Remote sensing based assessment of land cover and soil moisture in the Kilombero floodplain in Tanzania

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Fridah Kirimi

aus

Nairobi, Kenia

Bonn, May 2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Bernd Diekkrüger
2. Gutachter: Prof. Dr. Karl Schneider

Tag der Promotion: 16.07.2018

Erscheinungsjahr: 2018

Declaration

I declare that this thesis is my own original work and that it has not been presented and will not be presented to any other University for a similar or any other degree award.

Fridah Kirimi

Acknowledgment

The Federal Ministry of Education and Research (BMBF) supported the research under the project GlobE Wetlands project number BMBF FKZ.031A250A-H. The optical images in the research were obtained from Blackbridge (RapidEye), ESA Sentinel portal (Sentinel-2) and the USGS website (Landsat). The SAR images were acquired from German Aerospace Center-DLR (TerraSar-X), MacDonald, Dettwiler, and Associates-MDA (RadarSat-2) and the Sentinel portal (Sentinel-1). The MODIS data was downloaded from USGS data portal, the SMOS from ESA, ERA-Interim from ECMWF and FLDAS_NOAH and VIC from the NASA LDAS website.

My sincere gratitude goes to Prof. Dr. Bernd Diekkrüger for taking up the supervisory role after the demise of my previous supervisor. I thank him for his guidance and unwavering support towards shaping my work and bringing it to completion. I appreciate the late Prof. Dr. Gunter Menz, for believing in me and recognizing my potential in accomplishing the research tasks by agreeing to take me as his Ph.D. student. I appreciate Dr. Carsten Montzka for his valuable comments and advice.

I acknowledge the GlobE lead coordinators, which include Prof. Dr. Mathias Becker, Prof. Dr. Bernd Diekkrüger, Prof. Dr. Barbara Reichert, Prof. Dr. Matthias Langensiepen, Prof. Dr. Thomas Heckelei, Prof. Dr. Salome Misana, Dr. Constanze Leemhuis, Dr. Andrea Rechenburg, Dr, Daniel Kyalo, and Dr. Volker Ermert. I thank my fellow GlobE Ph.D. students Bisrat, Björn, Carmen, Claudia, Erick, Esther, Geofrey, Innocent, Irene, Jane, Johanna, John, Julius, Kai, Katrin, Kristian, Landry, Larisa, Lalya, Matian, Ramadhan, Simon, Sonja, Sophie-Bo, Susanne, Viviane. Appreciation also goes to the RSRG and ZFL members Agatha, Amit, Dr. Andy, Anna, Annette, Bärbel, Carsten, Di, Ellen, Fabián, Dr. Frank, Gohar, Henrik, Javier, Jessica, Dr. Johannes, Kilian, Dr. Konrad, Kuria, Dr. Olena, Dr. Reginald, Ruben, Sascha, Stefanie, Tobias, Tomasz, Dr. Valerie and to my field assistants Goodluck, Seleman, Charles, Juma.

I thank all my friends for their encouragement and support. Dr. Georgina, Dr. Oliver, Dr. Evelyn, Dr. Naftali, Veronica, Mr. Hillary, Mr. Ngugi, Maggie, Mr. Gone, Bettie, Carol, Larissa, Jackie, Nancy W, Nancy M, Dr. Benson, Mr. Dismas, Julius M, Dr. Jude, Benard and everyone else who made an impact in this journey.

Special thanks go to my mother Caroline K, my late father John K, my sisters, Fides K and Florence K for their never-ending love, support and encouragement. To all my extended family members a big thank you for your support. Above all am grateful to God for seeing me through this journey.

Table of Content

List of Figures

List of Tables

List of Abbreviations

AMSR E	Advanced Microwave Scanning Radiometer Earth observing system
ASCAT	Advanced SCATterometer
ATBD	Algorithm Theoretical Basis Document
AVHRR	Advanced Very High-Resolution Radiometer
CHIRPS	Climate Hazards group Infra-Red Precipitation with Stations
CPC	Climate Prediction Center
ECMWF	European Centre for Medium-range Weather Forecasts
ECV	Essential Climate Variables
ERA-Interim	ECMWF Re-Analysis Interim
ERS AMI	European Remote Sensing Active Microwave Instrument
ESA	European Space Agency
FEWSNET	Famine Early Warning Systems NETwork
FLDAS	FEWSNET Land Data Assimilation System
GDAS	Global Data Assimilation System
GES DISC	NASA Goddard Earth Science Data and Information Services Center
GLDAS	Global Land Data Assimilation System
IGBP	International Global Biosphere Programme
L-MEB	L band Microwave Emission of the Biosphere
LP DAAC	NASA Land Processes Distributed Active Archive Center
LPRM	Land Parameter Retrieval Model
MERRA	Modern ERA Reanalysis for Research and Applications
METOP	METeorological Operational Satellite
MIR_SMUDP2	MIRAS Soil Moisture Use Data Product Level 2
MIRAS	Microwave Imaging Radiometer with Aperture Synthesis
MODIS	MODerate-resolution Imaging Spectrometer
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
RFE	African Rainfall Estimation
RFI	Radio Frequency Interference
SCA	Single Channel Algorithm

SMAP	Soil Moisture Active Passive
SMMR	Scanning Multichannel Microwave Radiometer
SMOS	Soil Moisture and Ocean Salinity
SRTM	Shuttle Radar Topography Mission
TESSEL	Tiled ECMWF Scheme for Surface Exchanges over Land
TMI	TRMM Microwave Imager
TRMM	Tropical Rainfall Measuring Mission
UMD	University of MarylanD
USAID	US Agency for International Development
VIC	Variable Infiltration Capacity

Abstract

Wetlands provide important ecological, biological, and social-economic services that are critical for human existence. The increasing demand for food, arable land shortage and changing climate conditions in East Africa have created a paradigm shift from upland cultivation to wetland use due to their year-round soil water availability. However, there is need to control and manage the activities within the wetlands to ensure sustainable use while negating any negative effects caused by these activities. This is implemented through the decisions made by the land managers within the wetlands. Providing the users of the wetlands with scientific knowledge acts as a support tool for policy-making geared towards the sustainable use of the wetlands. The overall research contains two main components: First, the need for timely land cover maps at a reasonable scale, and secondly, the assessment of soil moisture as a major contributor to agricultural production.

The objectives of the study were to generate land cover maps from multi-sensor optical datasets and to assess the performance of single-polarized Sentinel-1 Gray Level Co-occurrence Matrix (GLCM) texture and Principal Component Analysis (PCA) features by applying multiple classification algorithms in a floodplain in the Kilombero catchment. Furthermore, soil moisture spatial-temporal patterns over three hydrological zones was assessed, estimation of soil moisture from radar data and generation of soil moisture products from global products was investigated. The correlation of the merged products to Normalized Difference Vegetation Index (NDVI) measures was also investigated.

RapidEye, Sentinel-2 and Landsat images were used in determining the areal extents of four major land cover classes namely vegetated, bare, water and built up. The acquisition period of the images ranges from August 2013 to June 2015 for the RapidEye images, December 2015 to August 2016 for the Sentinel-2 images and 2013 to 2016 Landsat-8 images were included in the land cover time series dynamic study. However, the major challenge arising was cloud coverage and hence Sentinel-1 images were tested in the application of Synthetic Aperture Radar (SAR) in wetland mapping. Variograms were used in spatial-temporal assessment of soil moisture data collected from three hydrological zones, riparian, middle and fringe. A roughness parameter was derived from a semi-empirical model. Soil moisture was retrieved from TerraSAR-X and RadarSAT-2 with the retrieved roughness parameter as an input in a linear regression equation. Triple collocation was applied in error assessment of the global soil moisture products prior to development of a merged product. Cross-correlation was applied in relating NDVI to soil moisture.

Optical data (RapidEye, Landsat-8, and Sentinel-2) generated land cover maps used in assessing the land cover dynamics over time. The land cover ratios were related to depth to groundwater. As the depth to groundwater reduced in June the bare land coverage was 45-57% while that of vegetation was 34-47%. In December when the depth to groundwater was highest, bare land coverage was 62-69% while that of the vegetated area was 27-

25%. This indicates that depth of groundwater and vegetation coverage responds to seasonality. During the dry season, 68-81% of the total vegetation class is within the riparian zone.

In the classification of the SAR images, the overall accuracies for the single polarized VV images ranged from 54-76%, 60-81% and 61-80% for Random Forest (RF), Neural Network (NN) and Support Vector Machine (SVM) respectively. GLCM features had overall accuracies of 64-86%, 65-88% and 65-86% for RF, NN, and SVM respectively. PCA derived images had similar overall accuracies of 68-92% for NN, RF, and SVM respectively. The PCA images had the highest overall accuracy for the entire time series indicating that reduction in the number of texture features to layers containing the maximum variance improves the accuracy.

The standard deviation of soil moisture was noted to increase with increasing soil moisture. Soil texture plays a key role in soil moisture retention. The riparian fields had a high water content explained by the high clay and organic matter content. A roughness parameter was derived and utilized in the retrieval of soil moisture from SAR resulting to R^2 of 0.88- 0.92 between observed and simulated soil moisture values from co-polarized RadarSAT-2 HH and TerraSAR-X HH and VV.

Merged soil moisture product from FEWSNET Land Data Assimilation System_NOAH (FLDAS_NOAH), ECMWF Re-Analysis Interim (ERA-Interim) and Soil Moisture and Ocean Salinity (SMOS) and FLDAS_Variable Infiltration Capacity (VIC), ERA-Interim and SMOS had similar patterns attributed to FLDAS_NOAH and FLDAS_VIC forced by the same precipitation product (RFE). Cross-correlation of Moderate-resolution Imaging Spectrometer (MODIS) NDVI and the merged soil moisture products revealed a 2-month lag of NDVI. Hence, the relationship is useful in determining the Start of Season from soil moisture products.

In conclusion, the successful land cover mapping of the study area demonstrated the use of satellite imagery for wetland characterization. The vast coverage and frequent acquisitions of optical and microwave remotely sensed data additionally make the approaches transferable to other locations and allow for mapping at larger scales. Soil moisture assessment from point data revealed varied soil moisture patterns whereas global remotely sensed and modeled products rather provide complementary information about growing conditions, and hence a situational assessment tool of potential of physical availability dimension of food security. This study forms a baseline upon which additional monitoring and assessment of the Kilombero wetland ecosystem can be performed with the current results marked as a reference. Moreover, the study serves as a demonstration case of remote sensing based approaches for land cover and soil moisture mapping, whose results are useful to stakeholders to aid in the implementation of adapted production techniques for yield optimization while minimizing the unsustainable use of the natural resources.

Zusammenfassung

Feuchtgebiete erbringen wichtige ökologische, biologische und sozial-ökonomische Dienstleistungen, welche entscheidend für das menschliche Dasein sind. Der steigende Bedarf an Nahrung, der Mangel an landwirtschaftlichen Nutzflächen und die Veränderung der klimatischen Bedingungen in Ostafrika haben zu einem Paradigmenwechsel vom Anbau im Hochland hin zur Nutzung von Feuchtgebieten geführt. Allerdings sind Kontrolle und Management der Aktivitäten in Feuchtgebieten notwendig, um die nachhaltige Nutzung zu sichern und negative Effekte dieser Aktivitäten zu vermeiden. Die Implementierung erfolgt durch die Landverwalter in den Feuchtgebieten. Den Nutzern von Feuchtgebieten wissenschaftliche Erkenntnisse bereitzustellen dient als Hilfsmittel zur politischen Entscheidungsfindung für die nachhaltige Feuchtgebietsnutzung. Die Forschung im Rahmen der Dissertation beinhaltet zwei Hauptkomponenten: erstens den Bedarf an aktuellen Landbedeckungskarten auf einer angemessenen Skalenebene und zweitens die Erfassung der Bodenfeuchte als wichtiger Einflussfaktor auf die landwirtschaftliche Produktion.

Das Ziel der Untersuchung war, Landbedeckungskarten auf Grundlage von multisensorischen optischen Daten zu erstellen und die Eignung der Textur der einfach polarisierten Sentinel-1 Grauwertmatrix (GLCM) sowie der einer Hauptkomponentenanalyse (PCA) bei Anwendung unterschiedlicher Klassifikationsalgorithmen zu beurteilen. Des Weiteren wurden raum-zeitliche Bodenfeuchtemuster über drei hydrologische Zonen hinweg modelliert, die Bodenfeuchte aus Radardaten abgeleitet sowie die Erstellung von Bodenfeuchteprodukten auf Basis von globalen Produkten untersucht. Die Korrelation der Bodenfeuchteprodukte mit dem Normalisierten Differenzierten Vegetationsindex (NDVI) wurde ebenfalls analysiert.

RapidEye, Sentinel-2 und Landsat Bilder wurden genutzt um die räumliche Ausdehnung der vier Hauptklassen (Vegetation, freiliegender Boden, Wasser und Bebauung) der Landbedeckung zu ermitteln. Für die Zeitreihenanalyse der der Landbedeckungsdynamik wurden RapidEye-Daten von August 2013 bis Juni 2015, Sentinel-2-Bilder von Dezember 2015 bis August 2016 und Landsat-8-Bilder von 2013 bis 2016 verwendet. Die größte Herausforderung war jedoch die Wolkenbedeckung, weshalb die Anwendung von Synthetic Aperture Radar (SAR) für die Feuchtgebietskartierung getestet wurde. Die gemessene Bodenfeuchte wurde mittels Variogrammen für die drei hydrologischen Zonen (Uferzone, Mitte und Randgebiete) raum-zeitlich interpoliert. Ein Rauhigkeitsparameter wurde aus einem semi-empirischen Modell hergeleitet. Die Bodenfeuchte wurde aus TerraSAR-X und RadarSAT-2- Bildern unter Verwendung des Rauhigkeitsparameters als Eingangsgröße in einer linearen Regression abgeleitet. Vor der Zusammenführung der Produkte wurde das globale Bodenfeuchteprodukt mithilfe von dreifacher Kollokation auf Fehler überprüft. Die Kreuzkorrelation zwischen NDVI und Bodenfeuchte wurde berechnet.

Optische Daten (RapidEye, Landsat-8 und Sentinel-2) wurden genutzt, um die zeitliche Dynamik der Landbedeckung zu bestimmen. Die Landbedeckungsverhältnisse wurde mit der Höhe des Grundwasserspiegels korreliert. Ein hoher Grundwasserstand im Juni resultierte in 45-57% unbedecktem Boden, während der Anteil der Vegetation 34-47% betrug. Im Dezember, als der Grundwasserspiegel seinen Tiefststand hatte, erhöhte sich der Anteil des freiliegenden Bodens auf 62-69% und der Anteil der Vegetation verringerte sich auf 27-25%. Das zeigt, dass Grundwasserspiegel und Vegetation saisonalen Schwankungen unterworfen sind. Während der Trockenzeit liegen 68-81% der gesamten als Vegetation klassifizierten Fläche innerhalb der Uferzone.

In der Klassifikation der SAR-Bilder liegt die Gesamtgenauigkeit der einfach polarisierten VV-Bilder im Rahmen von 54-76%, 60-81% und 61-80%, entsprechend für Random Forest (RF), Neuronale Netze (NN) und Support Vector Machine (SVM). Die GLCM ergab eine Gesamtgenauigkeit von 64-86%, 65-88% und 65-86% für RF, NN und SVM. Die über eine PCA abgeleiteten Bilder erreichten eine ähnliche Genauigkeit von 68-92% für NN, RF und SVM. Die PCA-Bilder weisen die höchste Gesamtgenauigkeit der gesamten Zeitreihe auf, was darauf hinweist, dass eine Reduktion von Textureigenschaften auf Layer der maximalen Varianz enthalten, die Genauigkeit erhöht.

Die Standardabweichung der Bodenfeuchte stieg mit zunehmender Bodenfeuchte. Die Bodentextur spielt dabei eine Schlüsselrolle für das Wasserhaltevermögen des Bodens. Die Uferzone wies einen hohen Wassergehalt auf, was durch den hohen Anteil von Ton und Humus zu erklären ist. Die beobachteten und simulierten Bodenfeuchtewerte von co-polarisierten RadarSAT-2 HH, TerraSAR-X HH und VV Daten korrelieren mit einem R^2 von 0.88 - 0.92.

Die zusammengesetzten globalen Bodenfeuchteprodukte von FLDAS_NOAH, ERA-Interim sowie SMOS und FLDAS_VIC, ERA-Interim und SMOS zeigen ähnliche Muster wie FLDAS_NOAH und FLDAS_VIC, was über die Verwendung desselben Niederschlagsproduktes (RFE) zu erklären ist. Die Kreuzkorrelation von MODIS NDVI und den zusammengeführten Bodenfeuchteprodukten ergab eine zeitliche Verzögerung des NDVI von zwei Monaten. Dieser Zusammenhang kann daher bei der Bestimmung des Saisonbeginns aus Bodenfeuchtigkeitsprodukten nützlich sein.

Zusammengefasst hat die Studie gezeigt, wie Satellitenbilder zur Charakterisierung von Wetlands genutzt werden können. Die große Abdeckung und häufige Aufnahme der optischen und Mikrowellen-Fernerkundungsdaten ermöglichen darüber hinaus die Übertragung der Ansätze auf weitere Gebiete und Kartierung auf größeren Skalen. Die Punktmessungen zeigen kleinräumige Muster der Bodenfeuchte, während globale Fernerkundungsprodukte und Modelle Informationen über die Wachstumsbedingungen liefern und somit ein Bewertungsinstrument der Ernährungssicherheit darstellen können. Weiterhin bildet die Studie eine Basis, auf der ein weitergehendes Monitoring und eine Bewertung des Feuchtgebietsökosystems durchgeführt werden kann. Sie ist ein Beispiel für fernerkundungsbasierte Ansätze zur Landbedeckungs- und Bodenfeuchtekartierung; ihre Ergebnisse sind nützlich, um Akteuren bei der Implementierung von Produktionstechniken zu unterstützen, welche die Erträge maximieren und gleichzeitig die nicht nachhaltige Nutzung der natürlichen Ressourcen minimieren.

1. Introduction

Sub-Sahara Africa is expected to account for 20% of the world population by 2050 (Alexandratos and Bruinsma, 2012). This gives rise to the need for increased food production to meet the demand of the increasing population. Sustainable Development Goal 2 (SDG2) is geared towards eradication of hunger through an increase in food production (Le Blanc, 2015). Climate change is a challenge to increased food production causing extreme events such as flooding or prolonged dry seasons and droughts, which lead to low yields or total crop failure (Nicol *et al.* 2015). Reported global warming modifies the agro-ecological zones subsequently altering the plant growing seasons which boils down to ambiguity in planting times. Unpredictable seasonality due to climate variability is a vulnerability experienced in the agriculture sector (National Bureau of Statistics, 2015). Fertile soils and constant water supply make wetlands preferred ecosystems for expanding and intensifying agricultural production.

Wetlands are multi-functional ecosystems with diverse agricultural, ecological, hydrological, logistical and social roles such as water reservoirs, water consumption, livelihoods (e.g., agriculture, pastoralism, fishery, domestic use, and industrial use), a buffer for flooding, carbon sequestration, and habitats for species (Moser *et al.* 2014). They control the source, amount, and temporal and spatial distribution of sediment and nutrient movements and influence distribution of flora and fauna.

However, the utilization of wetlands necessitates monitoring and any developments within them controlled to avoid degradation. It is with this framework that the 1971 Ramsar Convention on Wetlands of International importance was signed into a treaty with the aim of conservation and wise utilization of wetlands. The Kilombero wetland in Tanzania was declared a Ramsar site in the year 2002, and it is the largest seasonal freshwater lowland floodplain in East Africa.

About 80% of the people of Tanzania depend on subsistence agriculture (Shetto *et al.* 2007, Mombo *et al.* 2011). The Kilombero floodplain was identified as a potential agricultural area to enhance food security (Kato, 2007). Though the wetland was declared a Ramsar site, the land uses within it threaten its existence if activities are left uncontrolled. Moreover, the implementation of the Southern Agricultural Growth Corridor of Tanzania (SAGCOT) with an interest in agriculture development further emphasizes the need for land use planning (Milder *et al.* 2013). Major threats to African wetlands include land use activities such as reclamation for agricultural production and developmental pressures (Gardner *et al.* 2009, Junk *et al.* 2013, Moser *et al.* 2014).

Soil moisture information plays a vital support role in agricultural activities such as irrigation scheduling, precision farming, or drought risk assessment (McNairn *et al.* 2012, Wilhite and Pulwarty 2018). Monitoring soil moisture over vast areas poses a challenge due to its high spatial and temporal variability. Moreover, in situ soil moisture measurements are tedious and costly to collect at repeated intervals. Synthetic Aperture Radar (SAR) satellite imagery have a high response to soil dielectric properties associated with soil moisture content and are thus suited in soil moisture derivation (Ulaby *et al.* 1996). Examples of SAR imagery in soil moisture retrieval include RadarSat-2 and TerraSar-X (Baghdadi, Camus, *et al.* 2011, Gherboudj *et al.* 2011, Baghdadi *et al.* 2012, Chai *et al.* 2015, Gorrab, Zribi, Baghdadi, Mougenot, Fanise, *et al.* 2015, Yue *et al.* 2016). At large scales such as in the Kilombero catchment, however, multiple acquisitions covering the entire catchment would be required. The temporal acquisition interval of SAR images is rather long (RadarSat-2 24 days, TerraSar-X 11 days) and hence acquisitions within the catchment would be on different dates, introducing a challenge in estimation of the dynamic soil moisture. Additionally, collecting in situ data for validating the retrieved soil moisture is impractical at large scales. This prompts the application of large-scale remotely sensed observations (passive and active satellite sensors) and land surface model soil moisture products due to their high temporal resolution and their availability in the processed form. Land surface models use multiple land models, vegetation, and meteorological inputs to generate ensemble estimates of land surface states and fluxes such as soil moisture (McNally *et al.* 2017). Passive sensors (radiometers) record brightness temperatures, have a low spatial resolution and a high temporal resolution. The active sensors on the other hand record backscatter; they have a low temporal and high spatial resolution (Wang and Qu 2009, Chen *et al.* 2014, Srivastava *et al.* 2016).

Start of Season (SOS) marks the beginning of the photosynthetic phase of vegetation and is a crucial parameter in monitoring food security since a delayed planting time translates to yield reduction since the crops do not optimally utilize the growing period necessary to grow to maturity (Brown and de Beurs 2008, Vrieling *et al.* 2013). Associating the SOS with soil moisture illustrates the potential of determining the planting times from soil moisture. SOS derived from soil moisture will thus offer solutions to the uncertainty introduced by changing seasonality explained by changing climatic patterns by monitoring its variability over time hence identify the suitable planting times.

Science-based knowledge on the ecosystem forms the influential basis upon which management plans in support of sustainable use of wetlands can be established, more so considering the agricultural potential of the Kilombero wetland. The research herein was developed based on that concept.

1.1 Remote sensing and groundwater assessment of East African wetlands

Wetlands provide a wide range of ecosystem services among which are the provisioning services including food, water, and fuel (Gardner *et al.* 2009, Ramsar Convention on Wetlands *et al.* 2014). Due to water reliability and fertile soils, wetlands form suitable sites for agricultural production. Increased demand for food caused by increasing population subsequently places wetlands under threat of degradation and hence planning wetland use is critical in sustaining the food water ecosystem nexus (Leemhuis *et al.* 2017). Knowledge of land cover and its spatiotemporal changes plays an important role in planning the sustainable use of wetlands. Additionally, knowledge of water dynamics and spatiotemporal variability of soil moisture, a critical variable in agriculture, is key in determining the potential of vegetation growth ultimately translated to the potential of food production in a region.

The synoptic view, spatial and temporal capabilities of remote sensing sensors provide timely and cost-effective techniques of data acquisition for wetland mapping (Lillesand *et al.* 2004). Earth observation sensors include optical and microwave operating in the visible and microwave parts of the electromagnetic spectrum, respectively. Numerous land cover (Furtado *et al.* 2016, Morandeira *et al.* 2016, Moser *et al.* 2016, Tian *et al.* 2016, p. 20, Chatziantoniou *et al.* 2017, Mahdavi *et al.* 2017) and soil moisture mapping applications (Dabrowska-Zielinska *et al.* 2010, 2016, Jacome *et al.* 2013, Touzi and Cloude 2013) of remote sensing in wetlands have been reported worldwide.

Studies on mapping of wetlands in East Africa are reported though they apply low and medium resolution satellite imagery. Advanced Very High-Resolution Radiometer (AVHRR) data for 1983 to 1984 was used in assessing vegetation dynamics at regional scale over East Africa with results revealing that NDVI is suitable in determining phenology in areas of high spatial variability (Justice *et al.* 1986). Serneels et al. (2001) applied AVHRR in assessing the land cover changes in Mara National Reserve ecosystem in Kenya for the periods between 1981 and 1994. Reports indicated that expansion of agriculture within the buffer of the national park has resulted in decreased vegetation though the reserve is stable in terms of land cover changes since it is a protected area (Serneels *et al.* 2001). Landsat imagery was used in assessing the Omo River delta changes whereby results indicated an increase in sedimentation and decreased water levels in Lake Turkana due to increased agriculture and livestock farming between the years 1979 to 1989 (Haack 1996). Landsat was also applied in assessing the changes in the extent of Lake Baringo catchment in Kenya (Kiage *et al.* 2007). Pressure due to increasing human livestock population was reported to cause a reduction in forested areas and increased sediment deposition between 1986 and 2000.

In the East Africa region, Kuria et al. (2014) mapped the Malinda wetland using optical and radar data for the dry and wet season and assessed the phenological changes in the two-time epochs. High conversion rates for rice-farmed areas, cleared lands and homesteads were reported between May and August 2012 (Kuria *et al.* 2014). Mwita et al. (2013) mapped land use in small (<500 ha) inland valleys and floodplains in Tanzania (Usambara Highlands and Pangani floodplain) and Kenya (Mt. Kenya Highlands and Laikipia floodplain). The results revealed that floodplains were mainly used for cultivation of subsistence crops and cattle grazing whereas highlands are dominated by horticultural farming (Mwita *et al.* 2013). It was also reported that encroached natural vegetation is used as grazing grounds for cattle in the dry season and drained for use as arable land in the wet season. Within the same study, Sakané et al. (2011) characterized the highlands and floodplains based on the biophysical and socio-economic attributes. Physical accessibility, seasonality of flooding, population, market prospects, changes in rural livelihoods, and government policies were reported to influence the use of wetlands for agricultural production (Sakané *et al.* 2011).

Climate changes effecting changes in precipitation and evaporation cause ultimate shifts in soil moisture deficits, surface runoff and groundwater recharge (Calow and MacDonald 2009). With continued unreliability of rainfall and surface water, demand and pressure on groundwater is envisaged to increase (Braune and Xu 2010). There is a need to understand the recharge patterns, which will aid in policy formulation in sustainable use of this renewable water resource (Calow *et al.* 1997). (Moukana and Koike 2008) investigated long-term changes (1987-

2003) in groundwater levels within a catchment using a multivariate model to estimate reduction in groundwater based on land cover. Results indicated that land cover influences groundwater recharge reporting conversion from rice field to urban reduces recharge since infiltration of groundwater is halted by material used for constructing the urban areas (Moukana and Koike 2008). (Pan *et al.* 2011) assessed the impact of land use change on groundwater recharge based on a water-balancing model between the years 1980 and 2005 (Pan *et al.* 2011). Similar to results by (Moukana and Koike 2008), decreased cropland and grassland and increased urban and rural settlements resulted in a decrease in groundwater recharge. In Africa, (Ndou *et al.* 2017) modeled the relationship between depth to groundwater and evapotranspiration with reports indicating high evapotranspiration is expected in areas of low depth to groundwater whereas evapotranspiration decreases as the depth to groundwater increases (Ndou *et al.* 2017).

Understanding the land cover changes occurring within shorter periods is relevant in determining suitable cropping seasons, which in turn provides information to enable planning of cropping season for maximum crop yield. The study herein, sought to generate maps at shorter time intervals showing the changes that occur in the wetland throughout the year and thereby focused on the changes of land cover taking place within land use classes (e.g., an agricultural field left bare during the dry season and cultivated during rainy season). The relationship of groundwater and land cover was assessed additionally. However, the focus was at a smaller areal coverage within the Ifakara region where the piezometer distribution for the groundwater level measurements were distributed.

1.2 Soil moisture monitoring in East African wetlands

Soil moisture influences vegetation dynamics since it is directly associated with their photosynthetic activity. Knowledge of soil moisture distribution in space and time contributes to understanding the land surface systems, which subsequently influence the agricultural management options for improved productivity. For the spatiotemporal variability analysis, classical near-surface soil moisture measurements are obtained by techniques such as Time Domain Reflectometry (TDR) probes. In situ, soil moisture measurements are required in the validation of satellite-based moisture observations. Due to the disparity in spatial resolution of point and satellite footprint, a need to determine the number of soil moisture measurements required to obtain a representative mean value matching the footprint of the satellite arises (Famiglietti *et al.* 2008). Moreover, locations with an accurate representation of the average soil moisture over the entire sampling period require a limited number of measurements in future sampling campaigns (Cosh *et al.* 2008). Soil moisture monitoring with point measurements is time-consuming and tedious hence SAR is adopted for the continued estimation of soil moisture.

Soil moisture studies in East African wetlands included a study by Böhme et al. (2013) who calibrated a soil moisture sensor for wetland conditions (Böhme *et al.* 2013). The regression equation established for estimating volumetric soil moisture from sensor readings reported a lower root mean square error (RMSE) as compared to the RMSE resulting for the use of manufacture's calibration constants. Daniel et al. (2017) analyzed the soil physical properties at 0-20 cm and 20-40 cm depths in the Kilombero floodplain (Daniel *et al.* 2017). Results indicated that bulk density increased with increasing depth due to compaction. Soil organic carbon was reported to be higher in the top 0-20 cm though the quantity was insufficient for arable farming. Higher saturated soil water content was observed at the riparian zone (close to the Kilombero River) as compared to sampling sites further from the river. Kirimi et al. (2016) retrieved soil moisture from TerraSar-X imagery by applying a vegetation index derived from UAV photos in the Malinda wetland (Kirimi *et al.* 2016). Eliminating the effects of vegetation improved the accuracies of retrieved soil moisture in the two-time epochs evaluated.

Distinct from the previous studies, the current study aims at assessing spatial-temporal patterns of soil moisture from repeated measurements over two-time epochs (March-May 2015 and December 2015-February 2016). The temporal stability of the soil moisture data is assessed. Moreover, the minimum number of sampling points necessary to capture field means is examined. SAR application in roughness parameter and soil moisture estimation is tested.

1.3 Application of global soil moisture and NDVI products in East Africa

Soil moisture influences water availability for crop production whereas the growing conditions can be tracked using NDVI. In the assessment of the influence of water availability on vegetation dynamics, soil moisture is preferred as opposed to precipitation because much of the precipitation is lost through interception and runoff while soil moisture is the actual water available for the plant influencing the photosynthetic processes (Chen *et al.* 2014). Soil moisture and NDVI are datasets informing management bodies of vegetation dynamics from which declining yields can be perceived prompting humanitarian actions in cases where a food crisis is envisaged. Remotely sensed and globally modeled products are critical in information provision in data scarce regions.

Soil moisture is a key parameter in studies such as agriculture and hydrology as it regulates infiltration, runoff, and groundwater recharge and determines the water available to plants. Environmental factors influencing soil moisture include meteorological factors (precipitation, temperature), soil properties (such as soil texture, organic matter content), topography (e.g. slope influencing infiltration and runoff), and land use (e.g. vegetated land controlling rate of evaporation from the soil) (Petropoulos and Albergel 2013).

Approaches for soil moisture retrieval from spaceborne satellites include passive and active microwave sensors. Microwave remote sensing is of particular importance in soil moisture studies due to its day and night and all-weather availability Microwave sensors are suitable in soil moisture retrieval since there is a large contrast in dielectric constant of water (~80) and soil particles (~4) resulting in a strong dependency of the dielectric constant on volumetric soil moisture (Moran *et al.* 2004, 2004, Kong and Dorling 2008, Barrett *et al.* 2009).

Transmitted energy returned from the target determines the brightness of radar backscatter. Characteristics of the radar system as well as the target surface influence this energy. These include the incidence angle, wavelength, polarization, surface roughness and geometry of the surface. The roughness of a surface is dependent on the wavelength and incident angle. Long wavelengths penetrate through the canopy whereas shorter wavelengths reflect on the surfaces of canopies. Surfaces appear smooth with longer wavelengths and with increased incidence angles. Smooth surfaces have specular reflection; thus, a small amount of energy returns to the radar, and hence they appear dark. The soil moisture changes the electric properties of the soil influencing the absorption, transmission, and reflection of microwave energy. Reflectivity and image brightness increases with increased soil moisture content (Trudel *et al.* 2012). Decomposition of signals into individual causal components is a challenge due to a combination of surface characteristic in the total signal recorded. This makes microwave sensors ill-posed problem as any combination of signal due to roughness, soil moisture, vegetation can result in the same magnitude of recorded signal (Dawson *et al.* 1997, Mattia *et al.* 2006, Paloscia *et al.* 2008, Álvarez-Mozos *et al.* 2009, Pierdicca *et al.* 2014).

For passive microwave radiometers, naturally emitted radiation within their field of view is measured and recorded as brightness temperature. The field of view of passive sensors must be large to detect enough energy to record a signal. Passive sensors, therefore, have a low spatial resolution and a broad spatial coverage, more suited for analysis of soil moisture over large areas (Petropoulos and Albergel 2013). Inactive approaches, the satellite sensor sends a microwave pulse to a target feature on the ground, on interacting with the feature, the pulse is sent back to the sensor. The signal sent out and that received is compared to determine backscatter of the surface. Spaceborne radiometers platforms (sensors) include AQUA (AMSR-E) from 2002 to 2011, SMOS (MIRAS) launched in 2009, GCOM (AMSR2) launched in 2012 and SMAP launched in 2014. The low-frequency microwave region, 1-10 GHz, is mainly used to obtain soil moisture from near surface layers. The L band is most suitable for observation of soil moisture as it has a high vegetation penetration. Higher frequencies are affected by atmospheric effects and vegetation cover (Barrett *et al.* 2009, Nichols 2011). Active SAR sensors include TerraSar-X (X band wavelength 9.6 GHz) and RadarSat-2 (C band wavelength 5.3 GHz) launched in 2007, Sentinel-1 (C band wavelength 5.3 GHz) launched in 2014.

Soil moisture retrieval models which can be used at local scales include physical based theoretical models, semi-empirical models and empirical models (Nichols 2011). Empirical models are generated from site-specific frequency, incidence angles, and surface roughness and are thus not applicable to different sets of conditions (Srivastava *et al.* 2016).

Theoretical models are based on theoretical perspectives on the trend of backscatter in response to soil moisture or roughness. The Integral Equation Model (IEM) is a theoretical model, and is a function of radar frequency, polarization, incidence angle, dielectric constant; rms surface height, correlation length and the autocorrelation function to quantify the backscatter coefficient (Bindlish 2000, Lievens and Verhoest 2011, Wang *et al.* 2011). The IEM model does not account for double or multiple scattering and is suitable for soil moisture estimation in bare surfaces (Barrett *et al.* 2009).

The semi-empirical combined rules from the theoretical and empirical models e.g. water cloud model, Oh, Shi, Dubois model. The Oh model relates ratios of backscattering coefficients in different polarizations to soil moisture and surface roughness. The advantage of this model is that only rms height is required and there is limited need for field data when using multi-polarized data is used. The Dubois model works well with co-polarized coefficients as they are less sensitive to system noise, are easier to calibrate and are more accurate as compared to cross-polarized backscattering coefficients (Panciera *et al.* 2014, Choker *et al.* 2017). The water cloud model represents backscatter from the vegetation canopy and underlying soil during the phenological cycle of the plant. The total backscatter is a sum of contribution from vegetation, soil and two-way attenuation of vegetation layer (Bindlish and Barros 2001, Lievens and Verhoest 2011, Bai *et al.* 2017).

Other approaches include statistical analysis techniques where estimated soil moisture from radar images are correlated with the actual field measurements and mathematical relationships established (Azad Hossain and Easso 2009). Artificial neural networks which use complex mathematical equations to design desirable correlation between radar backscatter and in situ soil measurements are also applied in soil moisture retrieval (Santi *et al.* 2013, Notarnicola and Solorz 2014, Hassan-Esfahani *et al.* 2015, Hossain and Easson 2016). Change detection approaches involving the use of multi-date imagery to determine the changes from which parameters are derived are applied in soil moisture retrieval (Baghdadi, Saba, *et al.* 2011, Lievens and Verhoest 2012, Gorrab, Zribi, Baghdadi, Mougenot, and Chabaane 2015). Normalized Backscatter Soil Moisture Index (NBMI) is obtained from backscatter measurements at two different times over the same location. For the delta index, the backscatter difference between wet and dry reference image is divided by the dry image backscatter. This index accounts for roughness and vegetation provided they remain unchanged though the imagery must be acquired with the same wavelength and viewing geometry for the backscatter changes to be attributed to changes in soil moisture (Jacome *et al.* 2013). Another change detection technique is principal component analysis where correlated variables are transformed to a reduced number of uncorrelated principal components, some of which are related to soil moisture (Kong and Dorling 2008).

Rainfed agriculture is practiced in the Kilombero catchment (Kato 2007). The onset of rains marks the beginning of the planting season. Aligning water availability with planting times ensures maximization of the growing season to meet the crop water requirement, which results in a good harvest. Therefore, for optimal growth, precise determination of start of season (SOS) is crucial as it dictates the growing period (Brown and de Beurs 2008). In Africa, SOS, end of season (EOS) and length of crop cycle influence the quantity of yields produced by subsistence farmers (Vrieling *et al.* 2013). The SOS and EOS are indicators of reduction in yields and hence early warning parameters in food security monitoring. The length of season (LOS) limits crop choices under rainfed agriculture since the LOS has to fulfill the optimal growing period for the crops selected to complete its cycle to maturity (Vrieling *et al.* 2013). Sparse weather stations and data gaps in climate data limit its use in the determination of start and end of growing season. Remotely sensed NDVI, on the other hand, has a wide coverage, is available for a long time series depending on the launch of satellite sensors, and is thus a preferred approach in seasonality determination and detection of anomalies from long-term averages.

(Rojas *et al.* 2011) identified SOS and evaluated the probability of occurrence of drought using vegetation indices from AVHRR over cropped areas in Africa from 1982 to 2010. Drought probability maps were generated at continental and regional level. The method correctly indicated the 2009/10 drought in East Africa. 8 km GIMMS NDVI3g based on 1981 to 2011 data was analyzed to determine the length of growing period over Africa (Vrieling *et al.* 2013). Over lower East Africa, in Tanzania and extending to northern Mozambique, LOS ranged from 180 to 300 calendar days with the SOS and EOS reported as November-December and May-June (Vrieling *et al.* 2013).

Models relating rainfall based and remotely sensed SOS in west Africa were investigated (Brown and de Beurs 2008). The results indicated MODIS NDVI based SOS at 16-day composite provided the closest fit to observed SOS while shorter NDVI composites presented a challenge in SOS determination. Shorter compositing periods introduce artifacts in the composited NDVI due to cloud residue. A seasonal lag of 2-6 weeks of vegetation response to rainfall was reported for the wet savanna region in East Africa where AVHRR NDVI and outgoing longwave radiation, an indicator of rainfall, from NOAA were applied in the investigation (Shinoda 1995).

Soil moisture has a more direct link to plant growth than precipitation since some amount is lost through runoff or seepage to groundwater. In situ soil moisture networks for validation and calibration of remotely sensed and modeled soil moisture products over Africa are scarce raising interest in the application of globally available products. African Monsoon Multidisciplinary Analysis (AMMA) network in West Africa was among the data assessed in testing the performance of remotely sensed ASCAT and SMOS and ECMWF modeled SM-DAS-2 soil moisture product (Albergel, de Rosnay, Gruhier, *et al.* 2012). The ten in situ observation stations located in West Africa reported a correlation of 0.69, 0.74 and 0.73 for SM-DAS-2, ASCAT and SMOS soil moisture products.

In East Africa, (McNally *et al.* 2016) compared CCI-SM with modeled NOAH and VIC over rainfed agricultural areas. Pixel-wise Spearman rank correlations of CCI soil moisture from 1992 to 2013 was 0.76 and 0.85 for NOAH and VIC. The focus was mainly in Kenya and Ethiopia. Caveats to address while applying global products over large regions include low spatial and temporal coverage which ultimately determine the analysis methods adopted (McNally *et al.* 2016). Other soil moisture networks in Africa include the COSMOS network in Kenya and CARBOAFRICA in Sudan (Zreda *et al.* 2012, Ardö 2013).

While the studies mentioned before were at a regional and continent scale, the current study focuses on catchment scale. The lack of soil moisture networks in the area supported the use of globally available moisture data. Different moisture products portray varied values due to the differences in acquisition sensors and retrieval algorithms. Understanding the discrepancy across independent datasets poses a challenge, as there are no networks to assess the accuracy of the products. Triple collocation analysis assessed the relative errors of the independent datasets. Least square merging was applied in the generation of a merged soil moisture product. The impact of soil moisture on vegetation dynamics is assessed by estimating the variability in the relationship between soil moisture and vegetation. Unlike the study by (Chen *et al.* 2014), who used low-resolution AVHRR

NDVI in quantifying soil moisture NDVI relationship, the current study derives vegetation information from higher resolution 16-day MODIS NDVI composites at 250 m spatial resolution (Chen *et al.* 2014).

1.4 Aim of study

Wetland regions are highly dynamic systems. Lack of maps describing the allocation of different land cover types throughout the year is a major impediment to managing wetland activities. Knowledge of land cover types and their occurrence time within the year is required to monitor water availability and devise techniques of increasing production to improve food security and ensure sustainable use of wetlands. Land cover dynamics maps act as spatial records that aid in decision planning with respect to land uses. Documenting the short-term land cover patterns, however, aids in informed decision making on seasonal land use planning. Groundwater provides base flow for rivers and is a source of water for agricultural production in the face of changing climate. Understanding the land use-groundwater relationship gives insight into periods over which alternative water sources are required to ensure continued agricultural production. The current study sought to generate land cover maps from optical imagery over the study area and qualitatively examine the effect of land cover on depth to groundwater. Presence of clouds poses a challenge in mapping from optical images. Inconsistency in usable optical images due to cloud cover and limited data availability of single sensor systems motivated the use of a multi-sensor approach in mapping land cover whereby three sensors; RapidEye, Sentinel-2, and Landsat, for the periods 2013 to 2016 was adopted.

The cloud coverage challenge prompted the use of all-weather, free access Sentinel-1 SAR imagery. However, for the period under study, consistent acquisitions were only available as single polarized images for 2015 and 2016. The study therefore assessed the performance of multiple classification algorithms on consistently acquired Sentinel-1 images.

Understanding soil moisture variation in space and time at the plot scale requires high resolution spatial and temporal measurements subsequently aiding in decision making concerning farm management practices. With the planned SAGCOT agricultural intensification strategies, knowledge on field-specific soil properties is crucial in fostering precise farming activities such as in application of fertilizers, whereby infield variations are put into consideration to ensure enhanced crop growth since applications are as per the soil and crop needs (Milder *et al.* 2013). Subsequently, increased yields improve food security as well as the livelihoods of the population.

In situ soil moisture data is also required for validation of satellite-based soil moisture products. Satellite-based soil moisture is in form of pixels whereby soil moisture within the pixel is averaged to one value. In situ soil moisture, on the other hand, could have several measurements within the pixel coverage. Discrepancies in spatial resolution between the satellite footprint and field point data prompt the determination of number of sample points required to adequately capture the mean within a defined areal coverage (Jacobs 2004, Famiglietti *et al.* 2008). Identification of fields whose average is stable in time reduces the number of measurement points in future field campaigns for validating satellite-based soil moisture products (Cosh *et al.* 2008, Brocca *et al.* 2009).

Extending the soil moisture assessment to a larger site prompts the application of SAR imagery in soil moisture retrieval due to its vast coverage as soil moisture point data collection is time-consuming and impractical for large areas. Soil moisture and soil roughness influence backscatter recorded by SAR over bare soils. Soil roughness data collection is time-consuming and requires skills to analyze. Moreover, the small farms with varying management practices in Ifakara instigate the development of an a priori soil roughness parameter. In this study, a semi-empirical model by (Baghdadi *et al.* 2016) generated by incorporating data from multifrequency SAR data for soil moisture retrieval is inverted to derive the roughness parameter (Baghdadi *et al.* 2016).

Soil moisture data scarcity necessitates the application of remote sensing based monitoring techniques and physically based land surface models over the Kilombero catchment. Due to different retrieval algorithms, the available independent soil moisture products are characterized by systematic differences. The reliability of the products is impossible to evaluate due to lack of in situ soil moisture networks. This prompts the generation of merged soil moisture products that take into account the discrepancy estimates between the products.

Estimating variability in the relation between NDVI and soil moisture portrays the impact soil moisture has on vegetation dynamics. Continued monitoring of soil moisture from globally available products to determine the SOS offers a solution to the confusion in planting dates that the farmers experience due to seasonal differences in timing and amount of rainfall, caused by climate variability and climate change (Brown *et al.* 2010). Ultimately, precision in planting times ensures maximum utilization of rainwater in meeting the plant water requirements thus ensuring maximum yields are obtained. Moreover, a delay in SOS is an indicator of declined crop production, which is critical in food security monitoring (Brown and de Beurs 2008).

This study was executed under the BMBF funded GlobE project "Wetlands in East Africa – Reconciling future food production with environmental protection" whose goal is assessing the potential of utilizing wetlands for increased and sustainable food production (https://www.wetlands-africa.uni-bonn.de/). The multi-disciplinary project consists of a series of diverse groups with backgrounds in agronomy, hydrology, economics, ecology, social sciences and remote sensing.

The research presented herein is within the remote sensing discipline. The main purpose of the research was to assess land cover changes within the tropical floodplain in the Kilombero wetland and evaluate the soil moisture information based on in situ data, SAR and global datasets derived from different sources.

The research seeks to:

i. Assess the land cover dynamics to understand the land cover changes throughout the year using optical data.
ii. Examine the performance of VV single polarized backscatter images, grey level co-occurrence matrix (GLCM) texture images and their PCA derivatives in monitoring land cover at high spatial and temporal resolution.

iii. Explore the variation of soil moisture over space and time. Derive a roughness parameter as an input for estimating soil moisture from SAR data.
iv. Evaluate the performance of global soil moisture products and estimate the variability in the relation between soil moisture and SOS

The hypotheses guiding the study are as follows:

i. Depth to groundwater has an influence on spatiotemporal land cover dynamics.
ii. Single polarized SAR images have sufficient information enabling land cover classification.
iii. There exists a variation in soil physical properties along a hydrological gradient within the Kilombero floodplain.
iv. Surface roughness has a big effect on the accuracy of soil moisture derived from SAR imagery.
v. Global soil moisture products can be utilized to infer the start of planting seasons.

1.5 Organizational structure of the thesis

The thesis is organized into seven chapters (Figure 1.1).
Chapter 1 has presented information about the research objectives, hypotheses, and the justification of the study.

Chapter 2 introduces the description of the Kilombero catchment in terms of location, climate, soils and geology, vegetation, population and economic activities.

Chapter 3 focuses on land cover dynamics using optical imagery. Time series of RapidEye, Sentinel-2, and Landsat images were classified with the aim of understanding changes in land cover with respect to depth to groundwater over a period of three years (2013-2016). Cloud coverage remained a great challenge for the optical images and thus radar images were adopted. The areal scale of the study was at the extended plot scale (Table 1.1). Ifakara area is within the stated extent and is comprised of varying hydrological zones, subdivided, based on their flooding characteristics. The riparian zone is completely flooded during the long rainy season, the middle zone is partially flooded whereas the fringe is not flooded (Burghof *et al.* 2017).

Chapter 4 deals with the use of single-polarized free access Sentinel-1 radar images for the period 2015-2016 in monitoring the wetlands. For the study period, dual polarized images were hardly available and hence methods of improving classification accuracies were tested. Grey Level Co-occurrence images and Principal Component Analysis derivatives of the texture images are thus employed in delineating land covers using multiple classification algorithms to test their performance in wetland mapping.

Chapter 5 addresses spatial-temporal variability of the soil moisture at plot scale. In situ soil moisture was collected within selected sample fields in three hydrological zones, the fringe, middle and riparian zones. The spatial variability of physical properties including texture, organic carbon, and bulk density are analyzed premised on the notion that varying soil texture influences the soil moisture behavior.
At large scales, however, collection of in situ soil moisture is tedious. Moreover, soil moisture is highly dynamic and thus a fast measurement method required. Microwave satellite imagery from RadarSat-2 and TerraSar-X Sensors are thus employed in the derivation of soil moisture. Roughness information is crucial in soil moisture estimation from radar images. Roughness parameters are derived are derived from semi-empirical models and their performance tested. The assessment in this objective was at plot scale.

Chapter 6 is dedicated to assessing the performance of global soil moisture products over the Kilombero catchment. Global soil moisture products provide a solution in high temporal soil moisture acquisition at large scales. Determining the Start of Season from global moisture products is also evaluated to assess the plausibility of determining the SOS from soil moisture data.

Chapter 7 and 8 offers an overall synthesis of the topics covered. A conclusion of the research and recommendations for further research are given.

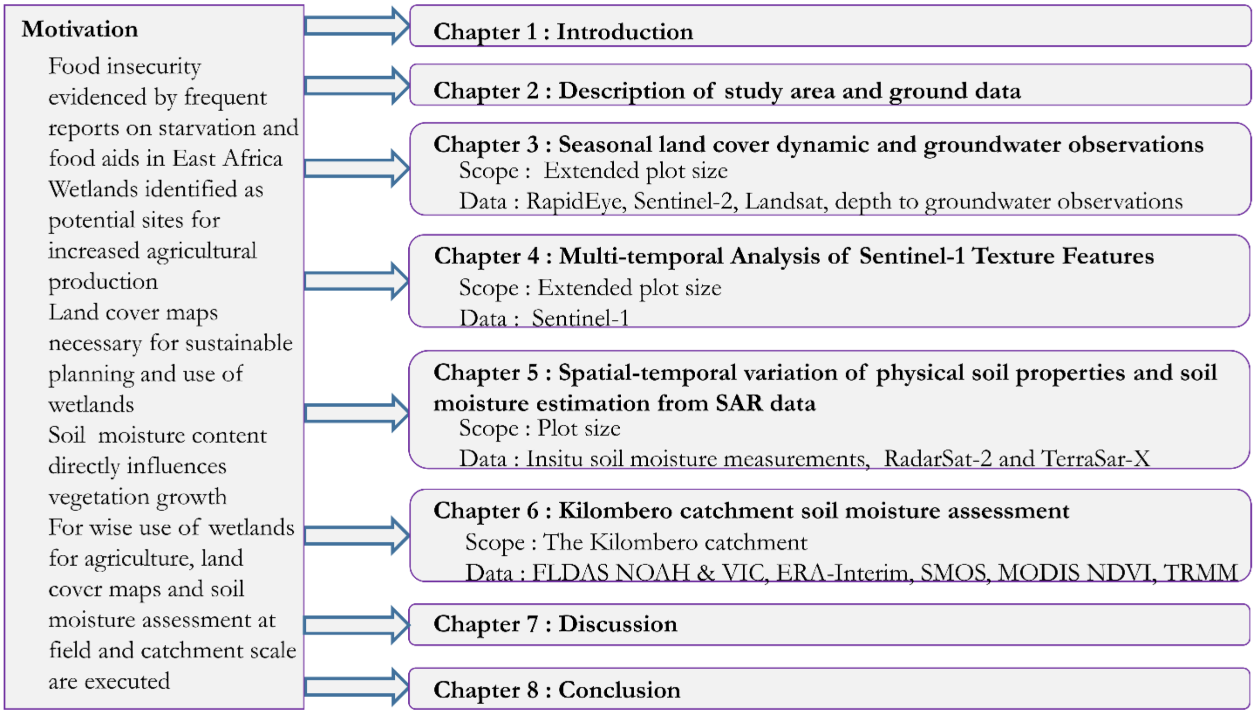


Figure 1.1: Organigram of research. Optical and Radar Remote Sensing (RS) in land cover mapping and Soil Moisture (SM) assessment formed the core of research

Table 1.1: Spatial scales addressed in the study

Spatial unit	Size	Aspects covered
Plot scale	734-19,206 m^2	Soil moisture spatial temporal variability of in situ data Soil moisture estimation from SAR data
Extended plot scale	13*5 km^2	Land cover dynamics using optical and SAR data
Catchment	40,240 km^2	Determining phenological parameters from global soil moisture products and NDVI

2. Study Area and Ground Data

2.1 Study Area

2.1.1 Location

The Kilombero valley floodplain is the largest seasonal freshwater lowland floodplain in East Africa situated in the Kilombero and the Ulanga Districts in the Morogoro region, south-central Tanzania. The floodplain is about 260 km and 52 km covering approximately 626,500 ha at high water. The area lies between 210 to 400 m a.s.l. To the northeast, there are the Udzungwa Mountains that have elevations up to 2580 m a.s.l. and to the southwest, it is bordered by the Mahenge highlands with elevations of up to 1520 m a.s.l. Ifakara, the largest settlement in the floodplain is located within the Kilombero wetland between 36°41'E, 08°05'S and 36°44' E, 08°12' S with an elevation of 230-270 m a.s.l. (Figure 2.1).

The Kilombero River divides the Ulanga and the Kilombero districts. The Ruhudji River and the Pitu River originating from Mbeya and Iringa mount ranges border the floodplain to the south ("Information Sheet on Ramsar Wetland (RIS)," 2002). The Ruhudji River receives water from several tributaries and then subdivides on the floodplain. West of the Kihansi River exists a zone of permanent swamps approximately 45 km long extending up to four kms from the bank. At Ifakara, rivers flow into the Kilombero River following the Udzungwa escarpment. The Kilombero River then leaves the floodplain to the south-east on the border of Selous Game Reserve and borders Mikumi national park to the north. The river joins the Luwegu River 56 km on as Rufiji River. This then forms the Rufiji River Delta at the outlet into the Indian Ocean. The delta comprises of seagrass, coral reef and the largest mangrove forest in East Africa. The surrounding mountains and highlands are important catchment areas crucial to the hydrology of the wetland.

The rivers flow through the evergreen forests and the Miombo woodlands and then enter the floodplain forming seasonal and perennial rivers, oxbow lakes, swamps, ponds, and lakes. Perennial rivers flow from the north to the south-flowing into the Kilombero River, which flows from west to east.

The floodplain has a series of swamps, ponds and lakes, river, riverine forest, grassland habitats with a section of Miombo woodland and groundwater evergreen forest. The major part of the wetland has flooded grassland and bushland.

The ecosystem regulates the flow of the Rufiji River as it maintains the slow rate of rise and fall of the water levels and provides nutrients and sediments for downstream areas at Mafia Rufiji mangrove, seagrass and coral reef complex. The annual flooding is important for the maintenance of soil fertility creating rich agricultural lands (Kato 2007). Ponds and flooded grasslands upstream are important breeding sites for fish during the wet season.

The study was carried out at different scales i.e. catchment (Figure 2.1), extended plot scale (Figure 2.2) and plot scale (Figure 2.2). The catchment was selected due to its potential in increased food production (Kato 2007). Due to the vast extent of the catchment, it would be impractical to set up experiments at all locations of the catchment. Therefore additional studies focused on an extended plot size in Ifakara chosen because of the high variability of the water table with extensive flooding during the rainy season, subsequent land use patterns (Burghof *et al.* 2017). Moreover, the alluvial fan at Ifakara study site acts as a natural dam that retains water during the rainy season. The selected study site is also within the area of focus of the project GlobE Wetlands of East Africa (https://www.wetlands-africa.uni-bonn.de/). The plot scales were selected such that they were within the three hydrological zones namely Riparian, Middle, and Fringe. At the extended plot and plot scale, the farms are privately owned by individual farmers, plots are highly fragmented and used for small-scale arable farming and are highly variable in management practices.

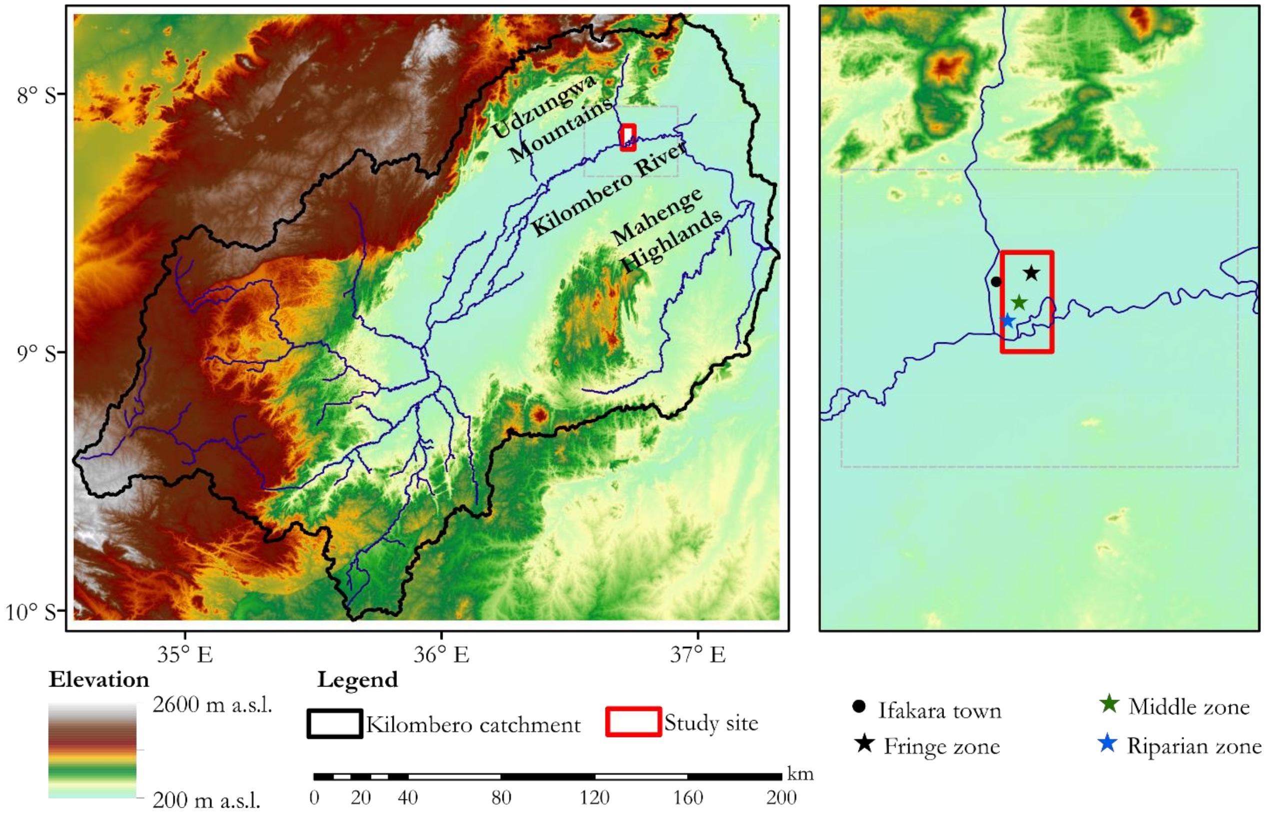


Figure 2.1: The Kilombero catchment and extended plot scale maps. Elevation information is from a TanDEM-X image. The locations of Ifakara town and the locations of the sample fields are shown.

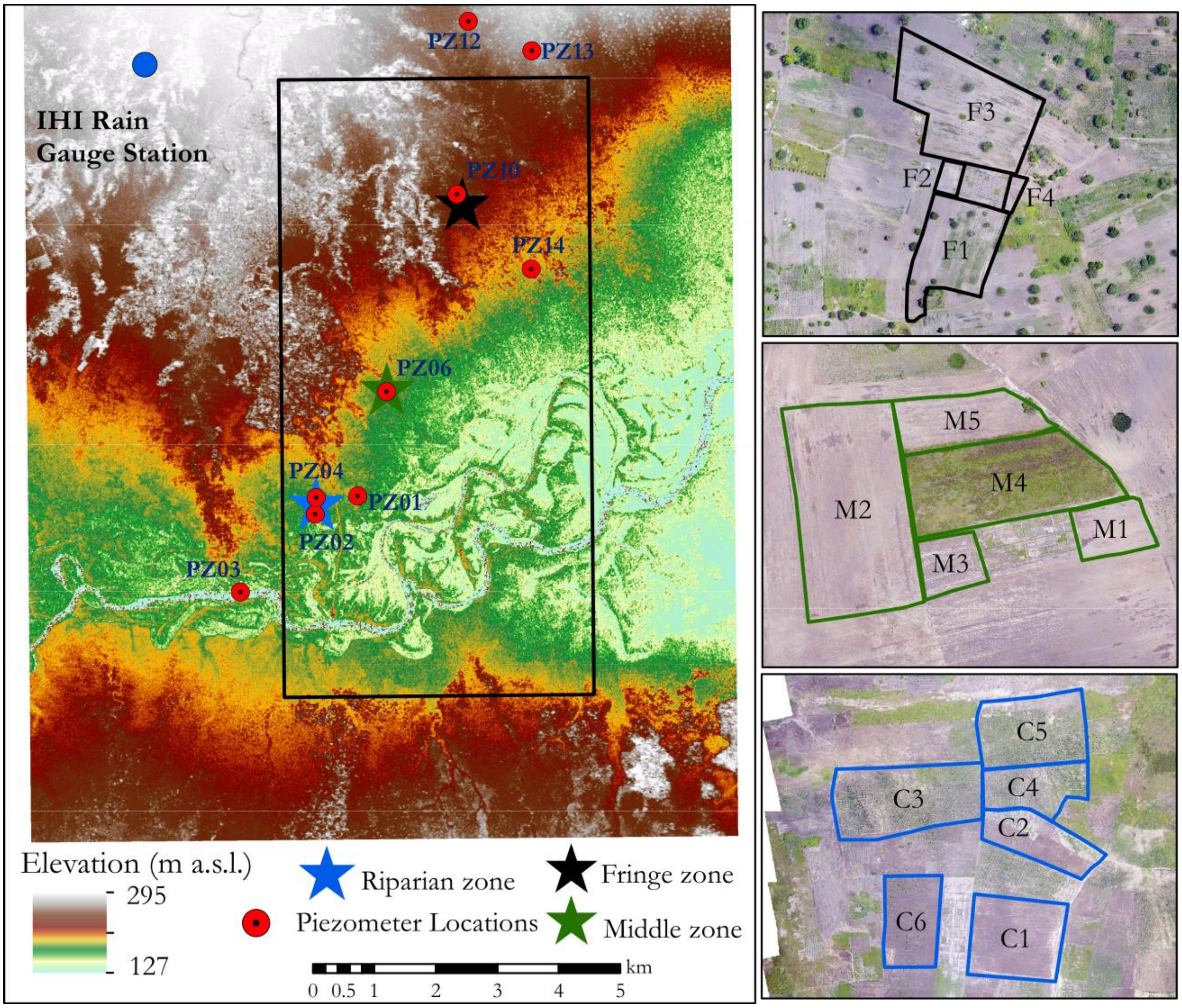


Figure 2.2: Location of the Ifakara study site within the Kilombero Valley, Tanzania. The riparian, middle and fringe zones are hydrological zones categorized based on the flooding pattern during the rainy season. Elevation information, piezometer (PZ) locations and the location of the rain gauge installed at Ifakara Health Institute (IHI) is presented. The sample fields used in the study are indicated. C1-C6 are riparian fields, M1-M5 are middle fields while F1-F4 are fringe fields. The fields are overlaid on UAV photos taken in September 2014

2.1.2 Climate

The climate of the Kilombero catchment falls under the tropical savanna climate as per the Köppen-Geiger classification (Peel *et al.* 2007). The Kilombero catchment has bimodal rainfall with mean annual rainfall of between 1200 – 1400 mm with an average daily temperature of 22-23°C with the climatic conditions characterized as subhumid (Koutsouris *et al.* 2016). The rain season is between November and May (Kangalawe and Liwenga 2005). The region experiences a high intensity of rains between March and May and a lower rain intensity between October and December. Flood peaks occur during the April-May while the dry season occurs between June and October. The wettest months being March and May while the driest months are in August to September (Koutsouris *et al.* 2016). However, from historical data, the driest month is August whereas the greatest precipitation occurs in April. The average annual temperature is 24.8 °C with the warmest month being November and the lowest temperature occurring in July (Climate-data.org 2015) (Figure 2.3)

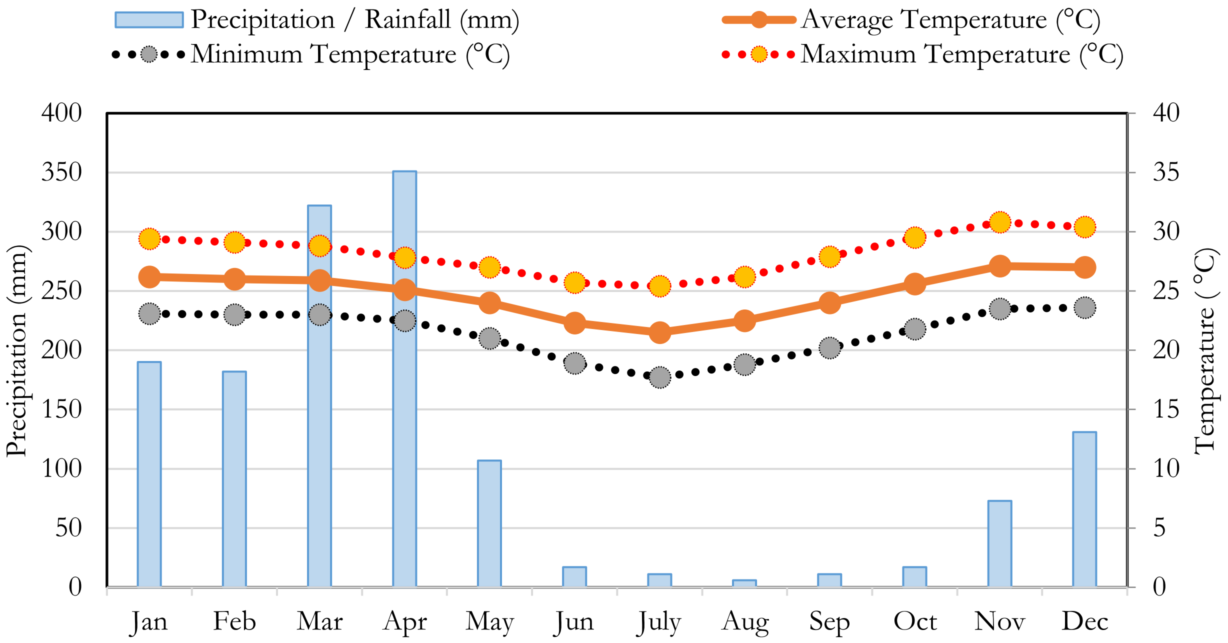


Figure 2.3: Climate data from climate-data.org based on data for 1982 to 2012

2.1.3 Vegetation

The Kilombero catchment falls under the alluvial plains agro-ecological zone (National Bureau of Statistics 2015). There are evergreen forests in the Udzungwa Mountains on the northern part of the Kilombero catchment which act as a natural water storage to the Kilombero floodplain. In the foothills of the mountain, there are Miombo woodlands with patches of closed forests (Kangalawe and Liwenga 2005). Agricultural lands, open forests and the wetlands are at the lower altitudes of the Kilombero Valley. On the edges of the floodplain, open woodlands, evergreen forests and teak plantations are present. A land cover map of 2014 is presented in Figure 2.4.

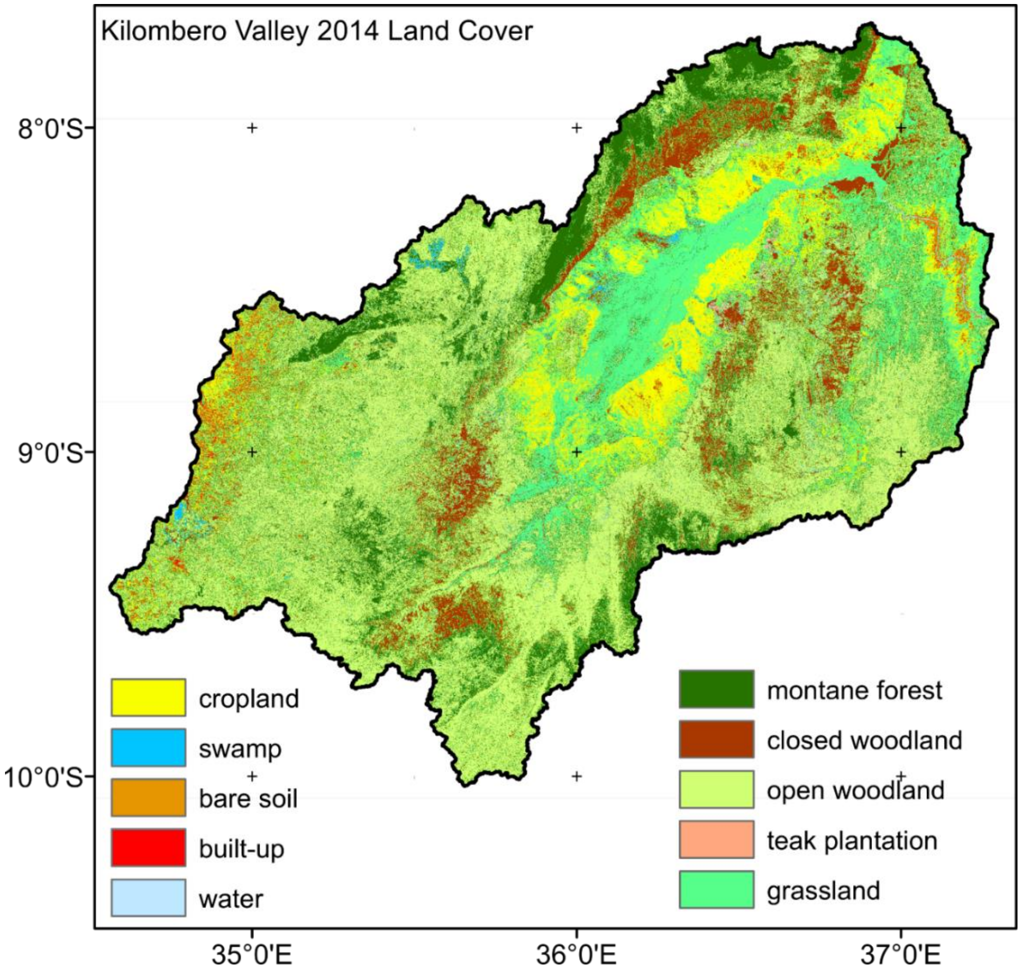


Figure 2.4: The Kilombero Valley land cover map **(Leemhuis *et al.* 2017)**

2.1.4 Geology and soils

The southern side of the catchment is largely covered by Acrisols (Figure 2.5 a). The lower section is well-drained soils. Acrisols are acidic with low availability of nutrients. They have a weak physical structure and a high susceptibility to rainfall erosion.

Arenosols and fluvisols are located within the floodplain (Figure 2.5 a). The floodplain is poorly to moderate excessively drained. Arenosols have a high sand content with low organic matter and low nutrient availability. They have a low available water capacity and are susceptible to leaching. Fluvisols are fertile and do not impair availability of nutrients. Their texture varies from coarse sand to heavy clays in basin areas.

The northern part of the catchment is characterized by Nitisols, is well drained, has gneiss and migmatite rocks and is within a medium gradient landform since it forms the base of the Udzungwa Mountains (Figure 2.5 b, c). Nitisols have a high clay content with a good physical structure. The available water capacity is high with high organic matter.

The eastern part of the catchment is characterized by cambisols, with sandstone lithology, is well drained and mainly lies within the floodplain (Figure 2.5 b). Cambisols have a good structural stability, a high porosity with a good water holding capacity. The soil textures are loamy to clayed and the organic matter is moderate to good. The western part of the catchment is characterized by lixisols and small percent coverage of nitisols, it is well drained and lies within a plateau. The physical characteristics of rocks found here are granodiorite and acidic pyroclastic (Figure 2.5 b, d). Lixisols are strongly weathered with low levels of available nutrients and low nutrient availability. They are free draining and lack evidence of saturation.

To the northeast, a small percentage of the catchment has leptosols soils that are excessively drained lying in a high gradient escarpment. Leptosols are free draining and have low water holding capacity due to their stoniness (Bouwman 1990).

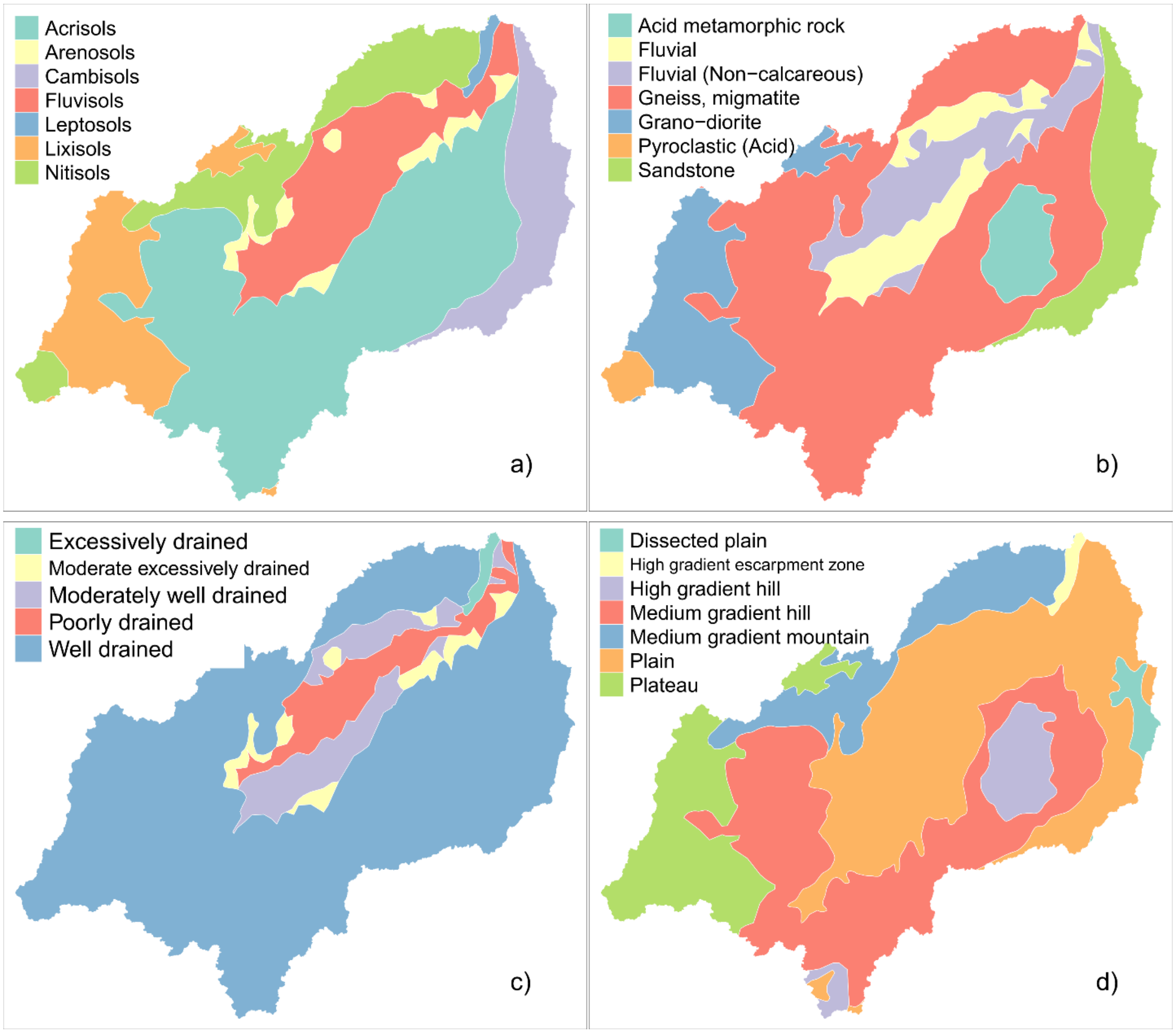


Figure 2.5: Soil and Terrain features for the Kilombero catchment. a) Soil map, b) Lithology map, c) Drainage map, d) Landforms. ("e-SOTER | ISRIC," 2018)

2.1.5 Population

The 1967 population census recorded a population of 12,313,469 while that for 1978 was 17,512,610. In 1988, the population was 23,174,336, more than double the population recorded in 1967. More specifically to regions, Morogoro (consisting of Morogoro Municipality, Gairo, Kilombero, Kilosa, Morogoro Rural, Mvomero and Ulanga districts) has a population of 939,264 and 1,279,931 in 1978 and 1988 respectively. The growth rates in Morogoro were 2.9 and 2.6 in 1967/1978 and 1978/1988 respectively. In 2002, Morogoro recorded a population of 1,753,362 and in 2012, an increased to 2,218,492 people. Narrowing down to Kilombero region, 187,593; 321,611 and 407,880 people were reported in the 1988, 2002 and 2012 census. This marked growth rates of 3.9

in 1988/2002 and 2.4 in 2002/2012 census (National Bureau of Statistics 2015, Tanzania National Bureau of Statistics 2018).

2.1.6 Economic activities

Subsistence agriculture is the main activity in the Kilombero Valley. Over the years, the wetland and surrounding buffer areas are for agriculture, livestock keeping and fishing. Rice cultivation is common in naturally inundates areas whereas in the Miombo region, teak and maize are mainly grown, maize being the flood recession crop during the dry season Livestock keeping, fishing, brick making, forest products, business and casual laboring are other livelihood options for people in the Kilombero Valley (Siima *et al.* 2012). Fishing locations include the Kilombero River, in small rivers in the wetland and in flooded crop fields during the rainy season. Teak plantations by the Kilombero Valley Teak Company (KVTC) play a critical role in forest management and protection of the wetland ecosystem (KVTC 2009). Additionally, the woodlands provide refuge for the seasonal wildlife migrations from the adjoining Selous Game Reserve located to the south-west of the Kilombero Valley. The Kilombero Valley is suitable for sugarcane farming with the Kilombero Sugar Company Limited (KSCL) owning vast sugarcane plantation (ISG 2016). However most crops are produced by small-scale farmers (Kangalawe and Liwenga 2005, National Bureau of Statistics 2015).

The Kilombero floodplain is important in conservation and biological diversity and is an important tourist destination for those visiting the Udzungwa Mountains National Park, Mikumi National Park, Kilombero game controlled area and Selous Game Reserve. It is home to the rare Puku Antelope and its survival depends on the floodplain dynamics. It forms a dry season habitat for large mammals, particularly the elephant and buffalo from the Selous ecosystem (KVFP 2002, Mombo *et al.* 2011). Fish in the myriad of rivers within the floodplain is a common economic activity which provides food and a source of livelihood for the local population. During the harsh dry season, the area is vital in provision of water and food for the large mammals and during the dry seasons, it acts as a spawning area for fish. The ecosystem holds more than 20,000 waterbirds in the wet season.

In terms of agricultural development, for the Kilombero Valley, the Tanzania government initiated Kilimo Kwanza ("Agriculture First") to address national food security and economic growth. It is implemented through Agricultural Sector Development Program (ASDP) and Southern Agricultural Growth Corridor of Tanzania (SAGCOT). SAGCOT is a private-public partnership launched in 2010 at the world economic forum to support and mobilize private sector investment to realize the goals of Kilimo Kwanza (Agriculture First) which is agricultural development. Another multi-agency program, feed the future of the US supports the development of irrigation schemes, infrastructural improvements, technological training such as System of Rice Intensification (SRI) and nutritional education.

Local organizations working with farmers to support sustainable resource management include: the Kilombero Valley Environment and Development Organization (KIVEDO), Association of Kilombero high quality Rice

Growers (AKIRIGO), Rural-Urban Development Initiative (RUDI), Mtandao wa Vikundi vya Wakulima Tanzania (MVIWATA) (National Networks of Farmers Groups in Tanzania) (http://www.mviwata.org/).

2.2 Ground Data collection

2.2.1 Training and validation data

Training and validation data for the classification was collected using GPS devices and UAV photo campaigns. A LumixGX 1 camera was mounted on an Unmanned Aerial Vehicle (UAV) and the MaVinci Desktop software was used for flight planning. The flight periods were 29th May to 1st June 2014, 24th to 26th September 2014 and the last flight campaign was from 27th February to 2nd March 2015. Post-processing of the geotagged UAV photos was performed with the Agisoft PhotoScan software.

Training and validation sample points identifying the land cover classes were determined with the Garmin Etrex 30 GPS with a horizontal positional accuracy of <3m (Garmin International 2005) for the periods between March and May 2015 and December 2015 to February 2016.

2.2.2 Monitoring depth to groundwater

Despite the large size of the Kilombero Valley, the study site focused on a small area over which the piezometers were installed. Depth to groundwater was monitored hourly by pressure sensors installed in 10-cm-diameter piezometers at each hydrological position from March 2015 to June 2016 (Gabiri *et al.* 2018). This research included observations from nine piezometers.

Piezometer 1, 2 and 4 were in the riparian zone, piezometer 3 was located a few meters from the Kilombero River, piezometer 6 was in the middle zone and piezometer 10 and 14 in the fringe zone while piezometer 12 and 13 were located on higher regions above the study site.

2.2.3 Soil moisture sampling design

Soil samples were collected at three hydrological zones namely the riparian, middle and fringe. The riparian is closest to the Kilombero River whereas the fringe is located furthest away from the Kilombero River. The experimental campaigns involved the use of Time Domain Reflectometry (TDR) probes in the measurement of 5cm soil moisture. The soil moisture sensor was the ML3 ThetaProbe and HH2 recording device (Delta-T Devices).

The data was collected during the wet season between March and May 2015 and the beginning of wet season between December 2015 and February 2016. In situ measurements collected 8 times between March and May 2015. Repetitions were made 10 times between December 2015 to February 2016. In each of the fields, 10x10 m grids were established for measurement of soil moisture. The sizes of sample fields were 4,272 to 8,290 m^2

for the riparian, 1,523 to 10,727 m^2 for the middle and 734 to 19,206 m^2 for the fringe fields. In addition, soil samples were collected by core method at 10cm depths for determination of percentages of clay, sand, silt, soil organic carbon (SOC) and bulk density. Sampling locations varied from 9 in the middle and riparian zones and 20 in the fringe zones determined by the shapes of hydrological zones. Each sampling location was designed not more than 750m away from the subsequent sampling spot. Texture was determined using the laser method, bulk density by core method and SOC was determined using the modified Walkley-Black method as described by Daniel (Daniel *et al.* 2017).

2.2.4 Soil roughness sampling design

Roughness information was collected a day prior or after the satellite overpass. A 1m pin profiler, with 41 pins at 25mm spacing aided in roughness data collection. A pin profiler consists of equally spaced pins lowered into the surface to give an impression of the surface roughness. The heights of the pins are read against a white background attached to the profiler (Cihlar *et al.* 1987).

One set of roughness profile measurements was taken in fields with areas of less than 2,000 m^2, while 2 sets were taken in every other field. A set of roughness profile consisted of 2 measurements in the NS direction and two in the EW direction such that the total roughness length measured was 2m in the NS and EW direction.

3. Assessing seasonal land cover dynamic and groundwater observations

3.1 Introduction

Sub-Saharan Africa is expected to account for 20% of the world population by 2050 (Alexandratos and Bruinsma 2012). Sustainable Development Goal 2 (SDG2) is geared towards eradication of hunger through increased food production (Le Blanc 2015). Agricultural production systems are adversely affected by climate variability causing extreme weather events (increased amounts and variation in rainfall, increased temperature) and changing seasons (Peltonen-Sainio *et al.* 2008, Food and Agriculture Organization of the United Nations 2014, Vizy *et al.* 2015). Low crop productivity due to water stress and degraded soil fertility are among the consequences of the changing climate. Land is under increasing pressure driven by urbanization and population growth. Subsequently, need for land to expand agricultural production has led to deforestation and encroachment in protected areas like wetlands, negatively affecting the regulatory functions of the ecosystem.

Tanzania is an agricultural country with 80% of the population depending on subsistence agriculture (Shetto *et al.* 2007). Kilombero Valley was identified as a potential agricultural area to enhance food security and meet the SDG 2 (Beck 1964, Kato 2007). A large seasonally flooded alluvial floodplain is within the Kilombero Valley. Expansion of agricultural land in the floodplain is restricted due to the existence of protected areas i.e. the Kilombero Game Controlled Area, the Selous game reserve, and the Udzungwa National park. Furthermore, changing climatic patterns affect agricultural production causing extreme periods of dryness, floods, and changes in seasonality.

The increasing demand for food, shortage in arable land and unpredictable climate conditions in East Africa have recently created a shift from upland cultivation to Kilombero Valley. This highlights the need for wetland monitoring and the generation of seasonal land cover maps for sustainable conservation policies (Eitel *et al.* 2011, Seki *et al.* 2017). Within the study area, few studies on the status of the wetlands are existing. Monitoring spatial phenomenon through ground survey method is tedious and not cost effective. With their synoptic view, spectral and temporal capabilities of remote sensing sensors provide an alternative timely and cost-effective method of data acquisition.

Remote sensing of wetlands in East Africa has had a long history of study. The evaluation of vegetation dynamics at regional scale over East Africa with Advanced Very High Resolution Radiometer (AVHRR) data for 1983 to 1984 revealed that NDVI is suitable in determining phenology in areas of high spatial variability (Justice *et al.* 1986). (Haack 1996), reported an increase in sedimentation and decreased water levels in Lake Turkana due to increased agriculture and livestock farming between the years 1979 to 1989. AVHRR was applied in assessing the land cover changes in Mara National Reserve ecosystem in Kenya for the periods between 1981 and 1994. Reports indicated that expansion of agriculture within the buffer of the national park has resulted in decreased vegetation though the reserve is stable in terms of land cover changes since it is a protected area (Serneels *et al.* 2001). Within the Lake Baringo catchment in Kenya, pressure due to increasing human livestock population were reported to cause reduction in forested areas and increased sediment deposition between 1986 and 2000 (Kiage *et al.* 2007).

More recently however, within the Kilombero Valley, land cover change was assessed by (Ntongani *et al.* 2014) using local knowledge. From their research, the local population reported a conversion of forested areas and grasslands to cultivated areas for a period of more than 30 years. Recently, (Leemhuis *et al.* 2017) analyzed long term land changes (1994-2004 and 2004-2014) in the whole Kilombero Valley while (Seki *et al.* 2017) assessed land cover changes (1990, 1998 and 2011) in the Kibasira Swamp in the Kilombero Valley with both studies revealing an increase in agricultural land. While these studies cover periods between 8-13 years, they are relevant in portraying the long-term changes. The land cover dynamic within one year is important for determining the suitable cropping seasons in this rain fed agricultural region. Understanding the land cover dynamics will therefore aids in recognizing the short-term land use patterns subsequently promoting monitoring programs to ensure sustainable use of the wetland to increase agricultural production (SAGCOT n.d., Milder *et al.* 2013). Furthermore, to our knowledge, there is no published work on land cover mapping at high temporal resolutions over East African wetlands. Knowledge on seasonal land cover is important to farmers as it aids in understanding the patterns experienced throughout the year within the wetland and hence aid in identification of times and locations with optimum vegetation growth.

Groundwater provides base flow for rivers, dilutes effluents and is a source of water for domestic and commercial uses (Lerner and Harris 2009). A strong feedback between groundwater and land use exist influencing water availability for plants and groundwater recharge (Scanlon *et al.* 2005, Zhang and Schilling 2006, Pan *et al.* 2011, Prabhakar and Tiwari 2015). With changing climate, the value of groundwater is expected to increase to sustain agriculture and domestic use needs (Taylor *et al.* 2012). A review of land cover relation to groundwater by (Owuor *et al.* 2016) revealed a need for site-specific groundwater monitoring networks due to the highly varied environmental and geological landscapes of Africa. The importance of the land use-groundwater relation is emphasized due to the increased demand for water to meet the needs of the rising population (Owuor *et al.* 2016).

In this study, analysis of multispectral imagery (RapidEye, Landsat and Sentinel-2 images spanning from 2013 to 2016) was conducted to map seasonal land cover variability. First, RapidEye images were selected due to their uniqueness of the red edge band capable of detecting drastic reflectance changes in vegetation (Rodriguez *et al.* 2006, Schönert *et al.* 2015). The open access, free of charge and consistent availability of European Space Agency's (ESA) high-resolution Sentinel-2 images favored their selection as the second set of imagery for the study. The sub-humid wetland lies in a tropical region experiencing the presence of clouds exacerbated during the rainy season witnessed from March to May. To address the problem of cloud coverage, the third set of images used were Landsat to densify the acquisitions acquired over the study area despite their lower spatial resolution.

The nonparametric statistical learning technique, Support Vector Machine (SVM) was applied in the land cover characterization (Mountrakis *et al.* 2011). The concept of SVM is the generation of hyperplanes which best differentiate the input features. Learning involves iterative selection of optimal boundaries minimizing the misclassification. SVMs ability to give good accuracies despite limited training samples prompted its use for the research (Mountrakis *et al.* 2011, Paneque-Gálvez *et al.* 2013).

In addition to mapping seasonal land cover dynamics, other vegetation dynamics explanatory variables such as depth to groundwater and meteorological conditions are assessed (Li *et al.* 2014). The specific objectives of this study are:

i) To assess seasonal land cover changes using multi-temporal multi-spectral satellite imagery with an aim of understanding how land cover reflects the impact of water balance components.

ii) To generate cumulative seasonal land cover maps from all the classified images.

iii) To evaluate the temporal patterns of precipitation and spatio-temporal patterns of land cover and their relationship to depth to groundwater along the hydrological zones of a floodplain.

The information obtained is important in understanding spatio-temporal patterns of water availability that drive plant growth.

3.2 Materials and datasets

The hydrological zones were subdivided based on extent and duration of flooding during the long rainy season (Daniel *et al.* 2017). The fringe zone is not flooded; the middle zone is partially flooded whereas the riparian zone is completely flooded during the long rains. The riparian fields are located closest to the Kilombero River, the middle fields located between Riparian & Fringe fields have elevation slightly higher than the riparian fields, and the fringe fields are located furthest from the river at a higher elevation compared to the other two test sites. Elevation increases with increasing distance from the river. Subsequently, the flooding depths are higher in the riparian zone.

3.2.1 Satellite Data

The multi-spectral data used in this study comprised of two high-resolution (RapidEye and Sentinel-2) and the medium resolution (Landsat) imagery (Figure 3.1). Classification was performed on each individual image to generate a series of land cover maps on the dates of image acquisition. RapidEye has five spectral bands with at nadir pixel resolution of 6.5 m by 6.5 m resampled to 5 m by 5 m. The data acquired from http://blackbridge.com (Planet 2016) was level 1B in which the radiometric and sensor corrections have been applied. RapidEye images have a resolution of 5m for all bands hence no resampling was applied. Fourteen level 1B RapidEye scenes from August 2013 to June 2015 were analyzed.

Nine Sentinel-2 images were downloaded from https://scihub.copernicus.eu/dhus/ (ESA Sentinels Scientific Data Hub 2016). The images have a spatial resolution of 10m with acquisitions from December 2015 to August

2016. Sentinel-2 has 13 bands which include aerosol detection bands, vegetation classification bands, water vapor and cloud discrimination bands (Drusch *et al.* 2012). In this research, bands 2 to 8 designed for vegetation classification were utilized. Sentinel-2 bands 2, 3, 4 and 8 have a 10m resolution while bands 5, 6, 7 and 8A have a 20 m resolution. Resampling was performed on the 20 m bands to obtain an image with uniform 10m resolution.

Seven Landsat images were downloaded from the USGS Earth Explorer platform (https://earthexplorer.usgs.gov/). For the period spanning 2013 to 2016, bands 2 to 8 were used in the analysis as the other images had a large cloud coverage. Landsat Operational Land Imager multispectral bands 1-7 have 30 m spatial resolution whereas panchromatic band 8 has a resolution of 15m. Resampling was performed on the 15m band for uniformity in the Landsat bands. Resampling to a higher and common pixel size for the three sets of imagery would introduce a smoothing effect and a subsequent loss of spatial accuracy and the ability to discriminate features. The spectral properties of the bands in the classification are presented in Table 3.1 (Drusch *et al.* 2012, Schuster *et al.* 2012, Barsi *et al.* 2014).

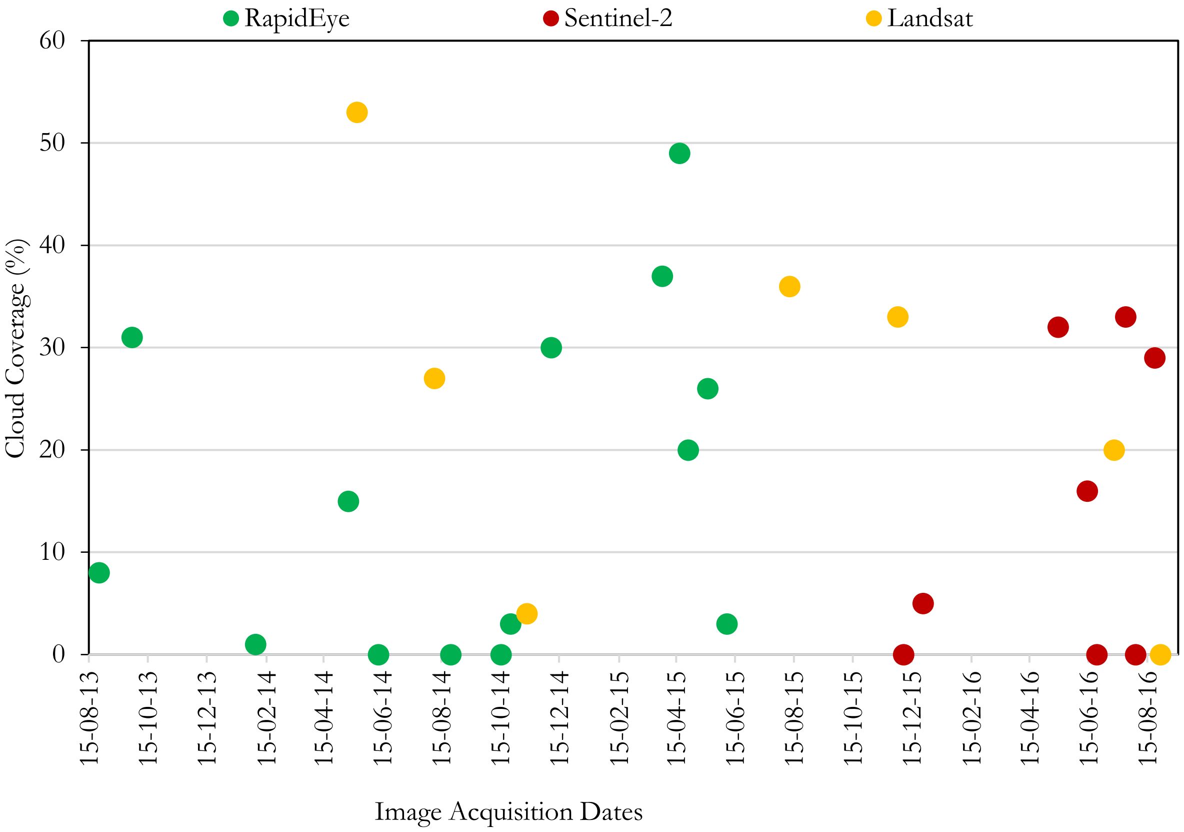


Figure 3.1: Temporal distribution of the RapidEye, Landsat and Sentinel-2 image acquisitions. Percentage coverage of cloud on the images is indicated.

The spectral band selection (visible, Near Infra-Red, Short Wave Infra-Red) are reported to be optimal bands in wetland mapping due to their spectral response on the electromagnetic spectrum. The strong reflectance in the Red Edge and Near Infra-Red related to plant biochemical properties are suitable for mapping vegetation (Adam *et al.* 2010).

Table 3.1: Attributes of the spectral bands of the sensors used in classification of the wetland

Sensor	Band	Name	Central Wavelength, micrometers	Spatial resolution (m)
Landsat	2	Blue – visible spectrum	0.482	30
	3	Green– visible spectrum	0.561	30
	4	Red– visible spectrum	0.655	30
	5	Near Infra-Red	0.865	30
	6	Short Wave Infra-Red 1	1.610	30
	7	Short Wave Infra-Red 2	2.200	30
	8	Panchromatic	0.590	15
Sentinel-2	2	Blue– visible spectrum	0.490	10
	3	Green– visible spectrum	0.560	10
	4	Red– visible spectrum	0.665	10
	5	Vegetation Red Edge	0.705	20
	6	Vegetation Red Edge	0.740	20
	7	Vegetation Red Edge	0.783	20
	8	Near Infra-Red	0.842	10
	8A	Vegetation Red Edge	0.865	20
RapidEye	1	Blue– visible spectrum	0.475	5
	2	Green– visible spectrum	0.555	5
	3	Red– visible spectrum	0.657	5
	4	Red Edge	0.710	5
	5	Near Infra-Red	0.805	5

3.2.2 Ground Data collection

3.2.2.1 Study design

During the data collection campaigns, the fields had rice plantations with varying plant heights. The plantation in the fringe had heights of 1 to 1.5 meters, the middle between 0.5 to 1m while the lowest heights were in the

Riparian ranging from 0.3 to 0.7 m. Plantations in the riparian fields were submerged by floods to a height of 0.5m. Scattered maize plantations were also present in the riparian fields.

3.2.2.2 Training and validation data

Knowledge of study area, false color composition (FCC) and NDVI images generated from the optical images formed the basis of generation of the training and validation points. Utilization of images collected in the past with no corresponding field data and non-accessibility of some areas within the wetland necessitated the inclusion of NDVI and FCC in the identification of land cover class. FCC used for clear identification of the classes under study was NIR: R: G hence for RapidEye 5:3:2, for Sentinel 8:4:3 and for Landsat 5:4:3. Land cover type auxiliary information from GPS data (March to May 2015 and December 2015 to February 2016) and UAV photos (May and September 2014 and February 2015) were collected to create a database for generating the training and validation data for the 2013-2016 thematic maps.

A LumixGX 1 camera was mounted on an Unmanned Aerial Vehicle (UAV) and the MaVinci Desktop software was used for flight planning. The flight periods were 29th May to 1st June 2014, 24th to 26th September 2014 and the last flight campaign was from 27th February to 2nd March 2015. Post processing of the geotagged UAV photos was performed with the Agisoft PhotoScan software.

Training and validation sample points identifying the land cover classes were determined with the Garmin Etrex 30 GPS with horizontal positional accuracy of <3m (Garmin International 2005). Points located in a clouded area on the images were removed.

Acquisition time for the satellite images, the GPS points and the UAV photos varied. Field data points were directly obtained from GPS points and others selected from UAV photos close to an image acquisition. Random sampling divided the points into training (70%) and validation (30%). However, in some cases, there was no field data (GPS points or UAV flights) corresponding or taken close to image acquisition date while other areas were inaccessible. Hence, NDVI was computed and random sampling performed on the each of the NDVI images to generate 80 to 150 reference points. The sampling design tool on ArcGIS 10.3.1 randomly subdivided these points into training (70%) and validation (30%). Therefore approximately 2/3 (50-100) and 1/3 (25-50) points were available for training and validation respectively. In such cases, inference from NDVI ranges, FCC and local knowledge of the study area were used in identifying the land cover classes.

3.2.2.3 Monitoring depth to groundwater

Despite the large size of the Kilombero Valley, the study site focused on a small area over which the piezometers were installed. Depth to groundwater was monitored hourly by pressure sensors installed in 10-cm-diameter piezometers at each hydrological position from March 2015 to June 2016 (Gabiri *et al.* 2018). This research included observations from nine piezometers (Figure 2.2).

Piezometer 1, 2 and 4 were in the riparian zone, piezometer 3 was located a few meters from the Kilombero River, piezometer 6 was in the middle zone and piezometer 10 and 14 in the fringe zone while piezometer 12 and 13 were located on higher regions above the study site.

In addition to monitoring depth to groundwater, a rain gauge station (Figure 2.2) installed at the Ifakara Health Institute (IHI) within the floodplain provided daily precipitation data.

3.3 Analysis Methods

3.3.1 Work Flow for Satellite Image Processing

The RapidEye images were orthorectified using the 1-arc second Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) to ensure they are located in their corresponding x, y locations. The orthorectification was performed in Erdas Imagine inbuilt Rational Polynomial Coefficients that relate image and ground coordinates taking into account the satellites orbital position. ATCOR (Atmospheric CORrection) model removes the influence of the atmosphere, solar illumination, sensor viewing geometry and terrain information from optical images in the extraction of physical earth surface parameters (Richter and Schläpfer 2013). This is important especially when the comparison of images acquired from different satellite sensors is required. Atmospheric correction implemented with ATCOR 2 in Erdas Imagine applies for multispectral imagery acquired over flat terrain. Bidirectional Reflectance Distribution (BRDF) defines the reflectance of a target as a function of sun and sensor viewing geometry. In small, flat areas with undulating topography, the BRDF effect is small as differences in illumination are negligible and therefore BRDF correction was not applied since our study site is 13*5km with an elevation difference of approximately 170 m (Vaudour *et al.* 2008).

Cloud coverage greatly affects optical imagery over the tropical region (Figure 3.1). Cloud removal process involved setting thresholds of brightness, secondly selecting of areas identified as either clouds or cloud shadows was performed by application of the thresholds, thirdly, generating masks from the cloud and cloud shadow layer and finally, clipping the RapidEye images to eliminate the cloud areas. Cloud masking of Top of Atmosphere Sentinel-2 and Landsat was performed using the F-mask (Zhu *et al.* 2015). Orthorectification and co-registration followed in the processing chain. Ground control points and google earth images were references for the coregistration implemented by autosync tool in Erdas Imagine assessed. The positional accuracies were less than 2m for the Sentinel-2 and less than 5 m for Landsat. Clipping to the size of the study area extent was the final step in preprocessing.

3.3.2 Classification

SVM is a machine-learning algorithm that separates classes by generating hyperplanes for optimal class separation. The method was selected as it has resulted in high accuracies in monitoring wetland dynamic (Chatziantoniou *et al.* 2017, Pande-Chhetri *et al.* 2017). The radial basis function, that fits data in a higher dimensional space to increase class separability was implemented in the EnMap toolbox (van der Linden *et al.* 2015). All optical images were classified using the training data after which performance metrics were assessed by computing the overall and Cohen's-Kappa coefficient (Congalton and Green 2009). Separation of training and validation points offered an independent unbiased accuracy assessment. Figure 3.2 shows the processing chain of the optical images.

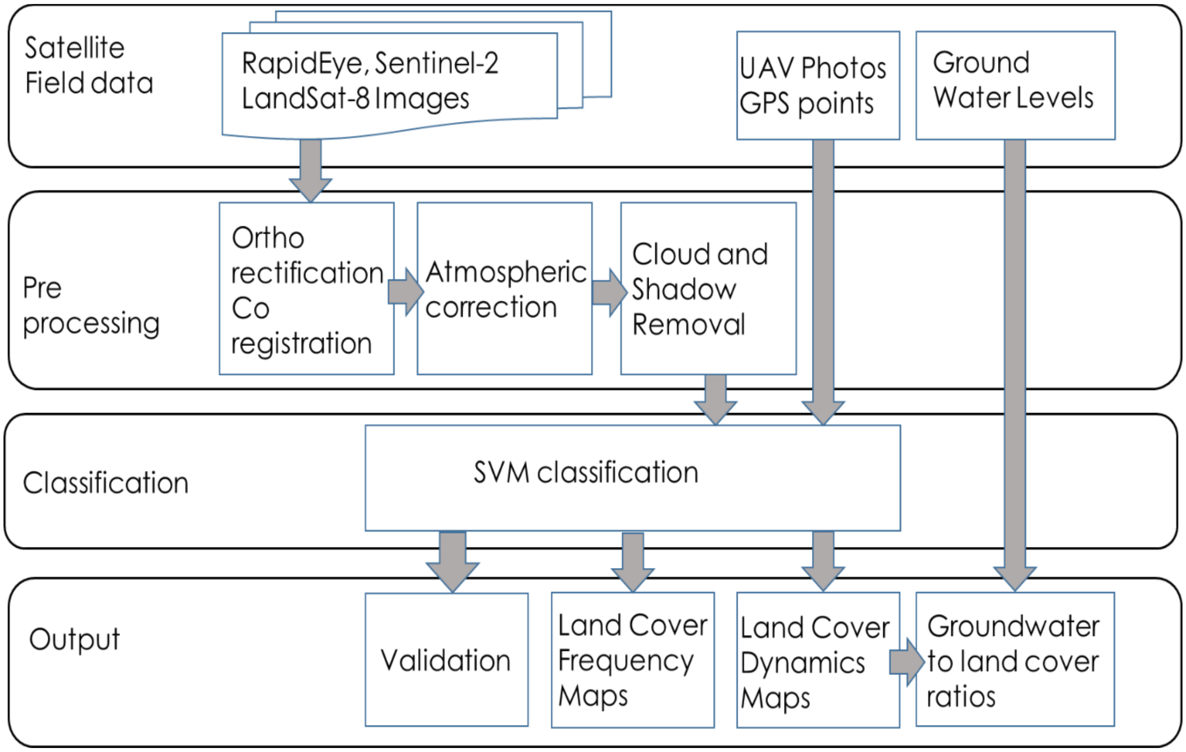


Figure 3.2: Data process flow. Auxiliary training and validation points for the Support Vector Machine (SVM) classification were obtained using the Unmanned Aerial Vehicle Photos (UAV) and the Global Positioning System (GPS) points

The analysis was as follows:

- Support Vector Machine classification algorithm was run on the series of multi-spectral images: RapidEye, Sentinel-2, and Landsat. Built up areas, bare land, vegetation cover and water extent coverage were the classes of interest. The tropical region is vastly affected by clouds thus a separate class of all masked out areas was included. The total area of classification was 5000 hectares. The sensing period for the RapidEye ranged from August 2013 to June 2015 and for Sentinel-2 from December 2015 to August 2016. To increase the temporal resolution of the time series, cloud-free and minimum-clouded Landsat images acquired within the three-year study period were included in land cover mapping. The RapidEye, Sentinel-2 and Landsat 5m, 10m and 30m resolution were retained for the classification as resampling to a higher pixel size will lower the classification accuracy (Fisher *et al.* 2017).

- The validation dataset was used to calculate the overall and Kappa coefficient as accuracy assessment tests for the classification. The overall accuracy is the percentage of correctly classified pixels divided by the total number of test pixels. The Kappa coefficient estimates the error reduction generated by the classier verses the error of a random classifier (Congalton and Green 2009). It is given by

  $$\widehat{K} = \frac{p_c - p_e}{1 - p_e}$$

  In which $\widehat{K}$ is the Cohen's-Kappa coefficient, p_c is correct classification giving the observed level of agreement, and p_e is expected agreement by chance.
- Cumulative seasonal land cover maps were generated from all the classified images to identify the spatial distribution of the most frequent classes over the study area. The spatial resolution of the classified RapidEye images was 5 m, Sentinel-2 was 10 m and Landsat was 30m. All the maps were resampled to a common resolution of 5 m. Resampling to the smallest resolution ensured that the classes assigned to the pixels are retained since the bigger pixels in Landsat and Sentinel-2 are subdivided into smaller sizes. The resampled classified maps were separated into individual classes i.e. bare, water, vegetated and built up. A stack was generated for each land cover class. The number of occurrence of each class through the stacked image was then calculated. This gave the frequency of recurrence of the given class presented as the cumulative seasonal maps, whose resolution is 5 m. The study selected images taken between January to June (time epoch 1 in the rainy season) and July and December (time epoch 2 in the dry season) in generation of the frequency maps.
- Groundwater levels were qualitatively related to land cover as an additional validation method explaining the occurrence of land covers and groundwater levels in the dry and wet season. Ratios generated were constrained to one such that $bare_{ratio} + vegetated_{ratio} + water_{ratio} + builtup_{ratio} = 1$

3.3.3 Geostatistical Analysis of depth to groundwater

A limited number of piezometers observations provided groundwater levels information. To obtain values for all areas in the study site, the piezometer readings were used in generation of groundwater surfaces. For interpolating the groundwater level measurements, several methods are available from which geostatistical methods are the most sophisticated approaches as they analyze the spatial structure of the data. This involves estimation and modeling of spatial correlation (covariance or semi-variance) of measured point data (Hengl *et al.* 2007, Hengl 2009).

Semi-variograms are plots of semi-variances against separation distances of point observations. Points are averaged for a standard distance called lag (experimental/sample variogram). Fitting of the variogram was done using linear, Exponential and Gaussian models. Weights assigned to semivariograms are iterative least square estimation based on a number of point pairs and distance (indicative of true spatial autocorrelation structure). A number of pairs at a certain lag divided by the square of the distance mainly gives the weight. Thus, a higher weight assigned to many pairs and at short distances. The semivariance is defined by the average squared difference of data points separated by lag distance h. Semivariance increases with increasing distance. When the distance between samples is less than the range, there is a spatial relationship between the samples. A separation distance greater than the range implies random variation (Trangmar *et al.* 1986). The semivariograms $\gamma(h)$ is calculated by

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i}^{n(h)} [\left(z(s_i) - z(s_i + h)\right)^2]$$

$$n(h)\ is\ number\ of\ pairs\ of\ observations\ over\ a\ separation\ distance\ h,$$

$$z(s_i)\ is\ the\ soil\ moisture\ observation\ at\ s_i,$$

$$and\ z(s_i + h) is\ the\ measured\ soil\ moisture\ value\ at\ s_i + h$$

With n point observations, $n * (n-1)/2$ semivariance pairs can be computed. Plotting the semivariances and their separation distance produces Figure 3.3 (Hengl 2009).

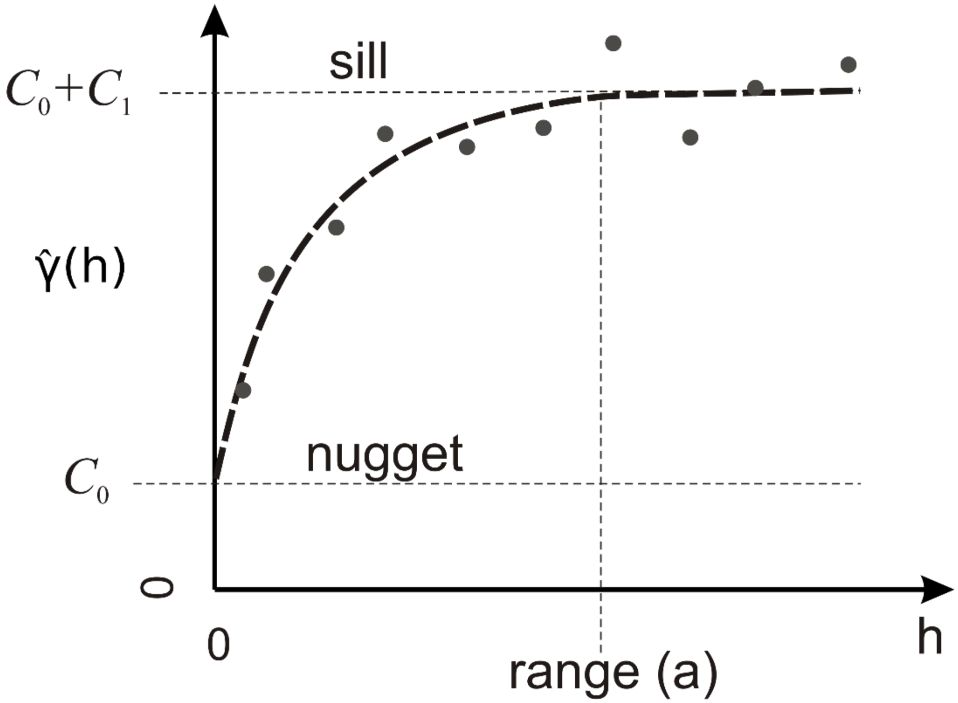


Figure 3.3: Components of a semivariogram

The Exponential and Gaussian variograms were fitted.

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C0 + C1(1 - e^{-h/a}) & h > 0 \end{cases}\ Exponential$$

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C0 + C1(1 - e^{-h^2/a^2}) & h > 0 \end{cases}\ Gaussian$$

$$C0 \geq 0 \; is\ the\ nugget\ variance\ and\ C1\ is\ the\ stuctural\ variance,$$
$$C0 + C1 = sill, a\ is\ the\ range$$

Nugget variance is the small-scale variability representative of variation due to random factors such as measurement errors not detectable at the sampling scale used. The sill gives an indication of total variance of the system, comprising of randoma and structural variance. Structural variance can be caused by intrinsic factors such as climate, parent material, topography, soil type and occurs at the point where the variogram flattens out (Rodríguez Martín *et al.* 2016). The range is the maximum distance at which spatial correlations exist. Beyond the range, correlation is minimal. The number of lags and lag tolerance were adjusted to produce the best fit for the data. The Exponential and Gaussian models were implemented in variogram modeling The degree of spatial dependence indicates the level of correlation between the measured point data different spatial locations.

$$degree\ of\ spatial\ dependence = \frac{C0}{C0 + C1}$$

A spatial dependence lower than 25% indicates a strong spatial correlation between the groundwater levels measured at different locations, 25% to 75% describe a moderate correlation, and a value greater than 75% indicates a weak spatial dependence (Jing *et al.* 2014). Monthly groundwater raster maps were generated for the period under study.

3.4 Results and Discussion

3.4.1 Time series of land cover characterization using multi-source imagery

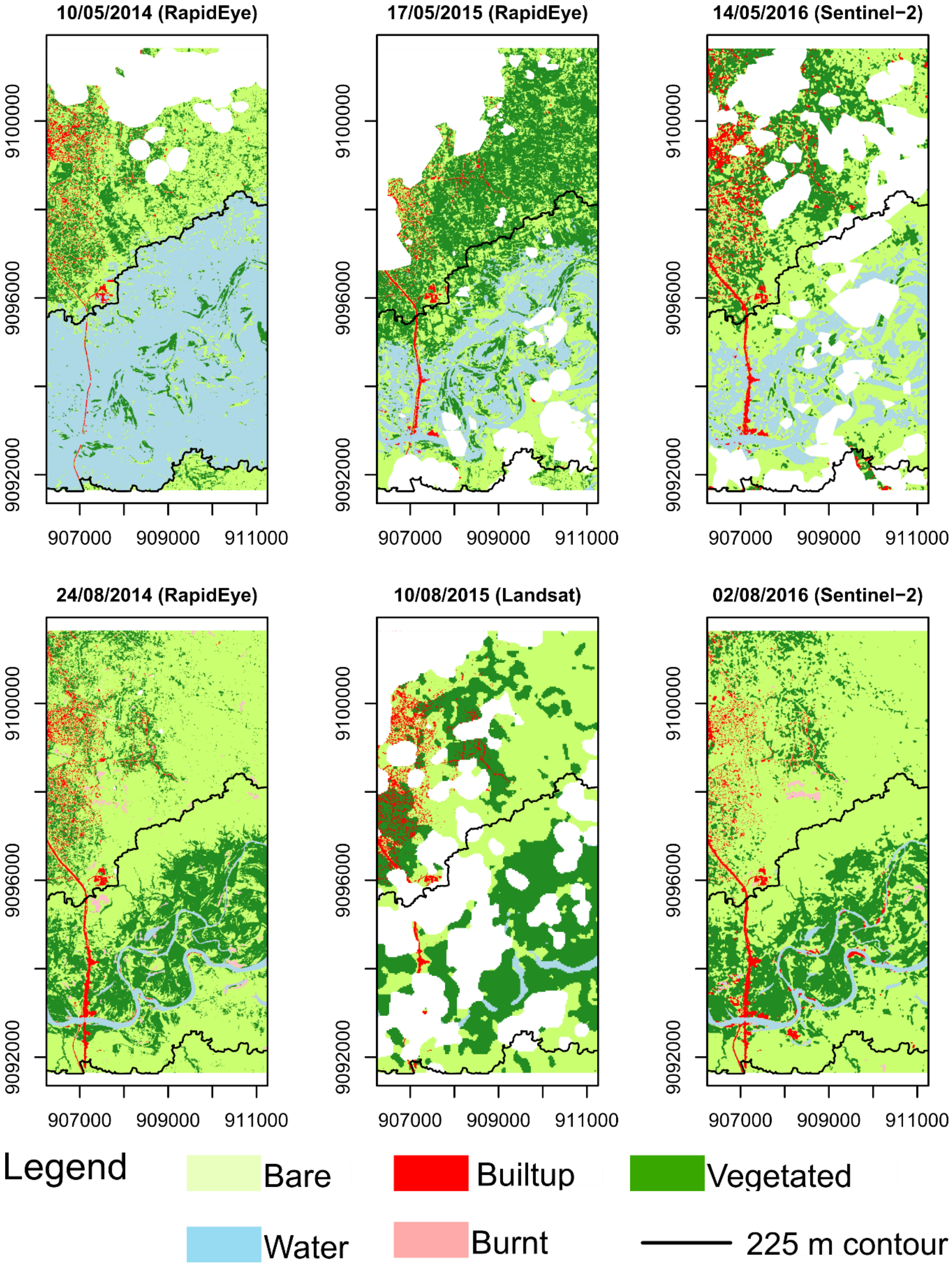
The Support Vector Machine classification algorithm was applied to multi-temporal multi-sensor RapidEye, Sentinel-2, and Landsat satellite imagery. Built up areas, bare land, vegetation cover and water extent coverage were the classes of interest (Figure 3.4, 3.5). The tropical region is vastly affected by clouds thus a separate class of all masked out areas was included. The total area of classification was 5000 hectares. The overall classification accuracies for the RapidEye, Sentinel-2, and Landsat ranged from 70 to 96% while the Cohen´s-Kappa coefficients ranged from 0.61 to 0.95 (Table 3.2).

Table 3.2: Accuracy assessment of the Support Vector Machine classification

Image Date	Sensor	Overall Accuracy	Kappa Coefficients	Image Date	Sensor	Overall Accuracy	Kappa Coefficients
25-08-13	RapidEye	93.8	0.92	17-05-15	RapidEye	81.6	0.77
28-09-13	RapidEye	94.1	0.93	06-06-15	RapidEye	87.9	0.85
03-02-14	RapidEye	87.8	0.85	10-08-15	Landsat	91.0	0.88
10-05-14	RapidEye	93.6	0.92	30-11-15	Landsat	82.3	0.76
19-05-14	Landsat	70.5	0.61	06-12-15	Sentinel-2	89.0	0.85
10-06-14	RapidEye	91.5	0.89	26-12-15	Sentinel-2	85.6	0.82
07-08-14	Landsat	90.0	0.87	14-05-16	Sentinel-2	91.7	0.89
24-08-14	RapidEye	96.7	0.95	24-05-16	Sentinel-2	88.4	0.86
15-10-14	RapidEye	85.4	0.82	13-06-16	Sentinel-2	87.3	0.84
25-10-14	RapidEye	82.1	0.78	23-06-16	Sentinel-2	95.0	0.93
11-11-14	Landsat	86.1	0.81	11-07-16	Landsat	90.8	0.88
06-12-14	RapidEye	95.6	0.94	23-07-16	Sentinel-2	78.9	0.74
31-03-15	RapidEye	84.9	0.81	02-08-16	Sentinel-2	84.3	0.8
18-04-15	RapidEye	92.7	0.91	22-08-16	Sentinel-2	79.5	0.75
27-04-15	RapidEye	91.2	0.89	28-08-16	Landsat	83.3	0.78

Despite the varying spatial resolution for Landsat (30m), Sentinel-2 (10m) and RapidEye (5m), the overall accuracies were all above 70%. The potential of each of the satellite sensors in delineating bare, vegetated, water and built up areas portrays its use in inventorying and monitoring land cover over the study area. The maps developed show temporal changes in land cover patterns providing seasonal change information for land use planning. Different climate-smart technologies can be adopted depending on the spatial seasonal patterns from the land cover maps and time of year.

The current study was limited to general land cover classes and did not assess crop type mapping. The major challenge was the small sizes of the farms having varying crop types. Object-based classification offers a solution to crop type and field scale mapping. However, it would require very high-resolution imagery capable of delineating the actual field boundaries.



10/05/2014 (RapidEye)

17/05/2015 (RapidEye)

14/05/2016 (Sentinel−2)

24/08/2014 (RapidEye)

10/08/2015 (Landsat)

02/08/2016 (Sentinel−2)

Figure 3.4: Classified images of the rainy (May) and dry season (August). The masked areas (white) are due to cloud coverage; the 225m contour line shows the delineated riparian zone. RapidEye has a resolution of 5 m, Sentinel-2 10m and Landsat 30 m.

The use of different sensors with slightly varying number of spectral bands and range is a possible cause of error in classification. To minimize further differences and likely bias in the data acquired from the three satellite sensors, the same classification algorithm was performed on each individual image. The interpretation of each image is independent of the subsequent images. Varying spatial resolution introduces uncertainties particularly for the Landsat 30m resolution images. High-resolution satellite imagery is preferable for classification of small sites with small and heterogeneous field sizes and small features present like in the current study (Fisher *et al.* 2017). Features with sizes smaller than the low Landsat resolution pixels were not distinguishable.

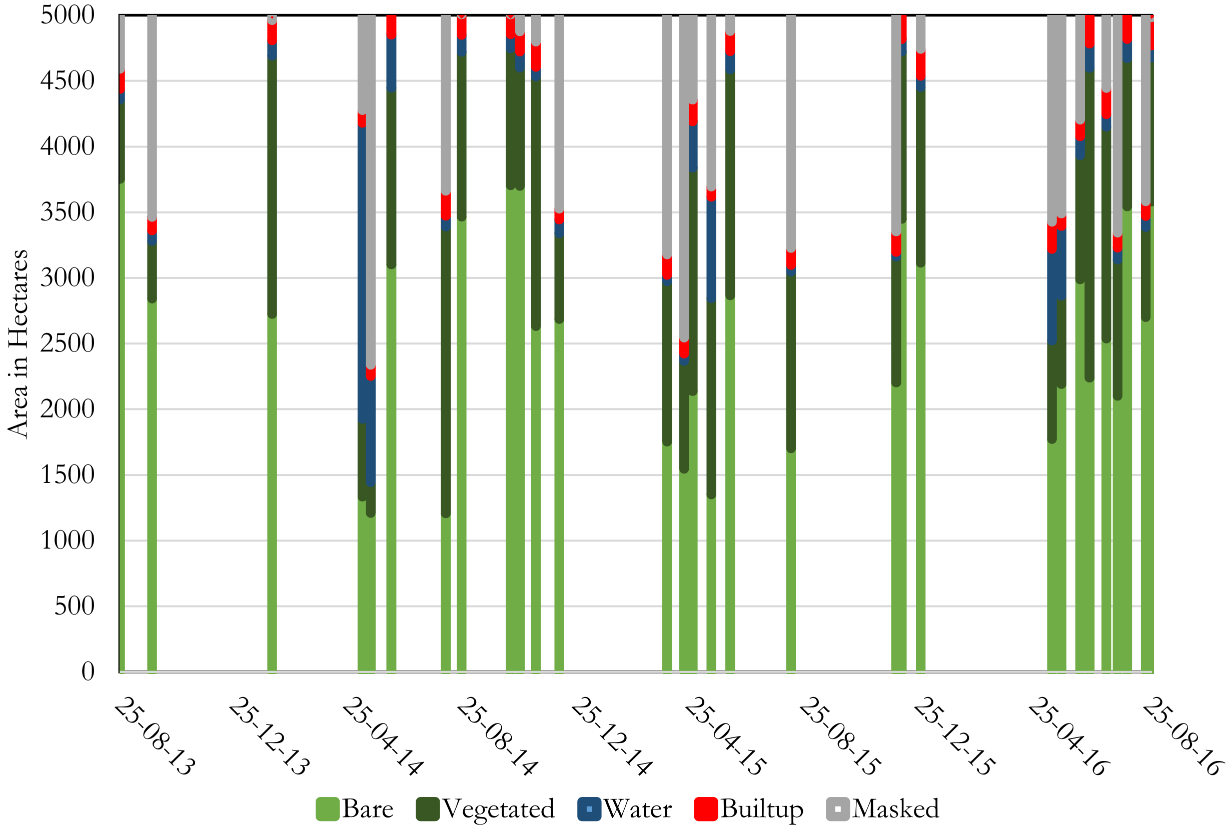


Figure 3.5: Areal coverage of land classes in hectares. The spacing between acquisitions indicates the inconsistent optical image availability.

The bare cover class is greater than 70% in the months of August to October, which coincides with the dry season. It is a common practice for the farmers to clear their fields by burning in preparation for the subsequent short rainy season.

From December to February, the bare land coverage decreases as the rainy season begins. Flooding along the Kilombero River is a periodical event though the extent varies from year to year (Figure 3.6). In June, bare land coverage increases due to sediment accumulation following receding waters after the floods. The location of vegetated land cover shifts depending on the season. In the rainy season, areas outside of the riparian zone are vegetated while vegetation coverage within the riparian zone increases in the dry season. The built up class remains constant with the slight variations due to changes in masked out areas. The road heading to the Kilombero River is slightly higher than the surrounding area. During the rainy season, flooding partly submerges this road. As a result, the width of the road varies between the rainy and the dry season as seen on the optical images (Figure 3.6). During the dry season, the road width remains constant and is clearly visible in the optical images. Cloud coverage ultimately prevents that the areas masked out due to clouds and cloud shadows fall in any other class and thus land cover change analysis was not possible (Figure 3.1).

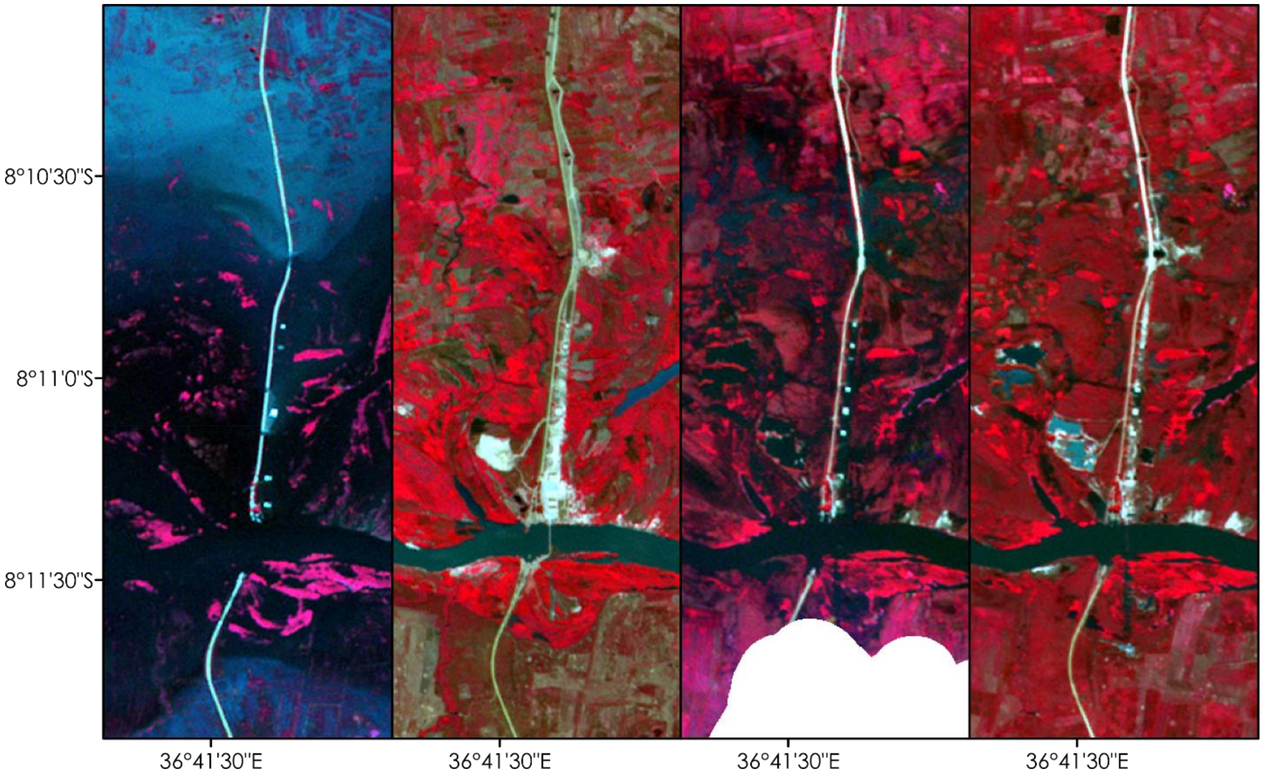


Figure 3.6: RapidEye images for 10-05-14 (rainy season), 24-08-14 (dry season), 17-05-15 (rainy season) and 06-06-15 (dry season). The right side of the road towards the Kilombero River submerges during the rainy season.

The farmers in the study area mainly practice agriculture dependent on rain. The major land cover in the dry season between June and October is bare. High temperatures and low soil water content attribute to the bare land cover. At the beginning of the rainy season in December, the vegetation cover increases and reaches a peak in May-June (Figure 3.4). Areas in the north of the study area, not affected by floods are vegetated at this time. Hence, images acquired in May to June indicate that there is vegetation in the northern sections. Images acquired in August indicate that the northern parts are bare explained by little or no vegetation since this marks the dry season.

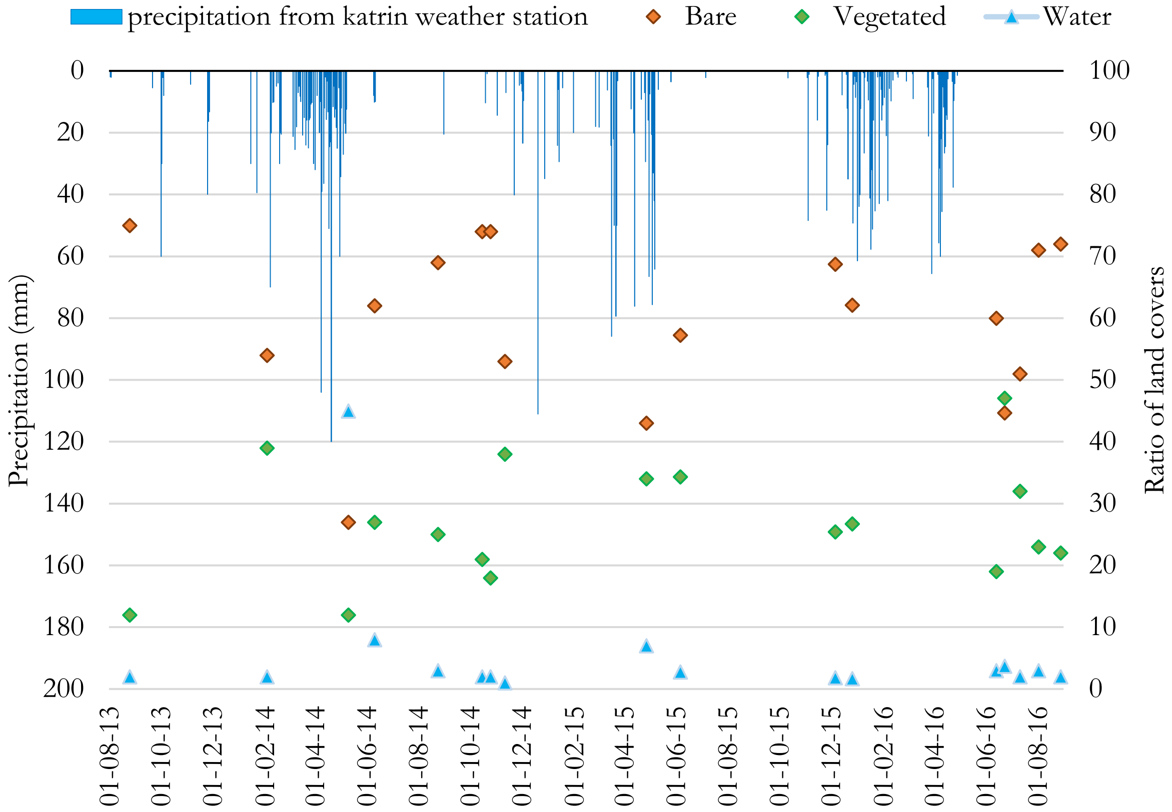


Figure 3.7: Relationship of precipitation and ratio of land cover.

The maximum flooding extent during the rainy season reaches up to 3 km on both sides of the Kilombero River as was experienced in May 2014. However, the El Niño phenomenon caused the exceptional flooding (Figure 3.4). The flooded zone corresponds to the 225m contour. The high vegetation coverage in the middle and fringe areas in May and June indicates the presence of suitable conditions to support vegetation growth in areas out of the riparian zone (Figure 3.4). The riparian zone has a bare land cover in June due to the accumulation of sediments following the rains and subsequent floods.

The pattern of land cover change over the three years is consistent and is determined by the seasons. It is evident that with increased precipitation, there is a reduction of bare land cover and increase in vegetated cover. Consequently, low precipitation results in increased bare cover class (Figure 3.4). Knowledge derived from the seasonal land cover thematic maps are a basis for operational monitoring of land conditions which aid in establishing land management practices to increase agricultural production. Vegetation covers the middle and fringe zones for most of the rainy season. In the dry season, the riparian zone, having close proximity to the river is vegetated implying the potential existing in this area for crop production. Conversely, the areas away from the river are bare for most of the dry season. This implies that when it comes to decisions making on land use for increased food production during the dry season, farms in the middle and fringe zones would require additional water supply to produce crops in the dry season.

From the time series of the RapidEye, Landsat, and Sentinel-2 images, frequencies of land cover classes over the entire period were determined (Figure 3.8). The bare land cover through the investigated period dominates the areas to the northeast, central and south of the study area. The region within the riparian zone is frequently vegetated postulating sufficient soil moisture for continued agriculture.

Information from cumulative seasonal maps aid in detailing locations where agricultural improvement mechanisms such as setting up of water pumps and irrigation planning can be spatially located. Areas with a low bare cover recurrence rate, can adopt less cost-effective methods (such as manual pumps) whereas areas such as to the northeast where bare land cover is dominant require permanent methods (such as drip irrigation systems) that ensure constant and continuous availability of water to support vegetation growth (Figure 3.8).

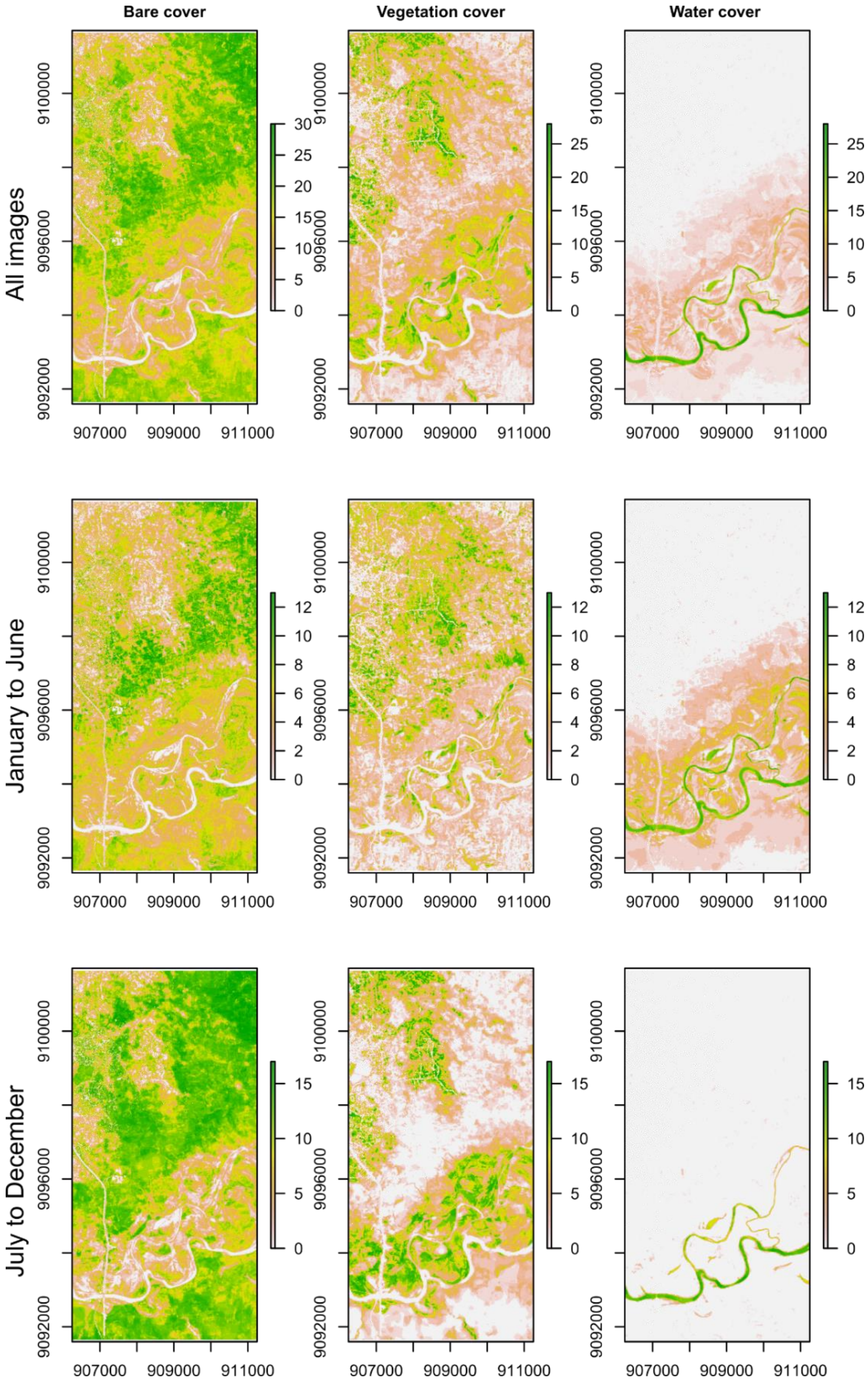


Figure 3.8: Recurrence of bare cover (left column), vegetation (central column) and water (right) column for all images (upper row), rainy season (central row) and dry season (bottom row). The legend indicates the number of images that generated the cumulative seasonal maps.

The spatial-temporal changes occurring between January to June and July to December are assessed by computing the cumulative season duration maps of each class from the classification maps produced (Figure 3.8). Observations indicate an increase in bare cover to the south and northeast in the second half of the year. In the first half of the year, the central and northern regions are vegetated with no vegetation cover observed in these central and northeast regions during the second half of the year. The static structures visible in both the wet and dry season vegetation images located towards the lower north are trees. The extent of water coverage, with maximum extent observed in May 2014 when there was excessive flooding. In the second half of the year, the water extent is shown to reduce, consistent with the dry season experienced in that time of the year.

Analyzing land cover changes in the riparian zone aided in the understanding of its unique seasonal dynamics due to its proximity to the river. This area covers 2500 ha (half of the total study site). Vegetation increased during the dry season contrary to the areas in the middle and fringe zones. During the rainy season, the riparian zone has less vegetation as compared to the middle and fringe zones because it is largely flooded.

Vegetation within the riparian zone was greater than 50% of the total vegetation coverage of the study area from August to October (Figure 3.9). High vegetation coverage in the riparian zone in the dry season indicates the presence of a high soil water content capable of supporting the growth of crops. The bare land cover was greater than 50% of the total coverage in April to June representative of exposure of sediment accumulation following receding floods in the riparian zone. The results are consistent with a report by (Wilson and Lewis 2015) who indicated that rainfed lowland rice is prone to submergence under floods. The receding floods in the riparian zone expose bare land whereby any previously existing vegetation is buried under a layer of silt washed down from the surrounding farms.

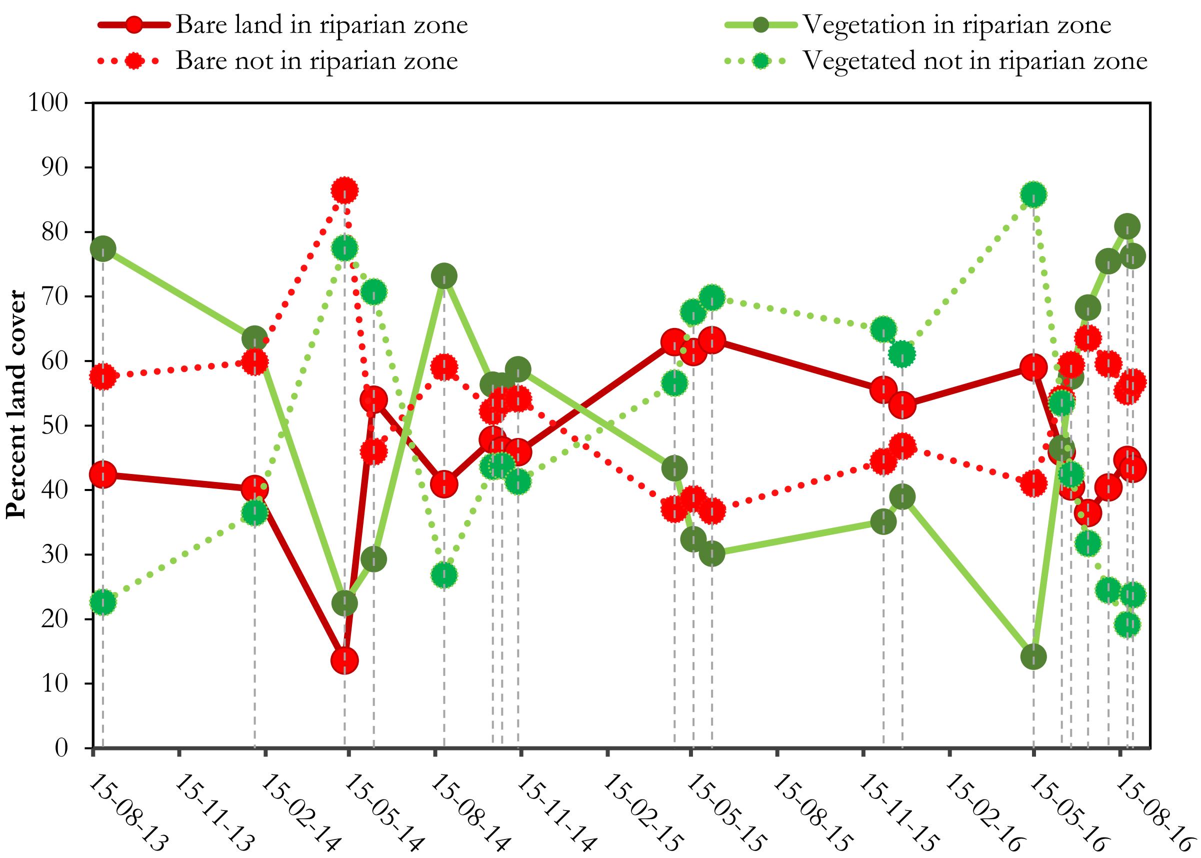


Figure 3.9: Ratio of vegetation and bare land cover within the riparian zone as a function of total vegetation and total bare land cover of the study area in percentage. The percentage not in the riparian zone is also shown.

Varying yearly flooding patterns in the riparian zone pose a challenge in the determination of planting times. From the results (Figure 3.4), a delay in planting time is recommended until the floods recede. In the dry season, the presence of vegetation within the riparian zone infers the agricultural potential of this area. However, in the selection of a suitable crop to plant in the riparian zone during the dry season, factors such as the high temperatures experienced during the dry season should be taken into consideration (Nicol *et al.* 2015).

The seasonal land cover dynamics indicates the farmers currently adopt a balance between risk and chance in crop growth. Given the current reports on changing weather patterns, a challenge exists of knowledge on expected flooding patterns. Farmers within the study area take a chance by prepare their fields during the rainy season to maximize the use of the available water for rice plantations. With this opportunity comes the risk of crop loss in case excess flooding destroys the planted rice. Based on the current land cover seasonal maps, forecasting to aid in planning the planting times and spatial locations is recommended.

3.4.2 Relationship between land cover and depth to groundwater

The hourly depth to groundwater observations were averaged to daily readings. Table 3.3 highlights the highest and lowest piezometer readings for the period under study. The riparian zone had the highest recorded value of depth to groundwater (3.3 m). The lowest depth to groundwater level at the middle zone was 2.83 m where that at the fringe was 3.38 m. Negative values stipulate flooding represented by the highest groundwater levels. The highest value was in the riparian zone at -3.98 m. The middle and fringe zones had -0.89m and -0.08 m as the highest values respectively.

Table 3.3: Piezometer readings for the period of observations running from March 2015 to June 2016

Name	Highest measured water level	Lowest measured water level	Low depth to groundwater occurrence periods	High depth to groundwater occurrence periods
PZ01	-1.01	3.33	Greater than -0.5 30/04/-21/05/15 23/04-14/05/16	Below 3.0 07/10/15-18/01/16
PZ02	-2.79	1.88	Greater than -0.5 15/04/-18/05/15 13/01-09/03/16 09/04-03/06/16	Below 1.5 18/10/15-18/01/16
PZ03	-3.98	1.80	Greater than -2.0 24/04/-27/05/15 11/02-28/02/16 14/04-30/05/16	Below 1.5 26/08/15-01/01/16
PZ04	-1.94	2.50	Greater than 0.1 22/01-01/03/16 10/04-01/06/16	Below 1.5 28/08/15-05/01/16
PZ06	-0.89	2.83	Greater than 0.1 15/04/-09/06/15 03/04-12/06/16	Below 2.0 23/09/15-22/01/16
PZ10	-0.08	1.30	Greater than 0.1 21/03/-24/05/15 20/01-26/02/16 31/03-12/05/16	Below 1.0 27/09/15-16/12/15

PZ12	3.09	3.38	Consistently between 3.15 and 3.25 in June to December 2015	Increased level between 3.1 and 3.15 from January to June 2016
PZ13	0.20	2.35	Greater than 0.5 29/04/-14/05/15 07/02-12/02/16 06/04-05/05/16	Below 2.0 27/10/15-07/01/16
PZ14	-0.04	2.05	Greater than 0.2 23/03/-21/04/15 27/01-23/02/16 31/03-09/05/16	Below 1.6 05/09/15-05/01/16

Groundwater depths for the middle and fringe show a similar pattern whereas that of the fringe shows that the water is close to the surface throughout the year (Figure 3.10). The fringe receives groundwater from the mountains as base flow while the groundwater level in the riparian and the middle zone depend on the river water level. The depth to groundwater is higher for the riparian zone compared to the middle in the rain season. Rainfall pattern is reflected directly in the fringe piezometers whereas the response in the middle and riparian fields is very little and is delayed (Gabiri *et al.* 2018). A steep decrease in depth to ground water is observed in March and January, usually taking between 1 to 2 months to reach a peak. A slight increase in depth to groundwater is observed in February coinciding with the reduced precipitation. An increase in precipitation in March subsequently results in a decrease in depth to ground water. In the dry season, the increase in depth to ground water has a gentle slope taking approximately 7 months to reach its lowest level. The precipitation and depth to ground water pattern indicate that the groundwater responds to precipitation patterns. This implies that some portion of the rainwater infiltrates into the ground in rainy seasons resulting to a fast rise in the water level. The rate of rise of the water level (recharge) is faster than the rate at which the depth to groundwater increases (groundwater drains).

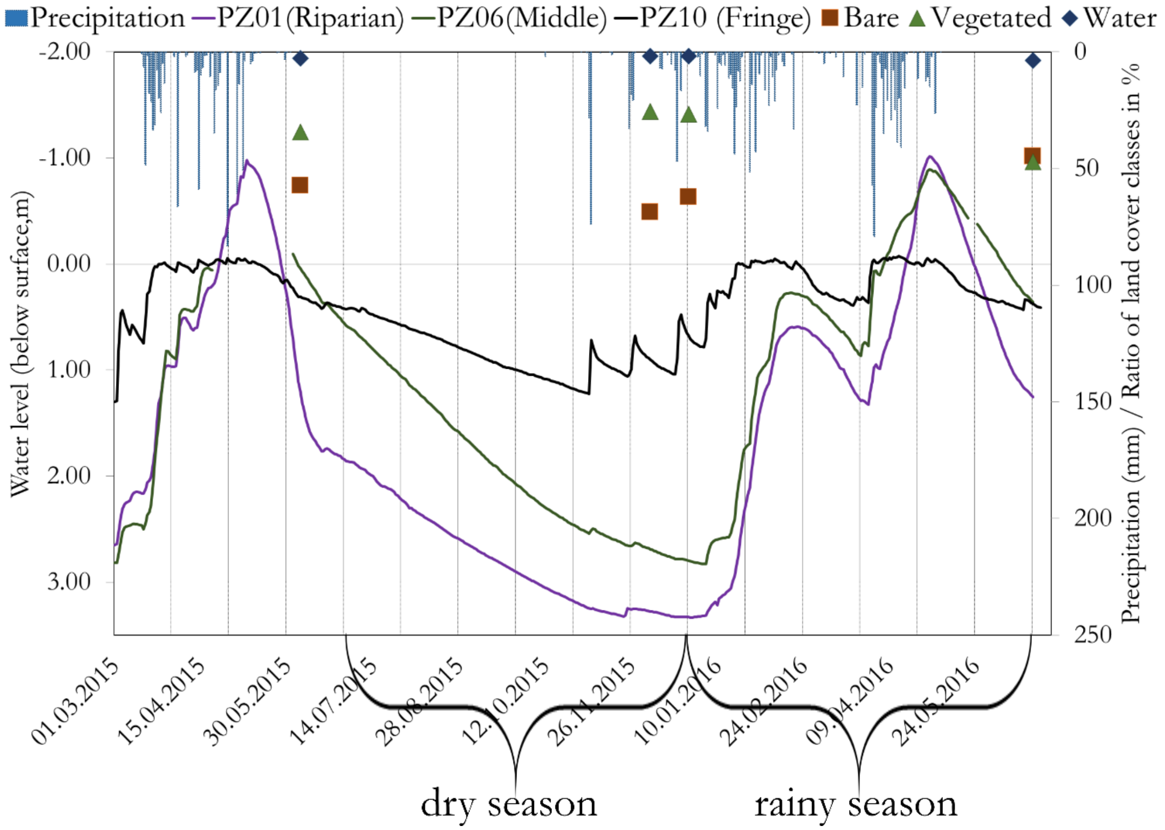


Figure 3.10: Groundwater depths in the riparian (PZ 01), middle (PZ 06) and fringe (PZ 10) zones and the corresponding ratios of land cover. The effects of changes in precipitation on depth to groundwater levels is shown.

Figure 3.10 presents the ratios of the land cover classes of images with less than 10% cloud cover together with the depth to groundwater. The ratio of areas for the water class remained constant for the images analyzed. Towards the end of May, the depth to groundwater level increased and bare land cover within the riparian zone increased as flooding recedes. The vegetated regions fall in the middle and fringe fields. The groundwater depth at the beginning of December was low ranging from 3.3, 2.7 and 0.9 for the riparian, middle and fringe zones. The high areal coverage of bare land is a result of the dry season experienced in the preceding months. By the end of December, the depth to groundwater was 3.3, 2.8 and 0.7m. Despite the lower groundwater towards the end of December, vegetation cover increased, attributed to the occurrence of precipitation. In June 2016, depth to groundwater was 1.26, 0.36 and 0.38 for the riparian, middle and fringe zones. Comparing the vegetated and bare cover of 2015 and 2016, it is evident that the vegetated coverage in 2016 is higher due to earlier recession of flooding water in 2016. In 2015, the decrease in depth to groundwater level began on 27th of May for the Riparian and 15th May for the fringe. However, in 2016, the decline began on 23rd of May in the Riparian zone and 6th of May in the fringe zone. Saturated soils had drained excess water by 23rd of June 2016 and hence the soil water content was suitable to support vegetation growth. The results show that the land covers respond to depths to groundwater.

Climate influences groundwater via precipitation infiltration and evapotranspiration rates depending on temperature (Li *et al.* 2014). Comparison of the precipitation data from the IHI station and climate data in Ifakara (climate-data.org) indicate that average temperatures are similar (Figure 3.12). Precipitation recorded at IHI is however slightly lower by approximately 1.2mm per day. Increased precipitation reduces the depth to groundwater, occurring between March and April 2015 and October 2014 to April 2016 (Figure 3.10). High temperatures as a proxy of high evapotranspiration rates affect shallow groundwater as reported by (Ndou *et al.* 2017). Consequently, the high temperature period results in a decrease in groundwater level. Lower temperatures, occurring between October and December 2015 in the rainy season, result in a rise in groundwater level. Another episode of low temperature between February and April 2016 subsequently resulted to a rise in groundwater level. Both temperature and precipitation influence the groundwater levels in the study area.

Point groundwater information are not sufficient to explain the spatial pattern of the land cover image and hence, application of Kriging interpolation method on the observed groundwater levels generated groundwater maps containing information at sampled and non-sampled locations. Variogram parameters that generated the kriged images are shown in Table 3.4. The depth to groundwater patterns from March 2015 to June 2016 are as shown in Figure 3.11.

Table 3.4: Variogram model parameters for generating kriged images

Dates	Model	Nugget variance	Sill	Degree of Spatial Dependance (%)	Range
15-03-15	exponential	0.2	0.9	17.0	5199.1
15-04-15	exponential	0.1	0.9	6.3	6498.7
15-07-15	exponential	1.0	0.1	88.2	10469.0
15-08-15	exponential	1.0	0.1	88.1	10469.0
15-09-15	exponential	0.9	0.2	84.1	5199.1
15-10-15	exponential	1.0	0.2	79.6	5199.1
15-11-15	exponential	0.7	0.5	60.0	5199.1
15-12-15	exponential	0.9	0.4	67.8	5199.1
15-01-16	exponential	0.3	1.1	19.4	5199.1
15-02-16	exponential	0.3	0.7	32.1	4957.1
15-03-16	exponential	0.3	0.7	27.3	5199.1
15-04-16	exponential	0.1	1.6	6.8	10412.4
15-05-16	exponential	0.0	1.3	0.0	9193.2
15-06-16	exponential	0.3	1.5	15.3	8803.3

The degree of spatial dependence for the groundwater depth raster images generated indicated that in the second half of the year (July-December) had data has a low spatial correlation, i.e. spatial dependence (Table 3.4). There was a high correlation for the groundwater depths observed in the first half of the year. From this analysis, it is evident that during the dry season when the groundwater level is low, there is a low spatial correlation of the groundwater level measured at the piezometers distributed in the study area. During the first half of the year, the groundwater depth rises and a strong spatial dependence is evidenced. Figure 3.10 shows the pattern of groundwater depth over the study period.

Depth to groundwater reduces in March and a further reduction in depth is observed in May (Figure 3.10). From July to December, the depth to groundwater increases, within this time a vast amount of land has bare cover and there is a rare occurrence of precipitation (Figure 3.10). At the onset of the rains in December, the depth to groundwater is observed to increase from January to June, with the peak observed in the riparian region in April and May. The negative depth to groundwater portrays flooding (Figure 3.10). The flooding is observed to commence on the southwestern side of the study area. There is a reduction in bare class and an increase in vegetated land cover from January to June.

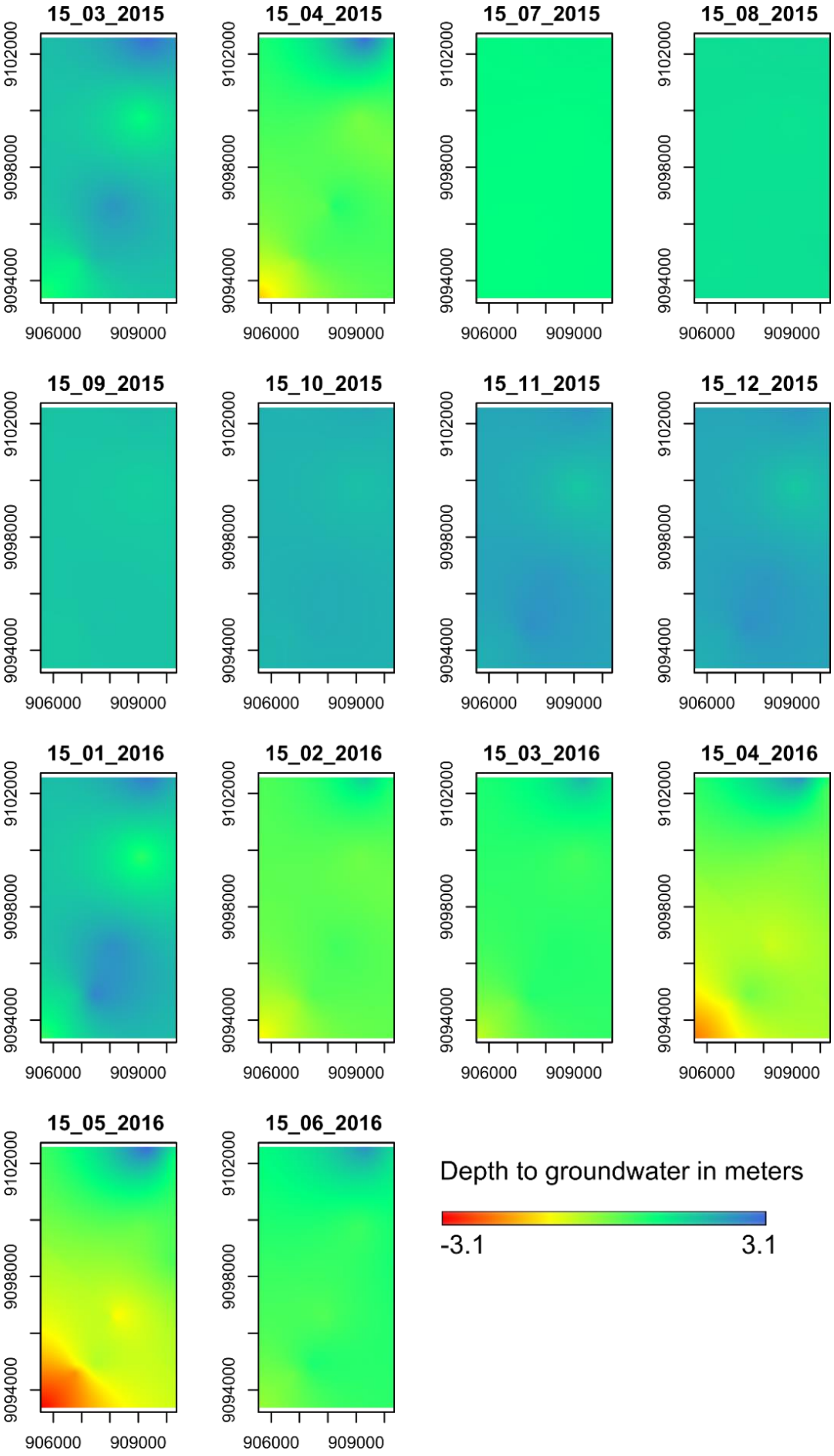


Figure 3.11: Groundwater levels interpolated using the Kriging method and changes in groundwater level between 2015 and 2016. Negative values indicate flooding.

The areas, particularly to the south of the study area, have a higher rise in water level as opposed to the northern parts of the study area. The Kilombero River drives the groundwater in the southern part whereas lateral flow from the Udzungwa Mountains drives the groundwater in northern part (Gabiri *et al.* 2018). Floods are variable and discernable on the optical images. Lateral flow is deep-seated and more stable.

The current study showed that a link exists between land cover and depth of groundwater. High groundwater depths result in a high percentage of vegetated land cover class whereas bare land cover class mainly occurs in low groundwater depths and after flooding events. However, in the riparian zone, despite a high depth to groundwater in the dry season, vegetation thrived due to moisture emanating from the Kilombero River.

A more detailed assessment of thematic land cover was not possible due to the limited number of cloud-free images. Differentiation of crop types and mapping of the field boundaries was not possible due to the highly fragmented farm sizes and high variability of crops grown. To capture the field boundaries, very high-resolution imagery would be required. Furthermore, the limited number of matching cloud-free images and groundwater level restricted the analysis of the relationship between land cover and groundwater water depth.

3.5 Conclusions

The study understanding the land cover dynamics within Ifakara, in the Kilombero River floodplain, by generating land cover maps and cumulative seasonal land cover maps. The dynamics of land cover and depth to groundwater at three hydrological zones were evaluated. The first major finding was that in the rainy season, groundwater level is high and a large amount of land in the middle and riparian zone are vegetated. The riparian zone is flooded with a maximum observed extent of 3 km on both sides of the Kilombero River during the El Niño rains in 2014. However, the flooding intensities vary from year to year. The second major finding was that in the dry season, groundwater level is low, the area within the riparian zone is vegetated and the other zones are mainly bare. This availability of moisture capable is supporting vegetation growth in the margins of the riparian zone indicates that the moisture in soils emanates from the water in the Kilombero River. The high proportion of vegetated areas in the middle and fringe zones during the rainy season point to the good moisture conditions, which favor crop growth. Most sections of the riparian zone are classified as bare after the rainy season due to accumulation of soil sediments. The cumulative season maps revealed low vegetation coverage in the central and northeast regions of the study area. Groundwater responds to precipitation and the rate of groundwater recharge is higher than the rate of draining.

This research enhanced our understanding of land cover changes and their role in agricultural production in rainfed agricultural systems by providing spatiotemporal explicit land use information. Seasonal land cover information will help in land use planning with respect to suitable crops and farming practices that will lead to increased yield and therefore an increase in food security. Future studies investigating the economic capability of the farmers and agricultural development research institutions to adopt climate-smart agriculture is recommended. Continued monitoring is recommended to assess the impact of adopted technologies on agricultural yields. The effect of land use on groundwater quality should be assessed in the face of increasing pressure on water demands as it forms an additional source of water for the growing population and therefore any contamination points should be controlled and halted.

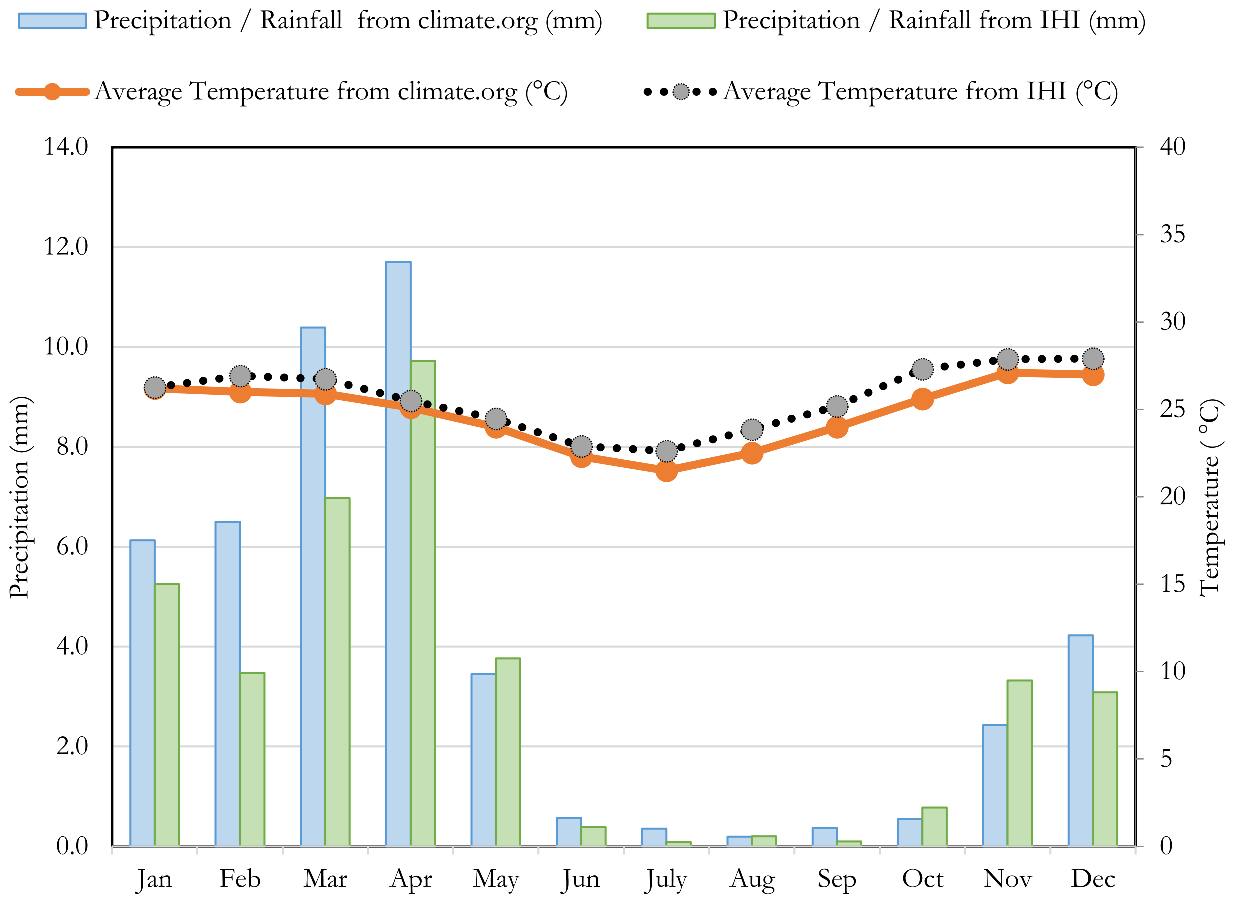


Figure 3.12: Comparison of climate data from IHI and climate data from climate.org for Ifakara

4. Multi-temporal Analysis of Sentinel-1 Texture and Principal Component Analysis features

4.1 Introduction

Tanzania is mainly an agricultural country with 90% of the rural population involved in the agricultural sector (Kalinga and Shayo 1997). Due to increased population and elevated food requirements, wetlands are utilized for agricultural activities, making them highly dynamic systems (Mombo *et al.* 2011). Lack of records showing the spatial extents and use of the wetlands affects their sustainable and wise use (Kalinga and Shayo 1997). Wetland monitoring is thus essential in ensuring controlled utilization of resources and activities to warrant a balanced ecosystem. Furthermore, development and implementation of policies regarding wetland use require detailed knowledge of wetland dynamics (Frenken and Mharapara 2001, Eitel *et al.* 2011). Wetlands undergo rapid changes over time and cover vast regions.

Therefore, monitoring such spatial phenomenon through ground survey methods is not cost effective. Synoptic viewing and temporal capabilities of remote sensors provide an alternative timely and cost-effective data acquisition platform (Guo *et al.* 2017). Longer wavelength microwave images are sensitive to the structure of target features whereas optical images are sensitive to the chemical composition of the features (Shang *et al.* 2009, Schuster *et al.* 2015). However, the presence of clouds adversely affects optical remote sensing imagery resulting in inconsistency in available satellite data. Due to this shortcoming, weather independent SAR observations are favored especially for tropical regions. Over the years, various studies have been carried out in mapping habitats and monitoring agricultural areas using SAR images (Evans *et al.* 2010, Turkar *et al.* 2012, Betbeder *et al.* 2015).

The scattering mechanisms of polarimetric parameters have aided in wetland monitoring in several studies. However, polarimetric decompositions are only possible in cases where fully polarized or co polarized images are available. Flooded vegetation is characteristic of double bounce whereas for open water specular reflection with lower backscatter values is dominant (White *et al.* 2015). Vegetated lands are dominated by volume scattering due to the interaction of the waves with the stems and leaves (Moser *et al.* 2016). These scattering mechanisms aid in discrimination of land cover types within the wetlands (White *et al.* 2015). (Morandeira *et al.* 2016) utilized the H/A/α decomposition parameters in mapping plant types in a wetland. Classification of model-based decomposition yielded higher accuracies as compared to Eigen decompositions in a wetland in the Amazon (Furtado *et al.* 2016). The Kennaugh decomposition matrix for dual polarized TerraSar-X images resulted in higher classification accuracies as opposed to the single polarized images in wetland monitoring (Moser *et al.* 2016). In another study, TerraSar-X Shannon Entropy resulted in higher classification accuracy compared to the single polarized HH and VV data (Betbeder *et al.* 2014). Similar backscatter values can occur for single polarized images and thus discrimination of features is low (Mosleh *et al.* 2015). Single polarized images require more acquisitions than multi-channel images (Shao *et al.* 2001).

Polarimetric elements are generated from dual polarized and/or quad-polarized Single Look Complex (SLC) products that have amplitude and phase information and requires two or more orthogonal polarizations (Cloude 2010). The Sentinel-1 images for the study area over the study period, however, are single polarized and thus pose a challenge with respect to generation of this polarimetric information. Incorporation of spatial dependence of neighborhood pixels was shown to increase land cover classification accuracy (Ghimire *et al.* 2010). Spatial structuring of Landsat TM using the Getis statistic detected changes in spectral values between objects otherwise not perceived in a comparative maximum likelihood classification (Wulder and Boots 1998). (Myint 2003) utilized the fractal approach, spatial autocorrelation which included Moran's I and Geary's C, mean and standard deviation in Landsat classification with the spatial autocorrelation resulting in highest accuracy. Variance, Moran's I and fractal dimension was used in another study in urban land cover classification (Emerson *et al.* 2005). The fractal approach proved to be effective in improving classification accuracy in urbanized areas.

The calculation of Grey Level Co-occurrence Matrix (GLCM) texture features is another method to spatially analyze neighboring pixel values. GLCM texture shows tonal changes within a set distance on an image. Utilizing information on arrangements of objects and their spatial relationships enables better discrimination of land cover classes (Ghimire *et al.* 2010, Rodríguez-Galiano *et al.* 2011). GLCM texture has been shown to improve overall accuracies in analyzing the health of trees using multispectral Ikonos imagery (Wang *et al.* 2015), in mapping wetland communities in the Everglades National Park (Szantoi *et al.* 2013), and in field-based classification using single polarized HH GLCM texture features from TerraSar-X images (Mahmoud *et al.* 2011). They have been shown to be correlated to forest biomass (Eckert 2012) and attempts have been made to map vegetation structure (Wood *et al.* 2012). GLCM texture was utilized in discriminating oil palm and rubber plantations from other land cover classes (Torbick *et al.* 2016). In that study, contrast and homogeneity were deemed to be the most important variables. (Inglada *et al.* 2016) utilized LandSat-8 and Sentinel-1 in crop monitoring where GLCM texture was generated from Sentinel-1 data. Energy, entropy and angular second moment were found to be the most important GLCM texture features. VV polarization was found to be more important input as opposed to the VH polarization. Texture information from ALOS and Sentinel-1 Dual polarized data were classified using the SVM and Maximum likelihood with SVM giving best accuracy (Homa Zakeri *et al.* 2017). Random Forest and SVM had highest accuracies as compared to Neural network and Maximum likelihood in mapping mountainous forests with spectral derived parameters, SAR derived parameters among which was GLCM and a combination of all GLCM features (Attarchi and Gloaguen 2014). Therefore it is evident from this classification literature, different classifiers have different strengths resulting in different accuracies for the same classes (Lu and Weng 2007).

Principal Component Analysis reduces the dimensionality of a large number of interrelated variables generating uncorrelated bands that maintain the variation in the data (Gupta *et al.* 2013). GLCM texture features have bands with potentially redundant information and thus PCA was applied to the texture features.

Chamundeeswari *et al.* (2009) assessed the PCA applicability of texture features in classification accuracy (Chamundeeswari *et al.* 2009). It was reported that PCA based classification yielded higher accuracies. The study, however, involved the k mean unsupervised classification.

Addition of GLCM features in wetland mapping has been assessed by studies such as Landsat GLCM increased the accuracy of change detection the Mediterranean coastal wetland indicating the importance of texture in providing additional information in areas with high spatial variation (Berberoğlu *et al.* 2010). Application of high-resolution Quickbird in the classification of riparian and forest ecosystem resulted in an increased accuracy of 2-19% for vegetation classes when texture information (Johansen *et al.* 2007). Texture from the red band of high-resolution IKONOS was incorporated in mapping plant communities in a wetland reporting high accuracies where texture features were applied (Barbosa 2010). (Kim *et al.* 2011) incorporated scale and texture in high-resolution aerial imagery in mapping salt marsh features (Kim *et al.* 2011). They reported that GLCM features improved between class discrimination yielding higher classification accuracies. Hyperspectral imagery was utilized in object-based mapping of the Everglades wetlands using GLCM texture features (Zhang and Xie 2012). In the study, minimum noise fraction was used to reduce the dimensionality of data prior to application of multi classifiers. High spatial and spectral resolution, the data transformation and combining GLCM spatial information were reported to increase accuracies in vegetation classification. Still, within the Everglades, the addition of NDVI and texture features to aerial photo spectral bands was reported to increase discrimination of plant communities with SVM yielding higher accuracies as compared to Maximum Likelihood (Szantoi *et al.* 2013). (Chatziantoniou *et al.* 2017) examined the performance of Sentinel 2 (S2) transformations (Principal Component Analysis and Minimum Noise Fraction), S2 spectral bands, S2 NDVI texture values, Sentinel-1 SAR in wetland mapping in the Mediterranean reporting that the texture features from NDVI improved the vegetation and bare accuracies by 2-4% (Chatziantoniou *et al.* 2017).

The Kilombero Wetland in Tanzania is mainly used for crop production. However, agricultural practices in use give low yields. As reported by (Ntongani *et al.* 2014), a local information survey showed that more land is cleared to make room for further agriculture (Ntongani *et al.* 2014). A quantification of the spatial extent of these agricultural fields is vital in regional agricultural management and decision making on how to produce more yield for the increasing population.

Flooding events are a common occurrence along the River Kilombero during the rainy season resulting to loss of crop. The farmers subsequently have reduced production for subsistence use. The weather independent image acquisition capability of SAR systems warrant their use in water extent mapping. Moreover, SAR's unique response in discriminating water and land favor its application in flood mapping. Flat surfaces reflect the signal away from the sensor decreasing the radiation recorded by the SAR sensor and thus dark pixels appear for water surfaces as opposed to the brighter land surfaces (Kussul *et al.* 2011, Refice *et al.* 2018). Monitoring the flood extent is an asset to farmers to determine yield expected based on flooding duration of their plots. Presence of clouds in this tropical region hinders the utilization of optical images in monitoring the area. However, these images are useful in selection of reference data. Within the study period, very few cloud-free scenes were available.

In this study, Sentinel-1 GLCM texture measures and PCA derivatives are evaluated to assess their suitability in mapping a floodplain wetland in Ifakara, Tanzania. The purpose is to show the potential of temporally consistent Sentinel-1 VV acquisitions in wetland monitoring which is crucial in understanding land cover changes in time thus deriving seasonal land cover cycle. Moreover, the response of different classification schemes is assessed. The methods applied include Support Vector Machine (SVM), Artificial Neural Network (NNET) and Random Forest (RF).

To test the quality of the classification, sensitivity, positive predictive values, kappa and overall are calculated according to (Congalton 1991). The unique feature of the study is the evaluation of GLCM texture features with their corresponding PCA images applying three different classification algorithms to a time series of SAR imagery over a tropical wetland. Moreover, a method that utilizes few available optical images for generation of reference data is tested since the area is adversely affected by cloud coverage. The main objective of the current study is the assessment of the potential and performance of single-polarized Sentinel-1 VV images, GLCM texture features and PCA images derived from the texture features with minimal reference data in continuous monitoring of the Kilombero wetland by applying different classification algorithms. Binary change masks are applied in selection of the training and validation data. Change detection involves identifying changes of an area between acquisition dates. Binary change detection deals with separating the areas into change and non-change areas (Hu and Ban 2014). To minimize the probability of erroneous changes between acquisitions, the images in the study are all from the Sentinel-1 sensor and were all acquired at the same incidence angle (ascending mode at 39°) and had the same temporal resolution (24 days). Additionally, flooding frequencies are assessed based on the classified images.

4.2 Materials

4.2.1 Datasets

4.2.1.1 Satellite Data

Sentinel-1 is one of the European Space Agency (ESA) Copernicus satellite systems with SAR sensors. It operates in C-band at a frequency of 5.4 GHz, an altitude of 693 km and an orbital inclination angle of 98.18°. The operational modes of the satellite are the Strip Map mode with 80 km swath with 5x5 m^2 spatial resolution. The Interferometric Wide Swath mode has a swath width of 250 km and a 5x20 m^2 spatial resolution. The Extra Wide Swath mode has a 400 km swath width and a 25x100 m^2 spatial resolution. Finally, the Wave mode has a 20x20 km^2 swath width and a 5x20 m^2 spatial resolution (European Space Agency 2013). Single Look Complex (SLC) and Ground Range Detected (GRD) data products are available for the four modes. The images selected over Ifakara were acquired in the Interferometric Wave mode single polarizations (VV) at an incidence angle of

39° in ascending mode at 15:53hrs. A time series of Sentinel-1 GRD images were acquired from the Sentinel's scientific data hub (https://scihub.copernicus.eu/dhus/).

A level 1B RapidEye image acquired on 27th April 2015 from http://blackbridge.com (Planet 2016) and two Sentinel-2 images acquired on 26th December 2015 and 5th March 2016 from https://scihub.copernicus.eu/dhus/ (ESA Sentinels Scientific Data Hub 2016) were the reference images from which the training and validation data was derived. The RapidEye and Sentinel-2 images were orthorectified and coregistered ensuring correct positional accuracy prior to generation of NDVI and reference points for classification.

4.2.1.2 Reference Data

Field data collection is vital in determining how accurate a classification has been performed. This motivated the need to gather ground reference information to be used in assessing the accuracy of the classifications.

Reference data was obtained from ground measurements and remote sensed derived products. The field reference data was collected using Etrex 30 Garmin GPS between February and May 2015 and December 2015 and February 2016. During field planning, random points were selected on a base map before the actual fieldwork. During the fieldwork, these points were visited to confirm the actual land class. Moreover, a UAV flight campaign was performed in February 2015.

To secure a sufficient number of reference points, RapidEye and Sentinel-2 optical images were included in the generation of reference data. NDVI was also generated as a means of improved land class interpretation. The water and built up areas were relatively easy to identify as the town is located towards the middle of the study area and the Kilombero River flows on the lower section of the study area. At least 400 reference points for each of the optical image dates (27th April, 26th December 2015) were randomly selected in ArcGIS 10.3.1 (Figure 4.1). The land cover classes were interpreted by visual inspection of the false color composite and corresponding NDVI and with the aid of the UAV photos and GPS points. Photographic records of different land covers were also acquired. The reference data is listed in Appendix A1.

Challenges faced during the GPS data collection included inaccessibility of some locations as they were within the protected area. Moreover, the lands are communally owned and hence permission to access the lands was not always possible. In other cases, the road network proved challenging to reach the locations. This prompted the use of the remote sensing derived reference points. However, cloud coverage was a major challenge in the use of NDVI for additional reference point generation. The small sizes of farms accompanied by different farm management practices resulting to highly varied land cover confined the study to four broad classes. Collection of spectral signatures from individual crops was therefore not part of the current study.

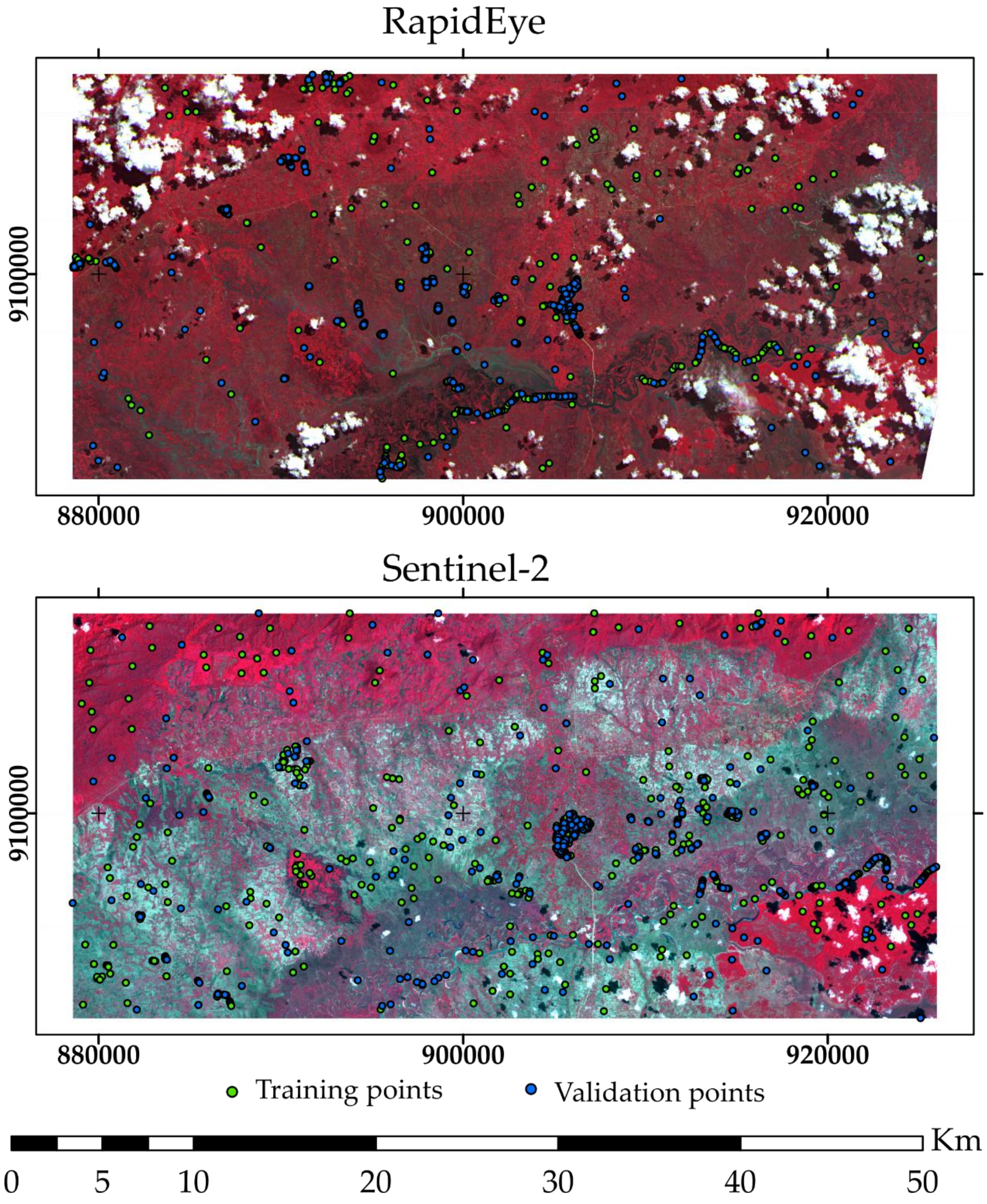


Figure 4.1: Training and validation points over RapidEye (27-04-2015) and Sentinel-2 (26-12-2015) scenes. The colors of the multispectral satellite imagery appear according to the band combinations 5-3-2 (RapidEye) and 8-4-3 (Sentinel-2). The white areas are clouds.

Differences in image appearance are due to sensors and time of acquisition. The RapidEye image was acquired on 27th April 2015, during the rainy season when the fields are vegetated. The Sentinel-2 image was acquired on

26th December 2015 during the beginning of the rainy season just after the dry season. The fields at this time are mainly bare. The display for both images used the band combination NIR: R: G.

4.3 Methods

Availability of consistent coverage of single-polarized Sentinel-1 data over the Kilombero floodplain during the period of study enables continuous monitoring of the wetland. The influence of utilizing GLCM texture features is assessed by applying different classification schemes and evaluating uncertainties and classification errors. A flow chart is given in Figure 4.2.

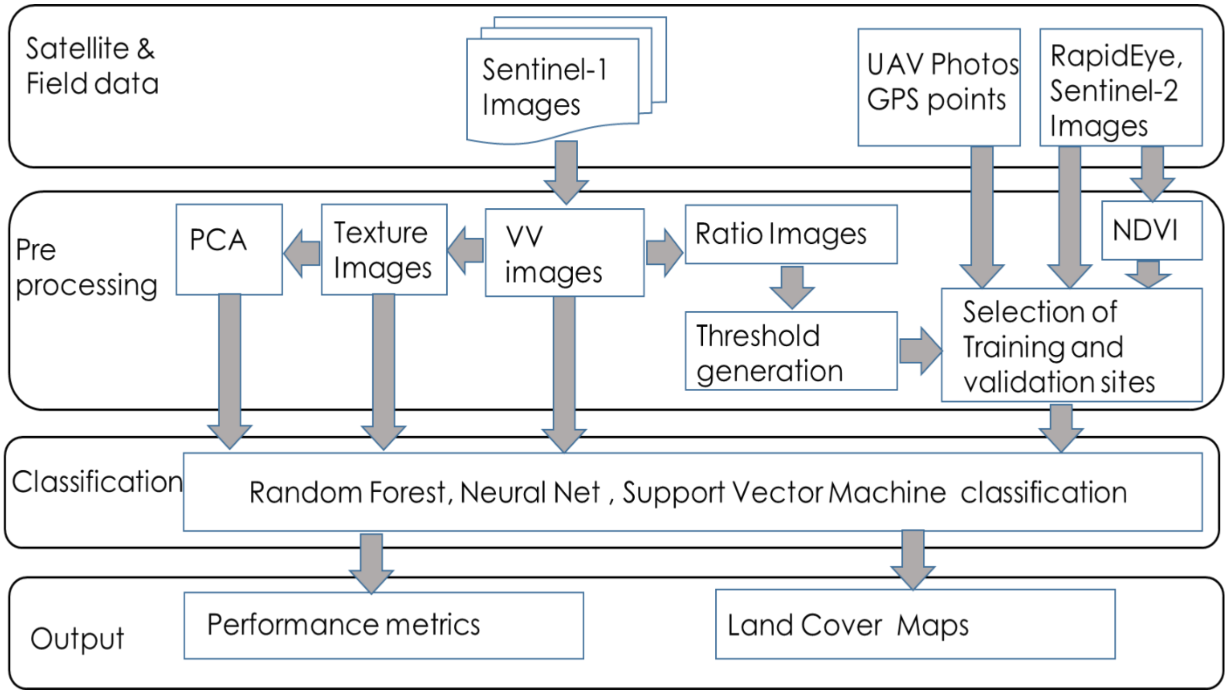


Figure 4.2: Data process flow

Radar data processing is performed with the SNAP Toolbox (Sentinel-1 Toolbox n.d.). The IW GRD single VV polarized Sentinel-1 data is radiometrically and geometrically calibrated using the 3-second SRTM DEM to correct for topographic variations and resampled to 10 m spatial resolution. Lee refined speckle filtering is applied for noise removal. However, application of speckle filtering on images before texture extraction has been an issue of controversy with some studies applying it (Pultz and Brown 1987, Pierce *et al.* 1998) while others do not (Treitz *et al.* 2000, Arzandeh and Wang 2002). The intensities were converted to Sigma 0 backscatter values in decibels. The region of interest was extracted and used for subsequent analysis.

4.3.1 Grey Level Co-occurrence Matrix (GLCM)

Grey Level Co-occurrence Matrix developed by (Haralick *et al.* 1973) was designed to estimate how often variation of grey level intensity of a pixel occurs within a given moving window size (defining local window within which the GLCM is computed) in a given direction (angle and distance defining neighbourhood relationship for the pixels whose GLCM is being computed) (Haralick 1979). The generation of GLCM texture involves generation of spatial relationships of the reference and neighbor pixels, the decision on the window size or separation between the pixels (*d*), generating symmetrical matrices and normalizing to obtain the probability of occurrence (Pradhan *et al.* 2014). Interpixel distance selected for the GLCM computations included 5, 7 and 9. The orientations selected, based on the position of the reference pixel, included 0° (considers pixels to the east and west of reference pixel), 45° (considers pixels to the northeast and south-west of reference pixel), 90° (considers pixels to the north and south of reference pixel), 135° (considers pixels to the north-west and south-east of reference pixel) and a selection considering all the pixels surrounding the reference pixel (direction invariant).

Taking $I(B_x * B_y)$ to be the satellite image with spatial extents in x- (B_x) and y-direction (B_y), $(x_1, y_1)(x_2, y_2)$ are reference and neighborhood pixels whose texture is being computed and separation between the pixels is d, then GLCM is given by (Haralick *et al.* 1973):

$$GLCM\ (i,j) = \{((x_1,y_1),(x_2,y_2)) \in (B_x * B_y) \times (B_x * B_y)$$
$$x_1 - x_2 = 0, |y_1 - y_2| = d, for\ 0°\ direction$$
$$x_1 - x_2 = d, y_1 - y_2 = -d, for\ 45°\ direction$$
$$|x_1 - x_2| = d, y_1 - y_2 = 0, for\ 90°\ direction$$
$$x_1 - x_2 = d, y_1 - y_2 = d, for\ 135°\ direction$$
$$I(x_1,y_1) = i, I(x_2,y_2) = j\}$$
$$d(x_1,y_1),(x_2,y_2) = \{|x_1 - x_2|, |y_1 - y_2|\}$$

GLCM is the probability of occurrence of a pixel with value j for all values from 1 to B occurring at a distance d with respect to a pixel i for all values from 1 to B on an image (the number of times gray tones i and j have been neighbors). Different weights are then assigned to the GLCM probability values depending on the output information required (Haralick 1979, Hall-Beyer 2007) (Table 4.1). GLCM texture features can be arranged in three main groups: The contrast and orderliness groups as well as the descriptive statistics group. The contrast group is designed to enhance variations of the pixel intensities. Weights of dissimilarity increase linearly whereas those of contrast increase exponentially highlighting variation present in the image. For homogeneity, the weights decrease exponentially with increased variation of the neighboring pixels indicating large values in areas with small differences in grey levels. The orderliness group is designed to enhance regularity or commonness within a window. It comprises of Angular Second Moment which uses the square of the probability as a weight for itself whereas the square root of this is the energy/uniformity. ASM measures uniformity or repetitions of pixel pairs such that large values indicate less local texture variation. Entropy is obtained by multiplying the GLCM logarithm with itself, measuring image disorder. Small entropy values indicate textural uniformity of the image. Lastly, the general statistics are defined in the descriptive group based on the GLCM matrix values. The GLCM mean, variance and correlation are computed based on the frequency of occurrence in relation to the neighboring pixels. Mean incorporates the grey level of each line in the texture calculation. Variance is a measure of heterogeneity which increases when the grey level values differ from the mean. Correlation is a measure of grey tone linear dependencies in the image. High correlation values indicate linear relationships between grey levels of pixel pairs (Arzandeh and Wang 2002). Table 4.1 lists the GLCM texture feature formulation.

Table 4.1: The GLCM texture features extracted from each VV SAR image

	Texture feature	Formulas
Contrast group	Contrast	$\sum_i \sum_j GLCM_{i,j}(i-j)^2$
	Dissimilarity	$\sum_i \sum_j GLCM_{i,j}\lvert i-j\rvert$
	Homogeneity (Inverse Difference Moment)	$\sum_i \sum_j \frac{GLCM_{i,j}}{1+(i-j)^2}$
Orderliness group	Angular Second Moment (ASM)	$\sum_i \sum_j {GLCM_{i,j}}^2$
	Energy (Uniformity)	$\sqrt{\sum_i \sum_j {GLCM_{i,j}}^2}$
	Maximum Probability (MAX)	Gives the largest probability within the selected window
	Entropy	$-\sum_i \sum_j GLCM_{i,j}(\log GLCM_{i,j})$
Descriptive statistics	GLCM Mean	$\mu = \sum_i \sum_j i(GLCM_{i,j})$
	GLCM Variance	$\sigma^2 = \sum_i \sum_j GLCM_{i,j}(i-\mu)^2$
	GLCM Correlation	$\sum_i \sum_j GLCM_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{({\sigma_i}^2)({\sigma_j}^2)}}\right]$

The experiments conducted on the set of images included:

Experiment 1: Classification of multi-temporal GLCM features: The aim of this experiment is to make use of the temporal stability of GLCM texture features in generating multi-temporal time series with the aim of understanding the land cover patterns over the study site for a period of two years. Moreover, three classification schemes are applied. The aim is to identify the best classification scheme from the texture features over a floodplain overlooked by mountainous ranges. Comparative analysis of classification performance of the GLCM images and PCA images. The first three PCA features were selected following (Chamundeeswari *et al.* 2009) as these components capture the maximum variance in the texture features. The aim is to assess the impact of utilizing PCA features as opposed to all the GLCM texture features in the classification endeavors.

Experiment 2: Water extent assessment: Based on the classification obtained from experiment 1, temporal changes in water extents were assessed. The aim was to identify areas prone to submergence and subsequent loss of crop yield.

4.3.2 Classification schemes

Optical imagery are sources of reference data in cases where actual ground data is unavailable. Due to the scarcity of optical data over the study area because of cloud coverage, change detection using binary change mask was adopted in the generation of training data for the land cover classification. The method involves performing a classification on the SAR image acquired on date 1 using the training data acquired from a corresponding optical image date. Next, the difference in backscatter between the SAR image on date 1 and SAR image on date 2 is analyzed by generating a ratio image (Kim 2016) (Figure 4.3).

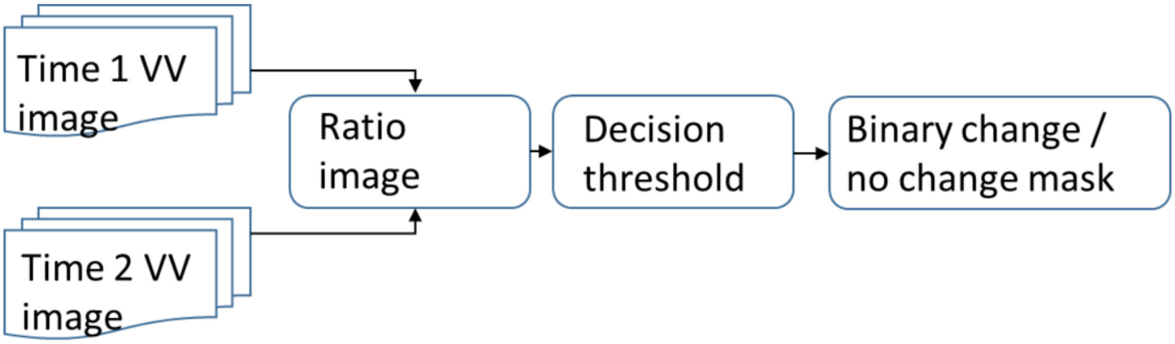


Figure 4.3: Generation of the binary mask images

Areas of "change" and "no change" are defined by applying decision threshold to the ratio image histogram. The ratio image is classified into a binary mask file of change and no change areas. The no change mask is overlaid on the training dataset. The training areas within the no change areas are selected and applied in classification of the SAR image acquired on date 2. The assumption in the method adopted is that the areas marked as changed are associated with land cover changes (Chen 2008). The main reason for selection of the method was the unavailability of ground truth information for all the time series of sentinel-1 image acquisitions spanning 2015-2016. Change detection by the pixel-based image ratioing is indicated by (Hussain *et al.* 2013);

$$I_{ratio} = I_{i+1}/I_i$$

$$where\ I_{i+1}\ is\ the\ VV\ polarized\ image\ at\ time\ 2\ and\ I_i\ is\ the\ image\ at\ time\ 1,$$
$$assuming\ availability\ of\ a\ corresponding\ optical\ image$$

The ratio is in the range of 1 whereby values greater than 1 indicate a higher VV backscatter in time 2, with values lower than 1 indicating a higher backscatter in time 1. Values of 1 indicate no change in backscatter between the two image acquisition times (Hussain *et al.* 2013). A histogram representing the distribution of the ratio values was generated. A binary mask was developed by setting values smaller than 20% percentile and values larger than 80% percentile to zero and all remaining values to 1. Taking time 1 to have a corresponding optical image from which reference areas (from section 2.2) are selected, then the reference points falling within the binary mask with a value of 1 are used in the classification of the time 2 image. The method is based on the selection of reference points from areas of no change between two images taken on different dates. For each of the images considered, stratified sampling was performed on the reference points such that 70% were training data and the remaining 30% were for validation of the classification (Appendix A1).

The GLCMs were stacked into one data set as it has been reported that using a single texture feature gives poor results (Arzandeh and Wang 2002). Classification of the texture features was carried out with three different classifiers. These included Random Forest, Support Vector Machine and Artificial Neural Network. The classification was implemented in R-Studio with package RStoolbox (Benjamin Leutner and Horning 2017). The classification models were imported from the caret package. The “superclass” function which requires the input image, the classification model, the response parameter, the training, and validation data.

SVM is based on the generation of hyperplanes which best differentiate the input features by acquiring the largest minimum distance to the training samples thus placing the features in the best possible class. To enlarge the feature space, nonlinear kernels are used. Adjusting the bandwidth and regularization parameter controls the SVM behavior. The radial basis function was used in this research.

NNET are a series of neurons that model non-linear relationships between predictors and response parameters. The algorithm learns a function by training the data. It consists of three layers, input, hidden and output. The hidden layer can have many nonlinear layers.

The non-parametric supervised ensemble classification algorithm Random Forest (RF) generates a series of trees where each contributes as a vote in deciding the classes to which the features in the input image belong (Breiman 2001). RF uses ensemble learning where samples are randomly selected from the training dataset to grow a so-called classification tree. The splits at the nodes of these trees are determined by purity measures given by the Gini index.

Entropy gives the degree of randomness in a set of predictions. This is used in assessing the similarity of the classification algorithms in assigning the classes. The entropy is zero if the classification of that pixel is the same in the three algorithms.

4.3.3 Validation of classifications

For the validation of the classifications, a set of validation metrics is calculated. Please note that classification is performed for SAR data only, where the multispectral data has been utilized to identify training and validation points. The separation of training points from validation points offers an independent unbiased accuracy assessment. Validation points to assess the accuracy of SAR classification at dates where no multispectral data is available are identified by implementing the same procedure introduced in section 3.2 to identify training points. The number of training and validation points for each date is listed in the Appendix A1. The classes to investigate are bare, built up, vegetated and water. This simplified analysis of four classes only is justified by the general aim of this study to investigate the temporal course of vegetation and water dynamics in the Kilombero wetlands. The temporal course of the validation metrics for the classifiers Random Forest, Neural Net and Support Vector Machine are compared for all classes jointly by kappa coefficient and overall accuracies (reference to these metrics) as well as by the positive predicted value and the sensitivity measure given by:

$$positive\ predicted\ value = True\ positive\ /(True\ positive\ \ + False\ positive)$$

$$sensitivity = True\ positive\ /(True\ positive\ \ + False\ negative)$$

4.4 Results and Discussion

4.4.1 GLCM temporal classification

Texture features were generated with the 9*9 moving window size in the invariant direction (the average of all four spatial arrangements). The 10 texture features were combined into a single image. PCA was generated from this dataset. Random Forest, Neural Net and Support Vector Machine resulted in the overall accuracy and kappa coefficients displayed in Table 4.2.

Table 4.2: Averaged Overall Classification Accuracy for Random forest (RF), Neural net (NNET), Support Vector Machine (SVM) classifiers. TEX, VV, PCA are texture, single polarized and Principle Component derived overall classification accuracies

	VV			TEX			PCA		
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
Kappa coefficients	0.49	0.59	0.58	0.68	0.68	0.68	0.78	0.78	0.78
Overall accuracy	0.62	0.70	0.70	0.77	0.77	0.77	0.84	0.84	0.84

Overall accuracies for the single polarized VV images were 0.62 for RF and 0.70 for NNET and SVM respectively. The GLCM featured resulted in accuracies of 0.77 for all the three classification algorithms. PCA features reported the highest accuracy of 0.84 for all the algorithms. The high overall accuracies of the PCA images indicate that reduction in dimensionality of texture data to its basic components containing the most variance increases the accuracy of classification. The GLCM are relatively large and therefore compressed with PCA. PCA texture features capture the variation of the land covers and are recommended for use when dealing with a large dataset.

Similar to the pattern in the overall accuracy results, the single polarized classifications exhibited low kappa coefficients. Single polarized kappa ranged from 0.62 to 0.70 for the three classifiers representing its low ability to identify features from single polarized SAR images. The Kappa coefficients for the GLCM texture features and PCA features were 0.77 and 0.84 respectively. The results indicate that PCA features have high potential in monitoring land cover in the Kilombero wetland yielding higher accuracies as compared to utilizing the single polarized sentinel VV images or their corresponding texture features.

Table 4.3: Average Sensitivity and positive predicted values for single polarized, texture features and Principal component classified images

	VV			GLCM			PCA		
sensitivity	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
bare	0.45	0.52	0.51	0.61	0.62	0.61	0.66	0.66	0.66
vegetated	0.50	0.64	0.68	0.66	0.67	0.67	0.84	0.84	0.84
built up	0.55	0.63	0.57	0.85	0.79	0.84	0.89	0.89	0.89
water	0.95	0.97	0.97	0.95	0.94	0.92	0.96	0.96	0.96
positive predicted value	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
bare	0.47	0.54	0.56	0.64	0.64	0.63	0.83	0.83	0.83
vegetated	0.47	0.56	0.54	0.63	0.65	0.63	0.74	0.74	0.74
built up	0.59	0.79	0.82	0.87	0.89	0.88	0.94	0.94	0.94
water	0.96	0.97	0.97	0.97	0.98	0.98	0.95	0.95	0.95

The sensitivity gives the probability that a class in the classified map is actually in that class on the ground indicating the effectiveness of a classifier to identify a land cover class correctly. The sensitivity of the bare class was lowest for the single polarized images with the application of the Random Forest classifier (Table 4.3). The three classification algorithms reported the highest sensitivity for the bare class (0.66) in the classification of the PCA features. The ability to identify vegetation was lowest for the single polarized images with the Random Forest classifier, while the highest vegetation discrimination ability is reported for the PCA images irrespective of the classifier (0.84). Texture features have a similar performance in vegetation discrimination as the single polarized SVM algorithm. Discrimination of built up areas is best for the PCA features regardless of the classifier while it is lowest for the single polarized images. The classifiers ability to discriminate built up areas from both the texture and PCA images was highly throughout the time series. Built up areas have high backscatter due to double-bounce scattering hence increasing their discrimination potential (Zhang *et al.* 2014). Water was easily distinguishable for all classifiers for the single polarized, texture and PCA images due to its distinctive low backscatter values.

The positive predicted value gives the probability that a land cover on the ground is classified correctly on the map. The bare prediction was lowest for the Random Forest single polarised images. A slight improvement was observed for the texture features with the best prediction ability recorded for the PCA features. Similarly, prediction ability for the vegetated land cover class was lowest for the single polarised images and highest for the PCA features. It was noted that the prediction for vegetated areas was slightly lower than for the bare area when dealing with the PCA features. All the classifiers correctly mapped built up areas at high precision, greater than 0.8, except for single polarized images Random Forest and NeuralNet. Classification of water was high with a precision of 0.9 to 1 due to its unique low backscatter characteristics. The complete tables of overall accuracies, kappa coefficients, sensitivity and positive predicted values are in the Appendix A2-A11.

The classified maps appear identical (Figure 4.4 a, b, c) nonetheless the uncertainty map (Figure 4.4d) reveals the areas with differences between the models. The areas along the Kilombero River have a high uncertainty rate represented by the high values of entropy. Mixed pixels could be a causal effect of the high entropy values as varied covers including submerged vegetation is located within this riparian zone.

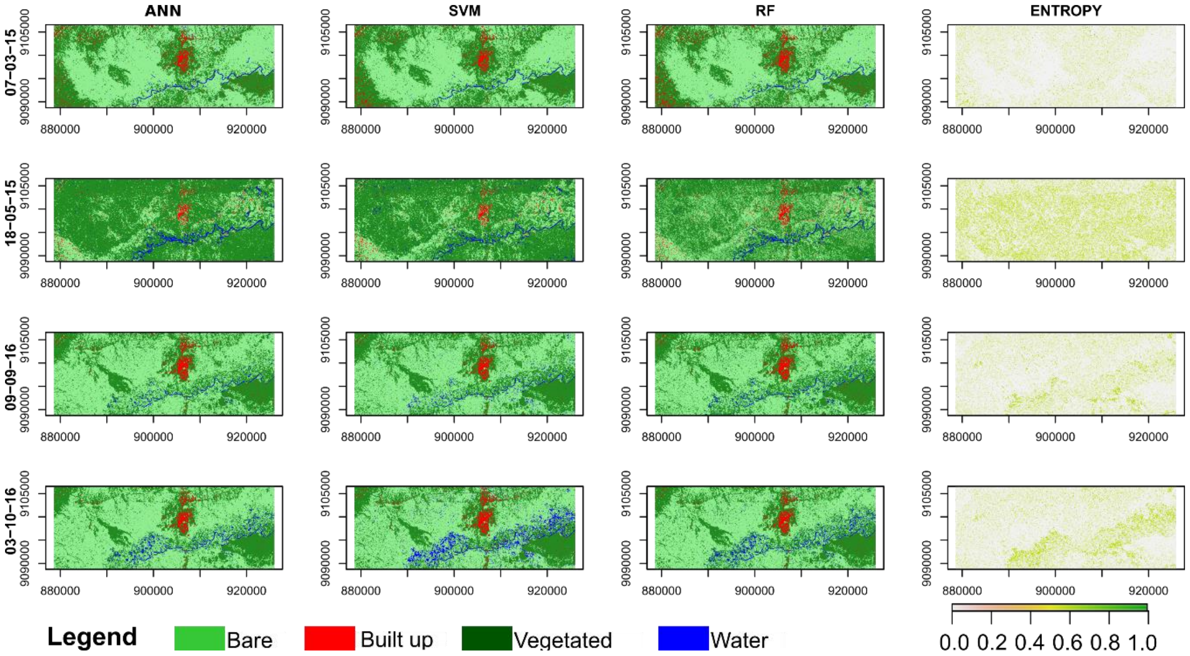


Figure 4.4: Land cover classifications of selected dates based on Artificial Neural Network (ANN), Support Vector Machine (SVM) and Random Forest (RF). The corresponding Entropy image indicates difference in the assignment of land use classes using the three classifiers low vales (light areas) have common class assigned whereas the high vales (green areas) were different for the three algorithms.

SAR backscatter constitutes contribution from vegetation, soil water content and surface roughness. SAR is an ill-posed problem as different combinations influencing factors can give the same value. The method adopted for selection of training sites involved ratioing backscatter values from two dates. Hence in the selection of non-change areas, land cover changes could have occurred but with no significant magnitude thus assumed to have the same land cover. Inability to decompose the total backscatter into individual contributory elements could introduce errors in classification since the changes recorded could be due to, for example, a change in soil moisture and not necessarily a change in the land cover class. This is a shortcoming of the method adopted though it worked well in general. Additionally, the GLCM texture features are derived from the backscatter values and hence are also subject to influence by vegetation, soil water content and soil roughness (Kurvonen and Hallikainen 1999).

4.4.2 Flooding extents and farmer information

The highest flooded area was computed by superimposing the water class from the series of classified images. Multiple SAR images thus enabled the identification of maximum flood extent. Temporal assessment of water extent indicates that the areal coverage of water along the riparian region is highest in the month May which is consistent with the precipitation pattern of the area (Figure 2.3). The farm located up to 3 km from the Kilombero River to the west of Ifakara and Minepa wards (circle 1), south of the Kilombero River in Minepa ward (circle 2) and up to 6km to the East of Kiberege ward (circle 3) has the highest likelihood of submergence in the rainy season (Figure 4.5). Farms located in this area, on one hand, are at risk of a reduced cultivation period and a reduced yield. On the other hand, temporary floods transport necessary nutrients onto the fields and improve the soil conditions. The resulting map time series provides useful information to plan general management strategies by the authorities and to individually support farmers in this area where rainfed subsistence agriculture is the main economic activity. In addition, setting up insurance policies for farmers currently adopting a balance between risk and chance of crop growth is required. The analysis at hand is based on two-year data, which makes it prone to general variability between the years, e.g. according to teleconnection processes such as El Nino Southern Oscillation. Therefore, the study needs to be extended to gain general decadal flood probabilities as soon as the SAR record is long enough.

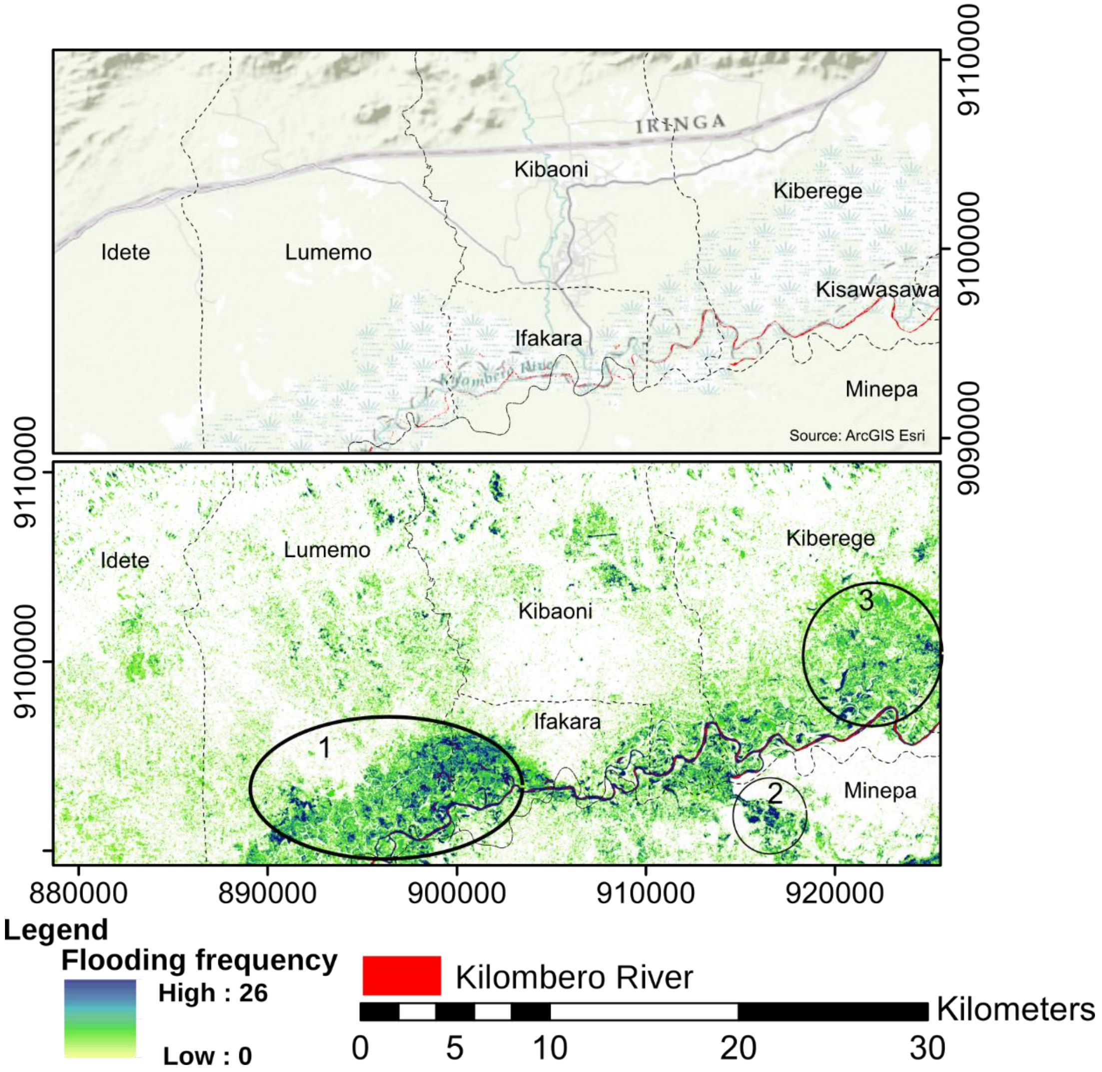


Figure 4.5: Flood-prone extents derived from the multiple SAR images. The number of images used in generating the cumulative land cover map were 26. The circles show example regions with floods extending more than 3km from the Kilombero River.

The continuous monitoring of land cover change with the cloud independent SAR aids in assessing the vegetation patterns, which in turn are linked to the yield expected. The study has shown the potential of Sentinel-1 in assessing the patterns of land cover in a region where data on dynamics is scarce. Land cover maps and patterns are useful to individual farmers as they provide a means of inventorying areal coverage and patterns of land use on their farms. This provides a basis upon which evaluations on suitable mechanisms to improve agricultural yield are adopted. At the management level, the land cover maps are preliminary databases providing information on vegetated land. With further analysis, cropped areas and rates of growth in areal coverage and their link to yields produced give insight into establishing measures of sustained use of the wetland.

The flooding extents aid in the identification of flood susceptible fields. For the farmers, the information aids in preventing absolute loss by delaying the planting time till after the heavy precipitation. The agricultural support organizations in the area would consequently develop flood-combating mechanisms helpful to farmers to prevent total loss of food crops. Organizations could also offer technical advice and training to farmers in flood-prone areas on alternative crops and planting times.

4.5 Conclusions

Utilization of texture features in delineating land covers is tested in the Kilombero Valley, Tanzania. Texture features were extracted from consistently available single polarized Sentinel-1 VV time series to monitor their performance in land cover mapping. Moreover, PCA of the texture features was generated to assess the effect of the use of data whose dimensionality has been reduced while retaining its variation. SVM, NNET and RF classification algorithms were applied on the single polarized, texture and PCA images. Due to the scarcity of optical images from which reference data is sought, ratioing the VV images corresponding to optical image dates was adopted to derive non-change areas. Reference data within the no change areas for each subsequent image in the time series was applied in classification. Single polarized images yielded low classification accuracies as compared to the texture and PCA images. PCA had higher overall accuracy values (84%) as compared to the texture features (77%). This indicates that the reduction in dimensionality of the data has a positive effect on the accuracies. The three classification algorithms had similar and higher performances for the PCA suggesting that reducing the dimensionality of the texture features into components containing most variance is beneficial for improved classification. Additionally, irrespective of the classifier applied for the PCA features, the ensuing performance is similar. In terms of correct identification of classes, built up and water had high sensitivity and positive predictive values. This is explained by their distinctively high and low backscattering coefficients values respectively in SAR imagery. The relatively lower bare and vegetation sensitivity and positive predictive values indicate other factors such as soil moisture, surface roughness, and the vegetation attenuation had an influence of the SAR backscatter recorded. Therefore, the training sites selected from image ratioing could have contributed to interclass confusion during the land cover classification. Entropy maps indicate the level of uncertainty between the classification algorithms applied. The areas of most disagreements were in the riparian zone, near the Kilombero River. Submerged vegetation is a common occurrence in this zone thus contribution to variations in backscatter values subsequently causing confusion in the classifications. Areal extents of flooding indicated that maximum flooding occurs up to 3 km from the Kilombero River. The implication to the farmers is the high probability of loss of crops due to flooding. Flood extent information is useful to the authorities in terms of offering support to farmers whose crops are affected by floods. Moreover, the inundation patterns are valuable as inputs in flood modeling geared towards reduction of crop loss. With respect to the wetland and sustainable use, matching the crops that can do well given the water available is recommended. Additionally, analysis of the precipitation patterns will give more accurate information particularly with respect to inundation duration, appropriate planting and harvesting periods.

5. Spatial-temporal variation of physical soil properties and soil moisture estimation from SAR data

5.1 Introduction

Increased food demand caused by high population growth has led to encroachment of wetlands in East Africa. Studies by (Rebelo *et al.* 2010, Mombo *et al.* 2011) highlighted the importance of the Kilombero wetland for food production. Assessment of soil properties at field scale captures the infield variation which could influence farm management practices for higher yields (Gebbers and Adamchuk 2010, Stoorvogel *et al.* 2015). Soil moisture influences groundwater recharge, surface, and subsurface runoff and is important in understanding hydrological responses in a watershed (Korres *et al.* 2015). Knowledge of soil moisture distribution is important in agricultural management and hydrological studies (Hupet and Vanclooster 2002, Moran *et al.* 2004, Henderson and Lewis 2008, Ziadat *et al.* 2010, Nichols 2011, Brocca *et al.* 2017).

Soil moisture is highly varied in space and time and hence analyzing the spatial and temporal persistence of soil moisture contributes to the understanding of land surface systems and the development of management options for improved productivity such as the implementation of irrigation schemes. Spatial-temporal variation of soil moisture can be analyzed via statistical methods involving means, standard deviation, the coefficient of variation and relative differences. Geostatistical analysis, on the other hand, examines the spatial structure of point data (Brocca *et al.* 2007). At field scales, spatial variations are due to differences in topography, the texture of the soil and temporal variability due to management practices such as tilling and cropping (Jenny 1971, Famiglietti *et al.* 1998, Brocca *et al.* 2009). In wet conditions, soil moisture variability was influenced by porosity and hydraulic conductivity whereas in dry conditions it was influenced by elevation, aspect and clay content in a study along a hillslope transect (Famiglietti *et al.* 1998). Field scale variation is site-specific and therefore transferability to other study sites needs investigation depending on the locations physiographic parameters (Entin *et al.* 2000, Brocca *et al.* 2007). Analysis of temporal consistency aids in the probability of upscaling the point data to a mean value comparable to a remote sensing product (De Lannoy *et al.* 2007). Studies on spatial-temporal variability of soil moisture at field scales with emphasis on determining stability of soil moisture at certain locations within the test fields are important in agricultural management. Brocca et al. (2007, 2009) assessed the spatial structure of soil moisture in an inland valley at three locations of varying elevation in Italy reporting a high negative correlation between elevation and soil moisture (Brocca *et al.* 2007, 2009). Within agricultural fields, Heathman et al. (2012) and Zhang et al. (2013) utilized permanently installed moisture sensors to assess the spatial-temporal patterns of soil moisture for future applications in hydrological modeling and irrigation management (Heathman *et al.* 2012, Zhang *et al.* 2013). In both studies, the potential of utilizing soil moisture networks in representing local field scale averages was revealed.

Geostatistics examines the spatial structure of point data, which involves estimation and modeling of spatial correlation (covariance or semi-variance) (Hengl *et al.* 2007, Hengl 2009). Spatial prediction aids in determining values in unvisited locations surrounded by existing observations. The analysis is a function of the distance between observed measurements, overall coverage and area occupied by each sample (Western and Blöschl 1999). Field scale soil moisture distribution maps aided in investigating the moisture patterns while highlighting the in-field variability, important in identifying water-stressed regions (Zhang *et al.* 2013).

The classical methods of soil moisture measurements including TDR probes, gravimetric measurements are time-consuming and costly over large areas. This has led to the use of remote sensing imagery in the determination of soil moisture. The strengths in the use of microwave imagery lie in their extensive areal coverage and independence from effects of solar irradiation and cloud coverage. Moreover, the consistent availability of SAR imagery provides a potential source of regular soil moisture estimates that are vital in agricultural management such as planning for auxiliary sources of water for plants in cases of low soil moisture. Active radar imagery have high spatial resolution and are thus preferred in the estimation of soil moisture over extended areas.

In situ soil moisture is also required to authenticate any estimates of soil moisture computed from satellite imagery. The mismatch in spatial resolution between point data and satellite footprints (e.g. TerraSar-X, 25-50 km) and the high spatial-temporal variation of soil moisture introduces challenges in the aggregated use of point and satellite data in soil moisture estimation over large areas. Differences in spatial resolution between field and satellite footprint soil moisture prompt a need to determine the minimum number of points required to give an average representative value of soil moisture over a defined areal coverage comparable to the satellite footprint (Jacobs 2004, Brocca *et al.* 2007). Therefore, for future auxiliary data collection in the study site, the minimum number of points required to obtain soil moisture within error values of ±1 and ±2% volumetric soil moisture at 95% confidence interval is computed. Identification of fields whose average soil moisture is stable in time will further reduce the number of measurements required during field campaigns. Temporal persistence of soil moisture patterns assesses fields which are stable in time and therefore could be used as long-term test sites for determining soil moisture dynamic used as reference point for remote sensing products.

SAR imagery has been used in soil moisture derivation in a large number of studies (Das and Paul 2015a, Zhang *et al.* 2015, El Hajj *et al.* 2016, Yue *et al.* 2016, Kim *et al.* 2017). Backscatter recorded by the SAR imagery is influenced by roughness, soil moisture, and presence of vegetation (Petropoulos *et al.* 2015). The backscatter from bare soils is a factor of roughness and soil moisture (Ulaby *et al.* 1996). Methods of collecting roughness information include photogrammetric methods (Taconet and Ciarletti 2007, Grims *et al.* 2014), Light Detection and Ranging (LIDAR) methods (Fernandez 2010, Turner *et al.* 2014) with the most common being the use of pin profiler. Parameterization of roughness is a great source of error in soil moisture retrieval (Verhoest *et al.* 2008, Álvarez-Mozos *et al.* 2009) and hence methods seeking to utilize roughness from alternative sources have been on the rise. Attempts to reduce the effect of roughness include the application of rainfall reducing factor (Jackson *et al.* 1997), and the use of effective roughness parameters (Zribi and Dechambre 2003, Lievens *et al.* 2009, Lievens and Verhoest 2011).

The practicality of roughness parameter collection by the local population at Ifakara is minimal. The pin profiler requires time and skill to operate and analyze the data to obtain the root mean square height (rmsh) and correlation length. Additionally, presence of high field variability, small and uneven farms and different management practices dependent on the farm owners call for the development of a priori roughness information as a solution to the challenges arising in the collection of in situ roughness measurements. Moreover, representation of farm soil parameters with averages of point data gives rise to errors subsequently propagated to the soil moisture retrieval and thus a representative roughness parameter is required. Previous studies have obtained apriori roughness information through inversion of empirical models (Mattia *et al.* 2006), multi polarized images (D'Urso and Minacapilli 2006), multi-angular images (Zribi and Dechambre 2003, Rahman *et al.* 2008, Sahebi and Angles 2010, Wang *et al.* 2016), co-polarized images (Magagi & Kerr, 2001) and the use of polarimetric parameters (Hajnsek *et al.* 2003).

In summary, the study generates temporal moisture maps from TDR probe measurements showing variation in the moisture in 14 agricultural fields located in different hydrological zones within a floodplain, investigates the spatial-temporal variation of soil moisture in small-scale fields that are approximately 700 to 19,000 m^2 and the influence of soil physical properties on the moisture content. The study also investigates the minimum number of sampling points required to accurately estimate the field average soil moisture and assesses the temporal stability of the measured soil moisture. Finally, the study simulates soil moisture from SAR imagery by application of apriori roughness parameter.

5.1.1 Soil properties

Soil consists of mineral particles, organic matter, microorganisms, water and air all of which are important in terrestrial plant growth (Hillel 1998). The particles have varying mineral composition, texture, and structure resulting in different aggregates. Water and air fill gaps between the particles. The proportion of water compared to soil particles is the moisture content.

$$m_v = \frac{v_w}{v_t} \quad m_g = \frac{m_w}{m_{ds}} \quad \rho_w = \frac{m_w}{v_w} \quad \rho_b = \frac{m_{ds}}{v_t} \quad \rho_w = \frac{m_v \rho_w}{\rho_b}$$

$$m_v \; is\ volumetric\ soil\ moisture\ content, v_w \; volume\ of\ water, v_t \; is\ total\ volume\ of\ soil\ sample,$$

$$m_g \; is\ gravimetric\ soil\ moisture\ content, m_w \; mass\ of\ water\ in\ the\ sample,$$

$$m_{ds} \; is\ mass\ of\ dry\ sample,$$

$$\rho_w \; is\ density\ of\ water, \rho_b \; is\ the\ bulk\ density\ of\ soil\ sample$$

Topography, climate, soil texture and vegetation influences the soil moisture availability and variability across fields. Ground measurement methods in soil moisture determination include the use of time domain reflectometry (TDR) probes. For vast areas, however, it is tedious and time consuming to carry out field-based measurements. Hence, satellite-based methods are used in conjunction with the TDR measurements. These measurements are mainly used as training and validation data for soil moisture retrieved over large areas using remote sensing imagery. These instruments measure dielectric properties which are related to the soil moisture content (Ventrella D. *et al.* 2008).

Soil roughness is the variance in height of the soil surface. Soil roughness varies depending on agricultural practice and precipitation. The roughness parameters represented by the root mean square height (rmsh), correlation length and autocorrelation function (ACF) describe statistical properties of a randomly rough surface.

$$rmsh = \sqrt{\frac{\sum_{i=1}^{n}\left(h_i - \bar{h}\right)^2}{n-1}}$$

$$h_i\ is\ the\ vertical\ height\ at\ location\ i,$$

$$\bar{h}\ is\ the\ mean\ vertical\ height\ of\ the\ soil\ surface\ for\ n\ samples$$

The rmsh is the standard deviation of height variation with respect to a selected level surface. The ACF describes the statistical independence of two locations on a surface separated by a distance, x. The shape of the autocorrelation function determines the behavior of the backscatter. The Exponential ACF was used in describing surface roughness.

$$\rho(l) = {}^{1}/_{e} = \approx 0.3678$$

$$\rho_{corr}(x) = e^{-x/l}$$

ACF describes the degree of correlation between the height h_i at location i and the height h_{i+x} at a horizontal distance x from i.

$$\rho_{corr} = \frac{\int h_i h_{i+x}\, di}{\int h_i^2\ \ di}$$

The limits of integration extend over the overlapping segment of the height profiles. The correlation length is the horizontal distance/displacement at which the ρ_{corr} drops below $1/e$ (the surface profile is auto correlated at a lag equal to $1/e$). The exponential ACF is usually used for smooth surface and where there are long sampling intervals.

Surface roughness is wavelength dependent. Surfaces appear rough in short wavelengths and appear smoother as the wavelength increases. It is thus essential that normalization is done with the wave number k:

$$k = \frac{2\pi}{\lambda}, giving\ normalised\ rmsh\ (khrms)$$

The possible roughness classification includes random roughness and oriental roughness. Random roughness is due to the organization of soil clods whereas oriented roughness is dependent on tillage tool used (Marzahn *et al.* 2012). Methods used for soil roughness measurements include using laser scanners, photogrammetry methods, pin profilers, photogrammetric method. Roughness is usually a challenge to measure and thus deriving roughness purely from satellite imagery is investigated.

5.1.2 Synthetic Aperture Radar (SAR)

SAR operates in the microwave region of the electromagnetic spectrum. Active sensors emit electromagnetic radiation towards the earth; the radiation interacts with the earth features and the backscattered energy recorded. Active systems are all weather and unaffected by clouds, rain (Azad Hossain & Easso, 2009).

Swaths/tracks are the satellites flight paths. Cross track or range is perpendicular to flight direction whereas along track or azimuth is parallel to the direction of motion of the satellite (Figure 5.3). Time delay separates echoes in the across track direction and Doppler history separates echoes in the along-track dimension. Using time delay and Doppler histories, SAR image resolutions are independent of distance from scene to imaging sensor (Van Zyl and Kim 2011).

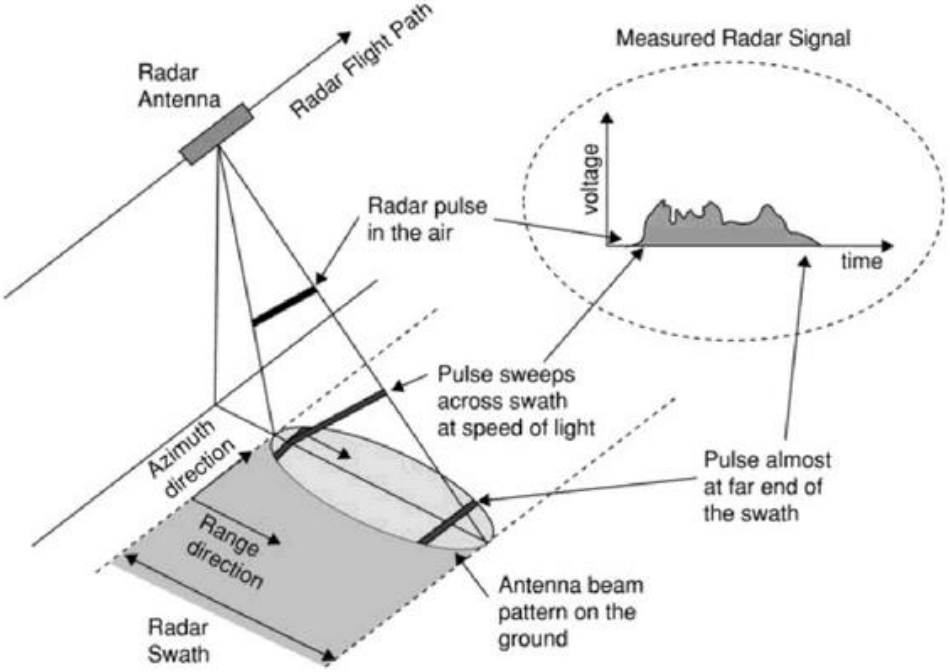


Figure 5.1: SAR satellite geometry (Van Zyl and Kim 2011)

SAR transmits polarized pulses / high bandwidth phase waves across the imaged scene. The incident energy is scattered, reflected, absorbed or it penetrates the surface it is in contact with. The pulses returned at the sensors are processed together to simulate a long aperture with a high resolution in direction of flight. The waves received are a sum of detected amplitude and phase information recorded as backscatter for each pixel forming the radar image. Echoes from targets further away in the along track direction are received at later times.

5.1.2.1 *Radar equation*

Backscatter is the amount of radiation reflected from a target surface measured in a unit area in radar cross section (Ulaby *et al.* 1996). In a monostatic configuration, the transmitter and receiver are located at the same location whereas bistatic configuration, the transmitter and receiver are spatially separated (Skolnik 1990).

A radar image consists of digital numbers at each pixel, determined by the strength of radar signal reflected from the target surface. The interaction of the incident radiation with the earth surface determines variations in brightness of the radar image, which is useful in discriminating features on the ground. The radar equation describes the radar power returned from each transmitted radar pulse. This equation explains the radar propagation, system, and target parameters.

Suppose a radiator at point R produces power outwards, and the power radiated by a point source (P_t) is spread out uniformly over the surface of a sphere (isotropic radiation), then the portion of this power radiated in a given direction is given by

$$P = \frac{P_T}{4\pi R^2}\ Wm^{-2}$$

In radar, an antenna concentrating the power in a given direction is used instead of an isotropic radiator. Hence power transmitted is given by

$$P = \frac{P_T G_T}{4\pi R^2}\ Wm^{-2}$$

Gain is the ratio of power density produced in a preferred direction to the power density produced by an isotropic radiator. The range is the distance from the satellite to the target feature on earth. The radar cross section per unit area ($\sigma\ m^2$) describes the power the target extracts from the incoming radiation. Hence the power radiated back is given by

$$P_R = \frac{P_T G_T \sigma^0}{4\pi R^2} * \frac{1}{4\pi R^2}\ W\ m^{-2}\ due\ to\ isotropic\ propagation$$

For the actual power received, the equation is multiplied by antenna s aperture A_r.

$$A_r = \frac{G_R\,\lambda^2}{4\pi}$$

Power received by the radar system after scattering from the target becomes;

$$P_R = P_T(\sigma^0)\left(\frac{G_T G_R \lambda^2}{(4\pi)^3 R^4}\right)W$$

$$P_R \; power\ received, P_T \; power\ transmitted,$$

$$\lambda \; wavelength, R \; range$$

$$G_T \; gain\ of\ transmitting\ antenna, G_R \; gain\ of\ receiving\ antenna$$

$$\sigma^0 \; radar\ backscattering\ coefficient$$

The total power returned to a satellite platform from a pixel is given by

$$P_R = P_T(\sigma^0)\left(\frac{G_T G_R \lambda^2}{(4\pi)^3 R^4}\right) * r_a * r_g$$

$$r_a \; r_g \; azimuth\ and\ ground\ range\ resolution$$

To convert the backscatter to decibels the formula applied is as shown below,

$$\sigma^0(dB) = 10 \; \log \sigma^0$$

5.1.2.2 System and target parameters

Backscatter coefficient is a measure of reflectivity of a target surface and is influenced by the sensor (frequency, polarization, incidence angle) and target (soil moisture, surface roughness) parameters (Sahebi and Angles 2010, Balenzano 2013, Notarnicola and Solorz 2014). To measure the moisture or roughness from backscatter, it is necessary to separate their effect on the total backscatter recorded (Van Zyl and Kim 2011).

a) Incidence Angle

The sensitivity of backscatter to soil moisture decreases with increasing incident angle (Beaudoin 1990, Coppo 1995). Higher incidence angles are more sensitive to soil roughness. Hence, steep incidence angles are preferred to reduce the contributions from soil roughness in soil moisture retrieval (Baghdadi *et al.* 2002, Holah *et al.* 2005).

b) Polarization

Polarization is the orientation of the electric field of the electromagnetic wave. Linear polarization consists of HH, HV, VV, VH with the first term being the emitted radiation and the second received radiation. The composition of polarization provides information on the form and orientation of scattering on the target surface. HH and VV were found to be the most sensitive in soil moisture retrieval as opposed to cross polarizations (Ulaby and Batlivala 1976). Higher uncertainty is reported in the use of cross polarization than co polarization in soil moisture retrieval (Jiankang Ji et al., 1996).

c) Wavelength

Wavelength affects the penetration depth of the incident waves. Longer wavelengths will result in deeper penetration depth of the incidence radiation thus more information on soil profile (Ulaby *et al.* 1996). Sensitivity to roughness is higher in longer wavelengths (lower frequencies) as compared to shorter wavelengths (higher frequencies) (Mattia *et al.* 1997). Surface parameters sensitive to wavelength hence they are normalized with $k=2\pi/\lambda$ (Kornelsen and Coulibaly 2013).

d) Soil moisture

Soil moisture has a unique characteristic with respect to dielectric properties of water (~80) and dry soil particles (~4). Microwave sensors have the capability of detecting changes in dielectric constant occurring in soil-water mixtures as the moisture changes. Backscatter increases with increasing soil moisture until 35% vol (Das and Paul 2015b). The use of radar backscatter in soil moisture estimation is limited to soils with less than 35% vol as the backscatter becomes insensitive to soil moisture at higher values (Dobson and Ulaby 1981, Chanzy 1993). The soil moisture changes the electric properties of the soil influencing the absorption, transmission, and reflection of microwave energy. Reflectivity and image brightness increases with increased soil moisture content.

e) Soil roughness

The surface of the soil affects the backscatter response of electromagnetic waves. There is higher sensitivity of roughness at high incidence angles (Holah *et al.* 2005). Target roughness varies depending on the wavelength and incidence angle. The Rayleigh criterion determines the smoothness or roughness of a surface by the following equations;

$$h > \frac{\lambda}{8 * \cos\theta} \quad rough\,surface$$

$$h < \frac{\lambda}{8 * \cos\theta} \quad smooth\,surface$$

$$\lambda\ wavelength, h\ mean\ height\ of\ surface\ variations, \theta\ incidence\ angle$$

Surfaces appear smooth with longer wavelengths and with increased incidence angles. Smooth surfaces have specular reflection; thus, a small amount of energy returns to the radar, and hence they appear dark on the radar image.

5.1.2.3 Active microwave sensing of soil moisture retrieval

The greatest challenge in the estimation of soil moisture is determining the portion of backscatter from the water content in the soil (Barrett *et al.* 2009). To retrieve soil moisture from bare surfaces, there is need to eliminate the scattering contribution of surface roughness from the total measured backscatter value.

$$\overset{\circ}{\sigma}_{soil\ moisture} = \overset{\circ}{\sigma}_{total\ backscatter} - \overset{\circ}{\sigma}_{\substack{roughness \\ contribution}}$$

Soil moisture retrieval models from active microwave sensors at local scales include physical based theoretical models, semi-empirical models and empirical models (Nichols 2011). Empirical models are generated from site-specific data i.e. backscatter, soil moisture and surface roughness and are not applicable to different sets of conditions.

Physical models are based on theoretical perspectives on the trend of backscatter in response to soil moisture or roughness e.g. Kirchhoff approximation model. The models assume the apriori knowledge of root mean square height (rmsh), correlation length and backscatter. The Integral Equation Model (IEM) is another example of a theoretical model, which requiring apriori knowledge of radar frequency, polarization, incidence angle, backscatter coefficient, root mean square height, correlation length and the autocorrelation function for computation of the dielectric constant. The model does not account for double or multiple scattering and is suitable for soil moisture estimation in bare surfaces (Barrett *et al.* 2009).

The semi-empirical models combine rules from the theoretical and empirical models e.g. water cloud model, Oh, Shi, Dubois model. The Oh model relates ratios of backscattering coefficients in different polarizations to soil moisture and surface roughness (Oh *et al.* 1992). The advantage of this model is that relatively few field measurements e.g. surface roughness are required, more so in cases where multi polarized imagery is used. The Dubois model works well with co-polarized coefficients as they are less sensitive to system noise, are easier to calibrate and are more accurate as compared to cross-polarized backscattering coefficients (Dubois *et al.* 1995). The Shi model is derived using the L band with incidence angles ranging from 25°-70° is valid for co-polarized backscatter coefficients and is based on regression analysis of simulated backscatter using the single scattering term of the IEM (Petropoulos *et al.* 2015). The water cloud model represents backscatter from the vegetation canopy and underlying soil during the phenological cycle of the plant (Bindlish and Barros 2001). The total backscatter is a sum of contribution from vegetation, soil and two-way attenuation of vegetation layer. Other approaches include statistical analysis techniques where relationships between in situ soil moisture, backscatter from radar, soil and vegetation variables are established and applied in the retrieval of soil moisture e.g. linear regression equations (Nichols 2011). The linear dependence of backscatter and soil moisture over bare soils have been reported to be reliable in soil moisture retrievals (Baghdadi *et al.* 2007, Le Morvan *et al.* 2008, Gorrab, Zribi, Baghdadi, Mougenot, and Chabaane 2015). The linear relationship is of the form;

$$\sigma^{0}{}_{dB} = Am_v + f(R)$$

$$A\ is\ the\ radar\ signal\ sensitivity$$

$$m_v\ is\ the\ soil\ moisture$$

$$f(R)\ is\ a\ function\ of\ roughness$$

5.2 Materials

The study area consists of riparian, middle and fringe fields located in different hydrological zones. The characteristics of these zones vary based on flooding duration in the rainy season (Burghof *et al.* 2017). Riparian fields are closest to the Kilombero River, have the lowest elevation and are completely flooded in the rainy season. The middle fields have non-flooded as well as some patches of flooded regions in the rainy season. Fringe fields are located furthest from the Kilombero River and are rarely flooded.

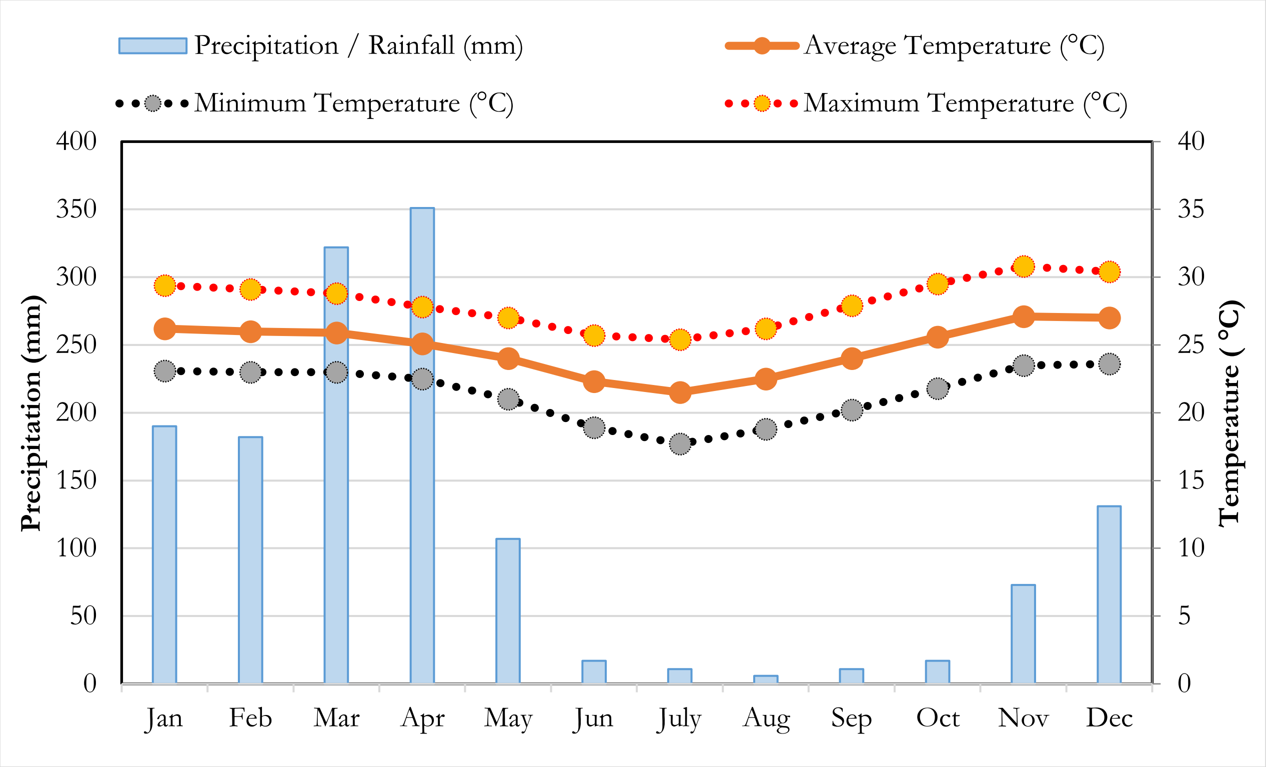


Figure 5.2: Climate data for Ifakara based on weather data collected between 1982 to 2012 (Climate-data.org 2015)

The soils in the study site are mainly sand with gravel and silt sediments (Geological Survey of Tanzania 2007). The main economic activity is rainfed agriculture mainly growing rice, maize, and sugarcane. Additionally, fishing is a common practice in Ifakara.

5.2.1 Data

5.2.1.1 Soil moisture sampling design

Soil samples were collected at three hydrological zones namely the riparian, middle and fringe (Figure 2.2). The experimental campaigns involved the use of Time Domain Reflectometry (TDR) probes in the measurement of 5cm soil moisture. A TDR probe constitutes a high frequency electromagnetic wave propagated along a cable with a probe inserted into the soil. Time elapsed between sending and receiving the pulse is measure by the sensor. Propagation velocity is inversely proportional to the dielectric constant, providing the volumetric soil moisture content (Petropoulos and Albergel 2013). The automated and non-destructive nature of the TDR probe and fast volumetric measurements made it a preferred method for the in situ soil moisture measurements. The soil moisture sensor was the ML3 ThetaProbe and HH2 recording device (Delta-T Devices).

The data was collected during the wet season between March and May 2015 and the beginning of wet season between December 2015 and February 2016. In situ measurements collected 8 times between March and May 2015. Repetitions were made 10 times between December 2015 to February 2016. In each of the fields, 10x10 m grids were established for measurement of soil moisture. The sizes of sample fields were 4,272 to 8,290 m^2 for the riparian, 1,523 to 10,727 m^2 for the middle and 734 to 19,206 m^2 for the fringe fields. In addition, soil samples were collected by core method at 10cm depths for determination of percentages of clay, sand, silt, soil organic carbon (SOC) and bulk density. Sampling locations varied from 9 in the middle and riparian zones and 20 in the fringe zones determined by the shapes of hydrological zones. Each sampling location was designed not more than 750m away from the subsequent sampling spot. Texture was determined using the laser method, bulk density by core method and SOC was determined using the modified Walkley-Black method as described by Daniel (Daniel *et al.* 2017).

5.2.1.2 Soil roughness sampling design

Roughness information was collected a day prior or after the satellite overpass. A 1m pin profiler, with 41 pins at 25mm spacing was used in roughness data collection. A pin profiler consists of equally spaced pins lowered into the surface to give an impression of the surface roughness. The heights of the pins are read against a white background attached to the profiler (Figure 5.2).

One set of roughness profile measurements was taken in fields with areas of less than 2,000m^2, while 2 sets were taken in every other field. A set of roughness profile consisted of two measurements in the NS direction and two in the EW direction such that the total roughness length measured was 2m in the NS and EW direction.

Figure 5.3: Roughness measurements using the pin profiler

5.2.1.3 *Satellite data*

Soil moisture estimation from SAR is accomplished using dual polarized RadarSat-2 and TerraSar-X (Table 5.1). Three TerraSar-X and five RadarSat-2 images were acquired for periods of soil moisture less than 35% volume as the radar signal becomes less sensitive to soil moisture beyond this value (Dobson and Ulaby 1981, Holah *et al.* 2005, Verhoest *et al.* 2008).

Table 5.1: Acquisition parameters for SAR data utilized in soil moisture retrieval

Sensor	TerraSar-X	RadarSat-2
Polarization	HH_VV	HH_HV
Mode	spotlight ascending	fine beam ascending, descending
Wavelength (cm)	3	5.5
Frequency	X	C
Product type	Single Look Slant Range Complex (SSC)	Fine Beam Single Look Complex (SLC)
Incidence angle	22.73 °	30.31° – 36.44 °
Acquisition dates	07-03-2015 29-12-2015 09-01-2016	16-12-2015 19-12-2015 23-12-2015 09-01-2016 12-01-2016

5.3 Methodology

5.3.1 Spatial and temporal variability

5.3.1.1 Statistical methods

Soil moisture averages, standard deviations, and coefficient of variations (cv) were computed for the three hydrological zones for each measurement dates. The cv gave an indication of the variation between the measured data. Analysis of variance to test differences in soil property content was conducted for clay, sand, silt, organic carbon and bulk density at 5% significance level. The coefficient of variation is given by;

$$cv_j = \frac{\sigma(\bar{S}_j)}{\bar{S}_j} = \frac{\sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\left(S_{i,j} - \bar{S}_j\right)^2}}{\bar{S}_j}$$

$$where\ \bar{S}_j\ is\ the\ mean\ of\ each\ sampling\ day\ and\ \sigma(\bar{S}_j)\ the\ standard\ deviation$$

The equations are according to (Gilbert 1987) for determining the number of points required for estimating the mean within a given relative error and confidence interval are as follows;

$$n = \left(t_{1-\frac{\alpha}{2},n_0=1} * \frac{cv}{d_r}\right)^2$$

$$d_r = \frac{|\bar{x} - \mu|}{\mu}$$

$$d_r\ is\ the\ relative\ error,$$

$$t_{1-\frac{\alpha}{2},n_0=1}\ is\ the\ student\ t\ distribution\ at\ confidence\ interval\ 1 - \frac{\alpha}{2}$$

Relative error is allowable deviation of a value from the mean. The use of the coefficient of variation is preferred because it is less variable than the use of standard deviation (Gilbert 1987) in the computation of the number of sampling points required. Initially, the value of t is derived from normal distribution $Z_{1-\frac{\alpha}{2}}$ since no prior value of n is available. The n value obtained is input in the t distribution table to obtain the value of t variate cutting off $100\frac{\alpha}{2}$% of the upper tail with n-1 degrees of freedom. The process is continued iteratively until the $n_i - n_{i-1} \leq \varepsilon$. In this study, ε is 0. The relationship between soil moisture and coefficient of variation aided in quantifying the number of required samples.

Soil moisture mean (θ) and coefficient of variation are related through the exponential equation (Jacobs 2004, Brocca *et al.* 2009, 2012);

$$cv = A\exp(B\theta)$$

$$A\ is\ relative\ variability\ and$$

$$B\ is\ the\ dependence\ of\ variability\ on\ mean\ soil\ moisture\ conditions$$

5.3.1.2 *Relative differences*

The relative difference relates the differences in soil moisture measured over time with the mean of the repetitions. Given m measurement campaigns and n measurement points, the soil moisture measured at location i (given i=1…..n) and at time j (j=1…m) is given by $S_{i,j}$. Following Vachaud et al. (1985), the temporal stability by relative differences between soil moisture at individual locations and the mean of each sample is given by (Vachaud *et al.* 1985);

$$rd_{i,j} = \frac{S_{i,j} - \bar{S}_j}{\bar{S}_j} \; where\; \bar{S}_j = \frac{1}{n}\sum_{i=1}^{n} S_{ij} \;, \bar{S}_j \; is\; the\; field\; mean$$

The mean ($\overline{rd}_{i,j}$) and standard deviation ($\sigma(rd_{i,j})$) of relative difference for every location, i, is given by

$$\overline{rd}_{i,j} = \frac{1}{m}\sum_{j=1}^{m} rd_{i,j} \quad \sigma(rd_{i,j}) = \sqrt{\frac{1}{m-1}\sum_{j=1}^{m}\left(rd_{i,j} - \overline{rd}_{i,j}\right)^2}$$

A point is termed to have a high temporal stability and representative of the measured soil moisture if it has low values of $\overline{rd}_{i,j}$ and $\sigma(rd_{i,j})$. For such a point, the low relative difference indicates a small deviation of the soil moisture at location i, from the mean soil moisture. A high temporal stability realizes a possibility of estimation of soil moisture over larger scales with a limited number of observations (Brocca *et al.* 2010).

The root mean square error (RMSE) of the relative differences identified the fields with the highest temporal stability. The RMSE provided a metric for comparison of stability with other fields under study (Jacobs 2004).

$$RMSE_{i,j} = \left(\left(rd_{i,j}\right)^2 - \left(\sigma(rd_{i,j})\right)^2\right)^{\frac{1}{2}}$$

Geostatistical analysis was performed as explained in Chapter 3.3.3 using the Variogram Estimation and Spatial Prediction Software (Vesper 1.6) developed by Precision Agriculture Centre in Australia (ACPA, Sydney 2006). Soil moisture raster maps were produced by block kriging at 10m pixels. The spatial resolution was selected as it matched the sampling distance as well as the spatial resolution of the SAR images used in soil moisture estimation. The semivariogram model parameters produced were used in computing semivariances of all points. The semivariances are used in the generation of weights that estimate soil moisture values at all locations to produce a raster surface.

5.3.2 Soil moisture retrieval from SAR backscatter

SAR images have side-looking geometry (Figure 5.3) resulting in slant range images. The Single Look Complex RadarSat-2 and TerraSar-X images were processed in SNAP Toolbox (European Space Agency 2013). Radiometric calibration removes variations ensuring that the intensity values presented as backscatter values are true representations of target features on the ground (Moreira *et al.* 2013). Multi looking was performed to obtain square pixels. Geometric calibration using the SRTM 3-second digital elevation model later followed to correct for topographic variations. Suppression of speckle was performed with the Lee refined speckle filter. The backscatter values were converted from linear to decibels (dB) using the equation;

$$\sigma^0{}_{dB} = 10 * \log_{10} \sigma_{linear}$$

Due to the difference in incident angle for the RadarSat-2 images, normalization was carried out with the reference angle being 33° (van der Velde *et al.* 2008).

$$\sigma^0(33) = \frac{\sigma^0}{\cos^2\theta}$$

$$where\ \sigma^0\ is\ the\ backscatter$$

$$\theta\ is\ the\ incidence\ angle\ of\ the\ RadarSAT-2\ images$$

(Baghdadi *et al.* 2016) proposed a semi-empirical model for soil moisture derivation over bare surfaces based on the Dubois model. A large dataset from worldwide study sites with sensors in X, C, and L with incidence angles between 20 and 45 generated the model suitable for application in all surface conditions.

$$\sigma^0{}_{HH} = 10^{-1.287}(\cos\theta)^{1.227}10^{0.009\cot(\theta)\,mv}(kHrms)^{0.86\sin(\theta)}$$

$$\sigma^0{}_{VV} = 10^{-1.138}(\cos\theta)^{1.528}10^{0.008\cot(\theta)\,mv}(kHrms)^{0.71\sin(\theta)}$$

$$\sigma^0{}_{HV} = 10^{-2.325}(\cos\theta)^{-0.01}10^{0.011\cot(\theta)\,mv}(kHrms)^{0.44\sin(\theta)}$$

Inverting the Baghdadi equations;

$$A = 10^{-1.287}(\cos\theta)^{1.227}$$

$$B = 10^{-1.138}(\cos\theta)^{1.528}$$

$$C = 10^{-2.325}(\cos\theta)^{-0.01}$$

Inverting to obtain soil moisture (m_v);

$$m_v = \frac{\log(\sigma_{HH}/A * kHrms)}{0.009 * \cot\theta}$$

$$m_v = \frac{\log(\sigma_{VV}/B * kHrms)}{0.008 * \cot\theta}$$

$$m_v = \frac{\log(\sigma_{HV}/C * kHrms)}{0.011 * \cot\theta}$$

The root mean square height (rmsh) was computed for all pin profiler measurements with Matlab R2012b. The normalized rmsh (kHrms) was computed as;

$$kHrms = \frac{2\pi}{\lambda} * rmsh$$

For satellites with dual polarization, the soil moisture should be the same for two polarizations. Thus, solving the soil moisture equations as per the polarizations of the satellites simultaneously and inverting to obtain normalized root mean square height (kHrms);

$$kHrms_{sim} = \frac{\sigma_{VV} * A * 10^{0.009*\cot\theta}}{\sigma_{HH} * B * 10^{0.008*\cot\theta}} \; for\; TerraSAR - X$$

$$kHrms_{sim} = \frac{\sigma_{HV} * A * 10^{0.009*\cot\theta}}{\sigma_{HH} * C * 10^{0.011*\cot\theta}} \; for\; RadarSAT - 2$$

The linear dependence of radar on soil moisture for values of soil moisture between 5 and 35% vol has been reported by (Baghdadi *et al.* 2007, Baghdadi, Camus, *et al.* 2011). The equation used to retrieve soil moisture was of the form;

$$y = a * mv + b * kHrms\_sim + c$$

$$where\; a, b, c\; are\; coefficients$$

To assess the extent to which the model fits the data, the accuracy of the model, Residual standard error, r squared, and the F statistic were computed. The residual standard error is the average amount that the response variable (soil moisture), will deviate from the true regression line and thus measures the quality of the regression model. R squared measures how well the model fits the actual data. It provides a measure of the amount of variance in the response (soil moisture) variable that can be explained by the predictor variables (rhms and backscatter) (James *et al.* 2013). The adjusted r squared considers the number of variables in the model. The F-statistic provides an indication of the presence of a relationship between the predictor and response variables, with a higher value indicating the presence of a relationship.

To assess how well the model can predict soil moisture values, k-fold cross-validation was implemented. The measurements are divided into k groups/folds of equal size. The first fold, the validation set, and the model fit to k-1folds. The mean square error (MSE_1) is computed for this first fold. The procedure is repeated k times and each time an MSE is computed for the different folds (James *et al.* 2013). This procedure will result in MSE values equal k. The k fold cross validation is acquired by averaging the MSE values.

$$cross\ validation_{(k)} = \frac{1}{k}\sum_{i=1}^{k} MSE_i$$

5.4 Results and Discussion

5.4.1 Spatial and temporal variability

5.4.1.1 Statistical methods

The average soil moisture over the three hydrological zones are as shown in Figure 5.4, Table 5.2. In periods of excessive rainfall from mid of April 2015 and end of January 2016, the Riparian zone holds a higher capacity of water explained by the high percentage of clay content and organic carbon. The riparian fields have a high standard deviation in high moisture content.

Table 5.2: Spatial statistics; mean, standard deviation (sd), coefficient of variation (cv) and number of sampling points (num) of soil moisture

	Riparian				Middle				Fringe			
date	Mean (vol %)	sd (vol %)	cv (%)	num	Mean (vol %)	sd (vol %)	cv (%)	num	Mean (vol %)	sd (vol %)	cv (%)	num
07-03-15	18.0	6.7	37.2	525	20.9	13.9	66.3	219	19.5	11.2	57.5	337
31-03-15	37.0	18.2	49.0	436	42.8	14.1	33.0	221	40.5	12.1	29.81	336
20-04-15	48.5	16.7	34.4	177	46.8	11.2	23.9	211	45.2	8.7	19.3	299
23-04-15	54.2	15.2	28.1	146	51.1	9.4	18.4	177	48.6	7.9	16.3	342
27-04-15					54.0	8.2	15.3	181	46.9	8.7	18.4	269
01-05-15					54.3	8.1	15.0	183	48.3	7.0	14.5	250
14-05-15					52.0	8.1	15.5	178	49.3	7.5	15.2	266
17-05-15					49.9	9.6	19.3	191	48.5	8.7	17.9	295
16-12-15	8.6	3.6	41.8	477	6.5	3.1	47.7	250	6.7	3.5	52.5	419
19-12-15	6.2	3.2	50.5	244	8.6	3.3	37.9	125	15.9	6.9	43.5	226
23-12-15	21.3	6.8	31.9	248	23.5	7.7	32.7	137	25.7	8.9	34.7	206
29-12-15	9.8	3.4	34.6	246	7.3	3.4	47.4	122	11.8	5.4	45.7	194
09-01-16	22.3	6.4	28.7	254	20.8	10.2	49.1	146	23.8	8.3	35.1	238
12-01-16	27.7	7.0	25.3	253	28.9	10.9	37.6	148	31.5	8.2	25.9	242
20-01-16	39.5	8.3	21.1	250	44.6	10.7	24.1	117	42.1	11.3	26.9	189
31-01-16	44.3	12.0	27.1	103	46.0	10.9	23.8	104	43.3	8.7	20.0	164
06-02-16	48.1	13.8	28.7	111	44.2	11.5	26.0	114	45.6	7.6	16.6	130
11-02-16					49.8	8.6	17.4	113	46.9	7.2	15.4	135

Soil moisture was analyzed for variation in time and space. March to May marks the rainy season in the study area. Due to increased rainfall, the soil moisture levels increase to a maximum till May for the middle and fringe fields, after which the soil moisture recorded reduces for that first season. The riparian fields were flooded from the 27th April and thus no measurements were taken until the end of the season. The second collection of soil moisture was from December 2015 to February 2016 (Figure 5.4). The rains begin in December. The slight increase in soil moisture on 23rd December can be attributed to a precipitation experienced on 21st and 22nd. Due to the highly dynamic nature of soil moisture, the levels were observed to subside towards the end of the year. Frequent precipitation from January 2016 resulted in the rise of soil moisture exhibiting the strong dependency of moisture to precipitation.

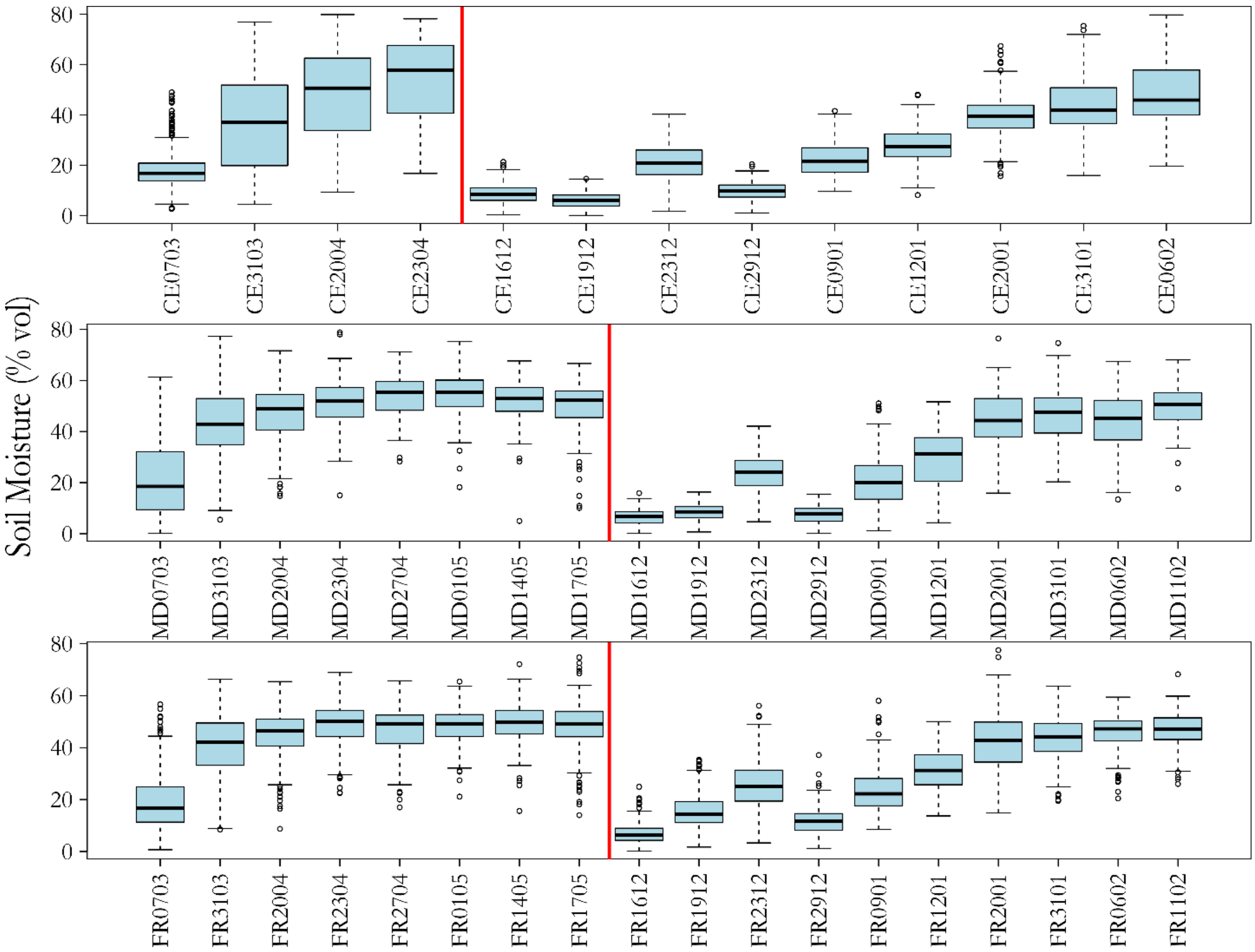
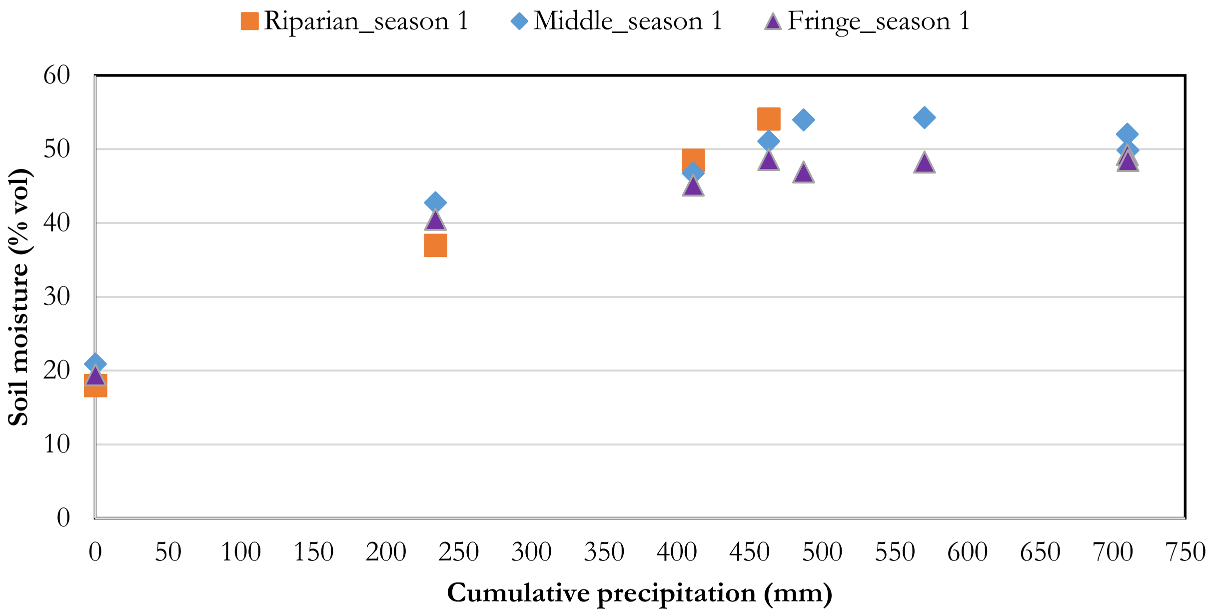


Figure 5.4: Mean temporal variation of soil moisture for riparian, middle and fringe fields from March to May 2015 and December 2015 to February 2016. The upper and lower whiskers are representative of standard deviation. CE are the riparian fields, MD the middle fields and FR the fringe fields. The whiskers above and below show the maximum and minimum values, the segment in the box shows the median value, the circles are probable outliers.

Figure 5.5 shows the relationship between the mean soil moisture observations and cumulative precipitation. A strong dependency of soil moisture on precipitation is noted for both study periods (March to May 2015 and December 2015 to February 2016).



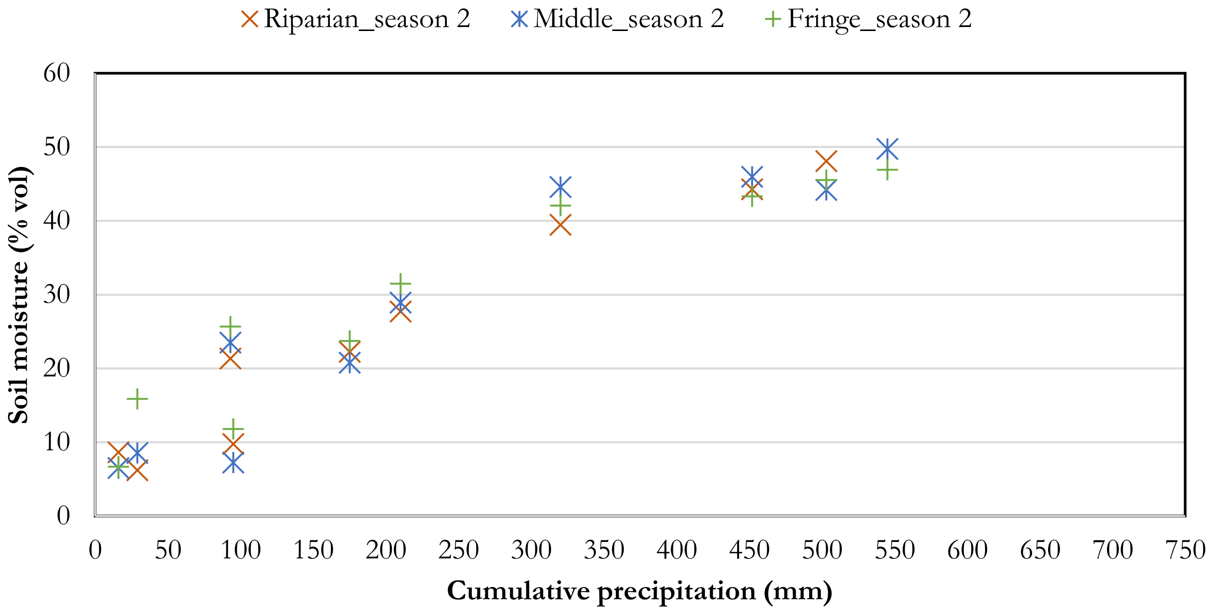


Figure 5.5: Mean soil moisture versus cumulative rainfall for March to May 2015 and December 2015 to February 2016

Table 5.3: Spatial variability of selected soil physical properties along the hydrologic gradient

	n	Clay%	Silt%	Sand%	Bulk Density g/cm^3	Organic Carbon g/kg	Texture
Riparian	16	35.0 ± 12.2 $^{a, b}$	56.2 ± 10.1	8.8 ± 18.2 ^{c}	1.03 ± 0.213 ^{d}	1.88 ± 0.742	silty clay loam
Middle	12	18.8 ± 5.3 ^{b}	58.3 ± 11.0	22.9 ± 14.9	1.14 ± 0.258	1.36 ± 0.609	silt loam
Fringe	8	13.9 ± 4.0 ^{a}	57.2 ± 7.3	28.9 ± 10.4 ^{c}	1.27 ± 0.136 ^{d}	1.42 ± 0.592	silt loam
p-value		3.5e-06 ***	0.86 ns	0.011 *	0.053 ns	0.11 ns	

Same letters in a column indicate significant difference at $p<0.05$

n is the number of mean soil moisture values included in the data analysis

ns non-significant difference at 0.05 probability level

* Significance at 0.05 probability level

*** Significance at 0.001 probability level

The clay percentage in the riparian zone is high and significantly different from the fringe and middle fields (Table 5.3, Figure 5.6). Organic carbon is highest in the riparian zone though not at a significant level as compared to middle and fringe. High organic content increases porosity and subsequently infiltration (Ziadat *et al.* 2010). Clay increases the retention capacity of the soils in the riparian zone (De Lannoy *et al.* 2006). There is a significant difference between the bulk density of riparian and fringe zones (Table 5.3, Figure 5.6). The riparian zone has a silty clay loam while the middle and fringe has silt loam texture according to the USGS soil classification.

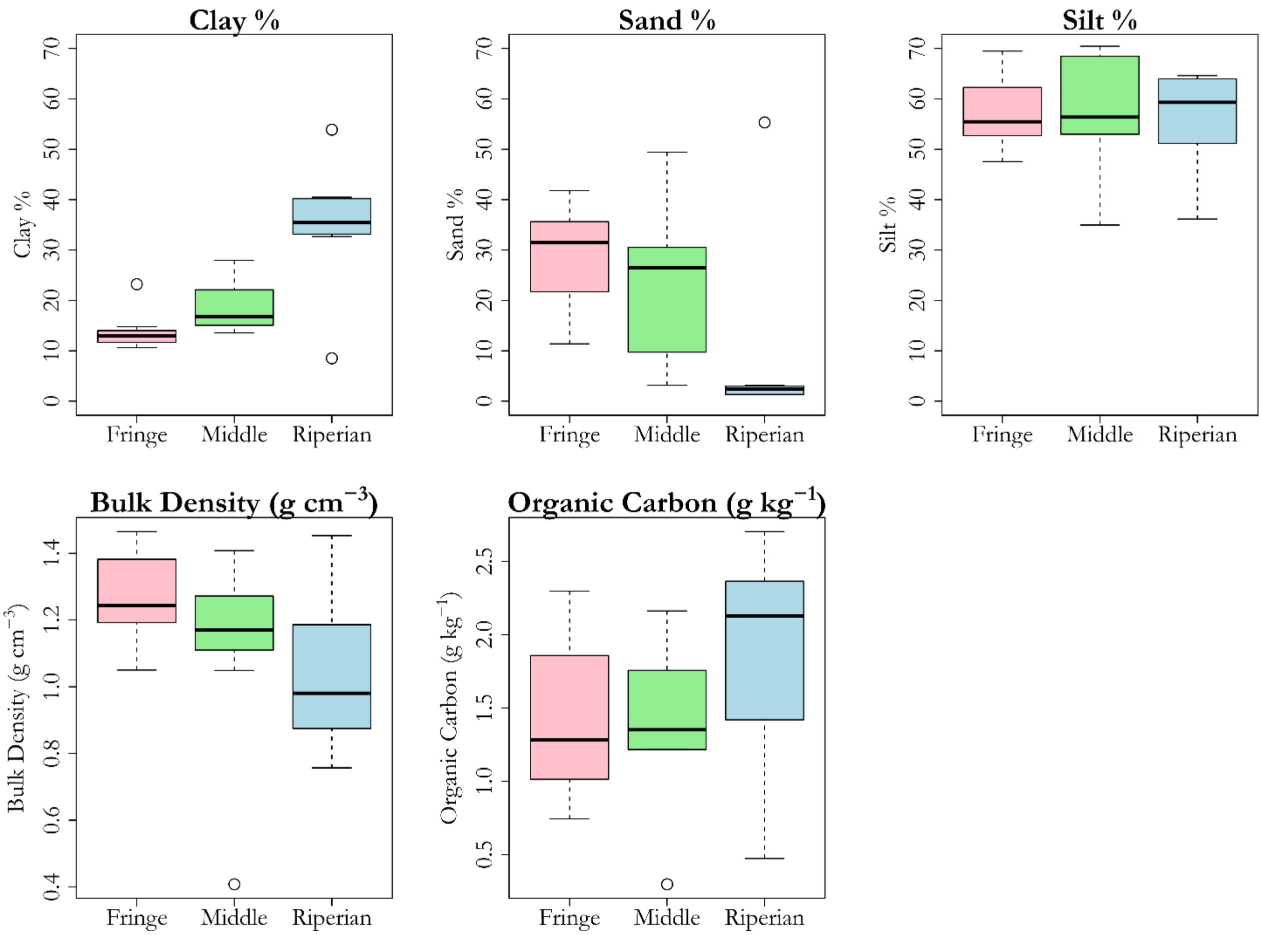


Figure 5.6: Whisker plots of soil physical properties for the Riparian, Middle and Fringe fields

5.4.1.2 Relative difference

Evaluation of temporal stability was performed by use of relative difference technique indicating how much moisture deviation from the mean over the data collection dates. A relative difference greater than 0 overestimates while less than 0 underestimates the averages despite the measurement times (Vachaud *et al.* 1985). The temporal mean relative difference was used in ranking. Figure 5.7, Table 5.4 shows the relative differences for the riparian, middle and fringe sites. The fields within each of the sites are shown in Figure 2.2.

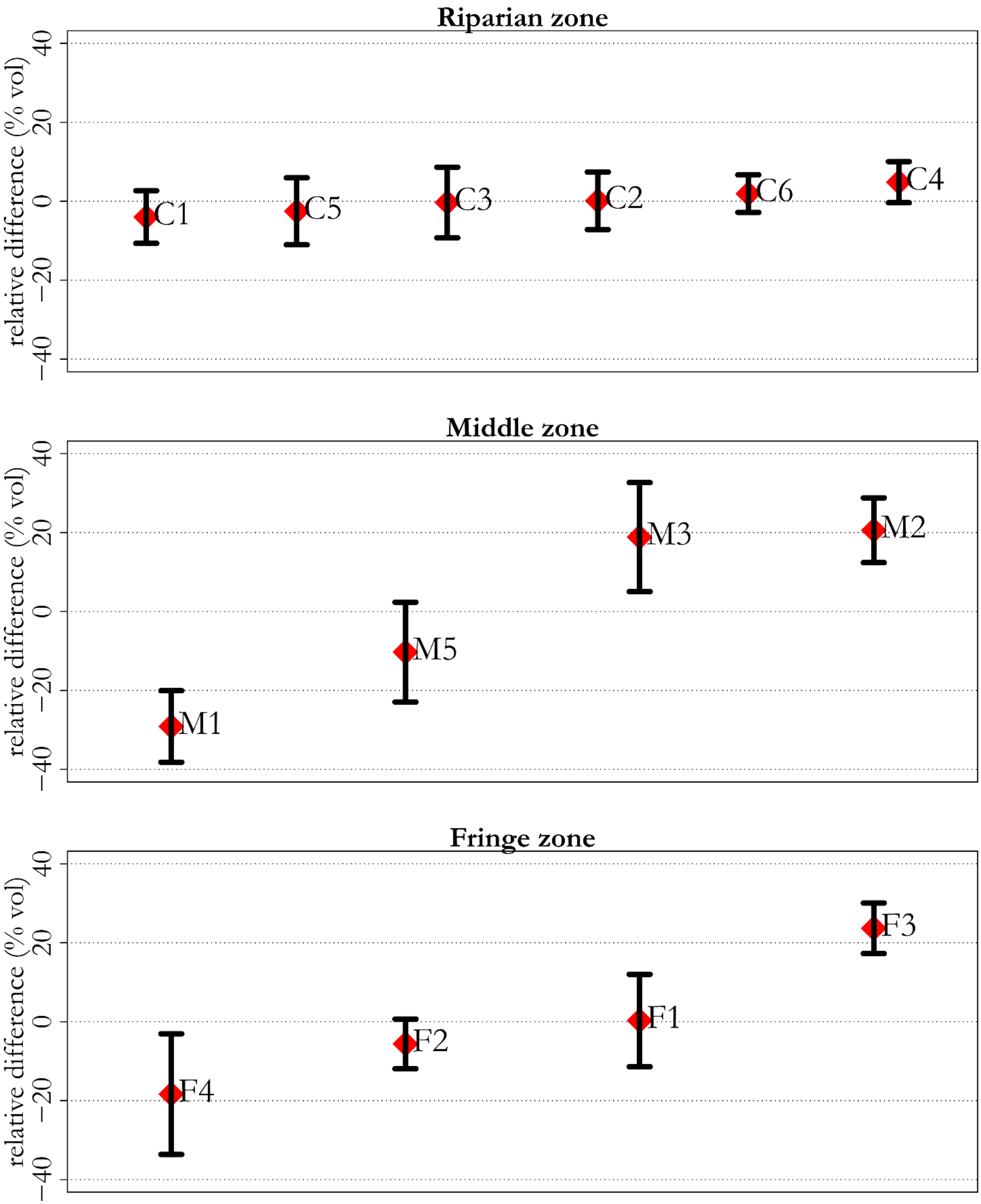


Figure 5.7: Relative difference of sample fields in the riparian, middle and fringe zones

Table 5.4: Relative difference, standard deviation of relative differences and RMSE

Field	Relative Difference%	Standard Deviation of Relative Difference%	Range	RMSE	Area (m^2)
C1	-4.00	6.67	13.34	0.08	6263.4
C2	0.11	7.29	14.58	0.07	3995.8
C3	-0.32	8.95	17.90	0.09	8289.8
C4	4.83	5.21	10.42	0.07	3856.6
C5	-2.54	8.49	16.98	0.09	6314.7
C6	1.93	4.76	9.52	0.05	4272.8
F1	0.30	11.70	23.40	0.12	11521.8
F2	-5.62	6.30	12.60	0.08	1126.7
F3	23.67	6.40	12.80	0.25	19204.3
F4	-18.35	15.28	30.56	0.23	732.8
M1	-29.13	9.07	18.14	0.30	1705.4
M2	20.58	8.20	16.40	0.22	10726.5
M3	18.85	13.81	27.62	0.24	1524.0
M5	-10.30	12.64	25.28	0.16	3210.1

From the results (Table 5.4), relative difference ranged between -4% to +4.83% vol for the riparian fields whereas the standard deviation ranged from 4.76 to 8.95%. Middle fields recorded quite high relative differences ranging from -29.13% to 20.58% with a standard deviation of 9.07 to 13.81% vol. The fringe fields exhibited a mean relative ranging from -18.35% to 0.30% with a standard deviation of 6.4 to 15.28% vol. Low standard deviation of the relative difference was observed for the riparian fields. The highest standard deviation was observed for all the middle fields.

The fields with the lowest relative difference and standard deviation of the relative difference are indicative of temporally stable fields. The riparian fields have the least deviations from mean throughout the measurement period. The representative sampling fields in the three zones were C6, M5, and F2 with lowest RMSE of 0.05, 0.16 and 0.08 respectively. Field C6 has a relative difference of 1.93±4.76% vol (range of 9.52) M5 has -10.30±12.64 (range of 25.3), while F2 has -5.62±6.30 (range of 12.6). Therefore, the representative fields in the riparian, middle and fringe hydrological zones were taken to be C6, M5, and F2. M3, M2, F3 were too wet to be representative fields whereas F4 and M1 were too dry. The high relative differences, particularly for M3, M2, F3, M1, and F4, could have been caused by random variation in terrain in addition to the actual variation of soil moisture with time (De Lannoy *et al.* 2007). Additionally, cropping patterns, variation in farm management, utilization of hand tools in land preparation could be causal factors of the temporal instability of mean field soil moisture.

The sample fields analyzed had varying sizes with some as small as 732 m^2 and the large ones as big as 19,000 m^2. The riparian fields, which indicated highest stability ranged from 3,800 to ~9,000 m^2. The average sizes of these fields as compared to other sample fields, the high clay content resulting in prolonged water retention and proximity to the Kilombero River, prompting moisture recharge may be linked to the low variability for the riparian fields.

The mean soil moisture values for the riparian zone, the middle and fringe are shown in Figure 5.8. This emphasizes the existence of variations of soil moisture explained by varying soil properties in the three hydrological sites.

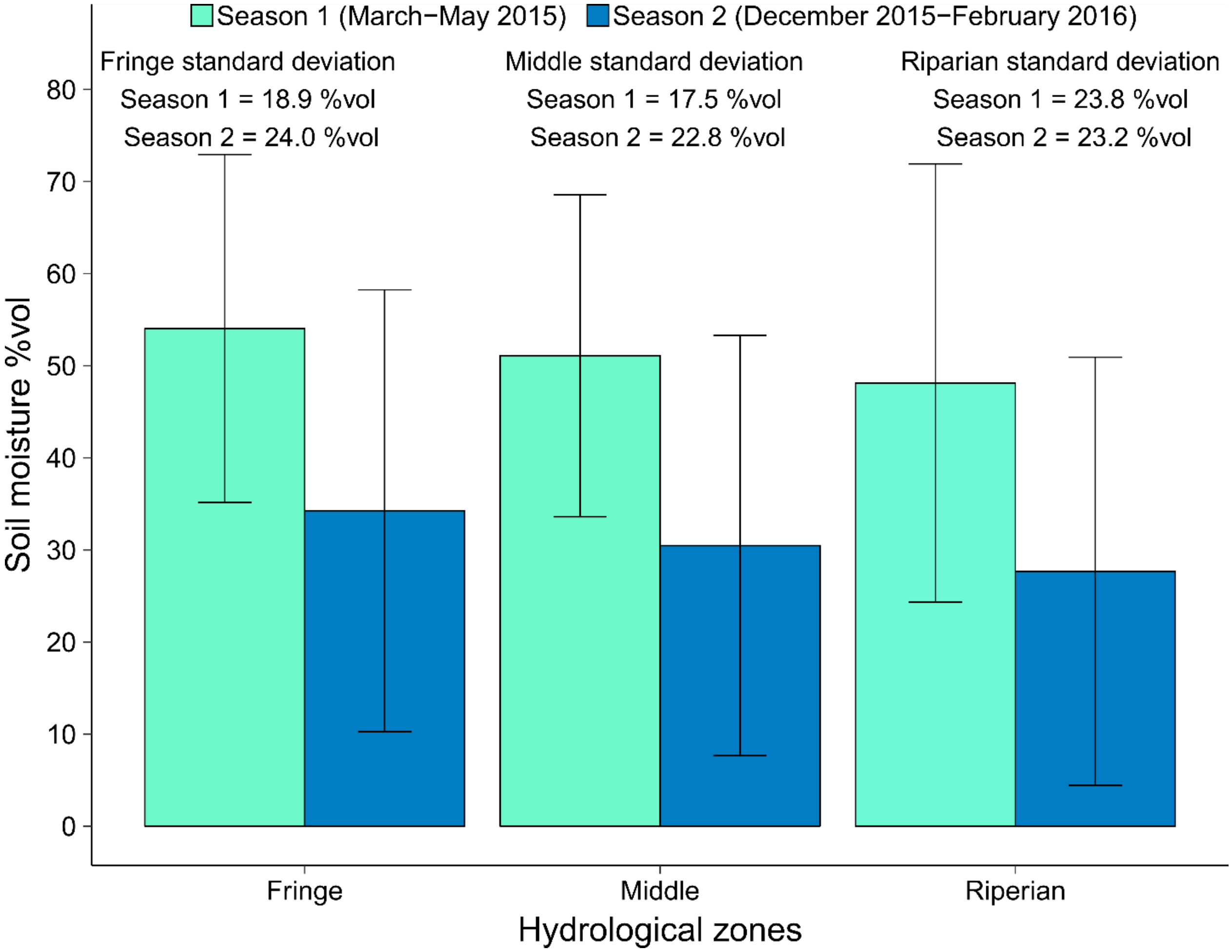


Figure 5.8: Average mean soil moisture for the three hydrological zones for March to May 2015 and December 2015 to February 2016

The cv from the mean soil moisture indicates that there is high variability in the dry season (season 2) falling between December and February. In the wetter season, the Fringe and middle exhibit low variability with the riparian having moderate variation (Table 5.5).

Table 5.5: Coefficient of variation of mean soil moisture for Riparian, Middle and Fringe fields

	Season1			Season2		
Zone	Riparian	Middle	Fringe	Riparian	Middle	Fringe
Means	48.1	51.1	54.0	27. 7	30.5	34.2
Standard Deviation	23.8	17.5	18.9	23.2	22.8	24.0
Coefficient of Variation	0.49	0.34	0.34	0.84	0.74	0.70

The cv values greater than 0.35 indicate high variability, which is observed for all the average measurements in the time between December and January (Season 2_dry season) as well as the time between March and May (Season 1_wet season) (Table 5.5). The wet season exhibits moderate variability for the middle and fringe fields, which has lower soil moisture values as compared to the riparian zone.

The standard deviation increases with increasing soil moisture (Figure 5.9). This is in agreement with a study by (Zribi and Dechambre 2003) where dry fields had fewer variations. Converse results, where fields with high moisture content had higher variations, were reported by (Hupet and Vanclooster 2002).

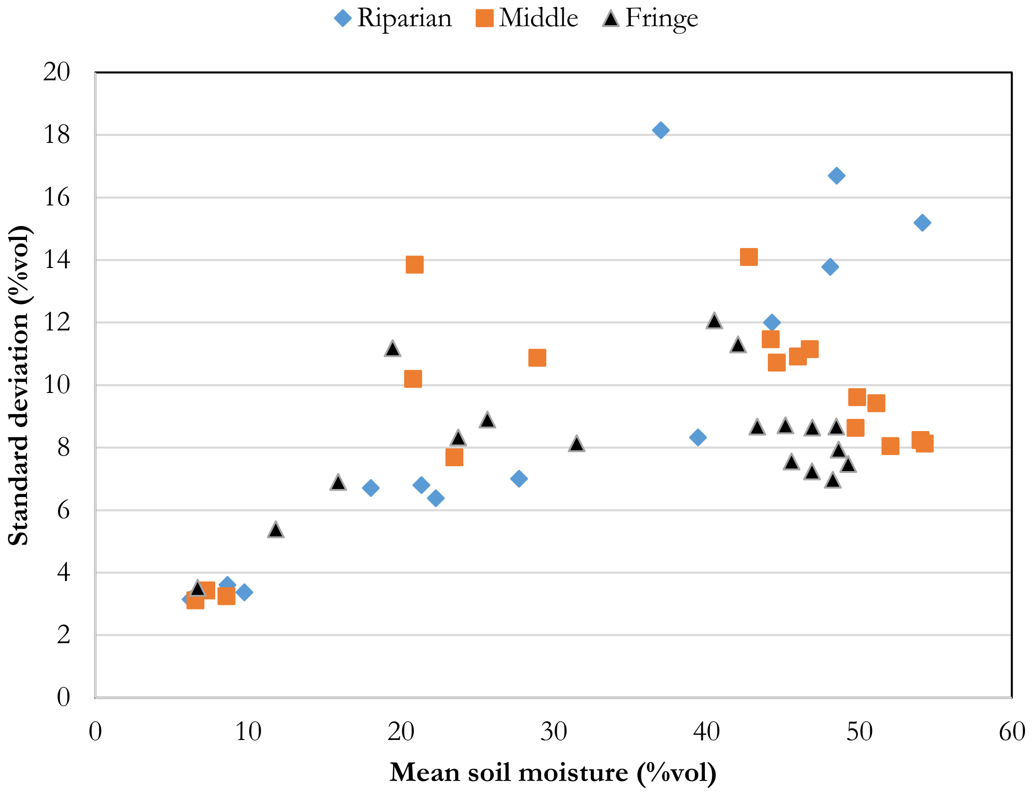


Figure 5.9: Soil moisture vs standard deviation for the riparian, middle and fringe fields

The coefficient of variation shows the relative magnitude of deviation as a function of the mean ensuring the possibility of comparison among different parameters (Kimble *et al.* 2000). CV =standard dev/mean. The coefficient of variation (cv) for 70% of the values was < than 35% concluding that was moderate variation in the measured soil moisture values (Figure 5.10). The maximum cv was reported for data measured on 07-03-15 in the middle (0.66) and fringe (0.58) fields. At high soil moisture values, the cv is more compact and concentrated at values between 10-40% whereas, for low soil moisture values, the range is wider from 20-70%. Generally, there is a decrease in coefficient of variation as soil moisture increases. This is consistent with (Bell *et al.* 1980, Famiglietti *et al.* 1999, Brocca *et al.* 2007, 2010) whose campaigns were also field based.

From the graphs (Figure 5.10), there is a high correlation between the cv of middle and fringe fields and their means. The riparian zone soil moisture was poorly correlated to the cv. This could be attributed to the variation in texture properties and fewer mean field measurements used in deriving the equations for the riparian zone.

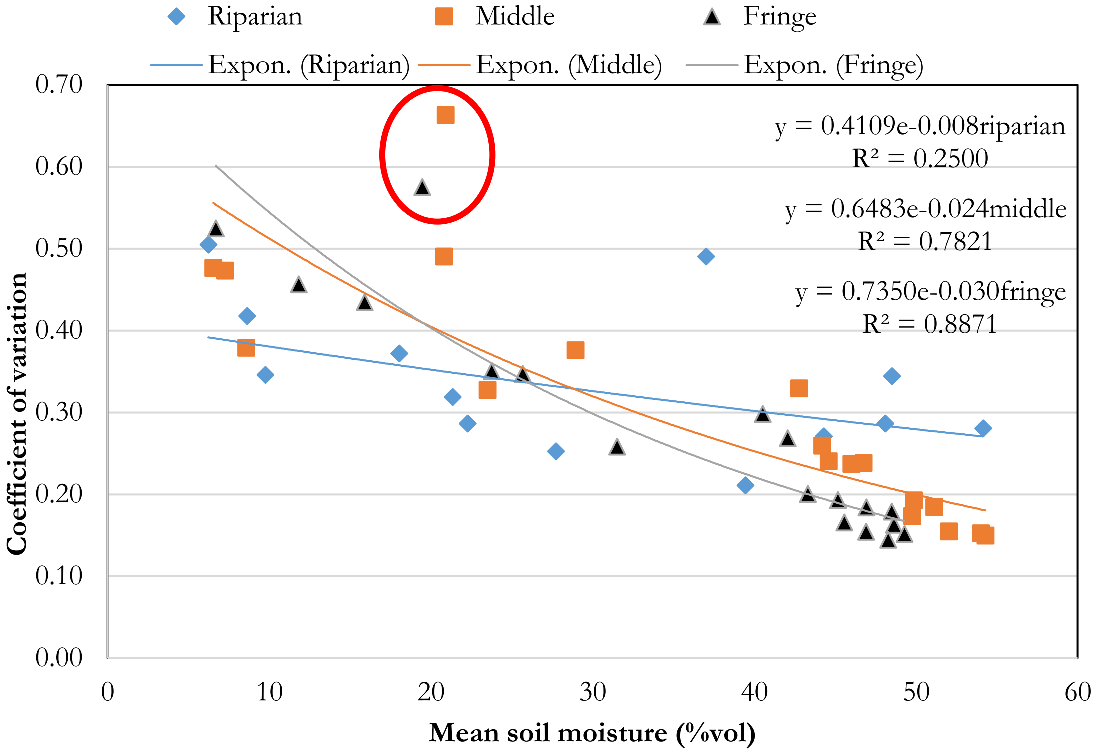


Figure 5.10: Soil moisture vs coefficient of variation for the Riparian, Middle and Fringe fields. The exponential equation showing the relationship between mean soil moisture and coefficient of variation is indicated. The circled points show the maximum coefficient of variation for data measured on 07-03-15 in the middle and fringe fields.

An exponential equation linking the mean soil moisture and coefficient of variation was established. The coefficients of the exponential equations are as shown in Table 5.6.

Table 5.6: Exponential relationship between mean soil moisture and coefficient of variation

Hydrologic zone	A	B	R^2
Riparian	0.411	-0.008	0.25
Middle	0.648	-0.024	0.78
Fringe	0.735	-0.030	0.89

A and B are relative variability and dependence of variation on the soil moisture means (Jacobs 2004). From the coefficients, it is evident that there was low relative variability for riparian fields and high relative variability for the fringe fields in dry conditions. The fringe fields had a relatively rapid decrease in variability (lowest dependence of variation on soil moisture) as the moisture increased. The equations linking cv and soil moisture are important in the determination of the number of moisture measurements required to give a mean value within a pre-specified margin of error and confidence interval. No further analysis concerning the number of points required was performed for the riparian fields due to the low correlation of the soil moisture to the coefficient of variation.

The number of points required at a relative error of 10 and 20% for the middle and fringe at a confidence level of 95% is shown in Figure 5.11. The drier periods (5 to 10% vol) need a higher number of measurements for areal coverage of ~17,000 and ~32,000 m² for the middle and fringe fields respectively.

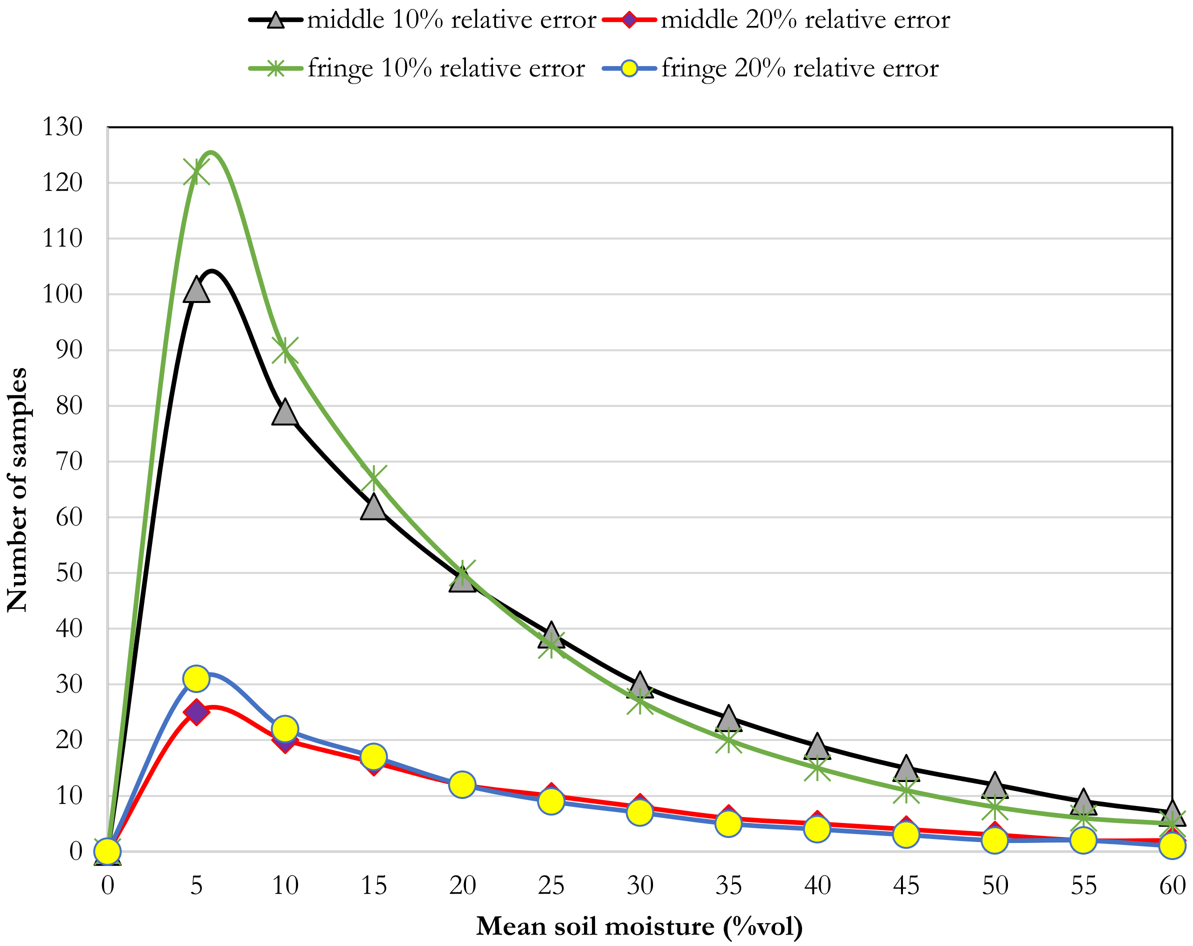
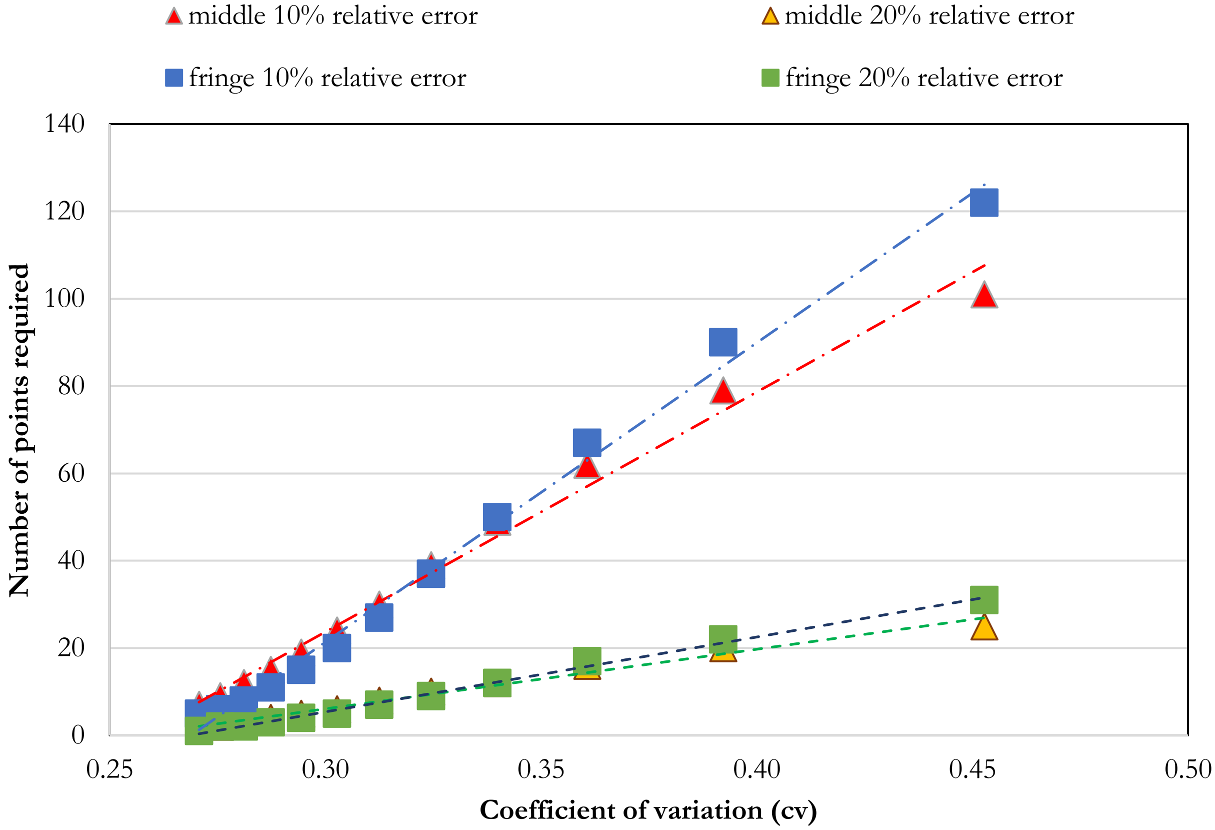


Figure 5.11: Mean soil moisture vs number of points required to obtain soil moisture with relative errors of 10 and 20% the middle and fringe fields.

A relative error of 10% (ranging from 105 to 125 for the 10% relative error) requires more points than 20% (from 25 to 30) in dry soil moisture conditions. The fringe fields have a higher number of points explained by the high values of the coefficients A and B in the exponential equation. The higher variation of the fringe fields indicates that more points are required to achieve the field averages within the given error margin. The sampling scale was 10m*10m, however from the results above, 105 and 125 points would have been sufficient for the middle and fringe fields to give mean field averages at 95% confidence level and a 10% relative error in dry soil moisture conditions (5%). As soil moisture increases, the number of points required reduces, falling to between 27-30 in soil moisture conditions of 30% vol.

The relationship between cv and number of points required is shown in Figure 5.12 where the regression equations were linear similar to a study by (Wang *et al.* 2008). The number of points are required to obtain average soil moisture in the middle and fringe fields at 95% confidence level with a relative error of 10% (md_0.1 and fr_0.1) and 20% (md_0.2 and fr_0.2) respectively at a sampling scale of 10m. The equations give a general idea of the number of samples that would be required for subsequent soil moisture sampling schemes of the study area.



$$y = 549.17(md\_0.1\ cv) - 141.64$$
$$R^2 = 0.9865$$

$$y = 136.52(md\_0.2\ cv) - 34.89$$
$$R^2 = 0.9838$$

$$y = 685.55(fr\_0.1\ cv) - 184.32$$
$$R^2 = 0.9930$$

$$y = 171.41(fr\_0.2\ cv) - 46.046$$
$$R^2 = 0.9947$$

Figure 5.12: Coefficient of variation vs number of measurement points required to acquire average soil moisture at 10 and 20% relative error at 95% confidence interval for the fringe and middle fields

5.4.1.3 *Geostatistical methods*

Quantile-Quantile (QQ) plots indicated the data was normally distributed, implying that the spatial pattern of the moisture was random (Brocca *et al.* 2007). The normal distribution requirement for kriging interpolation is met. The QQ plots for the complete soil moisture dataset is shown in the appendix (B.1).

Table 5.7: P-values from Shapiro Wilk Normality test. Values greater than the 0.05 indicate a normal distribution.

	Riparian	Middle	Fringe
	p-value	p-value	p-value
07-03-15	0.54	0.05	0.14
31-03-15	0.67	0.14	0.55
20-04-15	0.08	0.66	0.07
23-04-15	0.05	0.46	0.30
27-04-15		0.21	0.07
01-05-15		0.24	0.06
14-05-15		0.27	0.08
17-05-15		0.07	0.51
16-12-15	0.10	0.16	0.11
19-12-15	0.15	0.67	0.51
23-12-15	0.39	0.05	0.86
29-12-15	0.98	0.21	0.27
09-01-16	0.07	0.17	0.17
12-01-16	0.90	0.35	0.10
20-01-16	0.07	0.97	0.27
31-01-16	0.65	0.24	0.44
06-02-16	0.09	0.54	0.12
11-02-16		0.17	0.46

The ratio of nugget and sill indicated that weak spatial dependence occurred in relatively dry periods (Appendix B3). In the riparian fields, the range is within 0-40m in the drier periods whereas it ranges from 20-70m in the wetter season. In the middle fields, the range was between 80-140m in the dry season and 0-50 in the wet periods. In the fringe, the range was between 0-70m in the dry season and 0-40m in the wet season (Figure 5.16). This indicated that for the fringe and middle fields, the range was higher in the dry season with the converse observed for the riparian fields. The distance over which soil measurements are correlated at the fringe and middle increases in the dry season. Correlation is only at smaller distances in the riparian fields in the dry season. Sample spacing can be increased from 10 m to a value less than the range distance since the measurements below the range are correlated. Therefore, the sample spacing for the riparian fields will be lower (and hence more in number) than that of middle and fringe fields. The relationship between the deviation from the mean and the range indicate that low ranges in the dry season have low standard deviations in dry soils whereas high ranges corresponding to periods of high soil moisture have a higher standard deviation in wet soils. In the middle and fringe fields, the range and standard deviation seem to have an inverse relationship (Figure 5.13). At high soil moisture, the range is lower and the standard deviation is high. The distance of correlation decreases with increasing soil moisture implying that more sampling points are required in wet season which will subsequently reduce the deviation from the mean. The difference in the pattern of the distance of correlation, soil moisture, and standard deviation can be explained by soil texture properties (Vachaud *et al.* 1985). Following the assessment of the variograms and variogram parameters (Figure 5.14, Appendix B2, Table 5.8, Appendix B3), kriged images were generated with Vesper software. Examples are shown in Figure 5.15and Appendix B7-B9.

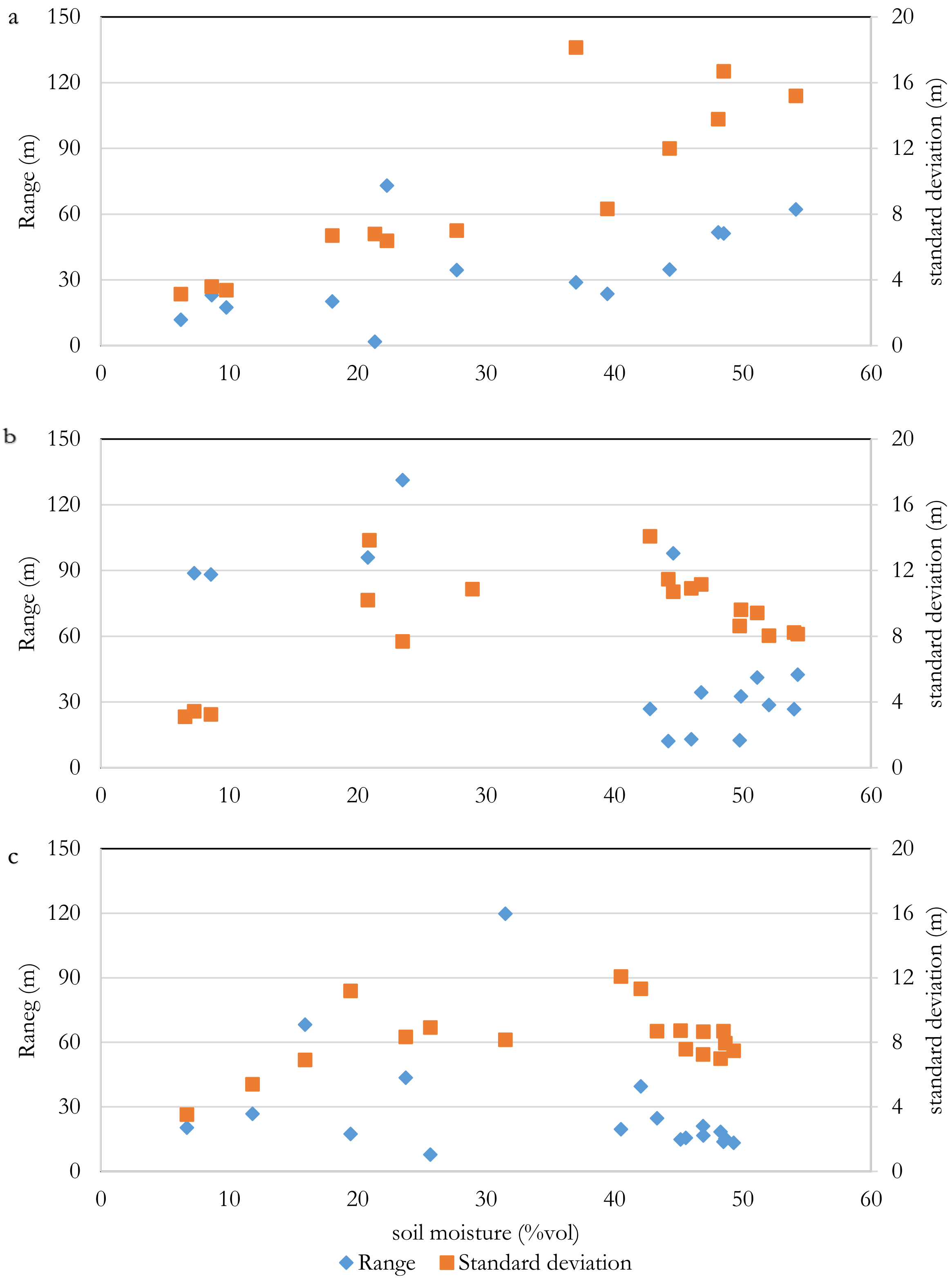


Figure 5.13: Soil moisture field means vs semi variogram range and standard deviation for a) riparian, b) middle and c) fringe zones

Spatial variogram parameters used in the generation of kriged soil moisture maps, the graphical representation of the variograms and a sample of kriged images are as shown. The complete list of variogram parameters is in the Appendix B3.

Table 5.8: Spatial semivariogram parameters. Strong spatial dependence is exhibited for values less than 25%, moderate for values between 25-50% and weak dependence for values greater than 75%.

fringe	model	nugget variance (c0)	partial sill (c1)	sill (c0+c1)	Spatial dependence (%)	range	rmse
09-01-16	Exponential	23.9	66.0	89.9	26.6	43.5	5.7
12-01-16	Gaussian	134.9	133.5	268.4	50.3	119.7	15.7
20-01-16	Exponential	155.7	659.2	814.9	19.1	39.5	24.3
31-01-16	Exponential	166.0	590.0	756.0	22.0	24.6	26.9

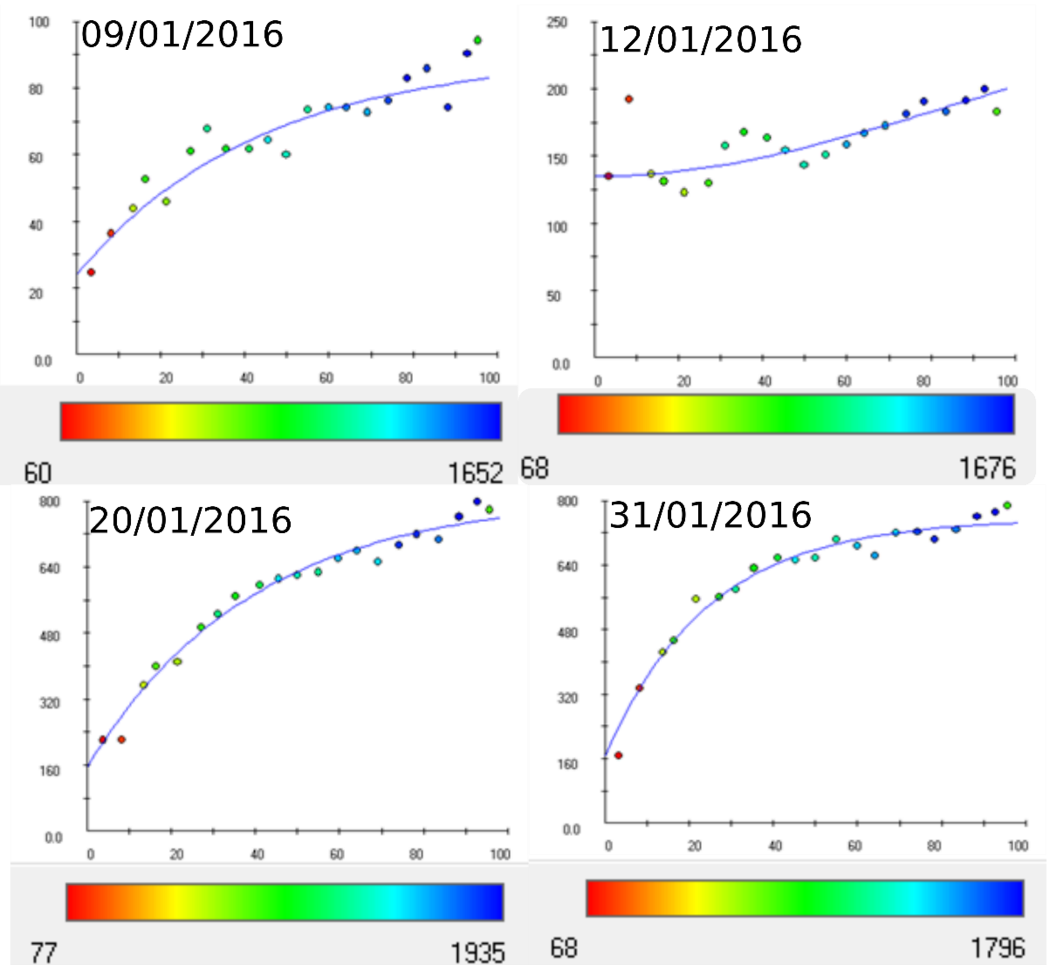


Figure 5.14: Semivariograms of soil moisture of selected fringe fields

The semivariogram shows the variance (y-axis) between all the points separated by distance (x-axis) as a series of points. The legend indicates the number of pairs of points used in the calculation of the variance at each lag (distance between points).

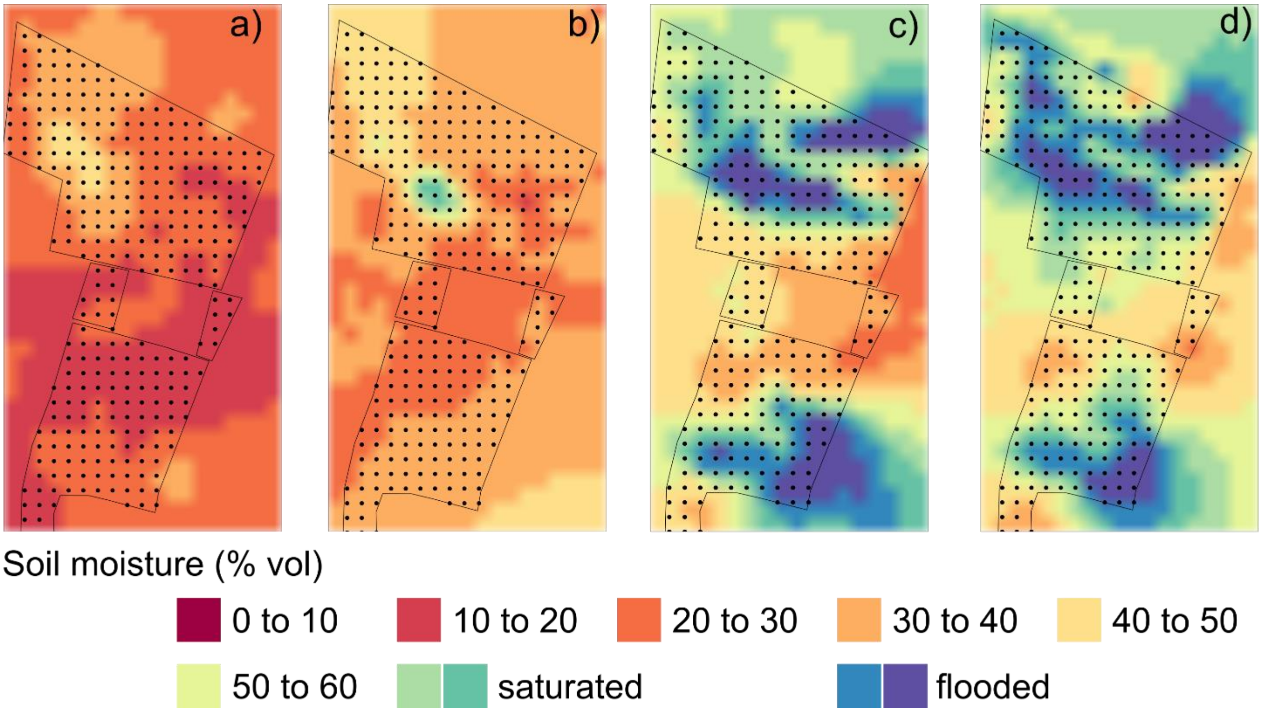


Figure 5.15: Soil moisture distribution maps for the fringe fields on 9th, 12th, 20th and 31st January 2016.

Temporal variograms at maximum 200 days were selected in the generation of the temporal variograms (Table 5.9). This number was selected based on the period between the first and second field survey campaign, which was approximately 200 days to capture the pattern of the second collection period. The range (number of days below which there is a correlation in the observed data) was 109 for the riparian, 65 for the middle and 71 days for the fringe.

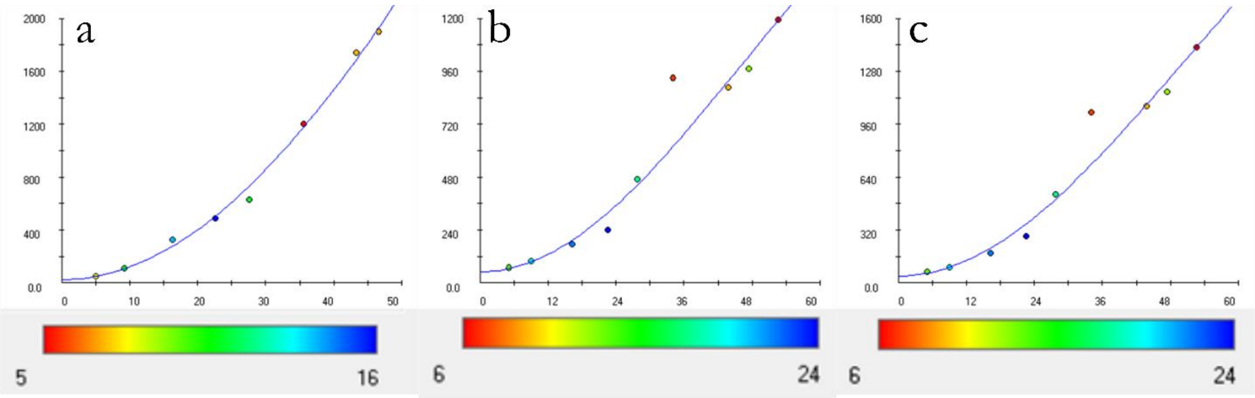


Figure 5.16: Temporal variograms for a) Riparian, b) Middle and c) Fringe zones

Table 5.9: Temporal Gaussian semivariogram parameters. Strong spatial dependence is exhibited for values less than 25%, moderate for values between 25-50% and weak dependence for values greater than 75%.

Zone	nugget (c0)	partial sill (c1)	sill (c0+c1)	spatial dependence	range	rmse
Riparian	24	11464	11488	0.2 (strong)	109.3	45.9
Middle	49.1	2424	2473	2.0 (strong)	65.8	110.6
Fringe	40.8	3287	3328	1.2 (strong)	71.1	113.1

5.4.2 Soil moisture retrieval from SAR backscatter

Field averages of soil moisture were computed for all the data collection dates. Only the fields with soil moisture <40% volume were considered due to the insensitivity of backscatter at high soil moisture values (Dobson and Ulaby 1981, Chanzy 1993, Holah *et al.* 2005). The coefficients of the linear equations are shown in Appendix B4 whereas the goodness of fit of the predictive linear equations are shown in Table 5.10. For all the dates shown, the ground was bare laying the prominence in using the method by (Baghdadi *et al.* 2016) in roughness retrieval.

Simulated kHrms values had very low correlations to the measured values. The coefficient of determination of simulated soil moisture from RadarSat-2 HH-HV was 0.6 while that from TerraSar-X HH-VV was 0.98 (Figure 5.17). Better performance of TerraSar-X is attributed to the higher sensitivity of roughness at high incidence angles (Holah *et al.* 2005). Roughness has a higher effect on RadarSat-2 at 30-36° as compared to TerraSar-X at 22.73°. Lower frequencies are more sensitive to roughness as compared to higher frequencies (Mattia *et al.* 1997). Moreover, penetration depth in RadarSat-2 C- band is more than that of TerraSar-X X-band. This translates to the shorter wavelength band X, having more roughness details than C, explaining the lower performance of C. This means that more roughness data would be required for soil moisture estimation using lower frequency radar data. Additionally, larger uncertainties exist for cross-polarized as compared to the co-polarized backscatter due to the poorer signal to noise ratio (Jiankang Ji *et al.* 1996).

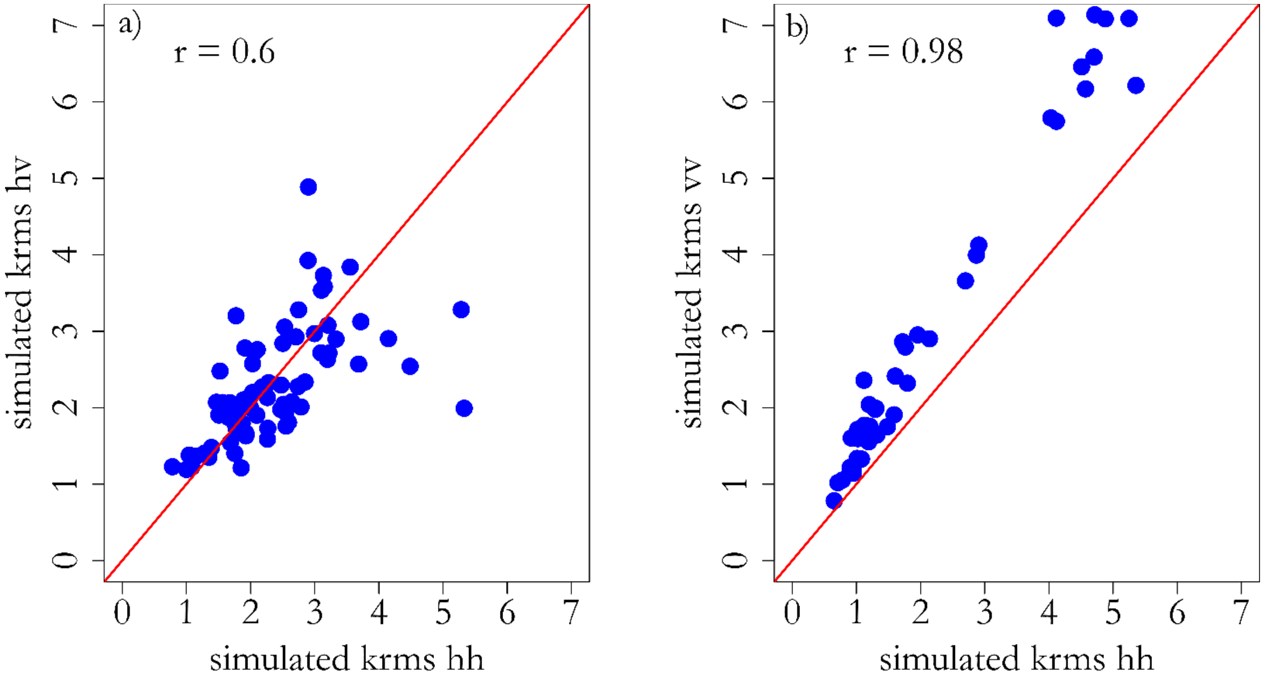


Figure 5.17: Simulated KHrms from RadarSat-2 HH and HV (a) and TerraSar-X HH and VV (b)

Linear regression equations estimated the soil moisture values based on the observed and simulated kHrms. The goodness of fit assessment of the regression equations indicated that the simulated kHrms from (Baghdadi *et al.* 2016) performed better than the observed roughness (Table 5.10). The coefficients are shown in Appendix B3.

Table 5.10: Goodness of fit of the predictive linear soil moisture equations

simulated soil moisture from	Residual standard error	degrees of freedom	Multiple R-squared	Adjusted R-squared	F-statistic	p-value
RadarSat-2 hh with measured khrms	7.61	41	0.36	0.33	11.77	0.00
RadarSat-2 hh with simulated khrms	2.47	41	0.93	0.93	285.70	< 2.2e-16
RadarSat-2 hv with measured khrms	8.24	41	0.25	0.22	6.98	0.00
RadarSat-2 hv with simulated khrms	2.37	41	0.94	0.94	312.40	< 2.2e-16
TerraSar-X hh with measured khrms	8.46	25	0.20	0.14	3.15	0.06
TerraSar-X hh with simulated khrms	3.30	25	0.88	0.87	90.28	0.00
TerraSar-X vv with measured khrms	8.42	25	0.21	0.15	3.30	0.05
TerraSar-X vv with simulated khrms	2.97	25	0.90	0.89	114.20	0.00

The residual standard errors for the models indicate that the simulated khrms produced better results (2.37 to 3.3) as opposed to the measured khrms (7.61-8.46). Additionally, the r squared from the models using simulated khrms (0.87-0.94) showed the data fitted the model better than the measured khrms (0.2-0.33). Moreover, the higher F-Statistic values of the model from simulated khrms indicated that there exists a relationship between the predictor (khrms, backscatter) and response (soil moisture) variables. In conclusion, it is evident that the model from simulated khrms performs better than the ones with the measured khrms as inputs.

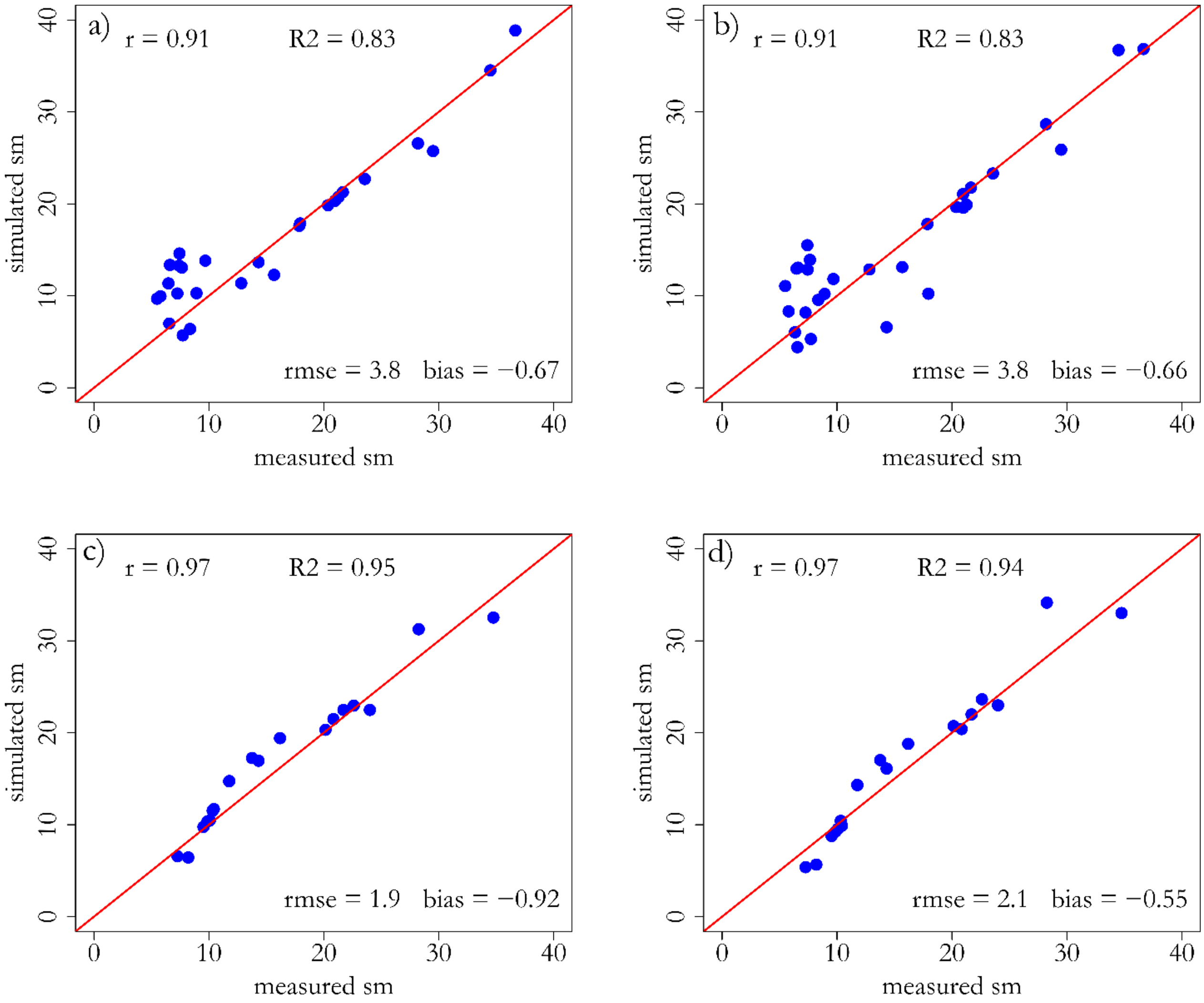


Figure 5.18: Retrieval of soil moisture on validation points from RadarSat-2 HH (a), RadarSat-2 HV (b), TerraSar-X HH (c), and TerraSar-X VV (d)

The predictive linear model was evaluated with the test data in a cross-validation process with 10-fold cross validation (Figure 5.19). The models from simulated kHrms indicated better performance as compared to using the roughness information measured in the field (Appendix B6).

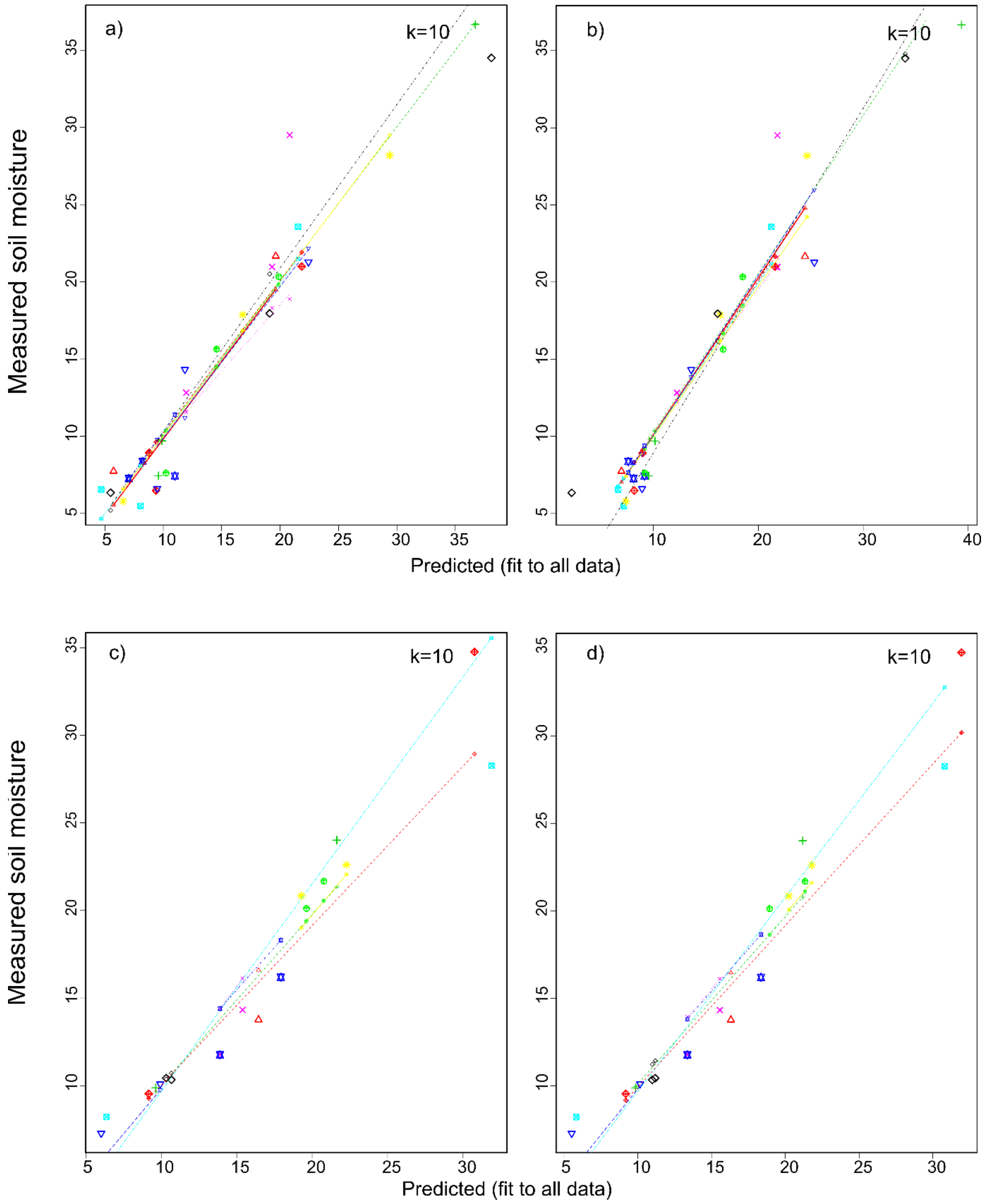


Figure 5.19: K fold cross-validation with simulated Khrms from RadarSat-2 HH (a), RadarSat-2 HV (b), TerraSar-X HH (c), and TerraSar-X VV (d)

The linear model estimated soil moisture with the simulated roughness parameter. Co-polarized bands resulted in a good performance as compared to the cross-polarized RadarSat-2 HV (Figure 5.20). Co-polarized imagery have been reported to be optimal in soil moisture retrieval (Ulaby and Batlivala 1976). Soil moisture from measured kHrms yielded poor results with R^2 from RadarSat-2 HH, HV, TerraSar-X HH, VV being 0.44, 0.3, 0.12 and 0.15 respectively (Appendix B.5).

Observed and simulated soil moisture from RadarSat-2 has a good agreement with underestimation observed in field F1 when the soil moisture is 15% vol (Figure 5.20). The middle fields M1, M2, M3, and M5 had an overestimation when the soil moisture is between 10 to 15% vol. There was a slight overestimation for F1 and F3 for soil moisture greater than 30% vol.

Soil moisture from RadarSat-2 HV yielded the poorest results (Figure 5.20) with a correlation coefficient of 0.33. There was an overestimation of simulated values of all the fields reaching values of 80% vol. This indicates that the predictive linear model based on HV polarization results to erroneous soil moisture values. Cross-polarized sensors have been reported to have a high sensitivity to soil moisture explaining the low performance of RadarSar-2 HV (Hirosawa *et al.* 1978).

TerraSar-X HH and VV exhibited similar patterns with a coefficient of correlation being greater than 0.9 for both HH and VV backscatter signals (Figure 5.21). The soil moisture for the fields at the fringe and riparian zone fit well whereas there was an underestimation of middle fields in low (<10% vol) and high (>25% vol) soil moisture conditions.

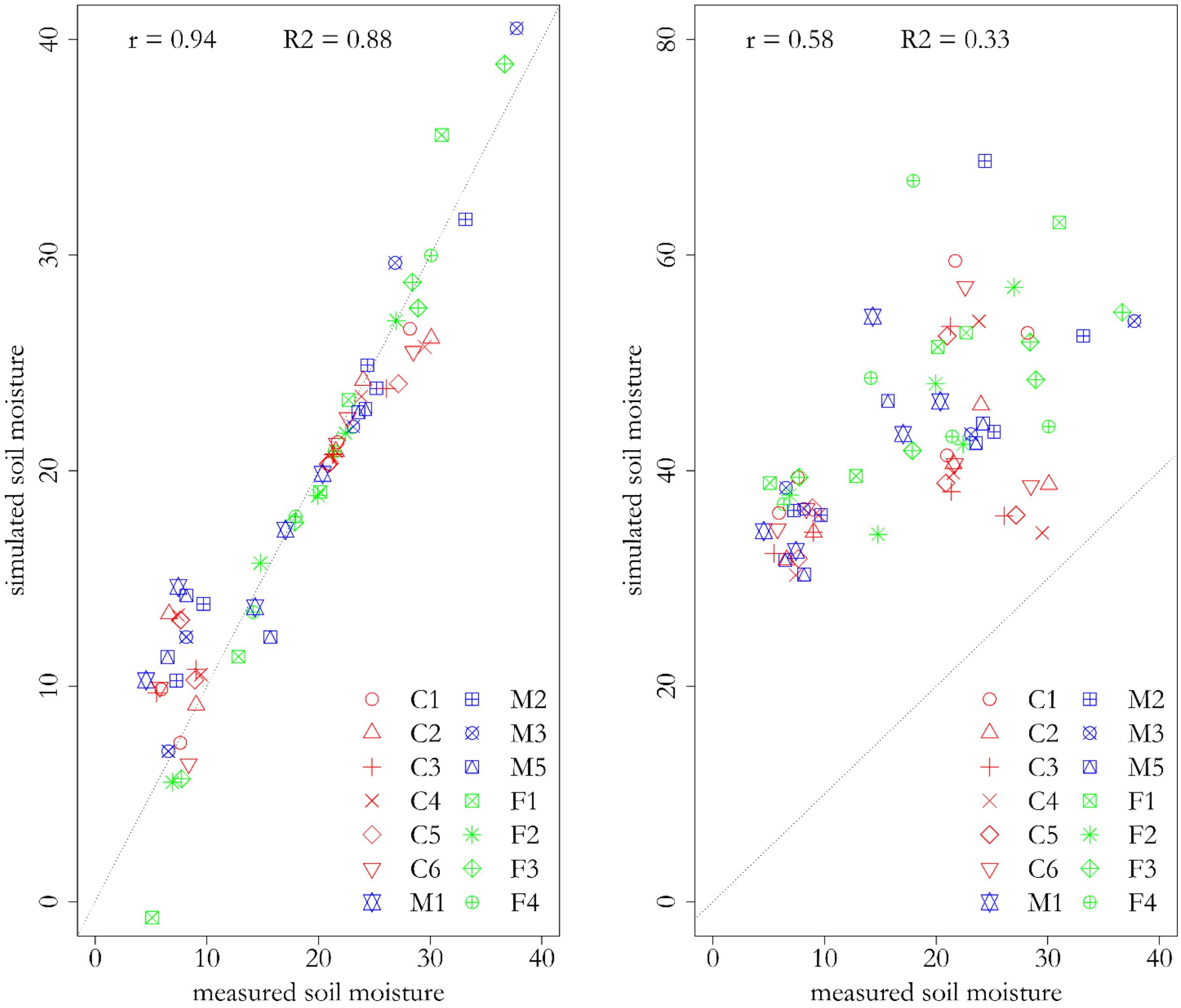


Figure 5.20: RadarSat-2 (a) HH and (b) HV simulated field soil moisture

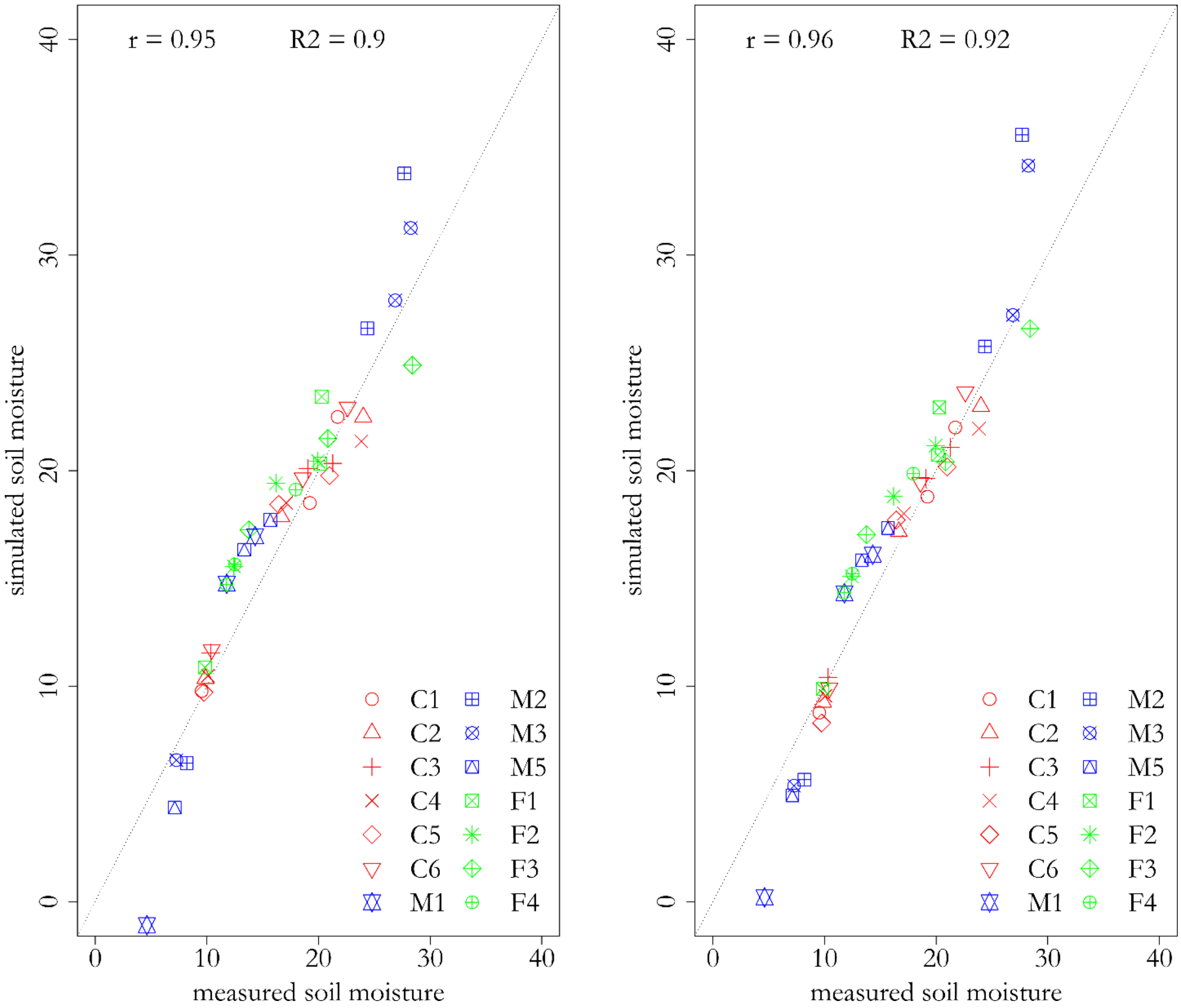


Figure 5.21: TerraSar-X (a) HH and (b) VV simulated field soil moisture

5.5 Conclusion

Knowledge of soil properties plays a crucial role in land management practices. Bulk density (weight of soil per unit volume) affects micro-organism activity, availability of plant nutrients and movement of air and water through the soil which is essential to plant growth. Increased bulk density reduces the air pores in soil subsequently reducing the water available to the plants. Bulk density is influenced by organic matter and texture in terms of percentage clay, sand, and silt. Organic matter is residue from plants returned to the soil in form of humus (USDA n.d.). An increase in organic matter reduces bulk density thus increasing the soils ability to hold water. A high sand content increases bulk density since the pores spaces are reduced. The results obtained from the analysis indicate that the riparian zone has the highest water holding capacity as the bulk density is low with high organic content.

In general, soils in the sample fields are within the densities recommended for plant growth, which is less than 1.4g/cm^3 (Daddow and Warrington 1983). However, the soils can be improved by increasing their water holding capacity. Increasing the organic content of the soil reduces bulk density increasing the air spaces and subsequently ability of motion of the air and water through the soils to the roots of plants (Bauer 1974). Fields in the study site are mainly used in the growth of rice during the wet season and are left bare until the next growing season. During this time, farmers can be encouraged to grow crops that grow well at low moisture content such as sweet potatoes. Their varying root length helps break down the compact layers of soil thus increasing aeration in the next growing season. Planting cover crops and application of fertilizers result in an increase in water holding capacity of the soil (Nicol *et al.* 2015). Soil disturbance when the soil is wet is likely to increase compactness thus reducing the spaces between the soil particles and thus should be avoided. Moreover, adopting drought-tolerant crops which require less growth water can be adapted to increase vegetation productivity (USDA n.d., Evanylo and McGuinn 2000, Nicol *et al.* 2015).

The field soil moisture dynamics and its relation to the coefficient of variation gives a basis upon which sampling designs can be set up. For the fields selected, increasing soil moisture results in an increase in deviation from the mean of the riparian fields. The converse is true for the middle and fringe fields. The coefficient of variation decreases with increasing soil moisture and was greater than 35% vol for the riparian, middle and fringe fields in low soil moisture, indicating the high variability of soil moisture across fields in the dry season. The implication of this with respect to assessing the field soil moisture against a remote sensed soil moisture product is that the actual soil moisture conditions are more reflective of the remote sensing product in wetter conditions since the cv is lower (Charpentier and Groffman 1992, Famiglietti *et al.* 1999). On the other hand, the spatial variograms have a weak spatial dependence of greater than 75 in the low soil moisture periods indicating there is little correlation between the measured soil moisture. This is in support of the results from the computations of the coefficient of variation. The preliminary information on the pattern of soil moisture in wet and dry conditions informs a decision of the number of points that would be required to correctly capture the mean moisture of fields at given margins of error and confidence levels. The number of points required increases with increasing soil moisture since the variation is high under such conditions as stipulated by the coefficient of variation. An increase of soil moisture from 20 to 60% vol reduces the number of points required by a factor of 6.5 for the middle and a factor of 11 for the fringe fields. The low correlation between the riparian and coefficient of variation may have been due to a lower number of sampling points as compared to the fringe and middle fields. The riparian fields were liable to flooding and hence in such cases, no sampling was done. This information is highly valuable for the planning of additional soil moisture data in which case the dry conditions require additional data collected whereas in wet conditions fewer moisture measurements suffice to represent the field mean (Hupet and Vanclooster 2002).

The relative differences and the standard deviation of the relative differences aided in identifying time stable fields that could be used in upscaling the soil moisture to a larger area coverage. Riparian fields were the most stable temporally for the period March to May and December to February. The soil texture in particular high clay content and its water holding capacity after rainfall events explains this temporal persistence. (Cosh *et al.* 2008) reported that soil type had a high influence on the temporal stability of soil moisture. The riparian fields could be used as representative fields in upscaling strategies for areas with similar soil texture properties. The point data used in upscaling is feasible in development of products comparable to modeled remote sensing products at satellite resolution and footprint (De Lannoy *et al.* 2007). Reduction of the uncertainty introduced in upscaled products can be accomplished by utilizing soil moisture related predictor variables such as NDVI or weather information (Crow *et al.* 2012).

Simulation of roughness parameter, kHrms, from RadarSat-2 and TerraSar-X yielded values that outperformed the observed roughness parameters. This indicated that infield variation of roughness is not captured by observations at several points in the field. Simulated soil moisture indicated that TerraSar-X had better results than RadarSat-2 explained by the difference in frequency, incidence angles, and polarization.

The overall context of the research was geared towards highlighting variations existing at the field scale. For this study, the variation existing in physical properties sand, silt, clay, organic carbon, bulk density for the three hydrological sites along the floodplain were assessed. These properties were assumed to be constant for the duration of the study. The variation in highly dynamic soil moisture was assessed for the periods between March and May 2015 and December 2015 to February 2016. It is evident that there are differences at the field scale for these soil properties. This subsequently translates to having different management options for the sites, for example, there is a distinct water accumulation pattern in the fields shown in Figure 5.15. If there is leaching or surface flow draining water out of his farm from this location, then to retain the soil nutrient availability, the farmer could apply additional manure at this excess water locations. The variations call for adoption of farm-specific technology adopted to manage the needs of individual farms. Soil moisture is a critical parameter in plant growth, however, it is tedious to collect every so often. Remote sensing imagery can be adopted in estimating the highly dynamic soil moisture.

In summary,

1. An increase in soil moisture resulted in an increase in standard deviation. The range increased with increasing moisture in the riparian whereas it reduced for the middle and fringe.
2. Temporal variation in soil moisture is due to precipitation events. However, the spatial variation of the soil moisture measured is explained by differences in soil texture. Moreso for the riparian fields which had higher water content due to high clay and organic matter content.
3. The coefficient of variation has an exponential relationship to mean soil moisture. Determination of required number of points at 10% confidence interval with an error limit of 10% and 20% indicated that higher number of points are needed for the fringe as compared to the middle. The riparian fields,

however, exhibited a poor correlation between the cv and mean soil moisture which can be explained by the lower number of measurement points included in deriving the relationship.

4. 105 and 125 measurement points in the fringe and middle fields suffice to estimate mean soil moisture at 95% confidence level and 10% relative error in dry moisture conditions. The values reduce to 27-30 measurement points for 30% vol soil moisture.
5. The riparian fields have a higher temporal stability as compared to the fringe and middle fields.
6. Temporal dependence of measured soil values is 109 days for the riparian, 65 days for the middle and 71 days.
7. RadarSat-2 and TerraSar-X yielded high performance in simulating soil moisture with simulated roughness parameters as inputs. Higher frequency X band outperformed the C band in the estimation of soil moisture from bare surfaces.

The temporally stable fields can be used as locations for setting up future soil moisture sensors. The potential use of such auxiliary data is in the validation of remote sensing data. Moreover, additional SAR sensors and synergy of different sensors in soil moisture retrieval is recommended for exploration. Data from additional soil moisture sensors can also be input in hydrological and rainfall-runoff models, all geared towards improved understanding of hydrological processes in Ifakara in the Kilombero Valley. Furthermore, it is recommended that supplementary research on the effect of vegetation on soil moisture spatial-temporal variability be investigated. Additionally, it is recommended that subsequent studies assess the infield distribution of chemical properties vital in crop growth such as Nitrogen, Phosphorus, and others at field scales.

6. Kilombero catchment soil moisture assessment

6.1 Introduction

Wetland management and conservation have been subjects of research especially due to the pressure exerted on the wetlands with increasing population (Whigham *et al.* 1993, Rebelo *et al.* 2010, Mombo *et al.* 2011)(Rebelo *et al.* 2010). Emphasis on the need for development and implementation of management policies geared towards protection and regulation of wetlands providing ecosystem services is on the rise (Millennium Development Goals and Nations Unies 2011, Sachs 2012, Leemhuis *et al.* 2017). However, for the sustainable development decisions, there is a need for data highlighting the overall interrelated functioning of the ecosystem. An important aspect of understanding an ecosystem is knowledge on the spatial location, extent, and dynamics of geophysical parameters. Complex nature of wetlands and high temporal variability favor the utilization of remote sensing in wetland assessments (Klemas 2011, Tiner *et al.* 2015). Vast extent and repeated coverage facilitate the continued monitoring of land parameters that influence wetland environments and their use (Guo *et al.* 2017). Water conservation and runoff regulation are among the wetland functions making wetlands a preferred ecosystem for increased food production (Chen and Wong 2016). Exploitation of the ecosystem potential for agriculture while neglecting the negative impacts of the natural resources degrades the functional capability of the wetland (Rockström and Karlberg 2010, Rockström *et al.* 2017, Rosa *et al.* 2017). Information on water availability to crops guides farmers on irrigation times whereas seasonal forecasts aid in the determination of planting and harvest times for maximum yield (Rosa *et al.* 2017). Historical study on the trends of geophysical parameters such as soil moisture could act as a plant distribution indicator geared towards plant species selection based on their tolerance to environmental stress (Li *et al.* 2013). Such decisions could include a shift in the cropping system and adaptation of climate resilient crops based on the soil water availability (Verhoeven and Setter 2010, Nicol *et al.* 2015).

Soil moisture is a crucial variable in hydrology and land-atmosphere interaction. It controls the meteorological and hydrological responses of the land surface. Soil moisture influences catchment drainage, groundwater recharge, and runoff. Remote sensed soil moisture products have vast spatial coverage and high temporal capability required to meet the demands of the agricultural sector (Lillesand *et al.* 2014). Soil moisture products previously in use include SMMR from 1978 to 1987 (C and X band), SSM/I launched in 1987 (Ku band), TRMM Microwave Imager in 1997 to 2001 (X band), AMSR E launched in 2002 to 2011 (C and X band), Windsat launched in 2003, Escat (L band), ASCAT (C band) onboard METEO A launched in 2006, SMOS launched in 2009 operating in the L band and SMAP launched in 2015 operating in L band (Das and Paul 2015b).

The Kilombero catchment and Tanzania by extension, lack precise knowledge on start and end of growing season. For a region that is dependent on subsistence rain-fed agriculture, precision on planting dates determines a good or failed harvest. Several studies have noted increased intensity in variabilities of extreme rainfall events (especially droughts) coupled with variation in length of growing seasons. (Peltonen-Sainio *et al.* 2008, Julien and Sobrino 2009, Shrestha *et al.* 2012) reporting longer growing seasons while (Vizy *et al.* 2015) reported a reduction in length of growing season. Precipitation is the main driver of soil moisture. However, vegetation, evaporation, soil type, irradiation, topography mainly influence the water present in the soil (Dorigo *et al.* 2012) and thus soil moisture is preferred to precipitation in the study of water availability for vegetation (Chen *et al.* 2016). However, non-existent in situ soil moisture monitoring network points to non-sustainability of rain-fed agriculture over the region hence famine and starvation. The lack of in situ soil moisture monitoring networks has made it impossible to validate the model derived soil moisture products, which could be analyzed to provide growing season information in conjunction with Normalized Difference Vegetation Index (NDVI). Since model soil moisture forecasts are normally available 3-6 months in advance (AghaKouchak 2014, Sheffield *et al.* 2014, Spennemann *et al.* 2017, Wilhite and Pulwarty 2018) it would be possible to determine the start of the season early enough if soil moisture values are available for validation.

Development and application of statistical methods to global soil moisture are important as they are utilized as inputs or for validation purposes for various applications. Some applications of global soil moisture products include numerical weather predictions, runoff forecasting/hydrological modeling and forecasting, vegetation and crop growth monitoring, ecological modeling and forecasting, epidemic risk assessment (drought monitoring), meteorological modeling and forecasting and societal risk assessment (Ochsner *et al.* 2013, Wagner *et al.* 2013).

Numerous studies have assessed and evaluated soil moisture products against in situ moisture measurements (An *et al.* 2016, Zhang, Zhang, *et al.* 2017). Low-frequency microwave region mainly used to obtain soil moisture from near surface layers. That is 1-10 GHz (Albergel, de Rosnay, Gruhier, *et al.* 2012). L band most suitable for observation of soil moisture as it has a high vegetation penetration (C band 0.5 -2 cm, L band 3-5 cm). Atmospheric effects and vegetation cover affect higher frequencies. Soil moisture distribution at large scale influenced by evaporation, soil type, irradiation, vegetation and topography and should be reflected in the soil moisture products. A challenge that remains in the use of large-scale satellite-derived soil moisture products is validation. This is especially so due to site-specific factors such as soil and vegetation that influences instrument performance, sampling networks for validation are limited in historic and current time (Dorigo *et al.* 2015). Moreover, in situ soil moisture is rare for the East African region as noted from the International Soil Moisture Network (Dorigo *et al.* 2011). In another analysis, (Albergel, de Rosnay, Gruhier, *et al.* 2012) assessed ASCAT, SMOS, ECMWF soil moisture-das-2 and in situ data for various parts of the world (Australia, France, US, West Africa, Spain, Italy, Germany) to test the ability of the global products to represent surface soil moisture. To avoid seasonal effects, monthly anomaly time series for soil moisture products were analyzed. It was reported that locally observed rainfall introduced discrepancies when compared to course resolution products particularly when local storm events occurred in summer. Over Europe performance of SMOS dependent on RFI effects. ASCAT and SMOS signals were influenced by vegetation with reduced sensitivity to soil moisture over dense vegetation canopies. The assimilated ECMWF SMDAS2 product was more consistent with in situ data compared with ASCAT and SMOS.

(Leroux *et al.* 2014) assessed the reliability of AMSR E VUA, ASCAT, SMOS, ECMWF soil moisture products against in situ data in four watersheds (Leroux *et al.* 2014). VUA and ASCAT grid products were first linearly interpolated on the SMOS grid. The linear correlation coefficient (*R*), standard error of estimate (SEE) between derived soil moisture and in situ measurements SEE, bias and root mean square error (RMSE) were computed for each data set with the ground measurements over each watershed at the corresponding overpass times. None of the four products had a significantly better correlation with in situ soil moisture. ECMWF had a lower SSE indicative of less scattered data, SMOS had a lower bias and RMSE while VUA bias was higher when the vegetation and soil temperature was low. This indicated that VUA has a high sensitivity to vegetation as compared to SMOS. AMSR E frequencies 6.9-10.7 GHz were reported to be more sensitive to vegetation and temperature than SMOS at 1.4 GHz. Moreover, the sensitivity was heightened by the effective temperature modeled at 36 GHz Temperature Brightness (TB). ASCAT soil moisture index was reported to be very noisy with rain events were observed in retrievals. SMOS retrievals were closest to the in situ measurements despite the in situ dynamics having a poor representation. The mean square error was least in the morning overpass as compared to the afternoon overpass.

The studies compare global soil moisture products against in situ measurements. However, additional global datasets in combination with these products aid in understanding the causal factors of the soil moisture recorded. Moreover, the comparison of in situ soil moisture with satellite products is deemed problematic due to the difference in scales of the two datasets that is point data against areal averages extending several kilometers. Remote sensed soil moisture is dynamic due to highly varied and constantly changing atmosphere influences. On the other hand, in situ moisture is measured without taking account of atmospheric influences with measurement depths determined by the sensors in use (Wagner *et al.* 2007). These challenges of in situ moisture have driven researchers to analyze the performance of global products without the inclusion of the reference soil moisture. Differences in systems, retrieval algorithms, and mission designs are causal factors of varied data values in global products and therefore studies geared towards assessing global product uncertainties and the inconsistencies between datasets are performed.

(Rötzer *et al.* 2015) analyzed ASCAT and SMOS satellite-based soil moisture products and ERA-Interim modeled soil moisture based on global spatial and temporal Koeppen Geiger climate classification (Rötzer *et al.* 2015). They reported a high correlation for all products in areas with low vegetation, especially in arid regions. Around the equator, SMOS conveyed low soil moisture in tropical rainforest whereas ASCAT and ERA-Interim had moderate values. Backscatter recorded by ASCAT indicated less sensitivity to soil moisture in densely vegetated tropical rain forests. ASCAT resulted in high fluctuations in time series spatial mean caused by the porosity data with high scattering and high variance indicating the large temporal changes in the soil moisture. High vegetation in tropical rainforest affected the ASCAT and SMOS retrieval of soil moisture. SMOS reported a weaker seasonal dependence on seasons than ASCAT. At L band, SMOS reported less sensitivity to vegetation effects than ASCAT at C band explaining less dependency of SMOS on seasons.

Vegetation time series trend is preferred to direct numerical observations between ground and satellite observations because of differences in acquisition time, vertical resolution, spatial resolutions and inter-observational periods (Owe *et al.* 2008). Assumptions made in vegetation seasonality analysis include correct representation of vegetation parameters by the sensor and vegetation signals change smoothly in time (Eklundh and Jönsson 2015). However, this is not the case as factors such as atmospheric absorption, cloud coverage, background color variations, viewing angles cause presence of noise in the data making the extraction of seasons difficult (Jonsson and Eklundh 2002). Application of cloud masks and atmospheric corrections reduce the noise however for a large number of images intended for time series assessment, the method is cumbersome. This resulted in the use of maximum value composites, which proved to be effective though some level of noise was still present in the data, calling for the development of smoothing algorithms. Universal start and end of seasons were not defined due to variations in climatic patterns, vegetation periods and varying number of seasons in different parts of the world (Eklundh and Jönsson 2015). Areas with big cycles of growth and decline are shown to have high amplitudes whereas low amplitudes are in areas where the change between highest and lowest NDVI values is small. In moist areas, amplitude is expected to be low due to weak seasonal variation as the area is green round the year (Jönsson and Eklundh 2004). Additionally, meteorological conditions affect the development of vegetation in particular presence of rainfall, which in turn support in assessment of drought (Rojas *et al.* 2011, Sheffield *et al.* 2014, Klisch and Atzberger 2016, Zhang, Chen, *et al.* 2017). Understanding the seasons of vegetation growth give a deeper insight into the interactions of vegetation and climate and subsequently give science-based support in decision making geared towards development and adaptation of coping mechanisms especially for rained agriculture-dependent areas.

(Dorigo *et al.* 2012) analyzed the trends of a merged remote sensed soil moisture product (consisted of SMMR, TMI, AMSR-E, ERS1/2, ASCAT) from 1988-2010 against ERA-Interim and GLDAS_NOAH soil moisture, Global Precipitation Climate Project precipitation and Global Inventory Monitoring and Modeling Studies NDVI (Dorigo *et al.* 2012). Significant trends of the merged product and the modeled soil moisture product and precipitation product resulted. The trend assessment with NDVI seemed complex due to the influence of vegetation type and changes. The merged product, however, had uncertainties introduced by differences in temporal resolution and coverage, which could have contributed to its ability to capture the NDVI trends. In Tunisia, an empirical model for predicting AVHRR NDVI from root zone soil moisture was developed (Zribi *et al.* 2010) following a high correlation of soil moisture in the rainy season with vegetation dynamics. However, the method was restricted for use in large-scale areas in semi-arid environments. (Chen *et al.* 2014) assessed the impact of soil moisture on vegetation over Australia reporting an NDVI lag period of one month behind soil moisture (Chen *et al.* 2014). Additionally, they reported enhancement of vegetation growth in dry regions and at the beginning stages of growth in wet regions. This emphasizes the importance of knowledge of the beginning of soil moisture increase. The research therefore seeks to identify the start of season (SOS) from NDVI and relate it to soil moisture measurements. This information is crucial to farmers as they are aware of the periods over which to plant to increase their yields since vegetation is highly sensitive to soil moisture at the beginning of the season (Chen *et al.* 2014). Moreover, the study hopes to offer a solution by addressing the uncertainties arising with respect to planting times due to the changing climatic conditions.

The objectives of the study were 1) to evaluate global soil moisture products over the Kilombero catchment, 2) to assess the error structure of each dataset 3) to develop a merged soil moisture product for the Kilombero catchment, and 4) to assess the applicability of the merged soil moisture product in determining the start and end of planting season. The merged soil moisture products can be used as a component of hydrological modeling, runoff models and food and water security monitoring in this data-scarce region. Modeled soil moisture products and a satellite soil moisture product have been selected for the analysis. In situ soil moisture products over East Africa were not included in the evaluation, as they are not available. Besides, the method in the application, Triple Collocation (TC), requires a large set of data (Scipal *et al.* 2010) and it evaluates relative errors of products. Moreover, the FLDAS_NOAH, FLDAS_VIC, and ERA-Interim were selected as they had a long coverage period (2010 to 2016) and had the highest spatial resolution as compared to other model products (0.1°, 0.25°, and 0.25° respectively). Spatial and temporal variability was analyzed using the Empirical Orthogonal Functions of the three global soil moisture products i.e. the FLDAS_NOAH, FLDAS_VIC, and ERA-Interim.

6.2 Global soil moisture products

6.2.1 Soil moisture satellite sensors

Passive microwave sensors measure radiation emitted from the surface of the earth as brightness temperatures. The components contributing to the magnitude recorded include the upwelling atmospheric emission, earth's surface emission attenuated by atmosphere, atmospheric downwelling atmospheric emission reflected at surface and attenuated along the upward path by the atmosphere and cosmic background emission attenuated by atmosphere reflected at the surface and attenuate again along the upward path by the atmosphere (Wigneron *et al.* 2017). Passive sensors have a high revisit period e.g. Soil Moisture and Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) 2-3 days. Communications and broadcast systems cause Radio Frequency Interference (RFI) resulting in abnormally high brightness temperatures. RFI signals affect brightness temperatures at a higher intensity in the C band as compared to the X band (Oliva *et al.* 2012). Radio frequency interference has led to degraded measurements in some parts of Europe, South Asia, Middle East and China.

The RFI has reduced through sensitization in concerned countries to protect the 1400-1427 MHz SMOS operation frequency (Oliva *et al.* 2012). Radiometer measurements are most accurate in areas of low vegetation. However, lower frequencies have higher penetrative ability suitable for measurements in vegetated areas. The penetration ability of L band radiometers makes it an asset in the retrieval of soil moisture. Additionally, L band has a deeper sampling depth, approximately 0-3 cm as compared to C band 0-1 cm and are affected less by atmospheric effects (Wigneron *et al.* 2017). Unique values are assigned to areas with water, ice snow, excess surface roughness and steep topography. Algorithms in L band soil moisture retrieval include:

1) 2-Parameter Retrieval Algorithm: SMOS soil moisture is based on this algorithm. Multi-angular dual polarized observations of brightness temperature are inverted using L Band Microwave Emission of the Biosphere (L-MEB) model to obtain soil moisture and optical depth at nadir. Incidence angles range from 0-55° with brightness temperatures for the 'optimal location' viewed from different angles. The value at the optimal location is used as a first approximation in the 2-parameter retrieval of soil moisture and optical depth at nadir. No estimates are required in the retrieval.
2) Single Channel Algorithm (SCA): SMAP is based on this algorithm. TB is measured at one incidence angle, 40°, at one polarization. TB are corrected for vegetation and roughness and finally, a dielectric mixing model is used to retrieve soil moisture.
3) Dual Channel Algorithm: This is an extension of the SCA where both polarizations, H, and V from SMAP are combined to retrieve soil moisture by minimizing the RMSE between simulated and observed Temperature Brightness (TB).
4) Land Parameter Retrieval Model (LPRM): Microwave Polarization Difference Index (MDPI) is computed from the TB and the $\tau - \omega$ model used to retrieve soil moisture. In densely vegetated areas, LPRM models do not converge and thus no data is recorded.
5) Multi-Temporal Dual Channel Algorithm: Combined consecutive TB observations taken in the early morning when the temporal changes in Vegetation Water Content (VWC) are constant are used in retrieval of soil moisture. The retrieved albedo is constant due to the stability of the surface temperature in early hours of the morning.

Table 6.1: A summary of the data used

Soil Moisture products	**Spatial resolution**	**Temporal resolution**	**Format**	**Period assessed**	**Data portal**
FLDAS_NOAH	0.1°	Daily	.NC	2010-2016	https://ldas.gsfc.nasa.gov/FLDAS/FLDASdownloadphp.
FLDAS_VIC	0.25°	Daily	.NC	2010-2016	https://ldas.gsfc.nasa.gov/FLDAS/FLDASdownloadphp.
ERA-Interim	0.125°	Daily	.NC	2010-2016	http://apps.ECMWF.int/datasets/data/Interim-full-daily/levtype=sfc/
SMOS	30-50 km	Daily	.NC	14/1/2011-2016	https://SMOS-diss.eo.esa.int/socat/SMOS_Open/search
Vegetation product	**Spatial resolution**	**Temporal resolution**	**Format**	**Period assessed**	**Data portal**
MODIS	250 m	16-day	.HDF	2010-2016	https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q
Rainfall product	**Spatial resolution**	**Temporal resolution**	**Format**	**Period assessed**	**Data portal**
TRMM_3B42	0.25°	Daily	.NC	2010-2016	https://disc2.gesdisc.eosdis.nasa.gov/data/TRMM_L3/TRMM_3B42_Daily.7/

6.2.1.1 SMOS

It was launched in 2009 has a spatial resolution of 30-50 km. Ascending mode in the early morning minimizes perturbation due to air, vegetation and soil temperature in the L band (Kerr *et al.* 2001). Moreover, L band has a high sensitivity to changes in soil moisture and salinity in the ocean and has a higher penetration depth up to 5 cm which performs well in dense vegetation (Kerr *et al.* 2001). On the other hand, high-frequency products have a penetration depth of 1-2 cm preventing penetration into the soil layer in vegetated environments. SMOS has multi-angular dual polarization capabilities hence any points on the surface are viewed frequently at different angles and polarizations. The angular information separates the different contributions to the signal. The angular signature retrieves soil moisture and vegetation optical depth (vegetation water content) which expresses the quantity of signal that is absorbed by the vegetation layer through minimization of a cost function between L band microwave emission of biosphere model (L-MEB) and the corresponding SMOS measured temperature brightness to estimate soil moisture and VOD. L-MEB is associated with certain land cover classes thus making it possible to quantify the contribution of these classes (Kerr *et al.* 2012). The retrieval algorithm consists of static (soil texture from FAO, ECOCLIMAP land use maps, topography index) and dynamic (rain, temperature, snow from ECMWF) datasets. Radio Frequency Interferences, vegetation opacity and surface roughness have regular updates for future inversions.

HQN model is used for roughness parametrization. For low vegetation, h_r is 0.1, for forests, it is set as 0.3, Q is 0 whereas N_v=0 and N_h=2. Dobson model was used in modeling dielectric constant before L2 v5.5. Currently, Mironov model is in use. Mironov improves retrievals over dry warm surfaces reducing extreme values in soil moisture. ECMWF determines the vegetation and soil temperature. $\tau - \omega$ model used to model vegetation. Albedo for low vegetation set to 1 while that for forests is set to 0.06-0.08. The vegetation structure is assumed to be isotropic i.e. tt_h=tt_v=1. Optical depth at nadir and soil moisture are byproducts of the 2 parameter retrievals by inversion of the LMEB model. Optical depth at nadir is modeled by a linear function of LAI.

The spatial extent of SMOS (on average, 40x40 km²) poses a challenge due to the introduction of significant spatial variability in soil moisture. Moreover, due to the spherical nature of the earth, the view angle changes therefore different areas covered at different angles (Kerr *et al.* 2012). Additionally, SMOS surface soil moisture retrieval maps are available though there are gaps associated with RFI, steep topography, dense vegetation, snow cover, frozen soils. Data products available are Level 1(TB) and level 2(ocean salinity over oceans or soil moisture/vegetation opacity over land) and level 3 (1d, 3d, 10d or particular month for the globe either morning or afternoon passes for soil moisture and vegetation opacity). To improve spatial resolution, SMOS data can be combined with other data such as MODIS data.

Albedo describes the fraction of solar energy reflected from the earth back into space. It is also optical brightness scaled from 0 to 1. The ratio of irradiance reflected to irradiance received by a surface. E.g. Ice albedo 0.5 to 0.7 means 50 to 70%of the incoming radiation is reflected. While for ocean 0.060 means only 6% of incoming solar radiation is reflected and the rest is absorbed.

Reflectivity describes the light reflected from the surface in relation to light incident upon it. The fraction of radiant energy reflected from a surface.

Emissivity is the ratio of energy radiated from a surface in relation to that radiated from a black body at the same temperature and wavelength under the same viewing angles.

The $\tau - \omega$ model computes emission from a two-layer soil water medium in H and V polarization using three terms; upwelling emission by vegetation, downward emission from the vegetation reflected by soil and attenuated by vegetation canopy and emission by soil attenuated by canopy (Wigneron *et al.* 2017). The terms account for influence of vegetation and soil roughness on angular measurements and polarization mixing. The temperature brightness is given by:

$$TB = emission\ by\ soil\ +\ emission\ by\ veg\ +\ emission\ from\ veg\ to\ soil\ to\ veg$$

$$TB = T_{soil}\left(1 - R_{soil,p}\right)VA_p + \left(1 - \omega_p\right)T_{veg}\left(1 - VA_p\right) + \left(1 - \omega_p\right)T_{veg}\left(1 - VA_p\right){R_{soil,p}}^{VA_p}$$

$$VA_p = \exp\left({}^{-\tau_p}/_{\cos\theta}\right)$$

$$T_{soil}\ is\ effective\ temp\ of\ soil, T_{veg}\ is\ effective\ temp\ of\ veg, R_{soil,p}\ is\ soil\ surface\ reflectivity,$$

$$\tau_p\ is\ veg\ optical\ thickness, \omega_p\ is\ albedo\ of\ veg\ canopy, VA_p\ is\ veg\ attenuation\ factor$$

Roughness modeling

The HQN model is used to model roughness.

$$R_{soil,p} = \left[(1 - Q)R_p + QR_q\right]exp^{-H_p\cos^{N_p}\theta}$$

$$p = H\ while\ q = V\ polarization\ \ , R_p \& \ R_q\ are\ the\ reflectivity\ of\ a\ smooth\ soil\ surface,$$

$$H_p\ is\ the\ rmsh\ or\ intensity\ of\ roughness, Q\ is\ the\ polarization\ mixing\ parameter,$$

$$N_p\ allows\ for\ dependence\ of\ roughness\ on\ \theta$$

$$Q = 0, N_p = 0\ and\ N_p = 2\ \ hence\ \ R_{soil,p} = R_p exp^{-H_p\cos^{N_p}\theta}$$

Roughness intensity increases with increasing roughness, Q, the difference between emissivity in horizontal and vertical polarization decreases as Q increases whereas Q shows an increase with increasing frequency.

Vegetation modeling

The vegetation optical thickness (height of vegetation canopy), is a function of optical thickness at nadir, incidence angle, and polarization (CBSA *et al.* 2014).

$$\tau_p = \tau_{NAD}\left(\sin^2\theta\ \ tt_p + \cos^2\theta\right)\ \ p = H, V$$

A value of tt>1 results to increasing τ_p as a function of θ. In cases where τ_p is independent of polarization and incidence angle, tt=1. The parameter tt accounts for effect of vegetation anisotropy on optical depth with respect to incidence angle, polarization. A good relationship was found between height of stems and tt v.

*property of being directionally dependent, which implies different properties in different directions

6.2.2 Modeled products

Modeled products are preferred due to the challenges that come with remotely sensed products such as algorithm changes when using multiple satellites and hence multiple sensors and algorithms (Liu *et al.* 2017). Additionally, short temporal coverages and gaps due to flight paths of the satellites prove acquisition of a continuous dataset for uninterrupted monitoring a task.

A Land surface model for food security monitoring should have a rainfall product, a long history of data and both drought and hydrological communities should utilize the datasets to estimate land surface states. GLDAS is a global terrestrial modeling system that produces optimal fields of land surface states by incorporating satellite and ground-based observations (Rodell *et al.* 2004). It drives NOAH, Mosaic, Community Model and Variable Infiltration Capacity (VIC). However, GLDAS utilizes a global rainfall dataset not optimized for data scarce food insecure regions (McNally *et al.* 2017) and hence the preference in utilizing the FLDAS products.

FEWSNET was created in 1985 by USAID as a provider of early warning and analysis of food insecurity issues. It acts as a support system as it provides data to help governments and agencies plan and respond to humanitarian crises. FLDAS is a customized LIS by NASA as a decision support tool in the functions of FEWSNET by producing models and estimates of hydroclimatic data over food insecure areas of Africa (McNally *et al.* 2017). FLDAS relies on the FEWSNET rainfall and is thus customized for the needs of data scarce food insecure regions in Africa. FLDAS constitutes of meteorological input, a land surface model input, and a post-processing evaluation component. NOAH and VIC Land Surface Models drive FLDAS.

FLDAS_NOAH is a 4-soil layer water and energy balance Land Surface Model with a spatial resolution of 0.1°. It available from 1982 to present, uses CHIRPS and MERRA2 and 1 KM NCEP modified IGBP land cover map from MODIS. NOAH soil moisture is available at daily, 5-day, 10-day, monthly temporal resolution. Evaporation in NOAH land surface models includes the weighted sum of canopy intercepted evaporation, transpiration from vegetation canopies and evaporation from bare soils based on the respective land surface coverage fractions (McNally *et al.* 2017).

The assumption in the calculation of potential evapotranspiration is a saturated ground surface with no canopy resistance. Canopy intercepted evapotranspiration scales PET by intercepted canopy water content and maximum canopy capacity. Transpiration scales PET with solar radiation, vapor pressure deficit, air temperature and soil moisture. Bare soil evaporation scales PET by a factor of current soil moisture availability. The surface runoff is computed with a two-layer conceptual approach from the Simple Water Balance Model. The upper layer produces surface runoff from excess precipitation when there is no soil moisture deficit. Subsurface runoff is produced as gravitational drainage from the bottom soil layer. Other parameters include monthly greenness fraction and albedo, soil texture datasets, vegetation and soil parameters.

In FLDAS_VIC, FLDAS uses VIC in water and energy balance to better represent surface fluxes (sensible heat, latent heat, ground heat, outgoing longwave radiation) where evapotranspiration is a greater proportion of rainfall than runoff. Run at a sub-daily time step, it closes the energy balance by iteratively adjusting the surface temperature and surface energy fluxes to balance incoming radiation. It has a spatial resolution of 0.25°. It is available from 2001 to present, uses RFE and GDAS, 1 KM UMD land cover classification from AVHRR. RFE is from NOAA CPC derived from infrared and microwave observations blended with WMO GTS data, available from 2000 at 0.1° with 1-day latency (McNally *et al.* 2017).

Evaporation in VIC land surface models includes the weighted sum of canopy intercepted evaporation, transpiration from vegetation canopies and evaporation from bare soils based on the respective land surface coverage fractions. Canopy intercepted evapotranspiration is calculated by scaling Reference evapotranspiration (ReET) with canopy architectural resistance for humidity and aerodynamic resistance for heat and water. Canopy transpiration is calculated by scaling ReET with canopy resistance. Bare soil evaporation is calculated by scaling ReET by the current soil moisture conditions, wilting point soil moisture, and field capacity. Subtracting runoff and infiltration capacity at 0-10cm soil layer determines the surface runoff. Other parameters include soil texture and bulk density, soil characteristic curve parameters, monthly LAI and UMD land cover classification.

ERA-Interim is a reanalysis product describes atmosphere, land surface and/or ocean conditions using forecast models and assimilation systems. The models extrapolate archived and current observations to generate estimates of the atmosphere or land surface parameters such as air temperature soil moisture (Dee *et al.* 2011). ERA-Interim is a global atmospheric reanalysis product from 1979. ERA-Interim is produced through sequential assimilation in 6-hour cycles (Dorigo *et al.* 2012). Estimation of the evolving state of the atmosphere and its underlying surface in each cycle encompasses a combination of available observations and prior information from a forecast model. ERA-Interim uses the ECMWFs Integrated Forecast System to constrain the state evolution within each analysis window and update bias correction estimate parameters. IFS comprise of three coupled components for atmosphere, land surface and ocean waves. The ERA-Interim forecast is provided by TESSEL land surface scheme. It has a resolution of 79km (Berrisford *et al.* 2011). Optimal assimilation technique produces estimates of temperature and relative humidity by combining 2m relative humidity and temperature over land with estimates from the recent analysis (Douville *et al.* 2000, Decker *et al.* 2012). The model produces gridded products of volumetric soil moisture content at 0.75°.

6.3 Datasets

NASA Land Data Assimilation System consists of land surface models forced with observations. The FLDAS_NOAH and VIC daily dataset covering a duration of 2010-2016 was downloaded from the NASA Land Data Assimilation System website. The resolution of NOAH is 0.1° where the resolution of VIC is 0.25°. The NOAH and VIC netCDF files were extracted in MATLAB software to acquire daily means of soil moisture within the catchment. The NOAH 0.1° data was resampled to 0.25° to be spatially consistent with the VIC data.

ERA-Interim is a reanalysis data set based on the ECMWF Integrated Forecast System. The downloaded dataset had a spatial resolution of 0.75°, at a depth of 0-7cm (Albergel, de Rosnay, Balsamo, *et al.* 2012). The dataset was selected as it has a high global spatial consistency (Scipal *et al.* 2008, Dorigo *et al.* 2010). The daily readings downloaded were at 06:00 UTC. The data assessed covered the period January 2010 to December 2016. The data were resampled to a resolution of 0.25°.

SMOS reprocessed level 2 soil moisture version 5.51 (SM_REPR_MIR_SMUDP2) Microwave Imaging Radiometer (MIR) NetCDF dataset is distributed by ESA. The ascending mode was considered as it is least affected by surface heating. Night time are preferred because of equilibrium conditions of surface soil, canopy and near-surface air (Owe *et al.* 2008). SMOS was selected for the study as it is one of the sensors dedicated for remotely sensed soil moisture retrievals operating at L band and has a relatively long acquisition period (from 2010) as compared to SMAP which was launched in 2015. At low frequencies, L band is sensitive to water content of soils and can penetrate through vegetation, which is an important factor in selection of the dataset given that the Kilombero catchment comprises of varying land covers (Entekhabi *et al.* 2010, Kerr *et al.* 2012).

The modeled soil moisture products (FLDAS_NOAH, FLDAS_VIC, and ERA-Interim) are produced daily. The satellite-derived moisture from SMOS however, suffers from data gaps explained by the flight swaths of the satellite and hence weekly composites were generated from the SMOS dataset. Inevitably, this translated to computing the weekly averages for the modeled datasets in preparation for the data merging (Section 6.4.2)

16-day composites MODIS NDVI on Terra satellite was obtained on the LP DAAC from the year 2010 to 2016. 16-day MODIS NDVI composites were adopted for this study because they had the least variation in phenometric dates when compared to the 8 day composite MODIS in a study to assesses the influence of different phenometric extraction methods (Wessels *et al.* 2009). There are therefore 23 images available per year. These were preferred to reduce the atmospheric effects and clouds. Dummy data was generated for 2009 and 2017 as the phenology is not calculated in the first and last year of assessment as seasonality is only possible for n-1 center most seasons (Eklundh and Jönsson 2017). The data are in the Sinusoidal projection. The study area is covered by tile h21v09. On downloading the data sets, they were batch projected with the Modis Reprojection Tool to geographic coordinates.

TRMM is a multi-satellite precipitation analysis product. An assembled precipitation estimate is generated from measurements of passive microwave measurements on multiple satellites. Infrared measurements are used to estimate precipitation. These are then calibrated against the assembled precipitation. Infrared estimates fill the gaps of the assembled precipitation. Precipitation from ground-based radar, rain gauge, and disdrometer are used in the calibration of the multi-satellite product. TRMMs, spatial (0.25°) and temporal (3h) resolutions favor its use over other precipitation products (Koutsouris *et al.* 2016). The daily timescale was selected for the analysis.

Daily precipitation is generated from the 3-h TRMM Multi-Satellite Precipitation Analysis TMPA (3B42). NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) produce the data. Datasets used covered the period 2010 to 2016.

6.4 Methodology

The flow diagram (Figure 6.1) describing the research process is shown below;

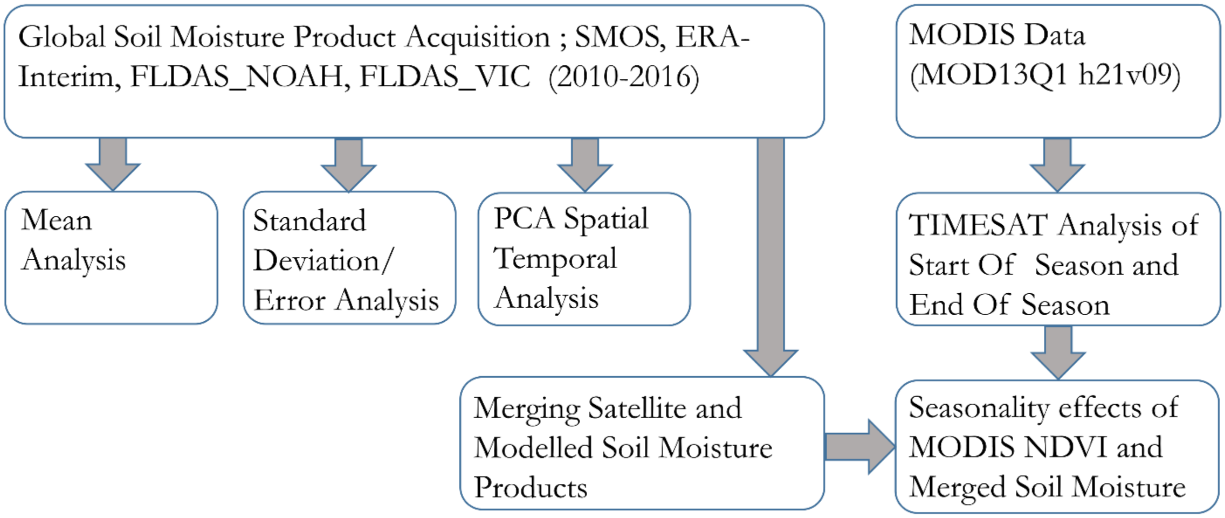


Figure 6.1: The data flow diagram in soil moisture assessment in the Kilombero catchment

6.4.1 Spatial-temporal averaging

The independent soil moisture products were resampled to a common grid of 0.25° to make the data comparable. The areal averages $\overline{sm}_i$, for each of the soil moisture products were computed as follows:

$$\overline{sm}_i = \frac{1}{n}\sum_{i=1}^{n} sm_i$$

$$where\ sm_i\ are\ the\ soil\ moisture\ values\ for\ each\ pixel\ per\ day,$$

$$n\ is\ the\ total\ number\ of\ pixels\ in\ the\ study\ site$$

6.4.2 Least square data merging

Least squares are an estimation theory that has been used in numerous studies and serves as a basis for modern data assimilation techniques. Following from (Yilmaz *et al.* 2012), given three independent realizations such as soil moisture, $S_x, S_y, and\ S_z$ of a variable along with their respective zero – mean errors $\epsilon_x, \epsilon_y, and\ \epsilon_z$ and error variances $\sigma_x^2, \sigma_y^2, and\ \sigma_z^2$. These realizations are represented as

$$S_x = \alpha S_t + \epsilon_x$$

$$S_y = \alpha S_t + \epsilon_y$$

$$S_z = \alpha S_t + \epsilon_z$$

Where S_t is the true value of the variable and α is a measure of the relation between these realizations and the assumed truth. The desired merged estimate S_m is obtained as

$$S_m = w_x S_x + w_y S_y + w_z S_z$$

Where w_x, $w_y$and w_z are the relative weights of $S_x, S_y, and\ S_z$, respectively. For the resulting merged product to be unbiased the sum of weights of all the products must be constrained to one ($w_x + w_y + w_z = 1$). The ultimate goal is to derive these weights as a function of the error variance of the three realizations and to find the error estimate of the merged product. The error estimate of the merged product is obtained as $\epsilon_m = S_m - \alpha S_t$ and the solution obtained from minimization of a selected cost function (J) in a mean squares sense. The cost function, J, is selected to be the error variance (σ_m^2) of the merged estimate in the form (Yilmaz *et al.* 2012).

$$J = \sigma_m^2 = w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + w_z^2 \sigma_z^2$$

$$J = \sigma_m^2 = w_x^2 \sigma_x^2 + (1 - w_x - w_z)^2 \sigma_y^2 + w_z^2 \sigma_z^2$$

Setting $\partial J / \partial w_z = 0$ and $\partial J / \partial w_x = 0$ in the above equation and solving for w_x,w_y and w_z, we obtain

$$w_x = \frac{\sigma_y^2 \sigma_z^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2}$$

$$w_y = \frac{\sigma_x^2 \sigma_z^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2}$$

$$w_z = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2}$$

The error variances of the merged product ($\sigma_x^2, \sigma_y^2, \sigma_z^2$) used in the computation of the weights was derived through triple collocation.

6.4.3 Triple collocation analysis (tca)

Triple collocation assesses errors in global geophysical products for example soil moisture. TC assumes that the unknown real soil moisture is linearly related to the estimated soil moisture through additive and multiplicative bias in three data sets. The advantage of TC is that it does not require reference data and it allows estimation of error variation in each of the three products.

The three data products (e.g. soil moisture) have independent error structures and no systematic biases a1, a2, a3 with N observations in the same geophysical process, and taking into account spatial and temporal collocation. Two approaches of estimating the error variances in the soil moisture products include the mean anomaly approach where the bias is incorporated into the mean and the TC is applied to the anomalies and the calibration constants approach where constants capturing the bias are introduced into the linear relationship between the three datasets and the hypothetical truth.

The mean anomaly approach decomposes the products into mean and anomaly components (Leroux *et al.* 2011).

$$a_1 = a + res_1$$

$$a_2 = a + res_2$$

$$a_3 = a + res_3$$

$$given\ res_1, res_2, res_3 are\ residual\ errors\ or\ anomalies\ relative\ to\ the\ hypothetical\ truth,$$

$$a\ has\ a\ variance\ of\ {\sigma_t}^2\ which\ is\ the\ true\ soil\ moisture\ variance\ with\ time$$

The difference in the data product value in relation to the residual errors is given by

$$a_1 - a_2 \ = res_1 - res_2$$

$$a_2 - a_3 = res_2 - res_3$$

$$a_3 - a_1 = res_3 - res_1$$

$$(a_1 - a_2)(a_2 - a_3) \ = (res_1 - res_2)(res_2 - res_3) = res_1 res_2 - res_1 res_3 - res_2 res_2 - res_2 res_3$$

$$(a_1 - a_2)(a_3 - a_1) \ = (res_1 - res_2)(res_3 - res_1) = res_1 res_3 - res_1 res_1 - res_2 res_3 - res_2 res_1$$

$$(a_2 - a_3)(a_3 - a_1) \ = (res_2 - res_3)(res_3 - res_1) = res_2 res_3 - res_2 res_1 - res_3 res_3 - res_3 res_1$$

Assuming the residuals are uncorrelated and mutually independent of the three datasets, the covariance terms become 0. The error variance is given by

$${\sigma_2}^2 = \langle res_2 res_2 \rangle = (a_1 - a_2)(a_2 - a_3)$$

$${\sigma_1}^2 = \langle res_1 res_1 \rangle = (a_1 - a_2)(a_3 - a_1)$$

$${\sigma_3}^2 = \langle res_3 res_3 \rangle = (a_2 - a_3)(a_3 - a_1)$$

$${\sigma_1}^2, {\sigma_2}^2, {\sigma_3}^2\ are\ error\ variances$$

In the calibration constant approach,

$$a_1 = \ \alpha_1 a_t + \epsilon_1,$$

$$a_2 = \ \alpha_2 a_t + \epsilon_2,$$

$$a_3 = \ \alpha_3 a_t + \epsilon_3,$$

Where $a_1, a_2, and\ a_3$ are soil moisture anomalies related to the true anomaly a_t through α_1, α_2 and α_3. In general, the products are expected to have different relationships with the truth i.e., $\alpha_1 \neq \alpha_2 \neq \alpha_3$ thus to ensure each data set has a consistent relationship with the truth, they are rescaled as below (Yilmaz *et al.* 2012);

$$a_1^* = \alpha a_t + \epsilon_1^*$$

$$a_2^* = \alpha a_t + \epsilon_2^*$$

$$a_3^* = \alpha a_t + \epsilon_3^*$$

Where a_1^*,a_2^* and a_3^* are the rescaled realizations and ϵ_1^*,ϵ_2^* and ϵ_3^* are the relative errors of the realizations with variances $\sigma_1^{*2}, \sigma_2^{*2}, and\ \sigma_3^{*2}$.

By arbitrarily setting any of the data sets as reference (e.g., by setting $\alpha = \alpha_1$) and assuming the error covariance between products are zero, and representativeness errors described in (Stoffelen 1998) are zero, these factors can be found as (Yilmaze et al 2012, Leroux 2011);

$$m_1 = 1$$

$$m_2 = \frac{\overline{a_1^* a_3^*}}{\overline{a_2^* a_3^*}} = \frac{\alpha_1}{\alpha_2}$$

$$m_3 = \frac{\overline{a_1^* a_2^*}}{\overline{a_3^* a_2^*}} = \frac{\alpha_1}{\alpha_3}$$

The rescaled values are related to the original estimates as $a_1^* = m_1 a_1,\ a_2^* = m_2 a_2, and\ a_3^* = m_3 a_3$ where $m_1, m_2, and\ m_1$ are the scaling factors. Where overbar indicate averaging in time. Assuming the errors in the products are independent from each other and from the truth, and assuming a mutual linear relationship between these estimates and the true soil moisture, the final error variances of the rescaled realization are found as (Leroux *et al.* 2011);

$$\sigma_1^{*2} = \overline{(a_1^* - a_2^*)(a_1^* - a_3^*)}$$

$$\sigma_2^{*2} = \overline{(aS_2^* - a_1^*)(a_2^* - a_3^*)}$$

$$\sigma_3^{*2} = \overline{(a_3^* - a_1^*)(a_3^* - a_2^*)}$$

The error variances are used to compute the weights required for least squares data merging. Different datasets result in different error variances. TC provides relative accuracies describing how a product compares to another product. Absolute values of error variances are acquired in cases where a reference dataset is available.

6.4.4 Principal component analysis / empirical orthogonal functions

PCA was used to analyze soil moisture spatial-temporal variability in the Kilombero catchment. EOF/PCA reduces the dimensionality of data to new variables that are linear combinations of the original ones (Lorenz 1956, Preisendorfer *et al.* 1988, Jolliffe 2002). The new data is representative of the maximum variability contained in the original data (Björnsson and Venegas 1997, Wilks 2006). Spatial patterns of variability, time variation, and importance of the variation patterns is an output of PCA.

Given a gridded dataset composed of a space-time field $X\ (t, s)$ with X being soil moisture at time t in spatial position s. The value of soil moisture at time t_i at location s_j is denoted as x_{ij} for i=1,…, n and j=1,…p. The observed data is stored in a matrix X.

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

$$where\ t = 1, \dots .n\ represents\ the\ soil\ moisture\ value\ at\ time\ t$$

There are n maps each being p points long. Each map is arranged into a row such that the size of matrix X is n by p. each of the p columns is a time series for a given location. EOF analysis is performed with X as the data matrix.

The time average of soil moisture at the ith spatial grid point is given by;

$$\bar{x}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ki}$$

The mean is removed from each of the p time series in matrix X.

The covariance matrix of X is calculated as

$$R = X^t X$$

The Eigenvalues and vectors are then computed

$$RC = C\Lambda$$

$$where\ \Lambda\ is\ a\ diagonal\ matrix\ containing\ the\ eigen\ values\ \lambda_i, of\ R,$$

$$C\ contains\ the\ eigen\ vectors\ c, of\ R$$

Each Eigenvalue λ_i has a corresponding Eigen vector c_i. The Eigen vectors are the maps / EOFs. The Eigen values on the other hand give a measure of fraction of total variance in R explained by the mode.

The Eigenvector matrix C is such that $C^tC = CC^t = I, with\ I\ being\ the\ identity\ matrix.$ In other words, Eigen vectors are orthogonal to each other indicating that they are uncorrelated to each other.

The time variability of the field (e.g. soil moisture) is calculated by

$$\overrightarrow{a_1} = X\overrightarrow{c_1}$$

The n components of vector $\overrightarrow{a_1}$ are the projections of the maps in X on EOF1 and the vector is a time series for evolution of EOF1. For every EOF_j there is a corresponding$\overrightarrow{a_j}$. $\overrightarrow{a_j}$ are referred to as the principal components.

In terms of PCA/EOF, the data matrix is given as

$$X = \sum_{j=1}^{p} \overrightarrow{a_j}(EOF_j)$$

The first N Eigenvectors capture the dynamic behavior of the data whereas the other Eigenvectors with smaller values are due to random noise (Björnsson and Venegas 1997, Hannachi *et al.* 2007).

In matrix form, PCA decomposes soil moisture matrix X into

$$X = B^TQ$$

Where B^T(standardized by diving by its standard deviation) is the temporal variability, Q (scaled by multiplying by standard deviation of B) is the spatial variability in m^3m^{-3}, and T denotes matrix transposition.

6.4.5 Phenology extraction

Seasonality describes annually occurring seasonal vegetation patterns. Fitted functions are used in extraction of seasonality parameters as satellite-derived observations have noise that introduces uncertainties and subsequent instability of measurements. SOS (EOS) is the day of year when the left (right) part of the fitted curve reaches a specified fraction of the amplitude. LOS is time from start to end of season.

Phenological parameters SOS, EOS, and LOS were extracted to assess their correlation with soil moisture and hence the potential application of soil moisture in determining the SOS in food production. Definitions are derived from TIMESAT seasonal amplitude (region between minimum and maximum value of each of the individual seasons) (Figure 6.2).

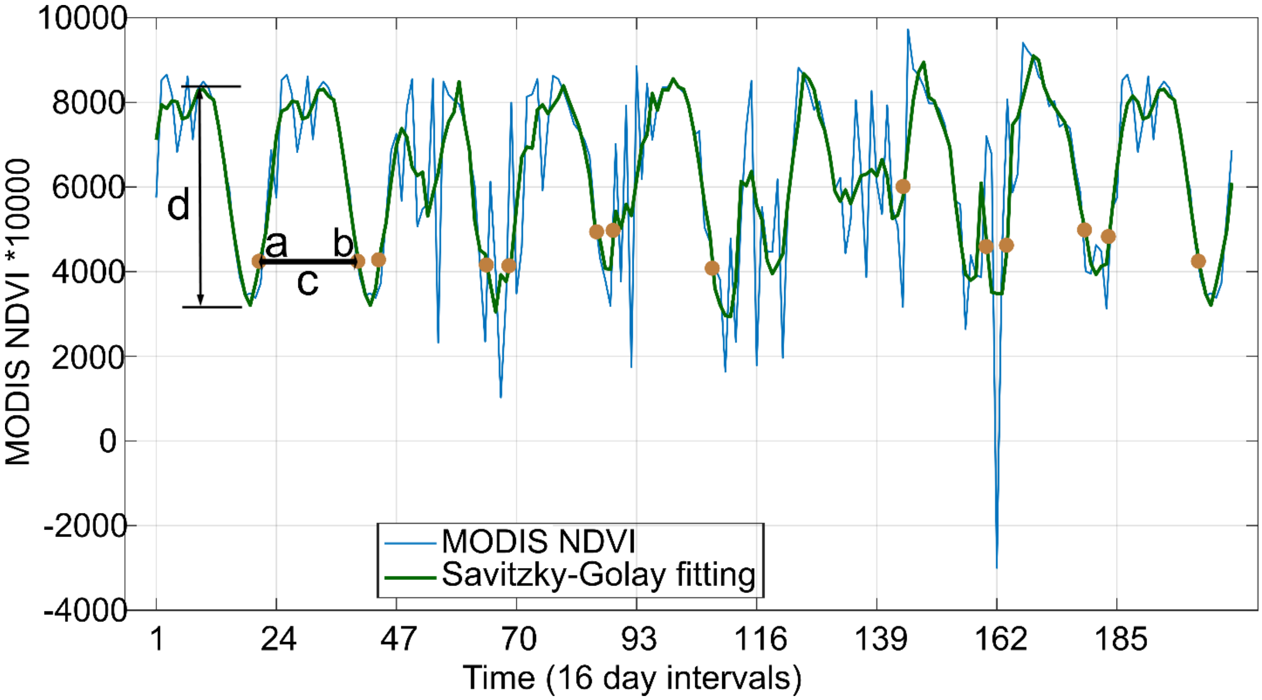


Figure 6.2: TIMESAT derived time series of the NDVI of a pixel. Raw NDVI data is shown against smoothed NDVI. Seasonality parameters include a; the start of season (SOS), b; end of season (EOS), c; length of season (LOS) and d; the amplitude and are determined from the Savitzky-Golay fitted function.

TIMESAT was used to calculate phenology metrics from the MODIS NDVI implemented in TIMESAT's graphical user interface in Matlab. The steps involved include computing the trend in the data using the Seasonal Trend Decomposition by Loess (STL), filtering to remove outliers and pixels with few data, calculation of a seasonal fit, smoothing the data with adaptive Savitzky-Golay filter to remove the noisy signals, extracting the seasonal parameters based on the smoothed data and finally generation of output. The output includes the start of season time, end of season time, length of season time of the middle of season, amplitude among others. The data above can be viewed as images. The data fitted well and thus no weights were assigned to the fitting function. The smoothing method adopted, the Savitzky-Golay closely follows the curve of the fitted data and is able to capture increases and decreases in the data suitable for locations with sudden changes in vegetation reflectance (Eklundh and Jönsson 2017). The selection of the threshold for the start and end of season was an iterative process based on determining the fractions of seasonal amplitude with the aim being capturing the seasons as known from the apriori knowledge of the study area. Seasonal amplitude is defined between the base level and maximum value for each individual season. A user defined threshold of 20% of the seasonal amplitude was adopted in determining the vegetation phenometrics following (Wessels *et al.* 2011). The start of season was defined as the date at which the seasonal amplitude on the left side of the curve is at 20%. End of season is defined as the point at which the right edge of the curve declines to 20% of the amplitude. TIMESAT assigns null values to pixels where no seasons are detected, where values are missing, where there is a failure in fitting function or the amplitude is low.

6.5 Results and Discussion

6.5.1 Soil moisture characteristics

The basin spatial average soil moisture from FLDAS_NOAH, FLDAS_VIC, and ERA-Interim showed seasonal variability fairly well (Figure 6.3). Soil moisture peaked towards the end of the year reaching a maximum in April-May. This represents the beginning of the short rains in December-January leading into the long rains in March-May. ERA-Interim had highest estimates, followed by FLDAS_VIC, and finally FLDAS_NOAH.

In addition to the spatial averages, the study further considered the temporal weekly averages from 2010 to 2016 of the soil moisture products (Figure 6.3). The products had largely uniform means across the basin with ERA-Interim having higher values (> than 0.25 m^3/m^3), closely followed by FLDAS_VIC and finally FLDAS_NOAH having values of less than 0.25. This is consistent with observations from the spatial averages, where ERA recorded highest spatial means Figure 6.3.

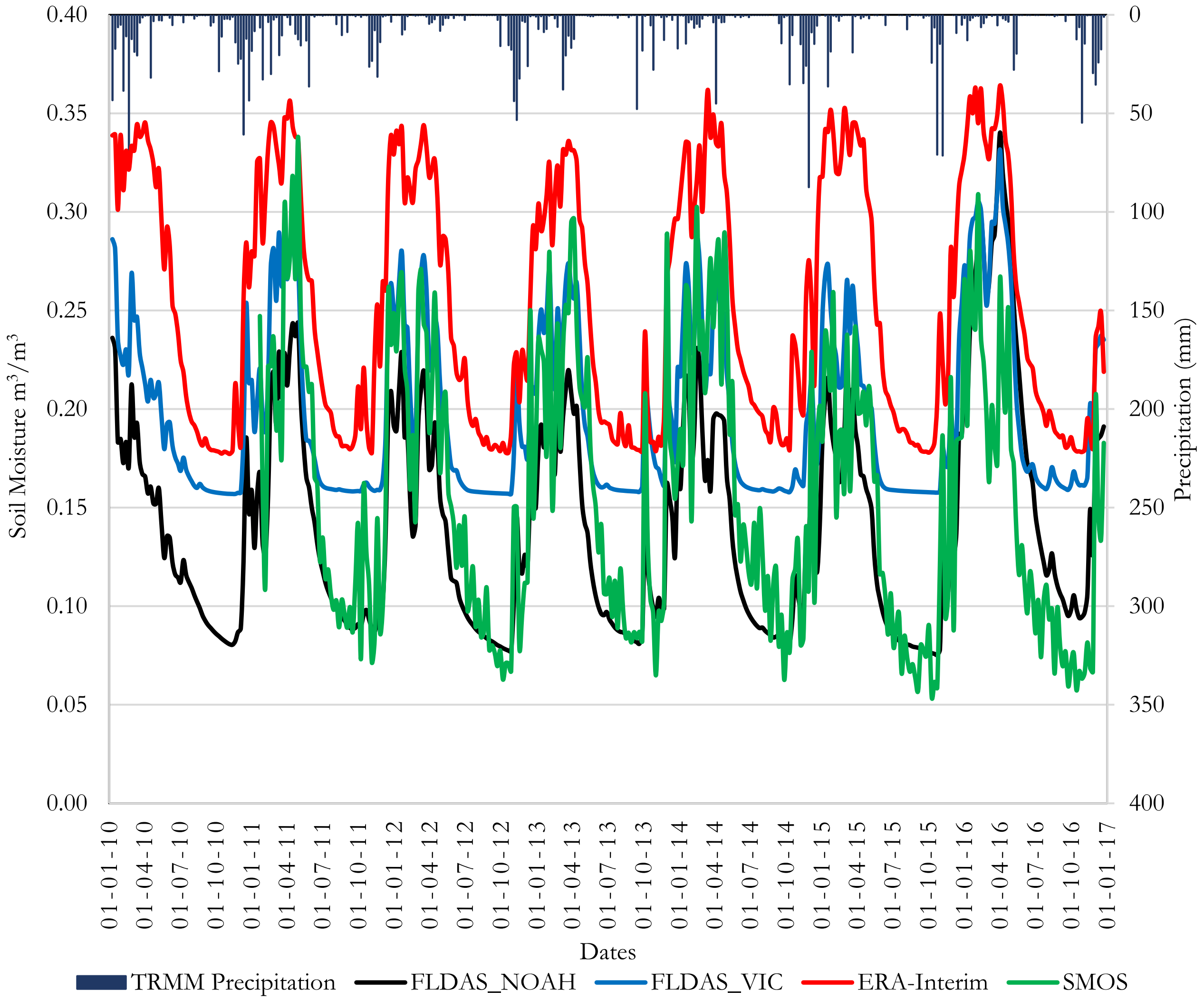
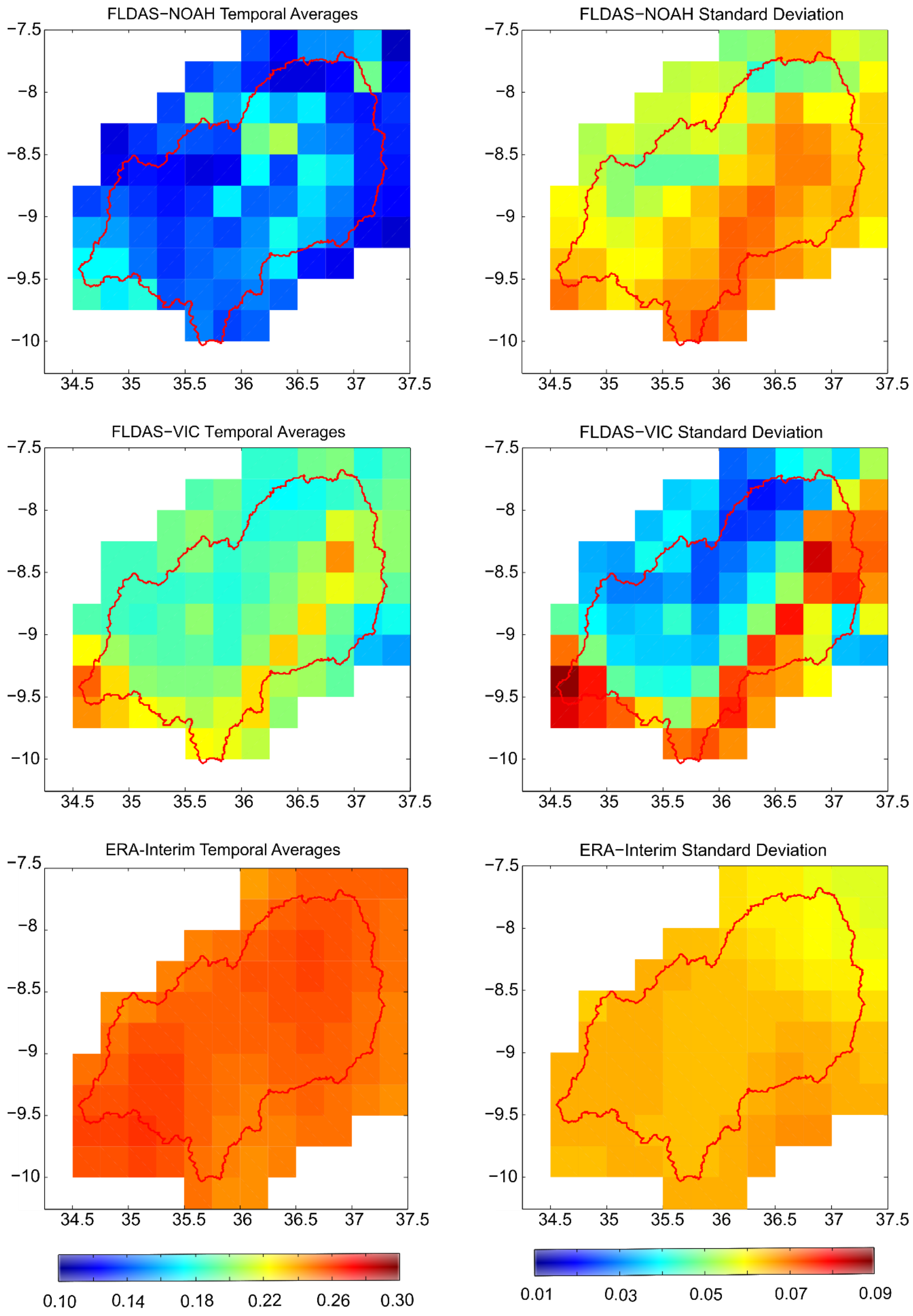


Figure 6.3: Weekly soil moisture spatial averages over the Kilombero catchment for the modeled and satellite-derived global soil moisture products. Cumulative weekly TRMM precipitation in mm is shown.

The soil moisture patterns are similar but not identical. Varying algorithms in computation of the soil moisture explain the differences in values. For example, ECOCLIMAP land cover, Global Land Cover Characteristics from AVHRR, University of Maryland land cover and IGBP land cover are inputs in soil moisture derivation for SMOS, ERA-Interim, FLDAS_VIC and FLDAS_NOAH respectively (Kerr *et al.* 2012, McNally *et al.* 2017).



FLDAS−NOAH Temporal Averages

FLDAS−NOAH Standard Deviation

FLDAS−VIC Temporal Averages

FLDAS−VIC Standard Deviation

ERA−Interim Temporal Averages

ERA−Interim Standard Deviation

Figure 6.4: Weekly soil moisture temporal averages and standard deviation for the modeled products from 2010 to 2016

The study further used the standard deviation of the soil moisture products to understand the uncertainty distribution over the basin (Figure 6.4). The spatial distribution of uncertainties in FLDAS_NOAH closely mirrored those of FLDAS_VIC where the eastern side, covered by the Mbarika Mountains, has higher uncertainties (standard deviations) compared to the western side. This similarity could be attributed to the fact that both the models (FLDAS_NOAH and FLDAS_VIC) are forced by same precipitation product; African Rainfall Estimation version 2.0 (RFE2;(McNally *et al.* 2017)). ERA-Interim, on the other hand, had an almost uniform distribution of uncertainties except for the northeastern tip of the basin with relatively lower uncertainties.

The SMOS weekly temporal mean and standard deviation portrayed low values in the Northern and Southern side where the Udzugwa Mountains and Mbarika Highlands are located (Figure 6.5). The low values are explained by the inconsistent coverages of the SMOS sensors. Satellite-derived records often have data gaps due to the swaths followed by the sensor as it moves around the earth acquiring data (Anderson *et al.* 2012). Moreover, passive sensors are also affected by dense vegetation cover which is observed in the mountainous regions of the Kilombero catchment. Appendix C3 shows the SMOS coverage over the Kilombero catchment.

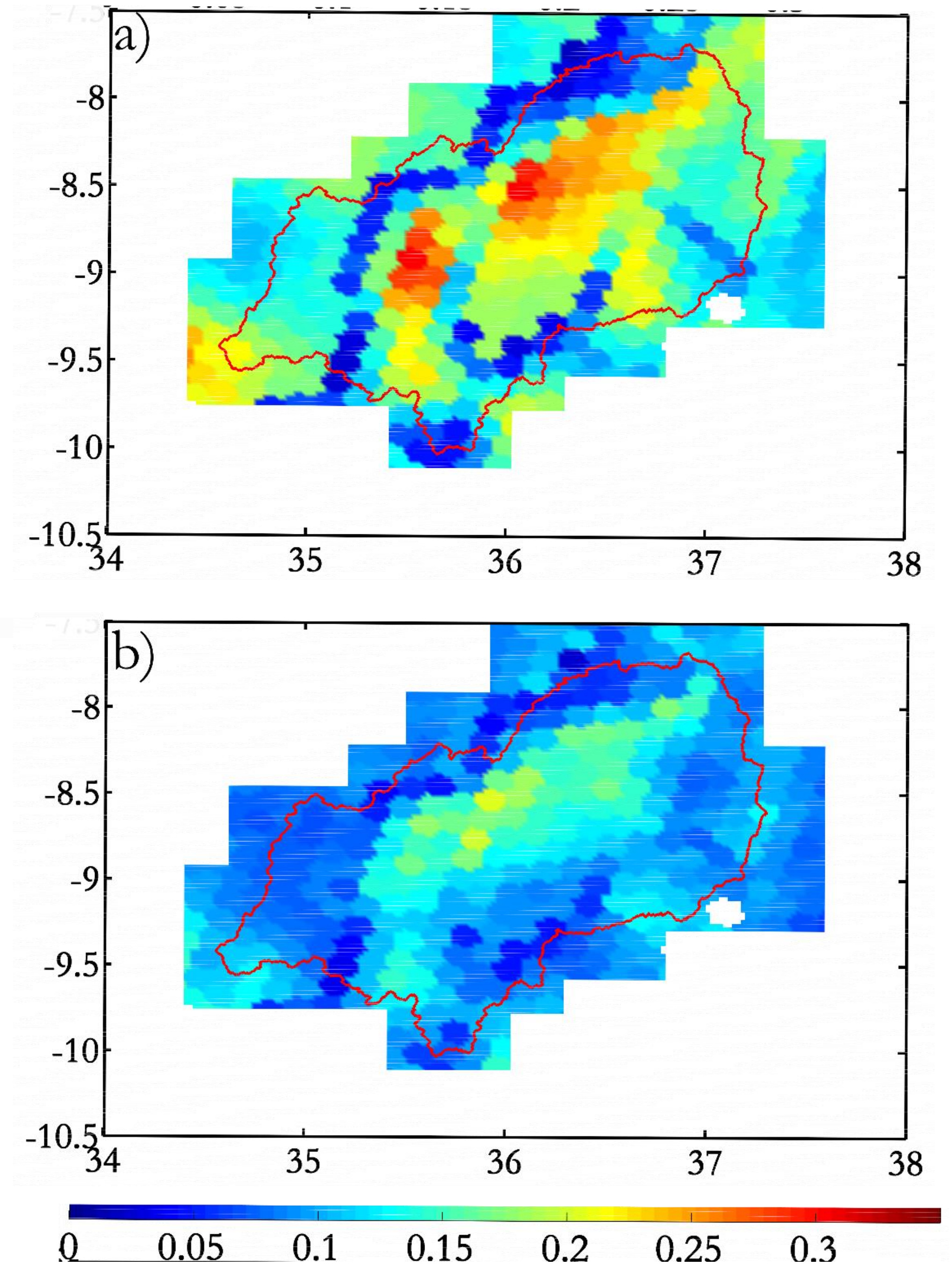


Figure 6.5: Weekly soil moisture temporal averages (a) and standard deviation (b) for SMOS from 2011 to 2016. The low values to the north, south and on the right side of the floodplain (covered by montane and closed woodland) are due to data gaps, common in densely vegetated areas.

6.5.2 Spatial and temporal variability

Finally, the study analyzed the spatial-temporal variability of the soil moisture products through principal component analysis. The PCA decomposition of the weekly soil moisture products resulted in a single dominant mode across all the products (Figure 6.6). This mode explained 84.65%, 77.84%, and 96.78% of the variability in FLAS_NOAH, FLDAS_VIC, and ERA-Interim, respectively (Table 6.2). The spatial variability (Figure 6.6) closely mirrored the spatial distribution of uncertainties (standard deviations) observed in Figure 6.5, while the temporal variabilities resemble the spatial averages in Figure 6.4. The PCA for SMOS was not performed because the data set suffered from data gaps. Satellite derived gaps are a product of overpass repeat cycle resulting to swaths missing data on a regular cycle (Anderson *et al.* 2012, McNally *et al.* 2016).

From the temporal variabilities, it is clear that the region has a bimodal moisture regime in response to bimodal rainfall (March-May and December – January). The temporal variability clearly depicts low moisture in 2010 especially from FLDAS_NOAH and FLDAS_VIC, this is a result of drought reported over the East Africa region in 2009/2010 (Masih *et al.* 2014). The temporal variabilities show relatively high moisture values over 2015/2016 (OND/MAM) duration, a reflection of 2015 El Niño rains. The other less dominant spatial-temporal modes are shown in the Appendix C 1.

Table 6.2: Percentage variabilities explained

%age variability explained	PC 1	PC 2	PC 3	PC 4	PC 5
FLDAS_NOAH	84.65	3.38	2.37	1.47	1.12
FLDAS_VIC	77.84	6.53	2.33	1.78	1.16
ERA_INTERIM	96.78	1.82	0.64	0.34	0.16

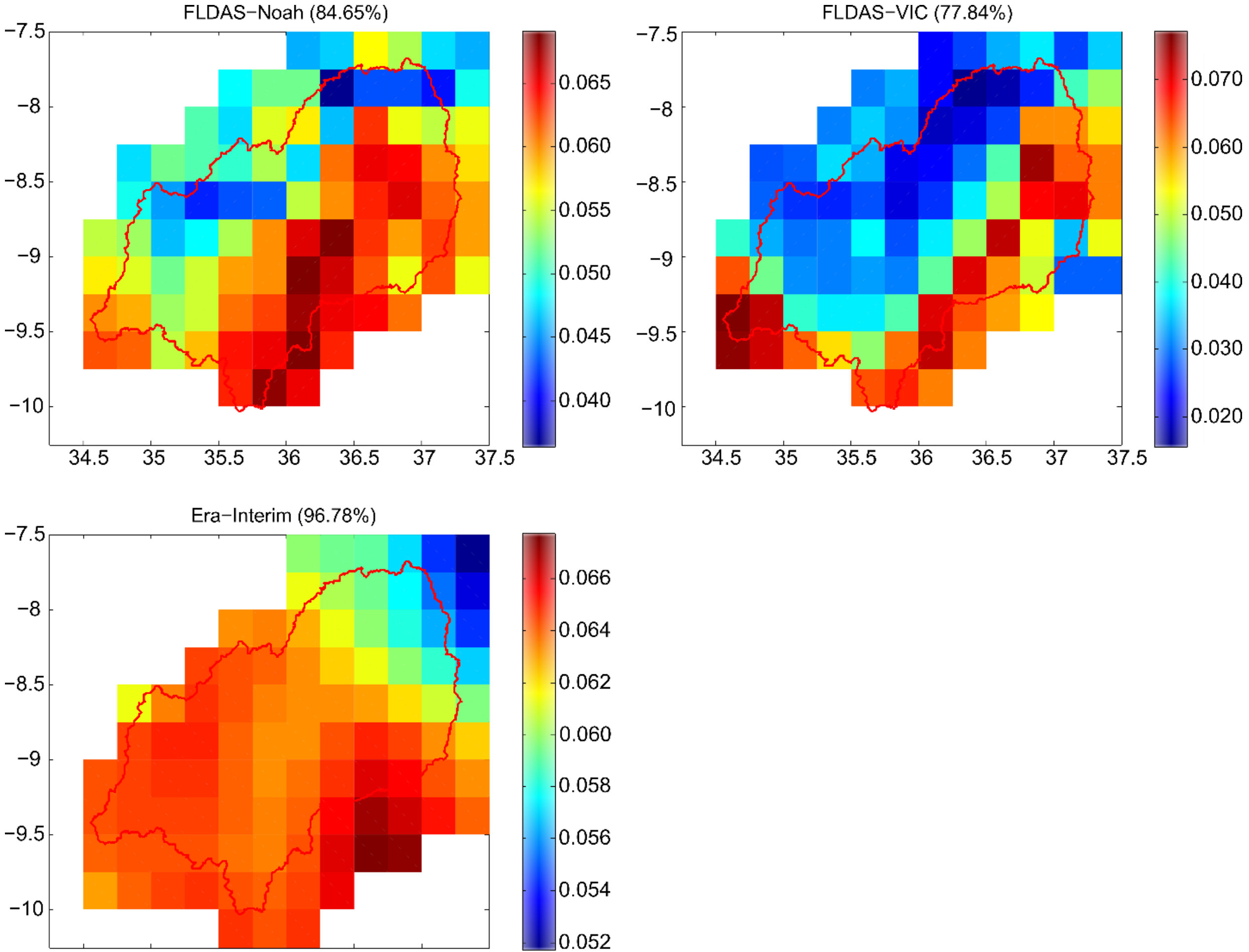


Figure 6.6: Spatial variabilities of EOF 1

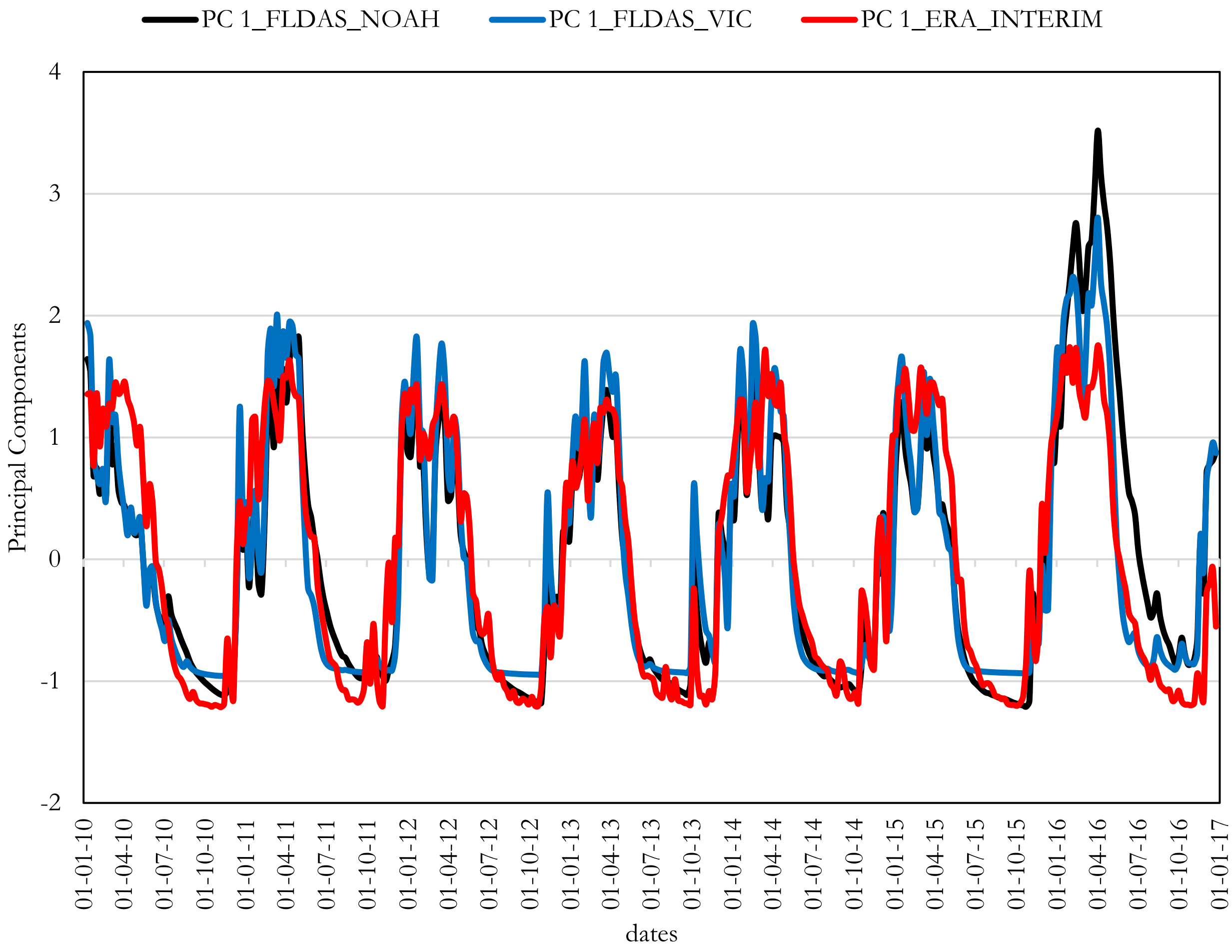


Figure 6.7: Moisture spatial-temporal variability from Principal Component 1

6.5.3 Merged products

The lack of in situ soil moisture over the study area implies that it is not possible to evaluate how well the model and satellite-based soil moisture products (FLAS_NOAH, FLDAS_VIC, ERA-Interim, and SMOS) captures the moisture of the region (Anderson *et al.* 2012). Due to varying computation algorithms, the moisture datasets have differences in soil moisture values and therefore merging the independent datasets to a consensus product reduces the systematic differences in the input datasets. Due to the critical role of soil moisture in agriculture and the fact that it's the only variable that indicates the plant available water between precipitation deficit and plant reaction. The chapter determined the rise and decline of soil moisture that can be linked to the start and end of planting season (AghaKouchak *et al.* 2015). Moreover, factors like vegetation cover also impede the proper presentation of soil moisture by the satellites. Therefore, objective merging of models (FLDAS_NOAH, FLDAS_VIC, and ERA-Interim) and satellite-derived (SMOS) soil moisture products based on their relative uncertainties provides the most suitable and reliable soil moisture product for the basin. In addition, based on spatial and temporal means, and spatial-temporal variability of soil moisture over the basin (Figure 6.7), it is plausible to work with basin spatial averages as the basin had largely uniform variation in the soil moisture products.

The selection of the modeled and passive product was such that the strengths and limitations of the methods are complimented in generation of the merged product. SMOS was selected as it is the current operational satellite dedicated to soil moisture measurements with the longest temporal coverage during the study period (2010 to 2016). Moreover, passive satellite moisture sensors are sensitive to surface soil moisture and are attenuated by vegetation coverage at shorter wavelengths. The advantage of SMOS is its operation in longer wavelength L band thus minimizing the attenuation of the microwave signals. A limitation to satellite-based moisture is the presence of data gaps due to the overpass swaths of the satellites and presence of dense vegetation.

The modeled datasets (FLDAS_NOAH, FLDAS_VIC, and ERA-Interim) were selected as they were daily products with the longest temporal coverage running from 2010 to 2016 and ensured consistency with the SMOS dataset. Modeled products have the advantage of consistent coverage, as they use a combination of existing data and newly acquired datasets in models for generation of the soil moisture. The models therefore are at an advantage of applying hydrological, climatic and land cover memory in computing soil moisture. FLDAS_NOAH and FLDAS_VIC were selected as they are forced by FEWSNET rainfall specifically designed for food security assessment in Africa and other data scarce regions while ERA-Interim has a high global spatial consistency (Scipal *et al.* 2008, Dorigo *et al.* 2010, McNally *et al.* 2017).

The assumptions made in generation of merged products by application of triple collocation are that the independent soil moisture products in consideration should be uncorrelated and should have a linear relationship (Scipal *et al.* 2008). Merged products one constituted FLDAS_NOAH, ERA-Interim, and SMOS while Merge 2 constituted FLDAS_VIC, ERA-Interim, and SMOS. The retrieval algorithms are different for the three sets of data in the two merged products thus the concept of uncorrelation is met. The linearity concept was tested (Figure 6.8), the p-value was less than 0.05 drawing a conclusion of significant correlation of the independent sets of data. The error variances were first computed followed by computation of weights that are assigned to each of the products to determine the contribution to the overall merged products (Table 6.3). Merged products are from the 2nd week of 2011 to 2016, making a total of 310 weeks. The reduction in the number of weeks was due to the time delay in the availability of SMOS soil moisture. To obtain a reliable estimation of the error variance, at least 100 observations are required for each of the three datasets (Dorigo *et al.* 2010, Leroux *et al.* 2011).

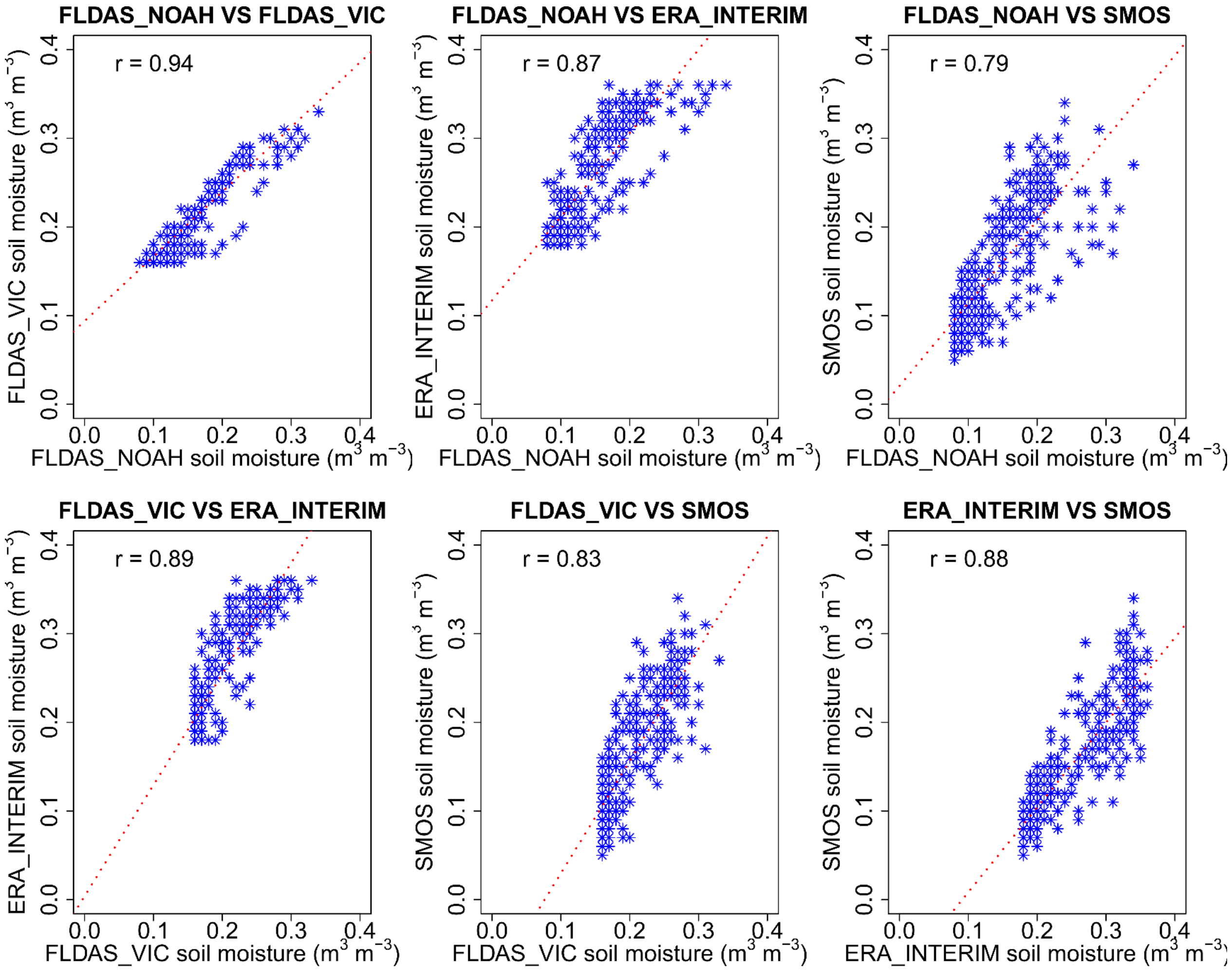


Figure 6.8: Linear relationship of the independent global soil moisture products used in generation of consensus merged soil moisture products.

In merge 1 and 2, ERA-Interim had the highest weight assigned, 0.77 and 0.52 respectively (Table 6.3). In both products, the satellite-derived soil moisture (SMOS) had the lowest weights assigned since it had the highest error variation. The microwave signals from passive sensors such as SMOS are sensitive to attenuation by vegetation cover explaining the high TC error values and subsequent low weight assignment for the consensus merged soil moisture products (Anderson *et al.* 2012). The dominant contributor in both merge 1 and 2 products was ERA-Interim reanalysis product which had a weight greater than 50%. The merged product 1 and 2 are shown in Figure 6.9.

Table 6.3: Average merging weights and TCA error variances for the Kilombero catchment

	Retrieval	Average Merging Weights	Average TCA error variances
Merge 1	FLDAS_NOAH	0.128	7.1×10^{-4}
	ERA-Interim	0.771	1.2×10^{-4}
	SMOS	0.102	9.0×10^{-4}
Merge 2	FLDAS_VIC	0.353	3.0×10^{-4}
	ERA-Interim	0.516	2.1×10^{-4}
	SMOS	0.131	8.1×10^{-4}

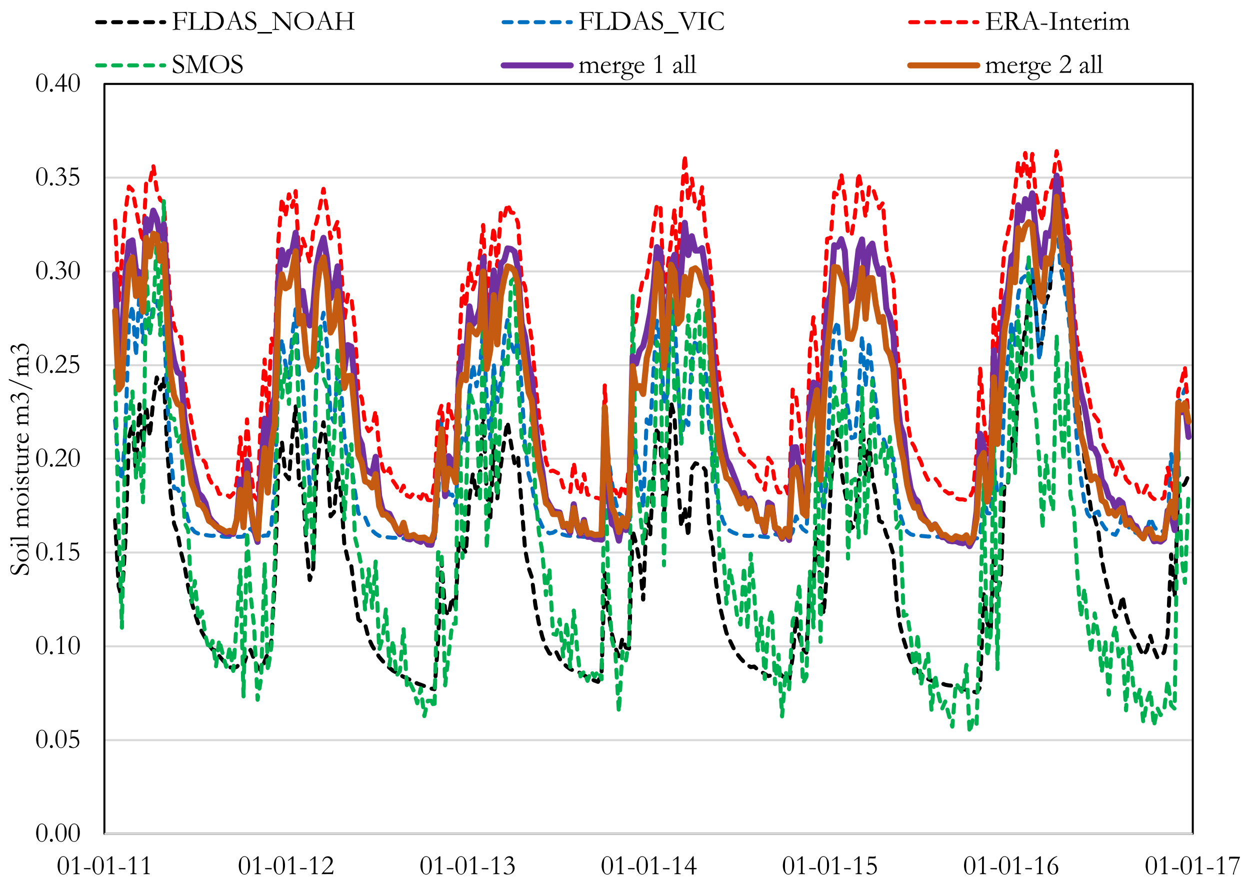


Figure 6.9: Soil moisture independent datasets (FLDAS_NOAH, FLDAS_VIC, ERA-Interim and SMOS), merge 1 (FLDAS_NOAH, ERA-Interim and SMOS) and merge 2 (FLDAS_VIC, ERA-Interim and SMOS) products

Interannual variability pattern over the Kilombero catchment is maintained for the independent soil moisture products and the two merged products. This indicates that the merged products sufficiently capture interannual timescale processes and are not affected by discrepancies between the datasets (Anderson *et al.* 2012).

The variation of the merged products from their original values is compared through a correlation analysis (Figure 6.10). The two merged soil moisture products had similar coefficient of determination and correlation. The coefficient of determination and correlation coefficient was 0.54 and 0.73 for merge 1 and 0.6 and 0.77 for merge 2. The similarities in the performance of merge 1 and 2 are explained by similarity in the RFE forcing in the retrieval concepts of FLDAS_NOAH and FLDAS_VIC soil moisture. However, the slight variations in their performance is attributed to the varying input parameters in generation of the global moisture FLDAS_NOAH and VIC.

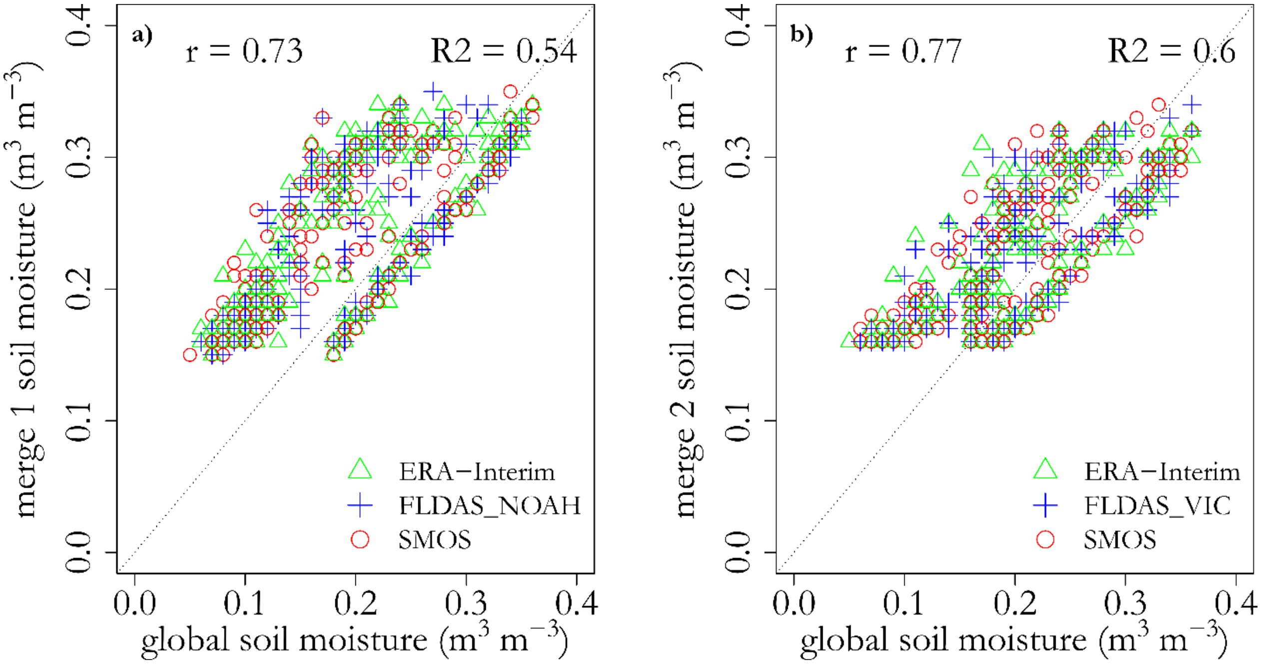


Figure 6.10: Relationship of merged soil moisture products a) 1 and b) 2 and the respective moisture products used in generation of the merged products.

6.5.4 MODIS NDVI seasonality assessment

Start, end, and length of the season for seven-year data assessed (2010-2016) are shown in Figure 6.12 to 6.14. However, dummy data was added as 2009 and 2017 to capture the seasonality of the years at the edge (2010 and 2016). The discussion focuses on data from 2010 to 2016. A land cover map is included in Figure 6.10 to understand the classes existing in the Kilombero catchment and link them to SOS and EOS.

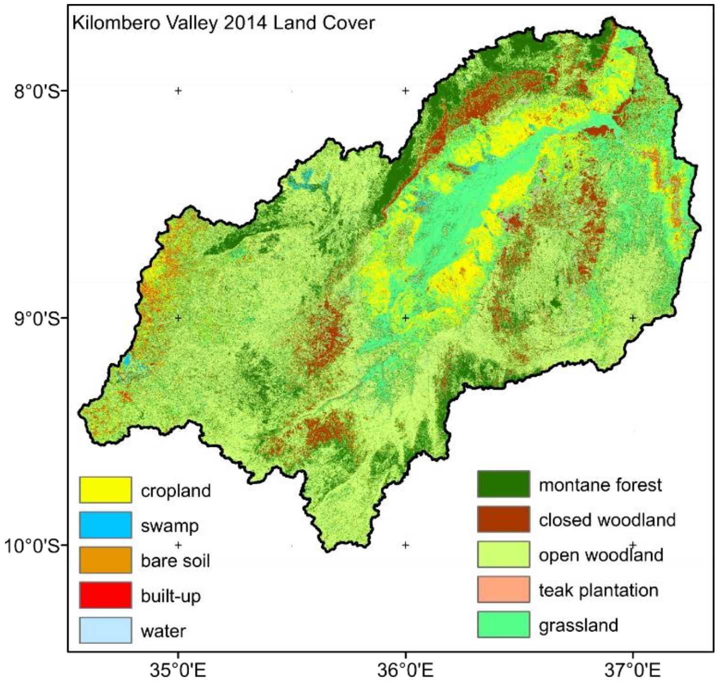


Figure 6.11: The Kilombero Valley land cover map

White areas (Figure 6.12 to 6.14) indicate pixels with missing data or where data could not be processed likely due to pixels with insufficient data in the time coverage of the seasonality computation (Eklundh and Jönsson 2017). Mahenge highlands are located to the south and the Udzungwa Mountains to the north of the catchment (Figure 6.11). These regions experience heavy presence of clouds. Clouds affect the data values recorded by the optical sensor and thus in the computations of the seasons, such pixels (outliers) are masked out. This could explain the pattern of missing values along the edge of the catchment to the north and south. The regions masked out are noted to be in high elevations and mainly covered by closed woodland and montane forests.

The phenology maps show the DOY when the season starts and ends, from which the length of season is retrieved. The Kilombero catchment, has a unimodal phenology pattern, consistent with a study by Vrieling who assessed the length of season from NDVI time series for the whole of Africa (Vrieling *et al.* 2013).

The analysis from TIMESAT software included maps showing start of season (SOS), end of season (EOS) and length of season (LOS). In order to determine the weeks within which the seasonality changes occur, the values of the NDVI pixels were extracted, and the period with the highest frequency of pixels for the SOS, EOS, and LOS were selected as the ultimate seasonality occurrence periods.

In the year 2010-2011, the start of season was observed to be between day 321-353 (mid-November to mid-December 2010) while the end of season was between day 225 to 257 (mid-August to mid-September 2011). In the year 2011-2012, SOS was observed to be between day 305 – 337 (November) while the EOS was between day 225 to 257 (mid-August to mid-September) (Figure 6.12).

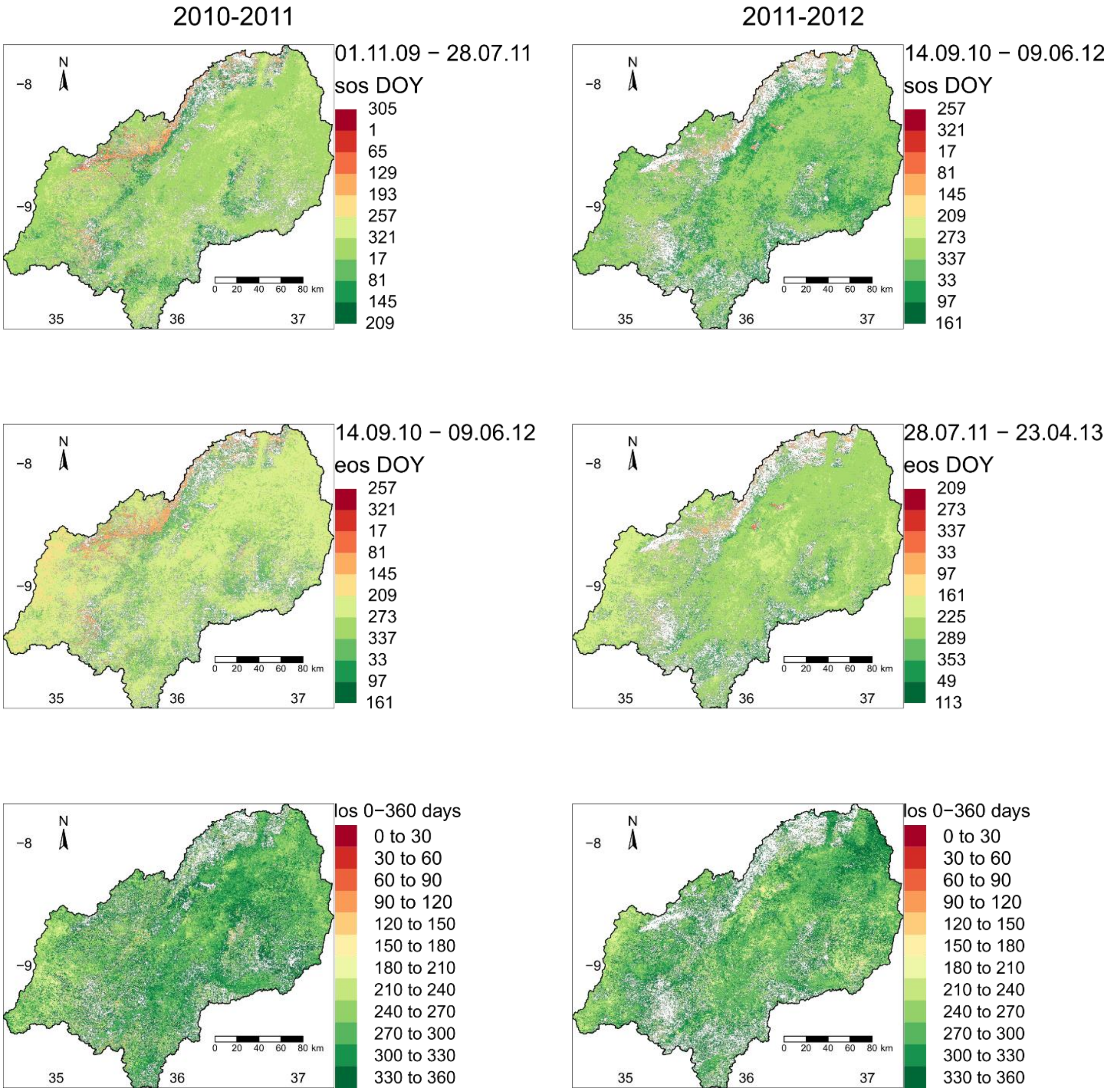


Figure 6.12: NDVI phenology for the Kilombero catchment from 2010 to 2012

In the year 2012-2013, the start of season was observed to be between day 289-321 (mid-October to mid-November 2012) while the end of season was between day 241 to 273 (September 2013). In the year 2013-2014, SOS was observed to be between day 305 – 337 (November) while the EOS was between day 225 to 257 (mid-August to mid-September) (Figure 6.12).

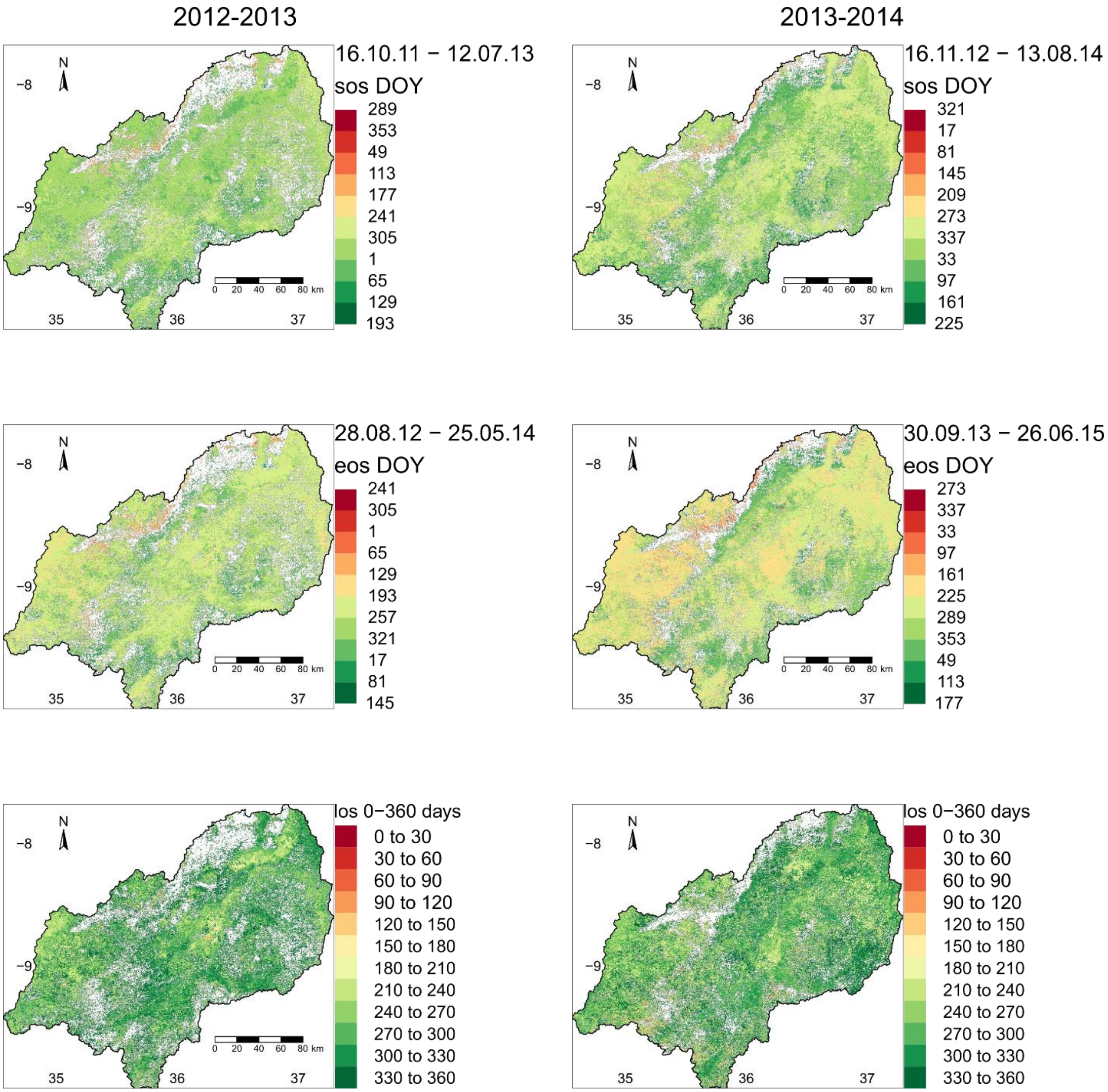


Figure 6.13: NDVI phenology for the Kilombero catchment from 2012 to 2014

In the year 2014-2015, the start of season was observed to be between day 289-321 (mid-October to mid-November 2014) while the end of season was between day 241 to 273 (September 2015). In the year 2015-

2016, SOS was observed to be between day 305 – 337 (November) while the EOS was between day 241 to 273 (September) (Figure 6.13).

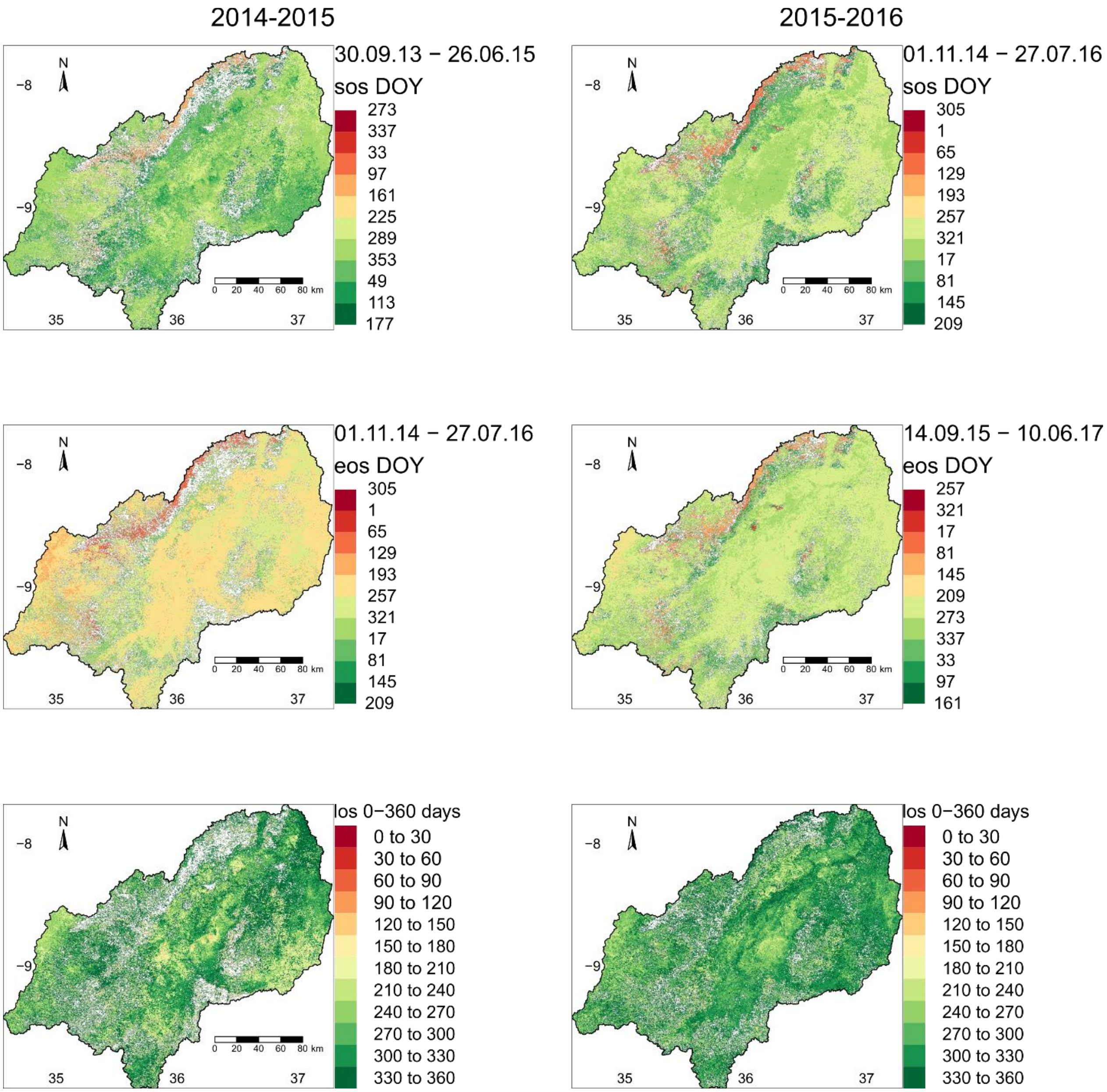


Figure 6.14: NDVI phenology for the Kilombero catchment from 2014 to 2016

In summary, the average SOS for the Kilombero catchment mainly occurs between November and December whereas the EOS occurs between August to September. On average, the LOS is therefore 290 days. Temporal averages of mean NDVI over the Kilombero catchment are shown in Figure 6.14. Included are the average distributions of start and end of season dates.

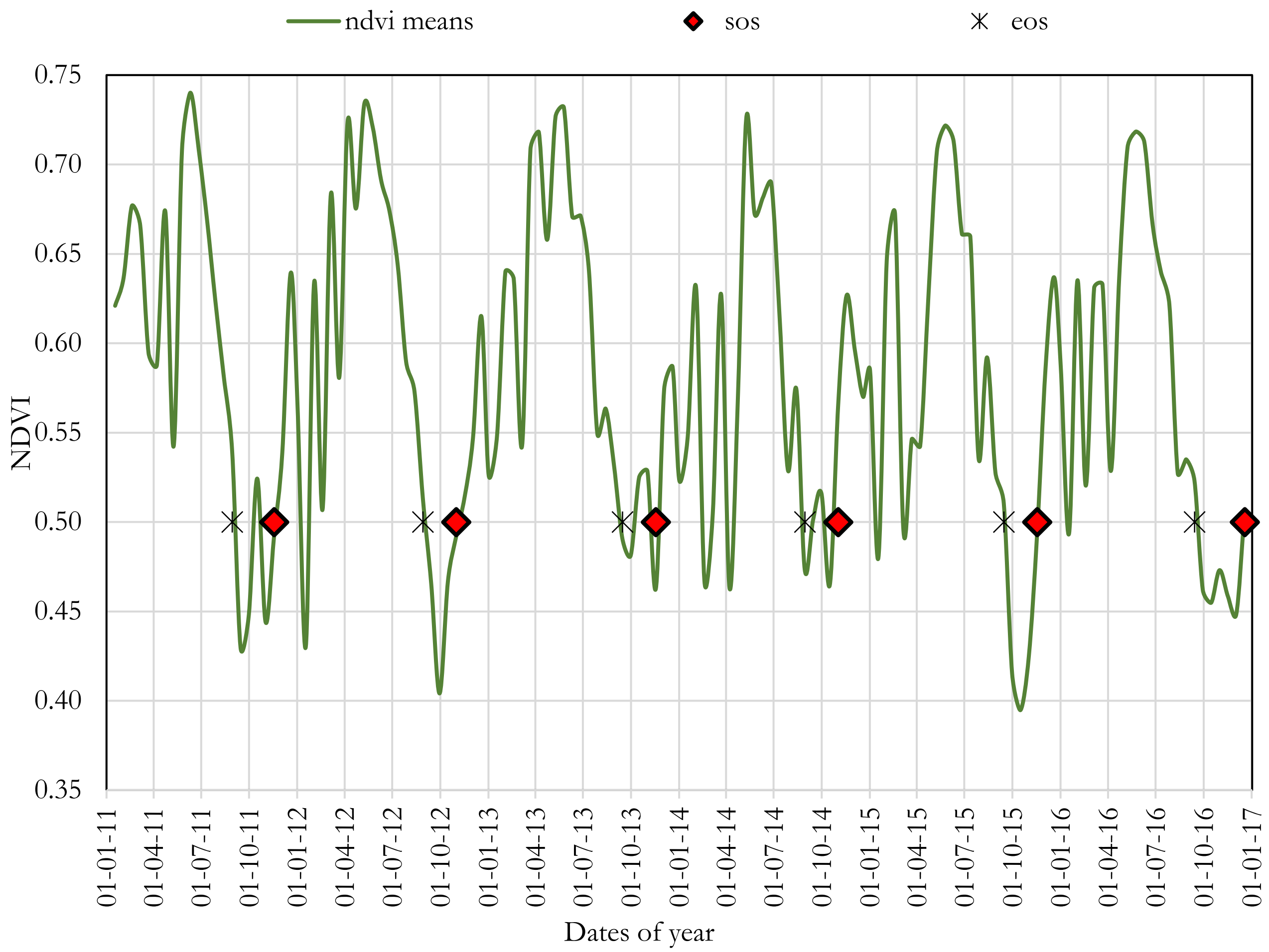


Figure 6.15: The Kilombero catchment mean NDVI, Start Of Season, and End Of Season

In 2012 and 2014, the season started in mid-October to mid-November as compared to 2011, 2013 and 2015 where it started in beginning November to beginning December. The latest start of the season was in 2010 from mid-November to mid-December. In 2011, 2012, 2014 the season ended towards mid-August to mid-September. In 2013, 2015 and 2016 the season ended slightly later i.e. in September (Table 6.4).

Table 6.4: Seasonal sequences of phenology based on MODIS NDVI time series

TIME	SOS	SOS	EOS	EOS	SOS	EOS	LOS
	date1	date2	date1	date2			
Year 1	17-11-10 (doy 321)	19-12-10 (doy 353)	13-08-11 (doy 225)	14-09-11 (doy 257)	03-12-10 (doy 337)	29-08-11 (doy 241)	269
year 2	01-11-11 (doy 305)	03-12-11 (doy 337)	12-08-12 (doy 225)	13-09-12 (doy 257)	17-11-11 (doy 321)	28-08-12 (doy 241)	285
year 3	15-10-12 (doy 289)	16-11-12 (doy 321)	29-08-13 (doy 241)	30-09-13 (doy 273)	31-10-12 (doy 305)	14-09-13 (doy 257)	318
year 4	01-11-13 (doy 305)	03-12-13 (doy 337)	13-08-14 (doy 225)	14-09-14 (doy 257)	17-11-13 (doy 321)	29-08-14 (doy 241)	285
year 5	16-10-14 (doy 289)	17-11-14 (doy 321)	29-08-15 (doy 241)	30-09-15 (doy 273)	01-11-14 (doy 305)	14-09-15 (doy 257)	317
year 6	01-11-15 (doy 305)	03-12-15 (doy 337)	28-08-16 (doy 241)	29-09-16 (doy 273)	17-11-15 (doy 321)	13-09-16 (doy 257)	301

SOS = start of season

EOS= end of season

LOS= length of season

The plotted values in Figure 6.15 are the middle dates of the weeks within which the SOS and EOS fall. It is noted that for the 6 years analyzed (Table 6.5), the SOS falls mainly between the first two weeks of November while the end of the season is mainly in the first two weeks of September. This is consistent with the rainfall patterns in the area. Despite the changing SOS patterns over the years, MODIS NDVI assessment gives indications of the beginning of a season when the NDVI value starts rising. This information is vital for the farmers, as they are made aware of the period to plant their crops to maximize the crop growth conditions resulting in increased yields. Moreover, the composite images are consistently available every 16 days increasing the temporal resolution of vegetation monitoring capacity.

The closeness of the short (November-January) and long rain (March-May) pattern seemed not to be picked by the seasonality assessment. This indicates that the sensitivity in identifying vegetation seasons with very close periods is low. The small decline in the NDVI curves between February and March is evident from the seasonality assessment but the magnitude is small such that it is not considered a season. In the study area, maize is planted in December and harvested towards the end of February to give way for rice plantations just before the beginning of the long rains in March. Quite a number of pixels had no values due to the presence of clouds, this implies that a complete coverage of the catchment in seasonality analysis is still challenging despite the utilization of the 16-day maximum composite images.

6.5.5 NDVI and soil moisture assessment

Vegetation growth is reliant on the availability of soil moisture. Detection of an increase in NDVI gives an estimate on the approximate planting time. However, a time lapse is required for the NDVI to be detected. This implies that the lag time when there is sufficient water in the soil for vegetation growth is lost when assessing the SOS with NDVI only. This calls for inclusion of soil moisture in the identification of SOS and additional analysis on the lag time between the presence of soil moisture capable of supporting vegetation growth and NDVI changes. This will aid in a higher precision in selection of SOS. Figure 6.16 show the Kilombero catchment average NDVI and average soil moisture values (results from the derived merged soil moisture product) for the Kilombero catchment.

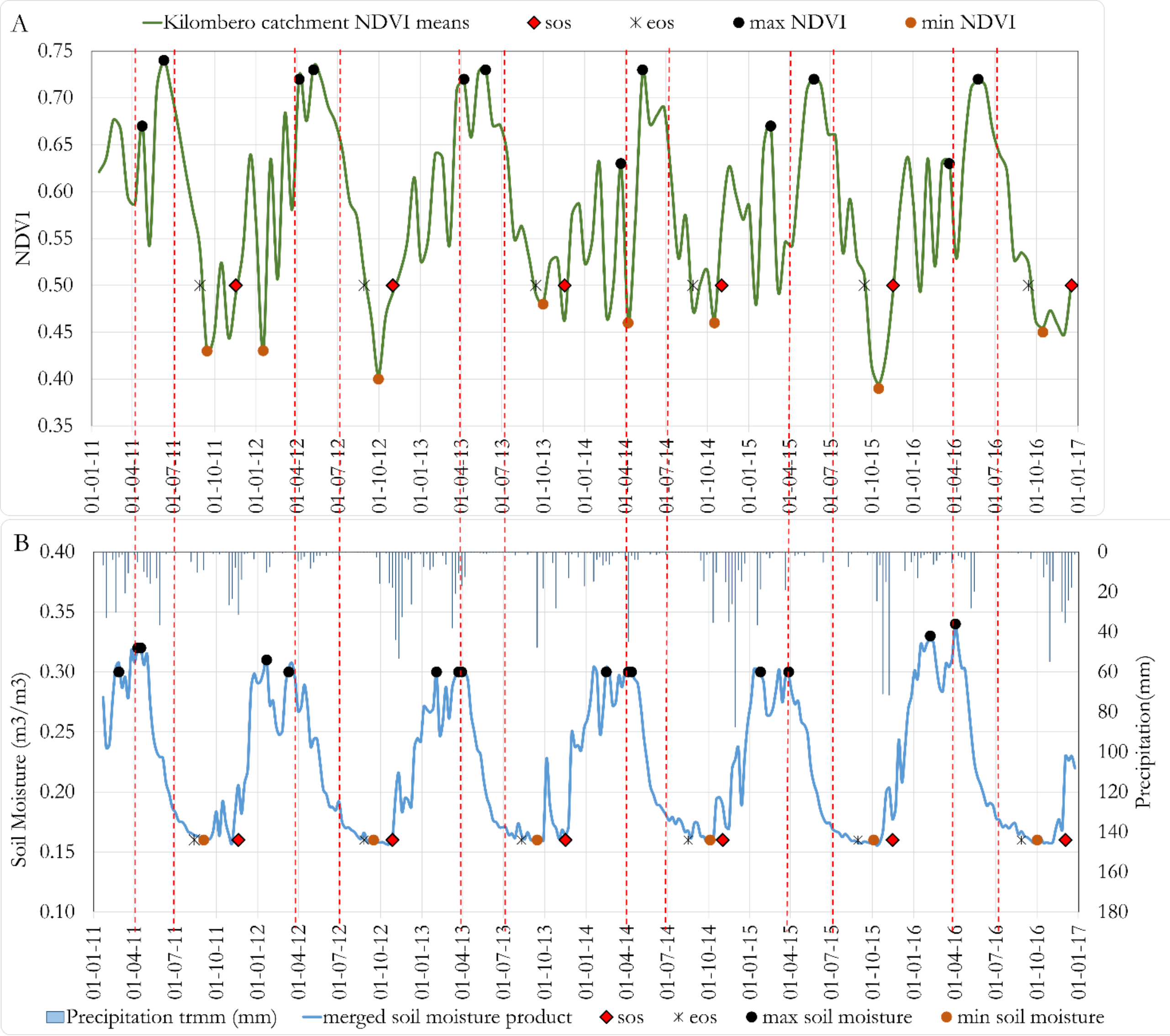


Figure 6.16: Comparison of NDVI and soil moisture patterns. Included is the daily TRMM precipitation

In 2011, there was an increase in soil moisture at the beginning of February. Towards the beginning of March, a deduction was experienced this subsequently resulted in a decrease in NDVI. Soil moisture later increased to its maximum values 0.3 to 0.31 for the month of April. During this time, NDVI increased to a value to a maximum of 0.74, occurring in mid of June. Given the maximum soil moisture was at the end of April and the max NDVI was in mid of June, there was a lag period of 1.5 months between the occurrence of maximum Soil moisture and NDVI. Soil moisture reduction started at the beginning of May reaching its lowest value of 0.15 from August to mid-September. NDVI started reducing from mid-June reaching the lowest value of 0.43in mid-September. There was an increase in Soil moisture in the first 2 weeks of October 2011 than a decline towards the last weeks of October. Subsequently, there was an increase in NDVI later reducing from mid-October. Soil moisture increased from the beginning of November to mid of January 2012 with a slight decrease from mid of

January to mid of February. The soil moisture increased again reaching a maximum in mid-March 2012. NDVI increased from November 2011 to April 2012. There were, however, several peaks and valleys at this time. The peak in mid of January resulted in a peak in NDVI at the beginning of January 2012. A slight decrease in soil moisture in mid-January resulted in a decrease in NDVI from beginning to mid of February while an increase in soil moisture from the end of February to mid-March increased NDVI from mid-March to beginning of April. A slight decrease of soil moisture at the beginning of April resulted in a decrease in NDVI from mid of April to the beginning of May 2012.

Soil moisture reduced from mid-April to end of October when the soil moisture values were lowest. Subsequently, NDVI reduced from mid-May to October 2012. Rainfall towards the end of October saw a rise in soil moisture reaching a small peak in mid-November before reducing again. The response time for the NDVI was one month with a slight peak occurring in mid-December. The slightly longer response time is due to the dry soils absorbing the moisture after being dry for five months. At the beginning of February, a peak in soil moisture results in a peak in NDVI in mid-February. An increase in soil moisture occurs from March 2013, reaching a peak from mid-March to beginning of April. NDVI reaches its peak at the beginning of April to June. Soil moisture shows a decreasing trend from April to mid-august. In this time, NDVI decreases, reaching its lowest value in mid-September.

At the beginning of October 2013, soil moisture increased with the NDVI reaching its peak at the beginning of November. There was a reduction in soil moisture, which increased towards the end of November. The NDVI response to the decline in soil moisture took one month. From the end of November to mid of February 2014, there was increasing soil moisture. The NDVI increases as well, reaching a peak in mid of March. A slight reduction in soil moisture after February followed with an increase occurring at the beginning of March. Peak NDVI values occurred from the end of March to the end of April. Subsequently, the rise in NDVI occurred from mid of April, reaching a peak in mid May. A reduction in soil moisture followed starting in May reaching lowest values in mid-August. There was a slight increase in soil moisture in mid-august. NDVI, on the other hand, reduced from mid-May to the beginning of September where it rose reaching a peak in October due to the soil moisture increase recorded. The lag time of the response was one month. From mid-October 2014 to mid-January, soil moisture increased resulting in an increase in NDVI in mid-February. The decreased soil moisture at the end of January resulted in decreased NDVI in mid-February. Soil moisture increased reaching its maximum in mid-March. Subsequently, NDVI peaked from mid-May to mid-June. 2015 had unique characteristics in that it recorded the earliest peak in short rains, which occurred in January, unlike the previously analyzed years. Additionally, max soil moisture occurred in March followed by an increased NDVI, which had a lag of 2 months to reach the peak. Moreover, the soil moisture reduction began in April reaching the lowest value at the beginning of August. The NDVI subsequently reduced from June to mid-October. Soil moisture increased from mid-October 2015 to February 2016. The corresponding NDVI increased till it reached a maximum in mid-March. Soil moisture decreased between March and may but increased till April. Subsequently, NDVI reached its peak in mid-May to mid-June 2016. Soil moisture then decreased to reach a minimum in August. NDVI subsequently reduced from mid-June to its minimum in October. Towards the beginning of November 2016, soil moisture had an upward pattern with NDVI showing a similar pattern of increase.

The study area has one season, which mainly begins towards the end of the year, extending to the next year and ending in September. The lowest soil moisture values occur between September and November. NDVI has several troughs and peaks especially as the NDVI increases from the SOS to its maximum value. The NDVI reduces between the months of May and June. The reduction is a smooth curve with minimal troughs and peaks. This is due to the changing weather in which case it gets hot and dry. The soil moisture at this time reduces drastically with no rainfall within the dry season hence subsequently; NDVI is stable as it reduces.

Table 6.5: Maximum and Minimum occurrence periods for the soil moisture and NDVI

Soil moisture		NDVI		SM-NDVI lag	SM-NDVI lag
Maximum Occurrence date	Minimum Occurrence date	Maximum Occurrence date	Minimum Occurrence date	Months	Months
26-02-11		23-04-11		2	
09-04-11	01-09-11	10-06-11	14-09-11	2	0.4
21-01-12		06-04-12		3	
10-03-12	15-09-12	08-05-12	29-09-12	2	0.5
02-02-13		07-04-13		2	
30-03-13	14-09-13	25-05-13	30-09-13	2	0.5
15-02-14		22-03-14		1	
12-04-14	04-10-14	09-05-14	16-10-14	1	0.4
24-01-15		18-02-15		1	
28-03-15	03-10-15	25-05-15	16-10-15	2	0.4
06-02-16		21-03-16		1	
02-04-16	01-10-16	24-05-16	15-10-16	2	0.5

A distinctive characteristic of SOS is its occurrence when soil moisture values are increasing. In the study site, SOS occurred when the soil moisture values are on average 0.17. On average, the end of season occurs when the soil moisture values reach 0.16.

A characteristic trend of the study period is the presence of constant rain that alternatively influences the behavior of soil moisture and NDVI response. The pattern observed indicates an increase in NDVI in towards the end of January or beginning of February. A slight reduction in NDVI follows which begins rising again in March, reaching a peak in May or June, depending on the rainfall occurrence (Figure 6.16). Between February and April when the soil moisture is highest, the lag time for the rises and falls of NDVI is 1 to 2 months (Table 6.5) though a lag time of 3 months was observed at the beginning of 2012. In the 2 months preceding January 2012, a relatively low amount of precipitation (15mm) had been recorded and hence causing a delay in the start of vegetation process to an amount detectable by NDVI satellite imagery. Towards the end of the dry season in September and October, any increase in soil moisture results in a lag time of approximately 0.5 months for NDVI to reach its peak.

The temporal resolution of the NDVI (16days from MODIS) and weekly merged soil moisture varies in time intervals. To have a consistent temporal resolution, monthly averages for both NDVI and soil moisture were computed for 2011 to 2016. Cross-correlation was applied to describe the lag of NDVI in relation to soil moisture (Figure 6.17).

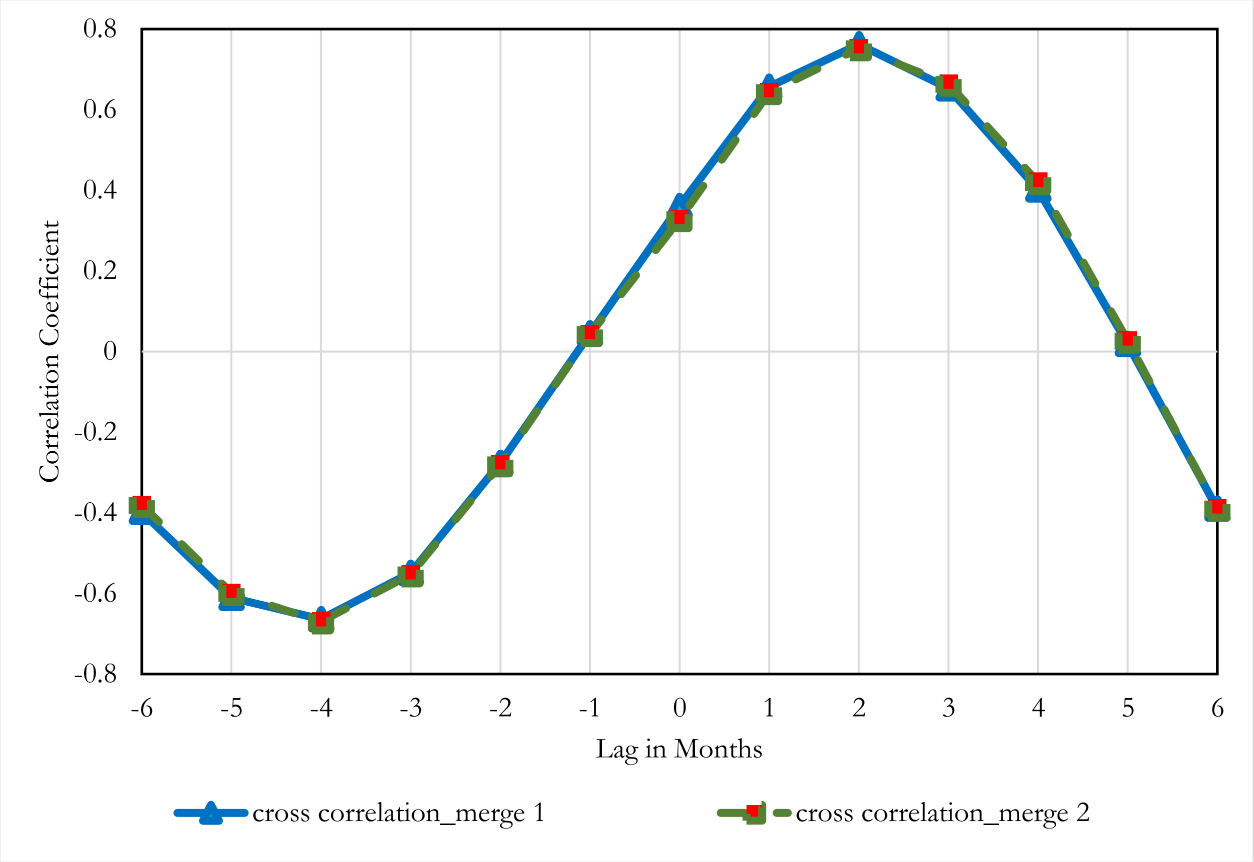


Figure 6.17: Cross-correlation of merged soil moisture and NDVI indicating the lag of NDVI to soil moisture

The maximum positive correlation value is 2 implying it takes 2 months for a change in soil moisture to be reflected in NDVI. Additional analysis was done to check the consistency in lag in the dry and wet seasons of the year and thus determine the effect of seasonality on the lag period.

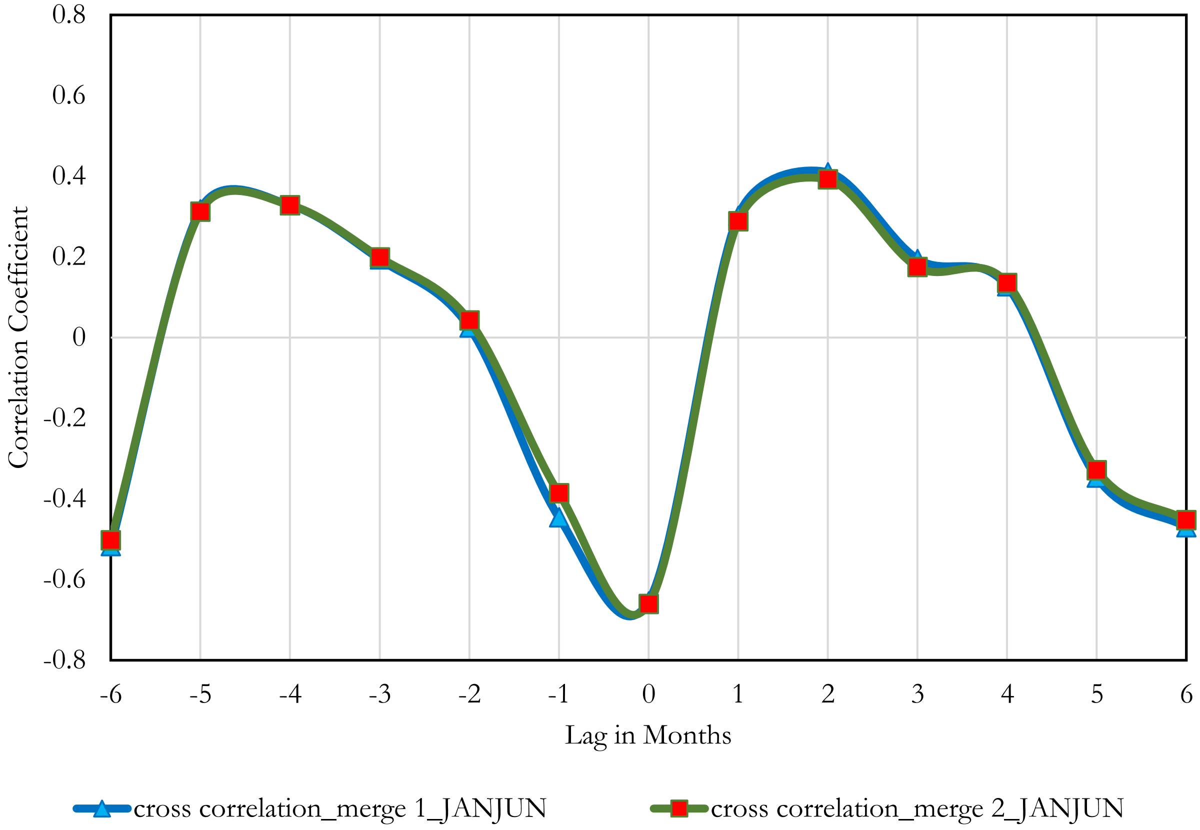


Figure 6.18: Cross-correlation of the merged soil moisture products and NDVI for January to June.

The correlation in the first half of the year which is rainy shows a two-month lag of NDVI to soil moisture, with a maximum correlation of 0.41. A month after the maximum correlation between the NDVI and soil moisture is reached, the rate of decline in correlation reduces. In the month that follows, the correlation seemingly stagnates, after which the rate of reduction in correlation steeply increases. The pattern of the graph is linked to the increase in soil moisture and subsequent NDVI towards the beginning of January. In February, the rain rate reduces and increases again in March. Since the vegetation responds to soil moisture availability, the pattern in reduction and increase in soil moisture is also exhibited in the soil moisture-NDVI lag relationship (Figure 6.16).

In the second half of the year, in the dry season, the lag of NDVI to soil moisture is one month (Figure 6.19) with a maximum correlation of 0.67. Similarly, (McNally *et al.* 2016) reported a better correspondence between NDVI and soil moisture in the dry period.

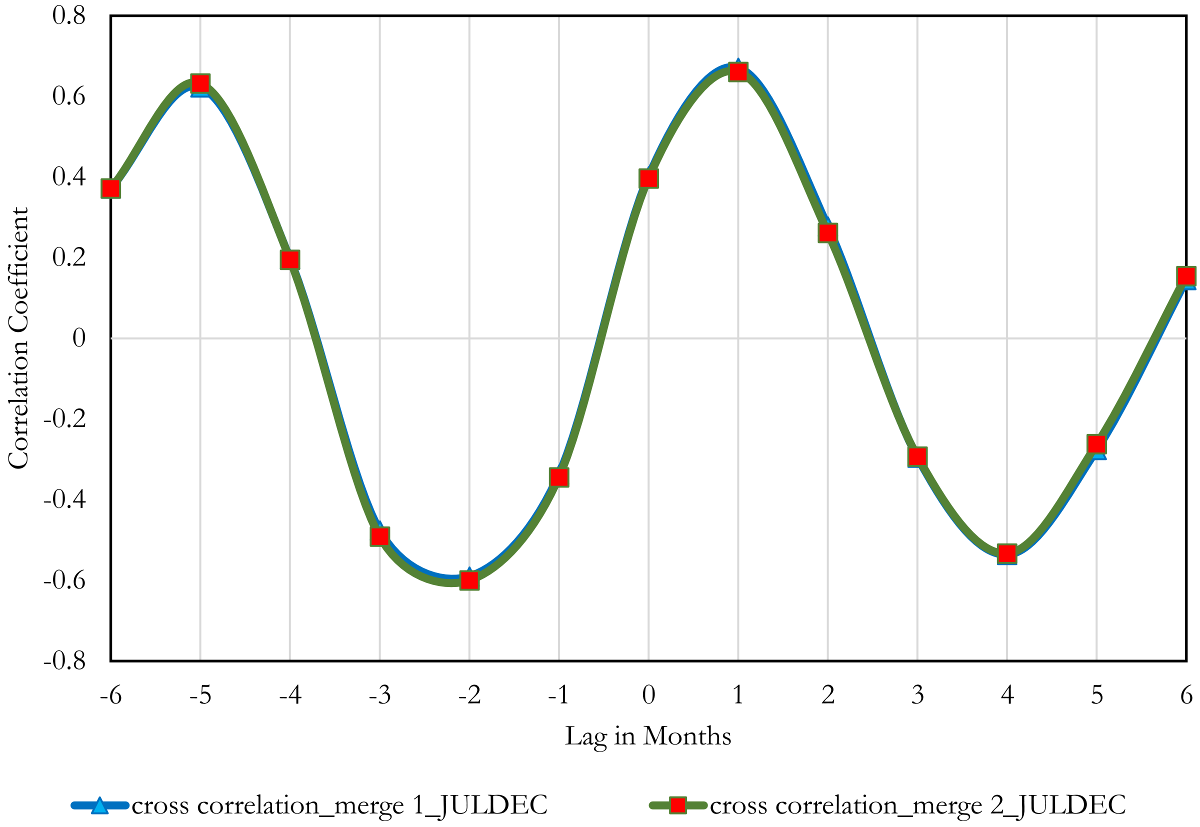


Figure 6.19: Cross-correlation of the merged soil moisture products and NDVI for July to December.

The outcomes presented show the sensitivity of NDVI to soil moisture changes similar to a study by (Chen *et al.* 2016). As noted in the results, the numerous peaks of NDVI are a consequence of a change in the soil moisture, the minimum lag period being 2 weeks with the maximum at 1 month when comparing the NDVI and merged soil moisture in their initial temporal resolution. However, on averaging the two sets of data, a lag period of 2 months is reported. In this case, however, some dynamics in the data are not detected since averages tend to smoothen out the data. Soil moisture products can be forecast up to 6 months in the future. However, NDVI images are acquired for the particular day in question. The dependence of NDVI on soil moisture, therefore, portrays an important use of the global soil moisture within the wetlands. The projected soil moisture is applicable in determination of planting times. This solves the problem of uncertainty faced by farmers with respect to suitable planting weeks more so in the face of changing seasons because of climate variations.

6.6 Conclusions

Wetlands are potential sites for increased food production to cater for the needs of the growing population. However, their use needs to be controlled and monitored to ensure their sustainability. Soil moisture is important for agricultural monitoring as it indicates plant water deficiencies earlier than conventional products such as the vegetation status from Normalized Difference Vegetation Index hence understanding its characteristics over a region such as the Kilombero catchment that depends on rain-fed agriculture largely contributes to improvement in agriculture. The study demonstrated the relevance of remotely sensed and modeled soil moisture products in identification of onset of vegetation growth.

In the current study, global soil moisture products were analyzed for the period 2010 to 2016. The spatial averages indicated that the ERA-Interim records the highest soil moisture as compared to FLDAS_VIC, whereas FLDAS_NOAH gives the lowest soil moisture. A similar pattern was observed for the uncertainties stipulated by the standard deviation of the three products.

PCA decomposition was employed in assessing the soil moisture spatial-temporal variability. All products had a leading dominant mode with 84.65%, 77.84% and 96.78% of the variation explained for FLDAS_NOAH, FLDAS_VIC, and ERA-Interim respectively. The spatial variation had a similar pattern to the standard deviation of the soil moisture products. The temporal variation, on the other hand, had similar patterns to the spatial averages of soil moisture indicating that the variation in mean soil moisture mainly influences the soil moisture variance.

The differences in the global soil moisture products called for the need to assess the error structure of the products. This was done using the triple collocation, whereby the error variances of each of the model products was determined. This acted as a weight in the generation of a merged soil moisture product. Two merged moisture products were produced. Merge 1 constituted of FLDAS_NOAH, ERA-Interim, and SMOS, whereas merge 2 constituted of FLDAS_VIC, ERA-Interim, and SMOS. The Merge product 2 had lower standard deviation and thus was adopted as the soil moisture product in the determining the start of growing season.

The end and start of growing season were calculated for the period corresponding to the stand-alone ensemble soil moisture product. It was noted that as the soil moisture rises, the NDVI begins increasing subsequently, marking the start of growing season, the peak in the NDVI however, occurs approximately 1-2 months after the peak of the soil moisture. The soil moisture increases from November with a slight reduction in February and an increase in March reaching a peak in April. The NDVI reaches a peak towards the end of May, from which a reduction in NDVI is experienced with the lowest occurring in October.

The study above has shown the potential of using global soil moisture and NDVI products in determining and monitoring the start of growing season. The wide spatial coverage and frequent temporal acquisitions of satellite-based and model-based products favor their use in the assessing changes over time. The study addresses the challenge experienced by farmers on uncertainty in planting times, which keeps changing year in year out due to climatic variations. The temporal resolution can identify changes occurring at minor scales. Moreover, a lag period of 2 to 4 weeks delayed vegetation response to soil moisture was identified. When the lag of monthly products was assessed, an NDVI lag of 2 months was detected, with the difference explained by the effects of averaging out of the soil moisture and NDVI. Soil moisture projections can be utilized in the identification of the precise periods in the start of growing season thus farmers fully utilize the moisture availability to plants especially since they depend on rain-fed agriculture.

It is recommended that future studies employ vegetation optical thickness, a product of SMOS instead of NDVI (Hornbuckle *et al.* 2016). This will account for vegetation water content for different vegetation types as different land covers have different responses to soil moisture (Van doninck *et al.* 2012). Incorporation of water requirement by different plants to assess the most suitable plants given the available soil moisture in the region warrants additional research. Soil texture determines the water holding capacity and thus its distribution recommended for future research.

7. General Discussion

7.1 Introduction

The study utilized remote sensing products to derive land cover maps and to assess the soil moisture conditions in the Kilombero catchment, Tanzania. The hypotheses at the commencement of the study were:

- Depth to groundwater has an influence on spatiotemporal land cover dynamics.
- Single polarized SAR images have sufficient information enabling land cover classification.
- There exists a variation in soil physical properties along a hydrological gradient within the Kilombero floodplain.
- Roughness has a big effect on accuracy of soil moisture derived from SAR imagery.
- Global soil moisture products can be utilized to infer the start of planting seasons.

The results of the study are presented in chapter 3 to 6. The current chapter gives the overall findings from the objectives assessed.

7.1.1 Seasonal land cover dynamics using optical data

Optical data was utilized in the generation of land cover maps in objective one. Additionally, depth to groundwater was related to seasonal land cover classes. The satellite imagery utilized included RapidEye, Sentinel-2, and Landsat-8. Reference information was obtained from aerial photographs, visual interpretation of the images and GPS points. The classification algorithm adopted was the Support Vector Machine (SVM). The land covers identified were open water, vegetation, bare land and built up areas. In the dry season, burnt areas were also identified.

The overall accuracies ranged from 70 to 96% with the Kappa index of 0.61-0.95 despite the varying resolutions of RapidEye (5m), Sentinel-2 (10m) and Landsat-8 (30m). RapidEye has accuracies ranging from 81 to 96%, Sentinel-2 78 to 95% while Landsat 70 to 91%. The assessment of effects of spatial resolution on classification accuracy resulted to slightly different accuracies from a SVM classification of the THEOS (Thailand Earth Observation System) 15m satellite imagery (90.65%) and Landsat 30m resolution imagery (89.00%) (Suwanprasit and Srichai 2012). Another study to map agricultural parcels using 2.5m, 5m, 10m SPOT and 30m Landsat reported that a reduction in mapping accuracy occurs for pixels above 10m (Alganci *et al.* 2013). This indicates that finer spatial resolution increases the potential of extraction of land cover features since information that is more detailed is retrieved during the classification process. For more detailed mapping of the small-scale parcel size and crop type mapping in the study area, higher resolution imagery is required as distinguishing small parcel less than 1.5 hectares requires imagery with resolutions less than 10m (Alganci *et al.* 2013).

The study revealed that the depth to groundwater was highest between July and December and was closest to the surface in March to June. This is in line with the precipitation pattern whereby precipitation is highest in

March to May from which the dry season sets in with minimal precipitation events between July and October. This indicates that precipitation plays a critical role in groundwater recharge. A study by Li et al. (2014) identified precipitation, temperature, groundwater exploitation and reservoir construction as the major factors influencing groundwater levels (Li *et al.* 2014). In this study however, we focused on precipitation although it would be interesting to study the influence of human activities on the groundwater dynamics especially with the increasing population and subsequently increasing demand for water.

There was a distinct variation in depth to groundwater in the three hydrological zones. The depth to groundwater at the fringe zone is close to the surface throughout the year. This zone receives recharge from the Udzungwa Mountains, located north of the study area (Gabiri *et al.* 2018). The riparian zone experiences the highest depth to groundwater between November and January after which the depth begins reducing due to groundwater recharge from the precipitation.

The bare land coverage was between 45-57% and increased to 62-69% in the January-June (wet) and July-December (dry) respectively. The vegetation cover was between 34-47% in June and reduced to 25-27% in December. Within this period, the depth to groundwater increased from June to December as the season shifted from wet to dry owing to the reduction in groundwater recharge. Land cover influences the rate of recharge as it affects the infiltration and evaporation rate as evidenced by Koike et al. (2008) who reported a reduction in groundwater owing to increased urban land cover (Moukana and Koike 2008). The rate of recharge (approximately 3 months) is higher than the rate of groundwater loss (approximately 7 months). When the rainy season begins in January, the land cover is mainly bare and conversion to vegetated cover occurs as moisture accumulates to support plant growth. Managed land cover such as annual crops have shallow rooting system and less extensive moisture extraction depths hence explaining the high groundwater infiltration rate (Owuor *et al.* 2016). (Scanlon *et al.* 2005) reported similar results stating that dry bare land increases ground recharge (Scanlon *et al.* 2005). In the dry season, which begins in June-July, evaporation reduces soil moisture subsequently resulting in increased depth to groundwater (Zhang and Schilling 2006).

The study also revealed the change of the land cover pattern at the riparian zone with respect to seasons. In the dry season, 68-81% of the total vegetation in the study area is within the riparian zone. This indicates that the zone relies on saturation from the Kilombero River and soil properties encompassing high clay content (Gabiri *et al.* 2018). A major limitation of the study was the availability of cloud-free images.

7.1.2 Performance of single-polarized VV, GLCM texture images and their PCA derivatives in monitoring land cover

The challenge of presence of clouds and thus inconsistent availability in the optical images motivated the second objective where weather independent SAR images are utilized in land cover mapping. During the study period (2015-2016), there was a consistent temporal acquisition of single-polarized Sentinel-1 images. Discrimination of features from single polarized images is low (Mosleh *et al.* 2015). This study generated Grey Level Co-occurrence Matrix (GLCM) data that accounts for the spatial dependence of neighboring pixels with an aim of increasing feature discrimination. GLCM produces a vast amount of data and hence Principal Component Analysis (PCA) was applied to reduce the GLCM features (Singh and Kaur 2011). Three classification algorithms were applied to the GLCM texture features and their corresponding PCA images.

The overall accuracies for the single polarized VV images ranged from 54-76%, 60-81%, and 61-80% for Random Forest (RF), Neural Net (NN) and Support Vector Machine (SVM) respectively. GLCM features had overall accuracies of 64-86%, 65-88%, and 65-86% for RF, NN, and SVM respectively. PCA derived images had similar overall accuracies of 68-92% for NN, RF, and SVM respectively. The PCA images had the highest overall accuracy for the entire time series indicating that reduction in the number of texture features to layers containing the maximum variance improves the accuracy. It was noted that the accuracies are the same despite application of different classification algorithms in the case of PCA images. The study, therefore, revealed that the PCA images capture the variation of the land covers subsequently improving the land cover mapping. None of the classification accuracies portrayed a distinctively better performance. The transformation of the set of texture features into new axes that capture significant information as is the case in PCA was reported to increase classification accuracy in mapping water, urban and vegetated areas using ERS-2 SAR imagery operating in C band (Chamundeeswari *et al.* 2009). A challenge posed in the study by Chamundeeswari when mapping small-scale agricultural plots that was similar to the current research of Ifakara was misclassified pixels in areas of heterogeneity due to mixed land cover classes. Mapping agricultural land using single C band polarizations of HH, HV and VV resulted in improved accuracies on inclusion of texture features (Treitz *et al.* 2000).

GLCM features from single polarized TerraSar-X were similarly reported to increase accuracies in crop type mapping on well-demarcated fields (Mahmoud *et al.* 2011). (Jia *et al.* 2012) observed improved accuracies on including texture features derived from TerraSar-X and ASAR in mapping cotton, non-crop and wheat areas (Jia *et al.* 2012). The texture features smoothened out the salt pepper appearance of the SAR images subsequently improving the classification accuracy. GLCM have been reported to be of a higher importance in mapping when compared to statistical neighborhood features (e.g. mean, variance) and structural features (e.g. length, width) in crop type identification (Inglada *et al.* 2016).

Variation in frequency of the radar images influences the discrimination capability of features as it affects the penetration depth of the microwave signal (Ulaby *et al.* 1996). Shorter wavelength X is attenuated by changes in plant structure as they grow, which plays a role in improving feature discrimination, while longer wavelength C can discriminate denser vegetation and it is affected by soil moisture and surface roughness in sparsely vegetated agricultural fields (Treitz *et al.* 2000, Sonobe *et al.* 2014). For crop type mapping therefore, X band is preferred and recommended for the small uneven fields in the study area. The higher spatial (3m) and temporal (11days) resolution of the TerraSar-X is an advantage in mapping small-scale farms as it capture the changes occurring between acquisitions and in monitoring the critical growth stages of crop growth (Lopez-Sanchez *et al.* 2011, Koppe *et al.* 2013).

7.1.3 Spatial-temporal variation of physical soil properties and soil moisture estimation from SAR data

Variation of soil moisture is a function of scale. The composition of soil texture, presense of topographic features, land cover and meterorological forcing influence moisture variability at field, watershed, regional and continental scale respectiverly (Crow *et al.* 2012). In the study, spatial-temporal distributions of soil physical properties were assessed as they influence moisture distribution in turn affecting field scale crop production. Soil moisture is dynamic and its change over time was assessed. The soil moisture field data was collected between March and May 2015 and December and February 2016. From the end of April to May, the riparian fields were flooded owing to the long rains.

Spatial distribution of texture (sand, silt, and clay), organic carbon content, bulk density and soil moisture indicated variations in distribution at the riparian, middle and fringe hydrological zones. Soil physical property distribution illustrated that the riparian zone has a high clay and organic carbon content and that the fringe has the highest level of sand and bulk density value. The highest silt content is in the middle zone. According to the USGS soil classification, the riparian zone has silty clay loam while the middle and fringe zones have silt loam.

Validation and upscaling soil moisture to a satellite footprint requires multiple measurements from multiple fields within the footprint. For example, the SMOS footprint is on average 43*43 km^2, representing one pixel. This presents two aspects in planning the data collection for validation. Firstly, several well-distributed SMOS pixels are required within the selected validation region, running several hundreds of km squared. Secondly, within each 43*43 km pixels, well-distributed fields from which observations of point moisture measurements will be carried out need to be identified. Understanding the soil moisture variability within the selected fields and subsequently applying sampling requirements to obtain average values representative of the footprint in crucial in the upscaling and/or validation process of satellite-derived moisture.

Understanding the variation in moisture is also important as it describes uncertainty associated with averaging point-based soil moisture observations to the extent of a satellite product. Results from the Ifakara soil moisture variability assessment indicated that the variation in moisture increases as the soil becomes drier. Brocca et al. (2009) assessed soil moisture in fields ranging from 5,000 to 8,800 m^2 at 10 by 10m sampling intervals and reported similar findings whereby coefficient of variation increases for drier fields (Brocca *et al.* 2009). Similarly, Famiglietti et al. (2008) reported a similar behavior in moisture pattern across varying scales among which was a 50km scale which is within the SMOS footprint (Famiglietti *et al.* 2008). However, he also reported that as the scale of sampling increases the uncertainty likewise increases. This implies that more measurements are essential to obtain a highly accurate footprint average at large areal extents and during drier conditions.

The number of sample points estimating means of soil moisture representative of the fields at the fringe (~370*200 m) and middle (~250*150 m) with an accuracy of 0.2% vol was 30 and 25 respectively. At field scales and considering high density of measurements taken at 10m intervals, soil texture was observed to play a role in field moisture variability, which influenced the coefficient of variation (CV) (Crow *et al.* 2012). The number of samples required was estimated using the Student t distribution whereby the CV governs the output following the principle that a high CV requires more sampling points. As the scale increases, however, moisture variability is influenced by topography and vegetation (Crow *et al.* 2012). Results by Famiglietti et al. (2008) stipulated that 30 samples are necessary at scales of 50*50km² to capture the soil moisture mean with an accuracy of ±0.03 cm^3/cm^3 while (Brocca *et al.* 2010) reported that 40 samples are required at an area of 60*60km² to capture moisture with an accuracy of ±2%. The higher number of sample requirement is explained by the larger areal extent and the higher accuracy requirement (Brocca *et al.* 2010).

Identification of temporally stable fields opens possibilities of reduction in the number of measurement samples required to obtain mean field soil moisture necessary in upscaling/validating satellite moisture. Temporal stability was assessed by applying relative difference technique to observed field moisture averages. In general, the riparian fields exhibited stable temporal patterns as thus were identified as representative fields. The high clay content increases the retention capacity of moisture contributing to the steady soil moisture and hence reduced variability exhibited for the riparian fields. The standard deviation of relative differences ranged from 4 to 15% for study fields ranging from approximately 730 to 19,200 m^2. Random patterns were observed in terms of areal extents and level of temporal variations while identifying the representative fields. Brocca et al. (2009) similarly reported randomness in terms of identifying representative locations in a flat field, reporting that topography influences the selection (Brocca *et al.* 2009).

More specific to the hydrological zones, the riparian fields, C1, C3 and C5 postulated negative values whereas C2, C4 and C6 had positive values. Negatives indicate underestimations from the average implying the fields are relatively drier than the mean. Positive values are representative of fields that are relatively wetter than the mean. In the middle fields, M2 and M3 have higher moisture contents than the mean, while M1 and M5 indicate drier than mean conditions over time. Moisture accumulation starts in M2 and M3 due to infield level differences that are visible from the soil moisture accumulation patterns in Appendix B8. M1 and M5 slightly elevated and hence in cases of precipitation events, once the soil becomes saturated, they experience surface runoff. In the fringe fields, F1 and F3 have wetter than average conditions. These two fields are large, and they have infield troughs where water accumulates first in precipitation events. F2 and F4 are smaller fields that are slightly elevated as related to F3 and F1 (Appendix B9). Temporally stable fields have a low value of relative difference as well as low standard deviation of relative difference. For the riparian, middle and fringe fields, C6, M2 and F2 with areal coverages of approximately 4,200, 1,100 and 10,700 m^2 were identified as temporally stable with relative differences of 1.93 (±4.76%), 20.58 (±8.2%) and -5.62 (±6.3%). A large standard deviation as was the case mainly for middle fields indicated that the soil moisture measured was not linearly related to the field averages (Cosh *et al.* 2008). Brocca et al. (2009) reported similar standard deviation values (6-18 and to 2-27%) for fields 5,000 and 8,800 m^2 respectively. From the results, however, there is a large variation from the mean, explained by the highly detailed and dense data collection approach adopted, which highlighted the sensitivity of texture on soil moisture at the field scale. More temporal persistence tests are recommended on additional fields with a less dense network of moisture samples. Additionally, a longer moisture collection period is required for conclusive definition of fields as temporally stable.

Temporal variograms were used in identifying temporal structure of the soil moisture. The results indicated that the number of days over which the moisture values were correlated was 109 days for the riparian, 65 days for the middle and 71 days for the fringe. Sampling dates comprised of two time epochs, one from March to May 2015 and the second from December 2015 to February 2016. Between the time epochs there was a measurement gap of 213 days. Assessing the temporal persistency in terms of days of year to determine if the moisture distribution is stable in time requires continuous data over a long period.

Deriving soil moisture from SAR imagery requires roughness parameters. However, since it is time-consuming and impractical to collect over large areas, a representative parameter was generated from the SAR backscatter. Linear regression equations were established from the polarization of the RadarSat-2 and TerraSar-X. Soil moisture was estimated from the regression equations and the simulated roughness parameter. The coefficient of correlation (r) between simulated and observed soil moisture was 0.94 and 0.58 for RadarSat-2 HH and HV respectively whereas it was 0.95 and 0.96 for TerraSar-X HH and VV. The lower sensitivity of backscatter to roughness at low incidence and low frequencies increases soil moisture retrieval accuracies as was the case for the TerraSar-X (Mattia *et al.* 1997, Holah *et al.* 2005). Poor signal to noise ratio for the cross-polarized backscatter explains the low moisture retrieval accuracy from the RadarSat-2 HV (Jiankang Ji *et al.* 1996).

7.1.4 Performance of the global soil moisture products and its relation to start of season

Estimating soil moisture from SAR over the Kilombero catchment requires intense and frequent collection of in situ soil moisture measurements. This is costly and impractical and hence global products are preferred. The objective was motivated by the need to have frequent soil moisture values at the catchment scale. Moreover, the use of soil moisture in determining the start of growing season was investigated.

The spatial average of FLDAS_NOAH, FLDAS_VIC, and ERA-Interim soil moisture revealed the seasonal variability with peaks towards the end of the year, reaching maximum in March to May. FLDAS_NOAH had the least soil moisture average while ERA-Interim had the highest estimates. The inconsistency in the values is attributed to varying retrieval algorithms, differences in spatial and temporal resolution. The standard deviation of the soil moisture products illustrated a similarity in the FLDAS_NOAH and FLDAS_VIC attributed to the fact that both are forced by RFE precipitation. Spatial-temporal variability was analyzed through Principal Component Analysis decomposition. Spatial variability explained in FLDAS_NOAH, FLDAS_VIC and ERA-Interim was 84.65, 77.84, and 96.78% respectively. The region has a bimodal soil moisture regime occurring in December to January and March to May.

Two merged soil moisture products were generated by applying the triple collocation approach, one from FLDAS_NOAH, ERA-Interim, SMOS, and the second from FLDAS_VIC, ERA-Interim, and SMOS. The moisture layers for FLDAS_NOAH and VIC was 0-10cm that of ERA-Interim was 0-7 cm whereas passive sensors detect moisture at depths of 3 cm. The vertical inconsistencies are handled in the rescaling performed in TC as long as the relationships are linear, which was the case in the products analyzed (Yilmaz *et al.* 2012).

Error variances in the global soil moisture products were determined by triple collocation. Using the error variances, weights were assigned to each of the products. The weights of product 1 were 0.13, 0.77 and 0.10 for FLDAS_NOAH, ERA-Interim, and SMOS. Weights of product 2 were 0.35, 0.52 and 0.13 for FLDAS_VIC, ERA-Interim, and SMOS. The higher weights for the ERA-Interim are consistent with studies by Scipal et al. (2008) and Dorigo et al. (2010) who reported consistency in the ERA global soil moisture product. The low weights assigned to the SMOS product are explained by the sensitivity of passive sensors to signal attenuation on vegetated areas (Anderson *et al.* 2012). The Kilombero catchment comprises of a series of vegetation covers, including forests and woodland areas (Figure 6.11). In the current analysis, the average values of the catchment irrespective of the land cover were assessed and hence the effect of vegetated areas lowered the performance of SMOS in formulation of the merged soil moisture products. The two merged products exhibited similar patterns through the study period as they are forced by the same precipitation product, RFE. Moreover, the ability of the two merged consensus products to portray similar interannual patterns as the independent soil moisture datasets boosts confidence in their use over the Kilombero catchment since the discrepancies between the datasets do not alter the annual variations in soil moisture (Anderson *et al.* 2012). The dissimilarities in the merged products are due to the difference in input parameters and derivation algorithms for the FLDAS_NOAH and FLDAS_VIC soil moisture. Some of the differences include the use of IGBP land cover classification from MODIS and use of green vegetation fraction to describe the vegetation phenology for the FLDAS_NOAH. FLDAS_VIC uses the UMD land cover classification derived from AVHRR and utilizes LAI to describe the vegetation phenology (McNally *et al.* 2016, 2017).

Start and end of the season (SOS and EOS) were determined from 16-day composites of MODIS NDVI from 2010 to 2016. Our findings indicated that the SOS occurs between November and December, EOS occurs between August and September and the LOS is 290 days. Despite the relatively short period of determining the phenological parameters, the SOS, EOS, and LOS were similar to a study by Vrieling et al. (2013) who assessed the phenological parameters for a period of 30 years over Africa. In the results reported over East Africa in the location of Kilombero catchment, the SOS and EOS was in December and August respectively while the LOS was 240-300 days.

The soil moisture temporal pattern was similar to that of NDVI though with a lag period. The SOS occurs in mid-November whereas the EOS occurs in August. The SOS and EOS occurred when the weekly soil moisture was 0.17 and 0.16 respectively. A lag of 2 months was observed for the NDVI for an increase in soil moisture. The results presented comprised of analysis within the Kilombero catchment unlike the results by McNally who focused on East Africa. (McNally *et al.* 2016) reported a 1-month lag in a moderately vegetated area (Mpala, Kenya), 1 to 2-month lag in a sparsely vegetated area (Tigray, Ethiopia), 1 month lag in a moderate to high vegetation area (Illubabor, Ethiopia) of GIMMS NDVI3g from CCI soil moisture between 1992 to 2013.

7.2 Conclusion

The chapter has revealed, that optical remote sensing facilitated the generation of seasonal land cover maps, which were linked to the depth to groundwater. Due to the limitation of cloud coverage, the study went ahead to show the use of free access Sentinel-1 radar images in continued monitoring of the floodplain.

Soil moisture is a dynamic and critical parameter in agriculture. Assessing the spatial-temporal variation of soil moisture in the riparian, middle and fringe hydrological zones was illustrated. Moreover, the soil physical parameters explaining the moisture variations were assessed. At the catchment scale, application of global soil moisture products was illustrated. The patterns of NDVI and soil moisture were assessed to show the potential use of soil moisture in identify the start of growing season, which is important to the farmers especially in light of the changing climate and seasons.

8. Synopsis

8.1 Conclusions and lessons learnt

The increasing demand for food, arable land shortage, and unpredictable climate conditions in East Africa have created a paradigm shift from upland cultivation to wetland use due to their fertile soils and year-round soil water availability (Ramsar Convention on Wetlands *et al.* 2014). The Kilombero catchment was identified as having potential for increased food provision to cater for Tanzania's growing population (Kato 2007, Tanzania National Bureau of Statistics 2018). Knowledge of land cover changes over time is essential in sustainable agricultural production in the wetland. In addition to understanding the characteristic land cover patterns in this rainfed agricultural region, knowledge on soil moisture is a key in improved production as it has a direct link to vegetation growth.

The research herein, therefore, sought to understand land cover patterns, soil moisture spatial-temporal variability at field and catchment scale and correlate the soil moisture with Start of Season. Multi-temporal multi-sensor optical images assessed the land cover changes occurring within the floodplain. Cloud cover presented challenges in the use of optical images and thus weather independent Radar images were assessed for their application in wetland land cover monitoring. For the period under investigation (2015-2016) however, single polarized Sentinel-1 images were available. A challenge in discrimination of features emerges with the utilization of single-polarized images and hence Grey Level Co-occurrence matrix images and the PCA derivatives of the texture images were assessed by applying multiple classification schemes.

Time Domain Reflectometry probes were taken in 18 soil moisture measurement campaigns on fields ranging from 700 to 19,000 m^2 at three hydrological zones between March and May 2015 and December 2015 to February 2016. The hydrological zones were defined by their variation in flooding patterns during the rainy season. The riparian zone is closest to the Kilombero River, has the lowest elevation and is usually flooded in the rainy season. The middle has patched floods and non-flooded regions whereas the fringe is not flooded in the rainy season. Statistical analysis, temporal persistence and geostatistical assessment were performed on the soil moisture data. The fields assessed have a bulk density less than 1.5g/cm^3, suitable for plant growth. There is high variability of soil moisture in drier soil conditions and hence it was observed that more sampling points are required in such conditions. The fields closest to the Kilombero river were the most temporally stable as compared to the middle and fringe fields.

Soil moisture is highly dynamic and requires continuous and consistent measurements. Field point measurements are tedious and expensive to acquire. Therefore the use of consistently available SAR imagery in soil moisture estimation is explored. Soil roughness influences the accuracy of the moisture retrieved from SAR. The practicality of roughness collection over study area is minimal and hence the study aimed at formulating apriori roughness parameter that can be applied in the retrieval of moisture. At the catchment scale, the spatial-temporal variability of global soil moisture products was assessed to understand the relative deviations in the products arising due to varied retrieval algorithms. Errors were assessed by applying triple collocation analysis, weights were generated and assigned to the products considered in generating a merged soil moisture product. Temporal patterns of NDVI and the merged soil moisture product were examined with the aim of establishing the time lag relationship. Applying the time lag relationship to soil moisture forecasts will aid in identifying the start of the season for agricultural production within the catchment. This information is crucial in the face uncertainty in planting times due to changing climate.

The hypothesis stating that the depth to groundwater has an influence on land cover was confirmed. The depth to groundwater declines from June after the rain season and increases in December at the beginning of the rain season. Bare land cover over the entire study area was 45-57% and 62-69%, while vegetation was 34-47% and 27-25% in June and December respectively. Vegetation cover decreases with increasing depth to groundwater.

The hypothesis stating that single polarized SAR images have sufficient information enabling land cover classification was rejected. The analysis from single polarized, texture images and PCA images revealed that increasing dimensionality of the single polarized VV images increases feature discrimination and thus results in improved classification accuracy.

The hypothesis that there exists a variation in soil physical properties along a hydrological gradient in Kilombero floodplain was accepted. Roughness has a big effect on the accuracy of soil moisture derived from SAR imagery was confirmed. Measured roughness parameters had a high variation due to the small sizes of the fields and representation of point measurements as field averages. Inversion of a model generated from multi-frequency multi-angular soil moisture measurements over different parts of the world under varying conditions resulted in a simulated roughness parameter applicable in the study area. Soil moisture estimation with the simulated roughness parameter resulted in R^2 of 0.83-0.95.

The hypothesis that global products such as soil moisture and NDVI can be used to determine the start of planting seasons was confirmed. A time lag of 2 months was revealed between soil moisture and NDVI increase. Hence identification of the Start of Season is possible with projected soil moisture data.

In summary, optical and microwave remote sensing have a high potential in monitoring wetland use through generating land cover maps. Soil physical properties play a role in soil moisture retention capability. SAR images have potential in deriving soil moisture given a representative roughness parameter. The use of global soil moisture product at the catchment scale aids in the identification of the start of planting season, which is important information to the farmers in the region.

8.2 Recommendations

The study can be improved by

- As this was the first study to understand the land cover changes in the study area, detailed crop type mapping was not carried out. With the knowledge about the land cover dynamics generated from the current research, more detailed mapping is recommended whereby farm management practices will be considered as they affect the reflectance and backscatter recorded by the satellite imagery. Continuous monitoring of backscatter and reflectance from large fields with specific crops such as sugarcane or rice to capture the growth pattern at different crop stages is recommended. The use of large-scale fields is recommended as several satellite imagery pixels will be contained within the fields ensuring values analyzed are within field boundaries. Assessment of the backscatter and reflectance values at different growth stages offers information such as start, end, and length of growing of the specific crop under study. This information from the remote sensing data coupled with crop models can then be used in yield prediction, an important aspect in food security assessment.

- Climate oscillations and telecommunications influence the precipitation patterns which in turn affect the vegetation growth. Incorporating climate phenomenon such as El Niño and La Niña in determining the start and end of season is recommended.
- The soil moisture - NDVI relationship was assessed over the entire catchment irrespective of the land cover. It is recommended that the same analysis will be repeated in cropped areas. However, since global products have a low spatial resolution, downscaling to the resolution of cropped areas will be required.

8.3 Outlook

The seasonal land cover maps were focused on a small area in Ifakara. Suggestions are made to extent the seasonal land cover generation to a larger area. The effectiveness of any technologies adopted with the aim of improving agricultural produces should be assessed by repeating the study and analyzing changes occurring in areal extent of vegetated land cover class.

Recommendations are made to have additional in situ soil moisture measurement campaigns distributed over the Kilombero catchment to test the performance of the merged soil moisture products as well as for application in upscaling and validating satellite-based soil moisture. Additionally, the potential of more global moisture products such as CCI, SMAP should be tested for their potential in agricultural management such as identifying declining moisture periods which can potentially be linked to drought occurrence.

9. References

ACPA, Sydney, 2006. Vesper 1.6.

Adam, E., Mutanga, O., and Rugege, D., 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management*, 18 (3), 281–296.

AghaKouchak, A., 2014. A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought. *Hydrology and Earth System Sciences*, 18 (7), 2485–2492.

AghaKouchak, A., Farahmand, A., Melton, F.S., Teixeira, J., Anderson, M.C., Wardlow, B.D., and Hain, C.R., 2015. Remote sensing of drought: Progress, challenges and opportunities: Remote Sensing of Drought. *Reviews of Geophysics*, 53 (2), 452–480.

Albergel, C., de Rosnay, P., Balsamo, G., Isaksen, L., and Muñoz-Sabater, J., 2012. Soil Moisture Analyses at ECMWF: Evaluation Using Global Ground-Based In Situ Observations. *Journal of Hydrometeorology*, 13 (5), 1442–1460.

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W., 2012. Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sensing of Environment*, 118, 215–226.

Alexandratos, N. and Bruinsma, J., 2012. *World agriculture towards 2030/2050: the 2012 revision*. ESA Working paper Rome, FAO.

Alganci, U., Sertel, E., Ozdogan, M., and Ormeci, C., 2013. Parcel-Level Identification of Crop Types Using Different Classification Algorithms and Multi-Resolution Imagery in Southeastern Turkey. *Photogrammetric Engineering & Remote Sensing*, 79 (11), 1053–1065.

Álvarez-Mozos, J., Verhoest, N.E.C., Larrañaga, A., Casalí, J., and González-Audícana, M., 2009. Influence of Surface Roughness Spatial Variability and Temporal Dynamics on the Retrieval of Soil Moisture from SAR Observations. *Sensors*, 9 (1), 463–489.

An, R., Zhang, L., Wang, Z., Quaye-Ballard, J.A., You, J., Shen, X., Gao, W., Huang, L., Zhao, Y., and Ke, Z., 2016. Validation of the ESA CCI soil moisture product in China. *International Journal of Applied Earth Observation and Geoinformation*, 48, 28–36.

Anderson, W.B., Zaitchik, B.F., Hain, C.R., Anderson, M.C., Yilmaz, M.T., Mecikalski, J., and Schultz, L., 2012. Towards an integrated soil moisture drought monitor for East Africa. *Hydrology and Earth System Sciences*, 16 (8), 2893–2913.

Ardö, J., 2013. A 10-Year Dataset of Basic Meteorology and Soil Properties in Central Sudan. *Dataset Papers in Geosciences*, 2013, 1–6.

Arzandeh, S. and Wang, J., 2002. Texture evaluation of RADARSAT imagery for wetland mapping. *Canadian Journal of Remote Sensing*, 28 (5), 653–666.

Attarchi, S. and Gloaguen, R., 2014. Classifying Complex Mountainous Forests with L-Band SAR and Landsat Data Integration: A Comparison among Different Machine Learning Methods in the Hyrcanian Forest. *Remote Sensing*, 6 (5), 3624–3647.

Azad Hossain, A.K.M. and Easso, G., 2009. Microwave Remote Sensing of Soil Moisture in Semi-arid Environment. *In*: P.-G. Peter, ed. *Geoscience and Remote Sensing*. InTech.

Baghdadi, N., Aubert, M., Cerdan, O., Franchistéguy, L., Viel, C., Eric, M., Zribi, M., and Desprats, J.F., 2007. Operational Mapping of Soil Moisture Using Synthetic Aperture Radar Data: Application to the Touch Basin (France). *Sensors*, 7 (10), 2458–2483.

Baghdadi, N., Aubert, M., and Zribi, M., 2012. Use of TerraSAR-X Data to Retrieve Soil Moisture Over Bare Soil Agricultural Fields. *IEEE Geoscience and Remote Sensing Letters*, 9 (3), 512–516.

Baghdadi, N., Camus, P., Beaugendre, N., Issa, O.M., Zribi, M., Desprats, J.F., Rajot, J.L., Abdallah, C., and Sannier, C., 2011. Estimating Surface Soil Moisture from TerraSAR-X Data over Two Small Catchments in the Sahelian Part of Western Niger. *Remote Sensing*, 3 (6), 1266–1283.

Baghdadi, N., Choker, M., Zribi, M., Hajj, M., Paloscia, S., Verhoest, N., Lievens, H., Baup, F., and Mattia, F., 2016. A New Empirical Model for Radar Scattering from Bare Soil Surfaces. *Remote Sensing*, 8 (11), 920.

Baghdadi, N., King, C., Bourguignon, A., and Remond, A., 2002. Potential of ERS and Radarsat data for surface roughness monitoring over bare agricultural fields: Application to catchments in Northern France. *International Journal of Remote Sensing*, 23 (17), 3427–3442.

Baghdadi, N., Saba, E., Aubert, M., Zribi, M., and Baup, F., 2011. Comparison between backscattered TerraSAR signals and simulations from the radar backscattering models IEM, Oh, and Dubois. *IEEE Geoscience and Remote Sensing Letters, IEEE - Institute of Electrical and Electronics Engineers*, 6 (8), 1160–1164.

Bai, X., He, B., Li, X., Zeng, J., Wang, X., Wang, Z., Zeng, Y., and Su, Z., 2017. First Assessment of Sentinel-1A Data for Surface Soil Moisture Estimations Using a Coupled Water Cloud Model and Advanced Integral Equation Model over the Tibetan Plateau. *Remote Sensing*, 9 (7), 714.

Balenzano, A., 2013. On the use of temporal series of L- and X-band SAR data for soil moisture retrieval. Capitanata plain case study. *European Journal of Remote Sensing*, 721–737.

Barbosa, I., 2010. Mapping wetland environments in the Brazilian Savannah from high resolution IKONOS image data. *In*: *ISPRS TC VII Symposium*. Presented at the 100 Years ISPR, Vienna, Austria: Available online at.

Barrett, B.W., Dwyer, E., and Whelan, P., 2009. Soil Moisture Retrieval from Active Spaceborne Microwave Observations: An Evaluation of Current Techniques. *Remote Sensing*, 1 (3), 210–242.

Barsi, J., Lee, K., Kvaran, G., Markham, B., and Pedelty, J., 2014. The Spectral Response of the Landsat-8 Operational Land Imager. *Remote Sensing*, 6 (10), 10232–10251.

Bauer, A., 1974. Influence of soil organic matter on bulk density and available water capacity of soils. *Farm Research*, 31 (5).

Beck, A.D., 1964. The Kilombero valley of south-central Tanganyika. *East African Geographical Review*, (2), 37–43.

Bell, K.R., Blanchard, B.J., Schmugge, T.J., and Witczak, M.W., 1980. Analysis of surface moisture variations within large-field sites. *Water Resources Research*, 16 (4), 796–810.

Benjamin Leutner and Horning, N., 2017. RStoolbox: Tools for Remote Sensing Data Analysis : R package version 0.1.8.

Berberoğlu, S., Akin, A., Atkinson, P.M., and Curran, P.J., 2010. Utilizing image texture to detect land-cover change in Mediterranean coastal wetlands. *International Journal of Remote Sensing*, 31 (11), 2793–2815.

Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., Uppala, S., and Simmons, A., 2011. The ERA-Interim archive Version 2.0, ERA Report Series 1, ECMWF, Shinfield Park. *Reading, UK*, 13177.

Betbeder, J., Rapinel, S., Corgne, S., Pottier, E., and Hubert-Moy, L., 2015. TerraSAR-X dual-pol time-series for mapping of wetland vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Betbeder, J., Rapinel, S., Corpetti, T., Pottier, E., Corgne, S., and Hubert-Moy, L., 2014. Multitemporal classification of TerraSAR-X data for wetland vegetation mapping. *Journal of Applied Remote Sensing*, 8 (1), 083648.

Bindlish, R., 2000. Multifrequency Soil Moisture Inversion from SAR Measurements with the Use of IEM. *Remote Sensing of Environment*, 71 (1), 67–88.

Bindlish, R. and Barros, A.P., 2001. Parameterization of vegetation backscatter in radar-based, soil moisture estimation. *Remote Sensing of Environment*, 76 (1), 130–137.

Björnsson, H. and Venegas, S.A., 1997. *A Manual for EOF and SVD Analyses of Climatic Data*. McGill University, No. 97–1.

Böhme, B., Becker, M., and Diekkrüger, B., 2013. Calibrating a FDR sensor for soil moisture monitoring in a wetland in Central Kenya. *Physics and Chemistry of the Earth, Parts A/B/C*, 66, 101–111.

Bouwman, A., 1990. Global distribution of the major soils and land cover types. *Soils and the Greenhouse Effect*, 33–59.

Braune, E. and Xu, Y., 2010. The Role of Ground Water in Sub-Saharan Africa. *Ground Water*, 48 (2), 229–238.

Breiman, L., 2001. Random Forests. *Machine Learning*, 45 (1), 5–32.

Brocca, L., Ciabatta, L., Massari, C., Camici, S., and Tarpanelli, A., 2017. Soil Moisture for Hydrological Applications: Open Questions and New Opportunities. *Water*, 9 (2), 140.

Brocca, L., Melone, F., Moramarco, T., and Morbidelli, R., 2009. Soil moisture temporal stability over experimental areas in Central Italy. *Geoderma*, 148 (3–4), 364–374.

Brocca, L., Melone, F., Moramarco, T., and Morbidelli, R., 2010. Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resources Research*, 46 (2), n/a-n/a.

Brocca, L., Morbidelli, R., Melone, F., and Moramarco, T., 2007. Soil moisture spatial variability in experimental areas of central Italy. *Journal of Hydrology*, 333 (2–4), 356–373.

Brocca, L., Tullo, T., Melone, F., Moramarco, T., and Morbidelli, R., 2012. Catchment scale soil moisture spatial–temporal variability. *Journal of Hydrology*, 422–423, 63–75.

Brown, M.E., de Beurs, K., and Vrieling, A., 2010. The response of African land surface phenology to large scale climate oscillations. *Remote Sensing of Environment*, 114 (10), 2286–2296.

Brown, M.E. and de Beurs, K.M., 2008. Evaluation of multi-sensor semi-arid crop season parameters based on NDVI and rainfall. *Remote Sensing of Environment*, 112 (5), 2261–2271.

Burghof, S., Gabiri, G., Stumpp, C., Chesnaux, R., and Reichert, B., 2017. Development of a hydrogeological conceptual wetland model in the data-scarce north-eastern region of Kilombero Valley, Tanzania. *Hydrogeology Journal*.

Calow, R. and MacDonald, A., 2009. What will climate change mean for groundwater supply in Africa. *Overseas Development Institute (ODI) Background Note*.

Calow, R.C., Robins, N.S., Macdonald, A.M., Macdonald, D.M.J., Gibbs, B.R., Orpen, W.R.G., Mtembezeka, P., Andrews, A.J., and Appiah, S.O., 1997. Groundwater Management in Drought-prone Areas of Africa. *International Journal of Water Resources Development*, 13 (2), 241–262.

CBSA, UoR, TV, and INRA, 2014. *Algorithm Theoretical Basis Document (ATBD) for the SMOS Level 2 Soil Moisture Processor Development Continuation Project Issue 3.9*. ESA, Technical Report P38-43 No. ASC_SMPPD_ 037.

Chai, X., Zhang, T., Shao, Y., Gong, H., Liu, L., and Xie, K., 2015. Modeling and Mapping Soil Moisture of Plateau Pasture Using RADARSAT-2 Imagery. *Remote Sensing*, 7 (2), 1279–1299.

Chamundeeswari, V.V., Singh, D., and Singh, K., 2009. An Analysis of Texture Measures in PCA-Based Unsupervised Classification of SAR Images. *IEEE Geoscience and Remote Sensing Letters*, 6 (2), 214–218.

Chanzy, A., 1993. Basic soil surface characteristics derived from active microwave remote sensing. *Remote Sensing Reviews*, 7 (3–4), 303–319.

Charpentier, M.A. and Groffman, P.M., 1992. Soil moisture variability within remote sensing pixels. *Journal of Geophysical Research*, 97 (D17), 18987.

Chatziantoniou, A., Petropoulos, G.P., and Psomiadis, E., 2017. Co-Orbital Sentinel 1 and 2 for LULC Mapping with Emphasis on Wetlands in a Mediterranean Setting Based on Machine Learning. *Remote Sensing*, 9 (12), 1259.

Chen, C., ed., 2008. *Image processing for remote sensing*. Boca Raton, Fla.: CRC Press.

Chen, R.Z. and Wong, M.-H., 2016. Integrated wetlands for food production. *Environmental Research*, 148, 429–442.

Chen, T., de Jeu, R.A.M., Liu, Y.Y., van der Werf, G.R., and Dolman, A.J., 2014. Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia. *Remote Sensing of Environment*, 140, 330–338.

Chen, T., McVicar, T., Wang, G., Chen, X., de Jeu, R., Liu, Y., Shen, H., Zhang, F., and Dolman, A., 2016. Advantages of Using Microwave Satellite Soil Moisture over Gridded Precipitation Products and Land Surface Model Output in Assessing Regional Vegetation Water Availability and Growth Dynamics for a Lateral Inflow Receiving Landscape. *Remote Sensing*, 8 (5), 428.

Choker, M., Baghdadi, N., Zribi, M., El Hajj, M., Paloscia, S., Verhoest, N., Lievens, H., and Mattia, F., 2017. Evaluation of the Oh, Dubois and IEM Backscatter Models Using a Large Dataset of SAR Data and Experimental Soil Measurements. *Water*, 9 (1), 38.

Cihlar, J., Dobson, M.C., Schmugge, T., Hoogeboom, P., Janse, A.R.P., Baret, F., Guyot, G., Le Toan, T., and Pampaloni, P., 1987. Review Article Procedures for the description of agricultural crops and soils in optical and microwave remote sensing studies. *International Journal of Remote Sensing*, 8 (3), 427–439.

Climate-data.org, 2015. Climate-data.org Historical average temperature and rainfall. *URL www. en. climate-data. org*.

Cloude, S., 2010. *Polarisation applications in remote sensing*. Oxford: Oxford University Press.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, 37 (1), 35–46.

Congalton, R.G. and Green, K., 2009. *Assessing the accuracy of remotely sensed data: principles and practices*. 2nd ed. Boca Raton: CRC Press/Taylor & Francis.

Cosh, M.H., Jackson, T.J., Moran, S., and Bindlish, R., 2008. Temporal persistence and stability of surface soil moisture in a semi-arid watershed. *Remote Sensing of Environment*, 112 (2), 304–313.

Crow, W.T., Berg, A.A., Cosh, M.H., Loew, A., Mohanty, B.P., Panciera, R., de Rosnay, P., Ryu, D., and Walker, J.P., 2012. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, 50 (2).

Dabrowska-Zielinska, K., Budzynska, M., Kowalik, W., and Turlej, K., 2010. Soil moisture and evapotranspiration of wetlands vegetation habitats retrieved from satellite images. *Hydrology and Earth System Sciences Discussions*, 7 (4), 5929–5955.

Dabrowska-Zielinska, K., Budzynska, M., Tomaszewska, M., Malinska, A., Gatkowska, M., Bartold, M., and Malek, I., 2016. Assessment of Carbon Flux and Soil Moisture in Wetlands Applying Sentinel-1 Data. *Remote Sensing*, 8 (9), 756.

Daddow, R.L. and Warrington, G.E., 1983. Growth limiting soil bulk densities as influenced by soil texture.

Daniel, S., Gabiri, G., Kirimi, F., Glasner, B., Näschen, K., Leemhuis, C., Steinbach, S., and Mtei, K., 2017. Spatial Distribution of Soil Hydrological Properties in the Kilombero Floodplain, Tanzania. *Hydrology*, 4 (4), 57.

Das, K. and Paul, P.K., 2015a. Soil moisture retrieval model by using RISAT-1, C-band data in tropical dry and sub-humid zone of Bankura district of India. *The Egyptian Journal of Remote Sensing and Space Science*, 18 (2), 297–310.

Das, K. and Paul, P.K., 2015b. Present status of soil moisture estimation by microwave remote sensing. *Cogent Geoscience*, 1 (1).

Dawson, M.S., Fung, A.K., and Manry, M.T., 1997. A robust statistical-based estimator for soil moisture retrieval from radar measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (1), 57–67.

De Lannoy, G.J.M., Houser, P.R., Verhoest, N.E.C., Pauwels, V.R.N., and Gish, T.J., 2007. Upscaling of point soil moisture measurements to field averages at the OPE3 test site. *Journal of Hydrology*, 343 (1–2), 1–11.

De Lannoy, G.J.M., Verhoest, N.E.C., Houser, P.R., Gish, T.J., and Van Meirvenne, M., 2006. Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE3). *Journal of Hydrology*, 331 (3–4), 719–730.

Decker, M., Brunke, M.A., Wang, Z., Sakaguchi, K., Zeng, X., and Bosilovich, M.G., 2012. Evaluation of the Reanalysis Products from GSFC, NCEP, and ECMWF Using Flux Tower Observations. *Journal of Climate*, 25 (6), 1916–1944.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137 (656), 553–597.

Dobson, M.C. and Ulaby, F., 1981. Microwave Backscatter Dependence on Surface Roughness, Soil Moisture, And Soil Texture: Part III-Soil Tension. *IEEE Transactions on Geoscience and Remote Sensing*, GE-19 (1), 51–61.

Dorigo, W., de Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., and Fernández-Prieto, D., 2012. Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture: Trends in Remotely Sensed soil moisture. *Geophysical Research Letters*, 39 (18).

Dorigo, W.A., Gruber, A., De Jeu, R.A.M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R.M., and Kidd, R., 2015. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment*, 162, 380–395.

Dorigo, W.A., Scipal, K., Parinussa, R.M., Liu, Y.Y., Wagner, W., de Jeu, R.A.M., and Naeimi, V., 2010. Error characterisation of global active and passive microwave soil moisture datasets. *Hydrology and Earth System Sciences*, 14 (12), 2605–2616.

Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T., 2011. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15 (5), 1675–1698.

Douville, H., Viterbo, P., Mahfouf, J.-F., and Beljaars, A.C.M., 2000. Evaluation of the Optimum Interpolation and Nudging Techniques for Soil Moisture Analysis Using FIFE Data. *Monthly Weather Review*, 128 (6), 1733–1756.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36.

Dubois, P.C., van Zyl, J., and Engman, T., 1995. Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing*, 33 (4), 915–926.

D'Urso, G. and Minacapilli, M., 2006. A semi-empirical approach for surface soil water content estimation from radar data without a-priori information on surface roughness. *Journal of Hydrology*, 321 (1–4), 297–310.

Eckert, S., 2012. Improved Forest Biomass and Carbon Estimations Using Texture Measures from WorldView-2 Satellite Data. *Remote Sensing*, 4 (12), 810–829.

Eitel, J.U.H., Vierling, L.A., Litvak, M.E., Long, D.S., Schulthess, U., Ager, A.A., Krofcheck, D.J., and Stoscheck, L., 2011. Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sensing of Environment*, 115 (12), 3640–3646.

Eklundh, L. and Jönsson, P., 2015. TIMESAT: A Software Package for Time-Series Processing and Assessment of Vegetation Dynamics. *In*: C. Kuenzer, S. Dech, and W. Wagner, eds. *Remote Sensing Time Series*. Cham: Springer International Publishing, 141–158.

Eklundh, L. and Jönsson, P., 2017. TIMESAT 3.3 with seasonal trend decomposition and parallel processing Software Manual.

El Hajj, M., Baghdadi, N., Zribi, M., Belaud, G., Cheviron, B., Courault, D., and Charron, F., 2016. Soil moisture retrieval over irrigated grassland using X-band SAR data. *Remote Sensing of Environment*, 176, 202–218.

Emerson, C.W., Lam, N.S., and Quattrochi, D.A., 2005. A comparison of local variance, fractal dimension, and Moran's *I* as aids to multispectral image classification. *International Journal of Remote Sensing*, 26 (8), 1575–1588.

Entekhabi, D., Reichle, R.H., Koster, R.D., and Crow, W.T., 2010. Performance Metrics for Soil Moisture Retrievals and Application Requirements. *Journal of Hydrometeorology*, 11 (3), 832–840.

Entin, J.K., Robock, A., Vinnikov, K.Y., Hollinger, S.E., Liu, S., and Namkhai, A., 2000. Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research: Atmospheres*, 105 (D9), 11865–11877.

ESA Sentinels Scientific Data Hub, 2016. Available from: https://scihub.copernicus.eu/dhus/#/home [Accessed 2 Oct 2016].

European Space Agency, 2013. Sentinel-1 User Handbook and Exploration Tools.

Evans, T.L., Costa, M., Telmer, K., and Silva, T.S.F., 2010. Using ALOS/PALSAR and RADARSAT-2 to Map Land Cover and Seasonal Inundation in the Brazilian Pantanal. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3 (4), 560–575.

Evanylo, G.K. and McGuinn, R., 2000. Agricultural management practices and soil quality: measuring, assessing, and comparing laboratory and field test kit indicators of soil quality atributes.

Famiglietti, J.S., Devereaux, J.A., Laymon, C.A., Tsegaye, T., Houser, P.R., Jackson, T.J., Graham, S.T., Rodell, M., and van Oevelen, P.J., 1999. Ground-based investigation of soil moisture variability within remote sensing footprints During the Southern Great Plains 1997 (SGP97) Hydrology Experiment. *Water Resources Research*, 35 (6), 1839–1851.

Famiglietti, J.S., Rudnicki, J.W., and Rodell, M., 1998. Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. *Journal of Hydrology*, 210 (1–4), 259–281.

Famiglietti, J.S., Ryu, D., Berg, A.A., Rodell, M., and Jackson, T.J., 2008. Field observations of soil moisture variability across scales. *Water Resources Research*, 44 (1), W01423.

Fernandez, J.C., 2010. Characterization of Surface Roughness of Bare Agricultural Soils using LIDAR. A Dissertation for the degree of Doctor of Philosophy. University of Florida.

Fisher, J.R.B., Acosta, E.A., Dennedy-Frank, P.J., Kroeger, T., and Boucher, T.M., 2017. Impact of satellite imagery spatial resolution on land use classification accuracy and modeled water quality. *Remote Sensing in Ecology and Conservation*.

Food and Agriculture Organization of the United Nations, 2014. *Climate-smart agriculture sourcebook.*

Frenken, K. and Mharapara, I., 2001. Wetland Development And Management In SADC Countries. *In*: *Proceedings of a sub-regional workshop*. Harare, Zimbabwe: FAO, SAFR.

Furtado, L.F. de A., Silva, T.S.F., and Novo, E.M.L. de M., 2016. Dual-season and full-polarimetric C band SAR assessment for vegetation mapping in the Amazon várzea wetlands. *Remote Sensing of Environment*, 174, 212–222.

Gabiri, G., Burghof, S., Diekkrüger, B., Leemhuis, C., Steinbach, S., and Näschen, K., 2018. Modeling Spatial Soil Water Dynamics in a Tropical Floodplain, East Africa. *Water*, 10 (2), 191.

Gardner, R., Connolly, K.D., and Bamba, A., 2009. African Wetlands of International Importance: Assessments of Benefits Associated with Designations under the Ramsar Convention. *The Georgetown International Environmental Law Review*, XXI (2), 257–294.

Garmin International, 2005. Garmin eTrex Series.

Gebbers, R. and Adamchuk, V.I., 2010. Precision Agriculture and Food Security. *Science*, 327 (5967), 828–831.

Geological Survey of Tanzania, 2007. Geological and Mineral Information System.

Gherboudj, I., Magagi, R., Berg, A.A., and Toth, B., 2011. Soil moisture retrieval over agricultural fields from multi-polarized and multi-angular RADARSAT-2 SAR data. *Remote Sensing of Environment*, 115 (1), 33–43.

Ghimire, B., Rogan, J., and Miller, J., 2010. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters*, 1 (1), 45–54.

Gilbert, R.O., 1987. *Statistical methods for environmental pollution monitoring*. New York: Wiley.

Gorrab, A., Zribi, M., Baghdadi, N., Mougenot, B., and Chabaane, Z., 2015. Potential of X-Band TerraSAR-X and COSMO-SkyMed SAR Data for the Assessment of Physical Soil Parameters. *Remote Sensing*, 7 (1), 747–766.

Gorrab, A., Zribi, M., Baghdadi, N., Mougenot, B., Fanise, P., and Chabaane, Z., 2015. Retrieval of Both Soil Moisture and Texture Using TerraSAR-X Images. *Remote Sensing*, 7 (8), 10098–10116.

Grims, M., Atzberger, C., Bauer, T., Strauss, P., and Mansberger, 2014. Low-cost Terrestrial Photogrammetry as a Tool for a Sample-Based Assessment of Soil Roughness. *Photogrammetrie - Fernerkundung - Geoinformation*, 2014 (5), 313–323.

Guo, M., Li, J., Sheng, C., Xu, J., and Wu, L., 2017. A Review of Wetland Remote Sensing. *Sensors*, 17 (4), 777.

Gupta, R.P., Tiwari, R.K., Saini, V., and Srivastava, N., 2013. A Simplified Approach for Interpreting Principal Component Images. *Advances in Remote Sensing*, 02 (02), 111–119.

Haack, B., 1996. Monitoring wetland changes with remote sensing: An East African example. *Environmental Management*, 20 (3), 411–419.

Hajnsek, I., Pottier, E., and Cloude, S.R., 2003. Inversion of surface parameters from polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (4), 727–744.

Hall-Beyer, 2007. GLCM Texture Tutorial [online]. Available from: http://www.fp.ucalgary.ca/mhallbey/tutorial.htm [Accessed 2 Oct 2016].

Hannachi, A., Jolliffe, I.T., and Stephenson, D.B., 2007. Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, 27 (9), 1119–1152.

Haralick, R.M., 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67 (5), 786–804.

Haralick, R.M., Shanmugam, K., and Dinstein, I., 1973. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3 (6), 610–621.

Hassan-Esfahani, L., Torres-Rua, A., Jensen, A., and McKee, M., 2015. Assessment of Surface Soil Moisture Using High-Resolution Multi-Spectral Imagery and Artificial Neural Networks. *Remote Sensing*, 7 (3), 2627–2646.

Heathman, G.C., Cosh, M.H., Han, E., Jackson, T.J., McKee, L., and McAfee, S., 2012. Field scale spatiotemporal analysis of surface soil moisture for evaluating point-scale in situ networks. *Geoderma*, 170, 195–205.

Henderson, F.M. and Lewis, A.J., 2008. Radar detection of wetland ecosystems: a review. *International Journal of Remote Sensing*, 29 (20), 5809–5835.

Hengl, T., 2009. *A practical guide to geostatistical mapping*. 2nd extended ed. Amsterdam: Hengl.

Hengl, T., European commission, Joint research centre, and Institute for environment and sustainability (Ispra, I., 2007. *A practical guide to geostatistical mapping of environmental variables*. Luxembourg: Publications Office.

Hillel, D., 1998. *Environmental soil physics*. San Diego, CA: Academic Press.

Hirosawa, H., Komiyama, S., and Matsuzaka, Y., 1978. Cross-polarized radar backscatter from moist soil. *Remote Sensing of Environment*, 7 (3), 211–217.

Holah, N., Baghdadi, N., Zribi, M., Bruand, A., and King, C., 2005. Potential of ASAR/ENVISAT for the characterization of soil surface parameters over bare agricultural fields. *Remote Sensing of Environment*, 96 (1), 78–86.

Homa Zakeri, Fumio Yamazaki, and Wen Liu, 2017. Texture Analysis and Land Cover Classification of Tehran Using Polarimetric Synthetic Aperture Radar Imagery. *Applied Sciences*, 7 (5), 452.

Hornbuckle, B.K., Patton, J.C., VanLoocke, A., Suyker, A.E., Roby, M.C., Walker, V.A., Iyer, E.R., Herzmann, D.E., and Endacott, E.A., 2016. SMOS optical thickness changes in response to the growth and development of crops, crop management, and weather. *Remote Sensing of Environment*, 180, 320–333.

Hossain, A.K.M. and Easson, G., 2016. Soil Moisture Estimation in South-Eastern New Mexico Using High Resolution Synthetic Aperture Radar (SAR) Data. *Geosciences*, 6 (1), 1.

Hu, H. and Ban, Y., 2014. Unsupervised Change Detection in Multitemporal SAR Images Over Large Urban Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (8), 3248–3261.

Hupet, F. and Vanclooster, M., 2002. Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field. *Journal of Hydrology*, 261 (1–4), 86–101.

Hussain, M., Chen, D., Cheng, A., Wei, H., and Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, 91–106.

Inglada, J., Vincent, A., Arias, M., and Marais-Sicre, C., 2016. Improved Early Crop Type Identification By Joint Use of High Temporal Resolution SAR And Optical Image Time Series. *Remote Sensing*, 8 (5), 362.

ISG, 2016. Illovo Sugar [online]. Available from: https://www.illovosugar.co.za/About-us/Tanzania [Accessed 2 Oct 2016].

Jackson, T.J., McNairn, H., Weltz, M.A., Brisco, B., and Brown, R., 1997. First order surface roughness correction of active microwave observations for estimating soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (4), 1065–1069.

Jacobs, J., 2004. SMEX02: Field scale variability, time stability and similarity of soil moisture. *Remote Sensing of Environment*, 92 (4), 436–446.

Jacome, A., Bernier, M., Chokmani, K., Gauthier, Y., Poulin, J., and De Sève, D., 2013. Monitoring Volumetric Surface Soil Moisture Content at the La Grande Basin Boreal Wetland by Radar Multi Polarization Data. *Remote Sensing*, 5 (10), 4919–4941.

James, G., Witten, D., Hastie, T., and Tibshirani, R., eds., 2013. *An introduction to statistical learning: with applications in R*. New York: Springer.

Jenny, H., 1971. *Factors of soil formation: a system of quantitative pedology*. McGraw-HillAnn Arbor: University Microfilms.

Jia, K., Li, Q., Tian, Y., Wu, B., Zhang, F., and Meng, J., 2012. Crop classification using multi-configuration SAR data in the North China Plain. *International Journal of Remote Sensing*, 33 (1), 170–183.

Jiankang Ji, van der Keur, P., Thomsen, A., and Skriver, H., 1996. Soil moisture retrieval using the Danish L- & C-band polarimetric SAR. IEEE, 1300–1302.

Jing, L., Qingwen, M., Wenhua, L., Yanying, B., C, D.B.G., and Zheng, Y., 2014. Spatial Variability Analysis of Soil Nutrients Based on GIS and Geostatistics: A Case Study of Yisa Township, Yunnan, China. *Journal of Resources and Ecology*, 5 (4), 348–355.

Johansen, K., Coops, N.C., Gergel, S.E., and Stange, Y., 2007. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment*, 110 (1), 29–44.

Jolliffe, I.T., 2002. *Principal component analysis*. 2nd ed. New York: Springer.

Jonsson, P. and Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (8), 1824–1832.

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

Julien, Y. and Sobrino, J.A., 2009. Global land surface phenology trends from GIMMS database. *International Journal of Remote Sensing*, 30 (13), 3495–3513.

Junk, W.J., An, S., Finlayson, C.M., Gopal, B., Květ, J., Mitchell, S.A., Mitsch, W.J., and Robarts, R.D., 2013. Current state of knowledge regarding the world’s wetlands and their future under global climate change: a synthesis. *Aquatic Sciences*, 75 (1), 151–167.

Justice, C.O., Holben, B.N., and Gwynne, M.D., 1986. Monitoring East African vegetation using AVHRR data. *International Journal of Remote Sensing*, 7 (11), 1453–1474.

Kalinga, G.. and Shayo, E., 1997. *Wetland characterization and classification for sustainable agricultural development*. Harare.

Kangalawe, R.Y.M. and Liwenga, E.T., 2005. Livelihoods in the wetlands of Kilombero Valley in Tanzania: Opportunities and challenges to integrated water resource management. *Physics and Chemistry of the Earth, Parts A/B/C*, 30 (11–16), 968–975.

Kato, F., 2007. Development of a Major Rice Cultivation Area in the Kilombero Valley, Tanzania. *African Study Monographs*, (36), 3–18.

Kerr, Y.H., Waldteufel, P., Richaume, P., Wigneron, J.P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S.E., Leroux, D., Mialon, A., and Delwart, S., 2012. The SMOS Soil Moisture Retrieval Algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (5), 1384–1403.

Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., and Berger, M., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, 39 (8), 1729–1735.

Kiage, L.M., Liu, K. - B., Walker, N.D., Lam, N., and Huh, O.K., 2007. Recent land- cover/use change associated with land degradation in the Lake Baringo catchment, Kenya, East Africa: evidence from Landsat TM and ETM+. *International Journal of Remote Sensing*, 28 (19), 4285–4309.

Kim, C., 2016. Land use classification and land use change analysis using satellite images in Lombok Island, Indonesia. *Forest Science and Technology*, 12 (4), 183–191.

Kim, M., Warner, T.A., Madden, M., and Atkinson, D.S., 2011. Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: scale, texture and image objects. *International Journal of Remote Sensing*, 32 (10), 2825–2850.

Kim, S.-B., van Zyl, J.J., Johnson, J.T., Moghaddam, M., Tsang, L., Colliander, A., Dunbar, R.S., Jackson, T.J., Jaruwatanadilok, S., West, R., Berg, A., Caldwell, T., Cosh, M.H., Goodrich, D.C., Livingston, S., Lopez-Baeza, E., Rowlandson, T., Thibeault, M., Walker, J.P., Entekhabi, D., Njoku, E.G., O’Neill, P.E., and Yueh, S.H., 2017. Surface Soil Moisture Retrieval Using the L-Band Synthetic Aperture Radar Onboard the Soil Moisture Active–Passive Satellite and Evaluation at Core Validation Sites. *IEEE Transactions on Geoscience and Remote Sensing*, 55 (4), 1897–1914.

Kimble, J.M., Follett, R.F., and Stewart, B.A., 2000. *Assessment methods for soil carbon*. CRC press.

Kirimi, F., Kuria, D.N., Thonfeld, F., Amler, E., Mubea, K., Misana, S., and Menz, G., 2016. Influence of Vegetation Cover on the Oh Soil Moisture Retrieval Model: A Case Study of the Malinda Wetland, Tanzania. *Advances in Remote Sensing*, 05 (01), 28–42.

Klemas, V., 2011. Remote Sensing of Wetlands: Case Studies Comparing Practical Techniques. *Journal of Coastal Research*, 27, 418–427.

Klisch, A. and Atzberger, C., 2016. Operational Drought Monitoring in Kenya Using MODIS NDVI Time Series. *Remote Sensing*, 8 (4), 267.

Kong, X. and Dorling, S.R., 2008. Near- surface soil moisture retrieval from ASAR Wide Swath imagery using a Principal Component Analysis. *International Journal of Remote Sensing*, 29 (10), 2925–2942.

Koppe, W., Gnyp, M.L., Hütt, C., Yao, Y., Miao, Y., Chen, X., and Bareth, G., 2013. Rice monitoring with multi-temporal and dual-polarimetric TerraSAR-X data. *International Journal of Applied Earth Observation and Geoinformation*, 21, 568–576.

Kornelsen, K.C. and Coulibaly, P., 2013. Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. *Journal of Hydrology*, 476, 460–489.

Korres, W., Reichenau, T.G., Fiener, P., Koyama, C.N., Bogena, H.R., Cornelissen, T., Baatz, R., Herbst, M., Diekkrüger, B., Vereecken, H., and Schneider, K., 2015. Spatio-temporal soil moisture patterns – A meta-analysis using plot to catchment scale data. *Journal of Hydrology*, 520, 326–341.

Koutsouris, A.J., Chen, D., and Lyon, S.W., 2016. Comparing global precipitation data sets in eastern Africa: a case study of Kilombero Valley, Tanzania: Comparing global precipitation data sets in Tanzania, East Africa. *International Journal of Climatology*, 36 (4), 2000–2014.

Kuria, D.N., Menz, G., Misana, S., Mwita, E., Thamm, H.-P., Alvarez, M., Mogha, N., Becker, M., and Oyieke, H., 2014. Seasonal Vegetation Changes in the Malinda Wetland Using Bi-Temporal, Multi-Sensor, Very High Resolution Remote Sensing Data Sets. *Advances in Remote Sensing*, 03 (01), 33–48.

Kurvonen, L. and Hallikainen, M.T., 1999. Textural information of multitemporal ERS-1 and JERS-1 SAR images with applications to land and forest type classification in boreal zone. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (2), 680–689.

Kussul, N., Shelestov, A., and Skakun, S., 2011. Flood Monitoring from SAR Data. *In*: F. Kogan, A. Powell, and O. Fedorov, eds. *Use of Satellite and In-Situ Data to Improve Sustainability*. Dordrecht: Springer Netherlands, 19–29.

KVFP, 2002. Kilombero Valley Floodplain | Ramsar Sites Information Service [online]. Available from: https://rsis.ramsar.org/ris/1173 [Accessed 2 Oct 2016].

KVTC, 2009. KVTC-TZ [online]. Available from: http://kvtc-tz.com/ [Accessed 2 Oct 2016].

Le Blanc, D., 2015. Towards Integration at Last? The Sustainable Development Goals as a Network of Targets: The sustainable development goals as a network of targets. *Sustainable Development*, 23 (3), 176–187.

Le Morvan, A., Zribi, M., Baghdadi, N., and Chanzy, A., 2008. Soil Moisture Profile Effect on Radar Signal Measurement. *Sensors*, 8 (1), 256–270.

Leemhuis, C., Thonfeld, F., Näschen, K., Steinbach, S., Muro, J., Strauch, A., López, A., Daconto, G., Games, I., and Diekkrüger, B., 2017. Sustainability in the Food-Water-Ecosystem Nexus: The Role of Land Use and Land Cover Change for Water Resources and Ecosystems in the Kilombero Wetland, Tanzania. *Sustainability*, 9 (9), 1513.

Lerner, D.N. and Harris, B., 2009. The relationship between land use and groundwater resources and quality. *Land Use Policy*, 26, S265–S273.

Leroux, D.J., Kerr, Y.H., Al Bitar, A., Bindlish, R., Jackson, T.J., Berthelot, B., and Portet, G., 2014. Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (3), 1562–1571.

Leroux, D.J., Kerr, Y.H., Richaume, P., and Berthelot, B., 2011. Estimating SMOS error structure using triple collocation. IEEE, 24–27.

Li, F., Qin, X., Xie, Y., Chen, X., Hu, J., Liu, Y., and Hou, Z., 2013. Physiological mechanisms for plant distribution pattern: responses to flooding and drought in three wetland plants from Dongting Lake, China. *Limnology*, 14 (1), 71–76.

Li, X., Li, G., and Zhang, Y., 2014. Identifying Major Factors Affecting Groundwater Change in the North China Plain with Grey Relational Analysis. *Water*, 6 (6), 1581–1600.

Lievens, H. and Verhoest, N.E.C., 2011. On the Retrieval of Soil Moisture in Wheat Fields From L-Band SAR Based on Water Cloud Modeling, the IEM, and Effective Roughness Parameters. *IEEE Geoscience and Remote Sensing Letters*, 8 (4), 740–744.

Lievens, H. and Verhoest, N.E.C., 2012. Spatial and temporal soil moisture estimation from RADARSAT-2 imagery over Flevoland, The Netherlands. *Journal of Hydrology*, 456–457, 44–56.

Lievens, H., Vernieuwe, H., Álvarez-Mozos, J., De Baets, B., and Verhoest, N.E.C., 2009. Error in Radar-Derived Soil Moisture due to Roughness Parameterization: An Analysis Based on Synthetical Surface Profiles. *Sensors*, 9 (2), 1067–1093.

Lillesand, T., Kiefer, R.W., and Chipman, J., 2014. *Remote sensing and image interpretation*. John Wiley & Sons.

Lillesand, T.M., Kiefer, R.W., and Chipman, J.W., 2004. *Remote sensing and image interpretation*. 5th ed. New York: Wiley.

van der Linden, S., Rabe, A., Held, M., Jakimow, B., Leitão, P.J., Okujeni, A., Schwieder, M., Suess, S., and Hostert, P., 2015. The EnMAP-Box—A toolbox and application programming interface for EnMAP data processing. *Remote Sensing*, 7 (9), 11249–11266.

Liu, Z., Teng, W., Ostrenga, D., Albayrak, R., Savtchenko, A., Yang, W., Vollmer, B., and Meyer, D., 2017. Global Precipitation Products at NASA GES DISC for Supporting Agriculture Research and Applications.

Lopez-Sanchez, J.M., Ballester-Berman, J.D., and Hajnsek, I., 2011. First Results of Rice Monitoring Practices in Spain by Means of Time Series of TerraSAR-X Dual-Pol Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4 (2), 412–422.

Lorenz, E., 1956. *Empirical Orthogonal Functions and Statistical Weather Prediction*. Massachusetts Institute of Technology, Scientific Report No. 1.

Lu, D. and Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28 (5), 823–870.

Mahdavi, S., Salehi, B., Amani, M., Granger, J.E., Brisco, B., Huang, W., and Hanson, A., 2017. Object-Based Classification of Wetlands in Newfoundland and Labrador Using Multi-Temporal PolSAR Data. *Canadian Journal of Remote Sensing*, 43 (5), 432–450.

Mahmoud, A., Elbialy, S., Pradhan, B., and Buchroithner, M., 2011. Field-based landcover classification using TerraSAR-X texture analysis. *Advances in Space Research*, 48 (5), 799–805.

Marzahn, P., Seidel, M., and Ludwig, R., 2012. Decomposing Dual Scale Soil Surface Roughness for Microwave Remote Sensing Applications. *Remote Sensing*, 4 (12), 2016–2032.

Masih, I., Maskey, S., Mussá, F.E.F., and Trambauer, P., 2014. A review of droughts on the African continent: a geospatial and long-term perspective. *Hydrology and Earth System Sciences*, 18 (9), 3635–3649.

Mattia, F., Le Toan, T., Souyris, J.-C., De Carolis, C., Floury, N., Posa, F., and Pasquariello, N.G., 1997. The effect of surface roughness on multifrequency polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (4), 954–966.

Mattia, F., Satalino, G., Dente, L., and Pasquariello, G., 2006. Using a priori information to improve soil moisture retrieval from ENVISAT ASAR AP data in semiarid regions. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (4), 900–912.

McNairn, H., Shang, J., Jiao, X., and Deschamps, B., 2012. Establishing crop productivity using RADARSAT-2. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B8, 283–287.

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C.D., and Verdin, J.P., 2017. A land data assimilation system for sub-Saharan Africa food and water security applications. *Scientific Data*, 4, 170012.

McNally, A., Shukla, S., Arsenault, K.R., Wang, S., Peters-Lidard, C.D., and Verdin, J.P., 2016. Evaluating ESA CCI soil moisture in East Africa. *International Journal of Applied Earth Observation and Geoinformation*, 48, 96–109.

Milder, J., Buck, L., Hart, A., Scherr, S., and Shames, S., 2013. A framework for Agriculture Green Growth: Greenprint for the Southern Agricultural Growth Corridor of Tanzania (SAGCOT). *Dar es Salaam: SAGCOT Centre*.

Millennium Development Goals and Nations Unies, 2011. *The Millennium Development Goals report 2011*. New York: United Nations.

Mombo, F., Speelman, S., Huylenbroeck, G.V., Hella, J., Pantaleo, M., and Moe, S., 2011. Ratification of the Ramsar convention and sustainable wetlands management: Situation analysis of the Kilombero Valley wetlands in Tanzania. *Journal of Agricultural Extension and Rural Development*, 3 (9), 153–164.

Moran, M.S., Peters-Lidard, C.D., Watts, J.M., and McElroy, S., 2004. Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian Journal of Remote Sensing*, 30 (5), 805–826.

Morandeira, N., Grings, F., Facchinetti, C., and Kandus, P., 2016. Mapping Plant Functional Types in Floodplain Wetlands: An Analysis of C-Band Polarimetric SAR Data from RADARSAT-2. *Remote Sensing*, 8 (3), 174.

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., and Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1 (1), 6–43.

Moser, L., Schmitt, A., Wendleder, A., and Roth, A., 2016. Monitoring of the Lac Bam Wetland Extent Using Dual-Polarized X-Band SAR Data. *Remote Sensing*, 8 (4), 302.

Moser, L., Voigt, S., Schoepfer, E., and Palmer, S., 2014. Multitemporal Wetland Monitoring in Sub-Saharan West-Africa Using Medium Resolution Optical Satellite Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (8), 3402–3415.

Mosleh, M., Hassan, Q., and Chowdhury, E., 2015. Application of Remote Sensors in Mapping Rice Area and Forecasting Its Production: A Review. *Sensors*, 15 (1), 769–791.

Moukana, J.A. and Koike, K., 2008. Geostatistical model for correlating declining groundwater levels with changes in land cover detected from analyses of satellite images. *Computers & Geosciences*, 34 (11), 1527–1540.

Mountrakis, G., Im, J., and Ogole, C., 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66 (3), 247–259.

Mwita, E., Menz, G., Misana, S., Becker, M., Kisanga, D., and Boehme, B., 2013. Mapping small wetlands of Kenya and Tanzania using remote sensing techniques. *International Journal of Applied Earth Observation and Geoinformation*, 21, 173–183.

Myint, S.W., 2003. Fractal approaches in texture analysis and classification of remotely sensed data: Comparisons with spatial autocorrelation techniques and simple descriptive statistics. *International Journal of Remote Sensing*, 24 (9), 1925–1947.

National Bureau of Statistics, 2015. Environmental Statistics in Tanzania Mainland.

Ndou, N.N., Palamuleni, L.G., and Ramoelo, A., 2017. Modelling depth to groundwater level using SEBAL-based dry season potential evapotranspiration in the upper Molopo River Catchment, South Africa. *The Egyptian Journal of Remote Sensing and Space Science*.

Nichols, S., 2011. Review and evaluation of remote sensing methods for soil-moisture estimation. *Journal of Photonics for Energy*, 028001.

Nicol, A., Langan, S., Victor, M., and Gonsalves, J., 2015. *Water-smart agriculture in East Africa*. IWMI.

Notarnicola, C. and Solorz, R., 2014. Integration of Remotely Sensed Images and Electromagnetic Models into a Bayesian Approach for Soil Moisture Content Retrieval: Methodology and Effect of Prior Information. *In*: M.S.F. Nezhad, ed. *Dynamic Programming and Bayesian Inference, Concepts and Applications*. InTech.

Ntongani, W.A., Munishi, P.K.T., More, S.R., and Kashaigili, J.J., 2014. Local Knowledge on the Influence of Land Use/Cover Changes and Conservation Threats on Avian Community in the Kilombero Wetlands, Tanzania. *Open Journal of Ecology*, 04 (12), 723–731.

Ochsner, T.E., Cosh, M.H., Cuenca, R.H., Dorigo, W.A., Draper, C.S., Hagimoto, Y., Kerr, Y.H., Njoku, E.G., Small, E.E., and Zreda, M., 2013. State of the Art in Large-Scale Soil Moisture Monitoring. *Soil Science Society of America Journal*, 77 (6), 1888.

Oh, Y., Sarabandi, K., and Ulaby, F.T., 1992. An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 30 (2), 370–381.

Oliva, R., Daganzo, E., Kerr, Y.H., Mecklenburg, S., Nieto, S., Richaume, P., and Gruhier, C., 2012. SMOS Radio Frequency Interference Scenario: Status and Actions Taken to Improve the RFI Environment in the 1400-1427-MHz Passive Band. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (5), 1427–1439.

Owe, M., de Jeu, R., and Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. *Journal of Geophysical Research*, 113 (F1).

Owuor, S.O., Butterbach-Bahl, K., Guzha, A.C., Rufino, M.C., Pelster, D.E., Díaz-Pinés, E., and Breuer, L., 2016. Groundwater recharge rates and surface runoff response to land use and land cover changes in semi-arid environments. *Ecological Processes*, 5 (1).

Paloscia, S., Pampaloni, P., Pettinato, S., and Santi, E., 2008. A Comparison of Algorithms for Retrieving Soil Moisture from ENVISAT/ASAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (10), 3274–3284.

Pan, Y., Gong, H., ZHou, D., Li, X., and Nakagoshi, N., 2011. Impact of land use change on groundwater recharge in Guishui River Basin, China. *Chinese Geographical Science*, 21 (6), 734–743.

Panciera, R., Tanase, M.A., Lowell, K., and Walker, J.P., 2014. Evaluation of IEM, Dubois, and Oh Radar Backscatter Models Using Airborne L-Band SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (8), 4966–4979.

Pande-Chhetri, R., Abd-Elrahman, A., Liu, T., Morton, J., and Wilhelm, V.L., 2017. Object-based classification of wetland vegetation using very high-resolution unmanned air system imagery. *European Journal of Remote Sensing*, 50 (1), 564–576.

Paneque-Gálvez, J., Mas, J., Moré, G., Cristóbal, J., Orta-Martínez, M., Luz, A.C., Guèze, M., Macía, M.J., and Reyes-García, V., 2013. Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. *International Journal of Applied Earth Observation and Geoinformation*, 23, 372–383.

Peel, M.C., Finlayson, B.L., and McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences discussions*, 4 (2), 439–473.

Peltonen-Sainio, P., Jauhiainen, L., and Hakala, K., 2008. Climate change and prolongation of growing season: changes in regional potential for field crop production in Finland. *Agricultural and Food Science*, 18 (3–4), 171.

Petropoulos, G.P. and Albergel, C., eds., 2013. *Remote sensing of energy fluxes and soil moisture content*. Boca Raton, Fla.: CRC Press.

Petropoulos, G.P., Ireland, G., and Barrett, B., 2015. Surface soil moisture retrievals from remote sensing: Current status, products & future trends. *Physics and Chemistry of the Earth, Parts A/B/C*, 83–84, 36–56.

Pierce, L., Xie, H., Dobson, M.C., and Ulaby, F., 1998. Texture features for classification with ERS/JERS composites. Presented at the Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS'98. 1998 IEEE International, IEEE, 348–350.

Pierdicca, N., Pulvirenti, L., and Pace, G., 2014. A Prototype Software Package to Retrieve Soil Moisture From Sentinel-1 Data by Using a Bayesian Multitemporal Algorithm. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (1), 153–166.

Planet [online], 2016. *Planet*. Available from: https://planet.com/ [Accessed 2 Oct 2016].

Prabhakar, A. and Tiwari, H., 2015. Land use and land cover effect on groundwater storage. *Modeling Earth Systems and Environment*, 1 (4).

Pradhan, B., Hagemann, U., Shafapour Tehrany, M., and Prechtel, N., 2014. An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. *Computers & Geosciences*, 63, 34–43.

Preisendorfer, R.W., Mobley, C.D., and Barnett, T.P., 1988. The principal discriminant method of prediction: Theory and evaluation. *Journal of Geophysical Research*, 93 (D9), 10815.

Pultz, T. and Brown, R., 1987. SAR image classification of agricultural targets using first-and second-order statistics. *Canadian Journal of Remote Sensing*, 13 (2), 85–91.

Rahman, M.M., Moran, M.S., Thoma, D.P., Bryant, R., Holifield Collins, C.D., Jackson, T., Orr, B.J., and Tischler, M., 2008. Mapping surface roughness and soil moisture using multi-angle radar imagery without ancillary data. *Remote Sensing of Environment*, 112 (2), 391–402.

Ramsar Convention on Wetlands, Food and Agriculture Organization of the United Nations, and International Water Management Institute, 2014. *Wetlands and agriculture: partners for growth*. Gland, Switzerland: Ramsar Convention on Wetlands; Rome, Italy: FAO; Colombo, Sri Lanka: International Water Management Institute (IWMI).

Rebelo, L.-M., McCartney, M.P., and Finlayson, C.M., 2010. Wetlands of Sub-Saharan Africa: distribution and contribution of agriculture to livelihoods. *Wetlands Ecology and Management*, 18 (5), 557–572.

Refice, A., D'Addabbo, A., and Capolongo, D., 2018. *Flood Monitoring through Remote Sensing*.

Richter, R. and Schläpfer, D., 2013. Atmospheric/Topographic C orrection for Satellite Imagery (ATCOR-2/3 user guide, version 8.3. 1, February 2014). *ReSe Applications Schläpfer, Langeggweg*, 3.

Rockström, J. and Karlberg, L., 2010. The Quadruple Squeeze: Defining the safe operating space for freshwater use to achieve a triply green revolution in the Anthropocene. *AMBIO*, 39 (3), 257–265.

Rockström, J., Williams, J., Daily, G., Noble, A., Matthews, N., Gordon, L., Wetterstrand, H., DeClerck, F., Shah, M., Steduto, P., de Fraiture, C., Hatibu, N., Unver, O., Bird, J., Sibanda, L., and Smith, J., 2017. Sustainable intensification of agriculture for human prosperity and global sustainability. *Ambio*, 46 (1), 4–17.

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J.K., Walker, J.P., Lohmann, D., and Toll, D., 2004. The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, 85 (3), 381–394.

Rodriguez, D., Fitzgerald, G.J., Belford, R., and Christensen, L.K., 2006. Detection of nitrogen deficiency in wheat from spectral reflectance indices and basic crop eco-physiological concepts. *Australian Journal of Agricultural Research*, 57 (7), 781.

Rodríguez Martín, J.A., Álvaro-Fuentes, J., Gonzalo, J., Gil, C., Ramos-Miras, J.J., Grau Corbí, J.M., and Boluda, R., 2016. Assessment of the soil organic carbon stock in Spain. *Geoderma*, 264, 117–125.

Rodríguez-Galiano, V.F., Abarca-Hernández, F., Ghimire, B., Chica-Olmo, M., Atkinson, P.M., and Jeganathan, C., 2011. Incorporating Spatial Variability Measures in Land-cover Classification using Random Forest. *Procedia Environmental Sciences*, 3, 44–49.

Rojas, O., Vrieling, A., and Rembold, F., 2011. Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. *Remote Sensing of Environment*, 115 (2), 343–352.

Rosa, M.F., Bonham, C.A., Dempewolf, J., and Arakwiye, B., 2017. An integrated approach to monitoring ecosystem services and agriculture: implications for sustainable agricultural intensification in Rwanda. *Environmental Monitoring and Assessment*, 189 (1).

Rötzer, K., Montzka, C., and Vereecken, H., 2015. Spatio-temporal variability of global soil moisture products. *Journal of Hydrology*, 522, 187–202.

Sachs, J.D., 2012. From Millennium Development Goals to Sustainable Development Goals. *The Lancet*, 379 (9832), 2206–2211.

SAGCOT, n.d. Southern Agricultural Growth Corridor of Tanzania : Concept Note.

Sahebi, M.R. and Angles, J., 2010. An inversion method based on multi-angular approaches for estimating bare soil surface parameters from RADARSAT-1. *Hydrology and Earth System Sciences*, 14 (11), 2355–2366.

Sakané, N., Alvarez, M., Becker, M., Böhme, B., Handa, C., Kamiri, H.W., Langensiepen, M., Menz, G., Misana, S., Mogha, N.G., Möseler, B.M., Mwita, E.J., Oyieke, H.A., and Wijk, M.T., 2011. Classification, Characterisation, and Use of Small Wetlands in East Africa. *Wetlands*, 31 (6), 1103–1116.

Santi, E., Paloscia, S., Pettinato, S., Notarnicola, C., Pasolli, L., and Pistocchi, A., 2013. Comparison between SAR Soil Moisture Estimates and Hydrological Model Simulations over the Scrivia Test Site. *Remote Sensing*, 5 (10), 4961–4976.

Scanlon, B.R., Reedy, R.C., Stonestrom, D.A., Prudic, D.E., and Dennehy, K.F., 2005. Impact of land use and land cover change on groundwater recharge and quality in the southwestern US. *Global Change Biology*, 11 (10), 1577–1593.

Schönert, M., Zillmann, E., Weichelt, H., Eitel, J., Magney, T.S., Lilienthalc, H., Siegmannd, B., and Jarmerd, T., 2015. The Tasseled Cap Transformation for RapidEye data and the Estimation of Vital and Senescent Crop Parameters. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Schuster, C., Förster, M., and Kleinschmit, B., 2012. Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *International Journal of Remote Sensing*, 33 (17), 5583–5599.

Schuster, C., Schmidt, T., Conrad, C., Kleinschmit, B., and Förster, M., 2015. Grassland habitat mapping by intra-annual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 34, 25–34.

Scipal, K., Dorigo, W., and deJeu, R., 2010. Triple collocation - A new tool to determine the error structure of global soil moisture products. IEEE, 4426–4429.

Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., and Wagner, W., 2008. A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophysical Research Letters*, 35 (24).

Seki, H.A., Shirima, D.D., Courtney Mustaphi, C.J., Marchant, R., and Munishi, P.K.T., 2017. The impact of land use and land cover change on biodiversity within and adjacent to Kibasira Swamp in Kilombero Valley, Tanzania. *African Journal of Ecology*.

Sentinel-1 Toolbox, n.d.

Serneels, S., Said, M.Y., and Lambin, E.F., 2001. Land cover changes around a major east African wildlife reserve: The Mara Ecosystem (Kenya). *International Journal of Remote Sensing*, 22 (17), 3397–3420.

Shang, J., McNairn, H., Champagne, C., and Jiao, X., 2009. Application of Multi-Frequency Synthetic Aperture Radar (SAR) in Crop Classification. *In*: G. Jedlovec, ed. *Advances in Geoscience and Remote Sensing*. InTech.

Shao, Y., Fan, X., Liu, H., Xiao, J., Ross, S., Brisco, B., Brown, R., and Staples, G., 2001. Rice monitoring and production estimation using multitemporal RADARSAT. *Remote Sensing of Environment*, 76 (3), 310–325.

Sheffield, J., Wood, E.F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., and Ogallo, L., 2014. A Drought Monitoring and Forecasting System for Sub-Sahara African Water Resources and Food Security. *Bulletin of the American Meteorological Society*, 95 (6), 861–882.

Shetto, R., Owenya, M., African Conservation Tillage Network, CIRAD, and FAO of the United Nations, eds., 2007. *Conservation agriculture as practised in Tanzania: three case studies*. Nairobi : Centre de Coopération international de recherche agronomique pour le développement ; Rome, Italy: African Conservation Tillage Network ; Food Agriculture Organization of the United Nations.

Shinoda, M., 1995. Seasonal phase lag between rainfall and vegetation activity in tropical Africa as revealed by NOAA satellite data. *International Journal of Climatology*, 15 (6), 639–656.

Shrestha, U.B., Gautam, S., and Bawa, K.S., 2012. Widespread Climate Change in the Himalayas and Associated Changes in Local Ecosystems. *PLoS ONE*, 7 (5), e36741.

Siima, S.B., Munishi, P.K.T., Ngaga, Y.M., and Navrud, S., 2012. Estimating direct use value of Kilombero Ramsar Site based on market price method. *Tanzania Journal of Forestry and Nature Conservation*, 81 (2), 133–146.

Singh, M. and Kaur, G., 2011. SAR Image Classification Using PCA and Texture Analysis. *In*: V.V. Das, G. Thomas, and F. Lumban Gaol, eds. *Information Technology and Mobile Communication*. Berlin, Heidelberg: Springer Berlin Heidelberg, 435–439.

Skolnik, M.I., ed., 1990. *Radar handbook*. 2nd ed. New York: McGraw-Hill.

Sonobe, R., Tani, H., Wang, X., Kobayashi, N., and Shimamura, H., 2014. Parameter tuning in the support vector machine and random forest and their performances in cross- and same-year crop classification using TerraSAR-X. *International Journal of Remote Sensing*, 35 (23), 7898–7909.

Spennemann, P.C., Rivera, J.A., Osman, M., Saulo, A.C., and Penalba, O.C., 2017. Assessment of Seasonal Soil Moisture Forecasts over Southern South America with Emphasis on Dry and Wet Events. *Journal of Hydrometeorology*, 18 (8), 2297–2311.

Srivastava, P.K., Petropoulos, G., and Kerr, Y.., 2016. *Satellite Soil Moisture Retrieval: Techniques and Applications.*

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research: Oceans*, 103 (C4), 7755–7766.

Stoorvogel, J.J., Kooistra, L., and Bouma, J., 2015. Managing soil variability at different spatial scales as a basis for precision agriculture. *Soil-Specific Farming: Precision Agriculture; Lal, R., Stewart, BA, Eds*, 37–72.

Suwanprasit, C. and Srichai, N., 2012. Impacts of spatial resolution on land cover classification. *Proceedings of the Asia-Pacific Advanced Network*, 33 (0), 39.

Szantoi, Z., Escobedo, F., Abd-Elrahman, A., Smith, S., and Pearlstine, L., 2013. Analyzing fine-scale wetland composition using high resolution imagery and texture features. *International Journal of Applied Earth Observation and Geoinformation*, 23, 204–212.

Taconet, O. and Ciarletti, V., 2007. Estimating soil roughness indices on a ridge-and-furrow surface using stereo photogrammetry. *Soil and Tillage Research*, 93 (1), 64–76.

Tanzania National Bureau of Statistics, 2018. Tanzania National Bureau of Statistics [online]. *Statistics for Development*. Available from: http://www.nbs.go.tz/nbstz/index.php/english/ [Accessed 12 Jan 2018].

Taylor, R.G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., Konikow, L., Green, T.R., Chen, J., Taniguchi, M., Bierkens, M.F.P., MacDonald, A., Fan, Y., Maxwell, R.M., Yechieli, Y., Gurdak, J.J., Allen, D.M., Shamsudduha, M., Hiscock, K., Yeh, P.J.F., Holman, I., and Treidel, H., 2012. Ground water and climate change. *Nature Climate Change*, 3 (4), 322–329.

Tian, S., Zhang, X., Tian, J., and Sun, Q., 2016. Random Forest Classification of Wetland Landcovers from Multi-Sensor Data in the Arid Region of Xinjiang, China. *Remote Sensing*, 8 (11), 954.

Tiner, R.W., Lang, M.W., and Klemas, V.V., 2015. *Remote sensing of wetlands: applications and advances*. CRC Press.

Torbick, N., Ledoux, L., Salas, W., and Zhao, M., 2016. Regional Mapping of Plantation Extent Using Multisensor Imagery. *Remote Sensing*, 8 (3), 236.

Touzi, R. and Cloude, S., 2013. *Applications of SAR Polarimetry on Land: Soil Moisture and Wetlands POLINSAR 2013*.

Trangmar, B.B., Yost, R.S., and Uehara, G., 1986. Application of Geostatistics to Spatial Studies of Soil Properties. *In*: *Advances in Agronomy*. Elsevier, 45–94.

Treitz, P., Howarth, P., Filho, O., and Soulis, E., 2000. Agricultural crop classification using SAR tone and texture statistics. *Canadian Journal of Remote Sensing*, 26 (1), 18–29.

Trudel, M., Charbonneau, F., and Leconte, R., 2012. Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarization data for estimating soil moisture over agricultural fields. *Canadian Journal of Remote Sensing*, 38 (4), 514–527.

Turkar, V., Deo, R., Rao, Y.S., Mohan, S., and Das, A., 2012. Classification Accuracy of Multi-Frequency and Multi-Polarization SAR Images for Various Land Covers. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5 (3), 936–941.

Turner, R., Panciera, R., Tanase, M.A., Lowell, K., Hacker, J.M., and Walker, J.P., 2014. Estimation of soil surface roughness of agricultural soils using airborne LiDAR. *Remote Sensing of Environment*, 140, 107–117.

Ulaby, F. and Batlivala, P., 1976. Optimum Radar Parameters for Mapping Soil Moisture. *IEEE Transactions on Geoscience Electronics*, 14 (2), 81–93.

Ulaby, F.T., Dubois, P.C., and van Zyl, J., 1996. Radar mapping of surface soil moisture. *Journal of Hydrology*, 184 (1–2), 57–84.

USDA, n.d. Soil Bulk Density /Moisture /Aeration : Soil Quality Kit - Guide for Educators. *Natural Resources Conservation Service*.

Vachaud, G., Passerat De Silans, A., Balabanis, P., and Vauclin, M., 1985. Temporal Stability of Spatially Measured Soil Water Probability Density Function1. *Soil Science Society of America Journal*, 49 (4), 822.

Van doninck, J., Peters, J., Lievens, H., De Baets, B., and Verhoest, N.E.C., 2012. Accounting for seasonality in a soil moisture change detection algorithm for ASAR Wide Swath time series. *Hydrology and Earth System Sciences*, 16 (3), 773–786.

Van Zyl, J. and Kim, Y., 2011. *Synthetic aperture radar polarimetry*. 1st ed. Hoboken, NJ: Wiley.

Vaudour, E., Moeys, J., Gilliot, J.M., and Coquet, Y., 2008. Spatial retrieval of soil reflectance from SPOT multispectral data using the empirical line method. *International Journal of Remote Sensing*, 29 (19), 5571–5584.

van der Velde, R., Su, Z., and Ma, Y., 2008. Impact of Soil Moisture Dynamics on ASAR σo Signatures and Its Spatial Variability Observed over the Tibetan Plateau. *Sensors*, 8 (9), 5479–5491.

Ventrella D., Castellini M., Di Giacomo E., Giglio L., Campi P., Palumbo A.D., and Mastrorilli M., 2008. Evaluation of different water content measurement methods to analyze soil water dynamics. *In*: Lamaddalena N., Santini A., Severino G., and Palladino M., eds. *Irrigation in Mediterranean agriculture: challenges and innovation for the next decades*. Bari : CIHEAM, 151–159.

Verhoest, N.E.., Lievens, H., Wagner, W., Álvarez-Mozos, J., Moran, M.S., and Mattia, F., 2008. On the Soil Roughness Parameterization Problem in Soil Moisture Retrieval of Bare Surfaces from Synthetic Aperture Radar. *Sensors*, 8 (7), 4213–4248.

Verhoeven, J.T.A. and Setter, T.L., 2010. Agricultural use of wetlands: opportunities and limitations. *Annals of Botany*, 105 (1), 155–163.

Vizy, E.K., Cook, K.H., Chimphamba, J., and McCusker, B., 2015. Projected changes in Malawi’s growing season. *Climate Dynamics*, 45 (5–6), 1673–1698.

Vrieling, A., de Leeuw, J., and Said, M., 2013. Length of Growing Period over Africa: Variability and Trends from 30 Years of NDVI Time Series. *Remote Sensing*, 5 (2), 982–1000.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., Zeil, P., and Rubel, F., 2013. The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications. *Meteorologische Zeitschrift*, 22 (1), 5–33.

Wagner, W., Naeimi, V., Scipal, K., de Jeu, R., and Martínez-Fernández, J., 2007. Soil moisture from operational meteorological satellites. *Hydrogeology Journal*, 15 (1), 121–131.

Walker, W., 2011. *Introduction to RADAR Remote Sensing for Vegetation Mapping and Monitoring*.

Wang, C., Zuo, Q., and Zhang, R., 2008. Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method. *Journal of Hydrology*, 352 (3–4), 309–321.

Wang, H., Allain-Bailhache, S., Meric, S., and Pottier, E., 2016. Soil Parameter Retrievals Over Bare Agricultural Fields Using Multiangular RADARSAT-2 Dataset. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9 (12), 5666–5676.

Wang, H., Zhao, Y., Pu, R., and Zhang, Z., 2015. Mapping Robinia Pseudoacacia Forest Health Conditions by Using Combined Spectral, Spatial, and Textural Information Extracted from IKONOS Imagery and Random Forest Classifier. *Remote Sensing*, 7 (7), 9020–9044.

Wang, L. and Qu, J.J., 2009. Satellite remote sensing applications for surface soil moisture monitoring: A review. *Frontiers of Earth Science in China*, 3 (2), 237–247.

Wang, S.G., Li, X., Han, X.J., and Jin, R., 2011. Estimation of surface soil moisture and roughness from multi-angular ASAR imagery in the Watershed Allied Telemetry Experimental Research (WATER). *Hydrology and Earth System Sciences*, 15 (5), 1415–1426.

Wessels, K., Steenkamp, K., von Maltitz, G., and Archibald, S., 2011. Remotely sensed vegetation phenology for describing and predicting the biomes of South Africa: Remotely sensed vegetation phenology of South Africa's biomes. *Applied Vegetation Science*, 14 (1), 49–66.

Wessels, K.J., Bachoo, A., and Archibald, S., 2009. Influence of composite period and date of observation on phenological metrics extracted from MODIS data. *In*: *Sustaining the Millennium Development Goals*. Presented at the 33rd International Symposium on Remote Sensing of Environment, Stresa, Lago Magglore, Italy.

Western, A.W. and Blöschl, G., 1999. On the spatial scaling of soil moisture. *Journal of hydrology*, 217 (3), 203–224.

Whigham, D.F., Dykyjová, D., and Hejný, S., 1993. *Wetlands of the world: Inventory, ecology and management Volume I: Africa, Australia, Canada and Greenland, Mediterranean, Mexico, Papua New Guinea, South Asia, Tropical South America, United States*. Dordrecht: Springer Netherlands.

White, L., Brisco, B., Dabboor, M., Schmitt, A., and Pratt, A., 2015. A Collection of SAR Methodologies for Monitoring Wetlands. *Remote Sensing*, 7 (6), 7615–7645.

Wigneron, J.-P., Jackson, T.J., O'Neill, P., De Lannoy, G., de Rosnay, P., Walker, J.P., Ferrazzoli, P., Mironov, V., Bircher, S., Grant, J.P., Kurum, M., Schwank, M., Munoz-

Sabater, J., Das, N., Royer, A., Al-Yaari, A., Al Bitar, A., Fernandez-Moran, R., Lawrence, H., Mialon, A., Parrens, M., Richaume, P., Delwart, S., and Kerr, Y., 2017. Modelling the passive microwave signature from land surfaces: A review of recent results and application to the L-band SMOS & SMAP soil moisture retrieval algorithms. *Remote Sensing of Environment*, 192, 238–262.

Wilhite, D.A. and Pulwarty, R.S., eds., 2018. *Drought and water crises: integrating science, management, and policy*. Second edition. Boca Raton: CRC Press.

Wilks, D.S., 2006. *Statistical methods in the atmospheric sciences*. 2nd ed. Amsterdam ; Boston: Academic Press.

Wilson, T. and Lewis, I., 2015. *The Rice Value Chain in Tanzania : A report from the Southern Highlands Food Systems Programme*. Food and Agriculture Organization of the United Nations.

Wood, E.M., Pidgeon, A.M., Radeloff, V.C., and Keuler, N.S., 2012. Image texture as a remotely sensed measure of vegetation structure. *Remote Sensing of Environment*, 121, 516–526.

Wulder, M. and Boots, B., 1998. Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. *International Journal of Remote Sensing*, 19 (11), 2223–2231.

Yilmaz, M.T., Crow, W.T., Anderson, M.C., and Hain, C., 2012. An objective methodology for merging satellite- and model-based soil moisture products. *Water Resources Research*, 48 (11).

Yue, J., Yang, G., Qi, X., and Wang, Y., 2016. Soil moisture retrieval in well-covered farmland by Radarsat-2 SAR data. IEEE, 1699–1702.

Zhang, C. and Xie, Z., 2012. Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. *Remote Sensing of Environment*, 124, 310–320.

Zhang, M., Li, M., Wang, W., Liu, C., and Gao, H., 2013. Temporal and spatial variability of soil moisture based on WSN. *Mathematical and Computer Modelling*, 58 (3–4), 826–833.

Zhang, X., Chen, B., Fan, H., Huang, J., and Zhao, H., 2015. The Potential Use of Multi-Band SAR Data for Soil Moisture Retrieval over Bare Agricultural Areas: Hebei, China. *Remote Sensing*, 8 (1), 7.

Zhang, X., Chen, N., Li, J., Chen, Z., and Niyogi, D., 2017. Multi-sensor integrated framework and index for agricultural drought monitoring. *Remote Sensing of Environment*, 188, 141–163.

Zhang, X., Zhang, T., Zhou, P., Shao, Y., and Gao, S., 2017. Validation Analysis of SMAP and AMSR2 Soil Moisture Products over the United States Using Ground-Based Measurements. *Remote Sensing*, 9 (2), 104.

Zhang, Y., Zhang, H., and Lin, H., 2014. Improving the impervious surface estimation with combined use of optical and SAR remote sensing images. *Remote Sensing of Environment*, 141, 155–167.

Zhang, Y.K. and Schilling, K.E., 2006. Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: A Field observation and analysis. *Journal of Hydrology*, 319 (1–4), 328–338.

Zhu, Z., Wang, S., and Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269–277.

Ziadat, F.M., Taimeh, A.Y., and Hattar, B.I., 2010. Variation of Soil Physical Properties and Moisture Content Along Toposequences in the Arid to Semiarid Area. *Arid Land Research and Management*, 24 (2), 81–97.

Zreda, M., Shuttleworth, W.J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R., 2012. COSMOS: the COsmic-ray Soil Moisture Observing System. *Hydrology and Earth System Sciences*, 16 (11), 4079–4099.

Zribi, M. and Dechambre, M., 2003. A new empirical model to retrieve soil moisture and roughness from C-band radar data. *Remote Sensing of Environment*, 84 (1), 42–52.

Zribi, M., Paris Anguela, T., Duchemin, B., Lili, Z., Wagner, W., Hasenauer, S., and Chehbouni, A., 2010. Relationship between soil moisture and vegetation in the Kairouan plain region of Tunisia using low spatial resolution satellite data: Soil Moisture and Vegetation with Satellite data. *Water Resources Research*, 46 (6).

A Appendix: Multi-temporal Analysis of Sentinel-1 Texture and Principal Component Analysis features

In this section of the appendix, additional figures and tables from objective 2 (Chapter 4) are shown. Classification methods used were Random Forest (RF), Neural Net (NN) and Support Vector Machine (SVM). Sentinel 1 single polarized VV, Grey Level Co-occurrence Features (GLCM Texture) and Principal Component Analysis (PCA) features were classified.

A 1: Training and validation points

	Training					Validation				
	bare	Vegetation	built up	water	SUM	bare	Vegetation	built up	water	SUM
01_180115_	145	145	187	164	641	63	62	77	74	276
02_110215_	72	67	59	71	269	36	31	32	27	126
03_070315_	75	75	79	84	313	32	32	30	38	132
04_310315_	118	129	130	170	547	50	54	54	64	222
05_240415_	146	145	184	166	641	62	62	80	72	276
06_180515_	145	148	185	165	643	63	59	79	73	274
07_110615_	107	118	138	157	520	49	54	59	81	243

08_050715_	98	130	143	165	536	42	54	61	73	230
09_290715_	77	130	147	168	522	33	54	62	70	219
12_091015_	29	227	150	193	599	12	96	64	83	255
13_021115_	36	228	152	194	610	13	98	65	82	258
14_261115_	197	244	186	193	820	90	104	79	83	356
15_201215_	196	250	185	193	824	91	98	80	83	352
16_130116_	63	149	59	153	424	26	62	25	65	178
17_060216_	62	179	55	153	449	33	77	23	66	199
18_010316_	135	273	140	211	759	60	116	61	93	330
19_250316_	37	136	50	178	401	20	56	21	77	174
20_180416_	63	175	55	75	368	23	73	24	32	152
21_120516_	57	163	53	110	383	21	69	23	47	160
22_050616_	40	147	55	160	402	32	64	24	68	188
23_290616_	63	158	52	168	441	29	67	22	72	190
24_230716_	78	135	47	200	460	28	58	20	86	192
26_090916_	78	117	39	176	410	35	51	17	76	179
27_031016_	78	125	44	154	401	32	51	19	66	168
28_271016_	83	144	41	150	418	34	60	17	64	175
30_141216_	77	167	38	13	295	32	72	16	6	126

A 2: Overall Classification Accuracy. RF, NNET, SVM are random forest, neural network and support vector machine classifiers. TEX, VV, PCA are texture, single polarized and PCA derived overall classification accuracies

	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	_TEX	_TEX	_TEX	_VV	_VV	_VV	_PCA	_PCA	_PCA
18-01-15	0.76	0.73	0.71	0.62	0.66	0.68	0.88	0.88	0.88
11-02-15	0.86	0.88	0.85	0.65	0.75	0.74	0.92	0.92	0.92
07-03-15	0.79	0.81	0.81	0.60	0.73	0.72	0.89	0.89	0.89
31-03-15	0.71	0.65	0.65	0.57	0.60	0.61	0.86	0.86	0.86
24-04-15	0.77	0.79	0.77	0.62	0.68	0.67	0.86	0.86	0.86
18-05-15	0.79	0.76	0.75	0.57	0.65	0.65	0.87	0.87	0.87
11-06-15	0.71	0.71	0.72	0.57	0.61	0.61	0.86	0.86	0.86
05-07-15	0.78	0.77	0.77	0.60	0.74	0.75	0.86	0.86	0.86
29-07-15	0.79	0.80	0.80	0.65	0.75	0.75	0.88	0.88	0.88

09-10-15	0.84	0.85	0.86	0.67	0.76	0.75	0.82	0.82	0.82
02-11-15	0.81	0.82	0.81	0.70	0.74	0.74	0.82	0.82	0.82
26-11-15	0.76	0.80	0.78	0.55	0.65	0.65	0.78	0.78	0.78
20-12-15	0.73	0.76	0.71	0.54	0.60	0.61	0.77	0.77	0.77
13-01-16	0.66	0.66	0.69	0.58	0.61	0.66	0.77	0.77	0.77
06-02-16	0.72	0.74	0.75	0.61	0.72	0.74	0.85	0.85	0.85
01-03-16	0.75	0.75	0.72	0.62	0.67	0.67	0.84	0.84	0.84
25-03-16	0.80	0.65	0.77	0.62	0.75	0.72	0.84	0.84	0.84
18-04-16	0.67	0.68	0.72	0.62	0.64	0.63	0.80	0.80	0.80
12-05-16	0.70	0.72	0.67	0.59	0.67	0.66	0.79	0.79	0.79
05-06-16	0.76	0.79	0.79	0.63	0.70	0.71	0.83	0.83	0.83
29-06-16	0.81	0.82	0.82	0.65	0.81	0.80	0.87	0.87	0.87
23-07-16	0.86	0.86	0.86	0.65	0.79	0.77	0.91	0.91	0.91
09-09-16	0.81	0.86	0.84	0.76	0.79	0.79	0.91	0.91	0.91
03-10-16	0.85	0.84	0.83	0.74	0.79	0.78	0.90	0.90	0.90
27-10-16	0.84	0.82	0.83	0.68	0.74	0.74	0.83	0.83	0.83
14-12-16	0.64	0.71	0.71	0.55	0.61	0.61	0.68	0.68	0.68

A 3: Kappa Coefficients. RF, NNET, SVM are random forest, neural network and support vector machine classifiers. TEX, VV, PCA are texture, single polarized and PCA derived kappa coefficients

	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	_TEX	_TEX	_TEX	_VV	_VV	_VV	_PCA	_PCA	_PCA
18-01-15	0.66	0.61	0.59	0.48	0.51	0.52	0.83	0.83	0.83
11-02-15	0.81	0.83	0.80	0.52	0.66	0.64	0.89	0.89	0.89
07-03-15	0.72	0.74	0.74	0.46	0.62	0.61	0.85	0.85	0.85
31-03-15	0.61	0.51	0.53	0.41	0.45	0.46	0.82	0.82	0.82
24-04-15	0.69	0.72	0.69	0.49	0.57	0.55	0.82	0.82	0.82
18-05-15	0.72	0.67	0.66	0.41	0.52	0.52	0.83	0.83	0.83
11-06-15	0.61	0.61	0.62	0.42	0.47	0.47	0.81	0.81	0.81
05-07-15	0.70	0.69	0.69	0.46	0.64	0.65	0.81	0.81	0.81
29-07-15	0.72	0.73	0.73	0.53	0.66	0.66	0.84	0.84	0.84
09-10-15	0.78	0.79	0.82	0.55	0.67	0.66	0.76	0.76	0.76
02-11-15	0.74	0.76	0.74	0.60	0.64	0.65	0.76	0.76	0.76

26-11-15	0.67	0.72	0.70	0.39	0.52	0.52	0.71	0.71	0.71
20-12-15	0.64	0.68	0.62	0.38	0.46	0.47	0.70	0.70	0.70
13-01-16	0.54	0.55	0.57	0.44	0.48	0.53	0.69	0.69	0.69
06-02-16	0.62	0.64	0.66	0.47	0.62	0.64	0.80	0.80	0.80
01-03-16	0.66	0.66	0.63	0.50	0.56	0.55	0.78	0.78	0.78
25-03-16	0.73	0.49	0.67	0.48	0.65	0.60	0.78	0.78	0.78
18-04-16	0.55	0.57	0.62	0.49	0.51	0.50	0.72	0.72	0.72
12-05-16	0.60	0.62	0.56	0.45	0.55	0.54	0.72	0.72	0.72
05-06-16	0.67	0.71	0.71	0.50	0.59	0.60	0.77	0.77	0.77
29-06-16	0.74	0.75	0.75	0.53	0.73	0.72	0.82	0.82	0.82
23-07-16	0.81	0.81	0.80	0.53	0.71	0.69	0.87	0.87	0.87
09-09-16	0.74	0.81	0.77	0.67	0.71	0.70	0.87	0.87	0.87
03-10-16	0.79	0.78	0.77	0.64	0.72	0.69	0.87	0.87	0.87
27-10-16	0.78	0.76	0.77	0.56	0.65	0.65	0.77	0.77	0.77
14-12-16	0.44	0.53	0.53	0.28	0.41	0.39	0.50	0.50	0.50

A 4: Bare class sensitivity

Bare	Sensitivity								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.24	0.04	0.00	0.36	0.04	0.00	0.80	0.80	0.80
11-02-15	0.91	0.94	0.91	0.60	0.86	0.89	0.94	0.94	0.94
07-03-15	0.82	0.82	0.82	0.45	0.76	0.76	0.85	0.85	0.85
31-03-15	0.52	0.15	0.38	0.29	0.27	0.25	0.77	0.77	0.77
24-04-15	0.72	0.75	0.69	0.51	0.59	0.58	0.83	0.83	0.83
18-05-15	0.76	0.61	0.69	0.44	0.76	0.76	0.78	0.78	0.78
11-06-15	0.50	0.62	0.48	0.40	0.37	0.27	0.77	0.77	0.77
05-07-15	0.48	0.48	0.43	0.29	0.36	0.40	0.71	0.71	0.71
29-07-15	0.53	0.55	0.42	0.37	0.47	0.42	0.68	0.68	0.68
09-10-15	0.66	0.90	0.86	0.55	0.59	0.38	0.59	0.59	0.59
02-11-15	0.43	0.60	0.47	0.37	0.53	0.53	0.63	0.63	0.63
26-11-15	0.59	0.59	0.62	0.52	0.59	0.52	0.79	0.79	0.79
20-12-15	0.60	0.63	0.56	0.36	0.31	0.25	0.58	0.58	0.58

13-01-16	0.45	0.85	0.52	0.36	0.03	0.00	0.42	0.42	0.42
06-02-16	0.47	0.69	0.75	0.31	0.56	0.75	0.61	0.61	0.61
01-03-16	0.70	0.86	0.88	0.42	0.94	0.91	0.56	0.56	0.56
25-03-16	0.42	0.00	0.21	0.25	0.00	0.00	0.25	0.25	0.25
18-04-16	0.61	0.55	0.55	0.55	0.27	0.42	0.61	0.61	0.61
12-05-16	0.48	0.28	0.38	0.35	0.43	0.00	0.48	0.48	0.48
05-06-16	0.53	0.53	0.47	0.34	0.50	0.47	0.41	0.41	0.41
29-06-16	0.61	0.61	0.61	0.39	0.64	0.67	0.64	0.64	0.64
23-07-16	0.79	0.86	0.88	0.48	0.69	0.81	0.79	0.79	0.79
09-09-16	0.80	0.91	0.82	0.73	0.89	0.89	0.84	0.84	0.84
03-10-16	0.93	0.93	0.89	0.73	0.89	0.91	0.84	0.84	0.84
27-10-16	0.77	0.77	0.84	0.66	0.73	0.84	0.64	0.64	0.64
14-12-16	0.60	0.74	0.76	0.67	0.57	0.57	0.45	0.45	0.45

A 5: Bare class positive predicted values

Bare	Positive Predicted Value								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.43	0.20	0.00	0.39	0.20	0.56	0.65	0.65	0.65
11-02-15	0.94	0.94	0.89	0.70	0.77	0.76	0.92	0.92	0.92
07-03-15	0.79	0.79	0.82	0.63	0.64	0.64	0.88	0.88	0.88
31-03-15	0.52	0.39	0.39	0.33	0.30	0.30	0.76	0.76	0.76
24-04-15	0.70	0.74	0.68	0.48	0.58	0.58	0.81	0.81	0.81
18-05-15	0.68	0.66	0.66	0.39	0.47	0.47	0.86	0.86	0.86
11-06-15	0.51	0.48	0.58	0.37	0.40	0.41	0.75	0.75	0.75
05-07-15	0.61	0.65	0.75	0.33	0.58	0.63	0.79	0.79	0.79
29-07-15	0.65	0.66	0.64	0.41	0.62	0.59	0.90	0.90	0.90
09-10-15	0.66	0.59	0.64	0.53	0.68	0.73	0.61	0.61	0.61
02-11-15	0.52	0.58	0.52	0.41	0.46	0.50	0.61	0.61	0.61
26-11-15	0.65	0.74	0.67	0.48	0.65	0.65	0.56	0.56	0.56
20-12-15	0.59	0.64	0.56	0.33	0.37	0.38	0.66	0.66	0.66
13-01-16	0.47	0.50	0.55	0.44	0.14	0.56	0.82	0.82	0.82

06-02-16	0.46	0.51	0.53	0.32	0.48	0.51	0.85	0.85	0.85
01-03-16	0.61	0.55	0.53	0.44	0.46	0.46	0.86	0.86	0.86
25-03-16	0.59	0.64	0.56	0.23	0.54	0.56	0.86	0.86	0.86
18-04-16	0.51	0.50	0.60	0.51	0.39	0.40	0.87	0.87	0.87
12-05-16	0.54	0.69	0.48	0.38	0.47	0.00	0.95	0.95	0.95
05-06-16	0.59	0.68	0.71	0.35	0.52	0.60	1.00	1.00	1.00
29-06-16	0.71	0.69	0.73	0.47	0.77	0.77	0.92	0.92	0.92
23-07-16	0.79	0.77	0.74	0.54	0.76	0.65	0.94	0.94	0.94
09-09-16	0.81	0.85	0.88	0.78	0.78	0.75	0.93	0.93	0.93
03-10-16	0.76	0.76	0.77	0.72	0.73	0.69	0.95	0.95	0.95
27-10-16	0.83	0.77	0.79	0.69	0.73	0.70	0.97	0.97	0.97
14-12-16	0.68	0.66	0.67	0.61	0.65	0.71	0.83	0.83	0.83

A 6: Built up class sensitivity

Built up	Sensitivity								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.78	0.74	0.74	0.48	0.26	0.26	1.00	1.00	1.00
11-02-15	0.76	0.76	0.72	0.48	0.28	0.20	1.00	1.00	1.00
07-03-15	0.71	0.71	0.71	0.33	0.29	0.19	1.00	1.00	1.00
31-03-15	0.85	0.64	0.79	0.48	0.45	0.45	1.00	1.00	1.00
24-04-15	0.80	0.80	0.82	0.57	0.62	0.51	0.92	0.92	0.92
18-05-15	0.76	0.76	0.64	0.36	0.33	0.33	0.94	0.94	0.94
11-06-15	0.90	0.90	0.87	0.38	0.36	0.36	0.97	0.97	0.97
05-07-15	0.92	0.79	0.74	0.59	0.59	0.54	0.95	0.95	0.95
29-07-15	0.88	0.90	0.93	0.67	0.64	0.62	0.98	0.98	0.98
09-10-15	0.94	0.96	0.94	0.62	0.68	0.66	0.98	0.98	0.98
02-11-15	0.94	0.92	0.96	0.79	0.79	0.73	0.98	0.98	0.98
26-11-15	0.82	0.91	0.91	0.50	0.53	0.62	0.91	0.91	0.91
20-12-15	0.83	0.86	0.84	0.52	0.65	0.59	0.97	0.97	0.97
13-01-16	0.84	0.80	0.72	0.56	0.88	0.80	0.80	0.80	0.80
06-02-16	0.88	0.83	0.83	0.54	0.83	0.79	0.83	0.83	0.83
01-03-16	0.85	0.91	0.93	0.66	0.76	0.76	0.88	0.88	0.88

25-03-16	0.91	0.00	0.86	0.59	0.73	0.59	0.91	0.91	0.91
18-04-16	0.79	0.88	0.96	0.67	0.75	0.71	0.79	0.79	0.79
12-05-16	0.92	0.88	0.88	0.77	0.88	0.81	0.81	0.81	0.81
05-06-16	0.78	0.78	0.83	0.52	0.61	0.52	0.87	0.87	0.87
29-06-16	0.86	0.86	0.95	0.68	0.77	0.64	0.82	0.82	0.82
23-07-16	0.92	0.88	0.96	0.54	0.67	0.63	0.83	0.83	0.83
09-09-16	0.82	0.71	0.76	0.59	0.59	0.59	0.82	0.82	0.82
03-10-16	0.89	0.83	0.89	0.61	0.83	0.78	0.89	0.89	0.89
27-10-16	0.82	0.76	0.82	0.53	0.65	0.65	0.65	0.65	0.65
14-12-16	0.81	0.81	0.75	0.31	0.81	0.50	0.75	0.75	0.75

A 7: Built up class positive predicted values

Built up	Positive Predicted Value								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	1.00	0.94	0.85	0.41	1.00	1.00	0.96	0.96	0.96
11-02-15	0.83	0.86	0.90	0.55	0.88	0.83	0.89	0.89	0.89
07-03-15	0.65	0.75	0.75	0.30	0.75	0.80	0.84	0.84	0.84
31-03-15	0.85	0.88	0.93	0.52	1.00	1.00	0.97	0.97	0.97
24-04-15	0.91	0.94	0.94	0.76	0.88	0.97	0.95	0.95	0.95
18-05-15	0.93	0.96	0.91	0.34	0.85	0.85	0.84	0.84	0.84
11-06-15	0.71	0.73	0.69	0.52	0.70	0.70	0.90	0.90	0.90
05-07-15	0.86	0.86	0.85	0.53	0.88	1.00	0.86	0.86	0.86
29-07-15	0.93	0.90	0.95	0.64	0.90	1.00	0.93	0.93	0.93
09-10-15	0.87	0.89	0.90	0.66	0.81	0.85	0.88	0.88	0.88
02-11-15	0.92	0.90	0.92	0.78	0.88	0.90	0.90	0.90	0.90
26-11-15	0.88	0.86	0.89	0.57	0.67	0.66	0.94	0.94	0.94
20-12-15	0.93	0.95	0.91	0.60	0.61	0.71	0.89	0.89	0.89
13-01-16	0.75	0.77	0.82	0.52	0.58	0.77	0.91	0.91	0.91
06-02-16	0.91	0.95	0.95	0.62	0.83	0.86	1.00	1.00	1.00
01-03-16	0.91	0.95	0.94	0.73	0.84	0.84	1.00	1.00	1.00
25-03-16	0.91	0.89	0.86	0.54	0.80	0.76	1.00	1.00	1.00
18-04-16	0.79	0.78	0.79	0.62	0.72	0.71	0.90	0.90	0.90

12-05-16	0.89	0.92	0.88	0.61	0.72	0.84	0.91	0.91	0.91
05-06-16	0.90	0.95	0.95	0.55	0.74	0.75	0.95	0.95	0.95
29-06-16	0.95	1.00	0.88	0.68	0.77	0.78	1.00	1.00	1.00
23-07-16	0.92	0.91	0.88	0.68	0.80	0.83	1.00	1.00	1.00
09-09-16	0.74	0.80	0.76	0.67	0.67	0.71	1.00	1.00	1.00
03-10-16	0.84	0.83	0.89	0.73	0.79	0.78	1.00	1.00	1.00
27-10-16	0.88	1.00	0.93	0.56	0.73	0.73	1.00	1.00	1.00
14-12-16	0.87	0.93	0.92	0.63	0.72	0.73	1.00	1.00	1.00

A 8: Vegetated class sensitivity

Vegetation	Sensitivity								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.83	0.89	0.93	0.50	0.91	0.98	0.74	0.74	0.74
11-02-15	0.82	0.85	0.88	0.65	0.88	0.88	0.85	0.85	0.85
07-03-15	0.73	0.78	0.78	0.65	0.81	0.84	0.86	0.86	0.86
31-03-15	0.60	0.77	0.55	0.53	0.67	0.68	0.77	0.77	0.77
24-04-15	0.68	0.71	0.67	0.57	0.64	0.72	0.80	0.80	0.80
18-05-15	0.65	0.69	0.71	0.37	0.41	0.41	0.82	0.82	0.82
11-06-15	0.55	0.43	0.67	0.52	0.71	0.81	0.79	0.79	0.79
05-07-15	0.70	0.77	0.85	0.51	0.90	0.93	0.80	0.80	0.80
29-07-15	0.80	0.78	0.85	0.62	0.87	0.92	0.90	0.90	0.90
09-10-15	0.73	0.62	0.70	0.55	0.73	0.83	0.68	0.68	0.68
02-11-15	0.73	0.69	0.69	0.60	0.62	0.65	0.62	0.62	0.62
26-11-15	0.73	0.75	0.73	0.39	0.61	0.64	0.61	0.61	0.61
20-12-15	0.57	0.64	0.55	0.37	0.54	0.68	0.63	0.63	0.63
13-01-16	0.49	0.20	0.59	0.49	0.66	0.88	0.88	0.88	0.88
06-02-16	0.53	0.44	0.44	0.49	0.49	0.42	0.91	0.91	0.91
01-03-16	0.49	0.24	0.18	0.47	0.03	0.01	0.95	0.95	0.95
25-03-16	0.76	1.00	0.82	0.42	0.92	0.92	0.97	0.97	0.97
18-04-16	0.50	0.52	0.62	0.45	0.67	0.55	0.88	0.88	0.88
12-05-16	0.59	0.83	0.61	0.39	0.51	0.95	0.93	0.93	0.93
05-06-16	0.65	0.78	0.83	0.53	0.58	0.70	0.98	0.98	0.98

29-06-16	0.75	0.78	0.73	0.48	0.75	0.78	0.95	0.95	0.95
23-07-16	0.72	0.67	0.62	0.49	0.67	0.59	0.97	0.97	0.97
09-09-16	0.58	0.70	0.67	0.58	0.48	0.45	0.91	0.91	0.91
03-10-16	0.54	0.54	0.54	0.51	0.43	0.37	0.89	0.89	0.89
27-10-16	0.76	0.73	0.65	0.43	0.54	0.41	0.97	0.97	0.97
14-12-16	0.63	0.70	0.70	0.53	0.53	0.65	0.88	0.88	0.88

A 9: Vegetated class positive predicted values

Vegetation	Positive Predicted Value								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.61	0.57	0.57	0.52	0.51	0.51	0.87	0.87	0.87
11-02-15	0.74	0.76	0.73	0.51	0.63	0.61	0.94	0.94	0.94
07-03-15	0.73	0.74	0.71	0.51	0.64	0.62	0.89	0.89	0.89
31-03-15	0.59	0.47	0.49	0.50	0.48	0.48	0.85	0.85	0.85
24-04-15	0.60	0.64	0.59	0.47	0.50	0.49	0.78	0.78	0.78
18-05-15	0.66	0.56	0.58	0.49	0.54	0.53	0.81	0.81	0.81
11-06-15	0.64	0.76	0.63	0.49	0.49	0.48	0.84	0.84	0.84
05-07-15	0.64	0.62	0.60	0.53	0.61	0.59	0.84	0.84	0.84
29-07-15	0.65	0.67	0.65	0.54	0.60	0.60	0.78	0.78	0.78
09-10-15	0.80	0.93	0.91	0.56	0.63	0.60	0.87	0.87	0.87
02-11-15	0.68	0.73	0.68	0.56	0.59	0.57	0.79	0.79	0.79
26-11-15	0.65	0.70	0.71	0.38	0.51	0.53	0.79	0.79	0.79
20-12-15	0.53	0.59	0.53	0.36	0.51	0.46	0.63	0.63	0.63
13-01-16	0.51	0.47	0.52	0.44	0.50	0.49	0.60	0.60	0.60
06-02-16	0.53	0.58	0.61	0.44	0.58	0.67	0.68	0.68	0.68
01-03-16	0.53	0.53	0.46	0.41	0.50	0.25	0.64	0.64	0.64
25-03-16	0.63	0.47	0.57	0.46	0.55	0.51	0.65	0.65	0.65
18-04-16	0.57	0.59	0.62	0.49	0.55	0.56	0.65	0.65	0.65
12-05-16	0.52	0.52	0.49	0.43	0.55	0.48	0.59	0.59	0.59
05-06-16	0.57	0.61	0.60	0.50	0.51	0.52	0.64	0.64	0.64
29-06-16	0.65	0.67	0.67	0.42	0.67	0.63	0.69	0.69	0.69
23-07-16	0.72	0.74	0.77	0.38	0.58	0.61	0.75	0.75	0.75

09-09-16	0.61	0.72	0.65	0.51	0.57	0.56	0.75	0.75	0.75
03-10-16	0.79	0.79	0.73	0.49	0.65	0.62	0.78	0.78	0.78
27-10-16	0.68	0.66	0.71	0.42	0.53	0.54	0.62	0.62	0.62
14-12-16	0.54	0.68	0.68	0.47	0.51	0.51	0.56	0.56	0.56

A 10: Water class sensitivity

Water	Sensitivity								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.98	0.95	0.88	0.98	0.98	0.98	1.00	1.00	1.00
11-02-15	0.95	0.95	0.85	0.95	0.95	0.90	0.90	0.90	0.90
07-03-15	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
31-03-15	0.92	0.98	0.94	0.92	0.94	0.96	0.98	0.98	0.98
24-04-15	0.90	0.94	0.94	0.87	0.89	0.87	0.92	0.92	0.92
18-05-15	1.00	0.98	0.94	1.00	0.96	0.96	0.98	0.98	0.98
11-06-15	0.94	0.96	0.89	0.92	0.92	0.91	0.94	0.94	0.94
05-07-15	0.98	0.98	0.96	0.96	0.96	0.96	0.98	0.98	0.98
29-07-15	0.91	0.93	0.91	0.87	0.91	0.91	0.93	0.93	0.93
09-10-15	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00
02-11-15	1.00	1.00	1.00	0.94	0.94	0.98	1.00	1.00	1.00
26-11-15	0.95	1.00	0.91	1.00	1.00	0.91	0.91	0.91	0.91
20-12-15	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00
13-01-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
06-02-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
01-03-16	0.97	1.00	0.93	0.99	0.97	1.00	0.99	0.99	0.99
25-03-16	0.98	1.00	0.96	0.98	1.00	0.98	1.00	1.00	1.00
18-04-16	0.96	0.96	0.91	0.96	0.96	0.96	0.91	0.91	0.91
12-05-16	0.97	1.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00
05-06-16	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00
29-06-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
23-07-16	1.00	1.00	0.96	0.96	1.00	0.94	0.98	0.98	0.98
09-09-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
03-10-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

27-10-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
14-12-16	0.50	0.00	0.00	0.50	1.00	1.00	0.75	0.75	0.75

A 11: Water class positive predicted values

Water	Positive Predicted Value								
	RF	NNET	SVM	RF	NNET	SVM	RF	NNET	SVM
	TEX	TEX	TEX	VV	VV	VV	PCA	PCA	PCA
18-01-15	0.98	0.98	0.97	0.98	0.98	0.98	1.00	1.00	1.00
11-02-15	1.00	1.00	1.00	1.00	1.00	1.00	0.95	0.95	0.95
07-03-15	1.00	0.96	1.00	1.00	1.00	1.00	0.96	0.96	0.96
31-03-15	0.94	0.94	0.94	0.85	0.94	0.94	0.91	0.91	0.91
24-04-15	0.98	0.95	0.98	0.95	0.98	1.00	0.95	0.95	0.95
18-05-15	0.98	0.98	1.00	0.95	0.98	1.00	0.98	0.98	0.98
11-06-15	0.96	0.94	0.98	0.89	0.96	0.96	0.96	0.96	0.96
05-07-15	0.98	1.00	1.00	0.88	0.96	1.00	0.93	0.93	0.93
29-07-15	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.96	0.96
09-10-15	0.98	0.98	0.96	0.91	0.96	0.96	0.84	0.84	0.84
02-11-15	1.00	0.98	1.00	0.98	1.00	1.00	0.88	0.88	0.88
26-11-15	0.95	0.96	0.91	0.96	0.96	0.95	0.95	0.95	0.95
20-12-15	0.97	0.97	0.97	0.98	0.97	0.98	0.94	0.94	0.94
13-01-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
06-02-16	1.00	1.00	1.00	1.00	0.98	0.98	1.00	1.00	1.00
01-03-16	1.00	0.99	1.00	1.00	1.00	0.99	1.00	1.00	1.00
25-03-16	1.00	0.94	0.98	1.00	1.00	1.00	1.00	1.00	1.00
18-04-16	1.00	1.00	1.00	1.00	0.96	1.00	1.00	1.00	1.00
12-05-16	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	1.00
05-06-16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
29-06-16	0.98	0.98	0.98	0.98	0.98	1.00	1.00	1.00	1.00
23-07-16	1.00	1.00	1.00	0.98	0.96	1.00	1.00	1.00	1.00
09-09-16	0.98	1.00	0.96	0.98	0.98	0.98	1.00	1.00	1.00
03-10-16	1.00	0.97	0.95	1.00	0.97	1.00	0.95	0.95	0.95
27-10-16	1.00	1.00	0.95	0.95	0.98	0.95	1.00	1.00	1.00
14-12-16	0.50	0.98	0.98	0.67	0.67	0.67	0.60	0.60	0.60

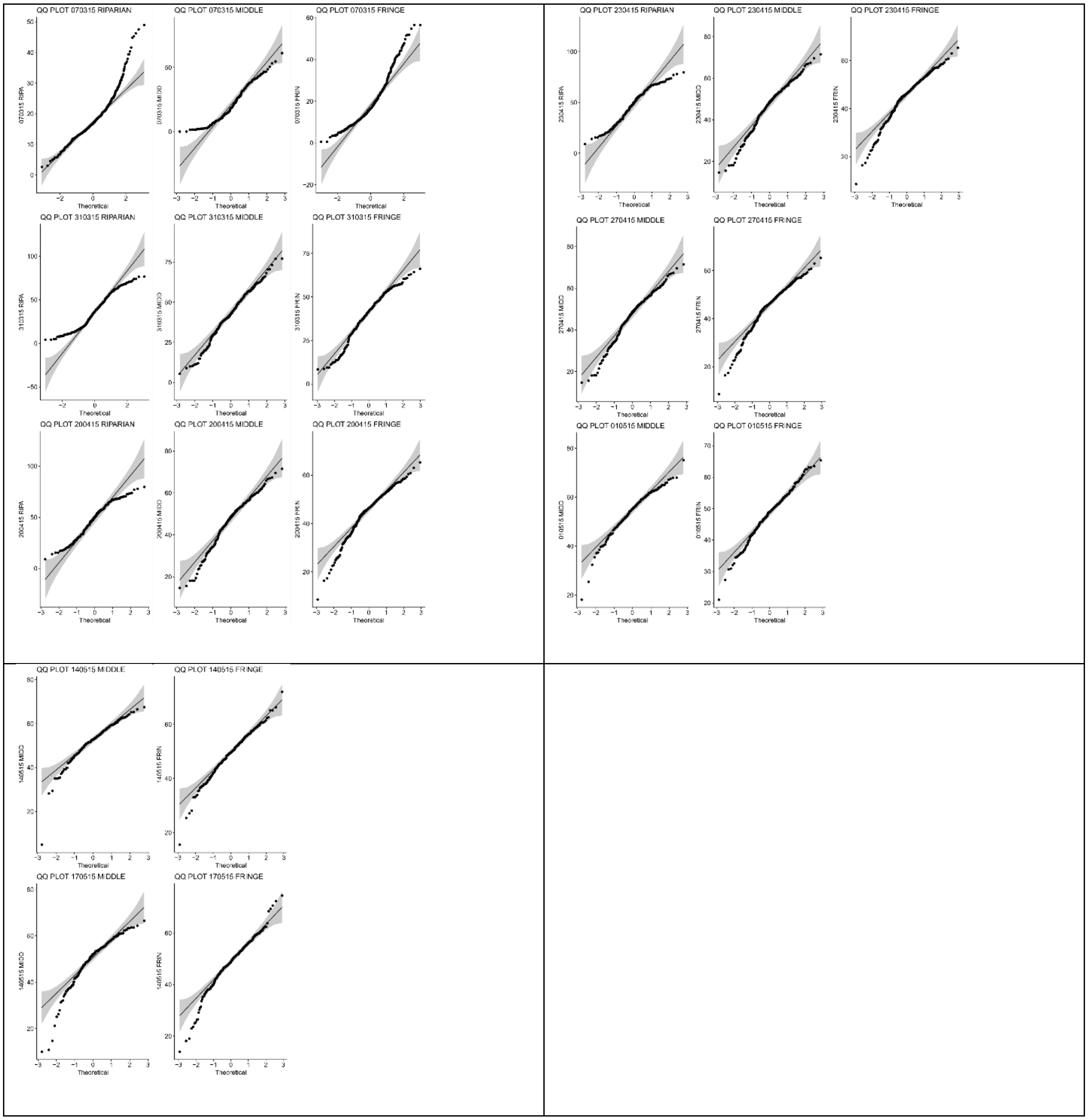
A 12: Confusion matrix 3rd October 2016. In italics are overall accuracies.

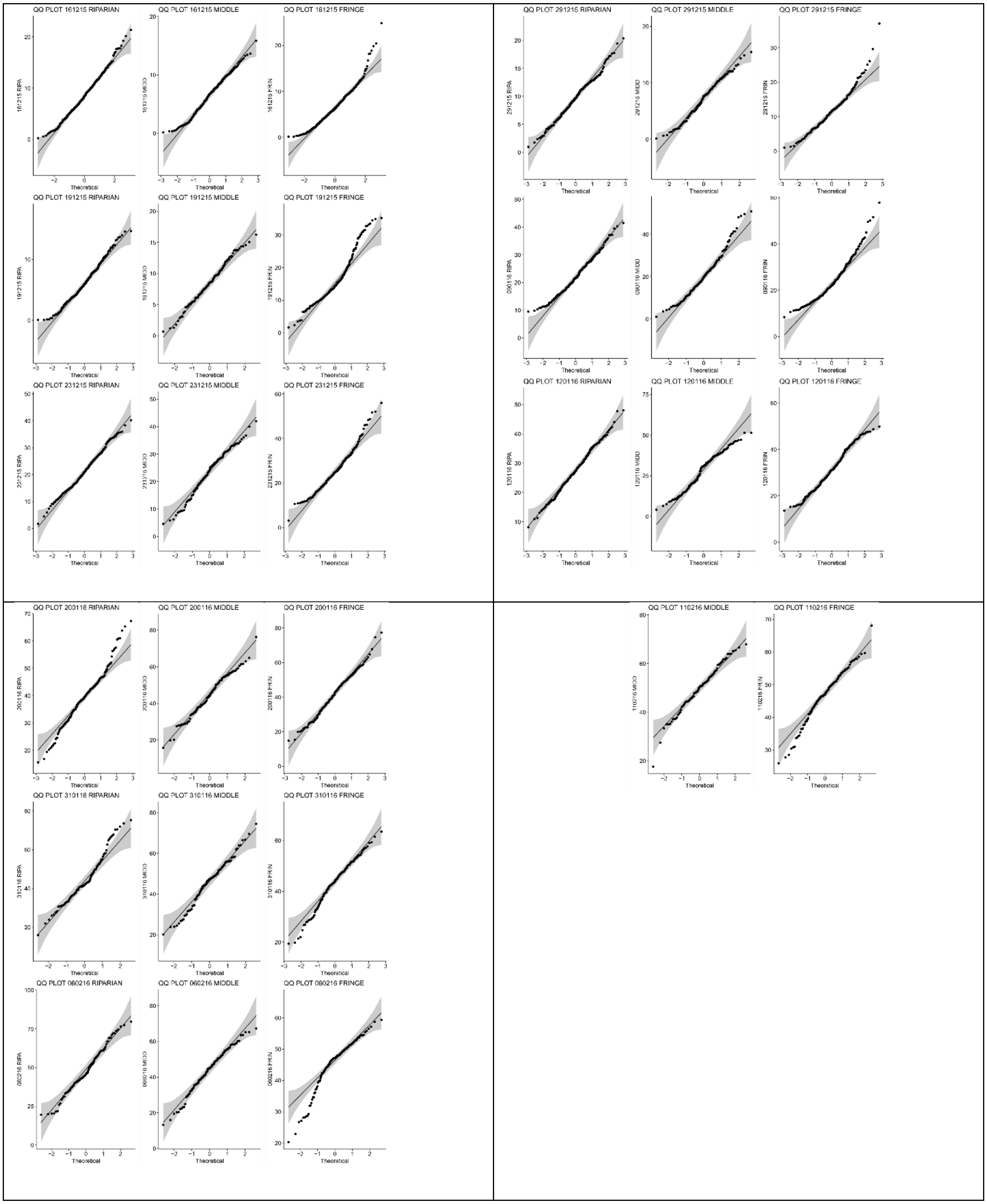
VV NET		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	40	0	5	0	0.89	45
	built up	0	15	3	0	0.83	18
	vegetation	15	4	15	1	0.43	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.73	0.79	0.65	0.97	*0.79*	
	No. of points	55	19	23	39		136
VV RF		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	33	0	12	0	0.73	45
	built up	0	11	7	0	0.61	18
	vegetation	13	4	18	0	0.51	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.72	0.73	0.49	1.00	*0.74*	
	No. of points	46	15	37	38		136
VV SVM		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	41	0	4	0	0.91	45
	built up	0	14	4	0	0.78	18
	vegetation	18	4	13	0	0.37	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.69	0.78	0.62	1.00	*0.78*	
	No. of points	59	18	21	38		136
GLCM NET		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	42	0	3	0	0.93	45
	built up	0	15	2	1	0.83	18
	vegetation	13	3	19	0	0.54	35
	water	0	0	0	38	1.00	38

	Pos predicted value	0.76	0.83	0.79	0.97	*0.84*	
	No. of points	55	18	24	39		136
GLCM RF		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	42	0	3	0	0.93	45
	built up	0	16	2	0	0.89	18
	vegetation	13	3	19	0	0.54	35
	water	0	0	0	38	1.00	38
		0.76	0.84	0.79	1.00	*0.85*	
		55	19	24	38		136
GLCM SVM		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	40	0	5	0	0.89	45
	built up	0	16	2	0	0.89	18
	vegetation	12	2	19	2	0.54	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.77	0.89	0.73	0.95	*0.83*	
	No. of points	52	18	26	40		136
PCA NET		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	38	0	7	0	0.84	45
	built up	0	16	2	0	0.89	18
	vegetation	2	0	31	2	0.89	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.95	1.00	0.78	0.95	*0.90*	
	No. of points	40	16	40	40		136
PCA RF		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	38	0	7	0	0.84	45
	built up	0	16	2	0	0.89	18
	vegetation	2	0	31	2	0.89	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.95	1.00	0.78	0.95	*0.90*	
	No. of points	40	16	40	40		136

PCA SVM		bare	built up	vegetation	water	Sensitivity	No. of points
	bare	38	0	7	0	0.84	45
	built up	0	16	2	0	0.89	18
	vegetation	2	0	31	2	0.89	35
	water	0	0	0	38	1.00	38
	Pos predicted value	0.95	1.00	0.78	0.95	*0.90*	
	No. of points	40	16	40	40		136

B Appendix: Spatial-temporal variation and estimation of soil moisture from Synthetic Aperture Radar imagery


QQ PLOT 070315 RIPARIAN
070315 RIPA
QQ PLOT 070315 MIDDLE
070315 MIDD
QQ PLOT 070315 FRINGE
070315 FRIN
QQ PLOT 310315 RIPARIAN
310315 RIPA
QQ PLOT 310315 MIDDLE
310315 MIDD
QQ PLOT 310315 FRINGE
310315 FRIN
QQ PLOT 200415 RIPARIAN
200415 RIPA
QQ PLOT 200415 MIDDLE
200415 MIDD
QQ PLOT 200415 FRINGE
230415 FRIN
QQ PLOT 230415 RIPARIAN
230415 RIPA
QQ PLOT 230415 MIDDLE
230415 MIDD
QQ PLOT 230415 FRINGE
230415 FRIN
QQ PLOT 270415 MIDDLE
270415 MIDD
QQ PLOT 270415 FRINGE
270415 FRIN
QQ PLOT 010515 MIDDLE
010515 MIDD
QQ PLOT 010515 FRINGE
010515 FRIN
QQ PLOT 140515 MIDDLE
140515 MIDD
QQ PLOT 140515 FRINGE
140515 FRIN
QQ PLOT 170515 MIDDLE
140515 MIDD
QQ PLOT 170515 FRINGE
140515 FRIN
Theoretical



QQ PLOT 161215 RIPARIAN

QQ PLOT 161215 MIDDLE

QQ PLOT 161215 FRINGE

QQ PLOT 191215 RIPARIAN

QQ PLOT 191215 MIDDLE

QQ PLOT 191215 FRINGE

QQ PLOT 231215 RIPARIAN

QQ PLOT 231215 MIDDLE

QQ PLOT 231215 FRINGE

QQ PLOT 291215 RIPARIAN

QQ PLOT 291215 MIDDLE

QQ PLOT 291215 FRINGE

QQ PLOT 120116 RIPARIAN

QQ PLOT 120116 MIDDLE

QQ PLOT 120116 FRINGE

QQ PLOT 200116 RIPARIAN

QQ PLOT 200116 MIDDLE

QQ PLOT 200116 FRINGE

QQ PLOT 310116 RIPARIAN

QQ PLOT 310116 MIDDLE

QQ PLOT 310116 FRINGE

QQ PLOT 060216 RIPARIAN

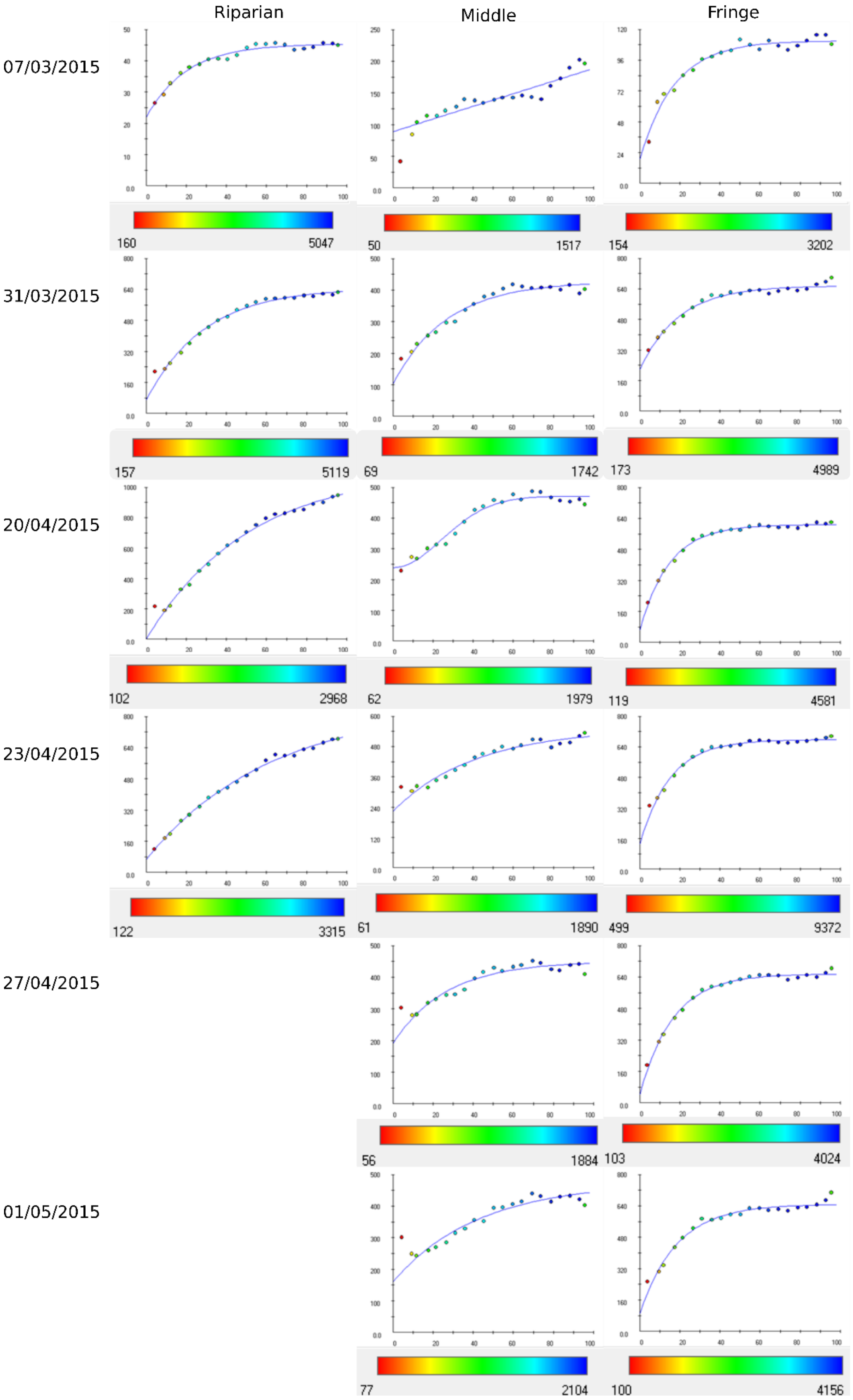
QQ PLOT 060216 MIDDLE

QQ PLOT 060216 FRINGE

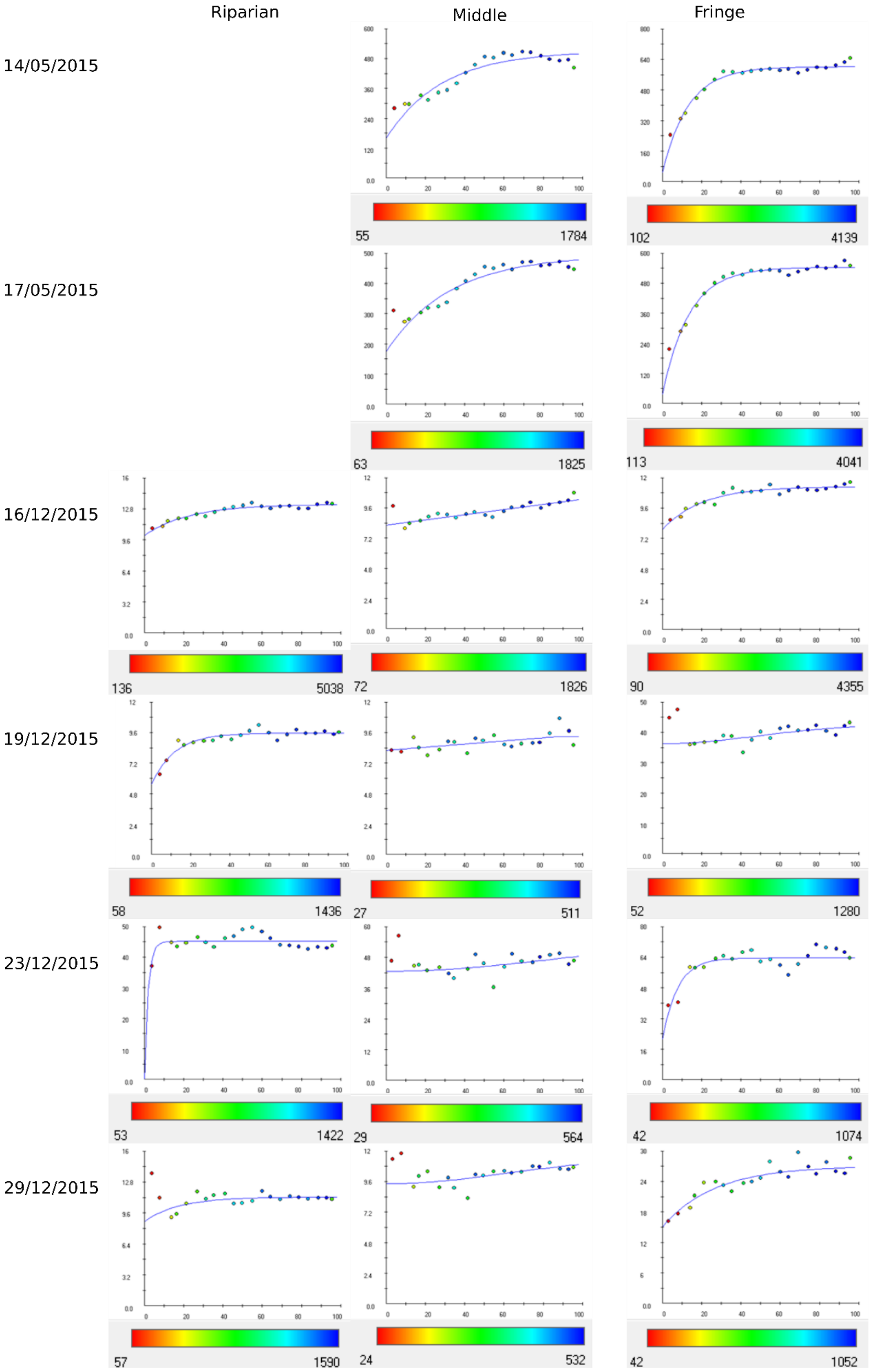
QQ PLOT 110216 MIDDLE

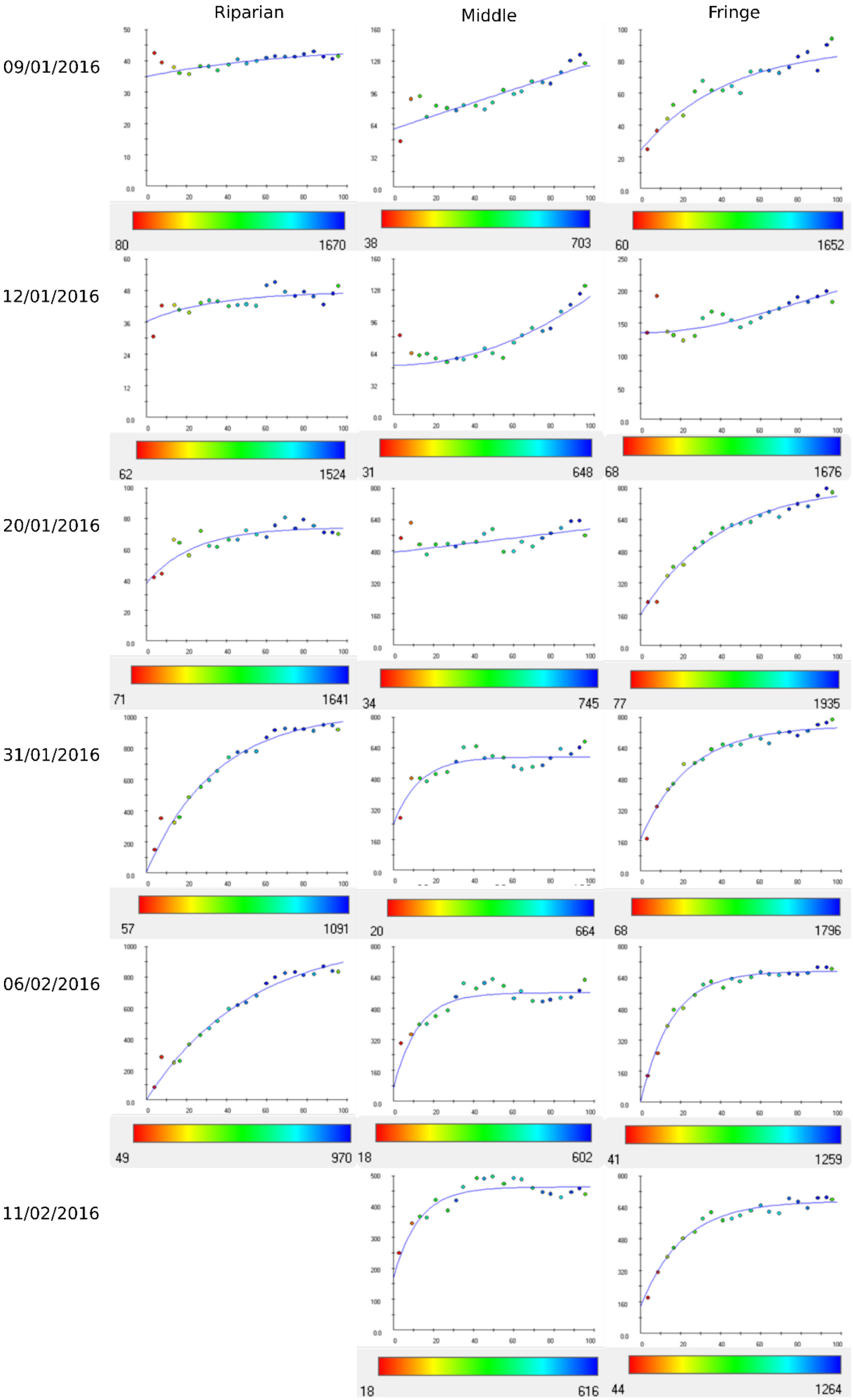
QQ PLOT 110216 FRINGE

B 1: Quantile-Quantile (QQ) plots revealing normally distributed soil moisture



Riparian
Middle
Fringe
07/03/2015
31/03/2015
20/04/2015
23/04/2015
27/04/2015
01/05/2015
0.0
10
20
30
40
50
0
20
40
60
80
100
160
5047
0.0
50
100
150
200
250
50
1517
0.0
24
48
72
96
120
154
3202
0.0
160
320
480
640
800
157
5119
0.0
100
200
300
400
500
69
1742
173
4989
0.0
200
400
600
800
1000
102
2968
62
1979
119
4581
122
3315
0.0
120
240
360
480
600
61
1890
499
9372
56
1884
103
4024
77
2104
100
4156


Riparian
Middle
Fringe
14/05/2015
17/05/2015
16/12/2015
19/12/2015
23/12/2015
29/12/2015
55
1784
102
4139
63
1825
113
4041
136
5038
72
1826
90
4355
58
1436
27
511
52
1280
53
1422
29
564
42
1074
57
1590
24
532
42
1052
0
20
40
60
80
100


Riparian
Middle
Fringe
09/01/2016
80
1670
38
703
60
1652
12/01/2016
62
1524
31
648
68
1676
20/01/2016
71
1641
34
745
77
1935
31/01/2016
57
1091
20
664
68
1796
06/02/2016
49
970
18
602
41
1259
11/02/2016
18
616
44
1264

B 2: Spatial Variograms for soil moisture for measurement dates between March to May 2015 and December 2015 to February 2016 for the Riparian, Middle and Fringe fields

B 3: Spatial semivariogram parameters. Strong spatial dependence is exhibited for values less than 25%, moderate for values between 25-50% and weak dependence for values greater than 75%.

riparian	model	c0 nugget variance	c1 partial sill	c0+c1 sill	Spatial dependence (%)		range	rmse
07-03-15	Exponential	21.98	23.47	45.45	48.36		20.28	0.90
31-03-15	Exponential	67.58	582.00	649.58	10.40		28.99	17.62
20-04-15	Exponential	0.00	1119.30	1119.30	0.00		51.24	31.25
23-04-15	Exponential	65.83	792.60	858.43	7.67		62.25	10.38
27-04-15								
01-05-15								
14-05-15								
17-05-15								
16-12-15	Exponential	10.06	3.22	13.28	75.78		23.11	0.22
19-12-15	Exponential	5.51	4.06	9.57	57.57		11.93	0.30
23-12-15	Exponential	0.00	45.17	45.17	0.00		1.89	2.33
29-12-15	Exponential	8.72	2.53	11.25	77.51		17.52	1.14
09-01-16	Exponential	34.97	9.89	44.86	77.95		73.13	1.95
12-01-16	Exponential	36.35	11.30	47.65	76.29		34.51	2.87
20-01-16	Exponential	37.68	36.59	74.27	50.73		23.69	5.14
31-01-16	Exponential	4.00	1031.10	1035.10	0.39		34.76	41.28
06-02-16	Exponential	0.00	1057.60	1057.60	0.00		51.69	39.63
11-02-16								
middle	model	c0 nugget variance	c1 partial sill	c0+c1 sill	Spatial dependence (%)		range	rmse
07-03-15	Exponential	88.90	50000.00	50088.90	0.18		50000.00	15.18
31-03-15	Exponential	102.80	325.00	427.80	24.03		26.97	15.44

20-04-15	Gaussian	238.60	233.00	471.60	50.59		34.40	14.51
23-04-15	Exponential	224.20	326.60	550.80	40.70		41.23	19.79
27-04-15	Exponential	191.70	259.10	450.80	42.52		26.78	21.61
01-05-15	Exponential	160.60	314.00	474.60	33.84		42.58	28.91
14-05-15	Exponential	160.70	350.30	511.00	31.45		28.74	31.66
17-05-15	Exponential	174.40	319.80	494.20	35.29		32.68	27.89
16-12-15	Exponential	8.23	1043.90	1052.13	0.78		50000.00	0.41
19-12-15	Linear with sill	8.21	1.09	9.30	88.31		88.27	0.53
23-12-15	Gaussian	42.59	13.76	56.35	75.58		131.30	4.20
29-12-15	Gaussian	9.43	2.16	11.59	81.40		88.86	0.83
09-01-16	Linear with sill	59.02	64.44	123.46	47.80		96.03	9.98
12-01-16	Gaussian	50.88	50000.00	50050.88	0.10		2610.50	9.78
20-01-16	Linear with sill	472.40	119.70	592.10	79.78		97.92	49.09
31-01-16	Exponential	240.30	352.20	592.50	40.56		13.06	44.80
06-02-16	Exponential	67.44	494.30	561.74	12.01		12.28	46.51
11-02-16	Exponential	169.20	294.70	463.90	36.47		12.60	25.08
fringe	model	c0 nugget variance	c1 partial sill	c0+c1 sill	Spatial dependence (%)		range	rmse
07-03-15	Exponential	18.62	92.60	111.22	16.74		17.38	4.04
31-03-15	Exponential	217.30	439.10	656.40	33.10		19.51	16.64
20-04-15	Exponential	64.12	544.60	608.72	10.53		14.79	9.29
23-04-15	Exponential	130.30	548.50	678.80	19.20		15.14	15.01
27-04-15	Exponential	42.25	614.00	656.25	6.44		16.60	13.22
01-05-15	Exponential	90.10	558.60	648.70	13.89		18.36	22.85
14-05-15	Exponential	46.92	555.00	601.92	7.80		13.29	20.29
17-05-15	Exponential	35.22	508.80	544.02	6.47		13.67	18.43
16-12-15	Exponential	7.90	3.41	11.30	69.85		20.32	0.25
19-12-15	Gaussian	36.31	6.38	42.69	85.06		68.18	3.44
23-12-15	Exponential	21.10	42.57	63.67	33.14		7.81	4.07
29-12-15	Exponential	15.04	12.03	27.07	55.56		26.76	1.40
09-01-16	Exponential	23.90	66.00	89.90	26.59		43.50	5.68
12-01-16	Gaussian	134.90	133.50	268.40	50.26		119.70	15.76

20-01-16	Exponential	155.70	659.20	814.90	19.11		39.53	24.32
31-01-16	Exponential	166.00	590.00	756.00	21.96		24.62	26.86
06-02-16	Exponential	0.00	674.00	674.00	0.00		15.55	18.83
11-02-16	Exponential	134.90	538.10	673.00	20.04		21.02	24.20

Standard error refers to the standard deviation of the coefficient estimate. The t value is a measure of the number of standard deviations the estimate is far away from 0. If it is far from zero, then a relationship between predictor and response variable exists. pr(>|t|) shows the probability of observing any value equal or larger than t. A small p value indicates it is highly unlikely to observe a relationship between predictor and response due to chance.

Example, for RadarSat-2 hh with measured khrms, the intercept is different to zero at 0.05 significance level, hh.linear is different to zero at 0.001level of significance while khrms is different to zero at 0.1 significance level. There the khrms, hh and intercept have small p values, concluding that there is a relationship between khrms, hh (backscatter) and soil moisture.

B 4: Coefficients for the linear equation for estimating soil moisture. HH, HV and VV are the coefficients for the backscatter, khrms is the measured normalized root mean square height representing the soil roughness, khrms_sim is the simulated roughness parameter

	estimate	standard error	t value	pr(>|t|)	
RadarSat-2 hh with measured khrms					
(intercept)	11.3	4.4	2.6	0.013	*
hh	86.5	17.9	4.8	0.000	***
khrms	-3.0	1.7	-1.8	0.084	.
RadarSat-2 hh with simulated khrms					
(intercept)	20.6	1.3	15.4	<2e-16	***
hh	169.3	7.3	23.2	<2e-16	***
khrms_sim	-13.8	0.7	-19.4	<2e-16	***
RadarSat-2 hv with measured khrms					
(intercept)	17.6	4.2	4.2	0.000	***
hv	419.7	113.0	3.7	0.001	***
khrms	-3.3	1.9	-1.8	0.085	.

RadarSat-2 hv with simulated khrms					
(intercept)	23.5	1.0	23.6	<2e-16	***
hv	993.9	42.0	23.7	<2e-16	***
khrms_sim	-13.4	0.6	-22.2	<2e-16	***
TerraSar-X hh with measured khrms	estimate	standard error	t value	pr(>|t|)	
(intercept)	20.7	5.3	3.9	0.001	***
hh	-27.5	15.6	-1.8	0.091	.
khrms	2.8	1.3	2.1	0.043	*
TerraSar-X hh with simulated khrms					
(intercept)	15.0	2.0	7.5	0.000	***
hh	81.6	9.8	8.3	0.000	***
khrms_sim	-6.9	0.5	-13.0	0.000	***
TerraSar-X vv with measured khrms					
(intercept)	20.6	5.1	4.0	0.000	***
vv	-32.5	17.7	-1.8	0.078	.
khrms	2.8	1.3	2.2	0.038	*
TerraSar-X vv with simulated khrms					
(intercept)	14.8	1.8	8.4	0.000	***
vv	108.0	10.8	10.0	0.000	***
khrms_sim	-11.9	0.8	-14.6	0.000	***

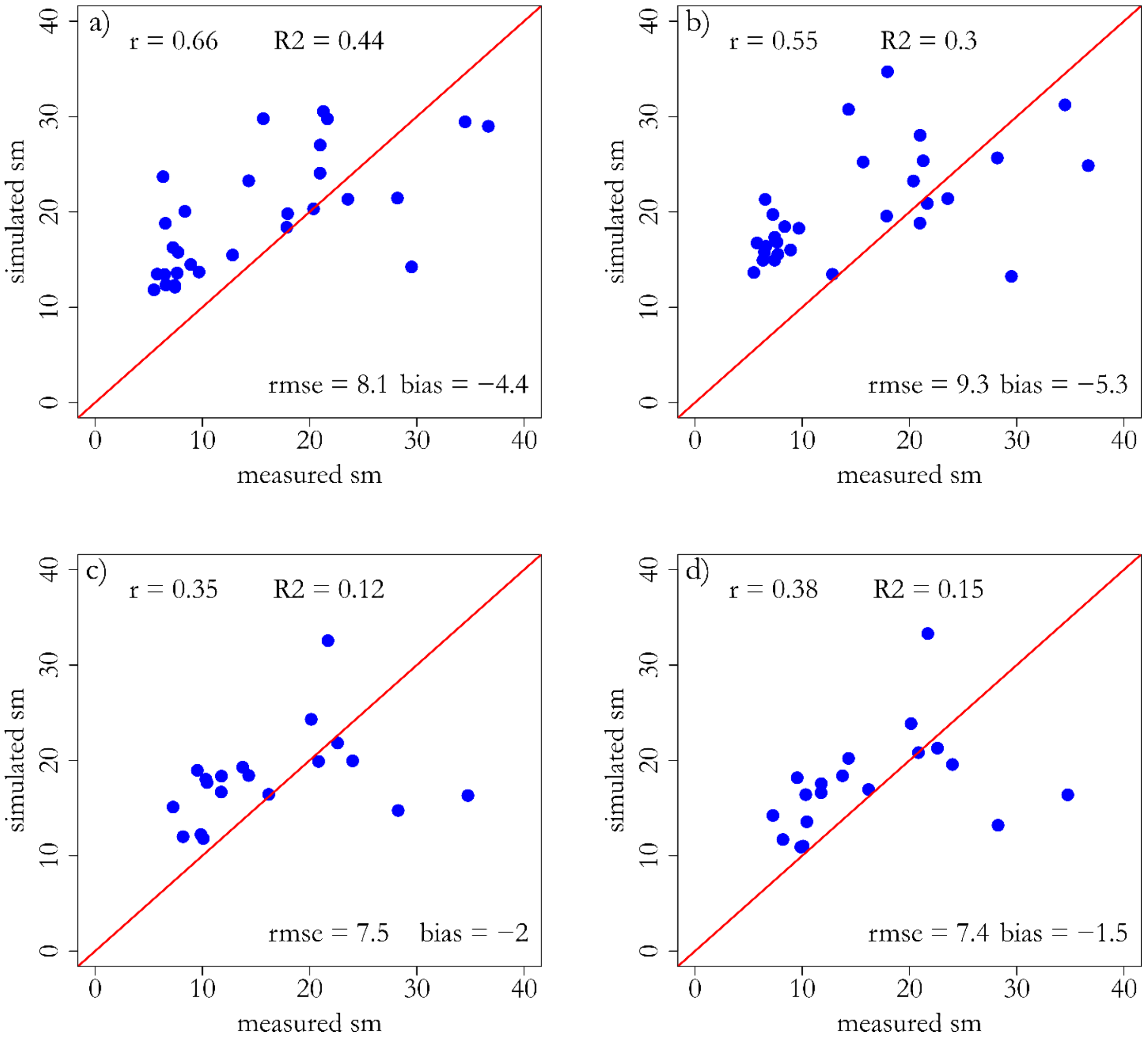
ns non-significant difference at 0.05 probability level

. Significance at 0.1 probability level

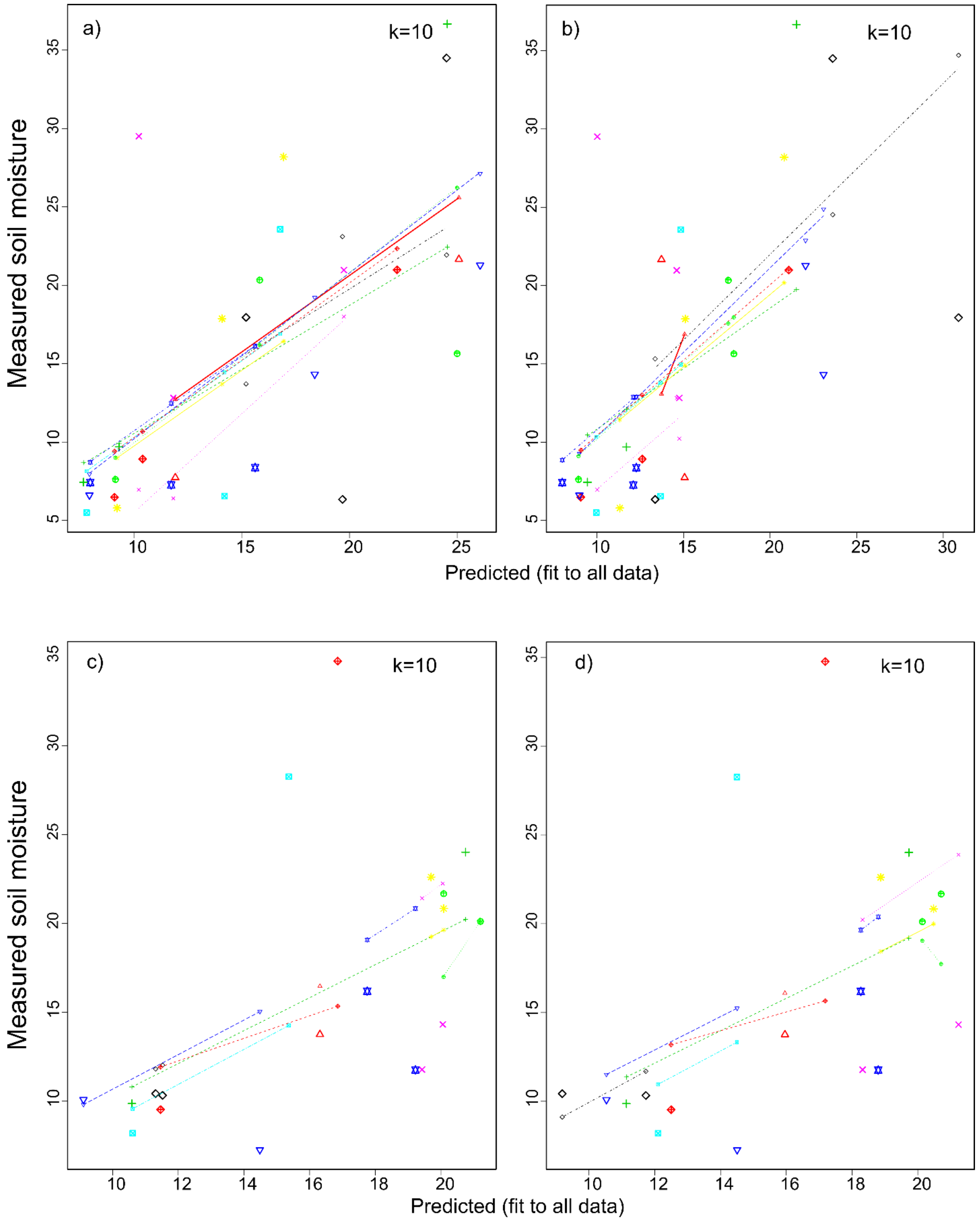
* Significance at 0.05 probability level

** Significance at 0.01 probability level

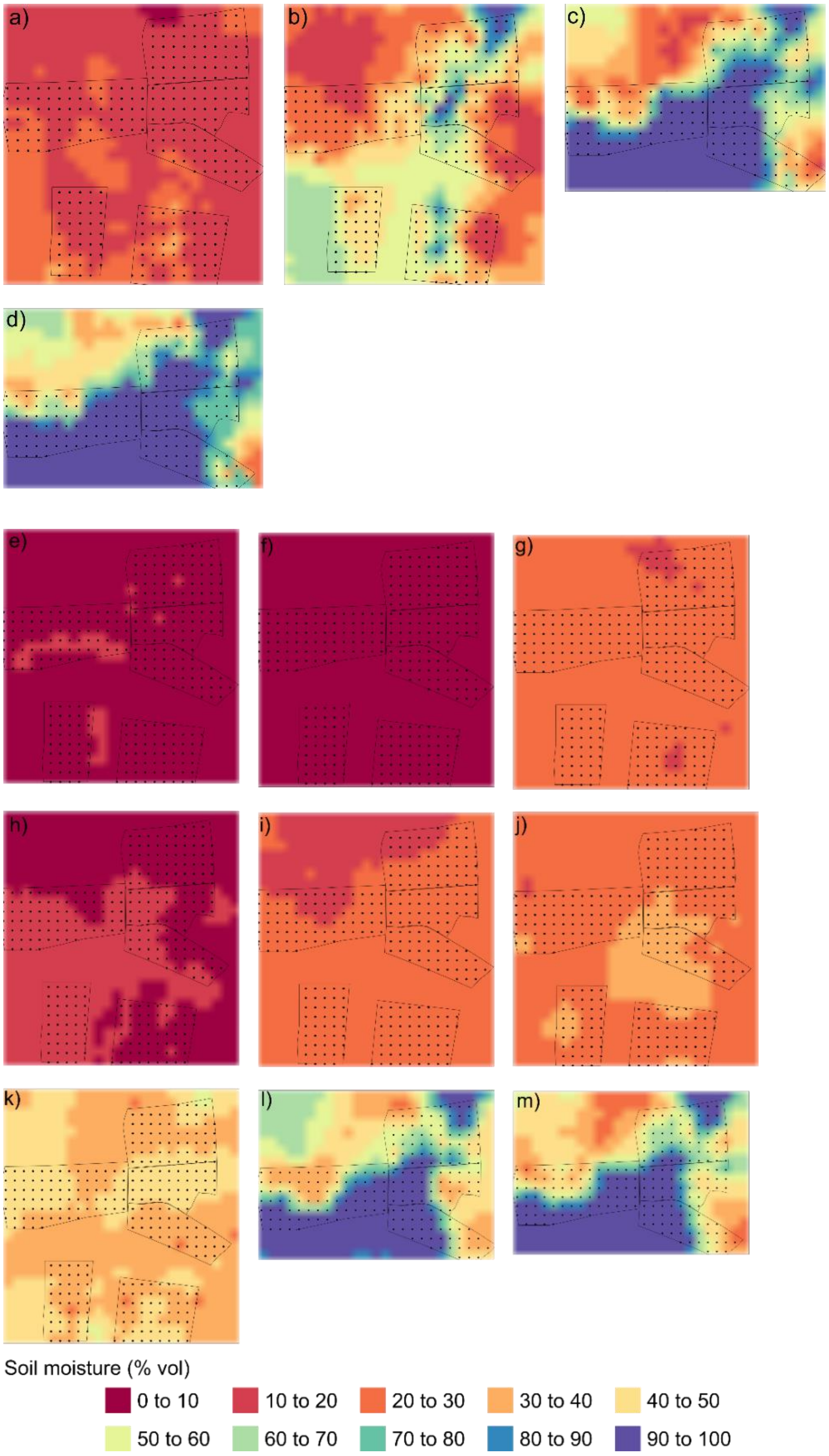
*** Significance at 0.001 probability level



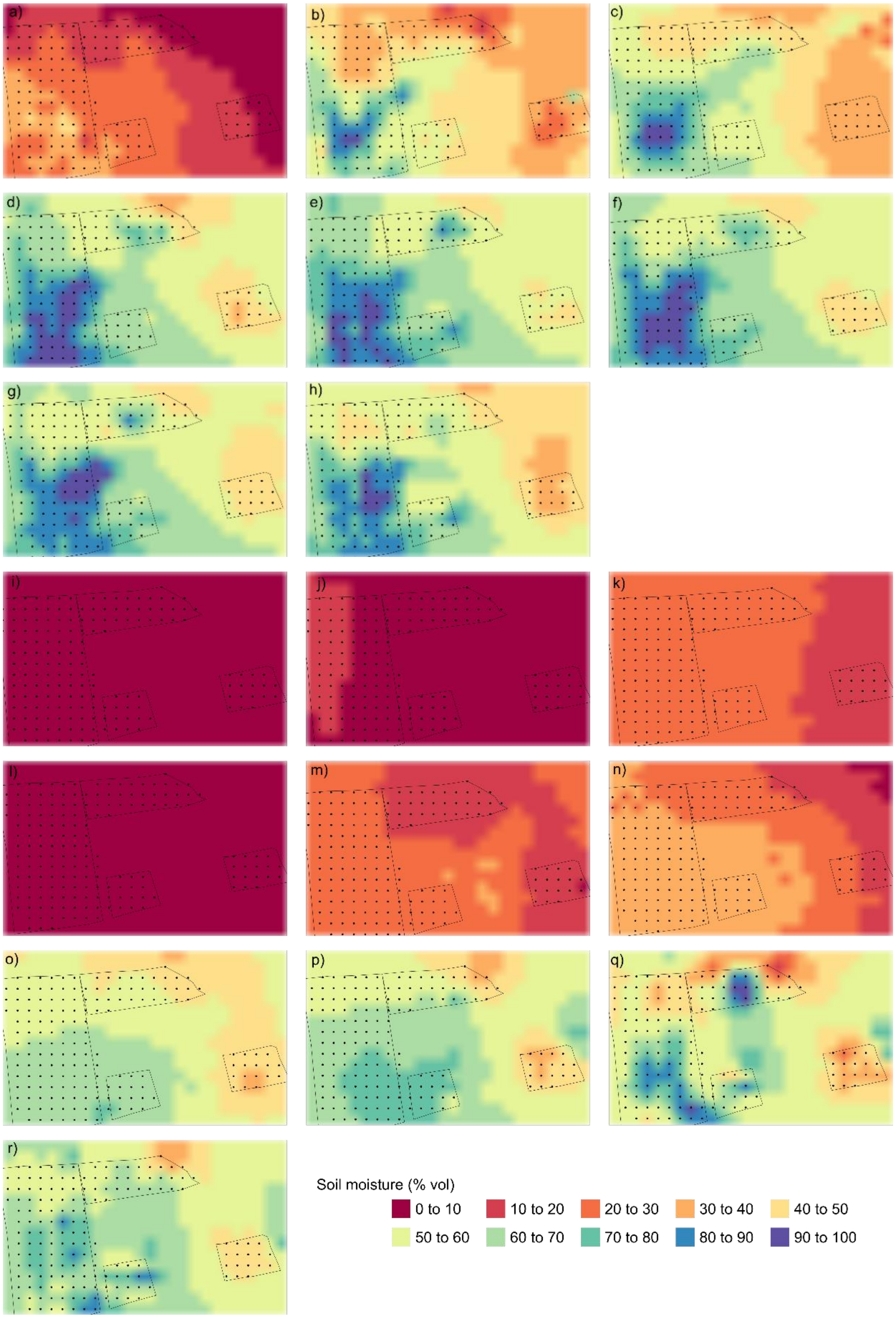
B 5: Soil moisture estimation from measured kHrms from (a) RadarSat-2 HH, (b) RadarSat-2 HV, (c) TerraSar-X HH, and (d) TerraSar-X VV



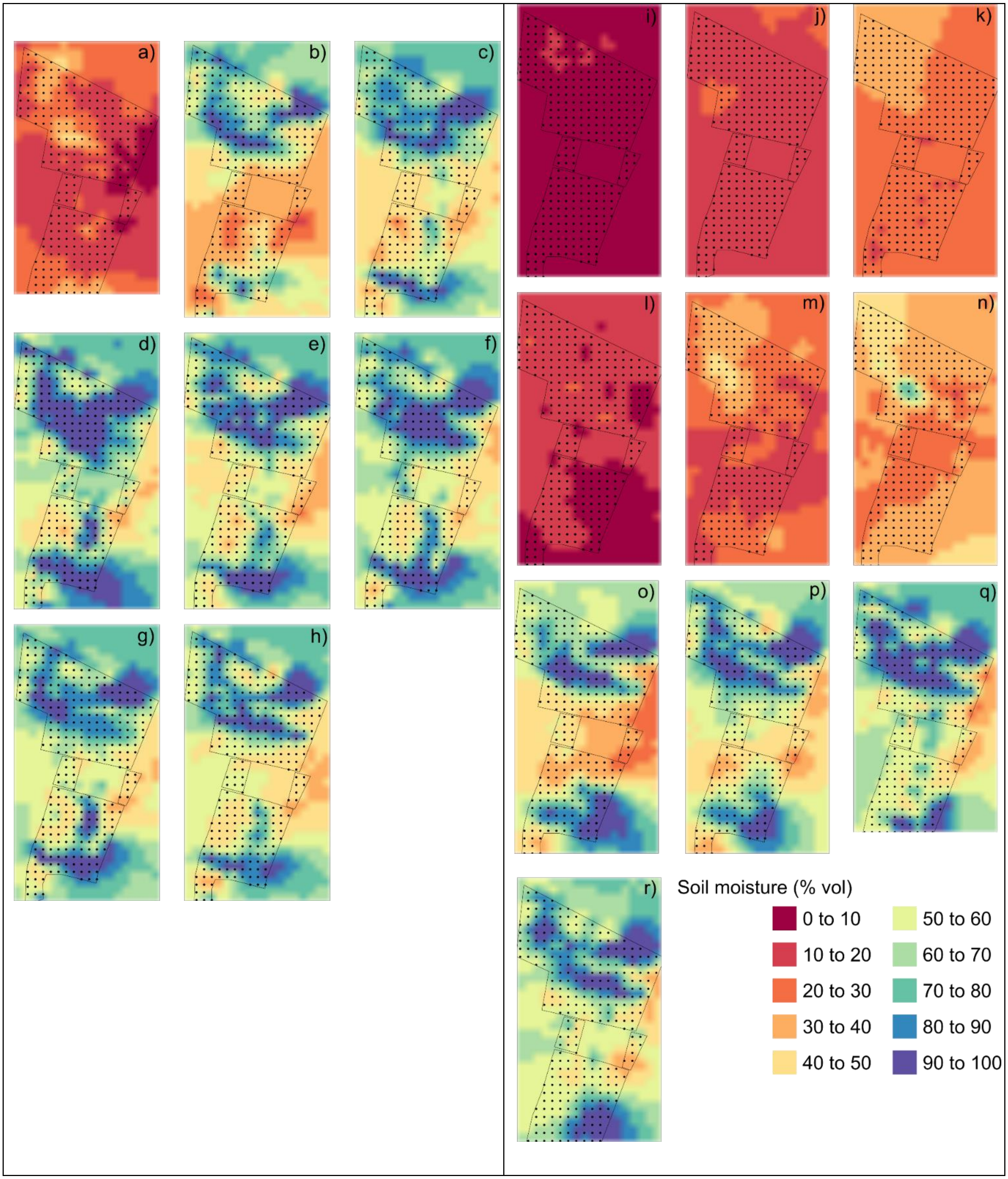
B 6: K fold cross validation with measured Khrms from (a) RadarSat-2 HH, (b) RadarSat-2 HV, (c) TerraSar-X HH, and (d) TerraSar-X VV



B 7: Soil moisture distribution maps for the riparian fields. a) 07-03-15, b) 31-03-15, c) 20-04-15, d) 23-04-15, e) 16-12-15, f) 19-12-15, g) 23-12-15, h) 29-12-15, i) 09-01-16, j) 12-01-16, k) 20-01-16, l) 31-01-16, m) 06-02-16

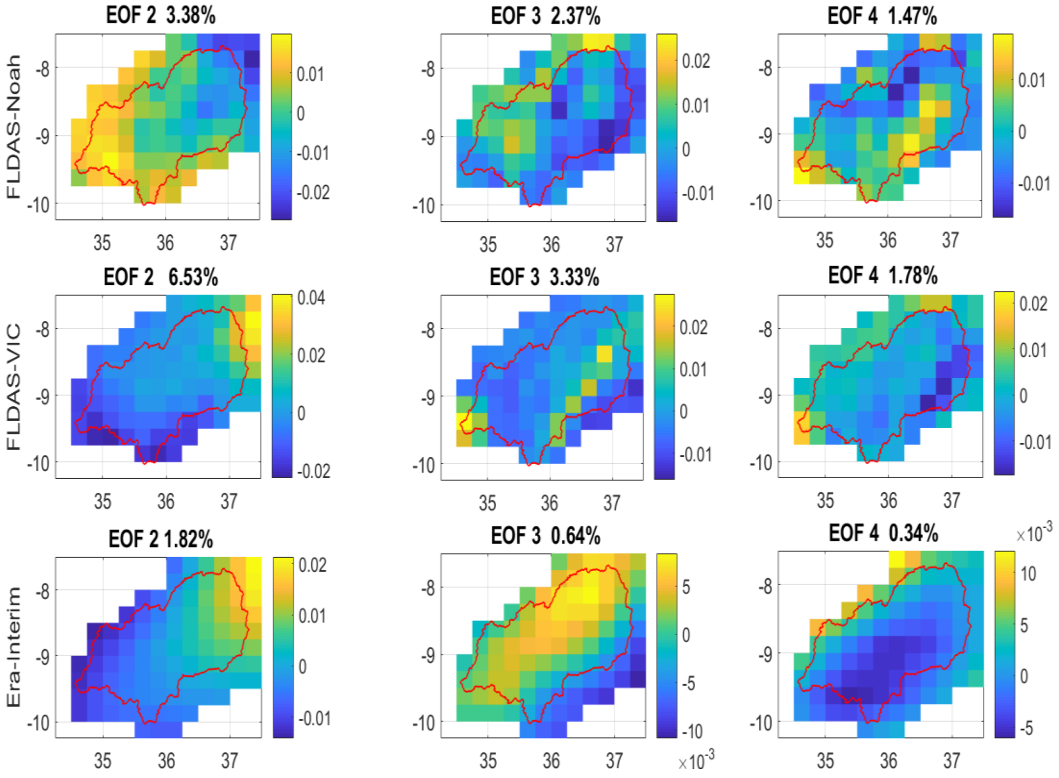


B 8: Soil moisture distribution maps for the middle fields. a) 07-03-15, b) 31-03-15, c) 20-04-15, d) 23-04-15, e)27-04-15, f) 01-05-15, g) 14-05-15, h) 17-05-15, i) 16-12-15, j) 19-12-15, k) 23-12-15, l) 29-12-15, m) 09-01-16, n) 12-01-16, o) 20-01-16, p) 31-01-16, q) 06-02-16, r) 11-02-16

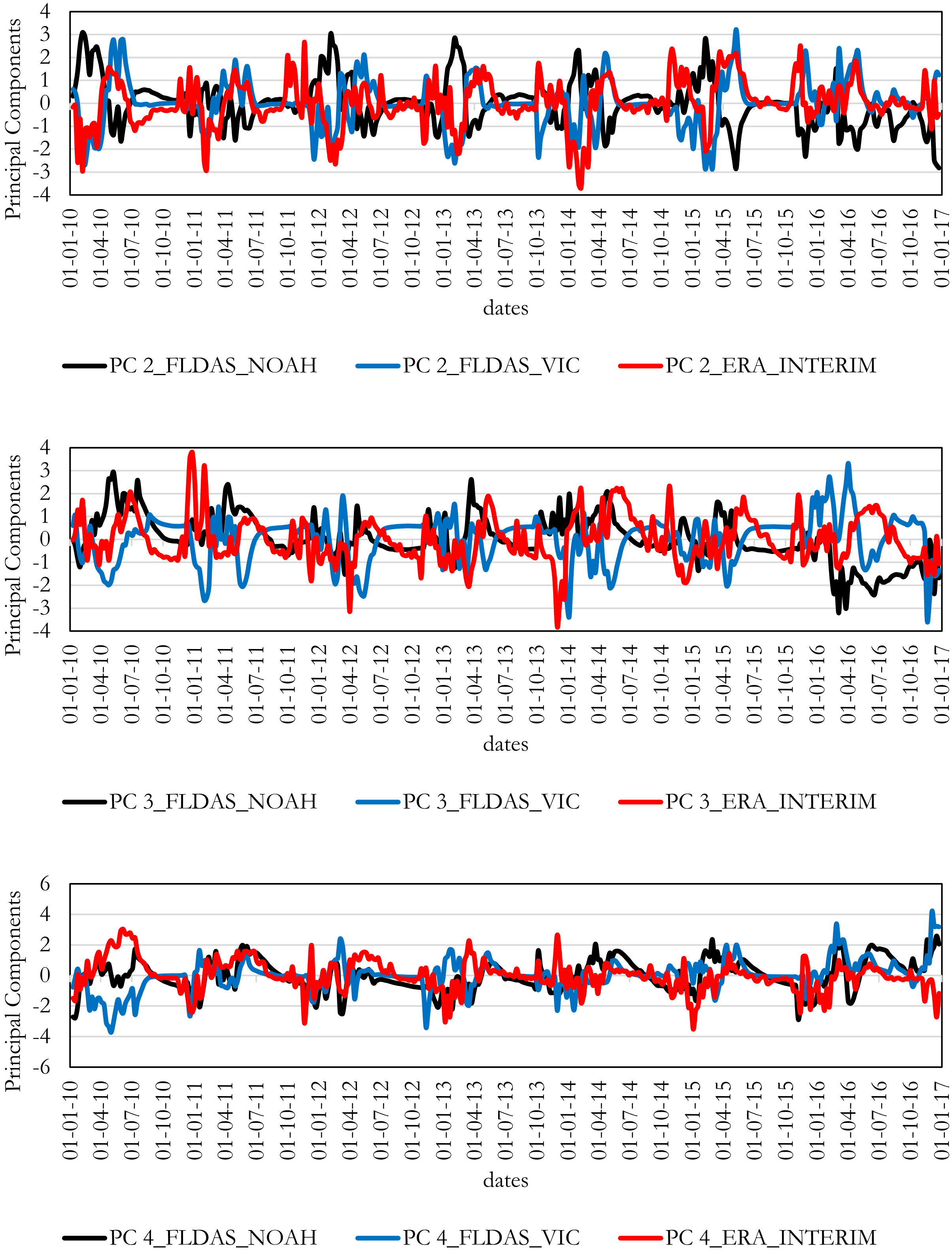


B 9: Soil moisture distribution maps for the fringe fields. a) 07-03-15, b) 31-03-15, c) 20-04-15, d) 23-04-15, e)27-04-15, f) 01-05-15, g) 14-05-15, h) 17-05-15, i) 16-12-15, j) 19-12-15, k) 23-12-15, l) 29-12-15, m) 09-01-16, n) 12-01-16, o) 20-01-16, p) 31-01-16, q) 06-02-16, r) 11-02-16

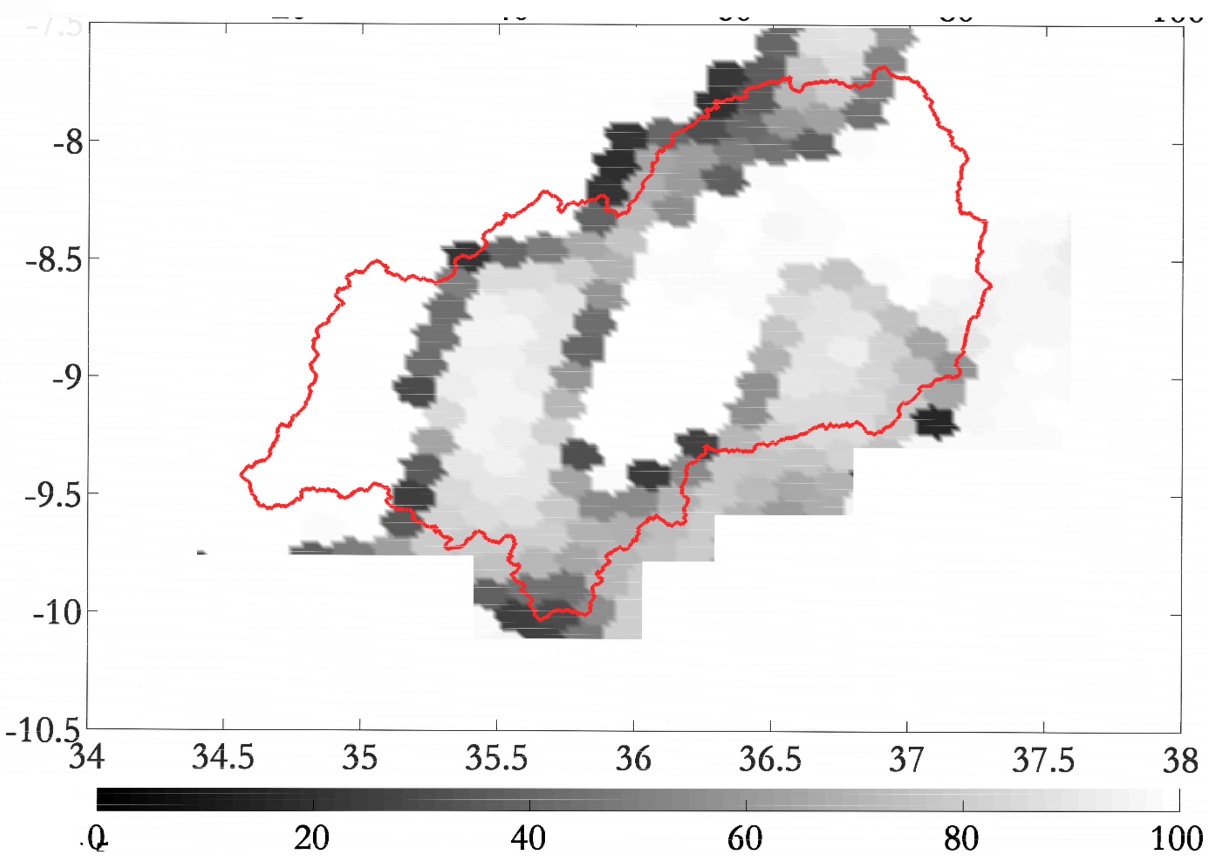
C Appendix: Global soil moisture assessment



C 1: The spatial variabilities as EOF's from PCA decomposition of the soil moisture products



C 2: Soil moisture spatial temporal variabilities derived from principal component analysis of the global moisture products



C 3: SMOS soil moisture coverage in percentage over the Kilombero catchment