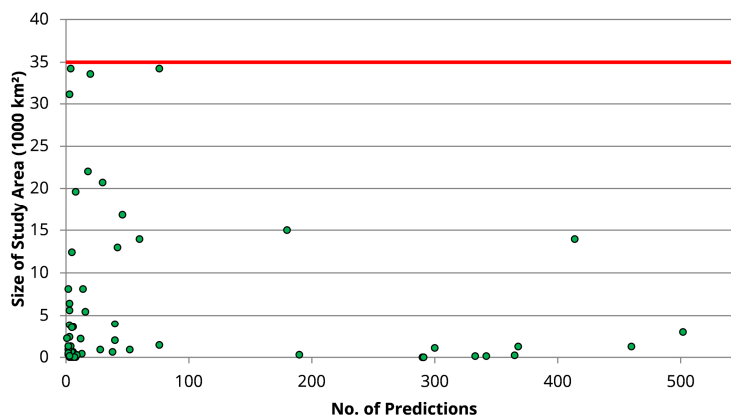
Figure 3.4: Overview of research on weighted function based spatio-temporal data fusion: number of predictions versus size of study area (in 1,000 km²); the red line marks the approximate size of one Landsat scene (35,000 km²).

In this work, the ESTARFM algorithm was used as a basis for the generation of dense time series. It proved to be highly suitable to reconstruct information on land surface phenology in a high spatial resolution (Gärtner et al., 2016; Zhu et al., 2010). Furthermore, it generally yields higher accuracies than other spatio-temporal data fusion algorithms in heterogeneous areas such as West Africa with its small-scale mosaic of agriculture and savanna (Jarihani et al., 2014). However, in order to be applicable for larger scales and in cloud-prone areas, where input remote sensing datasets are scarce, several improvements had to be made to the algorithm. These were developed within this thesis and are presented in chapter 5.

3.5 Classification Approaches

The general objective of remote sensing image classification is to derive a thematic map of identifiable features of the earth's surface from remote imagery by categorizing pixels into different LULC classes (Lillesand et al., 2004; Tso and Mather, 2009). In the history of remote sensing, numerous classification techniques have been developed for this purpose. The aim of this section is not a comprehensive review of classification approaches (which can be found in e.g. Tso and Mather, 2009) but shall give a short overview of the common methods and their respective advantages and drawbacks in order to describe the decision for the applied classifier in this work.

In general, image classification can be divided into two main approaches, supervised and unsupervised classification. The major difference between them is that in unsupervised methods, no reference data have to be gathered by the user for the training of the classifier (Tso and Mather, 2009). Here, the user specifies the number of target clusters (i.e. classes) and the classifier automatically combines pixels into classes by minimizing a predefined error function (Tso and Mather, 2009). Although, these methods do not need direct interaction with the user during the clustering, most often the user re-runs the algorithm multiple times by tuning given parameters until the result meets the user's expectation (Tso and Mather, 2009). The two most common unsupervised classification methods are the k-means algorithm and the ISODATA (Iterative Self-Organizing Data Analysis) technique (Lillesand et al., 2004).

In supervised classification methods, the user has to gather a representative reference database for each of the target classes (Tso and Mather, 2009). These classes have to be sufficiently distinguishable from each other on the basis of the given input remote sensing data. If this is not the case, the classification accuracy will significantly decrease. Thus, a proper definition of classes is crucial for a successful supervised classification. In addition, the selection of training areas is important which should be known sites of homogeneous land cover (Tso and Mather, 2009). The compilation of a suitable reference database for the classification can be a time consuming task, especially when the training areas are sampled at ground and not on the remote sensing data directly. However, for the generation of more sophisticated classifications, supervised methods are usually preferred over unsupervised ones since classes can be defined more accurately and they commonly yield higher accuracies (Tso and Mather, 2009).

The supervised methods can be grouped into two general types – parametric and non-parametric algorithms. The former ones are based on the assumption of a certain statistical distribution for each class, e.g. the normal distribution, requiring estimates of statistical parameters like the mean or variance (Forkuor, 2014; Schowengerdt, 2007). However, if this distribution does not exist the algorithm will fail to produce sufficiently high accuracies. This issue is especially the case in regions with heterogeneous classes like in West Africa where the

signature of e.g. agricultural area can vary considerably depending on the cultivation practices or the occurrence of the rainy season. Common parametric classifiers are the Level-Slice Classifier (or parallel-piped classifier), Minimum Distance Classifier or Maximum Likelihood Classifier (Schowengerdt, 2007; Tso and Mather, 2009).

Since non-parametric algorithms are not based on these theoretical assumptions and are generally easier to use, a lot of studies have determined their superiority and these algorithms are enjoying great popularity in recent years (Tso and Mather, 2009). The most common non-parametric methods are the Artificial Neural Networks, the Support Vector Machines and Decision Trees.

The development of Artificial Neural Networks (ANN) was inspired by the human brain's efficiency in pattern recognition learning how certain patterns in the dataset are shaped (Tso and Mather, 2009). As all non-parametric methods, they do not make assumptions on the statistical distribution of the input data but learn about regularities in the given training data. This 'knowle6dge' is used to create a ruleset which can subsequently be applied for the classification of other datasets or as in remote sensing the whole satellite image (Tso and Mather, 2009). Probably the most common ANN is the back-propagating multi-layer perception (Pal and Mather, 2003). Numerous studies highlighted the superiority of ANNs versus parametric methods like Maximum Likelihood (e.g. Pal and Mather, 2003) but there are also several drawbacks that hindered a more frequent application in the literature. As for other supervised methods a sufficient training of the classifier is crucial for a successful classification. For ANNs however, the manual training can be even more time consuming. The user has to define and built the architecture of the ANN including the definition of several parameters such as the learning rate which is not straightforward and difficult to learn for users new to ANNs (Tso and Mather, 2009). When the whole architecture is set up, the computing time with ANNs can still take several days for the classification of a single image rendering it impractical for the application in larger areas.

The second common non-parametric method is the Support Vector Machines (SVM) algorithm which was introduced by Vapnik (1979). During its training, SVMs thrive to define a hyperplane separating the dataset into a discrete predefined number of classes (Mountrakis et al., 2011). This is done by minimizing the probability of misclassifications in the training dataset. The major advantage of SVMs is their need for only a small training dataset in comparison to other parametric and non-parametric methods since. On the other hand, general limitations of this method are the determination of parameters like kernel functions which have a strong effect on the result of the classification (Forkuor, 2014; Tso and Mather, 2009).

Decision Trees use a hierarchical splitting (top-down) mechanism to separate the training dataset into different classes. In general, a decision tree works like an inverted tree composed of a so-called root node at the top, interior nodes and terminal nodes at the bottom (Tso and

Mather, 2009). The root node as well as the interior nodes each represent a decision, e.g. a threshold for a spectral input dataset, while the terminal nodes represent the final classes. In general, decision trees can be constructed manually or in an automatic way. For the manual design of a decision tree considerable user knowledge of the different classes and their characteristics of input features is necessary. This approach can be especially difficult to implement for a high number of classes and when the class boundaries are not distinct (Tso and Mather, 2009). Automatic approaches of decision trees enjoyed an increasing popularity in recent years since they outperform most other traditional classifiers and do not require the comprehensive user knowledge about class characteristics (Tso and Mather, 2009). In these approaches, a training database for each class is gathered and the algorithm automatically determines suitable thresholds for the different nodes. Probably the most frequently used decision tree classifier is the Random Forest (RF) algorithm (Breiman, 2001). This method automatically generates a large number of classification trees. Each tree is determined based on a randomly selected bootstrap sample of the input training dataset and a random sample of features at each node. Finally, each tree casts a vote for the most popular class and a majority vote of all trees determines the final classification result (Forkuor, 2014; Tso and Mather, 2009). Since its development in 2001, RF was used in a high number of studies because it has several advantages over other parametric and non-parametric methods. First of all, it is very user-friendly and easy to learn since it has only a few parameters and does not require extensive tuning (Breiman, 2001). Furthermore, it outperforms almost every other available complex classification algorithm in terms of processing time and it is generally robust to noise and overfitting while maintaining high levels of accuracy (Gislason et al., 2006; Wohlfart et al., 2016b). For these reasons, the random forest algorithm has been selected for the classification process in this work.

Besides the selection of a suitable classification method, it also important to decide if and how spatial patterns in the input dataset can be used to improve the classification result. General approaches in this context are the pixel-based, object-based or sub-pixel based classification. Sub-pixel based approaches are often applied when mixed pixels are frequent in the image meaning when a single pixel is actually composed of more than one target class. Within these approaches, the spectral signal of a pixel is attempted to be unmixed and split into several land cover classes (Lillesand et al., 2004). This approach is a common way of overcoming insufficient spatial resolution by e.g. estimating cover fractions for each pixel. However, an exact spatial delineation of objects is not possible with sub-pixel based approaches. Since a spatio-temporal data fusion is applied in this work to generate time series in a sufficiently high spatial resolution the implementation of a sub-pixel approach is not necessary. Object-based approaches on the other hand combine several neighbouring homogeneous pixels into one object or cluster by a segmentation of the image. Such an

approach can be superior if spatially distinguishable objects of sufficient size are available but the ideal object size has to be found prior to classification. However, this approach is of limited suitability for an application in the study area since West African agriculture is rather small-scale and objects are often difficult to identify. On the other hand, the region's semi-natural vegetation underlies gradual spatial changes rendering a suitable segmentation a difficult task. Thus, the pixel-based approach was chosen in this work which uses the feature space of each pixel separately without incorporating the spatial context of the pixel.

4 Data Basis and Pre-Processing

The following chapter presents the data basis of this thesis, describes how the data was acquired and the steps that were conducted for the pre-processing of the different remote sensing products.

As input for the spatio-temporal data fusion, high spatial resolution Landsat data (30 m) from the sensors TM (Thematic Mapper), ETM+ (Enhanced Thematic Mapper) and OLI (Operational Land Imager) were applied in combination with medium spatial resolution data from the MODIS (Moderate-Resolution Imaging Spectroradiometer) sensors (500 m) aboard the Terra and Aqua satellites. The Landsat data input was acquired as atmospherically corrected surface reflectance data from the USGS Landsat Archive (USGS, 2017b) and further processed. In addition, the MCD43A4 product was chosen from the MODIS portfolio, which contains nadir bidirectional reflectance distribution function (BRDF) adjusted reflectances in a daily temporal resolution for collection 6. In the following, section, the decision for these Landsat and MODIS sensor systems is explained.

Furthermore, besides the above-mentioned satellite datasets, several other data sources were applied and are presented in this chapter.

4.1 Selection of Suitable Satellite Sensors

The general spatio-temporal data fusion techniques are usually developed to be flexible in the use of sensor combinations. This also accounts for the basic ESTARFM algorithm which was applied and further developed in this work. In general, several different sensor systems of low to medium spatial resolution could be combined with a number of different high spatial resolution sensors (compare Table 3.1) which has been done in previous studies. In this work, the sensors TM, ETM+ and OLI of the Landsat satellite series were jointly used with datasets from the MODIS sensor since this combination has several advantages for the purpose of spatio-temporal data fusion and the subsequent derivation of agricultural area maps for multiple years.

First of all, sensors used as input for spatio-temporal data fusion should generally have overlapping spectral bands in order to establish a relationship between the two systems. Since MODIS has 36 bands in the visible and infrared spectrum, it can be combined well with a series of high resolution sensors. The overlap of Landsat bands in the visible and near

infrared spectrum with MODIS is highlighted in Figure 4.1. Although, MODIS generally has narrower bandwidths, the two sensor systems correspond well with each other and give a good flexibility in the selection of input bands for the fusion. In this work, the NDVI was chosen as an input to the data fusion which reduces discrepancies between different sensors as investigated by Jarihani et al. (2014) (compare section 3.4). Thus, an exact match in spectral bands of the two sensor systems is not necessary for the applied methodology. However, preliminary fusion tests were conducted with spectral bands as input to the fusion in order to confirm the findings of Jarihani et al. (2014). In addition, it could be an asset in future studies to use the spectral bands as input for the fusion depending on the specific purpose.
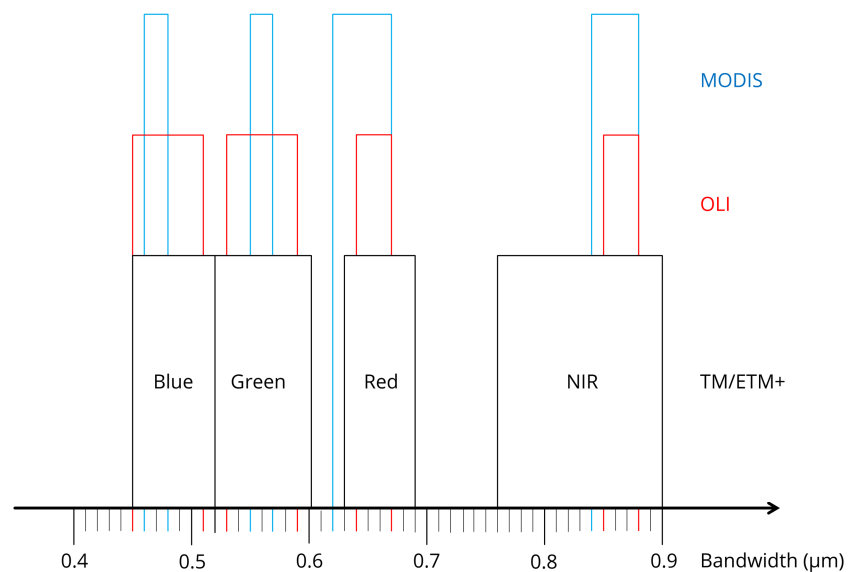
Figure 4.1: Comparison of band widths of selected sensors in the visible and near infrared region of the electromagnetic spectrum.

Another important reason for using the Landsat-MODIS combination is the long overlapping data record of both sensor systems. The MODIS sensors deliver data since the year 2000 (first full annual data record for 2001), while the Landsat time series go back to 1972. Thus, a fusion of time series starting in 2000/2001 is possible which allows the monitoring of agricultural area for one and a half decades up until today. In addition, the geographical coverage of Landsat data is better than for all other high resolution sensors. Despite some gaps in earlier years, Burkina Faso can be completely covered with annual Landsat datasets while data of other high spatial resolution sensors is often only sparsely available.

Lastly, it is important for work in Africa and an important requirement within the WASCAL project that the data and the subsequent developments are available without any cost. Thus, developed methodologies should be based on unrestricted remote sensing data. Data access

to MODIS and Landsat products is free of charge and they are easily downloadable via respective online platforms (NASA, 2017a; USGS, 2017c).

4.2 Landsat TM, ETM+, and OLI

In this work, multi-temporal data from the Landsat sensors TM, ETM+ and OLI were used as high spatial resolution data source for the spatio-temporal data fusion. The Landsat programme is under the lead of NASA (National Aeronautics and Space Administration) and USGS comprising seven successfully launched satellites (Landsat-1 to -5 and Landsat-7 to -8) since 1972 to this day (USGS, 2017d).

The TM sensor in its second generation was installed on the Landsat-5 platform which delivered image data in seven spectral bands and a spatial resolution of 30 m (6 bands) and 120 m (thermal infrared band) between 1984 and 2011. The transmission of TM data originally worked via TDRSS (Tracking Data Relay Satellite System) directly to a central ground receiving antenna in the USA. However, in 1987, this transmitter failed and subsequently, data could only be received and stored by a network of regional International Ground Stations (IGS) (Wulder et al., 2016). This lead to a decentralisation of TM data storage and resulted in the issue that more data was stored outside the USGS Landsat Archive than inside. Thus, the USGS started to consolidate the TM data in 2010 in the framework of the Landsat Global Archive Consolidation (LGAC) initiative with the aim of a single, central archive situated at the Earth Resources Observation and Science (EROS) Center in Sioux Falls, South Dakota (Wulder et al., 2016). Since this effort is still not completed, data for different regions and periods can still be missing in the archive.

The ETM+ sensor was launched in 1999 aboard the Landsat-7 satellite. It generally replicated the capabilities of the TM sensor including an additional panchromatic band with 15 m spatial resolution and an improved spatial resolution (60 m) of the thermal infrared band (NASA, 2017b). In May 2003, the Landsat-7's Scan Line Corrector (SLC) failed which resulted in missing data stripes covering about 25% of the images from that date on. Although, several post-processing efforts have been undertaken to fill these missing data gaps with e.g. data of other dates, the general use of ETM+ datasets after 2003 is limited.

In 2013, the latest satellite of the Landsat programme, Landsat-8 was launched with the OLI (Operational Land Imager) instrument aboard. Its major improvement over the TM sensor is that OLI is a push-broom instead of a whisk-broom sensor. While whisk-broom sensors use a moving mirror to scan a whole line of pixels, push-broom sensors have a series of detectors to scan a whole line of pixels at once. Though, these are more difficult to calibrate, push-broom systems decrease the danger of failure through the reduction of mechanical parts and the received data is also less affected by view angle effects (Knauer et al., 2016; Schmidt et al., 2015). Furthermore, the OLI sensor includes two new bands, one for the improved detection

of cirrus clouds and thus atmospheric correction and one for coastal zone observations (NASA, 2017b).

For the processing of whole Burkina Faso and the years 2001, 2007 and 2014, a total of 504 surface reflectance scenes (Table 4.1) from the Landsat sensors TM, ETM+ and OLI were downloaded from the USGS Landsat Archive (USGS, 2017b). The data covers 17 Landsat tiles (Figure 4.2, tile 194/53 is only included in the focus area of chapter 5) for the years 2001 and 2014 while for 2007 no data were available from the two southwestern tiles (path/row: 192/52, 193/52).

Table 4.1: Overview over the used Landsat datasets.

Study Area	Sensor	No. Tiles	Focus Year	Time Series
Burkina Faso	ETM+	17	2001	Dec. 2000-Jan. 2002
Burkina Faso	TM	15	2007	Dec. 2006-Oct. 2007
Burkina Faso	OLI	17	2014	Dec. 2013-Jan.2015
Focus Area	OLI	4	2014	Dec. 2013-Jan.2015
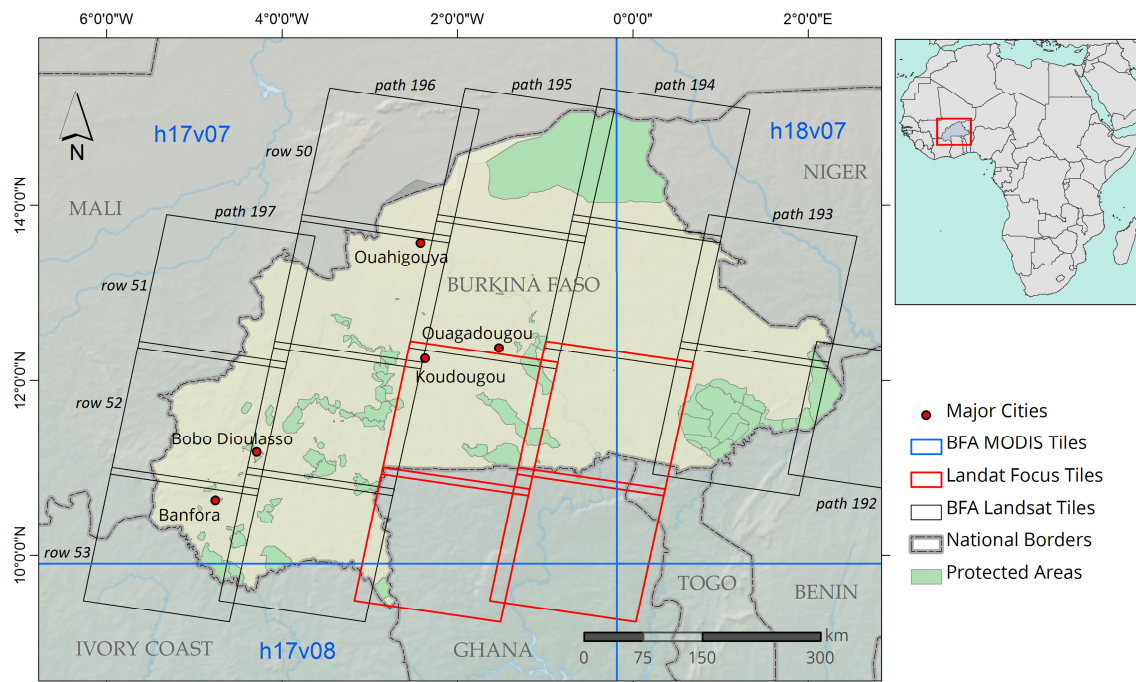
Figure 4.2: Overview over the used MODIS and Landsat tiles; red lines mark the Landsat tiles used for the focus area in chapter 5 (modified after Knauer et al., 2017).

The time series span the whole years 2001 and 2014 plus the respective preceding December and the following January for shoulder scenes in the data fusion process. In 2007, no Landsat data was available after October which also results in a shortened fused output time series. For the focus area of the algorithm development in chapter 5, time series of four southern Landsat tiles (red tiles in Figure 4.2) and the year 2014 were selected due to its spatial heterogeneity of land cover. The two eastern Landsat tiles of this focus area (194/52

and 194/53) were clipped to the extent of the MODIS tiles h17v07 and h17v08 (also compare Figure 5.1).

The datasets are provided in an atmospherically corrected form which is conducted with the L8SR algorithm (USGS, 2015). Additional cloud and cloud shadow masks, which are derived with the CFmask algorithm (USGS, 2015; Zhu and Woodcock, 2012), are available together with the reflectance datasets. Based on these additional datasets, cloud and cloud shadows were masked from the reflectances and the NDVI was calculated afterwards (Figure 4.4). It was decided to exclude datasets with more than 40% cloud cover since a comprehensive inspection indicated only a limited added value of datasets with higher percentages. This threshold however can be defined freely by the user of the ESTARFM framework and has to be found for every study region and time series separately (Knauer et al., 2016). Figure 4.3 shows the distribution of the obtained Landsat scenes over the year 2014. It is apparent that no scenes fulfilling the above mentioned 40%-threshold could be acquired between July and October which also corresponds to the major part of the rainy season in the focus area.
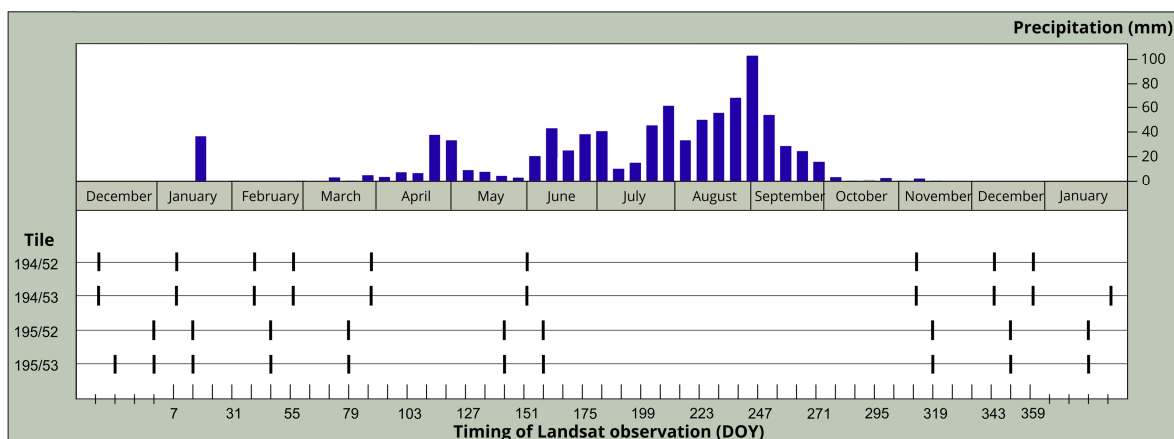
Figure 4.3: Overview of the distribution of the four Landsat tiles from the focus area over the year 2014 (December 2013 to January 2015) and the corresponding weekly rainfall sums averaged from three rainfall stations in the focus area (Bliefernicht et al., 2013; Quansah et al., 2015); figure modified after Knauer et al. (2016).

4.3 Terra/Aqua MODIS

The first MODIS sensor was launched with the Terra satellite (originally known as EOS AM-1) in December 1999. In 2002, a second MODIS sensor followed aboard the Aqua satellite (originally known as EOS PM-1; NASA 2017b). These two missions are an important part of NASAs EOS (Earth Observation System) programme and became a big success in the history of satellite remote sensing with frequent use in the literature (NASA, 2017c). The sensors cover a swath width of 2,330 km with 36 spectral bands between 0.46 µm and 14.39 µm with varying spatial resolutions of 250 m, 500 m and 1,000 m (NASA, 2017c). Due to its relatively high spatial resolution of up to 250 m for the red and NIR bands while delivering a daily temporal

coverage, MODIS enjoys great popularity for the monitoring of vegetation. Furthermore, a suite of high quality standard products like different biophysical parameters and vegetation indices (e.g. LAI - Leaf Area Index, FPAR - Fraction of absorbed Photosynthetically Active Radiation or NDVI) is derived from the MODIS sensor and provided for cost-free download (NASA LP DAAC, 2013).

In this work, the MODIS product MCD43A4 (collection 6) is used as the medium spatial resolution time series for the data fusion. This combined product of the MODIS sensors aboard the Terra and Aqua satellites contains 500 m nadir BRDF (Bidirectional Reflectance Distribution Function) adjusted reflectances which should improve correspondence to the Landsat reflectances due to decreased view angle effects (Fensholt et al., 2010; Walker et al., 2012). An important enhancement of the collection 6 over the previous collection is the higher production frequency which has been increased from 8 day intervals to a daily coverage. This high frequency basically guarantees the availability of MODIS data at each Landsat observation date and allows maximum flexibility in the definition of the temporal intervals of the fused output time series. The MCD43A4 product is generated as a rolling 16-day composite from which the best possible BRDF is derived. The date associated with the respective daily dataset corresponds to the centre of the moving 16-day window while it was the first day of this period for the previous collection 5. Further information on the most recent collection of the MCD43A4 product are found in the MODIS User Guide (Knauer et al., 2016; Professor Crystal Schaaf's Lab, 2016)

Corresponding to the temporal coverage of the Landsat time series (Table 4.1), datasets of the MODIS product MCD43A4 have been obtained for three MODIS tiles (h17v07, h17v08 and h18v07, Figure 4.2). The red and NIR bands of the MODIS datasets were used to calculate the NDVI which was subsequently processed with the software package TIMESAT (Eklundh and Jönsson, 2015) (Figure 4.4). NDVI outliers, deviating more than 0.2 times from the standard deviation in a moving window, were removed from the time series. The half width of the moving window is defined in TIMESAT as the number of values per year (here 365) divided by 7. The resulting gaps were linearly interpolated and residual noise in the time series was smoothed with the Savitzky-Golay filter (Savitzky and Golay, 1964) in TIMESAT. The window size of the filter was set to 30 days and the adaptation strength to a value of 2. All parameters of the TIMESAT software were determined on the basis of visual analyses investigating the modification of these values in the graphical user interface of TIMESAT (Eklundh and Jönsson, 2015). Further information on the application of the TIMESAT software package can be found in (Eklundh and Jönsson, 2015). After the smoothing of the time series, the MODIS data was reprojected to the same projection as the Landsat data (UTM 30N), resampled to the 30 m Landsat spatial resolution with the nearest neighbour method and in the same step automatically co-registered (Knauer et al., 2016).
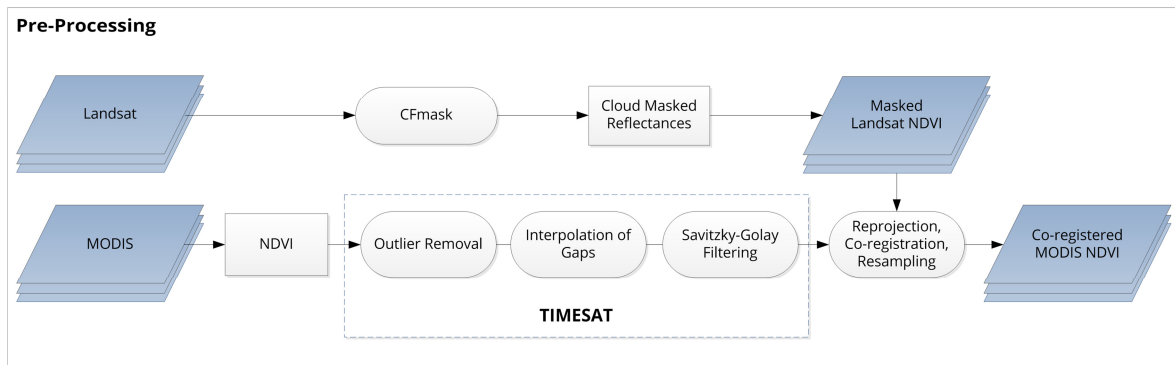
Figure 4.4: Workflow of the Landsat and MODIS pre-processing steps.

For the MODIS test dataset of the algorithm development (chapter 5), a different order of processing steps was originally applied. Here, the resampling and reprojection of MODIS data was conducted with the MODIS Reprojection Tool (MRT) prior to the filtering with TIMESAT (LP DAAC, 2015). However, since this processing order is more computation and storage intensive, the resampling and reprojection was processed with a self-developed code in IDL for the Burkina Faso wide processing. Furthermore, in the original MODIS test dataset, only data with a full BRDF inversion (labelled with the best quality flag) have been used and filtered. Although, this approach yielded good results, the derived time series may contain considerable data gaps for some periods and areas. Thus, it was discarded for the Burkina Faso wide processing so that all data was used but the filtering was adapted (as described above) to account for this difference in input data.

4.4 Reference Data

For the purpose of a high quality remote sensing based LULC classification, a suitable database of reference information for the training and validation of the classification model is essential. As for most African countries, such information is not comprehensively available for the entire area of Burkina Faso. However, in the framework of the WASCAL project, several field surveys have been conducted by the author and other project members for the collection of ground sampling points. These trips did not cover the entire country of Burkina Faso or all focus years but they helped building a broad understanding of the prevailing land cover, the land use and practices of the inhabitants and the general issues in the development and cultivation of the country. Based on the gathered knowledge of the region, very high resolution (VHR) imagery was visually analyzed in Google Earth for the generation of a comprehensive reference database for training and validation of the classifier. This is a common approach for remote sensing studies on larger scales and in emerging and developing countries in general where accurate ground truth information is missing and has thus been conducted by numerous authors before (Clauss et al., 2016; Gessner et al., 2015b; Knauer et al., 2017; Wohlfart et al., 2016b)

Eight common thematic LULC classes (Table 4.2) focusing on agriculture were defined for the collection of reference points based on the gathered expert knowledge and the description of existing regional LULC classifications (FAO, 2017; Gessner et al., 2015b).

Table 4.2: LULC classes for the derivation of agricultural area in Burkina Faso (modified after Knauer et al., 2017).

ID	Class name	Description
1	Rainfed Agricultural Area	Farmed land with partial tree cover (< 10%, potentially fruit trees) including annual crops, fallow and intensive pasture
2	Irrigated Agricultural Area	Croplands irrigated from water bodies (potentially 2 seasons)
3	Grassland	Herbaceous vegetation with less than 10% woody vegetation
4	Woody Vegetation (deciduous)	Areas with deciduous woody vegetation coverage of at least 10%
5	Plantation (evergreen)	Areas of evergreen broadleaved tree crops (mostly fruit & nut trees)
6	Water Bodies	Areas permanently covered with water
7	Temporarily Flooded	Temporarily flooded herbaceous or woody vegetation
8	Forest (evergreen)	Areas of dense evergreen broadleaved trees

In order to ensure that the distribution of sampling points is unbiased, the *Simple Random Sampling* approach was applied (Congalton and Green, 2009; Wohlfart et al., 2016a). For this purpose, an initial number of 700 points was generated, randomly spread over entire Burkina Faso. Each of these points was subsequently approached in Google Earth and the closest spatially homogeneous area with an approximate minimum size of nine Landsat pixels was delineated with a polygon. All VHR images were then analyzed for this polygon and the LULC classes for the three focus years (2001, 2007 and 2014) were deduced if sufficient information was available. The aim was to collect at least 100 polygons for phenologically dynamic classes such as rainfed agricultural area and a minimum number of 50 polygons for classes like water which are temporally more stable. These numbers were suggest by Congalton and Green (2009). One issue of the applied *Simple Random Sampling* approach is that fewer sampling polygons are obtained for less frequent LULC classes. Thus, in order to fulfill the above described minimum numbers for each class additional sampling had to be conducted. For this purpose, known areas of the respective under-sampled classes were targeted directly in Google Earth and classified. This approach might introduce small biases in the distribution of ground truth data, but it is essential for the collection of sufficient amounts of polygons. However, since this was only conducted for three classes (irrigated agricultural area, plantation and temporarily flooded) and a low number of polygons, the effect should be minor. The complete reference database for the training and validation of the classifier included 771 polygons for the year 2001, 836 polygons for 2007 and 1147 polygons for 2014 (Figure 4.5; Knauer et al., 2017).
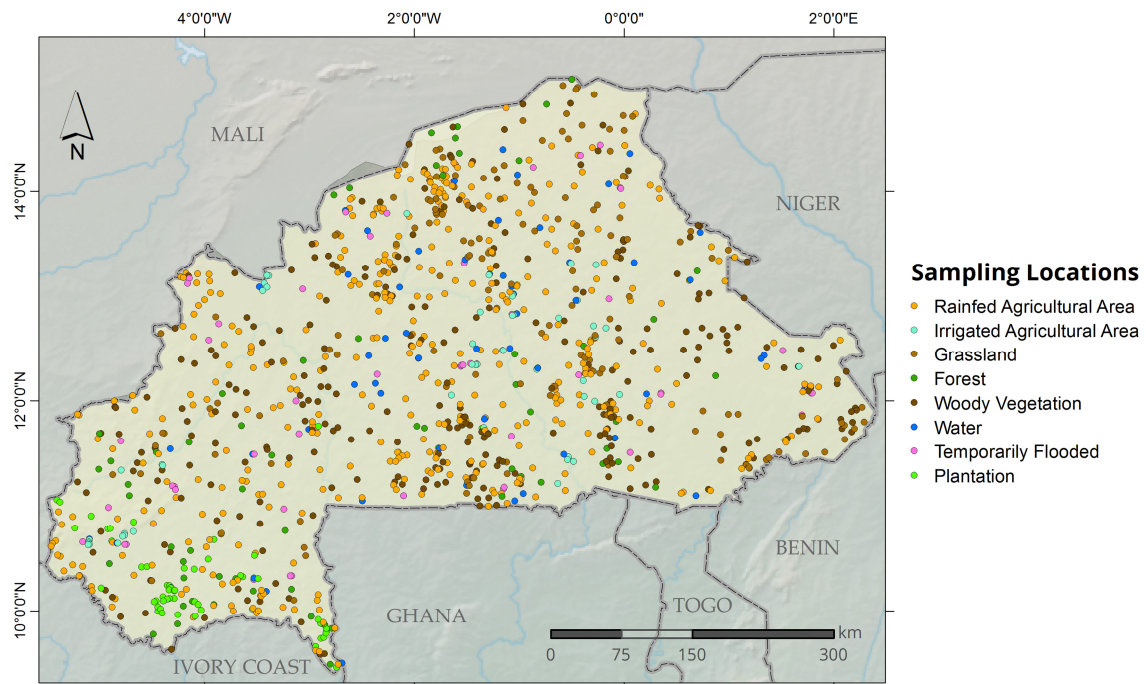
Figure 4.5: Reference points sampled for Burkina Faso.

4.5 Additional Datasets

For the analysis and interpretation of results, several auxiliary datasets were used in this work. In addition to the extent of MODIS and Landsat tiles, Figure 4.2 shows the location of formally protected areas in Burkina Faso. This dataset was obtained from the the Institut Géographique du Burkina Faso (IGB) and outlines the country's natural reserves representing the status of 2005 (IGB, 2016). These reserves are generally on different levels of protection and management from formally protected game ranches to strictly managed national parks. This dataset was used for the analysis of influence of agricultural area expansion on natural reserves in Burkina Faso (section 6.4.4). In addition, census data was downloaded from the Institut National de la Statistique et de la Démographie (INSD) of Burkina Faso and the included provincial rural population numbers (Figure 2.6) were used to analyse the relationship between population growth and agricultural expansion (section 6.4.3) (INSD, 2014).

In order to estimate the quality of derived statistics on agricultural area in section 6.5.1, datasets from FAOSTAT were downloaded for comparison (FAO, 2016). These datasets contain estimations of agricultural area and harvested area of primary crops on a national level for the years 2001 and 2013. Statistics for the year 2014 were so far not available in the FAO database. Generally, the included numbers especially the ones on agricultural area have to be interpreted with caution. In this variable, a constant number of 60,000 km² of 'permanent meadows and pastures' is included for the estimates of past decades. Personal

communication with FAO staff (Giorgia DeSantis 2016, pers.comm., 21 December) confirmed that this number is referring to an estimation of the year 1980 and thus highly outdated.

5 Development of a Novel Fusion Framework

Mapping and monitoring the spatio-temporal development of agricultural area and vegetation in general is a challenging task in cloud-prone and heterogeneous landscapes like West Africa (Knauer et al., 2014). Increasing the availability of consistent remote sensing time series in high temporal and spatial resolution could not only improve analyses of agricultural expansion but also various other studies on topics like land degradation, impact of climate change on vegetation and land surface phenology (LSP) or changes in vegetation composition. Anyway, until now, no single satellite sensor is capable to provide such time series for the application in challenging study areas like West Africa. While medium spatial resolution sensors like MODIS can provide sufficient information on the temporal development of vegetation, they lack the spatial detail to capture the small-scale mosaic of different land cover types as present in the study region. On the other hand, higher spatial resolution sensors such as aboard the Landsat satellites can provide a sufficient geometric resolution but they generally lack a suitable temporal resolution. This is not primarily due to the theoretic repeat cycle of Landsat (16 days), but rather due to the frequent cloud coverage in the region considerably decreasing actual data availability especially during the highly important growing season. In addition, the earlier Landsat missions experienced different system failures which further hamper data availability of these sensors (compare section 4.2). The high spatial resolution sensors aboard the two Sentinel-2 satellites will also increase the temporal resolution and thus data availability. However, since the series has just been completed in March 2017 with Sentinel-2B, this system does not allow for a consistent analysis of past years. In order to overcome the trade-offs between spatial and temporal resolution, the fusion of datasets from MODIS and Landsat has gained considerable popularity in recent years. The review of methods and their applications in section 3.4, especially of the ESTARFM algorithm shows the benefits of this approach to improve resolutions. However, the uneven global distribution of studies as well as their small focus areas also indicate that the methodology still comprehends limitations that so far hinder the application on larger scales and in some regions (Knauer et al., 2016).

In the following chapter, major drawbacks of the ESTARFM algorithm are identified and methodologies are further improved for a better applicability in West Africa and on larger

scales in general. A major step towards this is the development of a framework for ESTARFM facilitating and automating the time series generation as well as automatically accounting for regional differences in the heterogeneity of the focus area. Since the frequent cloud coverage in West Africa, especially during the rainy season, is substantially decreasing the number of cloud-free Landsat scenes, improving the usage of cloud-affected input scenes is also a major aim in the development of a framework (Knauer et al., 2016).

The implemented ESTARFM framework was tested for the generation of time series in a focus area in the border region between Burkina Faso and Ghana (section 5.1) in order to test and demonstrate its benefits for the application in the West Africa. The development of the ESTARFM framework was published within the WASCAL project as Knauer et al. (2016). This chapter partially contains descriptions and results of this publication.

5.1 Focus Area

The focus area of about 98,000 km² is located in the border region between southern Burkina Faso and northern Ghana (3°10'33"W - 0°11'31"W / 9°29'4"N - 12°24'9"N) (Figure 5.1). It is characterized by a tropical climate with strong rainy and dry seasons. Almost all of the annual rainfall, which ranges between 700 and 1200 mm, occurs between May and October. The monthly average temperatures are on a rather constant level between 25 and 32 °C. The focus area is part of the West Sudanian Savanna ecoregion, which is characterized by vegetation types with different densities of woody vegetation cover from open grassland, over closed woodland to dense forests (Olson et al., 2001). Higher percentages of woody vegetation cover are primarily found along streams or in the protected areas of the region (Kaboré-Tambi National Park, Mole National Park, Sissili classified forest, and Nazinga game ranch). The major part of the study area is under agricultural use, which has spread significantly during recent years: Gessner et al. (2015) identified an increase of agricultural area from 30 to 60% (2001-2013) in the vicinity of the Kaboré-Tambi National Park. The farming practices in the focus area are primarily small-scale rainfed agriculture with small field sizes (around 1ha) and pastoral agriculture (FAO, 2016; Forkuor, 2014). The productivity of rainfed agriculture is still on a very low level with about 1.15 t per ha and year cereals in Burkina Faso (FAO, 2016). The selected focus area is spatially highly heterogeneous since the mixture of small croplands and savanna vegetation forms a small-scale mosaic of different LULC types (Forkuor, 2014). An important landmark of this focus area is the Lake Bagré in the northeast. This reservoir was created by damming the White Volta River and is used for hydro power as well as irrigation of croplands downstream of the dam.

All in all, the focus area is very challenging for remote sensing, since it is heterogeneous in space with its small-scale agriculture and highly dynamic in time with strong rainy and dry seasons resulting in a distinct phenological change of vegetation throughout the year.
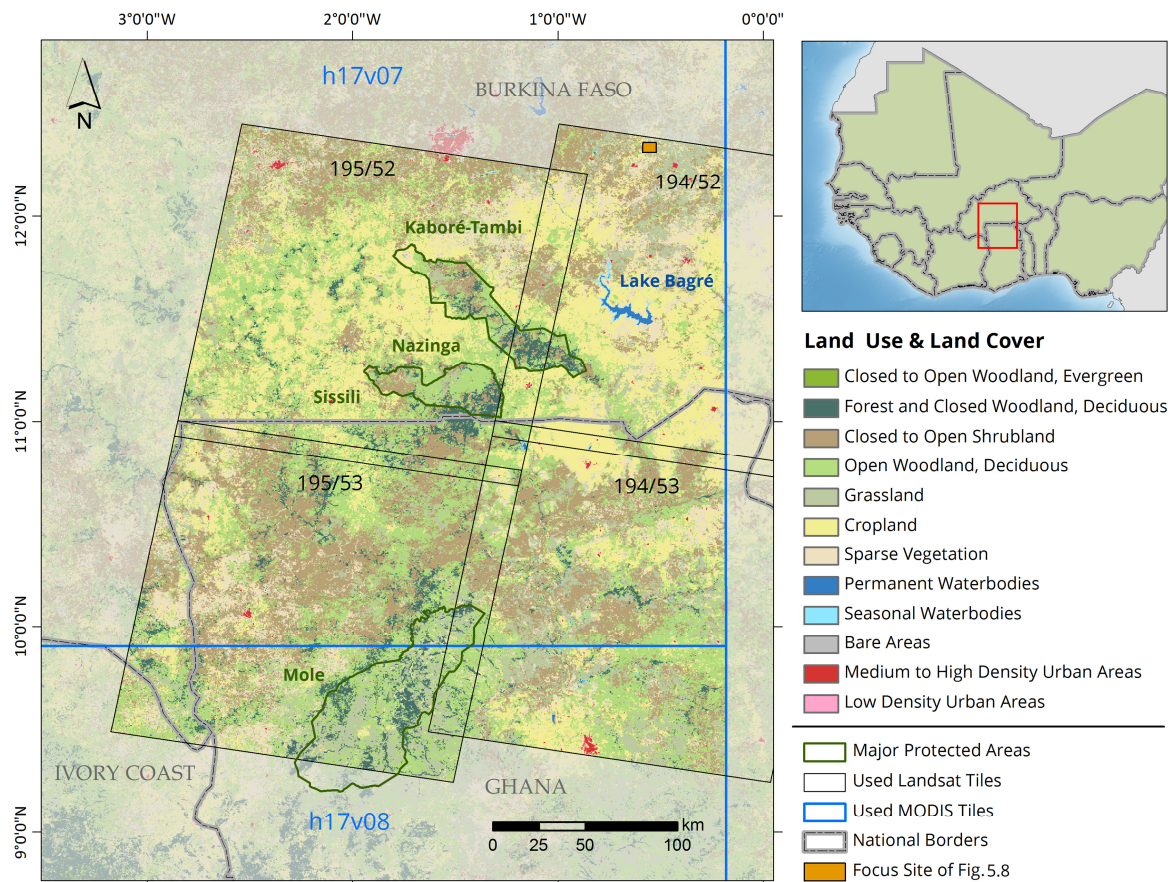
Figure 5.1: Overview over the focus area in the transboundary region between Burkina Faso and Ghana, outlined by the intersection of two MODIS tiles (h17v07, h17v08) and four Landsat-8 tiles (194/52, 194/53, 195/52, 195/53); In the background: regionally optimized land cover classification from the WASCAL project (Gessner et al., 2015b) with 250 m spatial resolution for 2006 (modified after Knauer et al., 2016).

5.2 Data Fusion for High Temporal Resolution Time Series

5.2.1 Original ESTARFM Algorithm

The ESTARFM fusion algorithm (Zhu et al., 2010) was used as a basis for the generation of high temporal and spatial resolution time series. Several adjustments were made to the original algorithm in order to enable the application for large areas, for heterogeneous landscapes and areas with frequent cloud cover. The newly developed ESTARFM framework was designed to minimize user input and to improve the automation of the processing.

The original ESTARFM generates synthetic Landsat-like images at any date where MODIS data is available. For the production of each synthetic scene, the algorithm uses a pair of Landsat and MODIS images before the date of prediction, hereafter called 'left shoulder', and one Landsat-MODIS pair after this date, called 'right shoulder'. For each Landsat pixel, ESTARFM establishes a relationship between Landsat and MODIS on both image pairs using

69

linear regression. For this approach, a moving window is used to collect pixels that are spectrally similar to the central pixel (Zhu et al., 2010). This similarity is defined by a threshold:

$$|R_s - R_c| \leq \sigma(Band) \times 2 / m \tag{5.1}$$

where R_s is reflectance of the candidate as similar pixel, R_c is the reflectance of the moving windows central pixel, $\sigma(Band)$ is the standard deviation of the whole band's reflectance and m is the number of LULC classes existing on the whole satellite scene. This number has to be estimated by the user.

The subsequently derived set of pixels is then used to fit the regression for the central pixel of the moving window. The slope of this regression is applied as a conversion coefficient between the central Landsat pixel and its corresponding MODIS pixel. On the basis of the conversion coefficient, the reflectance changes between the MODIS scenes at the pair dates versus the MODIS scene at the prediction date can be converted into reflectance change between the Landsat scenes at the pair dates and the new target Landsat scene at prediction date. The whole algorithm is described in detail in Zhu et al. (2010).

Even though ESTARFM is a highly sophisticated algorithm, it has some limitations hindering automated processing and the use for large areas (Zhu et al., 2010). Thus, the ESTARFM was further developed in this study towards the ESTARFM framework to overcome these limitations as outlined in the following sections.

5.2.2 Improved Selection of Similar Pixels

The above described method to select similar pixels is one limitation of the original ESTARFM. The selection of these pixels follows Equation (5.1) including the estimation of the number of LULC classes (m) of the satellite image. This number shall be increased by the user for images with a heterogeneous landscape in order to have a stricter threshold for the similar pixel selection (Zhu et al., 2010). This however leads to an undesired subjectivity in the selection process, since the user has to estimate such a number from visual inspection and expert knowledge. Furthermore, the use of a single number of LULC classes for the entire image hinders the consideration of spatial variations in the heterogeneity of the image or study region. If a study area is homogeneous in one part, i.e. has only a few LULC types and is rather heterogeneous in another part, e.g. having small-scale patterns of multiple LULC types, the original algorithm could not account for this and only one overall number of LULC classes would be applied. This issue is particularly substantial when the algorithm shall be applied for large areas as in the presented thesis. In order to overcome this, the developed ESTARFM framework follows a modified approach for the selection of similar pixels. In the applied method, the surface reflectance values of selected Landsat scenes are automatically clustered

on the basis of the ISODATA algorithm (Tou and Gonzalez, 1974). This general idea was already contemplated in the methodological article of the original ESTARFM by Zhu et al. (Zhu et al., 2010). However, it was not implemented in the original algorithm and no details of how to incorporate such a modification were worked out. The clustering in the ESTARFM framework is performed for every year of the time series and if available on the basis of multi-temporal Landsat scenes with several spectral bands. The acquisition dates of these Landsat scenes are selected by the user and should represent the most important phenological stages of the study area's vegetation. On the basis of this clustered image, the similar pixels are selected by the application of two conditions, defined in Equations (5.2) and (5.3). A pixel is selected as a similar pixel if it is in the same cluster as the central pixel of the moving window (Equation (5.2)) and the difference in reflectance between the candidate pixel and the central pixel has to be below or equal to one standard deviation of the corresponding cluster in the moving window (Equation (5.3)),

$$C_s = C_c \tag{5.2}$$

$$|R_s - R_c| \leq \sigma(C_{cWin}) \tag{5.3}$$

where C_s is the cluster the similar candidate pixel belongs to, C_c is the central pixel's cluster and $\sigma(C_{cWin})$ is the cluster's standard deviation of reflectances in the moving window. As a substitute for the clustering approach, a land cover dataset in the same spatial resolution could be used if available.

5.2.3 Incorporation of Partially Cloud Covered Landsat Scenes

The authors of the original ESTARFM algorithm (Zhu et al., 2010) suggested to use high spatial resolution scenes that are temporally close to the date of prediction. Thus, for the generation of a whole time series, a large number of input scenes would be necessary. It is expected, that high resolution scenes, which are temporally not close enough to the prediction date or even from another vegetative season, would decrease the quality of the prediction (Zhu et al., 2010). In earlier studies, the input data for the ESTARFM algorithm was either limited to cloud-free high resolution images or partially cloud covered images were used which then resulted in gaps in the synthetic output scenes (Bisquert et al., 2014; Tewes et al., 2015). Thus, in regions with frequent cloud cover, the application of ESTARFM is hindered since data availability is substantially reduced.

In this work, an automated filling of gaps in the high resolution scenes was implemented in order to be able to use Landsat images, even if they are partially cloud covered but without

reproducing these gaps on the synthetic output images. On the basis of the other modifications in the ESTARFM framework, the gaps are filled with fused images as described in the following (Figure 5.2).
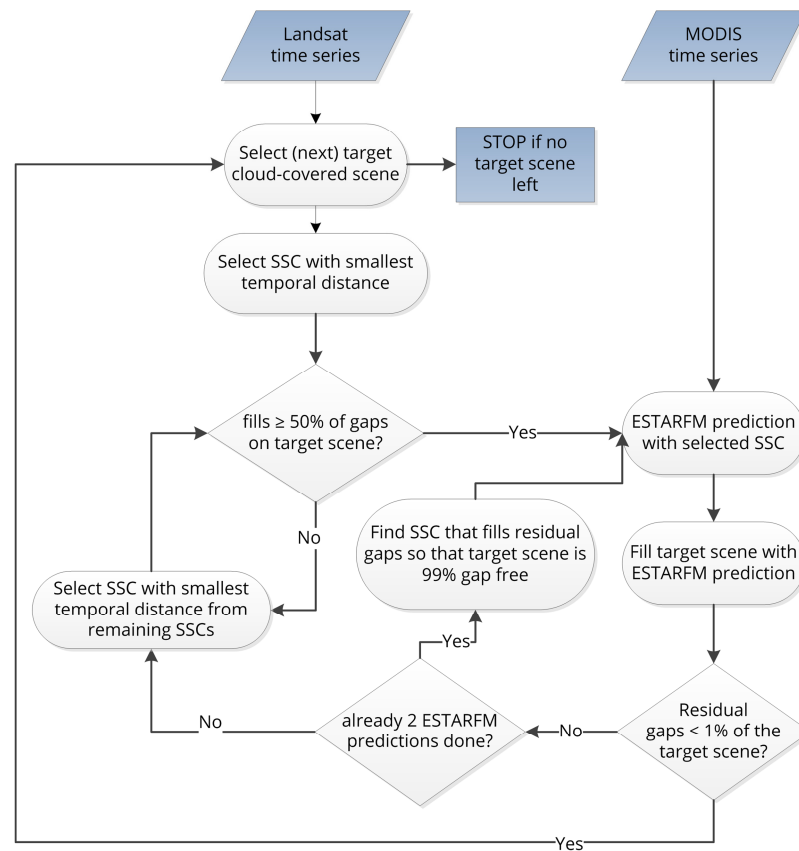
Figure 5.2: Automatic cloud-gap filling procedure for Landsat scenes. The gaps in the target scene are filled with ESTARFM predictions. The shoulder scene combinations (SSC) for this purpose are automatically selected after the ruleset described in the flowchart (modified after Knauer et al., 2016).

The gap-filling procedure starts with the automatic selection of the first cloud covered target scene. Then, the two shouldering Landsat scenes meaning the Landsat image before and the one after the target scene are determined (hereafter called 'shoulder scene combination', SSC) which fulfill two conditions: (1) the temporal distance between the images of the SSC has to be the smallest possible while (2) filling at least 50% of the target scene's gaps with the SSC prediction. After a suitable SSC is identified it is applied for the generation of an ESTARFM prediction which then fills the gaps of the target scene. If the remaining gaps cover less than one percent of the scene, the algorithm moves on to the processing of the next cloud-covered scene. If the gaps are still not below the threshold of one percent, the above procedure is repeated on the other possible SSCs. If the filling with a second prediction can still not decrease the residual gaps below one percent, the SSC is identified that directly fulfills the one percent-threshold. On the basis of this ruleset, not more than three predictions

have to be generated for the filling of one target scene and the maximum amount of processing time is limited.

5.2.4 Quality Layers

In the original ESTARFM algorithm, a possibility to assess the quality of the fusion is not included. However, knowing if and where a predicted image is based on stable regressions or not can indicate regions with potential prediction errors. Thus, the output of the ESTARFM framework has a set of quality layers included. One supplementary layer that comes with each predicted image contains the number of similar pixels used for the calculation of the conversion coefficient for each pixel. In addition for each input band, a layer with the coefficients of determination derived from the linear regressions between MODIS and Landsat is included. These additional information indicate the stability and quality of the regression and are thus a measure of the general well-functioning of the image fusion; they furthermore enable the analysis of regional differences in prediction quality.

5.2.5 Automation and Acceleration of the Processing

Since the original ESTARFM code was programmed for the prediction of single scenes, a generation of a whole time series would need manual input for every single time step. For this reason, an automatic framework was designed in IDL within this work. After the specification of all available MODIS and Landsat scenes, it generates a whole synthetic time series between a defined start and stop date and in a specified time step without any further input. The temporally closest shoulder scenes of MODIS and Landsat as well as the MODIS scene at the prediction date are chosen automatically for each fusion.

In addition, the processing speed was accelerated significantly so that the generation of long time series or larger areas can be accomplished in a reasonable amount of time. Most of the original ESTARFM functions only work with one Central Processing Unit (CPU) core which results in prediction times of more than 24 hours for a single Landsat scene with four bands (Table 5.1). In this work, the image fusion has been parallelized with the possibility to choose the number of CPU cores applied for the processing. The input images are split into tiles equal to the number of specified CPU cores and each tile is assigned to one core. Based on this scheme, the processing can be optimized for the user's individual system specifications and the time series generation is automatized and accelerated considerably (Table 5.1).

Table 5.1: Summary of processing times of original ESTARFM and the ESTARFM framework on the basis of nine CPU cores. The tests were conducted with four reflectance bands and a single layer of NDVI (modified after Knauer et al., 2016).

Input	Original ESTARFM	ESTARFM framework (9 cores)
4 bands	24 h 49 min	4 h 38 min
NDVI	15 h 32 min	1 h 45 min

5.2.6 Use Case of the ESTARFM Framework in the Focus Area

On the basis of the daily MCD43A4 MODIS product and the ESTARFM framework, the dates and time step of the synthetic output time series can be flexibly defined. Thus, a daily ESTARFM time series could be generated which would however implicate considerable processing times despite the acceleration of the algorithm and a high amount of data storage requirement. Since a lower temporal resolution is sufficient for various applications, an 8-day interval was chosen for the use case. The ESTARFM framework was applied to generate a 30 m NDVI time series for the year 2014 between day of the year (DOY) 7 and 359. For this purpose, the MODIS and Landsat datasets as described in sections 4.2 and 4.3 was used. Preceding studies determined that the use of NDVI as input to the fusion results in higher accuracies than using reflectances and deriving the NDVI from the output of the prediction, since the error propagation is reduced this way (Jarihani et al., 2014; Tian et al., 2013).
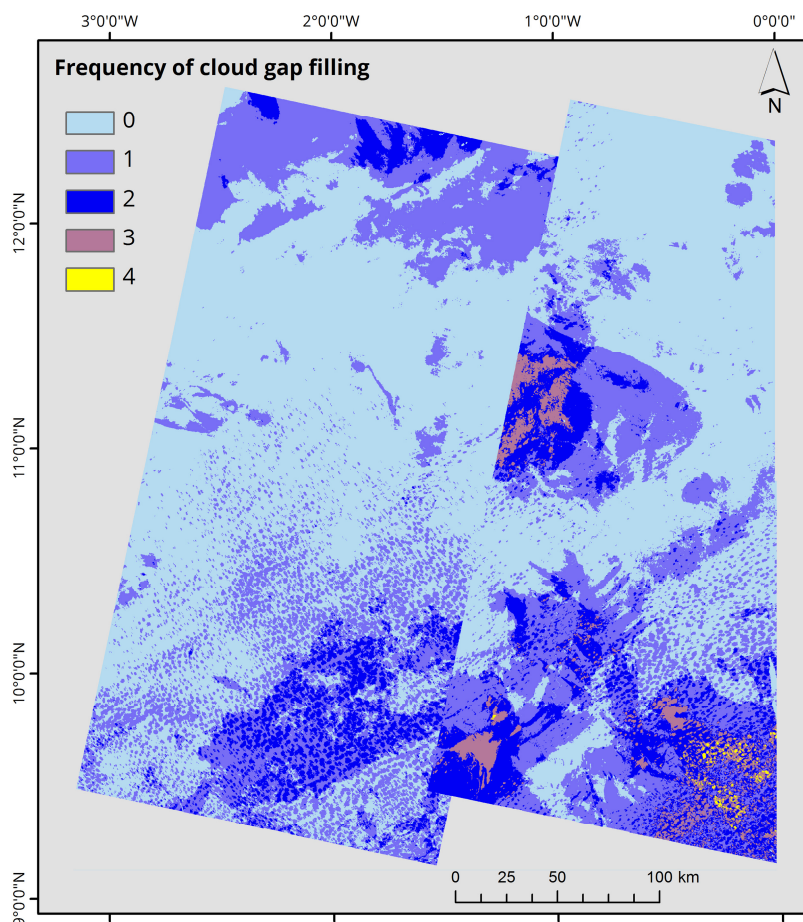

Figure 5.3: Map of the four Landsat tiles showing the frequency of gap filling for each pixel in the input time series of 2014 (modified after Knauer et al., 2016).

Out of the 30 Landsat scenes for this use case, 17 had data gaps due to cloud coverage but these could be filled with the novel feature of the ESTARFM framework (section 5.2.3) and

74

subsequently used as input for the time series generation. For the vegetative season of the focus area, only a single gap-free Landsat scene with cloud cover of less than 1% could be obtained for path 195. On Landsat path 194, no cloud-free scene was available during that period but several had a cloud cover below 40% and could be filled with the developed approach (Figure 5.3). Especially the time series of tile 194/53 in the Southeast was frequently cloud covered and some areas had to be filled four times over the course of the year. Without the filling of gaps in advance of the time series generation, i.e. with the original ESTARFM algorithm, these regions would maintain the gaps in the output time series as well.

5.3 Assessment of Prediction Accuracy

Several approaches were used for the accuracy assessment of the fused ESTARFM time series. The first method is comparable to a cross-validation: a Landsat scene was excluded from the input datasets and this scene was then predicted with the remaining images. Subsequently, the predicted scene was correlated with the actual Landsat scene and mean absolute errors (MAE) as well as root mean square errors (RMSE) were calculated. This was done for every Landsat image except for the respective first and last scenes of the time series because these dates cannot be predicted without additional shoulder scenes. This way, the accuracy of each tile's fused time series could be assessed. In order to facilitate this approach, the time steps of the ESTARFM output time series were matched to the 16-day repeat cycle of Landsat path 195 and thus the Landsat scenes of this path could be applied for validation as well (Figure 4.3). Since the eastern path 194 has a slightly different acquisition interval, additional predictions have been produced for this path at the dates of its Landsat scenes. In total, 195 fused images have been generated for the four tiles.

Because the input Landsat time series used for the predictions have a considerable temporal gap during the peak of the rainy season, eight Landsat scenes with up to 80% cloud cover were applied for the validation of this season. Though, the accuracy results of this period may not be representative for the whole focus area because of the dense cloud cover, this approach gives an estimation of the ESTARFM quality during the major part of the rainy season.

In addition, a subset of the focus area was examined over the different vegetative seasons in order to present the spatio-temporal variations of prediction quality and its relationship to land cover types and phenological phases.

Another important quality measure of the ESTARFM predications is the consistency of the fused time series in the overlapping areas of different Landsat paths which was analyzed in this work as well. Closely matching NDVI values in the overlapping areas of the two Landsat paths indicate a good prediction quality and facilitate the further application of the fused time series without any additional post-processing for the alignment of different paths.

5.4 Results

5.4.1 Temporal Development and Spatial Patterns of the ESTARFM Time Series

The ESTARFM time series is able to display the small-scale variations of the focus area very well (Figure 5.4). Areas of high mean NDVI values such as the residual patches of forest in the northwest of the focus area or the gallery forests alongside the rivers, like in the south of Bolgatanga or in the north of Wa, are clearly visible from the results. Cities and their direct vicinity like Tamale or Wa exhibit lower mean NDVI values and stand out as light round patches where they are surrounded by areas of higher NDVI. Particular attention should also be paid to the transition zones between the four different tiles where no edges or visible shifts
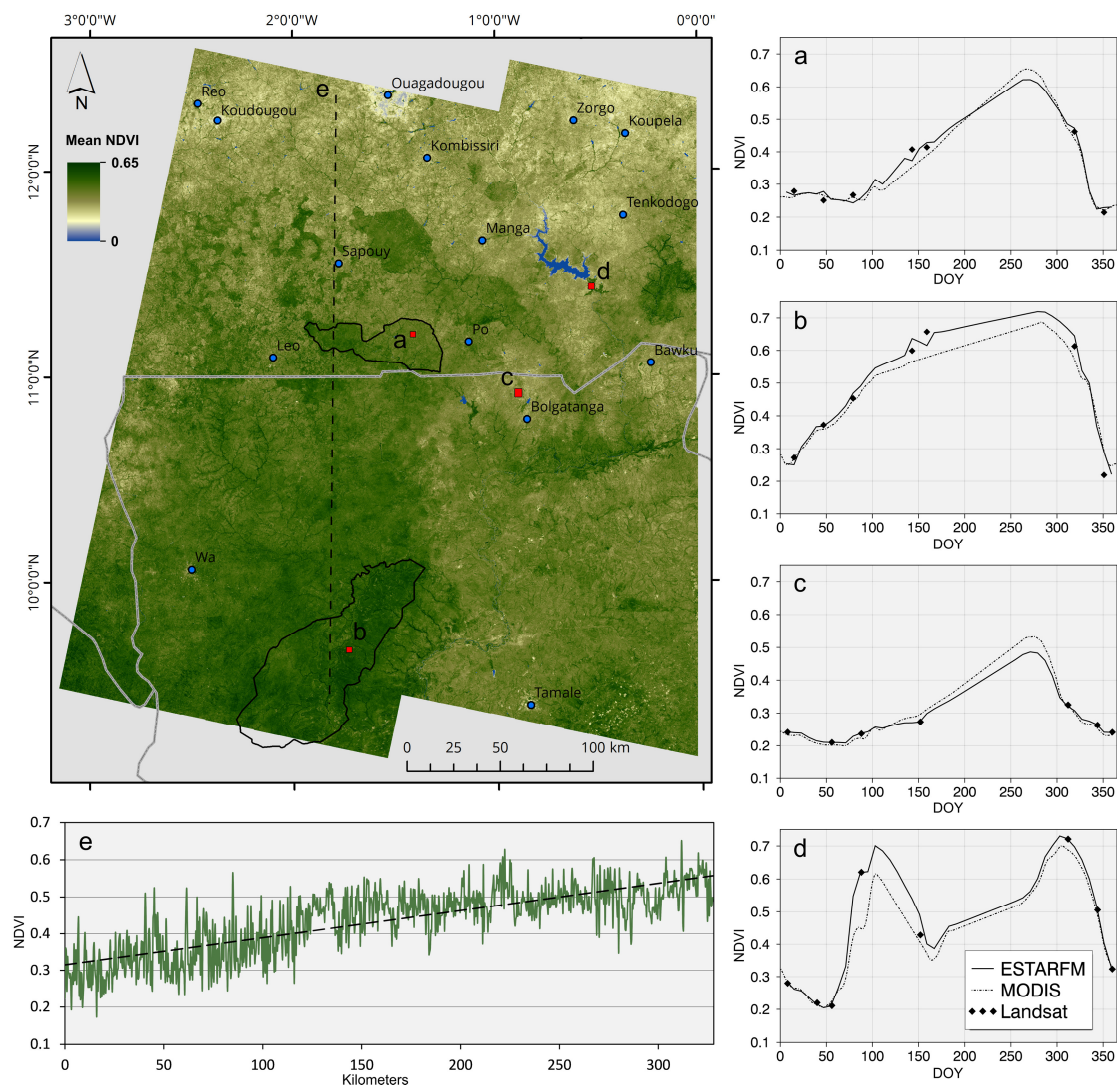
Figure 5.4: Annual average NDVI of the ESTARFM time series for the year 2014 (top left); temporal plots of MODIS, Landsat and ESTARFM time series for four selected focus sites and the year 2014 (red boxes in the map) – (a) a semi-natural savanna site on the Nazinga game ranch, (b) a semi-natural savanna site of high vegetation density in the Mole National Park, (c) a rainfed agricultural area and (d) an irrigated agricultural area; (e) a spatial transect crossing the study area from north to south (dashed line on the map); the dashed line of plot e marks the linear trend (modified after Knauer et al., 2016).

appear. This is an indicator for the well-functioning of the fusion with the improved ESTARFM framework. Such edges between paths would occur if the independent predictions of the different tiles would not match sufficiently (compare section 5.4.2).

The general vegetation gradient in West Africa is well reflected as shown in the transect of Figure 5.4e, which is crossing the study area from north to south. With mean annual rainfalls rising from north to south, the vegetation density expressed by its mean NDVI level increases as well. The temporal development of the three time series of ESTARFM, Landsat and MODIS is analysed in four graphs (Figure 5.4a-d): the first graph (Figure 5.4a) shows the NDVI's average course of the year over a part of the Nazinga game ranch, a semi-natural savanna site. In this area, the time series of ESTARFM and MODIS as well as the single observations from Landsat are well in line with each other. From these results it is also clearly visible that the Landsat time series misses some LSP features such as the peak of NDVI. Thus, the sole use of Landsat data would not be sufficient for phenological analyses in the focus area.

The focus area in the Mole National Park (Figure 5.4b) generally exhibits higher mean NDVI values than the Nazinga game ranch (Figure 5.4a) over the course of the year. It is also characterized by semi-natural savanna but its vegetation cover is denser and contains a higher percentage of trees. The differences between these two sites also reflect the north-south gradient of the overall focus area highlighted in the transect of Figure 5.4e. During the growing season, the Mole NDVI levels of MODIS and Landsat deviate from each other with higher values for the Landsat time series. If this is the case, the ESTARFM algorithm prioritises on the higher spatial resolution data of Landsat and the fused time series subsequently aligns better to these values.

In Figure 5.4c, the temporal development of rainfed agricultural area in the vicinity of Bolgatanga, Ghana is presented. In comparison to the plots of the two savanna sites, a lower mean NDVI level is apparent and the NDVI increase during the vegetative season (DOY 100/mid-April to DOY 320/mid-November) is not as strong as for these sites. This is related to the common agricultural practices resulting in a later green-up than for the semi-natural vegetation and a lower peak of NDVI (Knauer et al., 2016). The last plot (Figure 5.4d) displays the time series of an irrigated agricultural area at the Bagré dam in the south of Burkina Faso. Since this site is mostly independent of the timing and length of the rainy season, its course of the year follows a completely different behavior. The two visible NDVI peaks in mid-April (ca. DOY 100) and at the end of October (ca. DOY 300) reflect the two cropping cycles of this agricultural area.

Considerable differences can be observed by comparing the spatial patterns of mean annual NDVI from ESTARFM, MODIS and Landsat for three focus sites (Figure 5.5). The MODIS time series (second row) with its 500 m resolution is not able to depict the small-scale spatial features of this West African landscape. From the MODIS mean NDVI, neither the mosaic of agricultural areas and savanna in the rural focus site (first column) nor the irrigated croplands
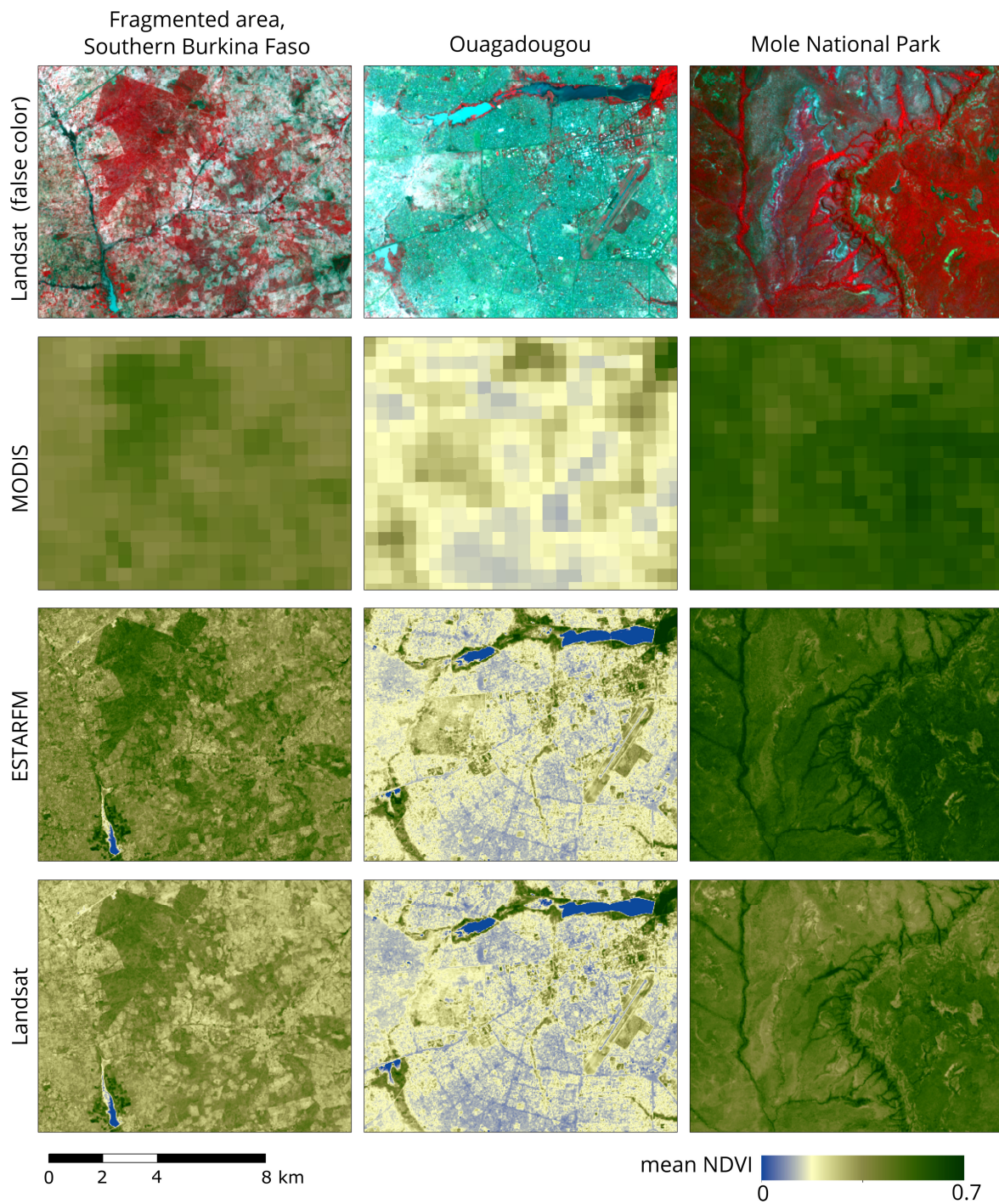
Figure 5.5: Overview of differences in mean annual NDVI for the year 2014 between MODIS (row 2), ESTARFM (row 3) and Landsat (row 4) time series for three selected focus sites: a highly fragmented landscape in the south of Burkina Faso (left column), Ouagadougou, the capital of Burkina Faso (middle column) and a site from the Mole National Park in northern Ghana (right column) (modified after Knauer et al., 2016).

at the southeastern lake can be distinguished. Further examples are the water bodies of Ouagadougou (second column) or the river courses in the Mole National Park (third column) which cannot be delineated properly at MODIS scale.

Nonetheless, since MODIS has a high temporal resolution delivering data for the whole phenological cycle, the level of mean annual NDVI values should be accurate. In the contrary, the mean NDVI values derived from the Landsat data of 2014 (fourth row) are substantially lower in the focus sites. This results from a lack of suitable scenes due to dense cloud coverage during the main vegetative season (July to October, compare Figure 4.3) where NDVI values are highest. This time of the year is almost not included in the annual mean NDVI of Landsat and values are too low in comparison to the more realistic ones of MODIS. However, on the basis of Landsat's spatial resolution, the above mentioned small-scale landscape features can be distinguished properly for all three focus sites. Finally, the ESTARFM time series (third row) preserves the strengths of both sensors, the high spatial resolution of Landsat, suitable for the delineation of small-scale patterns, and the high temporal resolution of MODIS producing realistic mean NDVI values.

5.4.2 Quality of the ESTARFM Time Series

The initial accuracy assessment of the ESTARFM predictions consists of three accuracy measures (Figure 5.6), the first one is the coefficient of determination (R^2) for the correlation between predicted and actual Landsat image. The second one is the mean absolute error (MAE) between these two images and the third is the root mean square error (RMSE). The north-eastern tile (194/52) yielded the overall highest correlations between predicted and actual NDVI values with an average R^2 of 0.83 for this tile. Average error values of about 0.02 for MAE and an average RMSE value of 0.03 were also the lowest for the tile 194/52. The northern tiles (row 52) in comparison to the southern tiles (row 53) generally produce lower error values and higher accuracies with minimum MAE values at 0.01 for both rows, ranging up to 0.04 MAE for the north and up to 0.07 MAE for the south.
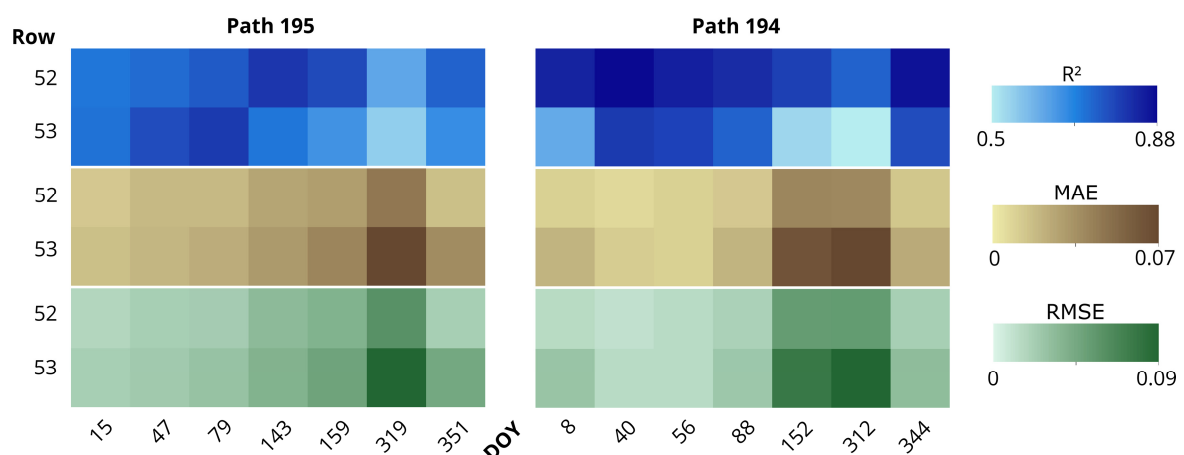

Figure 5.6: Quality and error measures for the ESTARFM time series: coefficient of determination (R^2), mean absolute error (MAE) and root mean square error (RMSE) between observed and predicted NDVI for the four tiles and the dates of 2014 (DOY) (modified after Knauer et al., 2016).

There are not only spatial differences in the quality but also in the temporal development over the course of the year 2014. During the dry season, i.e. the first part of the year until April (DOY 79 for path 195 and DOY 88 for path 194), low error levels can be observed for all tiles. For the subsequent vegetative season, the errors increase with the growing of vegetation (compare Figure 5.4a-d). The highest errors can be observed during the transition from the vegetative season to the next dry season around mid-November (DOY 319 for path 195 and DOY 312 for path 194) with a mean MAE of 0.04 and 0.07 for the paths 195 and 194 respectively. This period is characterized by rapid temporal changes with the browning of leafs, the subsequent decrease of NDVI and the first occurrences of wide-spread fires. During the last dates of the time series in December, the prediction accuracies considerably increase again to the level of the beginning of the year.

The spatial and temporal differences in the quality of prediction are also visible in the scatterplots between observed and predicted NDVI values (Figure 5.7). At the beginning of the year (DOY 40) during the dry season, the range of NDVI is comparatively low and the correlation values are highest with R^2 of 0.88 for row 52 and R^2 of 0.80 for row 53. With increasing NDVI ranges towards the vegetative season (DOY 152), the correlation values decrease with R^2 of 0.79 for row 52 and R^2 of 0.54 for row 53. At the start of the next dry season in the end of 2014 (DOY 344), the correlation values rise again with an R^2 of 0.87 for row 52 and R^2 of 0.77 for row 53 while the range of NDVI values is still high. From these
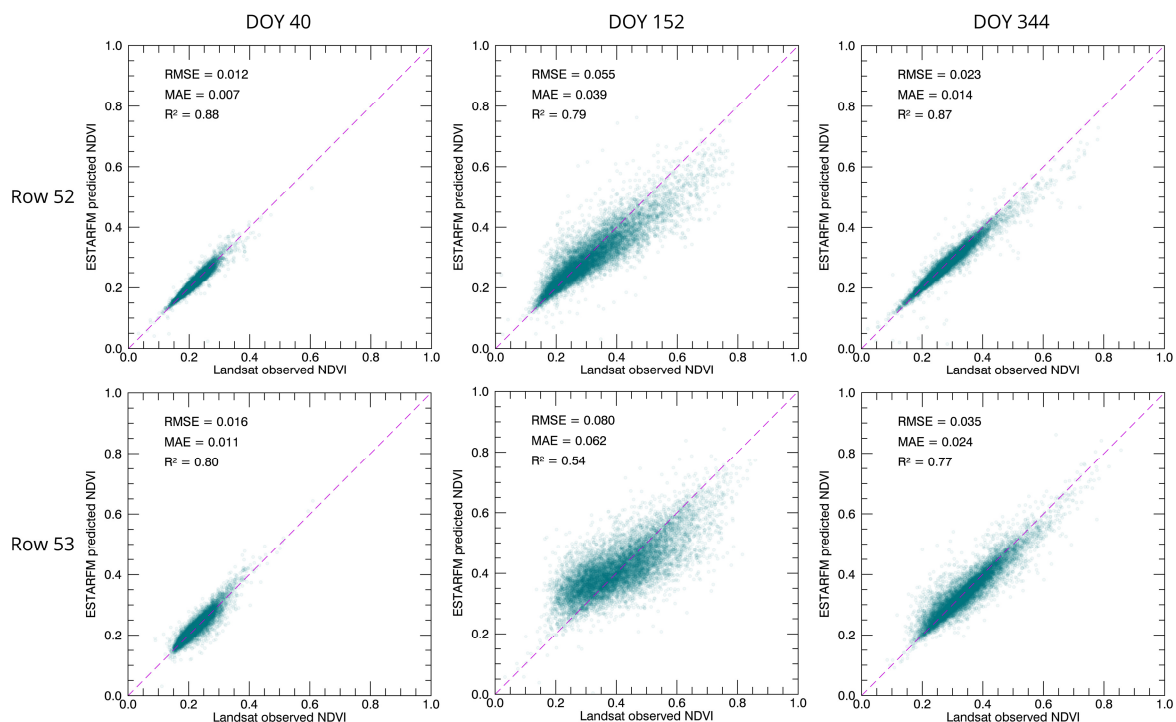

Figure 5.7: Scatterplots of observed versus predicted NDVI for Landsat path 194 and three dates of 2014: DOY 40 is located in the early dry season of 2014, DOY 152 is from the vegetative season and DOY 344 represents the next dry season at the end of 2014. Upper plots are from Landsat row 52 and lower plots from row 53 (modified after Knauer et al., 2016).

scatterplots, the north-south difference of quality can also be observed since the southern row 53 generally exhibits higher errors and more scattering in the plot than row 52.

Eight supplementary Landsat scenes (path 195) with data gaps of up to 80% were used for the estimation of the prediction accuracy of the most severely could-affected part of the rainy season (Table 5.2). The R² values for this period range between 0.3 and 0.76 with MAE between 0.03 and 0.10 and RMSE of 0.05 up to 0.12. Similar to the behavior of the general ESTARFM time series, error values for this period rise towards the climax of the rainy season at the beginning of October (DOY 280) and decrease after that point. Also, the northern tiles during that period exhibit higher correlation values than the respective southern tiles.

Table 5.2: Average quality and error values of ESTARFM predictions from the peak rainy season validated on the basis of eight additional Landsat scenes (path 195) with cloud cover of up to 80% (modified after Knauer et al., 2016).

Row	DOY 191			DOY 271			DOY 287			DOY 303		
	MAE	RMSE	R²	MAE	RMSE	R²	MAE	RMSE	R²	MAE	RMSE	R²
52	0.05	0.07	0.74	0.10	0.12	0.41	0.07	0.10	0.53	0.03	0.05	0.76
53	0.06	0.07	0.58	0.07	0.09	0.30	0.06	0.08	0.33	0.05	0.06	0.56

Within a further analysis, a focus site representing all relevant landscape types of the general focus area like woodlands, grasslands and croplands, was selected for the monitoring of absolute differences between predicted and observed NDVI (Figure 5.8). For the first part of the year, i.e. the late dry season (DOY 40), there is a high agreement between observed and predicted NDVI; only some minor differences appear over the water body in the north of the
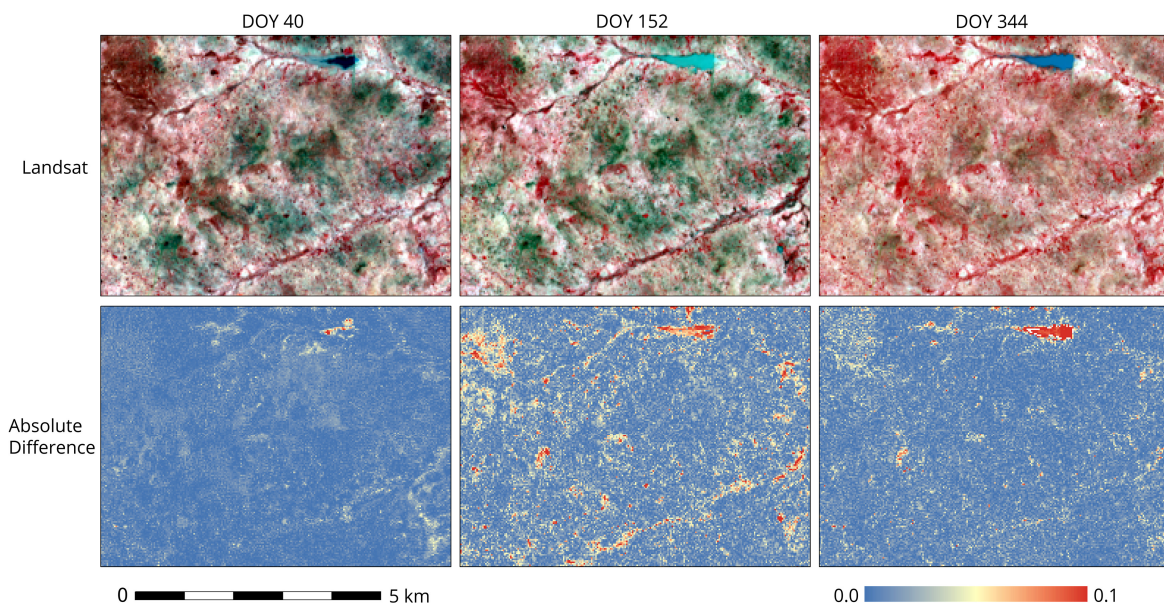

Figure 5.8: Absolute difference between observed and predicted NDVI values for a focus site on tile 194/52 (for exact location see Figure 5.1). The upper images show false colour composites of Landsat (RGB channel combination 5-4-3) for three dates of 2014 and the lower images show the corresponding absolute difference maps (modified after Knauer et al., 2016).

focus site. With the start of the vegetative season (DOY 152), the NDVI differences increase, primarily for areas with dense vegetation cover (marked in dark red colors on the Landsat false color images). The gallery forests alongside the rivers of the focus site are part of these areas. One reason for the differences here are temporarily inundated areas which are difficult to predict with ESTARFM. On the contrary, croplands and areas with less vegetation cover (brighter colors on the Landsat false color images) exhibit minor differences during that period (DOY 152). Towards the end of the year, during the beginning of the next dry season (DOY 344), the absolute differences decline again and only the water body maintains higher deviations between observed and predicted NDVI.

In addition, the fusion quality was assessed by the analysis of consistency between the different predicted Landsat paths. In general, the ESTARFM predictions in the overlapping areas of the two paths (Landsat path 195 and 194) exhibit low median deviations but with a changing pattern over the time series of 2014 (Figure 5.9).
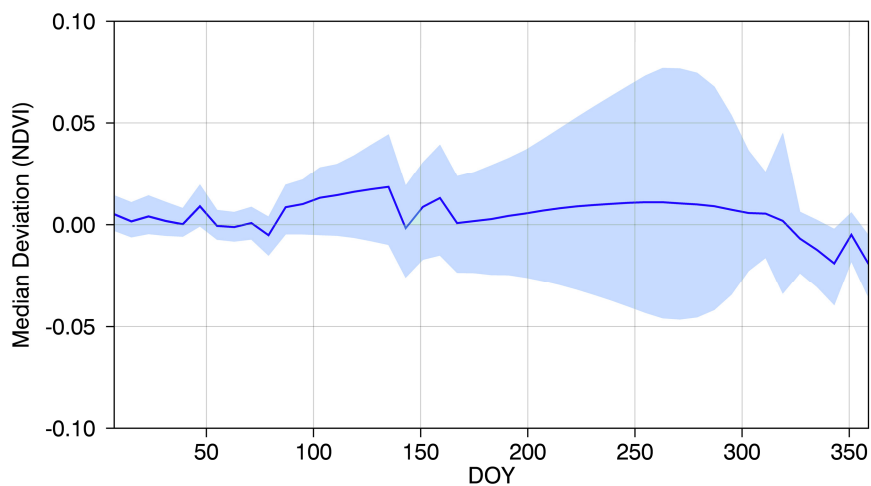
Figure 5.9: Analysis of NDVI differences in the overlapping area between path 195 and path 194 for 2014: Median deviation (dark blue line) and first to third quartile (blue area, modified after Knauer et al., 2016).

Similar to the development of errors identified in the previous analyses, the deviations increase during the rainy season between mid-April and mid-November (DOY 100–320). During this time, the first and third quartiles (light blue area) diverge from -0.05 to 0.08 deviation of NDVI.

5.4.3 Influence of Input Data Availability on Prediction Quality

In order to investigate the effect of temporal distance between Landsat shoulders and prediction date on the fusion quality, one Landsat scene (DOY 143, 195/52) was predicted with all possible shoulder scene combinations (SSC) and not just with the temporally closest one. For all these SSC, the MAE to the actual Landsat NDVI were calculated and plotted against the total temporal distance between the SSC (Figure 5.10). The errors rise with longer time

intervals between the SSC from below 0.03 MAE to more than 0.05 MAE. However, it is also obvious that there are certain groups of data points that have a similar error level over temporal distances of almost 100 days. These groups are from different right shoulders (marked by different symbols in Figure 5.10), i.e. the Landsat-MODIS combination following the prediction date. This indicates that the right shoulder is considerably more important for the prediction quality than the left shoulder preceding the prediction date. The reason for this clearly is that the left shoulder dates are from the dry season in the first part of the year where phenological changes are minor and thus NDVI is relatively stable.
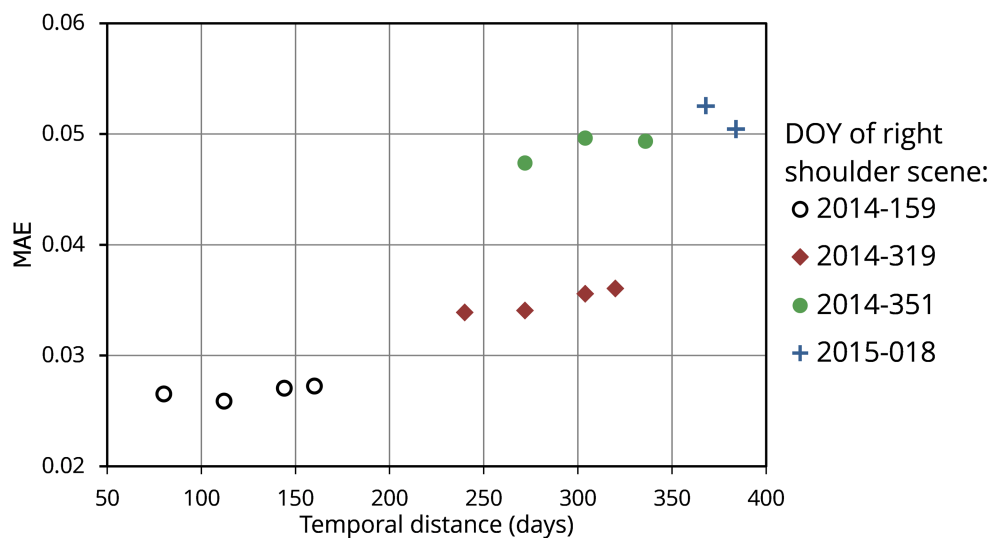
Figure 5.10: Relation between mean absolute error (MAE) and temporal distance of the shoulder scene combinations (SSC). All possible SSC were used for prediction of Landsat scene of DOY 143 (year 2014, tile 195/52) and subsequently the calculation of MAE between predicted scenes and actual Landsat scene. The symbols represent different right shoulder scenes: e.g. the three green dots represent three predictions of DOY 143 with the same right shoulder scene (DOY 319) but different left shoulder scenes (modified after Knauer et al., 2016).

On the contrary, the right shoulders are from different stages of vegetation development: DOY 159 falls within the same phenological phase as the target scene, the right shoulder from DOY 319 is situated at the end of the vegetative season with the browning of leaves, the scene of DOY 351 is widely covered by burned areas and DOY 18 (2015) falls within the next dry season. Subsequently, distinct increases in MAE can be observed every time the right shoulder scene and thus the phenological period changes. These results further indicate the importance to use shoulder scenes that are as close as possible to the prediction date and ideally are from the same phenological phase.

5.5 Discussion

5.5.1 Prediction Uncertainties with ESTARFM

In the presented ESTARFM framework, several improvements have been included that resolve shortcomings and issues mentioned as drawbacks of the original algorithm in previous studies (Fu et al., 2013; Jarihani et al., 2014; Zhu et al., 2010). Nonetheless, not all issues can be solved by the ESTARFM framework. The prediction quality is in a large part dependent on the data availability of the input MODIS and Landsat data. While the ESTARFM framework greatly improves the available set of Landsat input scenes, it is still essential to minimize the temporal distance between the shoulder scenes and the target date (Gao et al., 2015; Zhu et al., 2010). Furthermore, the analysis showed that it is even more important to have at least one shoulder scene from the same phenological phase as the prediction date (Figure 5.10). In addition to the availability of Landsat input data, also the quality of the MODIS input data plays an important role for the prediction with the ESTARFM framework. During the rainy season, the data availability of cloud-free MODIS observations is greatly reduced, especially in the southern parts of the study area. The use of a 16-day rolling composite for the MCD43A4 product diminishes this issue but for some areas southern Burkina Faso, periods of persistent cloud coverage can be even longer. In such a case, the MCD43A4 product of Collection 6 uses a backup database to fill in the gaps with information from the latest full BRDF inversion (Professor Crystal Schaaf's Lab, 2016). In addition, the data is filtered with the TIMESAT software package removing outliers as described in section 4.3. However, applying backup data and filters always introduces a certain amount of inaccuracy leading to discrepancies with the Landsat data and subsequently decreases the prediction accuracies for affected areas.

The land cover types in the area of prediction are another influence factor on the fusion quality (Figure 5.8). The results for the applied focus area indicate low absolute differences between predicted and observed NDVI for land cover types with a generally lower NDVI level such as grasslands or rainfed agricultural areas. On the other hand, land cover types that exhibit stronger phenological changes over the year, like the more densely vegetated regions of the focus area, also show higher absolute differences for the main part of the vegetative season. Such temporally dynamic land cover types are generally more challenging in the prediction with ESTARFM, especially when close shoulder scenes are missing (Emelyanova et al., 2013; Walker et al., 2014). This issue is also reflected in the north-south differences of fusion quality in the focus area with denser vegetation on the southern tiles in comparison to the northern ones. Furthermore, some land cover types with a high NDVI range also exhibit additional rapid changes through fires (e.g. savannas) or floods (e.g. gallery forests) decreasing the prediction quality. Another issue that can increase prediction uncertainties are cloud gaps in the MODIS time series. Even though these areas have been temporally

interpolated in the presented methodology, they can be less accurate in ESTARFM predictions. Land cover types with higher NDVI ranges should be more affected by this issue since the interpolation quality for these dynamic areas should be decreased. In the comparison studies by Jarihani et al. (2014) and Emelyanova et al. (2013), STARFM yielded better results than ESTARFM in such regions. Nonetheless, the overall accuracies for ESTARFM were higher than for STARFM and it also performed better in regions with a high spatial heterogeneity as in the focus area presented here (Emelyanova et al., 2013; Jarihani et al., 2014).

Furthermore, the quality of BRDF correction of the Landsat and MODIS input could negatively influence the prediction quality particularly for land cover types with a complex canopy structure. Walker et al. (2012, 2014) evaluated the STARFM fusion quality of the MODIS product MCD43A4 in combination with data from Landsat-5 Thematic Mapper (TM). Their results showed low correlations between STARFM predictions and actual Landsat data for the NIR band. As a likely reason for this issue, they identified that in contrary to MCD43A4, Landsat-5 TM scenes are by default not corrected for BRDF effects. According to Walker et al. (2014), this discrepancy should be even more pronounced in areas with tall and multi-layered canopies increasing the scattering of NIR radiation. Nonetheless, they also highlighted the value of the MCD43A4 product for the data fusion since other MODIS products may contain considerable BRDF effects because of the sensor's large swath width (Fensholt et al., 2007; Walker et al., 2012). Furthermore, view angle effects with the applied Landsat-8 sensor are minor since it is a push-broom and not a whisk-broom sensor. Although, in the following chapter, the whiskbroom systems Landsat-5 and Landsat-7 are applied as well, BRDF effects are also reduced by the use of NDVI as conducted in this work since the directional signatures of the red and NIR bands are very similar (Bréon and Vermote, 2012).

5.5.2 Added Value for Cloud-Prone Areas

As in many sub-tropical regions with distinct rainy seasons, the acquisition of cloud free remote sensing observations is challenging in West Africa (Fensholt et al., 2007). Higher spatial resolution images as from the Landsat systems are barely available during the rainy season and thus during the major part of the growing season. However, vegetation has a significantly different appearance and subsequently different remote sensing signal during the dry season in comparison to the rainy season. This is not only an issue for the discrimination of different LULC classes during the dry season as described above, but also for the ESTARFM prediction of rainy season dates from dry season dates. Thus, several studies highlighted the importance of Landsat input scenes being as close as possible to the target dates of the fusion in order to minimize prediction errors (Gao et al., 2015; Zhu et al., 2010). The results of this chapter confirmed these findings for the rapidly changing focus area. In addition, the value of Landsat input scenes from the same phenological season and the negative influence of abrupt events such as fires or floods on the prediction accuracy could be highlighted. Within the ESTARFM

framework, an automatic cloud gap filling of Landsat input scenes was implemented to avoid considerable temporal gaps between input and prediction dates. Thus, scenes with partial cloud cover can still be used without the propagation of gaps to the output images of the fusion. In the presented case study, this filling approach enabled the use of 30 Landsat input scenes for the years 2014 even though only 13 of them were cloud-free. This way, input images from several phenological stages of vegetation could be applied in the framework which was shown to significantly improve the quality of prediction.

At this point, it has to be mentioned as well that using the original STARFM (Gao et al., 2006) with two shoulder pairs as input can also avoid gaps in the fusion output as long as one of these shoulder pairs is cloud free. However, image areas with cloud coverage on both input pairs would still result in gaps in the output.

At higher latitudes, data availability could be increased significantly by using tiles of different Landsat paths jointly for the fusion in the ESTARFM framework. In the current implementation of the framework, each tile is treated individually for the generation of time series. However, in these regions the overlapping areas between Landsat paths are considerably bigger than in the focus region of this work. Thus, a pixel based approach which is using all available input data at the respective location could improve the generation of times series. However, on the one hand, such a development is not a trivial one in a data fusion approach like ESTARFM which is taking neighboring pixels into account. On the other hand, the development of the ESTARFM framework focused on the application in region of West Africa where such an approach would not be of great use due to minor overlaps between the paths. Nonetheless, this idea could be part of future improvements of the ESTARFM framework towards applications in other areas.

5.5.3 Added Value for Large-Scale Analyses

In preceding studies on the ESTARFM algorithm, its great potential was highlighted for the generation of time series in high spatial and temporal resolution (Fu et al., 2013; Jarihani et al., 2014; Zhu et al., 2010). However, also several limitations of the algorithm like the long processing times or the method for the selection of similar pixels were outlined that so far hinder large scales applications. The ESTARFM framework developed in this work considerably increases processing speed of the fusion through automation of the algorithm and parallelization of the processing. These are crucial improvements for the application on large scales. In addition, the above mentioned selection of similar pixels used for the regression between Landsat and MODIS was modified. The original approach included in ESTARFM using one global threshold for the entire study area does not account for spatial differences in the heterogeneity of land cover. If a study area is covered by several Landsat tiles with different landscape types and classes, a single threshold would not be sufficient for similar pixel selection. The newly developed selection approach takes regional differences into account,

independent of the size of the study area. So far, the original algorithm has only been applied for rather small study areas with an average size of about 6,600 km² (compare section 3.4.3) which is more than five times lower than the size of one Landsat scene (~35,000 km²). However, a variety of current research questions related to the mapping and monitoring of vegetation require spatially and temporally consistent information at national or supra-national level (Brandt et al., 2015; Fensholt et al., 2013). Especially the countrywide mapping and monitoring of agricultural area intended in this study can greatly benefit from such spatially consistent remote sensing time series. Previous sections have shown the need for high spatial resolution in order to delineate the small-scale farming practices for this purpose. At the same time, the phenological signatures produced with the ESTARFM framework allow for a discrimination of agricultural area from different vegetation types. All this is efficiently made possible on larger scales due to the improvements included in the ESTARFM framework. Thus, it is not only suitable to investigate different local processes but it can also produce results with great value for decision makers on national level.

5.6 Summary and Conclusions

In this chapter, the development of framework for the common fusion algorithm ESTARFM was presented which allows for an application in large, cloud-prone and heterogeneous landscapes. In the resulting ESTARFM framework, several improvements have been included focusing on automation, minimizing of processing times, general handling of the algorithm and on increasing the number of usable Landsat input scenes. The selection process of similar pixels for the regression between Landsat and MODIS was modified in order to account for spatial differences in the heterogeneity of the study area. In addition, an automatic filling of cloud gaps was included that allows using even partially clouded Landsat scenes as input without propagating gaps to the fused time series. Due to an automation and parallelization of the ESTARFM framework, the processing times could be accelerated significantly.

The framework was subsequently tested in a focus area (size: 98,000 km²) in the border region between Burkina Faso and Ghana. The results of this use case showed the high suitability of the ESTARFM framework for the generation of time series on large spatial scales in such a challenging heterogeneous and cloud-prone region. Averaged mean absolute errors (MAE) between predicted and actual Landsat images ranged from 0.02 for the dry season to 0.05 for the vegetative season. This generally indicated higher prediction uncertainties for the latter due to lower input data availability during that period. However, the overall results and the generated phenological signatures in particular showed the value of this method for the reconstruction of LSP in West Africa in a high spatial resolution.

The development of the ESTARFM framework lays the foundation for the mapping and monitoring of agricultural area in West Africa as envisaged in the next chapter. In addition, the framework can also be applied to and improve various other regional vegetation analyses on land degradation, changes in LSP or vegetation composition. With current and upcoming satellite missions such as ESA's Sentinel program, the interest in the analysis of dense time series will increase even more but due to persistent cloud coverage, several regions of the world will stay challenging focus areas. The ESTARFM framework could substantially contribute to the general trend and reconstruct time series also in such areas.

6 Agricultural Expansion in Burkina Faso

With a rapidly increasing population and a rather stagnating agricultural productivity, it is not surprising that agricultural area is spreading in Burkina Faso (compare chapter 1). General evidence for this expansion can be derived from national statistics and some annually updated global maps (FAO, 2016; Friedl et al., 2010). However, accurate information on the quantification of increase in agricultural area and its spatial distribution are missing for Burkina Faso. Statistics on cropped area are of limited use for this purpose since they are only available at provincial level and do not clearly reveal the locations of agricultural conversion. Furthermore, they are derived from household surveys by the National Ministry of Agriculture, Water and Fishery Resources and thus they include a considerable uncertainty (AGRISTAT, 2006). Agricultural extent can furthermore be deduced from various land cover products as outlined in section 3.3, but their usability is often limited for applications on regional scale as planned here because of their low spatial resolution or temporal availability, missing thematic complexity or low regional accuracy in general (Gessner et al., 2015b, 2012; Knauer et al., 2017; Leroux et al., 2014).

In order to derive accurate maps of agricultural area, remote sensing data of high temporal and spatial resolution is needed which cannot be provided by a single satellite sensor so far (compare chapters 1 and 3). The previous chapter presented the development of the ESTARFM framework which can be used to overcome this issue. In this chapter, the ESTARFM framework is applied to generate a high spatial (30 m) and high temporal (8-day) resolution time series for entire Burkina Faso and the years 2001, 2007, and 2014. From these time series, phenological metrics are derived and different types of agricultural area (rainfed agricultural area, irrigated agricultural area and plantations) are classified based on this dataset using the random forest algorithm. On the basis of these maps, the spatio-temporal development of agriculture is subsequently quantified and examined. In a further analysis, the relationship of agricultural expansion to population growth is investigated on provincial level in Burkina Faso outlining regional differences. Finally, the implications of this expansion for natural reserves with different protection status are presented based on a buffer analysis for the respective reserves and their surroundings (Knauer et al., 2017).

The mapping of agricultural expansion in Burkina Faso was already published in the context of the BMBF-funded WASCAL project (Knauer et al., 2017) in order to make it available for the research community. This chapter partially contains descriptions and results of this publication.

6.1 Time Series Generation and Classification Features

For the entire coverage of Burkina Faso and for the years 2001, 2007, and 2014, the ESTARFM framework was first used to automatically fill cloud-gaps in the input Landsat datasets. Hereafter, 8-day, 30 m NDVI time series were generated with the framework from the above described MODIS and filled Landsat data. This 8-day time interval was equally produced for all Landsat tiles starting at DOY 7 (7th January) and ending at DOY 359 (25th December), similar to the Landsat acquisition interval of path 195. This way, the resulting time series could easily be mosaicked after the production without further interpolation. All in all, more than 2000 ESTARFM Landsat-like scenes were predicted, distributed over 17 tiles (compare section 4.2). From these ESTARFM time series, several seasonal and land surface phenology (LSP) features were derived as input for the delineation of different LULC classes in Burkina Faso. Such LSP metrics are highly valuable for the discrimination of spectrally similar classes with differing phenological development. On the basis of single remote sensing observations, agricultural areas and different types of savanna are often difficult to distinguish from each other as preceding studies in the same latitudes of West Africa emphasised (Forkuor, 2014; Lambert et al., 2016; Vintrou et al., 2012b). This can be improved by the use of dense time series since these classes often show different phenological development over the course of the year.

For the classification of 2001 and 2014, a total of 57 features were used (Table 6.1). In 2007, not the complete ESTARFM time series could be generated due to missing Landsat data after October 2007 (compare section 4.2). Thus, a total of 41 features were analysed for the generation of a LULC map in 2007 as shown in Table 6.1. The feature set for classification comprises all available 8-day NDVI datasets (45 time steps for 2001 and 2014 and 33 time steps for 2007). Furthermore, eight LSP metrics were derived for 2001 and 2014 from the time series using the TIMESAT software package (Eklundh and Jönsson, 2015). Again, due to the shorter time series of 2007, only four LSP metrics could be derived for this year. In addition, three seasonal metrics were calculated from the time series of the three years: the mean NDVI, the maximum NDVI and the seasonal sum of NDVI (DOY 119-263). In order to account for a changing north-south gradient in the classes, a 1 km WorldClim rainfall dataset (Hijmans et al., 2005) of average long-term annual rainfall sums (1960-1990) was included as a classification feature after co-registration and resampling to the 30 m of the ESTARFM time series. Previous studies have shown that the changing north-south gradient in moisture is also

reflected in a changing land surface phenology of LULC classes (Bégué et al., 2014, 2011, Vintrou et al., 2012a, 2012b).

Table 6.1: Classification features for the derivation of LULC classes in Burkina Faso. Features marked with an asterisk could not be derived for 2007; features written in italic were generated with TIMESAT (Eklundh and Jönsson, 2015) (table modified after Knauer et al., 2017).

Feature	Description
Start of Season	Time for which the left part of the NDVI curve has increased to 30% of the seasonal amplitude measured from the left minimum level
*End of Season *	Time for which the right part of the NDVI curve has decreased to 50% of the seasonal amplitude measured from the right minimum level
*Length of Season *	Time from the start to the end of the season
Seasonal Amplitude	Difference between the peak value and the base level
Base Level	Average of the left and right minimum values
*Middle of Season *	Mean value of the times for which the left part of the NDVI curve has increased to the 80% level and the right part has decreased to the 80% level
Rate of Increase at the Beginning of Season	Ratio of the difference between the left 20% and 80% levels and the corresponding time difference
*Rate of Decrease at the End of the Season *	Ratio of the difference between the right 20% and 80% levels and the corresponding time difference
Maximum of NDVI	Annual maximum NDVI of the ESTARFM time series
Mean of NDVI	Annual mean NDVI of the ESTARFM time series
Seasonal Sum of NDVI	Sum of NDVI values (DOY 119 to DOY 263) of the ESTARFM time series
Mono-temporal NDVI	Single time steps of NDVI of the ESTARFM time series (45 scenes used for 2001 and 2014 and 33 scenes for 2007)
Precipitation	Long-term average of annual precipitation sums (WorldClim) (Hijmans et al., 2005)

6.2 Classification Procedure and Accuracy Assessment

As mentioned above, the random forest (RF) classification algorithm was used to derive the defined LULC classes from the ESTARFM time series (Breiman, 2001). Due to its ability to handle classes with non-normally distributed features, its robustness to overfitting and noise and its generally fast processing, it has been frequently used in the literature (Breiman, 2001; Gessner et al., 2015b; Wohlfart et al., 2016b). In this work, the randomForest package as implemented in the programming language R was applied (Liaw and Wiener, 2002). The final RF models were built independently for the three years on the basis of 500 decision trees each and the default value of 12 randomly sampled features at each split in the decision tree (Breiman, 2001).

As commonly recommended in the literature (Adelabu et al., 2015; Wohlfart et al., 2016b), the reference dataset was split into random sets of 70% and 30% of the total number of polygons. The 70% set was applied in the training of the RF classifier and the 30% set was used for validation purposes. Classification accuracy was then highlighted with confusion matrices which show the respective user's and producer's accuracy of each class in combination with

the overall accuracy of the classification. This way it can be estimated which classes are under- or overrepresented in the final classification (Congalton and Green, 2009). In addition, the feature importance is analysed with the randomForest package calculated as the total decrease in node impurities (Gini index) from splitting on the variable, averaged over all trees (Liaw and Wiener, 2002).

After the initial classification, several post-classification rules were applied to sharpen the agricultural focus and in order to improve the classification accuracy. In a first step, a *Savanna* class was created by merging the preliminary classes *Grassland*, *Woody Vegetation*, *Temporarily Flooded*, and *Forest*. Although these classes were not in the focus of this work, they were delineated separately because of their different spectral and phenological characteristics. In a second step, the *Irrigated Agricultural Area* class was improved by excluding areas with an annual mean NDVI below a threshold of 0.2. These areas mainly occur at shorelines of lentic water bodies which vary considerably in size over the course of the year. Some of these areas with temporary flooding were misclassified as *Irrigated Agricultural Area*. However, since their NDVI is too low for a cultivated area (additionally confirmed by visual analysis in Google Earth), these areas were reassigned to the *Temporarily Flooded* class. In a last step, the *Rainfed Agricultural Area* class was corrected within the boundaries of protected areas. Due to frequent uncontrolled fires during the dry season in these reserves, vegetation phenology is considerably altered resulting in some overestimations of this agriculture class. For the identification of such areas the MODIS burned area product (MCD45A1, Boschetti et al., 2013) was employed. In areas within the park boundaries that were affected by fires, *Rainfed Agricultural Area* was reclassified to *Savanna*.

6.3 Relationship to Population and Protected Areas

Since land in Burkina Faso is primarily cultivated for self-subsistence, it is important to know how these areas are developing in the light of increasing population pressure and if the situation for farmers and their livelihoods is potentially changing. Thus, possible relationships between agricultural expansion and rural population growth were analysed on provincial level. Changes of agricultural area and population between 2001 and 2014 were compared and regional differences in their development were highlighted. Furthermore, the correlation of these changes was analysed on provincial level. In addition, the respective datasets of the single years (2001 and 2014) of agricultural area and rural population numbers were correlated with each other. Their relation was investigated with the Pearson product-moment correlation coefficient (Kirch, 2008). Furthermore, in order to visualize the available farmland for self-subsistence, the changes of agricultural area per rural inhabitant between 2001 and 2014 were analysed on provincial level.

Another important issue in relation to agricultural expansion in Burkina Faso is the preservation of protected areas. When arable land is getting sparse, the pressure on the last areas of semi-natural vegetation should increase. For the analysis of this issue, two protected focus areas, the Kaboré Tambi National Park and the Tiogo classified forest were selected. The development of agricultural area in these parks and in their vicinity was investigated on the basis of a buffer analysis quantifying the density of agricultural area in the respective reserve and in five 2 km distance zones (0–2 km, 2–4 km, 4–6 km, 6–8 km, 8–10 km) around the reserve.

6.4 Results

In this section, the results of the Burkina Faso wide time series generation with ESTARFM as well as the results of the agricultural classification and their analysis with respect to population growth and protected areas are presented. In a first part, the NDVI time series generated with the ESTARFM framework as well as the relevance of individual features for the agricultural classification are shown. Then, the novel classifications of agricultural areas in Burkina Faso for 2001 and 2014 are presented. Several local examples are given highlighting regional differences and patterns and showing the development of the rather small scale changes of plantations and irrigated agricultural areas. Furthermore, the agricultural expansion between 2001 and 2014 is presented in the light of population growth and direct relationships are analysed on a provincial level. In a last section, the focus lies on two protected areas, investigating the agricultural development in their vicinity and deriving implications for their protection and habitat function. In general, the analysis of results in the following sections focuses on the years 2001 and 2014 since the classification of 2007 is not covering the whole country as outlined in section 4.2.

6.4.1 ESTARFM NDVI Time Series and Variable Importance

For each of the three focus years 2001, 2007, and 2014, a dense 8-day NDVI time series of 45 single time steps (33 for 2007) was generated on the basis of the ESTARFM framework. Figure 6.1 shows the Burkina Faso wide NDVI averaged over the ESTARFM time series of 2014. In general, the mean NDVI well reflects the north-south gradient of vegetation in Burkina Faso from the sparsely vegetated Sahara region to the sub-tropical southwest of the country. Even though this is not a classified map yet, multiple landscape features are apparent which have characteristic NDVI values standing out from their surroundings. The protected areas of Burkina Faso are one example exhibiting significantly higher mean NDVI values than their direct vicinity. Since they represent the (semi-) natural vegetation of the country, it can already be deduced from this map that their direct surroundings have been altered by anthropogenic influence. Further obvious man-made reduction of the mean annual NDVI can be observed

around several cities of Burkina Faso, especially around the capital Ouagadougou and the country's second biggest city Bobo Dioulasso. These patches represent the city area as well as the agricultural areas in their direct surroundings. How such agricultural areas are accumulated around urban areas can be well observed in the southwest of Ouahigouya, where several small towns formed spots of lower NDVI values that are considerably greater than the actual urban area. In the relatively dry Sahelian region in the north of Burkina Faso, the river network with its gallery forests stands out sharply from the surrounding savanna and agricultural areas.
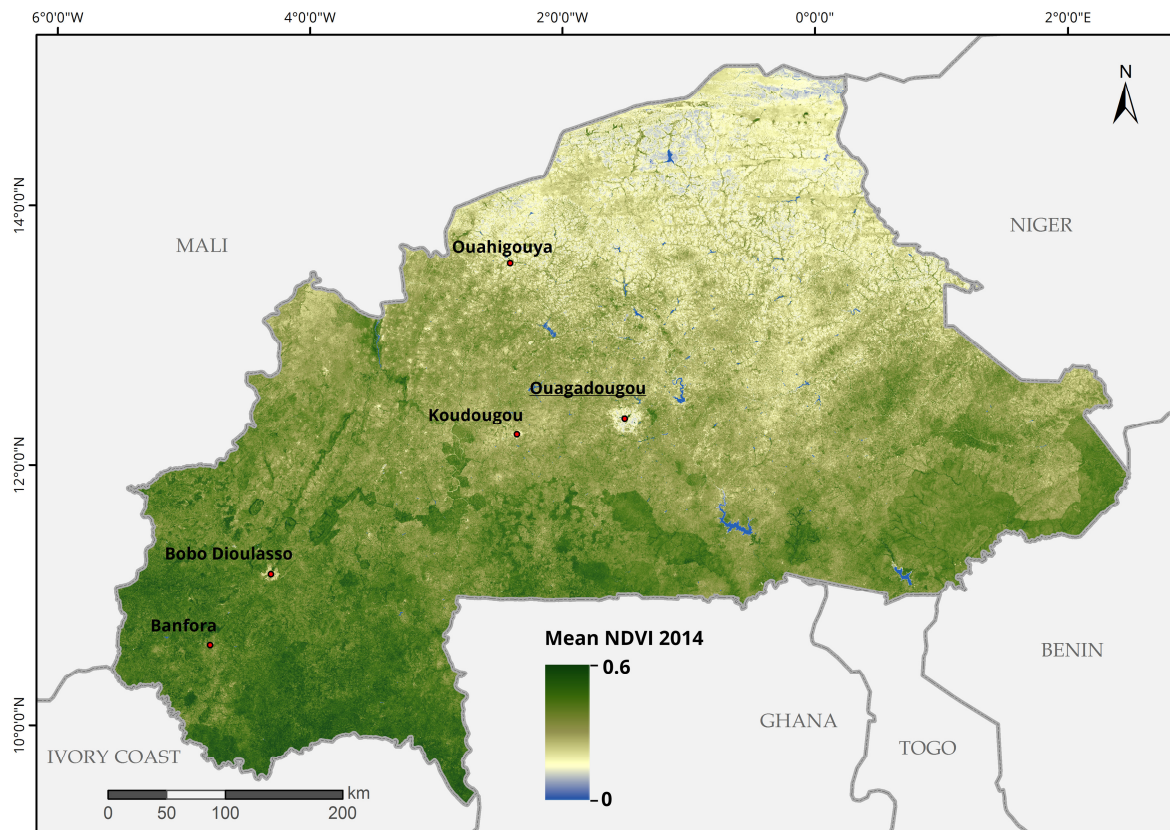
Figure 6.1: Mean annual NDVI for the year 2014 derived from the ESTARFM time series.

The mean annual NDVI is only one classification feature of up to 57 datasets used for the derivation of agricultural area from the ESTARFM input time series (compare section 6.1). Their respective importance for the classification result in the study area is highlighted in Figure 6.2. The long-term average precipitation derived from the WorldClim dataset is the most important classification feature representing the north-south gradient of climate and vegetation mentioned before. This gradient also reflects the general frequency of occurrence of different classes as well as the changing behaviour of some of them (Bégué et al., 2011; Knauer et al., 2017; Vintrou et al., 2012a, 2012b).
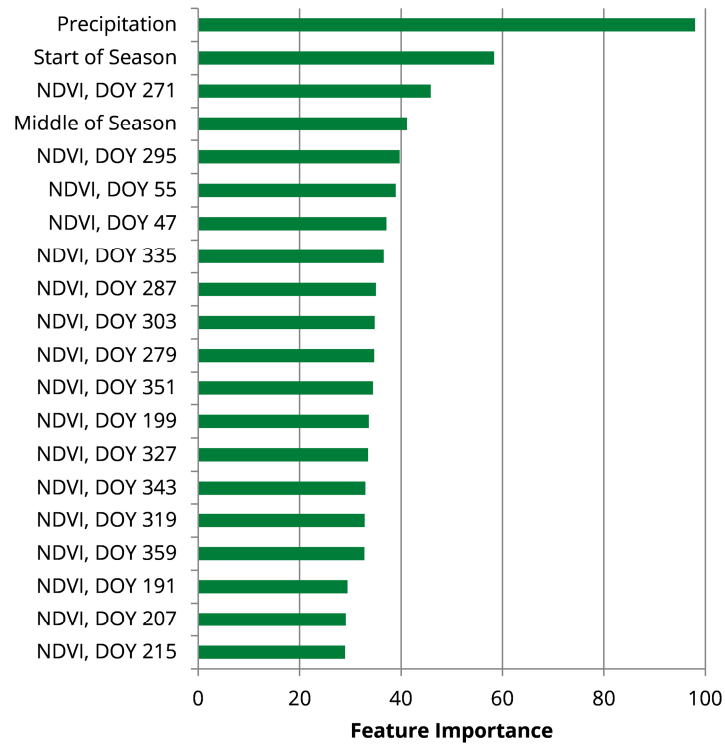
Figure 6.2: Average feature importance of the top 20 input variables on a scale between 0 and 100 for the years 2001 to 2014 (modified after Knauer et al., 2017).
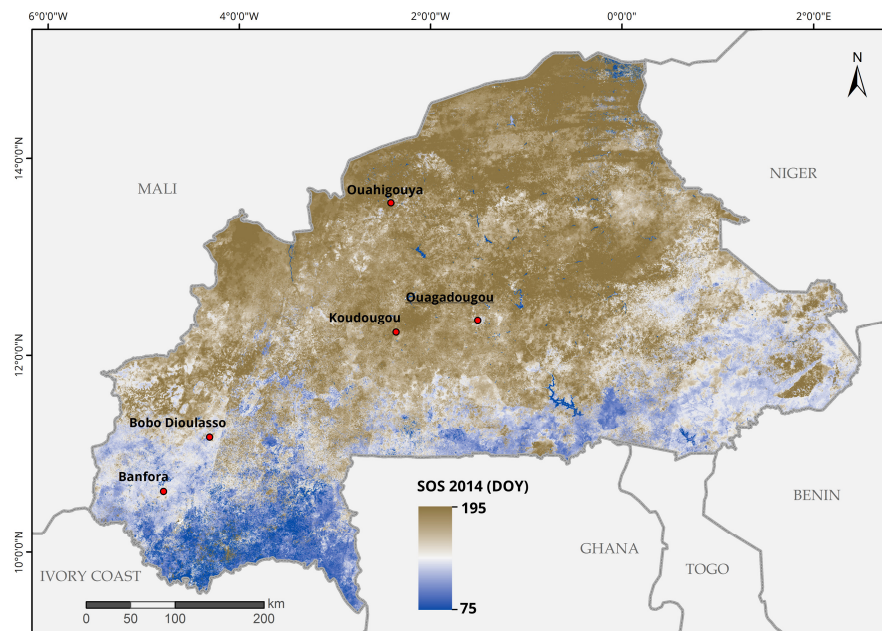
Figure 6.3: Start of Season for the year 2014 derived from the ESTARFM time series with TIMESAT.

The Start of Season (SOS) is the second most important feature for the classification of agricultural area (Figure 6.2 and Figure 6.3). The SOS again reflects the north-south gradient with differences in green up of vegetation from more than 100 days in Burkina Faso. It generally shows early dates for irrigated agricultural areas in the vicinity of water bodies being more independent of the onset of the rainy season.

Natural vegetation classes commonly display earlier SOS dates than rainfed agricultural areas which in parts can be seen by the comparison of protected areas and their surroundings. While natural vegetation responds quickly to the first rainfalls of the year, rainfed agricultural areas are typically sown later and thus present later green-ups (compare plots in Figure 5.4).

Another important feature for the classification of the different years is the NDVI at the end of September (DOY 271, Figure 6.4). Similar to the Middle of Season (MOS) feature derived with TIMESAT, the DOY 271 represents the peak of rainy season where vegetation cover is dense and NDVI values are highest. In general, classes with dense and multi-layered vegetation such as plantation and forests show higher NDVI values during this time than e.g. single-layered grasslands. Since these multi-layered LULC types also have a generally longer vegetative season, their MOS tends to be later as well.
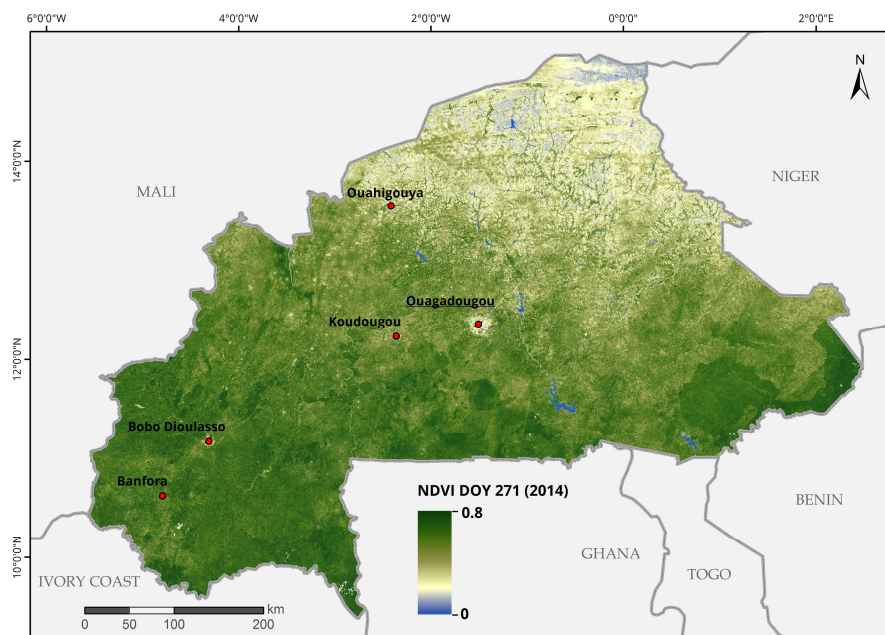
Figure 6.4: ESTARFM NDVI for end of September 2014 (DOY 271).

6.4.2 Agricultural Expansion between 2001 and 2014

Figure 6.5 shows the novel 30 m products of Burkina Faso wide agricultural area for the years 2001 and 2014. In 2001, rainfed agricultural area is covering about 60,441 km² corresponding to 22% of the national area of Burkina Faso (Table 6.2). The focus of agriculture in 2001 lies around the central cities of Ouagadougou and Koudougou as well as in the southeast of them. The outer regions bordering Niger and Nigeria in the east, the Sahel region in the north and the more densely vegetated southwest of Burkina Faso are almost not cultivated in 2001. The latter region however contains most of Burkina Faso's plantations which cover an approximate size of 561 km² in 2001. The climate in Burkina Faso's southwest with highest annual rainfalls is the most suitable for tree crops. In 2001, irrigated agricultural

areas only cover a small extent of about 78 km², established adjacent to artificial water reservoirs or alongside rivers and primarily located in the central areas of Burkina Faso.
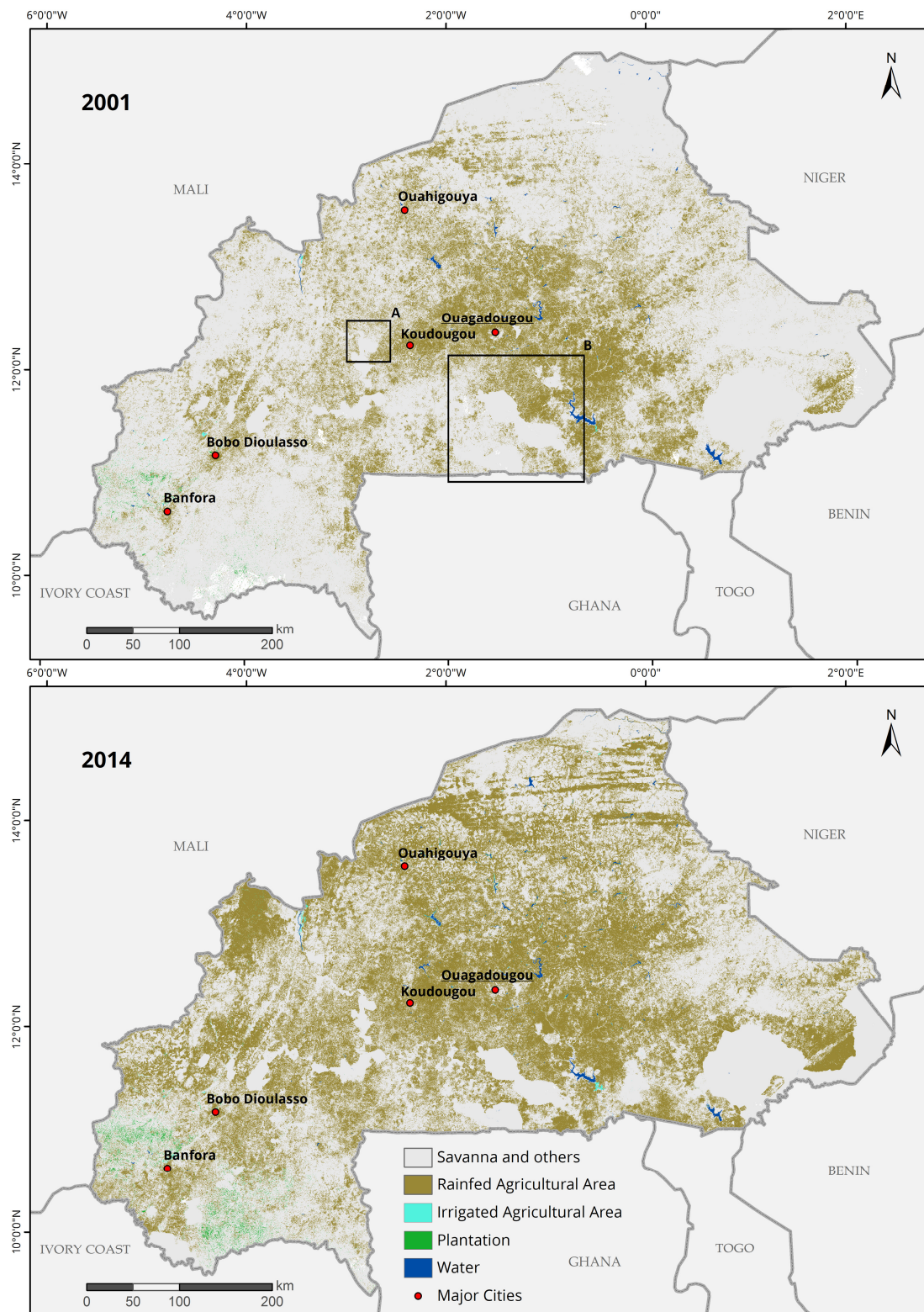
Figure 6.5: Classification results of agricultural areas in Burkina Faso for the years 2001 (top) and 2014 (bottom). The framed areas in the upper classification mark the extent of the focus sites of section 6.4.4, the classified forest of Tiogo (A) and the Kaboré Tambi National Park (B) (modified after Knauer et al., 2017).

Table 6.2: Extent of the three types of agricultural area for the years 2001 and 2014 and the respective change between these years (modified after Knauer et al., 2017).

| Class	2001 (km²)	2014 (km²)	Change
Rainfed Agricultural Area	60,441	114,994	+90%
Irrigated Agricultural Area	78	345	+344%
Plantation	561	1,568	+179%

Until 2014, the rainfed agricultural area spread to about 114,994 km² which is a plus of 90% and now corresponds to 42% of Burkina Faso's country area. In the central areas where large parts were already cultivated in 2001, the agricultural area generally concentrated even more. In addition, the rainfed agriculture spread from the central parts into the more remote regions of the north, east and southwest. In the Sahel region, the valorisation of major consolidated and immobilized sand dunes can be seen (Figure 6.5, bottom). Furthermore, the area covered by plantation almost tripled between 2001 and 2014, from 561 km² to 1,568 km². Two major agglomerations of plantations are now observable in the southwest of Burkina Faso, one is located in the southeast of Banfora, in the provinces of Comoé and Poni and another has established in the northeast of Banfora in the provinces Kénédougou and Léraba (compare Figure 6.9 for provinces).

Table 6.3: Confusion matrices for the years 2001 (top), 2007 (middle) and 2014 (bottom) highlighting user's (UA) and producer's accuracies (PA) for the agricultural classes as well as overall accuracies in the respective bottom right corner of the tables (modified after Knauer et al., 2017).

		Reference				
	2001	*Rainfed*	*Irrigated*	*Plantation*	*No Agriculture*	**UA**
Classification	*Rainfed*	1300	22	0	204	85.19%
	Irrigated	0	406	1	41	90.63%
	Plantation	0	2	288	8	96.64%
	No Agriculture	242	36	13	4969	94.47%
	PA	84.31%	87.12%	95.36%	95.16%	92.45%

		Reference				
	2007	*Rainfed*	*Irrigated*	*Plantation*	*No Agriculture*	**UA**
Classification	*Rainfed*	2146	40	0	262	87.66%
	Irrigated	18	244	0	37	81.61%
	Plantation	6	0	275	70	78.35%
	No Agriculture	210	116	17	5234	93.85%
	PA	90.17%	61.00%	94.18%	93.41%	91.05%

		Reference				
	2014	*Rainfed*	*Irrigated*	*Plantation*	*No Agriculture*	**UA**
Classification	*Rainfed*	3486	17	0	566	85.67%
	Irrigated	0	424	2	2	99.07%
	Plantation	0	0	57	17	77.03%
	No Agriculture	299	164	8	7237	93.89%
	PA	92.10%	70.08%	85.07%	92.52%	91.25%

The irrigated agricultural areas also increased significantly to a size of 345 km² which corresponds to a rise of 340%. Although the absolute extent is still minor in comparison to the other two types of agricultural area, the area increase shows intensification efforts to achieve crop yields more independent of the annual rainfall patterns.

Overall classification accuracies are listed in Table 6.3 with user's and producer's accuracies for the respective classes and for the three years of investigation. The accuracies are on a constantly high level with 92% for 2001, 91% for 2007, and 91% for 2014. Rainfed agricultural area is generally very well delineated with user's and producer's accuracies above 80% for all three years. Irrigated agricultural area exhibits lower producer's accuracies for the years 2007 and 2014 indicating an underestimation of this class. This is most likely caused by confusion with temporarily flooded areas in some regions. The user's and producer's accuracies for plantation areas range between 77% and 97%. Confusion primarily occurs with gallery forests, i.e. alongside rivers where plantations are overestimated.

Plantation Area

Within the 14 years of investigation, plantation area spread by an increase of 179% from about 56,000 ha to 157,000 ha (compare Table 6.2). 99% of the national plantation area is located in six southwestern provinces of Burkina Faso, namely Comoé, Kénédougou, Léraba, Poni, Houet, and Noumbiel (Figure 6.6 right).

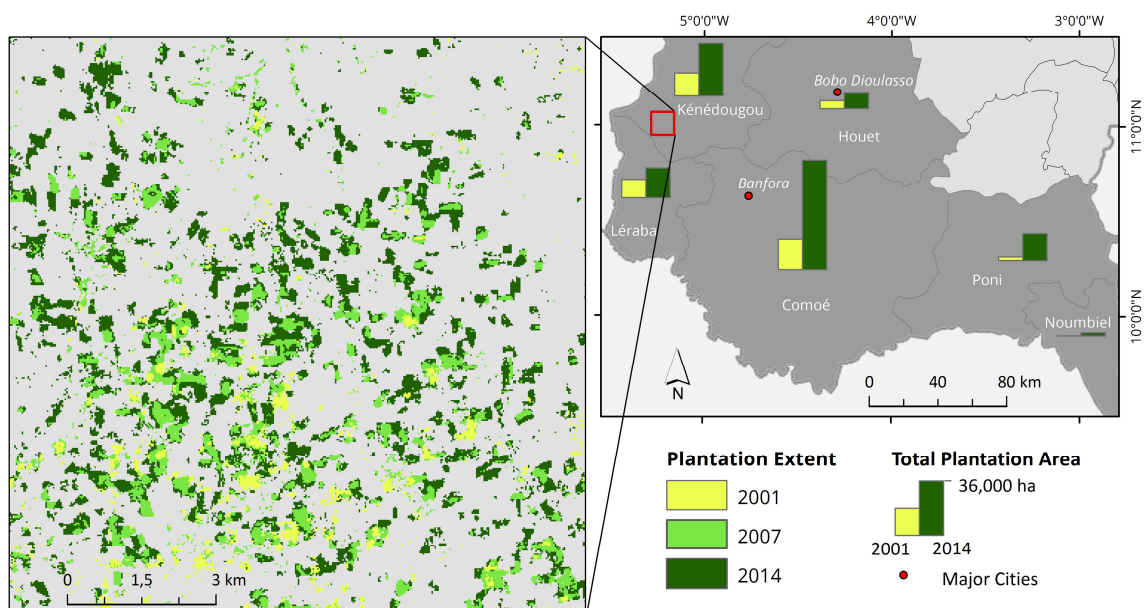Figure 6.6: Development of plantation areas in the Kénédougou province between 2001 and 2014 (left) and major plantation provinces in the Southwest of Burkina Faso and their plantation extent for 2001 and 2014 (right) (modified after Knauer et al., 2017).

Especially the Comoé province comprehends about 71,000 ha corresponding to 46% of the national plantation area. The detailed spatial development of plantation area is shown in

Figure 6.6 (left) for a focus area in the Kénédougou province, the province with the second biggest plantation area in Burkina Faso (34,000 ha).

Irrigation Area

In order to improve food security and to mitigate the negative impact of the increasing variation in the timing and length of the rainy season, an efficiently implemented irrigation system could play a central role in the agriculture of Burkina Faso (MAHRH, 2004). Despite an estimated potential of 233,500 ha irrigable land in Burkina Faso, only about 34,500 ha were actually irrigated in 2014 (IWMI, 2010; MAHRH, 2004; Sedogo and Bourgou, 2015). These irrigation areas are mainly located in the westernmost, central and northern provinces (Figure 6.7). The water for these cultivated areas mainly comes from small reservoirs, which developed through the damming of rivers, and less commonly from natural lakes. The distribution and density of water reservoirs in Figure 6.7 partially corresponds to the provinces with higher irrigation area. However, the smaller reservoirs are usually primarily used for household water supply and livestock.
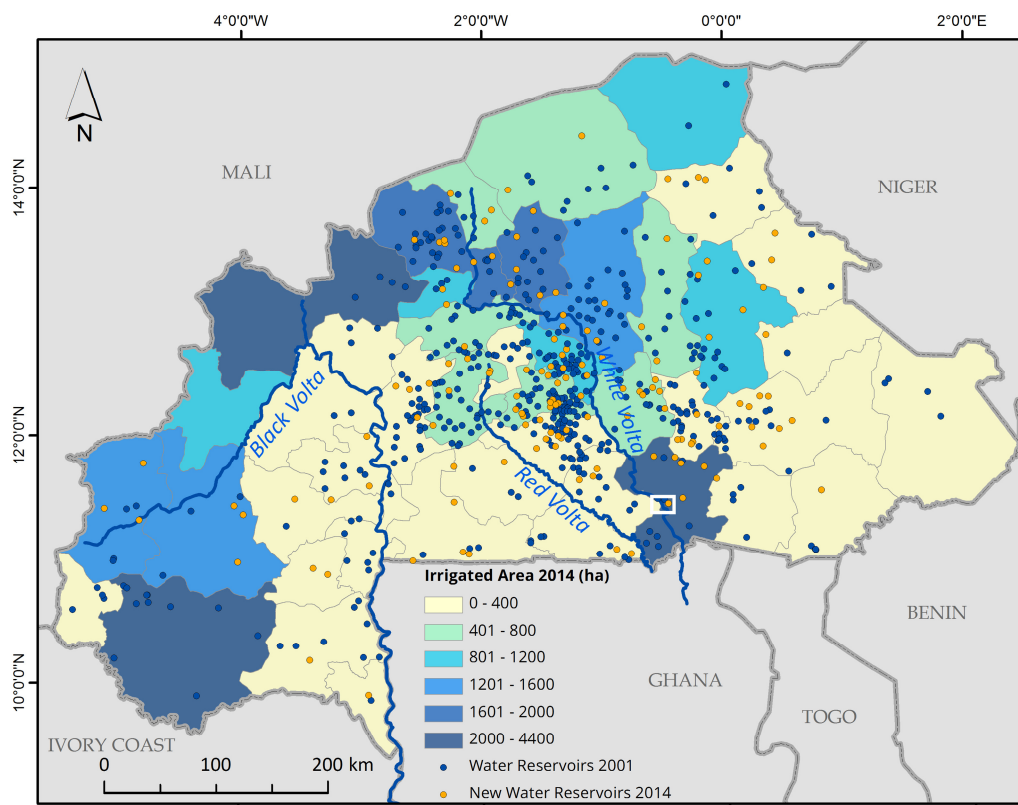

Figure 6.7: Irrigation area per province (in ha) for the year 2014 and positions of water reservoirs for the years 2001 and 2014. The latter were extracted by manual digitization from the classification results. The white rectangle in the Southwest of Burkina Faso marks the position of the focus site in Figure 6.8 (modified after Knauer et al., 2017).

Although, there are only few water reservoirs in the westernmost provinces, they are among the ones with the highest absolute irrigation area. The reason for this is that the highest values per province can be attributed to few major irrigation systems which was confirmed by visual inspection of these provinces. These professional irrigation sites receive their water supply directly from rivers like the Black Volta and their tributary streams without damming them. The irrigation area around Lake Bagré in the central southeastern province Boulgou (Figure 6.8) forms an exception of this rule. The dam was originally built for the generation of hydropower in 1994 retaining the water of the White Volta River. However in the following years, the surrounding area also developed to one of the biggest irrigation sites in Burkina Faso. In 2001, the Lake Bagré irrigation area covered about 9.26 km². Since then, this area spread to 21.67 km² (2007) and to 37.28 km² (2014) corresponding to an increase of about 300% during the time of investigation.
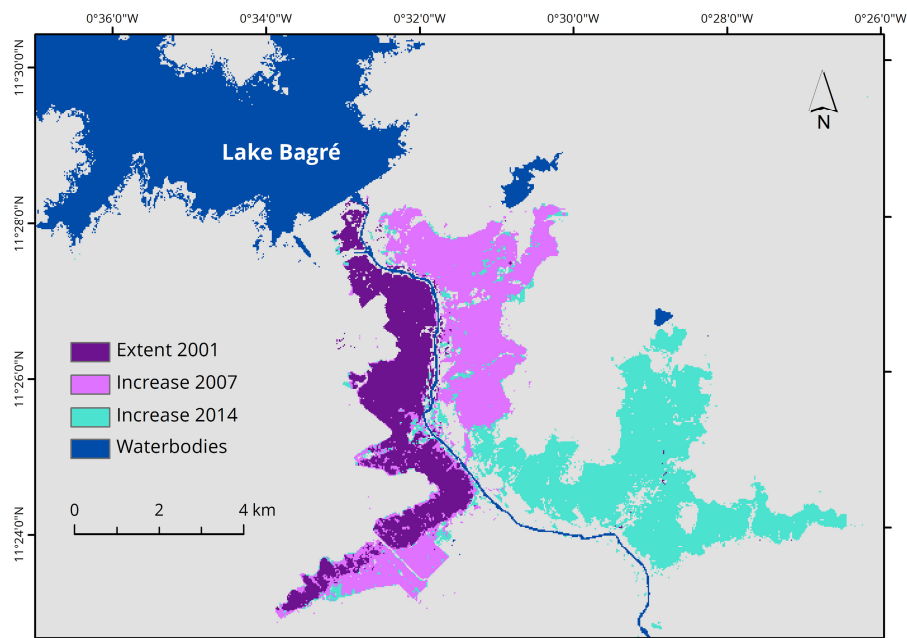
Figure 6.8: Development of the irrigation area downstream of the Bagré dam in southern Burkina Faso showing three stages of expansion (2001, 2007 and 2014); the location of this site is marked in Figure 6.7 (modified after Knauer et al., 2017).

6.4.3 Agricultural Expansion in the Light of Population Growth

Between 2001 and 2014, the absolute population of Burkina Faso experienced a rapid growth from about 12 million to more than 17 million people (FAO, 2016). The same trend is visible for the country's rural population in almost all provinces (compare Section 2.5). The provinces of Tapoa (TAP), Comoé (COM) or Houet (HOU) in the more remote regions of Burkina Faso experienced a rise of up to 244,000 people during this time (Figure 6.9 left). In the contrary, the central provinces, which had already high population numbers in 2001, and also the central southern provinces only exhibit a moderate population growth. The only

exception from the general trend constitutes the province of Zondoma (ZON) which exhibited a negative rural population development with a decrease of 58,000 rural inhabitants between 2001 and 2014. If this single drop in Zondoma is due to a specific reason or even an error in the dataset could not be determined.

The agricultural expansion between 2001 and 2014 in the provinces generally follows the patterns observed for the rural population growth (Figure 6.9 right). Agricultural areas emerged and spread into all regions of Burkina Faso even in the remote and previously almost not cultivated provinces. Mentioned above for their strong rural population growth, the provinces of Tapoa (TAP, 323,000 ha) and Comoé (COM, 315,000 ha) also range among the provinces with the largest increase in agricultural area. The two northern provinces Kossi (KOS) and Soum (SOM) experienced the highest numbers of agricultural expansion with increases of 392,000 ha and 386,000 ha, respectively. In line with the moderate population growth, only little increase in agricultural area can be observed for the central provinces with a plus between 600 ha and 60,000 ha for the provinces Koulpélogo (KOP) and Bazèga (BAZ), respectively.
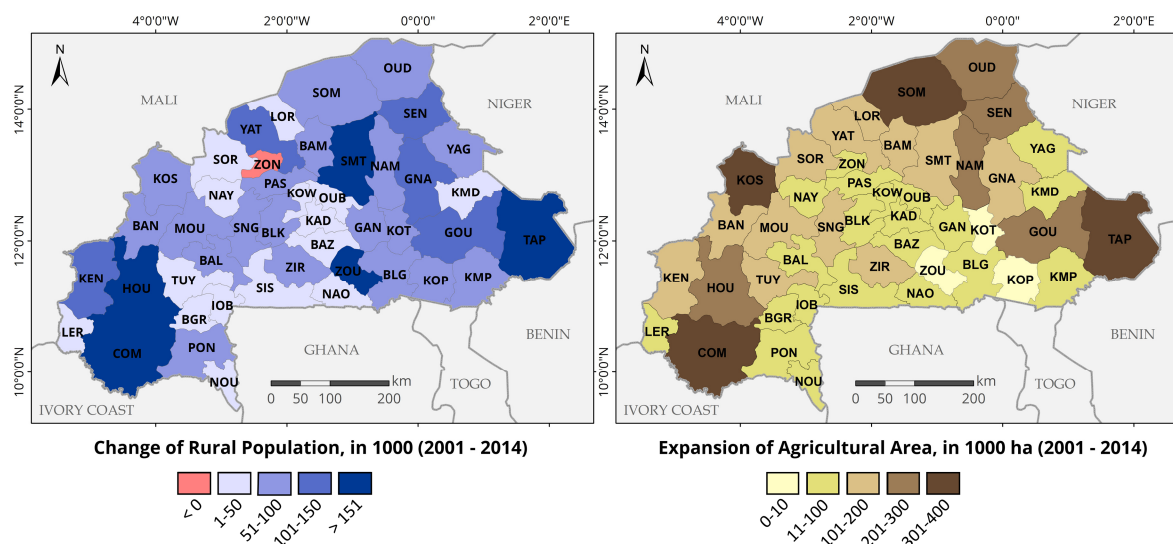
Figure 6.9: Provincial development of rural population in Burkina Faso between 2001 and 2014 (left) (INSD, 2014) and expansion of agricultural area by province extracted from the classification results (right); province names are abbreviated by three letters for readability (modified after Knauer et al., 2017).

By correlating the values of rural inhabitants and agricultural area, a significant relationship becomes apparent (Figure 6.10). In 2001, the rural population in the provinces correlates well with the respective extent of agricultural area (r = 0.84). The highest numbers in 2001 for both, rural population and agricultural area are observed for the large and rather central provinces of Boulgou (BLG), Sanmatenga (SMT), and Yatenga (YAT) (Figure 6.10a). In 2014, the correlation between population and agricultural area is still on the same level (r = 0.84) with a general increase for both variables (Figure 6.10b). The provinces with the highest

numbers of both variables in 2001 are still among the highest in 2014. In addition, the above mentioned provinces of Comoé (COM) and Soum (SOM) as well as several provinces from the region Est, namely Tapoa (TAP), Gourma (GOU), and Gnagna (GNA) joined the group of provinces with highest numbers for the two variables. Despite a rather low correlation coefficient (r= 0.49) for the relation of changes between 2001 and 2014, it is apparent that rural population growth and agricultural expansion go in line for several provinces (Figure 6.10c). However, numerous reasons and pathways are possible for provinces diverging from this general trend and resulting in different developments of agriculture and rural population. Especially the more remote provinces of the regions Sahel, Est or Boucle de Mouhoun show the strongest increases in both variables while the central and comparably smaller provinces of the regions Plateau Central, Centre-Est, and Centre-Sud exhibited only a rural population growth without a significant further expansion of their agricultural area.

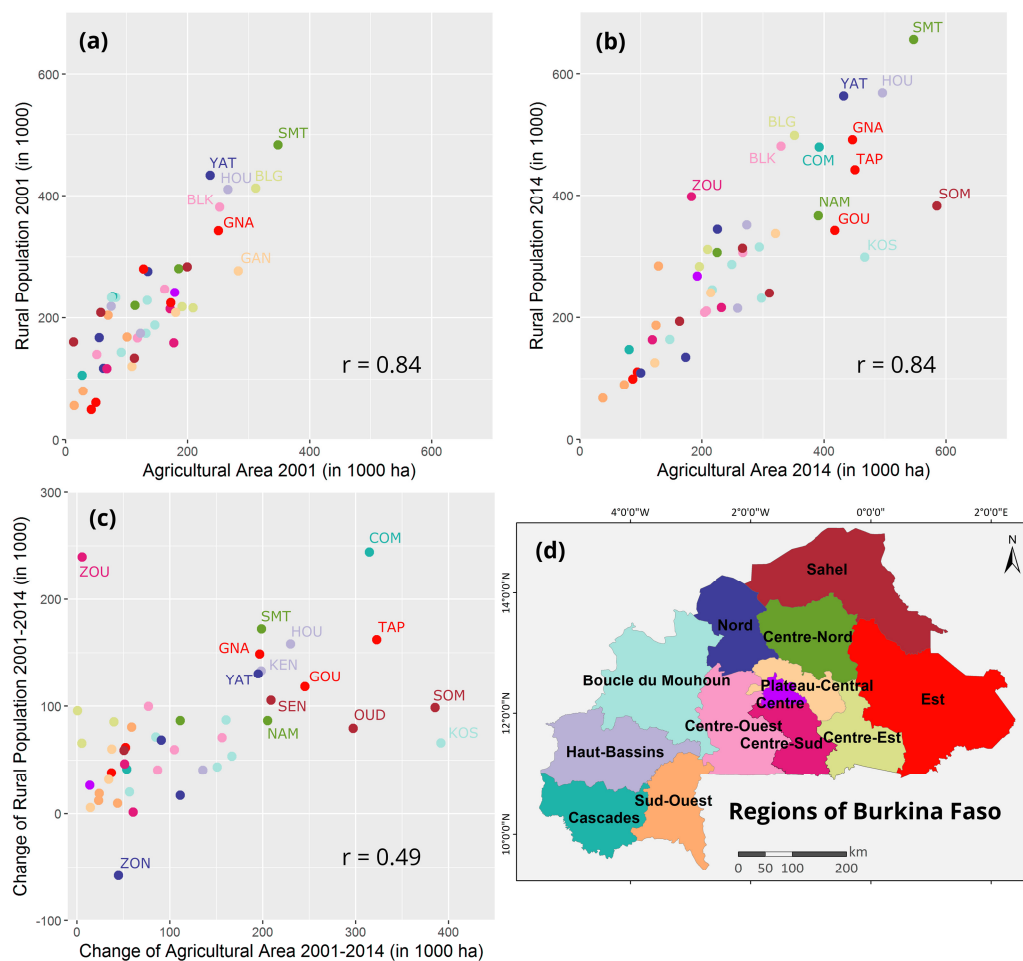Figure 6.10: Scatterplots of provincial agricultural area (in 1000 ha) versus rural population (in 1000 persons) for the years 2001 (a) and 2014 (b) and the changes between these two years (c); the points are coloured according to their region in Burkina Faso (d) (modified after Knauer et al., 2017).

Between 2001 and 2014, the mean agricultural area per rural Burkinabe increased from 0.63 ha to 0.88 ha (Figure 6.11). The remote northern provinces of Oudalan (OUD), Kossi (KOS), and Soum (SOM) experienced the highest increases among Burkina Faso's provinces with a plus of up to 1.24 ha/person. While their agricultural area considerably spread, they only observed a moderate population growth. Although, the per person agricultural area improved in most provinces during the period of investigation, several central and southwestern provinces observed either stable values or a decrease. Zoundwéogo (ZOU), Kouritenga (KOT), and Koulpélogo (KOP) are the provinces with the highest losses of up to 0.66 ha/person. These three provinces also exhibited the lowest numbers in agricultural expansion while their rural population numbers still rose by up to 240,000 people (ZOU) between 2001 and 2014 (compare Figure 6.9).
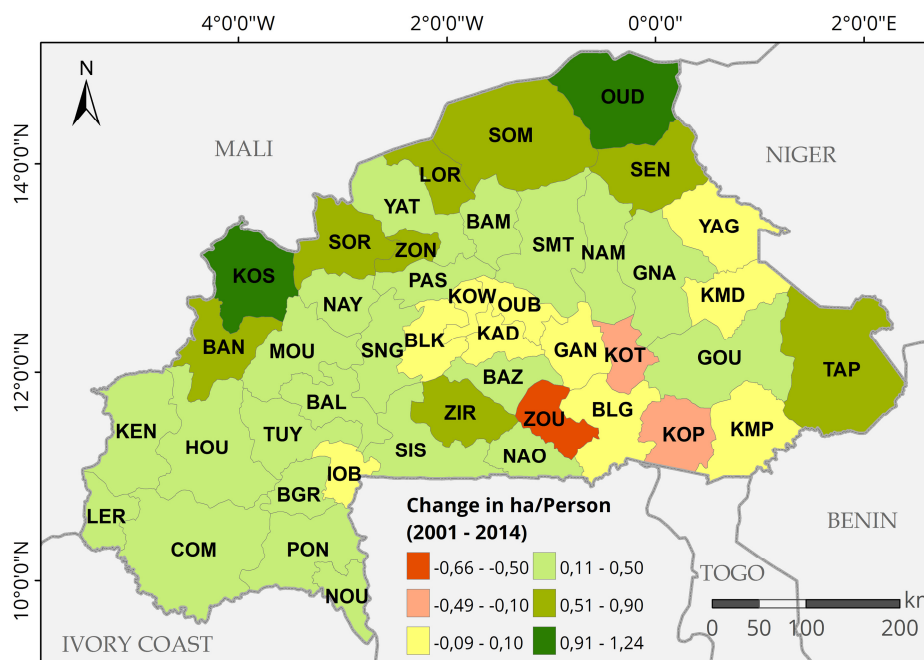
Figure 6.11: Provincial development of agricultural area per rural inhabitant in Burkina Faso between 2001 and 2014; province names are abbreviated by three letters for readability (modified after Knauer et al., 2017).

6.4.4 Impact of Agricultural Expansion on Protected Areas

The aim of this section is to show the development of agricultural area in two protected areas and their surroundings and to derive implications for the conservation (compare Figure 6.5). The first focus area consists of the Kaboré Tambi National Park (KTNP) founded in 1976 and the classified forest Nazinon which is attached to the KTNP in the northwest. They are situated about 60 km to the south of Burkina Faso's capital Ouagadougou (Figure 6.5 and Figure 6.12). At the beginning of the period of investigation in 2001, agricultural area within the boundaries of the protected areas was on a low level of 3% which was primarily due to some intrusions in the Nazinon forest. Relatively few agricultural activity could be observed in

the direct vicinity of the KTNP which is also a result of an obvious buffer zone around some areas of the park not included in the shapefile provided by the IGB (IGB, 2016). The agricultural area in 4 km distance and further is already covering more than 30% of the distance zones in 2001 since the KTNP lies in the central southern area of Burkina Faso where cultivation and population have a long history.

Figure 6.12: Development of agricultural area in the Kaboré Tambi National Park and its vicinity for the years 2001, 2007, and 2014 (top) and corresponding plots of the density of agricultural areas in the park and in 2 km distance zones surrounding it (bottom); location of this focus site in Burkina Faso is indicated in Figure 6.5 (modified after Knauer et al., 2017).

Until 2007, the agricultural activities within the protected areas stayed on the same low level but the surrounding distance zones experienced a rise of agricultural density with up to 43% in the 8–10 km zone. Between 2007 and 2014, a further increase by 8% can be observed for this zone and also the direct vicinity of the protected areas are becoming more and more cultivated. However, the KTNP itself remains unaffected by the agricultural expansion in its surroundings which indicates a proper management and protection of the park borders. Nonetheless, with rising population numbers and need for more farmland, the pressure on the park increases as well. In addition, this concentration of agricultural area in its direct vicinity increases the isolation of the KTNP and decreases natural patches in proximity to it. Such patches are important step stones for animal movement among habitat areas forming natural pathways and supporting genetic exchange between species populations.

The second focus area of this section is the classified forest of Tiogo which is located approximately 130 km in the west of Ouagadougou (Figure 6.5). This site is formally protected and in the direct vicinity of other areas with similar status of protection (Figure 6.13). In

comparison to the KTNP, the forest of Tiogo is not subject to a strict park management (Ouedraogo, 2014). In 2001, some agricultural activity can be observed in the northeastern part of the focus site with a so far minor coverage of 3% of the total park area. The agricultural area in the park vicinity is still on a moderate level with densities of up to 36% within the 4–6 km distance zone. Until 2007, the agriculture within the park boundaries increased considerably to 13% of the total area and intrudes further into the forest. All surrounding distance zones experienced a rise in agricultural density with up to 51% within the 8–10 km zone. Between 2007 and 2014, the agricultural area in Tiogo's northeast densified to 18% of the total protected area. Starting right at the park borders, dense agricultural area covers more than 50% in all distance zones. Even more seriously than for the KTNP, this situation in combination with insufficient protection resources highly endangers the remaining fragment of the Tiogo classified forest (Ouedraogo, 2014)
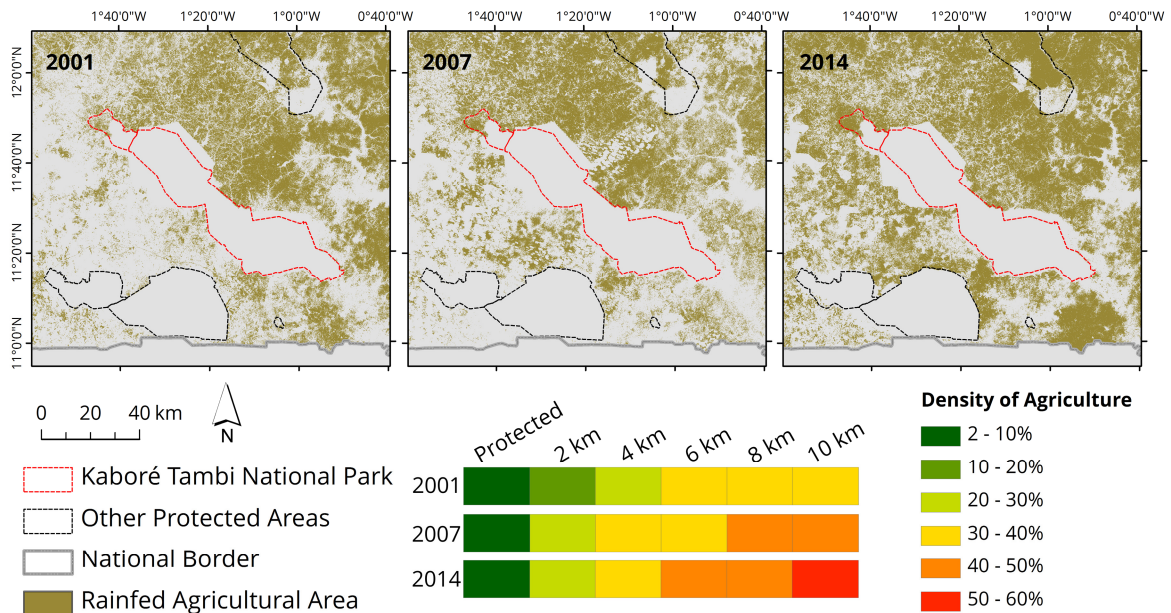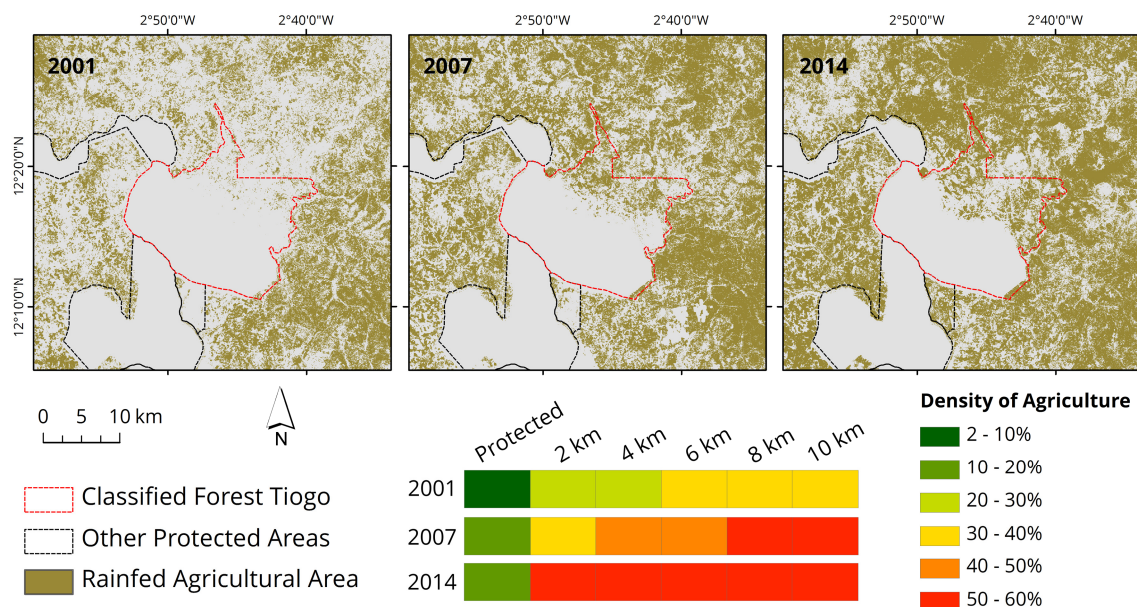

Figure 6.13: Development of agricultural area in the classified forest Tiogo and its vicinity for the years 2001, 2007, and 2014 (top) and corresponding plots of the density of agricultural areas in the forest and in 2 km distance zones surrounding it (bottom); location of this focus site in Burkina Faso is indicated in Figure 6.5 (modified after Knauer et al., 2017).

6.5 Discussion

6.5.1 Dense Time Series for the Classification of Agricultural Area

In this chapter, a novel methodology was applied in order to derive agricultural classifications in a high spatial resolution of 30 m for an entire country and three years. The ESTARFM framework, developed in this thesis, was used to generate comprehensive 8-day time series and these in combination with derived phenological metrics served as a basis for

106

the delineation of agricultural area in Burkina Faso. The overall classification accuracies of the three years (92% for 2001, 91% for 2007 and 91% for 2014) proved the suitability of the developed methodology. The use of dense time series with high spatial resolution instead of single remote sensing observations is a key feature of this approach. Frequent cloud coverage in West Africa considerably decreases the data availability of remote sensors (Fensholt et al., 2007) and mainly limits gap-free countrywide coverage to the dry season. However, during this period agricultural areas are harvested, have a very similar spectral signature to natural vegetation and are thus difficult to delineate during the dry season. This was determined by several studies who also showed the added value of multi-temporal input images or even the use of LSP features derived from time series for the classification of agricultural area in the region (Lambert et al., 2016; Liu et al., 2016; Zoungrana et al., 2015). However, not only the temporal component is important for a successful delineation but also a sufficient spatial resolution is crucial for this. The small-scale extensive farming systems in Burkina Faso with average field sizes below 1 ha require a high geometric detail for the discrimination of single fields. The novelty of the applied approach is that it brings together both, a high spatial and a high temporal resolution to optimally deal with the complex situation in West Africa for agricultural mapping. So far, existing region-wide LULC maps were only able to rely on one of the two components and thus either miss the spatial or the temporal resolution for a sufficiently accurate delineation of agricultural area. In addition, they mostly miss the thematic complexity of the region, i.e. do not cover the different relevant agricultural classes especially plantations. Thus, previous classifications are of limited use for detailed regional scale applications.

However, the partially low data availability from earlier Landsat generations such as Landsat-5 and -7 was also a limiting issue for the classification of multi-year countrywide agricultural area in this study. Although, the ESTARFM framework can significantly improve the number of usable scenes via an automated cloud-gap filling, there were not or not enough scenes available for some regions and years. In 2007, no Landsat-5 data were available for the southeast of Burkina Faso (tiles of path/row: 192/52, 193/52) and for several regions, the cloud coverage was on an exceptionally high level over the whole year. Subsequently, the results of this year were only used and presented here for local analyses. Anyway, the classification of 2014 has shown that this issue should not emerge for future applications of the developed approach since sufficient amounts of data are processed from the OLI sensor of Landsat-8. In addition, an improvement of data availability from Landsat-5 can still be expected with the completion of the Landsat Global Archive Consolidation (LGAC) mentioned in section 4.2. For future applications of the presented methodology, the Sentinel-2 satellites will also be an option and should improve input data availability (ESA, 2017).

Furthermore, it would be desirable for a repeated application of the methodology in the framework of a monitoring programme to automate the classification procedure. This could

also involve the automated generation or update of the reference database used for training and validation of the classification. If this could be implemented, the overall processing time from download of satellite scenes to the final maps would be improved significantly. However, such an operational monitoring framework is beyond the scope of the presented work which focused on the development of the fusion and classification approach and subsequently the delineation of agricultural area for certain years.

The mentioned reference data collection was a general challenge of the presented approach. A comprehensive reference database of LULC for multiple years for entire Burkina Faso is so far not available. Thus, these information were gathered from high resolution images in Google Earth on the basis of expert knowledge. Due to a common lack of multi-temporal datasets for a single year of interest, the discrimination of active cropland and crop-fallow rotation systems or even recently abandoned land can be challenging in Google Earth. Especially if only data from the dry season of the respective year is available, issues arise for the distinction between agricultural area and recently abandoned land. The still visible outline of a field may be misleading here and should result in some inaccuracies of the sampling. Nonetheless, the objective of this study was the delineation of agricultural area including fallow and intensive pasture without restricting the classification to harvested crop area only. This definition of agricultural area provides a valuable picture on the overall anthropogenic need of land for agricultural production in the country. However, if a suitable reference database would be available, the derived ESTARFM time series would have the potential to also distinguish harvested crop area from other types of agricultural area.

A comparison of statistics derived from the classification results to official statistics of agricultural area is difficult since these are either based on so called 'manual estimations' or are limited on harvested area (FAO, 2016). The FAO estimates an agricultural area of 105,700 km² in Burkina Faso for the year 2001 and 123,000 km² for the year 2013 while statistics for 2014 are currently not available (FAO, 2016). However, as mentioned in section 4.5, these values of agricultural area contain a constant number of 60,000 km² for 'permanent meadows and pastures' since 1984, when this value was estimated. In the rapidly changing environment of Burkina Faso, such a stable expanse of rangelands is of course highly unrealistic and not up to date. The FAO further estimates that 49,600 km² of primary crops were harvest in Burkina Faso in the year 2001 and 70,200 km² in 2013 with again missing numbers for the year 2014 (FAO, 2016). Although, these values are not directly comparable to the agricultural area statistics derived in this work, they indicate the strong expansion of agriculture in Burkina Faso during recent years, also visible in the results of this work.

6.5.2 Land Cover Conversion in the Context of Socio-Economic Development

For the considerable expansion of rainfed agricultural area, the rapid population growth seems to be the main driver. However, the development of plantations and irrigated

agricultural systems visible in the results of this work requires initial investments for which most Burkinabe lack the necessary capital. The growth of plantation area and the formation of distinct plantation outlines (compare section 6.4.2) indicate a considerable driving force of development and professionalization of tree cropping in Burkina Faso. And indeed, this change goes in line with the establishment of cashew nuts and mango businesses, which are predominantly Fairtrade, in Burkina Faso during the last decade. Since the year 2000, several external companies and organizations started investing in and promoting the formation of mango and cashew nut plantations. A primary example is the African Cashew initiative (ACi) (ComCashew, 2016) which was founded amongst others by the German Society for International Cooperation (GIZ). They started optimizing the entire value chain from cashew production, over the marketing in Europe to the distribution for several West African countries including Burkina Faso. In the mango production, professionalization was promoted by international projects and the establishment of local companies such as Fruiteq (Fruiteq, 2012; German Federal Ministry for Economic Cooperation and Development, 2016; van der Waal, 2011). They initiated the formation of farmer co-operatives, introduced trainings in mango cultivation, and improved the export paths and the Fairtrade marketing in Europe.
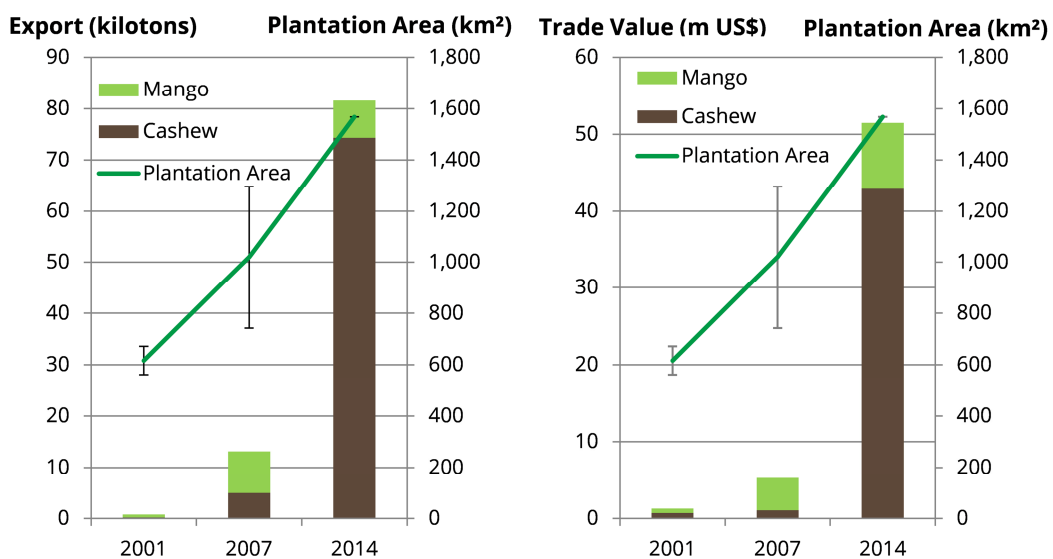
Figure 6.14: Total number in kilotons (left) and trade value (right) of exported mangos and cashew nuts from Burkina Faso and the total plantation area from the classification results for the years 2001, 2007, and 2014 (United Nations, 2016); the error bars determine the possible variation in plantation area due to data gaps in the respective classifications (modified after Knauer et al., 2017).

Although, mango and cashew nuts are not the only tree crops in Burkina Faso, others like citrus fruits or common nuts are clearly not responsible for the considerable growth in plantation area since their production statistics either stagnate or even decease over the last 14 years (FAO, 2016). Contrasting these crops, a rise in trade values and amount of exported

mangos and cashew nuts can be observed during recent years which reflects the success of these international efforts (Figure 6.14). In 2001, the amount of exported cashew nuts and mangos from Burkina Faso was on a rather low level accounting for 224 and 580 tons, respectively (United Nations, 2016). Towards the year 2014, the production of both crops rose rapidly to 74,303 and 7,256 tons. A similar development can be observed in the trade values increasing from 0.73 to 42.85 million US$ for cashew nuts and from 0.56 million to 8.54 million US$ for mangos (United Nations, 2016). Although, the growth of the cashew nut sector is even more outstanding than that of the mango business, a price increase (US$ per ton) for mangos from Burkina Faso can be observed. This development could be attributed to rising world market prices for mangos but might also indicate the improved quality of mangos from Burkina Faso and the success of Fairtrade marketing.

In addition to the remarkable expansion of plantation area, the increase in irrigated agricultural areas in Burkina Faso is also considerable (compare section 6.4.2). During the last decade, the government of Burkina Faso tried to promote the development of irrigation systems by investing considerable amounts of money within the framework of the National Programme for the Rural Sector (PNSR) (Burkina Faso, 2012). In addition to the governmental efforts, several international institutions and organizations have been supporting irrigated agriculture in local projects in Burkina Faso. A showcase project in this context is the above mentioned area around the Bagré dam located about 140 km southeast of Ouagadougou (Figure 6.8). For this region, the World Bank, the Global Water Initiative (GWI) and others are evolving irrigated agricultural systems and are advising and promoting smallholder farmers (IWMI, 2010; Sedogo and Bourgou, 2015; World Bank, 2010a). The considerable expansion of irrigated areas by about 300% to 37.28 km² in 2014 in the surroundings of the dam reflects the success of these efforts. Nonetheless, this number is still considerably lower than the estimated potential of irrigable land for the region of 299 km² (Sedogo and Bourgou, 2015).

6.5.3 Monitoring for Planning and Implementation of Measures

The links between agricultural expansion and the rapidly growing population of Burkina Faso are clearly visible from the results of this work. Between 2001 and 2014, the country's population numbers rose by about 40% in almost all provinces, while agricultural area almost doubled and spread into all outer and so far rather unused regions. Since 2001, the agricultural area per person in the rural regions of Burkina Faso experienced an overall increase from 0.63 ha to 0.88 ha. Anyway, several provinces with a dense cultivation in 2001 already show a decrease in agricultural area per rural person until 2014. Estimations foresee an end of possible agricultural expansion and the reaching of Burkina Faso's limits of arable land by the year 2030 (Mathys and Gardner, 2009). In addition, the total population is predicted to grow from currently about 17 million to more than 26 million people in 2030 and even to about 41 million in 2050 (FAO, 2016). Thus, it can be concluded that an overall positive

trend in agricultural area per person will not hold for a longer time. Furthermore, only the rural population was considered in the calculation of this factor since most of the farmers are smallholders and only cultivate for self-subsistence with rare access to the market (FAO, 2014). Nevertheless, a self-sustaining agriculture would be a key factor in Burkina Faso to improve food security and to resist fluctuating food prices during drought years. This however means that also the rising urban population needs to be taken into account in the calculation of such a factor and the result would be considerably lower numbers of agricultural area per person. In order to analyse regional migration and rural-urban development, the presented results in combination with additional statistical data could be valuable information for the planning of adaptation measures.

Nonetheless, the uncertainties of the used datasets of population numbers and also of the derived statistics of agricultural area have to be viewed critically. The population information are census data derived from household surveys in Burkina Faso (INSD, 2014). Such a survey in a developing country is always connected to a considerable amount of uncertainty and cannot be compared to the strict registrations of inhabitants conducted in industrialized countries. Although, the generally strong trend of population growth is undoubted, this fact may cause some inaccuracies in the conducted provincial analyses of agriculture and population. A further uncertainty in the underlying analysis is the annual variability of agricultural area depending on the amount of rainfall and the development of the rainy season. Although the numbers of agricultural area of each of the three years (2001, 2007 and 2014) are validated, it is unsure of what extent the area varies from year to year. If farmers are cultivating more area in a year of low rainfalls due to expected lower yields or if they might cultivate less in order to save resources could have an impact on the total agricultural area of a certain year. On the other hand, the amount of annual rainfall has a considerable influence on the phenology of a year and thus also on the successful delineation of agricultural area from semi-natural vegetation. Thus, an annual derivation of agricultural area over the period of investigation and a subsequent analysis of variation including rainfall data could increase the understanding of this influence factor and help in the interpretation of numbers of single years.

As indicated in section 6.4.4, the pressure on protected areas in Burkina Faso is constantly increasing due to the considerable population growth and agricultural expansion in their surroundings. Thus, some of the reserves with low protection measures have significantly decreased in size or even vanished completely. The densification of anthropogenic use through an increase of agricultural area and rural settlements in their vicinity endangers the suitability of these reserves as a habitat for a number of species. The preservation of a critical size and connectivity between natural habitats is crucial for many species to provide important habitat functions and to assure genetic exchange between populations and maintain biodiversity (Bennett and Saunders, 2010).

Since not enough wildlife rangers are available and accessibility of the forests is often difficult, it is important to involve the local communities around the protected areas if forests shall survive in their current state (Ouedraogo, 2014). The communities rely on the use of forests with cutting and collecting of wood for charcoal production or fires for cooking and try to use it for agriculture with a lack of alternative areas (Ouedraogo, 2014). Thus, a compromise has to be found to ensure the protection and regeneration of the forests and at the same time to guarantee the local population a sustainable living (Ouedraogo, 2014). This is currently under development on the basis of funds from the Forest Investment Program (FIP) of the Climate Investment Funds financed by the World Bank (World Bank, 2010b). Burkina Faso is one of eight pilot countries, given a 30 million dollar grant to reduce deforestation and promote sustainable forest management. The Tiogo classified forest, investigated in section 6.4.4 is one of six focus reserves of this program within Burkina Faso. The communities shall be involved in the management of the forest in order to protect it rather than exploit it. For this purpose, people shall learn how to grow crops amid the trees, to run nurseries or to implement sustainable rotation of forest grazing. Another potential driver of protection on the basis of this development is the income from carbon markets for which Burkina Faso is participating in the REDD+ program (Climate Investment Funds, 2013).

The further development of tourism and nature tourism in particular could also help with the protection of remnant forests by increasing the value of these areas for local communities. Entrance fees or the selling of local artworks and crafts to tourists could introduce alternative incomes to the ones exploiting the protected areas. However, although the general tourist numbers increased significantly during the beginning of the century from 126,000 people to 289.000 people in 2007, the numbers dropped again in recent years (191,000 people in 2014) (World Bank, 2016). Several factors such as political instabilities, terrorism or the occurrence of the Ebola virus in other countries of West Africa may have contributed to this negative development.

Against the background of population growth and the need to improve food production, the challenges for Burkina Faso to preserve the remaining protected areas are considerable. External funding will not guarantee the success of such an attempt rather the development of the country's economy and general infrastructure seems necessary. The results of this work indicate the suitability of the applied methodology to map and monitor the protection status of Burkina Faso's nature reserves. By a repeated application of the classification approach developed in this work for future years, the impact of Burkina Faso's efforts to preserve the remaining protected areas could be assessed.

Since Burkina Faso will soon reach its limits of arable land, there is a serious need to increase the sustainable and efficient use of existing agricultural areas. Anyway, as outlined in section 2.6, the cereal productivity of Burkina Faso only rose from about 0.4 tons per hectare and year to approximately 1.16 t/ha/year since the 1960s (FAO, 2016). In the same time, other

developing and thresholding countries like Brazil improved their cereal productivity from 1.34 t/ha/year to about 4.8 t/ha/year. For comparison, in Western Europe, productivity numbers of more than 7 t/ha/year are common. In several national and international projects (compare section 6.5.2), the development of efficient and more sustainable agricultural areas is promoted. These however mostly focus on geographically small areas and are not aiming at the majority of rainfed agricultural regions. Thus, a countrywide approach is necessary to support and develop the whole agricultural sector in Burkina Faso. This has been attempted by the government of Burkina Faso with considerable investments during the past decade. More than 10% of the national budget was assigned to the agricultural sector, especially in the framework of the Rural Development Strategy (SDR) (FAO, 2014). With this money, the production of staple crops was supported by subsidies on fertilizer, by distributing enhanced seeds and by improving and spreading of irrigation systems. Such efforts will primarily promote an intensification of existing agricultural areas and target at the increase of agricultural productivity. The promotion of local focus areas as the above mentioned Bagré Growth Pole Project is intended to attract private investment and trigger positive effects on the rest of Burkina Faso's economy (FAO, 2014). Anyway, the major part of national investments in the agricultural sector still goes into the production of cotton which is the primary agricultural export commodity in Burkina Faso. All in all, it is not sure yet if the described efforts have a significant impact on the broad agricultural community and on the general improvement of food security and sustainability in Burkina Faso. A repeated application of the developed approach could play a central role in the monitoring efforts from the government and external organizations to develop a sustainable land management that could furthermore also contribute to the above mentioned protection of natural reserves. On the basis of a regular monitoring, the country's focus areas and the development of the countrywide agricultural area could be surveyed. Thus, discrepancies between planning and implementation and potential hotspots of undesirable processes could be identified.

6.6 Summary and Conclusions

In this chapter, the expansion of agricultural area between 2001 and 2014 was quantified by the application of the previously developed ESTARFM framework. A high temporal (8-day) and high spatial (30 m) resolution time series was generated for the entire country of Burkina Faso comprising 17 Landsat tiles and three years. From the three synthetic time series, agricultural area was delineated in a 30m resolution for the whole country area (about 274,000 km²) and the target years 2001, 2007, and 2014. Until now, there are no other products available with such a temporal, spatial and thematic coverage and with a comparably high accuracy. Previous studies on the topic highlight the challenges in the delineation of agricultural area from remote sensing datasets in West Africa (Forkuor, 2014; Lambert et al.,

2016; Leroux et al., 2014; Vintrou et al., 2012b). Due to the predominantly small-scale rainfed subsistence farming and the mostly low productivity of fields, the discrimination of agricultural area from other LULC types is complicated. Despite these challenges, the overall accuracies of the presented classifications for the three years (92%, 91%, and 91% for 2001, 2007, and 2014 respectively) underline the suitability and well-functioning of the developed methodology. A major factor for the successful delineation of agricultural area in the region is the high temporal resolution. Thus, the whole phenological cycle of NDVI can be analysed to detect and use temporal differences between agricultural area and other classes.

In addition, the derived results give valuable insight in the development and professionalization of plantations, the establishment of irrigation systems, and the status and protection of natural reserves in Burkina Faso. Such information can be used by decision makers in order to check and quantify if recently introduced strategies and measures for the development of agricultural systems are having the planned effects. The developed methodology could further help with the evaluation of policies through mapping and monitoring of agricultural area in the future.

The results and analyses of this work indicate the major challenges of Burkina Faso regarding food production and food security which will further increase in the near future. The considerable expansion of agricultural area from 61,000 km² (2001) to 116,900 km² (2014) is rapidly consuming the last available land resources of the country spreading into all outer and partially rather unsuitable regions. This expansion is mostly driven by a likewise strongly increasing population. Since this population growth is expected to hold on in the future, also the expansion will continue if the productivity of agriculture will not be improved significantly. These trends are also increasing the pressure on the last natural reserves in Burkina Faso with illegal slash and burn practices and subsequent cultivation within park boundaries.

The dense, high spatial resolution time series generated with the ESTARFM framework also indicate its great potential for various other applications beyond the delineation of agricultural area. The combination of high temporal and high spatial resolution could also improve studies on land degradation, analysis of climate change impacts or changes in land surface phenology in West Africa.

7 Synthesis and Outlook

7.1 Summary and Accomplishment of Objectives

As outlined in the beginning of this work, West Africa is of major importance for global prospects due to its considerable population growth and the uncertainty of feeding these people against the background of low agricultural productivity in the region and climate change. With an expected number of 800 million people in 2050, West Africans will constitute a great part of the global population (FAO, 2016). However, most countries in the region are on a very low level of development which will impede their ability to cope with the negative effects of population growth, especially in terms of food security. So far, the agricultural productivity, i.e. the agricultural yield per hectare cultivated land, in West Africa is far behind other developing countries and was only slowly increasing in the past. This is primarily due to inefficient farming practices with small-scale manual farming, a low level of agricultural inputs, mechanization, fertilizer and pesticide application as well as irrigation. Since the slow increase in productivity is not sufficient to meet the food demand of a growing population, an expansion of agricultural area has taken place, especially during the last 20 years. The mapping and monitoring of this expansion is crucial in order to estimate its implications on agricultural production, livelihood of the people as well as environment. However, accurate information is missing due to insufficient statistical surveys. In addition, the delineation of agricultural area in West Africa is a very challenging topic for the application of remote sensing due to the described small-scale and inefficient farming practices as well as frequent cloud coverage decreasing data availability. These challenges have been highlighted by previous studies focusing on the delineation of agricultural area from remote sensing datasets in West Africa (Forkuor, 2014; Lambert et al., 2016; Leroux et al., 2014; Vintrou et al., 2012b). Thus, existing satellite based LULC products lack the necessary spatial, temporal or thematic coverage in the region for a comprehensive monitoring. Especially in Burkina Faso, in the center of West Africa, agricultural expansion is said to progress rapidly and is expected to reach the limits of arable land by the year 2030 (Mathys and Gardner, 2009).

The urgent necessity to monitor agricultural expansion in Burkina Faso and the difficulty of this task on the basis of remote sensing form the major objectives of the presented work. The first one (**I**) was to develop a suitable earth observation based methodology for the derivation of agricultural area in small-scale farming systems with high temporal variability at

high spatial resolution and national level. The second one (**II**) subsequently focused on the mapping, monitoring and analysis of agricultural expansion in Burkina Faso since the year 2001.

For the achievement of the **first objective (I)**, a **spatio-temporal data fusion framework** of medium and high spatial resolution remote sensing data was implemented. This way, the data availability of high resolution satellite data can be increased significantly in cloud-prone regions such as West Africa and consistent time series can be generated that allow for an accurate discrimination of agricultural areas from natural land cover types such as savannas and grasslands. For this purpose, the ESTARFM algorithm developed by Zhu et al. (2010) was employed and modified in order to make it applicable for the large-scale generation of time series in cloud-prone and heterogeneous landscapes. So far, this spatio-temporal data fusion approach was only applied for small study areas and in regions with proper input data availability. For the application on national level in West Africa, the ESTARFM framework was developed. First of all, an automatic filling of cloud gaps was included in this framework to be able to use even partially clouded Landsat scenes as input without propagating gaps to the fused time series. The developed filling procedure is an iterative approach generating ESTARFM predictions for the cloud gaps from other temporally close Landsat scenes. This implementation significantly increased the number of high spatial resolution input scenes which is an important asset to be able to use scenes that are temporally closer to the prediction date. As could be shown in this work, input scenes from the same vegetative season as the prediction date considerably increase the prediction accuracy and thus improve the consistency of the resulting time series. In addition, it was necessary to modify the selection of similar pixels used for the regression between Landsat and MODIS in order to account for spatial differences in the heterogeneity of the study area. The original approach included in ESTARFM using one global threshold for the entire study area does not account for spatial differences in the heterogeneity of land cover. If a study area is covered by several Landsat tiles with different landscape types and classes, a single threshold would not be sufficient for similar pixel selection. The newly developed selection approach takes regional differences into account, independent of the size of the study area. Furthermore, the ESTARFM framework considerably increases processing speed of the fusion through automation of the processing chain as well as parallelization of the data fusion. This way, the fusion process can be distributed on the number of allocated CPU cores and is approximately n times faster than the original algorithm where n is the number of assigned CPU cores. In order to accomplish the generation of time series on national level and for several years in a reasonable amount of time, this was a crucial improvement. An additional asset was the inclusion of quality layers that allow for an indication if the predicted images are based on stable regressions or not can thus indicate regions with potential prediction errors.

Especially when working on large scales, where a thorough analysis of single predictions is not feasible, these layers can give a quick indication of the well-functioning of the data fusion.

The framework was applied and its accuracy was assessed in a focus area in the border region between Burkina Faso and Ghana. The results of this use case showed the high suitability of the ESTARFM framework for the generation of time series on large spatial scales in such a challenging heterogeneous and cloud-prone region. Uncertainties remain when the input data availability is low and shoulder scenes for prediction have to be taken from another phenological season. Abrupt events such as fires or floods can also decrease the quality of prediction. Nonetheless, the overall accuracies for the predicted time series were on a high level and highlighted the well-functioning of the developed methodology. In addition, the time series derived with the ESTARFM framework indicated the added value for the delineation of agricultural area from different vegetation types. The produced phenological signatures with their high temporal resolution showed an improved discriminability in comparison to the ones from Landsat time series while at the same time maintaining a high spatial resolution.

All in all, the developed methodology has the potential to improve the mapping of dynamic land processes in West Africa and Burkina Faso in particular since it addresses the major prevailing difficulties for a successful performance of this task. Thus, the first objective focusing on the development of such a methodology could be accomplished.

In order to accomplish the **second objective (II)**, the **mapping, quantification and analysis of agricultural expansion** in Burkina Faso since the year 2001, the ESTARFM framework was applied to generate time series for the entire country. A high temporal (8-day) and high spatial (30 m) resolution NDVI time series for Burkina Faso comprising 17 Landsat tiles and the years 2001, 2007 and 2014 was produced. For this purpose, more than 500 Landsat scenes and 3000 MODIS scenes were processed with the automated framework. From the generated time series, seasonal and phenological metrics such as the start of season or the mean annual NDVI were extracted highlighting different characteristics of the prevailing LULC classes. All together with the single NDVI time steps, these metrics were used as the basis for the delineation of agricultural area in a 30 m resolution for the entire country area of about 274,000 km² and for the three focus years 2001, 2007, and 2014. This was conducted with the random forest classification algorithm and a comprehensive randomly sampled reference dataset generated based on very high resolution images from Google Earth and expert knowledge of the region. With this approach, overall classification accuracies of 92% for 2001, 91% for 2007, and 91% for 2014 could be achieved. This underlines the suitability and well-functioning of the developed methodology despite the prevailing challenges for this mapping. Until now, there are no other products available with such a temporal, spatial and thematic coverage and with a comparably high accuracy.

The results show a rapid expansion of agricultural area from 61,000 km² (2001) to 116,900 km² (2014) which is consuming the last available land resources of the country. At the beginning of the century, the dense agricultural area focuses on the central parts of Burkina Faso around the capital Ouagadougou and the areas in its southeast towards the border of Ghana and Togo. Until 2014, the agricultural areas also spread in the more remote regions of the northern Sahel, the moist southwestern regions, and the East bordering Niger and Nigeria. This expansion is mostly driven by a likewise strongly increasing population in Burkina Faso. While the population grew by about 40% in almost all provinces over the period of investigation, the agricultural area nearly doubled and spread over the entire country. This strong agricultural expansion currently results in a rise of agricultural area per rural inhabitant from 0.63 ha to 0.88 ha. However, especially in the central and southwestern provinces, which were already densely cultivated in 2001, almost no increase or a decrease in per capita agricultural area was detected. Such a negative trend can be expected for more provinces in the future when they reach their limits of arable land while the rural population continues growing.

In addition, the plantation area exhibits a strong increase from 561 km² to 1,568 km² (2001-2014) and the formation of two major plantation regions in the southwest of Burkina Faso: one in the provinces Kénédougou and Léraba and one in the Comoé and Poni. This expansion is driven by the professionalization of plantations with external investments and the establishment of a Fairtrade business mainly of cashew nuts and mangos. With a rising number of water reservoirs, also the irrigated agricultural area expanded considerably from 78 km² to 345 km². This increase primarily result from the efforts of the government of Burkina Faso and international organizations to develop irrigation systems in the country. These findings can be used by decision makers to check and quantify if recently introduced strategies and measures for the development of agricultural systems are having the planned effects. The developed methodology could further help with the evaluation of policies through mapping and monitoring of agricultural area in the future.

The presented agricultural expansion is also increasing the pressure on the last natural reserves in Burkina Faso. The analysis of this work could show the current situation in the surroundings of selected parks and emphasizes the importance of a working park management and protection as well as the necessity to include local communities in this progress. Where a protection is not functioning properly, illegal slash and burn practices and subsequent cultivation is taking place within the park boundaries due to an increasing need for agricultural area.

On the basis of the developed methodology an accurate mapping, quantification and analysis of agricultural expansion in Burkina Faso was conducted. Thus, the second objective could be accomplished as well.

7.2 Challenges and Future Opportunities

The presented methodology for the development of high spatial and temporal resolution remote sensing time series as well as the subsequent derivation of agricultural area provides numerous opportunities for future applications. In addition, some points in the methodology could be further developed depending on the requirements of the planned application.

The major issue determining the quality of spatio-temporal data fusion is the availability of input data. Although, this data availability could be increased significantly within the presented approach, there is still room for improvements. One idea to increase the number of input observations could be the use of tiles of different Landsat paths jointly in the ESTARFM framework. In the current implementation of the framework, each tile is treated individually for the generation of time series. However, in higher latitudes, the overlapping areas between Landsat paths are considerably bigger than in the focus region of this work and the use of different Landsat paths for the time series generation of a single location (i.e. pixel) could improve the fusion quality. Furthermore, such a joint use of input paths could also increase consistency and reduce discrepancies in the overlapping areas which can be an issue due to the different acquisition dates of input paths during certain periods of the time series. Nonetheless, such a development is not a trivial one in a data fusion approach like ESTARFM which is taking neighboring pixels into account and it is not of high value for regions closer to the equator having only small overlap in Landsat paths.

Furthermore, it would be desirable for a repeated mapping of agricultural area in the framework of a monitoring programme to automate the classification procedure. This could also involve the automated generation or update of the reference database used for training and validation of the classification. If this could be implemented, the overall processing time from download of satellite scenes to the final maps would be improved significantly.

The use of ESA's Sentinel satellites offers great opportunities for current and future large scale mapping applications. Since March 2017, the Sentinel-2 program now offers a 2-satellite comprises a constellation of two satellites, it and now has a temporal coverage of up to two days with a maximum spatial resolution of 10 m. This will significantly improve data availability for greater spatial regions and time series approaches even in cloud-prone areas. However, in such cloud-prone regions time series of these sensors will still have considerable temporal gaps. The ESTARFM framework could be used to fill these gaps and provide a consistent high spatial and temporal resolution time series. For the analysis of longer periods of time, ESTARFM time series derived from MODIS and Landsat covering spans periods prior to 2015 (the start of the first Sentinel-2 satellite) could be used jointly with Sentinel-2 time series thereafter this date.

The possibility to use gap-free, high spatial and high temporal resolution time series also provides new application opportunities for future research. The first and obvious one would

be the delineation of agricultural area for entire West Africa. This application could indicate national discrepancies in the expansion of agricultural area and would allow for a complete assessment of the state of West African agriculture. It would be particularly interesting to test the capability of the presented approach to identify plantations in West Africa's south where plantation area is considerably bigger. As outlined in section 3.1, applications of remote sensing data in this region are sparse, which primarily results from the even higher level of cloud cover in comparison to the Sahel and Sudanian area (Knauer et al., 2014). The application of the ESTARFM framework in southern West Africa would be a considerable challenge but a successful use would also be a great and valuable contribution to current research activities.

The accurate delineation of agricultural area in West Africa poses an important prerequisite for a more sophisticated assessment of yields as presented in section 3.1.3. If forecasts of annual food availability shall have a significant effect on the timely management and planning of food security and food aid, it is necessary to know the exact size and distribution of cultivated areas. In addition, consistent time series of land surface phenology can also significantly contribute to an improvement of yield prediction models. If proper in-situ observations of phenological events such as germination, tillering or flowering are available, these could be used to connect them with LSP metrics derived from high temporal and spatial resolution time series generated with the ESTARFM framework. Not only yield assessment could benefit from such analyses, but also large-scale investigations of plant phenology in West Africa and its relation to regional weather and climate change.

Finally, the application of consistent large-scale time series in high spatial and temporal resolution could promote the ongoing discussion about degradation and greening in the Sahel of West Africa. As outlined in section 3.1.4, this is still one of the key research topics in the region. The high spatial resolution of the ESTARFM time series would allow the connection to local analyses of degradation but at the same time a coverage of the whole Sahel area of West Africa. Thus, a more comprehensive analysis of degradation could be conducted for West Africa's Sahel in order to provide new insight in this complex topic.

These are just application examples for the region of West Africa but numerous research questions in other parts of the world could greatly benefit from the adaptation of spatio-temporal data fusion with the ESTARFM framework. Especially in the more cloud-prone areas, where remote sensing based research is complicated due to a lack of data, the framework could advance research.

The results and analyses of this work indicate the major challenges of Burkina Faso and West Africa in general regarding food production and food security. With a constantly growing population, the expansion of agricultural area will most likely continue until the limits of arable land are reached and no further enlargement is possible, which has been estimated to happen by the year 2030 for Burkina Faso. This is also the approximate time frame which is

left to significantly increase the agricultural productivity of the country via an intensification of farming on already cultivated land. If this cannot be accomplished in the near future, the booming population growth will result in severe food crises which could be amplified by a progressing unfavourable climate change.

8 References

Adelabu, S., Mutanga, O., Adam, E., 2015. Testing the reliability and stability of the internal accuracy assessment of random forest for classifying tree defoliation levels using different validation methods. *Geocarto International*. 30, 810–821.

African Statistical Coordination Committee, 2016. *African Statistical Year Book*.

AGRISTAT, 2006. Statistiques sur l'Agriculture et l'Alimentation du Burkina Faso [WWW Document]. URL http://agristat.bf.tripod.com/ (accessed 6.3.16).

Amoros-Lopez, J., Gomez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., Camps-Valls, G., 2013. Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *International Journal of Applied Earth Observation and Geoinformation*. 23, 132–141.

Anyamba, A., Tucker, C.J., 2005. Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. *Journal of Arid Environments*. 63, 596–614.

Bai, Z.G., Dent, D.L., Olsson, L., Schaepman, M.E., 2008. Proxy global assessment of land degradation. *Soil Use and Management*. 24, 223–234.

Bartholomé, E., 1988. Radiometric measurements and crop yield forecasting. Some observations over millet and sorghum experimental plots in Mali. *International Journal of Remote Sensing*. 9, 1539–1552.

Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*. 26, 1959–1977.

Bégué, A., Vintrou, E., Ruelland, D., Claden, M., Dessay, N., 2011. Can a 25-year trend in Soudano-Sahelian vegetation dynamics be interpreted in terms of land use change? A remote sensing approach. *Global Environmental Change*. 21, 413–420.

Bégué, A., Vintrou, E., Saad, A., Hiernaux, P., 2014. Differences between cropland and rangeland MODIS phenology (start-of-season) in Mali. *International Journal of Applied Earth Observation and Geoinformation*. 31, 167–170.

Bennett, A.F., Saunders, D.A., 2010. Habitat Fragmentation and Landscape Change. *Conservation Biology for all* 88–104.

8 References

Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., Chanussot, J., 2012. *Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches* 5.

Bisquert, M., Bégué, A., Poncelet, P., Teisseire, M., 2014. Evaluation of Fusion Methods for Crop Monitoring Purposes, in: *IGARSS 2014*. pp. 1429–1432.

Bliefernicht, J., Kunstmann, H., Hingerl, L., Rummler, T., Andresen, S., Mauder, M., Steinbrecher, R., Frieß, R., Gochis, D., Gessner, U., Quensah, E., Awotuse, A., Neidl, F., Jahn, C., Barry, B., 2013. Field and simulation experiments for investigating regional land-atmosphere interactions in West Africa: experimental setup and first results. *Climate and Land Surface Changes in Hydrology*. 359, 226–232.

BMBF, 2017. West African Science Service Center on Climate Change and Adapted Land Use [WWW Document]. URL www.wascal.org (accessed 2.21.17).

Bobée, C., Ottlé, C., Maignan, F., de Noblet-Ducoudré, N., Maugis, P., Lézine, A.-M., Ndiaye, M., 2012. Analysis of vegetation seasonality in Sahelian environments using MODIS LAI, in association with land cover and rainfall. *Journal of Arid Environments*. 84, 38–50.

Boko, M., Niang, I., Nyong, A., Vogel, C., Githeko, A., Medany, M., Osman-Elasha, B., Tabo, R., Yanda, P., 2007. Africa, in: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E. (Eds.), *Climate Change 2007*: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp. 433–467.

Bontemps, S., Defourny, P., Bogaert, E. Van, Arino, O., Kalogirou, V., Perez, J.R., 2011. GLOBCOVER 2009 *Products Description and Validation Report*. Louvain-la-Neuve, Belgium.

Boschetti, L., Roy, D., Hoffmann, A., Humber, M., 2013. MODIS Collection 5.1 Burned Area Product - MCD45 - User Guide.

Boschetti, M., Nutini, F., Brivio, P.A., Bartholomé, E., Stroppiana, D., Hoscilo, A., 2013. Identification of environmental anomaly hot spots in West Africa from time series of NDVI and rainfall. *ISPRS Journal of Photogrammetry and Remote Sensing*. 78, 26–40.

Bozzini, C., Maselli, F., 2002. Analysis of Multitemporal NDVI Data for Crop Yield Forecasting in the Sahel. *International Journal of Remote Sensing*. 17, 53–60.

Brandt, M., Mbow, C., Diouf, A.A., Verger, A., Samimi, C., Fensholt, R., 2015. Ground- and satellite-based evidence of the biophysical mechanisms behind the greening Sahel. *Gobal Change Biology*. 21, 1610–1620.

Brandt, M., Romankiewicz, C., Spiekermann, R., Samimi, C., 2014a. Environmental change in time series – An interdisciplinary study in the Sahel of Mali and Senegal. *Journal of Arid Environments*. 105, 52–63.

Brandt, M., Verger, A., Diouf, A., Baret, F., Samimi, C., 2014b. Local Vegetation Trends in the Sahel of Mali and Senegal Using Long Time Series FAPAR Satellite Products and Field Measurement (1982–2010). *Remote Sensing*. 6, 2408–2434.

Breiman, L., 2001. Random forests. *Machine Learning*. 45, 5–32.

Bréon, F.M., Vermote, E., 2012. Correction of MODIS surface reflectance time series for BRDF effects. *Remote Sensing of Environment*. 125, 1–9.

Brown, M.E., de Beurs, K.M., 2008. Evaluation of multi-sensor semi-arid crop season parameters based on NDVI and rainfall. *Remote Sensing of Environment*. 112, 2261–2271.

Brown, M.E., de Beurs, K.M., Vrieling, A., 2010. The response of African land surface phenology to large scale climate oscillations. *Remote Sensing of Environment*. 114, 2286–2296.

Brown, M.E., Hintermann, B., Higgins, N., 2009. Markets, climate change, and food security in West Africa. *Environmental Science & Technology*. 43, 8016–20.

Büdel, J., 1981. Klima-Geomorphologie. Schweizerbart Science Publishers, Stuttgart, Germany.

Burkina Faso, 2012. Programme national du secteur rural 2011-2015.

Butt, B., Turner, M.D., Singh, A., Brottem, L., 2011. Use of MODIS NDVI to evaluate changing latitudinal gradients of rangeland phenology in Sudano-Sahelian West Africa. *Remote Sensing of Environment*. 115, 3367–3376.

Callo-Concha, D., Gaiser, T., Webber, H., Tischbein, B., Müller, M., Ewert, F., 2013. Farming in the West African Sudan Savanna: Insights in the context of climate change. *African Journal of Agricultural Research*. 8, 4693–4705.

Capecchi, V., Crisci, A., Lorenzo, G., Maselli, F., Vignaroli, P., 2008. Analysis of NDVI trends and their climatic origin in the Sahel 1986–2000. *Geocarto International*. 23, 297–310.

Chen, B., Huang, B., Xu, B., 2015. Comparison of Spatiotemporal Fusion Models: A Review. *Remote Sensing*. 7, 1798–1835.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*. 103, 7–27.

CIESIN, 2016. Gridded Population of the World, Version 4 (GPWv4): Population Density Adjusted to Match 2015 Revision UN WPP Country Totals.

CILSS, 2016. Landscapes of West Africa - A Window on a Changing World.

Clauss, K., Yan, H., Kuenzer, C., 2016. Mapping Paddy Rice in China in 2002, 2005, 2010 and 2014 with MODIS Time Series. *Remote Sensing*. 8, 1–22.

Climate Investment Funds, 2013. Gazetted Forests Participatory Management Project for REDD+ (PGFC/REDD+) [WWW Document]. URL https://www-cif.climateinvestmentfunds.org/projects/gazetted-forests-participatory-management-project-redd-pgfcredd (accessed 11.14.16).

ComCashew, 2016. African Cashew initiative [WWW Document]. URL http://www.africancashewinitiative.org/ (accessed 11.13.16).

Congalton, R.G., Green, K., 2009. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd ed. CRC Press, Boka Raton, FL, USA.

Dardel, C., Kergoat, L., Hiernaux, P., Mougin, E., Grippa, M., Tucker, C.J., 2014. Re-greening Sahel: 30 years of remote sensing data and field observations (Mali, Niger). *Remote Sensing of Environment*. 140, 350–364.

de Beurs, K.M., Henebry, G.M., 2010. Phenological Research. Springer Netherlands, Dordrecht.

de Beurs, K.M., Henebry, G.M., 2010. Spatio-Temporal Statistical Methods for Modelling Land Surface Phenology, in: Hudson, I.L., Keatley, M.R. (Eds.), *Phenological Research*. Springer Netherlands, Dordrecht, pp. 177–208.

de Beurs, K.M., Henebry, G.M., 2004. Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment*. 89, 497–509.

Dent, D.L., Bai, Z.G., Schaepman, M.E., Olsson, L., 2009. Response to Wessels: Comments on "Proxy global assessment of land degradation." *Soil Use and Management*. 25, 93–97.

Descroix, L., Laurent, J.-P., Vauclin, M., Amogu, O., Boubkraoui, S., Ibrahim, B., Galle, S., Cappelaere, B., Bousquet, S., Mamadou, I., Le Breton, E., Lebel, T., Quantin, G., Ramier, D., Boulain, N., 2012. Experimental evidence of deep infiltration under sandy flats and gullies in the Sahel. *Journal of Hydrology*. 424–425, 1–15.

Descroix, L., Mahé, G., Lebel, T., Favreau, G., Galle, S., Gautier, E., Olivry, J.-C., Albergel, J., Amogu, O., Cappelaere, B., Dessouassi, R., Diedhiou, A., Le Breton, E., Mamadou, I., Sighomnou, D., 2009. Spatio-temporal variability of hydrological regimes around the boundaries between Sahelian and Sudanian areas of West Africa: A synthesis. *Journal of Hydrology*. 375, 90–102.

ECOWAS, 2017. ECOWAS Member states [WWW Document]. URL http://www.ecowas.int/member-states/ (accessed 2.1.17).

Eklundh, L., Jönsson, P., 2015. TIMESAT 3.2 with parallel processing Software Manual.

Eklundh, L., Olsson, L., 2003. Vegetation index trends for the African Sahel 1982–1999. *Geophysical Research Letters*. 30, 1430.

Emelyanova, I. V., McVicar, T.R., Van Niel, T.G., Li, L.T., van Dijk, A.I.J.M., 2013. Assessing the accuracy of blending Landsat–MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sensing of Environment. 133, 193–209*.

ESA, 2017. Introducing Sentinel-2 [WWW Document]. URL http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2 (accessed 1.20.17).

ESA, 2016a. ESA Climate Change Initiative Land Cover [WWW Document]. URL http://www.esa-landcover-cci.org/ (accessed 7.13.16).

ESA, 2016b. Land Cover CCI: Product User Guide.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenbock, H., Muller, A., Dech, S., 2013. Urban Footprint Processor - Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. *IEEE Geoscience and Remote Sensing Letters*. 10, 1617–1621.

FAO, 2017. Land Cover Classification System (LCCS): Classification Concepts and User Manual [WWW Document]. URL http://www.fao.org/docrep/003/x0596e/x0596e00.HTM (accessed 1.20.17).

FAO, 2016. FAOSTAT [WWW Document]. URL http://faostat3.fao.org/faostat-gateway/go/to/home/E

FAO, 2014. Burkina Faso - Country Fact Sheet on Food and Agriculture Policy Trends.

FAO, 2009. Harmonized World Soil Database.

Fensholt, R., Anyamba, A., Stisen, S., Sandholt, I., Pak, E., Small, J., 2007. Comparisons of Compositing Period Length for Vegetation Index Data from Polar-orbiting and Geostationary Satellites for the Cloud-prone Region of West Africa. *Photogrammetric Engineering & Remote Sensing*. 73, 297–309.

Fensholt, R., Rasmussen, K., Kaspersen, P.S., Huber, S., Horion, S., Swinnen, E., 2013. Assessing Land Degradation/Recovery in the African Sahel from Long-Term Earth Observation Based Primary Productivity and Precipitation Relationships. *Remote Sensing*. 5, 664–686.

Fensholt, R., Sandholt, I., Proud, S.R., Stisen, S., Rasmussen, M.O., 2010. Assessment of MODIS sun-sensor geometry variations effect on observed NDVI using MSG SEVIRI geostationary data. *International Journal of Remote Sensing*. 31, 6163–6187.

Fensholt, R., Sandholt, I., Rasmussen, M.S., Stisen, S., Diouf, A., Schultz, M., 2006. Evaluation of satellite based primary production modelling in the semi-arid Sahel. *Remote Sensing of Environment*. 105, 173–188.

Foley, J.A., DeFries, R.S., Asner, G.P., Barford, C., Bonan, G.B., Carpenter, S., Chapin, F.S., Coe, M.T., Daily, G.C., Gibss, H.K., Helkowksi, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, C., Ramankutty, N., Snyder, P.K., 2005. Global Consequences of Land Use. *Science*. 309, 570–574.

Forkuor, G., 2014. Agricultural Land Use Mapping in West Africa Using Multi-sensor Satellite Imagery. University of Wuerzburg.

Forkuor, G., Conrad, C., Thiel, M., Landmann, T., Barry, B., 2015. Evaluating the sequential masking classification approach for improving crop discrimination in the Sudanian Savanna of West Africa. *Computers and Electronics in Agriculture*. 118, 380–389.

Forkuor, G., Conrad, C., Thiel, M., Ullmann, T., Zoungrana, E., 2014. Integration of optical and synthetic aperture radar imagery for improving crop mapping in northwestern Benin, West Africa. *Remote Sensing*. 6, 6472–6499.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*. 114, 168–182.

Fritz, S., See, L., Rembold, F., 2010. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *International Journal of Remote Sensing*. 31, 2237–2256.

Fritz, S., See, L., You, L., Justice, C., Becker-Reshef, I., Bydekerke, L., Cumani, R., Defourny, P., Erb, K., Foley, J., Gilliams, S., Gong, P., Hansen, M., Hertel, T., Herold, M., Herrero, M., Kayitakire, F., Latham, J., Leo, O., McCallum, I., Obersteiner, M., Ramankutty, N., Rocha, J., Tang, H., Thornton, P., Vancutsem, C., van der Velde, M., Wood, S., Woodcock, C., 2013. The Need for Improved Maps of Global Cropland. *Eos Transactions American Geophysical Union*. 94, 31–32.

Fruiteq, 2012. Fruit du commerce équitable [WWW Document]. URL http://www.fruiteq.com/ (accessed 11.13.16).

Fu, D., Chen, B., Wang, J., Zhu, X., Hilker, T., 2013. An Improved Image Fusion Approach Based on Enhanced Spatial and Temporal the Adaptive Reflectance Fusion Model. *Remote Sensing*. 5, 6346–6360.

Fuller, D.O., 1998. Trends in NDVI time series and their relation to rangeland and crop production in Senegal, 1987-1993. *International Journal of Remote Sensing*. 19, 2013–2018.

Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., Yang, Y., 2015. Fusing Landsat and MODIS data for vegetation monitoring. *IEEE Transactions on Geoscience and Remote Sensing*. 47–60.

Gao, F., Masek, J.G., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*. 44, 2207–2218.

Gao, F., Wang, P., Masek, J., 2013. Integrating remote sensing data from multiple optical sensors for ecological and crop condition monitoring, in: Gao, W., Jackson, T.J., Wang, J., Chang, N.-B. (Eds.), *Remote Sensing and Modeling of Ecosystems for Sustainability X*. p. 886903.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N. V, Myneni, R., Knyazikhin, Y., Yang, W., 2008. Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. *Journal of Geophysical Research*. 113.

Gärtner, P., Förster, M., Kleinschmit, B., 2016. The benefit of synthetically generated RapidEye and Landsat 8 data fusion time series for riparian forest disturbance monitoring. *Remote Sensing of Environment*. 177, 237–247.

German Federal Ministry for Economic Cooperation and Development, 2016. Organic mangoes from Burkina Faso [WWW Document]. URL https://www.developpp.de/en/content/organic-mangoes-burkina-faso (accessed 11.13.16).

Gessner, U., Bliefernicht, J., Rahmann, M., Dech, S., 2012. Land Cover Maps for Regional Climate Modelling in West Africa – A comparison of Datasets, in: Perakis, K., Moysiadis, A. (Eds.), *Annual EARSeL Symposium 2012*, 21.-25. Mai 2012. pp. 388–397.

Gessner, U., Knauer, K., Kuenzer, C., Dech, S., 2015a. Land Surface Phenology in a West African Savanna: Impact of Land Use, Land Cover and Fire, in: Kuenzer, C., Dech, S., Wagner, W. (Eds.), *Remote Sensing Time Series*. Springer International Publishing, pp. 203–223.

Gessner, U., Machwitz, M., Esch, T., Tillack, A., Naeimi, V., Kuenzer, C., Dech, S., 2015b. Multi-sensor mapping of West African land cover using MODIS, ASAR and TanDEM-X/TerraSAR-X data. *Remote Sensing of Environment*. 164, 282–297.

Gessner, U., Niklaus, M., Kuenzer, C., Dech, S., 2013. Intercomparison of Leaf Area Index Products for a Gradient of Sub-Humid to Arid Environments in West Africa. *Remote Sensing*. 5, 1235–1257.

Gevaert, C.M., García-Haro, F.J., 2015. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment*. 156, 34–44.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. *Pattern Recognition Letters*. 27, 294–300.

Gonsamo, A., Chen, J.M., Price, D.T., Kurz, W.A., Wu, C., 2012. Land surface phenology from optical satellite measurement and CO2 eddy covariance technique. *Journal of Geophysical Research*. 117.

Gonzalez, P., 2001. Desertification and a shift of forest species in the West African Sahel. *Climate Research* 17, 217–228.

Gonzalez, P., Sy, H., Tucker, C.J., 2004. Local knowledge and remote sensing of forest biodiversity and forest carbon across the Sahel, in: Lykke, A.M., Kirkebjerg Due, M., Kristensen, M., Nielsen, I. (Eds.), *Proceedings of the 16th Danish Sahel Workshop*. Sahel-Sudan Environmental Research Initiative (SEREIN), University of Copehagen, Denmark, pp. 5–6.

Groten, S.M.E., 1993. NDVI — crop monitoring and early yield assessment of Burkina Faso. *International Journal of Remote Sensing*. 14, 1495–1515.

Guan, K., Medvigy, D., Wood, E.F., Caylor, K.K., Li, S., Jeong, S., 2014. Deriving Vegetation Phenological Time and Trajectory Information Over Africa Using SEVIRI Daily LAI. *IEEE Transactions on Geoscience and Remote Sensing*. 52, 1113–1130.

Hayward, D., Oguntoyinbo, J., 1987. Climatology of West Africa. Rowman & Littlefield Publishers.

Hein, L., 2006. The impacts of grazing and rainfall variability on the dynamics of a Sahelian rangeland. *Journal of Arid Environments*. 64, 488–504.

Hein, L., De Ridder, N., 2006. Desertification in the Sahel: a reinterpretation. *Gobal Change Biology*. 12, 751–758.

Hein, L., De Ridder, N., Hiernaux, P., Leemans, R., de Wit, A., Schaepman, M.E., 2011. Desertification in the Sahel: Towards better accounting for ecosystem dynamics in the interpretation of remote sensing images. *Journal of Arid Environments*. 75, 1164–1172.

Herrmann, S.M., Anyamba, A., Tucker, C.J., 2005. Recent trends in vegetation dynamics in the African Sahel and their relationship to climate. *Global Environmental Change*. 15, 394–404.

Herrmann, S.M., Tappan, G.G., 2013. Vegetation impoverishment despite greening: A case study from central Senegal. *Journal of Arid Environments*. 90, 55–66.

Heumann, B.W., Seaquist, J.W., Eklundh, L., Jönsson, P., 2007. AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. *Remote Sensing of Environment*. 108, 385–392.

Hickler, T., Eklundh, L., Seaquist, J.W., Smith, B., Ardö, J., Olsson, L., Sykes, M.T., Sjöström, M., 2005. Precipitation controls Sahel greening trend. *Geophysical Research Letters*. 32, L21415.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*. 25, 1965–1978.

Hilker, T., Wulder, M. a., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., White, J.C., 2009. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment*. 113, 1613–1627.

Huang, B., Song, H., 2012. Spatiotemporal Reflectance Fusion via Sparse Representation. *IEEE Transactions on Geoscience and Remote Sensing*. 50, 3707–3716.

Huber, S., Fensholt, R., 2011. Analysis of teleconnections between AVHRR-based sea surface temperature and vegetation productivity in the semi-arid Sahel. *Remote Sensing of Environment*. 115, 3276–3285.

IGB, 2016. Institut Géographique du Burkina Faso - Offical Homepage [WWW Document]. URL http://www.igb.bf/ (accessed 12.15.16).

INSD, 2014. Annuaire statistique 2013.

IUSS Working Group WRB, 2015. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps, *World Soil Resources Reports No. 106*. Rome.

IWMI, 2010. Burkina Faso National Consultation.

Jalloh, A., Rhodes, E.R., Kollo, I., Roy-Macauley, H., Sereme, P., 2011. Nature and management of the soils in West and Central Africa: A review to inform farming systems research and development in the region. Dakar, Senegal.

Jamali, S., Seaquist, J., Eklundh, L., Ardö, J., 2014. Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel. *Remote Sensing of Environment*. 141, 79–89.

Jarihani, A., McVicar, T., Van Niel, T., Emelyanova, I., Callow, J., Johansen, K., 2014. Blending Landsat and MODIS Data to Generate Multispectral Indices: A Comparison of "Index-then-Blend" and "Blend-then-Index" Approaches. *Remote Sensing*. 6, 9213–9238.

Jarlan, L., Tourre, Y.M., Mougin, E., Philippon, N., Mazzega, P., 2005. Dominant patterns of AVHRR NDVI interannual variability over the Sahel and linkages with key climate signals (1982–2003). *Geophysical Research Letters*. 32.

Kaspersen, P.S., Fensholt, R., Huber, S., 2011. A Spatiotemporal Analysis of Climatic Drivers for Observed Changes in Sahelian Vegetation Productivity (1982–2007). *International Journal of Geophysics*. 2011, 1–14.

Kirch, W. (Ed.), 2008. Pearson's Correlation Coefficient, in: Encyclopedia of Public Health. Springer Netherlands, Dordrecht, pp. 1090–1091.

Knauer, K., 2011. Monitoring ecosystem health of Fynbos remnant vegetation in the City of Cape Town using remote sensing. University of Wuerzburg.

Knauer, K., Gessner, U., Dech, S., Kuenzer, C., 2014. Remote sensing of vegetation dynamics in West Africa. *International Journal of Remote Sensing*. 35, 37–41.

Knauer, K., Gessner, U., Fensholt, R., Forkuor, G., Kuenzer, C., 2017. Monitoring Agricultural Expansion in Burkina Faso over 14 Years with 30 m Resolution Time Series : The Role of Population Growth and Implications for the Environment. *Remote Sensing*. 9, 1–25.

Knauer, K., Gessner, U., Fensholt, R., Kuenzer, C., 2016. An ESTARFM Fusion Framework for the Generation of Large-Scale Time Series in Cloud-Prone and Heterogeneous Landscapes. *Remote Sensing*. 8, 425.

Knudby, A., 2004. An AVHRR-based model of groundnut yields in the Peanut Basin of Senegal. *International Journal of Remote Sensing*. 25, 3161–3175.

Kuenzer, C., van Beijma, S., Gessner, U., Dech, S., 2014. Land surface dynamics and environmental challenges of the Niger Delta, Africa: Remote sensing-based analyses spanning three decades (1986-2013). *Applied Geography*. 53, 354–368.

Kumar, L., Rietkerk, M., van Langevelde, F., van de Koppel, J., van Andel, J., Hearne, J., de Ridder, N., Stroosnijder, L., Skidmore, A.K., Prins, H.H.T.T., 2002. Relationship between vegetation growth rates at the onset of the wet season and soil type in the Sahel of Burkina Faso: implications for resource utilisation at large scales. *Ecological Modelling*. 149, 143–152.

Lambert, M.J., Waldner, F., Defourny, P., 2016. Cropland mapping over Sahelian and Sudanian agrosystems: A Knowledge-based approach using PROBA-V time series at 100-m. *Remote Sensing*. 8.

Le Houérou, H.N., 1984. Rain use efficiency - A unifiying concept in arid-land ecology. *Journal of Arid Environments*. 7, 213–247.

Leroux, L., Jolivot, A., Bégué, A., Seen, D., Zoungrana, B., 2014. How Reliable is the MODIS Land Cover Product for Crop Mapping Sub-Saharan Agricultural Landscapes? *Remote Sensing*. 6, 8541–8564.

Li, J., Lewis, J., Rowland, J., Tappan, G.G., Tiszen, L.L., 2004. Evaluation of land performance in Senegal using multi-temporal NDVI and rainfall series. *Journal of Arid Environments*. 59, 463–480.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R news* 2, 18–22.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2004. Remote Sensing and Image Interpretation.

Liu, J., Heiskanen, J., Aynekulu, E., Maeda, E., Pellikka, P., 2016. Land Cover Characterization in West Sudanian Savannas Using Seasonal Features from Annual Landsat Time Series. *Remote Sensing*. 8, 365.

LP DAAC, 2015. MODIS Reprojection Tool [WWW Document]. URL https://lpdaac.usgs.gov/tools/modis_reprojection_tool (accessed 12.3.15).

Machwitz, M., 2010. Eine raum-zeitliche Modellierung der Kohlenstoffbilanz mit Fernerkundungsdaten auf regionaler Ebene in Westafrika. University of Wuerzburg.

Machwitz, M., Wegmann, M., Conrad, C., Dech, S., 2011. Remote Sensing based estimation of potential terrestrial carbon stocks in West Africa, in: *34th ISRSE (International Symposium on Remote Sensing of Environment)*. Sydney.

MAHRH, 2004. Politique nationale de développement durable de l'agriculture irriguée – Stratégie, plan d'action et plan d'investissement Horizon 2015 31.

Martínez, B., Gilabert, M.A., García-Haro, F.J., Faye, A., Meliá, J., 2011. Characterizing land condition variability in Ferlo, Senegal (2001–2009) using multi-temporal 1-km Apparent Green Cover (AGC) SPOT Vegetation data. *Global Planetary Change* 76, 152–165.

Martiny, N., Camberlin, P., Richard, Y., Philippon, N., 2006. Compared regimes of NDVI and rainfall in semi-arid regions of Africa. *International Journal of Remote Sensing*. 27, 5201–5223.

Maselli, F., Conese, C., Petkov, L., Gilabert, M.A., 1993. Environmental monitoring and crop forecasting in the Sahel through the use of NOAA NDVI data . A case study : Niger 1986-89. *International Journal of Remote Sensing*. 14, 3471–3487.

Maselli, F., Conese, C., Petkov, L., Gilabert, M.A., 1992. Use of NOAA-AVHRR NDVI data for environmental monitoring and crop forecasting in the Sahel. Preliminary results. *International Journal of Remote Sensing*. 13, 2743–2749.

Maselli, F., Romanelli, S., Bottai, L., Maracchi, G., 2000. Processing of GAC NDVI data for yield forecasting in the Sahelian region. *International Journal of Remote Sensing*. 21, 3509–3523.

Mathys, E., Gardner, A., 2009. USAID Office of Food for Peace Burkina Faso Security Country Framework FY 2010 - 2014. Washington, D.C.

Mbow, C., Fensholt, R., Nielsen, T.T., Rasmussen, K., 2014. Advances in monitoring vegetation and land use dynamics in the Sahel. *Geografisk Tidsskrift-Danish Journal of Geography*. 114, 84–91.

Meroni, M., Fasbender, D., Kayitakire, F., Pini, G., Rembold, F., Urbano, F., Verstraete, M.M., 2014. Early detection of biomass production deficit hot-spots in semi-arid environment using FAPAR time series and a probabilistic approach. *Remote Sensing of Environment*. 142, 57–68.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*. 66, 247–259.

NASA, 2017a. Reverb - The Next Generation Earth Science Discovery Tool [WWW Document]. URL https://reverb.echo.nasa.gov/reverb/ (accessed 1.30.17).

NASA, 2017b. A Landsat Timeline [WWW Document]. URL http://landsat.gsfc.nasa.gov/a-landsat-timeline/ (accessed 1.30.17).

NASA, 2017c. MODIS - Moderate Resolution Imaging Spectroradiometer [WWW Document]. URL https://modis.gsfc.nasa.gov/about/ (accessed 1.31.17).

NASA LP DAAC, 2013. MODIS Data Products Table [WWW Document]. URL https://lpdaac.usgs.gov/products/modis_products_table

NEPAD, 2013. Agriculture in Africa - Transformation and Outlook. Johannesburg.

Neumann, R., Jung, G., Laux, P., Kunstmann, H., 2007. Climate trends of temperature, precipitation and river discharge in the Volta Basin of West Africa. *International Journal of River Basin Management*. 5, 17–30.

Nicholson, S.E., Some, B., Kone, B., 2000. An Analysis of Recent Rainfall Conditions in West Africa, Including the Rainy Seasons of the 1997 El Nino and the 1998 La Nina Years. *Journal of Climate*. 13, 2628–2640.

Nicholson, S.E., Tucker, C.J., Ba, M.B., 1998a. Desertification, Drought and Surface Vegetation : An Example from the West African Sahel. *Bulletin of the American Meteorological Society*. 79, 815–829.

Nicholson, S.E., Tucker, C.J., Ba, M.B., 1998b. Desertification, Drought, and Surface Vegetation : An Example from the West African Sahel. *Bulletin of the American Meteorological Society*. 79, 815–829.

OEC, 2017. Burkina Faso Profile [WWW Document]. URL http://atlas.media.mit.edu/de/profile/country/bfa/ (accessed 2.1.17).

OECD, 2009. Regional Atlas on West Africa, Regional Atlas on West Africa, *West African Studies*. OECD Publishing, Paris.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J. a., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *Bioscience* 51, 933–938.

Olsson, L., Eklundh, L., Ardö, J., 2005. A recent greening of the Sahel—trends, patterns and potential causes. *Journal of Arid Environments*. 63, 556–566.

Ouedraogo, B., 2014. To limit forest loss, Burkina Faso brings communities into decision making [WWW Document]. Thomson Reuters Found. URL http://news.trust.org//item/20140617153754-d9zb2/?source=fiHeadlineStory (accessed 11.14.16).

Ouedraogo, I., Mbow, C., Balinga, M., Neufeldt, H., 2015. Transitions in Land Use Architecture under Multiple Human Driving Forces in a Semi-Arid Zone. *Land* 4, 560–577.

Paeth, H., Capo-Chichi, A., Endlicher, W., 2008. Climate Change and Food Security in Tropical West Africa - A Dynamic-Statistical Modelling Approach. *Erdkunde* 62, 101–115.

Paeth, H., Thamm, H.-P., 2007. Regional modelling of future African climate north of 15°S including greenhouse warming and land degradation. *Climate Change* 83, 401–427.

Pal, M., Mather, P.M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*. 86, 554–565.

Philippon, N., Jarlan, L., Martiny, N., Camberlin, P., Mougin, E., 2007. Characterization of the Interannual and Intraseasonal Variability of West African Vegetation between 1982 and 2002 by Means of NOAA AVHRR NDVI Data. *Journal of Climate*. 20, 1202–1218.

Prince, S.D., Brown de Colstoun, E.C., Kravitz, L.L., 1998. Evidence from rain-use efficiencies does not indicate extensive Sahelian desertification. *Gobal Change Biology*. 4, 359–374.

Prince, S.D., Wessels, K.J., Tucker, C.J., Nicholson, S.E., 2007. Desertification in the Sahel: a reinterpretation of a reinterpretation. *Gobal Change Biology*. 13, 1308–1313.

Professor Crystal Schaaf's Lab, 2016. MODIS User Guide V006 [WWW Document]. URL https://www.umb.edu/spectralmass/terra_aqua_modis/v006 (accessed 3.14.16).

Quansah, E., Mauder, M., Balogun, A.A., Amekudzi, L.K., Hingerl, L., Bliefernicht, J., Kunstmann, H., 2015. Carbon dioxide fluxes from contrasting ecosystems in the Sudanian Savanna in West Africa. *Carbon Balance and Management*. 10, 1.

Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles* 22, 1–19.

Rasmussen, K., Fensholt, R., Fog, B., Vang Rasmussen, L., Yanogo, I., 2014. Explaining NDVI trends in northern Burkina Faso. *Geografisk Tidsskrift-Danish Journal of Geography*. 114, 17–24.

Rasmussen, M.S., 1998a. Developing simple, operational, consistent NDVI-vegetation models by applying environmental and climatic information: Part I. Assessment of net primary production. *International Journal of Remote Sensing*. 19, 97–117.

Rasmussen, M.S., 1998b. Developing simple, operational, consistent NDVI-vegetation models by applying environmental and climatic information. Part II: Crop yield assessment. *International Journal of Remote Sensing*. 19, 119–139.

Rasmussen, M.S., 1997. Operational yield forecast using AVHRR NDVI data: Reduction of environmental and inter-annual variability. *International Journal of Remote Sensing*. 18, 1059–1077.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*. 5, 703–714.

Reed, B.C., Schwartz, M.D., Xiao, X., 2009. Remote Sensing Phenology, in: Noormets, A. (Ed.), *Phenology of Ecosystem Processes*. Springer-Verlag, pp. 231–246.

Reed, B.C., White, M., Brown, J.F., 2003. Remote Sensing Phenology, in: Schwartz, M.D. (Ed.), *Phenology: An Integrative Environmental Science, Tasks for Vegetation Science*. Springer Netherlands, Dordrecht, pp. 365–381.

Retzer, V., 2006. Impacts of grazing and rainfall variability on the dynamics of a Sahelian rangeland revisited (Hein, 2006)—new insights from old data. *Journal of Arid Environments*. 67, 157–164.

Ringelmann, N., Scipal, K., Bartalis, Z., Wagner, W., 2004. Planting date estimation in semi-arid environments based on Ku-band radar scatterometer data, in: *IGARSS 2004 - 2004 IEEE International Geoscience and Remote Sensing Symposium*.

Rojas, O., Vrieling, a., Rembold, F., 2011. Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. *Remote Sensing of Environment*. 115, 343–352.

Sasaki, N., Putz, F.E., 2009. Critical need for new definitions of "forest" and "forest degradation" in global climate change agreements. *Conservation Letters*. 2, 226–232.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*. 36, 1627–1639.

Schmidt, M., Lucas, R., Bunting, P., Verbesselt, J., Armston, J., 2015. Multi-resolution time series imagery for forest disturbance and regrowth monitoring in Queensland, Australia. *Remote Sensing of Environment*. 158, 156–168.

Schowengerdt, R.A., 2007. Remote sensing: models and methods for image processing. Academic Press.

Seaquist, J.W., Hickler, T., Eklundh, L., Ard, J., Ardö, J., Heumann, B.W., 2009. Disentangling the effects of climate and people on Sahel vegetation dynamics. *Biogeosciences* 6, 469–477.

Seaquist, J.W., Olsson, L., Ardö, J., Eklundh, L., 2006. Broad-scale increase in NPP quantified for the African Sahel, 1982–1999. *International Journal of Remote Sensing*. 27, 5115–5122.

Sedogo, S.A., Bourgou, T., 2015. Lier la demande et l'offre de conseil agricole autour des grands barrages - Le cas de Bagré au Burkina Faso.

Sjöström, M., Zhao, M., Archibald, S., Arneth, A., Cappelaere, B., Falk, U., de Grandcourt, A., Hanan, N.P., Kergoat, L., Kutsch, W.L., Merbold, L., Mougin, E., Nickless, A., Nouvellon, Y., Scholes, R.J., Veenendaal, E.M., Ardö, J., Grandcourt, A. De, 2013. Evaluation of MODIS gross primary productivity for Africa using eddy covariance data. *Remote Sensing of Environment*. 131, 275–286.

Song, H., Huang, B., 2013. Spatiotemporal Satellite Image Fusion Through One-Pair Image Learning. *IEEE Transactions on Geoscience and Remote Sensing*. 51, 1883–1896.

Tappan, G., Cushing, W.M., Cotilion, S.E., Mathis, M.L., Hutchinson, J.A., Dalsted, K.J., 2016. West Africa Land Use Land Cover Time Series: U.S. Geological Survey data release.

Tewes, A., Thonfeld, F., Schmidt, M., Oomen, R., Zhu, X., Dubovyk, O., Menz, G., Schellberg, J., 2015. Using RapidEye and MODIS Data Fusion to Monitor Vegetation Dynamics in Semi-Arid Rangelands in South Africa. *Remote Sensing*. 7, 6510–6534.

Thenkabail, P.S., Stucky, N., Griscom, B.W., Ashton, M.S., Diels, J., va der Meer, B., Enclona, E., 2004. Biomass estimations and carbon stock calculations in the oil palm plantations of African derived savannas using IKONOS data. *International Journal of Remote Sensing*. 25, 5447–5472.

Thornton, P.K., Bowen, W.T.T., Ravelo, A.C., Wilkens, P.W.W., Farmer, G., Brock, J., Brink, J.E.E., a.C. Ravelo, 1997. Estimating millet production for famine early warning: an application of crop simulation modelling using satellite and ground-based data in Burkina Faso. *Agricultural and Forest Meteorology*. 83, 95–112.

Tian, F., Wang, Y., Fensholt, R., Wang, K., Zhang, L., Huang, Y., 2013. Mapping and Evaluation of NDVI Trends from Synthetic Time Series Obtained by Blending Landsat and MODIS Data around a Coalfield on the Loess Plateau. *Remote Sensing*. 5, 4255–4279.

Tøttrup, C., Rasmussen, M.S., 2004. Mapping long-term changes in savannah crop productivity in Senegal through trend analysis of time series of remote sensing data. *Agriculture, Ecosystems & Environment*. 103, 545–560.

Tou, J.T., Gonzalez, R.C., 1974. Pattern recognition principles. United States.

Tso, B., Mather, P.M., 2009. Classification methods for remotely sensed data, Second. ed, Methods.

Tucker, C.J., Dregne, H.E., Newcomb, W.W., 1991. Expansion and Contraction of the Sahara Desert from 1980 to 1990. *Science*. 253, 299–301.

UCDAVIS Agricultural Sustainability Institute, 2018. What is sustainable agriculture? [WWW Document]. URL http://asi.ucdavis.edu/programs/sarep/about/what-is-sustainable-agriculture (accessed 5.20.18).

UNCCD, 2017. The Convention [WWW Document]. URL http://www.unccd.int/en/about-the-convention/Pages/Text-Part-I.aspx (accessed 8.15.17).

UNDP, 2015. UN Human Development Report.

United Nations, 2016. UN Comtrade Database [WWW Document]. URL http://comtrade.un.org/ (accessed 11.7.16).

United Nations, 2015. World Population Prospects, the 2015 Revision.

United Nations Statistics Divison, 2013. Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings [WWW Document]. URL http://unstats.un.org/unsd/methods/m49/m49regin.htm

USGS, 2017a. Global Food Security Analysis-Support Data at 30 Meters (GFSAD30) Project [WWW Document]. URL https://croplands.org/ (accessed 2.14.17).

USGS, 2017b. USGS Landsat Global Archive [WWW Document]. URL https://landsat.usgs.gov/usgs-landsat-global-archive (accessed 1.20.17).

USGS, 2017c. EarthExplorer [WWW Document]. URL https://earthexplorer.usgs.gov/ (accessed 1.30.17).

USGS, 2017d. Landsat Missions [WWW Document]. URL https://landsat.usgs.gov/ (accessed 1.30.17).

USGS, 2015. Product Guide - Provisional Landsat 8 Surface Reflectance Product, Product Guide Australia.

Van der Waal, H.-W., 2011. Meeting the challenges of exporting mangoes from Burkina Faso [WWW Document]. URL http://www.agriculturesnetwork.org/magazines/global/towards-fairer-trade/meeting-the-challenges-of-exporting-mangoes-from (accessed 11.13.16).

Vapnik, V., 1979. Estimation of Dependences Based on Empirical Data. Nauka, Moscow.

Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D., 2010. Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*. 114, 106–115.

Vintrou, E., Bégué, A., Baron, C., Saad, A., Lo Seen, D., Traoré, S., 2014. A Comparative Study on Satellite- and Model-Based Crop Phenology in West Africa. *Remote Sensing*. 6, 1367–1389.

Vintrou, E., Desbrosse, A., Bégué, A., Traoré, S., Baron, C., Lo Seen, D., 2012a. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. *International Journal of Applied Earth Observation and Geoinformation*. 14, 83–93.

Vintrou, E., Soumaré, M., Bernard, S., Bégué, A., Baron, C., Lo Seen, D., 2012b. Mapping Fragmented Agricultural Systems in the Sudano-Sahelian Environments of Africa Using Random Forest and Ensemble Metrics of Coarse Resolution MODIS Imagery. *Photogrammetric Engineering & Remote Sensing*. 78, 839–848.

Vlek, P.L.G., Le, Q.B., Tamene, L., 2010. Assessment of Land Degradation, its Possible Causes and Threat to Food Security in Sub-Saharan Africa, in: Lal, R., Stewart, B.A. (Eds.), Food Security and Soil Quality. *Advances in Soil Science*. Taylor & Francis, Boka Raton, pp. 57–86.

Vrieling, A., de Leeuw, J., Said, M.Y., 2013. Length of Growing Period over Africa: Variability and Trends from 30 Years of NDVI Time Series. *Remote Sensing*. 5, 982–1000.

Walker, J.J., Beurs, K.M. De, Wynne, R.H., 2014. Dryland vegetation phenology across an elevation gradient in Arizona , USA , investigated with fused MODIS and Landsat data. *Remote Sensing of Environment*. 144, 85–97.

Walker, J.J., de Beurs, K.M., Wynne, R.H., Gao, F., 2012. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment*. 117, 381–393.

Wang, P., Gao, F., Masek, J.G., 2014. Operational Data Fusion Framework for Building Frequent Landsat-Like Imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 52, 7353–7365.

Ward, D., 2005. Do we understand the causes of bush encroachment in African savannas? *African Journal of Range & Forage Science*. 22, 101–105.

Wessels, K.J., 2009. Comments on "Proxy global assessment of land degradation" by Bai et al. (2008). *Soil Use and Management*. 25, 91–92.

White, M., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R., van Leeuwen, W.J.D., Brown, J.F., de Wit, A., Schaepman, M.E., Lin, X., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., Lauenroth, W.K., 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006. *Gobal Change Biology*. 15, 2335–2359.

Wohlfart, C., Bevanda, M., Horning, N., Leutner, B., Wegmann, M., 2016a. Field Data for Remote Sensing Data Analysis, in: Wegmann, M., Leutner, B., Dech, S. (Eds.), *Remote Sensing and GIS for Ecologists: Using Open Source Software*. Pelagic Publishing, Exeter, pp. 136–149.

Wohlfart, C., Liu, G., Huang, C., Kuenzer, C., 2016b. A River Basin over the course of time: Multi-temporal analyses of land surface dynamics in the Yellow River Basin (China) based on medium resolution remote sensing data. *Remote Sensing*. 8.

World Bank, 2016. International Tourism, Number of Arrivals [WWW Document]. URL http://data.worldbank.org/indicator/ST.INT.ARVL?locations=BF (accessed 11.14.16).

World Bank, 2010a. Burkina Faso - Bagre Growth Pole Project.

World Bank, 2010b. Forest Investment Program - Burkina Faso.

World Food Summit, 1996. [WWW Document]. URL http://www.fao.org/WFS/ (accessed 5.20.18).

Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*. 185, 271–283.

Yu, L., Wang, J., Gong, P., 2013. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach. *International Journal of Remote Sensing*. 34, 5851–5867.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Liu, Z., 2005. Monitoring the response of vegetation phenology to precipitation in Africa by coupling MODIS and TRMM instruments. *Journal of Geophysical Research*. 110, 1–14.

Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*. 114, 2610–2623.

Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., Lefsky, M.A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sensing of Environment*. 172, 165–177.

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*. 118, 83–94.

Zhukov, B., Oertel, D., Lanzl, F., Reinhäckel, G., 1999. Unmixing-based Multisensor Multiresolution Image Fusion. *IEEE Transactions on Geoscience and Remote Sensing*. 37, 1212–1226.

Zoungrana, B.J.B., Conrad, C., Amekudzi, L.K., Thiel, M., Da, E.D., Forkuor, G., Loew, F., 2015. Multi-temporal landsat images and ancillary data for land use/cover change (LULCC) detection in the Southwest of Burkina Faso, West Africa. *Remote Sensing*. 7, 12076–12102.

Zurita-Milla, R., Clevers, J.G.P.W., Schaepman, M.E., 2008. Unmixing-Based Landsat TM and MERIS FR Data Fusion. *IEEE Geoscience and Remote Sensing Letters*. 5, 453–457.

Zurita-Milla, R., Clevers, J.G.P.W., Van Gijsel, J. a. E., Schaepman, M.E., 2011. Using MERIS fused images for land-cover mapping and vegetation status assessment in heterogeneous landscapes. *International Journal of Remote Sensing*. 32, 973–991.

Zurita-Milla, R., Kaiser, G., Clevers, J.G.P.W., Schneider, W., Schaepman, M.E., 2009. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sensing of Environment*. 113, 1874–1885.

Eidesstattliche Erklärung

Ich erkläre hiermit, dass die von mir eingereichte Dissertation zum Thema „Vegetation Dynamics in West Africa - Spatio-temporal Data Fusion for the Monitoring of Agricultural Expansion" selbstständig und ausschließlich unter Verwendung der angegebenen Literatur und Hilfsmittel verfasst wurde. Alle den angeführten Quellen wörtlich oder sinngemäß entnommenen Stellen habe ich als solche kenntlich gemacht.

Diese Arbeit ist keiner anderen Prüfungsbehörde vorgelegt worden.

_____

München, den 08. September 2017

Kim Knauer

geboren am 02.08.1985 in Erlenbach a.M.

Berufserfahrung & Praktika

Seit 2017 **Remote Sensing Specialist**, EOMAP GmbH & Co. KG.

2012-2013 **Wissenschaftlicher Mitarbeiter**, CRP-Gabriel Lippmann, Luxemburg.

2011-2012 **Wissenschaftliche Hilfskraft**, Lehrstuhl für Fernerkundung, Universität Würzburg in Kooperation mit dem DLR.

2010 **Praktikum**. RSS – Remote Sensing Solutions GmbH.

2009 **Praktikum**. AGIS –Anwender-Geo-Informations-Systeme GmbH.

2007/2008 **Praktikum**. DFD – Deutsches Fernerkundungs-Datenzentrum.

Ausbildung

2013-2017 **Dissertation**, Lehrstuhl für Fernerkundung, Universität Würzburg in Kooperation mit dem Deutschen Zentrum für Luft- und Raumfahrt (DLR). *»Remote Sensing of Vegetation Dynamics in West Africa: Spatio-Temporal Data Fusion for the Monitoring of Agricultural Expansion«*

2010-2011 **Diplomarbeit**, Lehrstuhl für Fernerkundung, Universität Würzburg. *»Monitoring ecosystem health of Fynbos remnant vegetation in the City of Cape Town using remote sensing«*

2004-2011 **Studium der Geographie**, Universität Würzburg. Nebenfächer: Geologie und Marketing.

2004 **Allgemeine Hochschulreife**, Julius-Echter-Gymnasium, Elsenfeld.

Veröffentlichungen

Knauer, K.; Gessner, U.; Fensholt, R. and Kuenzer C. (2017): Monitoring Agricultural Expansion in Burkina Faso over 14 Years with 30 m Resolution Time Series: The Role of Population Growth and Implications for the Environment. *Remote Sensing. 9 (132).*

Knauer, K.; Gessner, U.; Fensholt, R. and Kuenzer C. (2016): An ESTARFM Fusion Framework for the Generation of Large-Scale Time Series in Cloud-Prone and Heterogeneous Landscapes. *Remote Sensing. 8 (5).*

Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2015): Remote sensing of vegetation dynamics in West Africa: improved satellite time series for phenological analyses. *Poster presentation at the 36th ISRSE, Berlin, 11th-15th May 2015.*

Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2014): Challenges in the generation of LAI time series for West Africa: difficulties and potential improvements. Poster presentation at the GV2M, Avignon, 3rd-7th February 2014.

Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2014): Remote Sensing of Vegetation Dynamics in West Africa. *International Journal of Remote Sensing. 35 (17). 6357-6396.*

Knauer, K.; Gessner, U.; Dech, S. and Kuenzer, C. (2014): Remote sensing of vegetation dynamics in West Africa: improved satellite time series for phenological analyses. *Poster presentation at the 5th ESA Advanced Training Course on Land Remote Sensing, Valencia, 8th-12th September 2014.*

Gessner, U.; Knauer, K.; Kuenzer, C. and Dech, S. (2014): Land surface phenology in a West African savanna: impact of land use, land cover and fire. *In book: Remote Sensing Time Series revealing Land Surface Dynamics, Publisher: Springer, Editors: Claudia Kuenzer, Stefan Dech, Wolfgang Wagner, pp.203-223.*

Kuenzer, C. and Knauer K. (2013): Remote sensing of rice crop areas. *International Journal of Remote Sensing. 34 (6). 2101-2139.*

Knauer, K.; Klein, D.; Allsopp, N.; Baumhauer, R. and Dech, S. (2011): Monitoring ecosystem health of Fynbos remnant vegetation in the City of Cape Town using remote sensing. *South African Journal of Botany. 77 (572).*