PhD

3.º CICLO

FCUP
2019

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO
Statistical and physically based hyperspectral and multispectral reflectance modelling for agricultural monitoring: a case study in Vilankulo, Mozambique
Sosdito Estevao Mananze

Statistical and physically based hyperspectral and multispectral reflectance modelling for agricultural monitoring: a case study in Vilankulo, Mozambique

Sosdito Estevão Mananze

Tese de Doutoramento apresentada à
Faculdade de Ciências da Universidade do Porto
Departamento de Geociências Ambiente e Ordenamento do Território
Engenharia Geográfica
2019

STATISTICAL AND PHYSICALLY BASED HYPERSPECTRAL AND MULTISPECTRAL REFLECTANCE MODELLING FOR AGRICULTURAL MONITORING: A CASE STUDY IN VILANKULO, MOZAMBIQUE

Sosdito Estevão Mananze

Faculdade de Ciências da Universidade do Porto
Departamento de Geociências Ambiente e Ordenamento do Território
Plano doutoral em Engenharia Geográfica

2019

Orientador:

Professor Mário Cunha, Faculdade de Ciências da Universidade do Porto

Co-orientador:

Doutora Isabel Pôças, Centro de Investigação em Ciências Geo-Espaciais, da Faculdade de Ciências da Universidade do Porto

Nota prévia

Esta tese é resultado da sistematização de quatro trabalhos de investigação interligados, todos já publicados em revistas internacionais indexadas e com arbitragem científica. O candidato declara que foi o principal responsável pela conceção, recolha, análise de dados e discussão dos respetivos resultados e igualmente pela elaboração dos artigos publicados.

Os trabalhos foram realizados sob orientação do Doutor Mário Cunha, Professor Auxiliar do Departamento de Geociências, Ambiente e Ordenamento do Território da Faculdade de Ciências da Universidade do Porto e Investigador no Instituto de Engenharia de Sistemas Computacionais, Tecnologia e Ciência e colaborador no Centro de Investigação em Ciências Geo-Espaciais ambos da Universidade do Porto, e sob coorientação da Doutora Isabel Pôças Investigadora no Centro de Investigação em Ciências Geo-Espaciais da Faculdade de Ciências, Universidade do Porto.

A componente experimental dos trabalhos foi realizada em Moçambique, no distrito de Vilankulo, província de Inhambane e as restantes componentes foram maioritariamente conduzidas na Faculdade de Ciências da Universidade do Porto. Foi igualmente realizado um estágio de três meses no *Institute of Surveying, Remote Sensing and Land Information da University of Life Science and Natural Resources* de Viena, Áustria.

Este trabalho foi apoiado de uma ou de outra forma por várias entidades das quais se destacam a Fundação Calouste Gulbenkian através de uma bolsa de Doutoramento com a referência 140227, pelo Ministério de Ciência e Tecnologia, Ensino Superior e Técnico Profissional de Moçambique através do projeto HEST e pela Agência Austríaca para o Desenvolvimento, no âmbito do *Austrian Partnership Programme in High Education and Research for Development* (APPEAR), através do projeto número 176.

To my Family, with love!

Acknowledgements

I first want to thank GOD for my life, protection, and wisdom to conduct the research that gave rise to this thesis.

My deepest appreciation goes to my supervisor, Professor Mário Cunha, for his confidence, advice, encouragement, commitment, and unconditional support during the period of my study. He taught me how to be an independent scientist by letting me make my own choices at decisive stages along the way. No words can really describe my genuine thankfulness for his determinant supervision.

I would like to thank my co-supervisor, Doutora Isabel Pôças for her wisdom, continuous inspiration, and generosity. She was always ready to assist me, promptly answering my emails and questions. I highly appreciate her significant support, criticisms, expertise, and enthusiasm during the period of this work.

This work would not have been possible without the funding provided by the Gulbenkian Foundation through an individual doctoral scholarship (ref. 140227) and the Ministério de Ciência, Tecnologia, Ensino Superior e Técnico Profissional through the project HEST, which supported most of the expenses related with the research activities and my stays in Portugal.

I also want to appreciate the support given by the Eduardo Mondlane University, the Escola Superior de Desenvolvimento Rural. My colleagues were so kind, accepting to take over my teaching and administrative duties during my absences and supporting me during the field work. A special word to my director, Professor Simião Gabriel Balane for his support and comprehension!

I greatly thank my mom, brothers, and sisters for their unconditional moral support! To my church mates for their prayers!

I extended my gratitude to my children, Nkateko and Solange! Please, accept my apologies for not being able to be full-time dad you deserve thought I had to travel from time to time to Portugal for my work or although at home I had to be long time writing papers and thesis! Thanks for being two tolerant angels.

Lastly, but foremost, I cannot thank enough my beloved spouse, Edna! Baby, I am so sorry for all the time I had to be far away from you! You were so strong and intelligent, and you managed to take care of our children and to keep all our projects running properly! Your love, care, sympathy, and encouragement were so precious during the long way I had to run! Each day I love you more than the day before! My success is also yours!

Resumo

Esta tese tem enfoque na aplicação de dados e técnicas de deteção remota para a monitorização da agricultura no distrito de Vilankulo, província de Inhambane, região Sul de Moçambique. Nesta região, os sistemas de agricultura caracterizam-se sobretudo por parcelas de cultivo de pequenas dimensões, baixo nível de uso de fatores de produção, diversidade de culturas nas parcelas de cultivo, elevada heterogeneidade de condições biofísicas na parcela e entre parcelas e baixos níveis de intensidade fitotécnica, nomeadamente nos sistemas de irrigação. Estas características configuram um ambiente extremamente desafiador para a aplicação de dados e técnicas de deteção remota pois, na maioria dos casos, estes são calibrados e validados para aplicações em contextos de agricultura intensiva.

Neste contexto, o principal objetivo da presente tese é testar e validar metodologias cientificamente robustas para a monitorização de culturas agrícolas desenvolvidas em sistemas extensivos, de modo a apoiar o processo de tomada de decisão, tanto ao nível do agricultor, quanto ao nível dos decisores de políticas agrárias. Neste contexto, as metodologias consideradas nesta tese foram testadas e validadas para o milho, a principal cultura de subsistência na região, e tendo em conta os sistemas de agricultura e paisagem locais. O projeto de investigação realizado consistiu em quatro tópicos distintos porém interligados, que resultaram em quatro artigos científicos versando aspetos relevantes para a monitorização de parâmetros biofísicos de culturas agrícolas.

Artigo um: O tópico deste primeiro artigo está relacionado com a estimativa, em campo, da área foliar da cultura do milho. Foi calibrado e validado um modelo matemático com base em variáveis alométricas para a estimativa da área foliar do milho. O modelo em alusão consiste no produto entre o número de folhas na planta, o comprimento e a largura da folha maior, tendo-se mostrado adequado para a determinação da área foliar da cultura do milho em todas as fases do crescimento da cultura ($R^2 = 0.90$, $P < 0.000$).

Artigo dois: Este artigo explora diferentes tipos de dados e métodos de deteção remota para a estimativa do índice de área foliar. Para o efeito, dados multiespectrais (do sensor Sentinel-2) e hiperespectrais (obtidos em campo e simulados através de modelos de transferência radiativa) foram testados com base em métodos empíricos (paramétricos e não paramétricos) e físicos (inversão através da abordagem *Look-up table*, LUT). Os dados de campo para a validação das três abordagens foram obtidos através do modelo desenvolvido no primeiro artigo. Com relação ao método paramétrico, os índices de vegetação envolvendo três bandas do espectro eletromagnético foram os mais eficazes, nomeadamente: (i) índice da diferença entre bandas ($mDI_{d: 725; 715; 565}$) para dados hiperespectrais (RMSE = 0.58), (ii) razão entre bandas ($mSR_{c: 740; 705; 865}$) para os dados multiespectrais reamostrados (RMSE = 0.62), e (iii) $TBSI_{b:665;865;783}$ para dados multiespectrais do Sentinel-2 (RMSE = 0.18). Na aplicação de métodos não paramétricos foram usadas técnicas de *Machine learning*, tendo-se verificado que o algoritmo *relevant vector machine* foi mais eficaz quando aplicado aos dados hiperespectrais (RMSE = 0.50) e multiespectrais (RMSE = 0.53) do campo enquanto o algoritmo *Random Forest* (RMSE = 0.22) mostrou melhor desempenho com os dados do Sentinel-2. Usando a inversão dos modelos de transferência radiativa (método de base física), a função de custo *Bhattacharyya divergence* foi a mais eficaz. Os três métodos lograram estimar o índice de área foliar, tendo o método paramétrico obtido melhor eficácia em relação aos outros. A comparação geral revelou que o método paramétrico aplicado aos dados do Sentinel-2 e através do índice $TBSI_{b: 665; 865; 783}$ foi o mais eficaz.

Artigo três: O terceiro artigo teve como objetivo a estimativa da humidade do solo e a sua utilização na monitorização de indicadores de seca agrícola. A humidade do solo em três profundidades (15, 25 e 30 cm) foi estimada com recurso a um modelo de base física, o *Optical Trapezoid Model* (OPTRAM), e aplicada para monitorização multi- temporal e espacial da “seca agrícola” em campos agrícolas de milho e soja. O OPTRAM foi aplicado a dados espectrais do Sentinel-2 e os dados para a sua validação foram recolhidos no campo com base no método gravimétrico e em tensiómetros. O OPTRAM estimou com precisão a humidade do solo na zona radicular de 15 cm (RMSE = 0.025) e 25 cm (RMSE = 0.030) e a mesma foi usada com sucesso para determinar e mapear o nível de “seca agrícola” nas parcelas em estudo.

Artigo quatro: O quarto artigo trata do mapeamento do coberto do solo dando ênfase à avaliação da dinâmica espácio-temporal dos sistemas agrícolas no distrito de Vilankulo, maioritariamente dominados por pequenas áreas e com baixo nível de intensidade fitotecnica. O mapeamento foi desenvolvido na plataforma *Google Earth Engine* (GEE), aplicando o classificador *Random Forest* e usando como variáveis três bandas espectrais de imagens Landsat 7 ETM+ e Landsat 8 OLI/TIRS e cinco métricas de textura extraídas em 3 bandas espectrais. Para os três anos considerados – 2012, 2015 e 2018 –, os mapas de coberto do solo foram produzidos com um rigor (*overall accuracy*) igual ou superior a 89% e as estatísticas das áreas agrícolas derivadas dos mapas foram próximas às das estimativas produzidas pelas autoridades locais do sector agrícola. Os resultados deste estudo, nomeadamente no que respeita ao mapeamento da extensão, distribuição e características das áreas agrícolas, são extremamente importantes para o suporte de medidas para mitigar os impactos dos principais desafios societais em Moçambique, nomeadamente, a segurança alimentar, a produção agrária sustentável e a redução da pobreza. De facto, a informação consistente sobre as principais áreas de produção é necessária para suportar decisões relacionadas com o investimento, a logística, bem como a monitorização da produção. Esta informação quando obtida em tempo útil, pode ser usada em sistemas de aviso prévio de sistemas alimentares.

Os resultados desta tese têm um enorme potencial de aplicação tanto pelos agricultores como pelos legisladores e decisores ao nível de políticas agrárias e de segurança alimentar. Aos agricultores estes resultados podem ajudar na identificação atempada de áreas/situações de desenvolvimento anómalo das culturas, que poderão resultar da incidência de pragas, doenças, deficiências de nutrientes e/ou de água, permitindo-lhes tomar ações de reação em tempo útil e em áreas específicas. Os resultados poderão também apoiar uma melhor gestão do uso da água pelos agricultores através da disponibilização regular de informação acerca do teor de água no solo e identificação de áreas de “seca agrícola”. Por seu turno, os decisores podem utilizar a informação e as ferramentas desenvolvidas nesta tese para analisar de forma mais objetiva a evolução espacial e temporal das áreas agrícolas bem como o desenvolvimento das culturas ao longo dos ciclos culturais, com consequências sobre os níveis de produção, e por via disso, tomar decisões em relação a questões de segurança alimentar, necessidade de ajuda alimentar, importação ou exportação de alimentos. Por outro lado, a informação gerada sobre as dinâmicas dos diferentes tipos de cobertura do solo poderá apoiar a definição de estratégias relacionadas com a conservação de recursos naturais e da biodiversidade.

Palavras-chave: Índice de área foliar, teor de água no solo, híper- e multiespectral, machine learning, PROSAIL, vegetation indices

Abstract

This thesis focuses on the application of remote sensing (RS) data and techniques for agricultural monitoring in the district of Vilankulo, Inhambane province, in southern Mozambique. The main characteristics of agricultural systems in the region include the small size of cultivated areas, low level of input application, mixed crop systems and high biophysical heterogeneity within and between parcels. These characteristics shape a highly challenging environment for the implementation of RS techniques, which in many cases have been calibrated/validated in contexts of intensive agriculture.

The main aim of the thesis is to test and validate scientifically robust methods for crops monitoring in order to support the decision-making process at both farm and policy maker levels. In this context, the methodologies considered in this thesis were tested and validated for maize, which is the main subsistence crop in the region and also, considering the local agricultural systems and landscape pattern. A research project consisting of four different, but interconnected topics was implemented and yielded four research papers focusing on relevant issues of crop biophysical monitoring.

Paper one: this paper addresses the *in-situ* estimation of leaf area of maize. An allometric-mathematical model for non-destructive estimation of maize leaf area was calibrated and validated. The developed model is based on the number of leaves per plant, the major leaf length and the major leaf width of the greater leaf. This model showed to be suitable for estimating accurately maize leaf area in all crop growing stages ($R^2 = 0.90$, $P < 0.000$).

Paper two: the objective of this paper two was to remotely estimate maize leaf area index. Multispectral satellite (Sentinel-2), *in-situ* hyperspectral and physically simulated hyperspectral datasets were tested to retrieve maize leaf area index, by using empirically based (parametric and non-parametric) and physically based (look-up table inversion) processing methods. Ground truth data was estimated applying the model developed in paper one. For the parametric approach, three band spectral vegetation indices were the best performing: (i) the modified difference index ($mDI_{d:\ 725;\ 715;\ 565}$) for the in-situ hyperspectral dataset (RMSE = 0.58), (ii) the modified simple ratio ($mSR_{c:\ 740;\ 705;\ 865}$) for the multispectral dataset resampled from in-situ hyperspectral data (RMSE = 0.62), and (iii) the three band spectral vegetation index ($TBSI_{b:\ 665;\ 865;\ 783}$) for the Sentinel-2 dataset (RMSE = 0.18). For the non-parametric approach, the relevant support vector machine was the most accurate machine learning regression algorithm (MLRA) using the two datasets of field spectra [multispectral (RMSE =0.53) and hyperspectral (RMSE

=0.50)], whereas the MLRA Random forest (RMSE = 0.22) performed better for the Sentinel-2 data. When using the LUT inversion technique (physically based approach), the Bhattacharyya divergence cost function (RMSE = 0.20) were the best performing. Comparing all the approaches and datasets tested, the parametric one applied to Sentinel-2 data outperformed the other two approaches with the $\text{TBSI}_{b:\ 665;\ 865;\ 783}$ being the most accurate index.

Paper three: this paper aimed at remotely estimate soil moisture and subsequently monitor and map agricultural drought. Soil moisture at three depths (15, 25 and 30 cm) was retrieved through the Optical Trapezoidal Model (OPTRAM) and used for multi-year, multi-site monitoring of agricultural droughts over Maize and Soybean croplands in the study area. The OPTRAM was implemented using spectral data from Sentinel-2 and was validated against field soil moisture assessed by gravimetric methods and by Watermark Sensors in sandy-soils with very low water holding capacity. The OPTRAM accurately estimated the soil moisture at 15 cm (RMSE = 0.025) and 25 cm (RMSE = 0.030) and the estimates were successfully used as input to compute and map the soil water deficit index (SWDI), an indicator of agricultural drought.

Paper four: the paper is focused on the land cover mapping with emphasis on the assessment of smallholder agriculture dynamics. The Random Forest ensemble classifier, implemented within the Google Earth Engine (GEE) cloud computing platform, was used in the classification process. The input variables for classification consisted of three spectral bands of Landsat 7 ETM+ and Landsat 8 OLI/TIRS and five textural features extracted from the three non-correlated Landsat bands. For the three years considered "- 2012, 2015 and 2018 –", the land cover maps were produced with an overall accuracy equal or above 89% and the agricultural statistics derived from the maps were very similar to those estimated by the local agrarian entities. Mapping the cropland extent, distribution and characteristics is of great importance to support measures to mitigate the most challenging societal problems in Mozambique such as food security, sustainable agricultural production, and poverty reduction. In fact, accurate and consistent information on cropland is required to support future policy, investment, and logistical decisions, as well as production monitoring. Moreover, timely cropland information can directly feed early warning systems.

This thesis results have great potential for operational application by both farmers and agricultural and food security policy makers. For farmers, the results may assist in the early identification of growth anomalies caused by pests and diseases, nutrient or water deficit and thus enable them to decide and implement corrective actions on time and in specific areas. Also, the results may support a better management of water use in the cultivation areas by making available regular information about soil moisture and by identifying drought areas. For policy makers the results may be useful in accurately analyzing the spatial and temporal dynamics of agricultural

areas as well as the crop growth throughout the crops cycle, with direct impact on crop yields, and as such decide about food security, food aid and importation or exportation of food. Additionally, the information about land cover dynamics across time can help assisting the definition of conservation strategies for natural resources and biodiversity.

Key words: Leaf area index, soil moisture, hyper and multispectral, machine learning, PROSAIL, vegetation indices.

Contents

List of tables

List of figures

Abbreviation

ANN: Artificial Neural Network

CAADP: Comprehensive African Agriculture Development Program

CCRS: Canadian Center of Remote Sensing

CEOS: Committee on Earth Observation Satellites

CMI: Crop Moisture Index

DN: Digital Number

EMR: Electromagnetic Radiation

EOS: Earth Observation Satellites

ESA: European Space Agency

ETDI: Evapotranspiration Deficit Index

FAO: Food and Agriculture Organization

GDP: Gross Domestic Product

GEO: Group of Earth Observation

GEOGLAM: Global Agricultural Monitoring Initiative

Hz: Hertz

INE: Instituto Nacional de Estatística (National Institute of Statistics)

IUCN: International Union for Conservation of Nature

LAI: Leaf Area Index

LUT: Look-up Table

MICOA: Ministério para a Coordenação da Acção Ambiental (Ministry for Coordination of Environmental Affairs)

MLA: Machine Learning Algorithm

MW: Microwave

ND: Normalized Difference

NDVI: Normalized Difference Vegetation Index

NIR: Near Infrared

OBIA: Object-based image analysis

OPTRAM: Optical Trapezoid Model

PAW: Plant Available Water

PDI: Palmer Drought Index

PNISA: Programa Nacional de Investimento no Sector Agrário (National Program of Investment in Agrarian Sector)

PROSPECT: Leaf Optical Proprieties Spectra

RS: Remote Sensing

RTM: Radiative Transfer Model

SAIL: Scattering by Arbitrarily Inclined Leaves

SAR: Synthetic Aperture Radar

SDG: Sustainable Development Goals

SM: Soil Moisture

SMAP: Soil Moisture Active Passive

SMI: Soil Moisture Index

SMDI: Soil Moisture Deficit Index

SMOS: Soil Moisture and Ocean Salinity

SR: Simple Ratio

SSA: Sub Saharan Africa

STR: Shortwave Transformed Reflectance

SWD: Soil Water Deficit

SWDI: Soil Water Deficit Index

SWIR: Short Wave Infrared

TIR: Thermal Infrared

UAV: Unnamed Aerial Vehicle

UN: United Nations

UV: Ultraviolet

VI: Vegetation Index

VIS: Visible

WDI: Water Deficit Index

Introduction

The interest on using remote sensing (RS) techniques in agriculture has been increasing as a result of the potential of such tools to cope with some of the specificities of the agricultural activities (FAO 2011), including the strong seasonal patterns related to the biological lifecycle of crops, the dependence on the physical landscape (e.g., soil type), as well as climatic driving variables and agricultural management practices. Furthermore, as crop attributes can change within short time periods due to biophysical changing conditions, the agricultural systems monitoring needs to be timely, synoptic, cost efficient and repetitive coverage to provide valuable information. RS is suitable to fill this requirement as it can gather information over large areas with high temporal resolution (FAO 2011, Atzberger 2013).

The data recorded by RS sensors mounted in different platforms, integrating earth observation systems (EOS), can be used to assess several components of agriculture and agronomic process, including: crops yield (Doraiswamy, Sinclair et al. 2005, Cunha, Marçal et al. 2010, Bernardes, Moreira et al. 2012, Atzberger and Rembold 2013, Meroni, Marinho et al. 2013); acreage (Hannerz and Lotsch 2008, Fritz, See et al. 2015, Waldner, Fritz et al. 2015); crop phenological information (Sakamoto, Yokozawa et al. 2005, Rodrigues, Marcal et al. 2013, Shen, Wu et al. 2013); stress conditions (Gu, Brown et al. 2007, Atzberger and Rembold 2013, Rembold, Atzberger et al. 2013), trends and disturbances (Verbesselt, Hyndman et al. 2010); soil moisture and drought indicators (Soliman, Heck et al. 2013, Santos, Silva et al. 2014, Martínez-Fernández, González-Zamora et al. 2015, Sohrabi, Ryu et al. 2015, Martínez-Fernández, González-Zamora et al. 2016, Sadeghi, Babaeian et al. 2017, Mananze, Pôças et al. 2019), and biochemical and biophysical parameters (Blackburn and Ferwerda 2008, Li, Cheng et al. 2008, Bakhshandeh, Kamkar et al. 2011, Berjo, Cachorro et al. 2013, Clevers and Gitelson 2013, Pôças, Paço et al. 2014, Pôças, Paço et al. 2015, Verrelst, Rivera et al. 2015, Wenng 2017, Berger, Atzberger et al. 2018, Mananze, Pôças et al. 2018). The use of RS in agriculture can be extended from the conventional top view approach – either from airborne platform or EOS – to close data acquisition through ground measurements using various instruments (e.g. spectroradiometer) (Mananze, Pôças et al. 2018).

Amongst other things, the retrieved information can support the decision-making process, at farm level, to rapidly react and alleviate water stress, effects of pests, diseases, and/or nutrient insufficiency. Additionally, the retrieved information can support policy makers decision process to better anticipate and cope with the effects of disturbance events (e.g., climate-related hazards), to get an objective and unbiased spatial picture over large areas for risk assessment, and to define suitable strategies to ensure food security and natural resources management (Atzberger 2013).

Despite the recognized contribution of RS in both agronomic process inference and agricultural monitoring, its application in resource-poor smallholder farming systems, particularly in sub-Saharan Africa (SSA) has been limited. This is mostly because in a recent past, RS data acquisition and their processing were mostly associated to high costs and to complex computer software and processing algorithms, constituting a limiting factor mainly in SSA countries (Schimmelpfennig 2016, Say, Keskin et al. 2018). In addition, smallholder farming systems are characterized by very small sized fields, with mixed crops and including a few multipurpose trees, which present high temporal and spatial change dynamics; such conditions are very challenging to monitor using the freely available satellite imagery (Hannerz and Lotsch 2008, Schimmelpfennig 2016, Say, Keskin et al. 2018).

Due to these constraints, the farmers with poor resources are compelled to rely on intuition and highly uncertain methods to obtain information for decision making (Kpienbaareh, Kansanga et al. 2018). Nevertheless, these farming systems constitute the socio-economic backbone of many African countries and provide the basis for food security and thus, it is an imperative to adapt data and tools that enable an appropriate monitoring and mapping of these traditional agricultural systems. This is particularly crucial in the face of rising climate variability to which many African countries are highly exposed and vulnerable.

In this context, this thesis explores, examine and validate different types of RS data and methods for the retrieval of maize Leaf Area Index (LAI), Soil Moisture (SM), and Agricultural Drought, and for mapping the croplands extent and dynamics. Although the main applications were developed within the work setting of the Vilankulo district, southern Mozambique, they are expected to be generalized, at least, for other agricultural systems in the SSA.

The LAI is a quantitative measure of the amount of green leaf material present in the canopy per unit ground surface. It is defined as half of the total green leaf area (one-sided area for broad leaves) per unit ground surface (Chen and Black 1992). The LAI describes a fundamental property of the plant canopy in its interaction with the atmosphere, especially concerning radiation, energy and gas exchange (Monteith and Unsworth 1990). The plant' leaf area plays a key role in the absorption of radiation, in the deposition of photosynthates during the diurnal and seasonal cycles, and in the pathways of rates of biogeochemical cycling within the canopy-soil system (Bonan 1995). Hence, it is a crucial variable for crop growth evaluation and yield forecasting (Pan, Chen et al. 2018) and thus is widely applied in the field of agronomy, but is also applied in the fields of forestry, ecology, hydrology, eco-physiology and meteorology (Cattanio 2017, Fang, Ye et al. 2018).

The SM is recognized as a key variable in controlling the water and energy fluxes exchange between the earth surface and the atmosphere, through evaporation and transpiration processes (Wang, Qu et al. 2007, Corbari, Sobrino et al. 2010) and, has a crucial role in the biosphere as it guarantees the vegetation growth (Kerr 2007). The information about SM in the radicular zone has great application in studies related to crop growth modelling (Guérif and Duke 2000, de Wit and van Diepen 2007), meteorology (Betts and Ball 1998), global and regional climate modelling (Wetzel and Chang 1988, Beljaars, Viterbo et al. 1996). Moreover, SM is a suitable variable to monitor agricultural drought events, which are defined as a shortage of water to such a point that adversely affects the crop growth (Panu and Sharma 2002, Martínez-Fernández, González-Zamora et al. 2015).

Agricultural land use dynamics is the major cause of ecosystems degradation at global scale (Benayas and Bullock 2012) so, land use monitoring related to farming is important for land management (Schwilch, Bestelmeyer et al. 2011). The estimation of crops area extent and mapping of crop types and of crops dynamics provide crucial information for agricultural monitoring and management (Drusch, Del Bello et al. 2012), as well as to support agricultural investment policy and early warning systems related to food security (Hannerz and Lotsch 2008).

The main aim of the thesis is to test and validate scientifically robust RS methods for crops monitoring in order to support the decision-making process at both farm and policy makers levels.

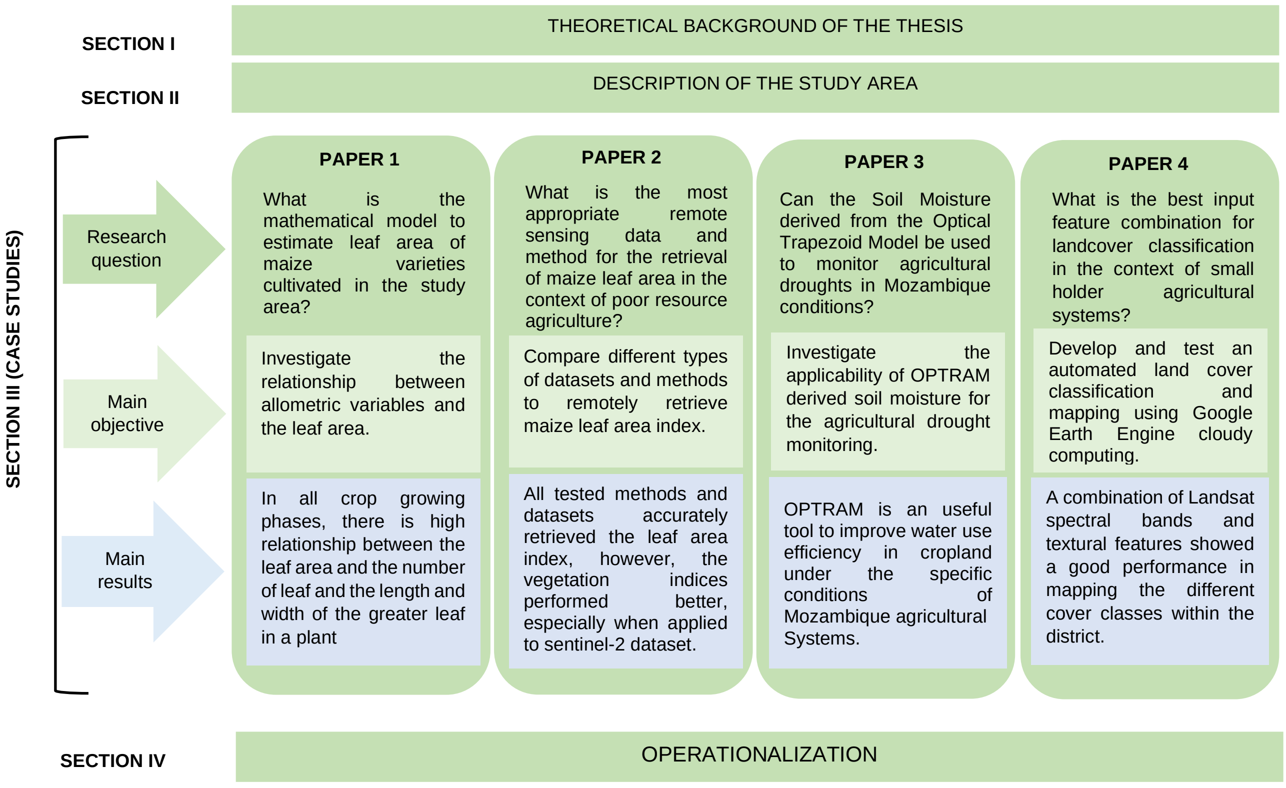
The Figure 1 outlines the structure of the thesis and provides the main content of each section.

In Section I, a conceptual framework supporting the research is presented while the section II briefly provides the biophysical, geographic and socio-economic context of Mozambique and the specific study area. In the section III, four peer-reviewed papers address different but integrated topics of agricultural monitoring and mapping using both in situ and remotely sensed data.

In **paper one,** an allometric based mathematical model for non-destructive estimation of maize leaf area was calibrated and validated. The developed model is suitable for estimating maize leaf area at different growing stages, involving the product of number of leaves per plant, the major leaf length and the major leaf width of the greater leaf. In **paper two,** multispectral satellite (Sentinel-2), in-situ hyperspectral and physically simulated hyperspectral datasets were tested to retrieve maize LAI, by using empirically based (vegetation indices and machine learning algorithms) and physically based (look-up table (LUT) inversion) processing methods. Ground truth LAI was estimated applying the model developed in **paper one**. Although the three methods accurately retrieved the LAI, the vegetation indices slightly outperformed the other two approaches, especially when applied to Sentinel-2 dataset. In **paper three** the emphasis was to remotely estimate SM and subsequently monitor and map agricultural drought. A physically based optical trapezoid model (OPTRAM) was applied to Sentinel-2 data and the ground truth SM data was collected through the gravimetric method and the water mark tensiometers. OPTRAM accurately derived SM at root zone depth, which was further used to successfully monitor agricultural drought at crop field level. The **paper four** focuses on land cover mapping giving emphasis to the evaluation of spatial and temporal dynamics of smallholder agriculture in a time period between 2012 and 2018. The classification was implemented within the Google Earth Engine (GEE) cloud computing applying the Random Forest ensemble classifier and the best suited input variables were tested. Reliable land cover maps were produced for 2012, 2015 and 2018, and agricultural area statistics derived from the maps were accurate when compared against local statistics.

In the section IV, the operational application of the Thesis's results is discussed, emphasizing their integration and relevance within the Agenda 2030 Objectives. Lastly, the section V presents a general discussion, as well as the conclusions and perspectives.

Fig. 1 - Thesis scheme with the organization of the main sections and the structure and scope of the case studies presented as scientific papers.



SECTION I THEORETICAL BACKGROUND OF THE THESIS

SECTION II DESCRIPTION OF THE STUDY AREA

SECTION III (CASE STUDIES)

	PAPER 1	**PAPER 2**	**PAPER 3**	**PAPER 4**
Research question	What is the mathematical model to estimate leaf area of maize varieties cultivated in the study area?	What is the most appropriate remote sensing data and method for the retrieval of maize leaf area in the context of poor resource agriculture?	Can the Soil Moisture derived from the Optical Trapezoid Model be used to monitor agricultural droughts in Mozambique conditions?	What is the best input feature combination for landcover classification in the context of small holder agricultural systems?
Main objective	Investigate the relationship between allometric variables and the leaf area.	Compare different types of datasets and methods to remotely retrieve maize leaf area index.	Investigate the applicability of OPTRAM derived soil moisture for the agricultural drought monitoring.	Develop and test an automated land cover classification and mapping using Google Earth Engine cloudy computing.
Main results	In all crop growing phases, there is high relationship between the leaf area and the number of leaf and the length and width of the greater leaf in a plant	All tested methods and datasets accurately retrieved the leaf area index, however, the vegetation indices performed better, especially when applied to sentinel-2 dataset.	OPTRAM is an useful tool to improve water use efficiency in cropland under the specific conditions of Mozambique agricultural Systems.	A combination of Landsat spectral bands and textural features showed a good performance in mapping the different cover classes within the district.

SECTION IV OPERATIONALIZATION

Section I |Conceptual Framework

1| Remote Sensing Concepts

RS is the science and, to some extent, the art of obtaining information about an object, area, or phenomena through the analysis of data acquired by a device that is not in contact with the object (Jensen 2000, Jones and Vaughan 2010). This is done by sensing and recording reflected or emitted energy and processing, analyzing, and applying that information (Engman and A. Schultz 2000, Basso, Cammarano et al. 2004). As such, the first requirement for RS is a source of energy, in the form of electromagnetic radiation (EMR), to illuminate the target. EMR consists of time-varying electric and magnetic fields that travel in the form of a wave at the speed of light and at a given frequency (Jones and Vaughan 2010).

The wavelength (λ) and frequency are the two EMR proprieties particularly important for understanding RS. The wavelength is the length of one wave cycle and can be measured as the distance between successive wave crests and expressed in meters (m) or some factor of meters such as nanometres (nm), micrometres (µm) or centimetres (cm). Frequency refers to the number of cycles of a wave passing a fixed point per unit of time and is normally expressed in hertz (Hz), equivalent to one cycle per second, and various multiples of hertz (Fonseca and Fernandes 2004, Bakker, Feringa et al. 2009, Jones and Vaughan 2010).

The total range of wavelengths of EMR is called electromagnetic spectrum. This spectrum is an overview of the continuum of the electromagnetic energy which can be subdivided in portions of extremely short wavelengths (cosmic and gamma rays: < 0.30 nm), short wavelengths (ultraviolet – UV: 0.03 – 0.4 µm), intermediate wavelengths (visible – VIS: 0.4 – 0.7 µm), long wavelengths (infrared: near infrared - NIR: 0.7 – 1.3 µm; short wave infrared – SWIR: 1.3 – 2.5 µm and thermal infrared - TIR: 2.5 – 100 µm) to extremely long wavelengths (radio and television waves: 10 cm – 10 km) (Fonseca and Fernandes 2004, Jones and Vaughan 2010). The Figure 2 presents the total range of wavelengths of EMR and highlights the regions of interest for remote sensing.

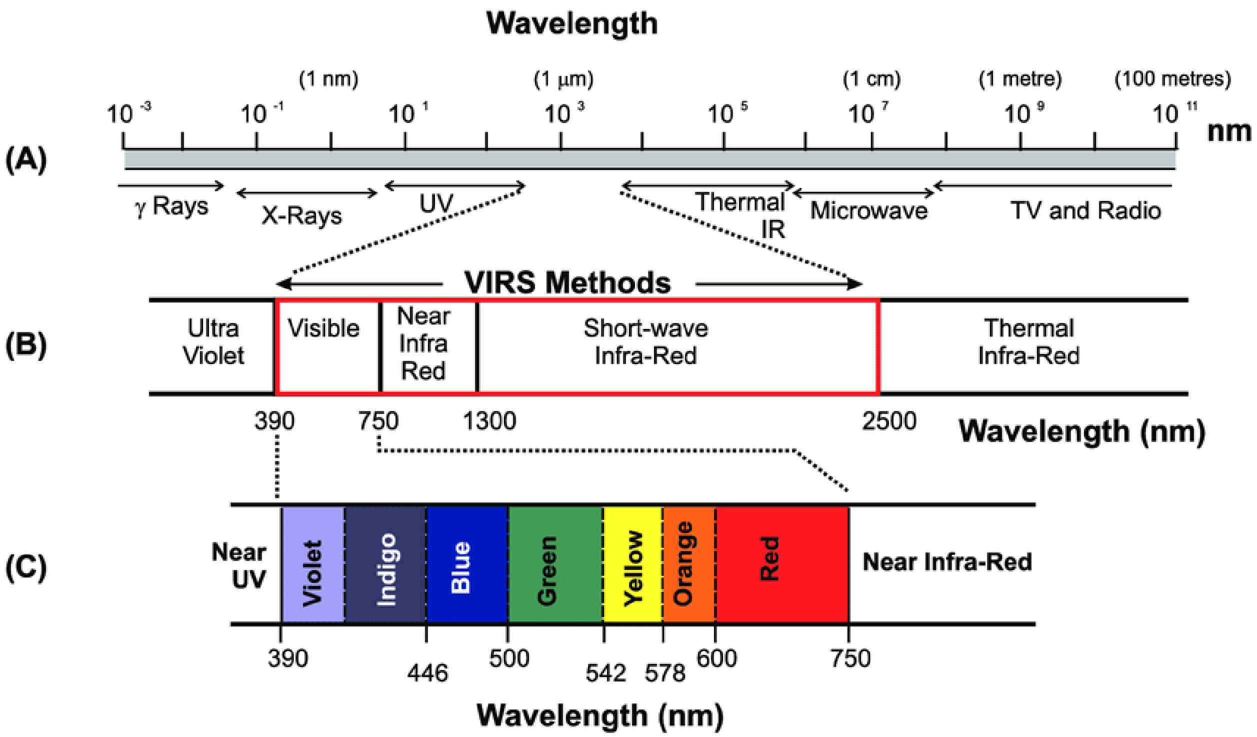


Fig. 2 - Electromagnetic spectrum highlighting the region of interest for remote sensing (Kerr, Rafuse et al. 2011)

Each portion of the electromagnetic spectrum presents different relevance to RS both in type of information that can be gathered and the volume of geospatial data acquisition. Most of the data is obtained in the visible and infrared ranges (Table 1). The UV portion is the region of shortest wavelength with practical use in RS while the microwave region presents the longest wavelength. In the other portions of the electromagnetic spectrum outside this range the radiation is strongly absorbed by the earth's atmosphere and thus not available for RS applications (Fonseca and Fernandes 2004, Jones and Vaughan 2010, CCRS).

The key phenomena associated with electromagnetic radiation interaction with matter include: **Absorption** – describes the magnitude of electromagnetic energy retention by a surface or an object; **Reflection** – the redirection of a beam of radiation when it encounters a surface or object; **Scattering** – the redirection of electromagnetic energy without absorption, when it encounters a surface or object; **Transmission** – when the electromagnetic energy passes through without interaction with a surface particles (Jones and Vaughan 2010).

Table 1 Main electromagnetic regions and indication of their availability for remote sensing applications. Adaptation from (Kerr, Rafuse et al. 2011, CCRS)

Electromagnetic region	Wavelength	Remote sensing applications
Gamma Ray	< 0.03 nanometers	Entirely absorbed by the Earth's atmosphere and not available for remote sensing.
X-ray	0.03 to 30 nanometres	Entirely absorbed by the Earth's atmosphere and not available for remote sensing.
Ultraviolet	0.03 to 0.4 micrometres	Wavelengths from 0.03 to 0.3 micrometres absorbed by ozone in the Earth's atmosphere.
Photographic Ultraviolet	0.3 to 0.4 micrometres	Available for remote sensing the Earth. Can be imaged with photographic film.
Visible	0.35 to 0.75 micrometres	Available for remote sensing the Earth. Can be imaged with photographic film.
Infrared	0.75 to 100 micrometres	Available for remote sensing the Earth. Can be imaged with photographic film.
Reflected Infrared	0.75 to 2.5 micrometres	Available for remote sensing the Earth. Near Infrared 0.75 to 1.3 micrometres and SWIR 1.3 to 2.5 micrometers. Can be imaged with photographic film.
Thermal Infrared	3.0 to 14 micrometres	Available for remote sensing the Earth. This wavelength cannot be captured with photographic film. Instead, mechanical sensors are used to image this wavelength band.
Microwave or Radar	0.1 to 100 centimetres	Longer wavelengths of this band can pass through clouds, fog, and rain. Images using this band can be made with sensors that actively emit microwaves.
Radio	> 100 centimetres	Not normally used for remote sensing the Earth.

The RS data are usually classified following four types of resolutions, namely: spectral, spatial, radiometric and temporal. **Spectral** resolution refers to the ability of a sensor to define fine wavelength intervals and is determined by the number and width of the individual spectral bands that are recorded. A RS dataset is said multispectral when is collected using sensors that detect radiation in a small number (3 to 10) of broad wavelength bands within the visible and infrared region of electromagnetic spectrum (Govender, Chetty et al. 2007, Jones and Vaughan 2010). The hyperspectral data are obtained by sensors that typically collect radiation from hundreds of narrow and contiguous spectral bands typically between 400 nm and 2500 nm. These contiguous, narrow bandwidths data potentially allow for in-depth examination of earth surface features which are difficult to assess when using multispectral data (Govender, Chetty et al. 2007, Darvishzadeh 2008). **Spatial** resolution refers to the size of the ground element represented by an individual pixel. It is dependent on the field of view and the height of the sensor used to capture the data. A high spatial resolution represents elements of small size while low resolution represents elements of large size (Jones and Vaughan 2010). **Radiometric** resolution refers to the number of distinct levels into which the intensity of the incident radiation is differentiated by a sensor (*grey levels*). It is usually expressed as the number of bits used for storing a digital number (DN). The **temporal** resolution refers to the time between successive data acquisition dates, i.e., the more frequent data acquisition, the higher the temporal resolution (Fonseca and Fernandes 2004, Bakker, Feringa et al. 2009, Jones and Vaughan 2010).

Broadly, the RS data can be obtained using three types of sensing platforms, namely, terrestrial, or roving platforms (hand-held or fixed spectroradiometers included), aircrafts and satellites. The **hand-held or fixed spectroradiometers** (proximal RS) are usually used to collect data for small-scales applications. The data from this type of platforms have the advantage that little if any atmospheric correction is needed and can be used for calibration and validation of large-scale data and radiative transfer models (Bakker, Feringa et al. 2009, Jones and Vaughan 2010).

The **aircraft** can be considered a flexible platform as the height, the direction, as well as the timing and frequency of flights can be adjusted for specific proposes. Different types of aircraft can be used, including unmanned aerial vehicles (UAVs), drones, balloons, helicopters, and light planes. However, data obtained from these platforms may suffer from various distortions due to drift, roll, and pitch and also the positioning of the aircraft may be slightly uncertain (Bakker, Feringa et al. 2009, Jones and Vaughan 2010).

The **orbiting satellites** provide very stable platforms that enable regular and periodic measurements to be made from precisely known positions. They are placed in orbits that are designed for the specific proposes of the mission, and to suit the characteristics of the instruments carried. Depending on their altitude and orbit the satellites may be divided in two categories: **Geostationary satellites –** placed at a height of about 36 000 km, above the equator (0^o inclination angle) and orbiting West to East in the same direction and orbital period as the earth rotation. They give continuous coverage over the same area day and night. This type of satellites is mainly used for communication and meteorological applications (e.g. INSAT satellites) (Fonseca and Fernandes 2004, Bakker, Feringa et al. 2009, Jones and Vaughan 2010, CCRS); **Near - polar satellites –** placed at heights ranging from 600 to 2000 km, with inclination angle between 80^o and 100^o and orbiting North – South at a rate such that it covers each area of the world at a constant time of the day, called local sun time, being the reason why many of this type of satellites are also known as sun-synchronous. This characteristic ensures similar illumination conditions while acquiring image over an area (Fonseca and Fernandes 2004, Bakker, Feringa et al. 2009, Jones and Vaughan 2010, CCRS).

The platforms carry **sensors**, which are the devices that gather energy, convert it into a signal and present it in a form suitable for obtaining information about the target object(s). The sensors used for RS can be broadly classified as those operating in the Optical-Infrared and those operating within the Microwave regions and, further subdivided into passive and active, depending on the source of energy. Active sensors use their own source of energy to illuminate the earth surface and a part of it is reflected and received to gather information, while passive sensors receive solar electromagnetic energy reflected from the surface or energy emitted from surface itself (Fonseca and Fernandes 2004, Bakker, Feringa et al. 2009, Jones and Vaughan 2010, CCRS).

2| Remote Sensing Applications focusing agriculture

2.1| Remote Sensing of Vegetation parameters

The use of RS for estimating biophysical and biochemical parameters of vegetation is underpinned by the optical proprieties of vegetation (behavior when interacting with electromagnetic radiation). The vegetation reflectance in the optical domain (400 – 2500 nm) is mainly governed by three factors: (i) the intrinsic optical proprieties of vegetation components, (ii) the arrangement of these components in the vegetation canopy and (iii) the optical proprieties of the soil underneath canopy and non-photosynthetic vegetation components (Dorigo, Zurita-Milla et al. 2007). In addition, the vegetation reflectance is strongly influenced by the viewing and illumination angles as well as the atmospheric conditions (Zhang, Lin et al. 2002, Croft, Anderson et al. 2012).

The spectral properties of vegetation elements are considered the main driving factor of canopy reflectance and influence the profile of the overall spectrum. Solar radiation is either absorbed by chemical components or dispersed by the structural elements such as cell walls (Wang, Skidmore et al. 2015, Féret, Gitelson et al. 2017). Chlorophyll and other pigments are the main absorbers in the visible domain (400 – 700 nm), leaf structure characteristics dominate in the near infrared region while, water, proteins and cellulose play the main role in the shortwave infrared and the canopy architecture mainly accounts for the magnitude and the directional variation in the reflected energy (Homolová, Malenovský et al. 2013, Houborg, Fisher et al. 2015).

The soil reflectance is non-Lambertian (Verbrugghe and Cierniewski 1997) and depends on soil composition (e.g. organic matter content, mineralogy), surface roughness, and soil water content (Baumgardner, Silva et al. 1986, Franceschini, Demattê et al. 2015, Franceschini, Demattê et al. 2018). The influence of the soil background over the total signal recorded by the sensor is wavelength dependent and largest in the near infrared region. Moreover, the presence of understory components (Eriksson, Eklundh et al. 2006) and non-photosynthetic elements (Verrelst, Schaepman et al. 2010), and the varying observation geometries (B. Lobell, Asner et al. 2002) influence canopy reflectance. Therefore, RS methods try to minimize these confounding effects (Knyazikhin, Schull et al. 2013) and enhance the sensitivity of reflectance data towards a trait or a propriety of interest (Haboudane, Miller et al. 2002, Govaerts, Verstraete et al. 2014).

Vegetation parameters comprise, among others, structural (e.g. plant height), physiological (e.g. photosynthesis rate), biochemical (e.g. nitrogen content) or phenological (e.g. leaf phenology) features. The vegetation parameters can be related to ecosystem functioning (Diaz, Hodgson et al. 2004, Wright, Reich et al. 2004, Orwin, Buckland et al. 2010), community ecology (Kraft, Valencia et al. 2008), plant response to environmental pressures (de Bello, Leps et al. 2006, Roumet, Fortunel et al. 2006), ecosystem services (Lavorel, Grigulis et al. 2011), plant biomass and yield (Yin, Dekker et al. 2014), vegetation vigor and stress (Zarco-Tejada, González-Dugo et al. 2013, Zhang, Pattey et al. 2017), plant disease (Mahlein, Oerke et al. 2012), among other. Currently, there are several plant parameters identified and measured by ecologists and agronomists (Kattge, Díaz et al. 2011) at the level of individual plants and further upscaled to canopy properties (Violle, Navas et al. 2007) and the data are often being compiled in various local (Kleyer, Bekker et al. 2008, Paula, Arianoutsou et al. 2009) and global (Kattge, Díaz et al. 2011) databases. Plant scientists agree on a set of key plant parameters, which are relatively easy to measure and have a stable and strong predictive response to ecosystem functions at various scales (Cornelissen, Lavorel et al. 2003, Diaz, Hodgson et al. 2004, Wright, Reich et al. 2004).

Even though data on key plant parameters can be obtained from field measurements, the data collection is onerous in terms of time and resources, it is limited to a certain space and time, and to a limited number of species, making it difficult to be used in support to the agronomic decision making process. Therefore, RS offers potential to complement or even replace field measurements of some plant parameters (Kokaly, Asner et al. 2009) at local and large scales.

The capabilities to retrieve plant parameters and canopy properties from optical RS have grown together with the technological development of RS sensors (Milton, Schaepman et al. 2009). Initial RS sensors providing infrequent data of coarse spatial and spectral resolutions have supported mostly vegetation classification into broad functional groups (Ustin and Gamon 2010) and the development of simple spectral vegetation indices (VIs) that were sensitive to broad variations in canopy properties (Turner, Cohen et al. 1999, Cohen and Goward 2004).

The following cohort of medium spectral and spatial resolution sensors together with the development of radiative transfer models (Liang 2004) have facilitated the quantitative estimation of some plant traits (e.g. chlorophyll content, (Dash and Curran 2004), water content, (Cheng, Zarco-Tejada et al. 2006)) and canopy properties (e.g. leaf area index, (Myneni, Hoffman et al. 2002, Mananze, Pôças et al. 2018)). The recent development of high spectral resolution imaging sensors encouraged even more the quantitative estimation of plant traits related to physiology and biochemistry.

The advantages of using RS are its capability to provide spatially contiguous and – for certain observations – high revisit frequency at the typical length scale of the targeted vegetation parameter. Moreover, it enables a broader sampling scheme to parameter mapping, determined by the combination of pixel size, spatial extent and revisit time of RS observations, than in situ measurements (Homolová, Malenovský et al. 2013).

Based on the above described vegetation optical proprieties, and other factors that influence the canopy reflectance, several algorithms have been developed to retrieve biophysical and biochemical variables from reflective RS data. They are commonly classified as statistical and physical approaches.

2.1.1| Statistical approach

In the statistical approach, the relationship between the spectral signature of vegetation and the biophysical or biochemical proprieties of interest is examined. In order to establish such relationships, biophysical / biochemical measurements need to be taken under varying field or laboratory conditions and for different plant species and/or phenological development stages. In fact, the validity and portability of the developed relationships greatly depends on the accuracy of the measurements and the range of conditions considered for the calibration of the relationships (Sims and Gamon 2002, Dorigo, Zurita-Milla et al. 2007). Several ways of relating spectral information to the targeted variable have been developed and applied including parametric and non-parametric methods:

(i) Parametric methods

Typically, in this method, a spectral band arithmetic formulation is defined as a spectral vegetation index (VI) and subsequently linked to the variable of interest based on a chosen fitting function (Thenkabail, Smith et al. 2000, Atzberger 2004, Pôças, Rodrigues et al. 2015, Verrelst, Rivera et al. 2015). The VI is a combination of a limited number of bands that enhance spectral features sensitive to a specific vegetation property while reducing undesired effects of background objects or other sources of disturbances (Dorigo, Zurita-Milla et al. 2007, Glenn 2008). According to (le Maire, François et al. 2004), the band combination can take different formulations, including: (i) single reflectance or difference between reflectances at two wavelengths; (ii) simple ratio of reflectances (SR); (iii) normalized difference ratios of reflectances (ND) and (iv) derivative of reflectance signatures. These VIs have been traditionally used based on two bands but more recently, additional formulations involving three bands were proposed (Wang, Yao et al. 2012). The possible bands, VI formulation, and fitting function can be based on a systematic assessment of all possible band combinations according to all possible VI formulations (Rivera, Verrelst et al. 2014, Verrelst, Rivera et al. 2015).

According to the measured vegetation properties, VIs can be divided into three main categories: structural, biochemical, and physiological. The measured structural properties include leaf area index (LAI) (Main, Cho et al. 2011, Mananze, Pôças et al. 2018), fractional cover (Barati, Rayegani et al. 2011, Ding, Zheng et al. 2016); the biochemical properties include water and pigments (chlorophyll, carotenoids, anthocyanins) (Stagakis, Markos et al. 2010, Alchanatis and Cohen 2011); and the physiological proprieties include stress-induced changes (Cong 2010, Panigada, Rossini et al. 2014, Gago, Douthe et al. 2015).

Many structurally oriented VIs rely on some combination of near-infrared (NIR) to red reflectance, such as the NIR to Red ratio, or simple ratio (SR) index. This is because increases in LAI correspond to increases in the chlorophyll absorption and NIR-scattering and to decreases in exposed substrate, resulting in decreasing Red reflectance and increasing NIR reflectance. Thus, equations for these VIs often compare reflectance at an absorbing wavelength to a non-absorbing wavelength (Roberts, Roth et al. 2011).

The biochemical and physiological indices rely on a similar comparison of absorbing and non-absorbing wavelengths, with varying absorbing wavelength according to the biochemical/physiological feature. For example, the canopy moisture/moisture stress indices include wavelengths associated with liquid water absorption (e.g., 970 and 1200 nm), while ligno-cellulose content indices utilize wavelengths from 1500 to 1800 and 2000 to 2350 nm, within the short-wave-infrared (SWIR) domain. By contrast, pigments (carotenoids, anthocyanins, and chlorophylls) absorb in the visible and ultraviolet domains, with distinct, but overlapping, absorption features. Chlorophyll primarily absorbs blue and red light. Anthocyanins absorb all but red light; and many carotenoids are yellow due to strong blue light absorption (Roberts, Roth et al. 2011).

An imperative issue when using spectral VIs for the estimation of biophysical parameters is the need of an efficient validated relationship between the VI and the variables(s) of interest. This implies a costly and time intensive field data collection program if relationships want to be valid for a wide range of species, canopy conditions, and view/sun constellations (Dorigo, Zurita-Milla et al. 2007). Hence, although this approach has the advantage of straightforwardness and often yields good results it has important drawbacks namely: (i) cannot be easily generalized and requires field data collection parallel to RS observations for calibration and validation purposes, and (ii) not all relevant spectral information is used for modelling, which potentially affects the estimation of variables with nonlinear relationship with spectral data (Weiss, Baret et al. 2000, Baret and Buis 2008, Atzberger, Darvishzadeh et al. 2015, Berger, Atzberger et al. 2018).

(ii) Non-parametric methods

The non-parametric methods normally consider all available spectral data and do not assume any type of relationship between the variable of interest (predicted) and the predictors. These characteristics make non-parametric methods potentially more convenient for the retrieval of complex and nonlinear properties of vegetation (Clevers, Kooistra et al. 2010, Kuhn and Johnson 2013, Doktor, Lausch et al. 2014, Pôças, Gonçalves et al. 2017). Numerous non-parametric regression algorithms are available in the statistics and machine learning literature, and they are being increasingly applied for vegetation biophysical parameters retrieval. They are calibrated to predict a variable of interest using a training dataset of input-output data pairs, derived from synchronized measurements of the parameter and the corresponding reflectance observations (Hastie, Tibshirani et al. 2009, Verrelst, Muñoz et al. 2012, Caicedo, Verrelst et al. 2014, Verrelst, Rivera et al. 2015, Pôças, Gonçalves et al. 2017). Regression Trees (Breiman, Friedman et al. 1984), Partial least square regression (Wold, Martens et al. 1983), Random forest (Breiman 2001), Relevance vector machine (Tipping 2001) and Neural network (Ripley 1996) are some of the statistical and machine learning regression algorithms that are frequently applied for the retrieval of biophysical variables from spectral data (Atzberger, Guérif et al. 2010, Clevers, Kooistra et al. 2010, Yao, Huang et al. 2015, Pôças, Gonçalves et al. 2017).

2.1.2| Physically based approach

In the physical approach, Radiative Transfer Models (RTM's) are used to simulate leaf and canopy reflectance which are afterward inverted to estimate leaf or canopy biophysical or biochemical variables. The RTM's simulate the interactions between radiation and the canopy components based in physical laws. The most reliable canopy reflectance modelling approaches involve the combination of a leaf optical model with a canopy and a soil reflectance model to calculate the top-of-canopy reflectance (Dorigo, Zurita-Milla et al. 2007, Darvishzadeh, Skidmore et al. 2008, Darvishzadeh, Matkan et al. 2012).

Several approaches have been developed and successfully applied to describe leaf scattering and absorption. These models can be distinguished by the complexity level that is taken into account and the underlying physics and the most commonly used are: (i) N-flux model (Fukshansky, Fukshansky-Kazarinova et al. 1991, Richter and Fukshansky 1996), which describes the leaf as a simple block of diffusing and absorbing material; (ii) Plate model (Fourty, Baret et al. 1996), which describes the leaf as a pile of N plates separated by N – 1 air spaces and assumes Lambertian scattering and leaf absorption components (biochemicals) homogeneously dispersed throughout the leaf, (iii) Stochastic and radiative transfer models that simulate the optical properties of the leaf by using a Markov chain or by directly using a radiative transfer equation and (iv) Ray tracing model (Govaerts, Jacquemoud et al. 1996), which requires the optical proprieties of the leaf material and a detailed description of the internal structure of the leaf. This model is generally based on Monte Carlo simulations and is often considered the most realistic model given that accounts for the complexity of internal leaf structure.

The N-flux models present high simplicity, but it is very difficult to link the overall leaf absorption coefficient with the specific absorption coefficients and the concentrations of the plant chemicals for inversion proposes (Fukshansky, Fukshansky-Kazarinova et al. 1991). The plate models incorporate the leaf chemicals as inputs so, they are suitable for inversion (Fourty, Baret et al. 1996). The RTM and mainly the ray tracing models are characterized by high complexity, making the computational implementation very intensive and model inversions numerically difficult (Govaerts, Jacquemoud et al. 1996).

For the canopy reflectance simulations, there are several RTM's which are normally classified according to their approach and level of complexity. According to (Goel 1988, Berger, Atzberger et al. 2018), there are four categories of canopy reflectance models:

(i) Geometrical models (Chen and Leblanc 1997): the canopy is described as a ground surface with opaque geometrical objects of known spatial distribution and optical properties. It is especially suited for forests or sparse canopies where the impact of multiple scattering and shading is negligible;

(ii) Turbid medium models (Verhoef 1984): assume the canopy as a turbid medium composed by small and randomly distributed absorbing and scattering elements. These models are better suited for dense canopies with the precondition that the vegetation elements are small compared to canopy height (Berger, Atzberger et al. 2018);

(iii) Hybrid models, representing a combination of (i) and (ii) (Goel and Grier 1988, Gastellu-Etchegorry, Demarez et al. 1996, Huemmrich 2001, Schlerf and Atzberger 2006): they are complex but versatile models and can be used to characterize canopies that are neither dense nor sparse;

(iv) Monte-Carlo ray tracing models (North 1996): describe the radiation regime in vegetated canopies more closely to reality. Obviously, these models are the most complex and computationally intensive.

Soil reflectance is a crucial element in modeling canopy reflectance, as it corresponds to the lower boundary condition and has its own spectral properties. The importance of this element is higher when sparse or low vegetation density canopies are to be simulated. The soil reflectance is typically measured in the field, extracted from satellite images, or simulated using soil reflectance models (Jacquemoud, Bacour et al. 2000, Atzberger, Jarmer et al. 2003, Atzberger 2004, Hapke 2008).

Inverting a canopy reflectance model consists in finding the set of input parameters that leads to the best match between the bi-directional reflectance simulated by a canopy reflectance model and the reflectance measured by the sensor (Combal, Baret et al. 2003). The inverse problem consists, therefore, in estimating the set of variables that leads to the observed spectrum. Different methods have been developed to solve this problem:

(i) the iterative optimization technique (Jacquemoud, Bacour et al. 2000, Casa, Baret et al. 2010): iteratively runs the canopy reflectance model with different sets of input variables, searching through a cost function, for the best fit between the simulated and the measured reflectance (Dorigo, Zurita-Milla et al. 2007). This method presents three main drawbacks: it requires an initial guess of the solution (Qiu, Gao et al. 1998); the accuracy of the solution depends on the merit function and is computationally too intensive to be suited for operational use or for the inversion of complex three dimensional canopy reflectance models (Kimes, Knyazikhin et al. 2000);

(ii) the lookup table (LUT) approach (Kimes, Knyazikhin et al. 2000, Weiss, Baret et al. 2000, Dorigo, Richter et al. 2009): the canopy reflectance model is used in a direct way to simulate a large number of possible spectra, depending on different combinations of input parameter values. Subsequently, a cost function searches in the LUT for the spectrum that has the highest similarity to the measured one. The set of variables used to simulate this spectrum is the final solution of the inversion. The LUT approach has the advantage that it is considerably faster than the iterative optimization technique, although time can significantly increase if large LUTs are used. The technique can also deal with more complex canopy reflectance models, such as ray tracing models, since the model is used only in the direct mode (Kimes, Knyazikhin et al. 2000, Combal, Baret et al. 2003, Verrelst, Rivera et al. 2015, Sehgal, Chakraborty et al. 2016). However, the inversion results may be seriously compromised due to poor choice of the distribution nature of the variables and the step size between successive values of the variables in the LUT may lead to an over or under-representation of certain variable ranges and thus to a non-representative result (Combal, Baret et al. 2003);

(iii) inversion based on machine learning algorithms (MLA) such as artificial neural network (ANN) (Atzberger 2004, Verger, Baret et al. 2011, Verrelst, Rivera et al. 2015): is seen as the most prominent hybrid approach that combines physical and statistical models. Like in the LUT approach, a canopy reflectance model is used in the direct mode to build a large synthetic data set that represents a wide variety of canopy conditions. While the LUT approach searches for the simulated spectrum closest to the measured one, the ANN minimizes the distance between canopy variables (Bacour, Baret et al. 2006, Martínez, Camacho et al. 2013). ANNs are fast after the training stage and less sensitive to model uncertainties than the iterative optimization and LUT approach (Combal, Baret et al. 2003, Dorigo, Zurita-Milla et al. 2007). When compared to LUT inversion, the MLA are typically more computationally efficient and thus they can be quickly and automatically applied to large data sets, whilst providing comparable retrieval accuracies. The main problem of the MLA is their ‘black box’ nature, the need to specify and tune the network architecture, and their tendency to perform unpredictably when inputs strongly deviate from their training data (Verrelst, Muñoz et al. 2012, Verrelst, Rivera et al. 2015).

A major and general drawback of the inversion of physically based canopy reflectance models is the ill-posedness problem. The solution is not necessarily unique because of compensation between several variables that could affect canopy reflectance in the same way (Fourty, Baret et al. 1996, Combal, Baret et al. 2003, Berger, Atzberger et al. 2018). This problem can be minimized incorporating prior information (Knyazikhin, Glassy et al. 1999), knowing the temporal dynamic of the biophysical variables and the typical distribution of the input variables (Koetz, Baret et al. 2005).

The figure 3 summarizes the empirically and the physically based approaches for biophysical parameters retrieval.

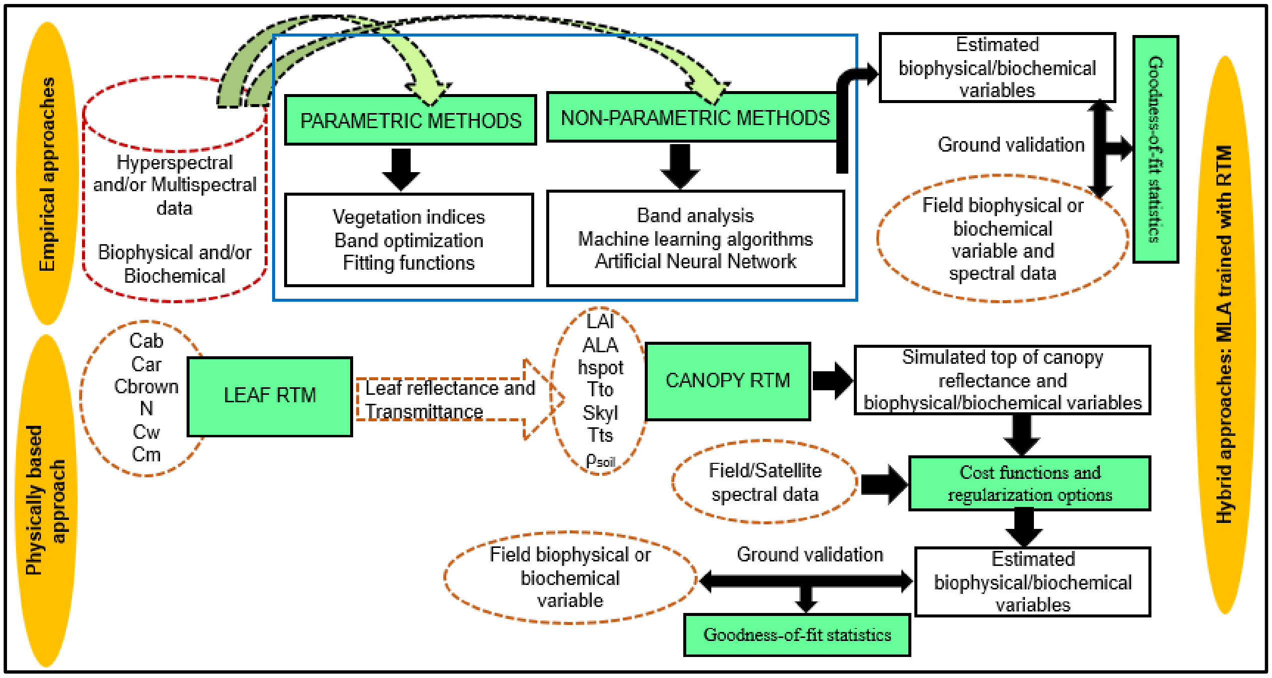


Fig. 3 – scheme summarizing the empirical and physically based approaches

2.2| Remote Sensing of surface soil moisture

Several regions of the electromagnetic spectrum have been applied for soil moisture (SM) retrieval. In fact, various approaches have been tested to remotely estimate surface SM, most of them relying on the fact that soils get darker when wet and thus becoming less reflective, while others linked to latent and sensible heat fluxes and thus exploring the fact that wet soils present higher thermal inertia than dry soils (Kerr 2007, Petropoulos, Ireland et al. 2015). Also, methodologies based on microwave data have been considered due to the sensitivity of this spectrum domain to variations in SM at soil surface layers (Paloscia, Pettinato et al. 2013). The next sections present the main concepts and examples of optical and microwave sensing used in surface soil moisture estimation and a summary of the main advantages and disadvantages of those methods is provided in the table 2.

2.2.1| Optical Sensing of surface soil moisture

Optical RS methods usually apply the Visible, Near Infrared (NIR) and Shortwave Infrared (SWIR) data for SM modeling. Studies exploring the relationship between SM and reflectance revealed that reflectance decreases with the SM increase (Lobell and Asner 2002, Weidong, Baret et al. 2002, Gao, Xu et al. 2013, Nocita, Stevens et al. 2013). The SWIR reflectance is sensitive to leaf liquid water content and has been shown to be more effective than the visible and NIR for monitoring SM changes at content less than or equal to 50% of the volumetric water content (Lobell and Asner 2002). Most of these optical studies have been based on empirical VIs depicting vegetation spectral properties (e.g., growth, biomass amount, biotic and abiotic stress) and the level of vegetation moisture stress, thereby allowing indirect estimates of SM (Jackson, Idso et al. 1981, Xiao, Chen et al. 1994, Ghulam, Qin et al. 2007, Zhang, Yang et al. 2015).

Most indices are based on the triangular or trapezoidal spaces of pixels distribution relative to optical observations in different electromagnetic frequency bands. For example, the Triangle Soil Moisture Index, TSMI (Amani, Parsian et al. 2016), is derived from the red − NIR triangular space. However, the relationship between satellite-derived VIs and SM is complicated because of the time lag of vegetation response to soil moisture. Also, as the content and distribution of soil characteristics affecting the soil spectral behavior vary greatly from site to site, these empirically based approaches are site specific, meaning that their application is very limited to the conditions to which they have been calibrated (Huan-Jun, Yuan-Zhi et al. 2009, Petropoulos, Ireland et al. 2015).

To overcome the empiricism of optical indices, a physically-based trapezoid model, Optical Trapezoid Model (OPTRAM), was recently developed (Sadeghi, Babaeian et al. 2017). The OPTRAM is grounded on a physical relationship between the surface or root-zone SM in bare or vegetated soils and the shortwave infrared transformed reflectance (STR) (Sadeghi, Jones et al. 2015). The OPTRAM is based on the interpretation of pixels distribution within the STR-NDVI (Normalized difference vegetation index) space. By solely relying on optical reflectance data and thus only depending on surface properties, the STR-NDVI space is nearly invariant over time, which allows a generalized parameterization of the model over the time for a specific location (Sadeghi, Babaeian et al. 2017). The OPTRAM was successfully applied to retrieve SM in regions with diversified soil types and vegetation conditions (Sadeghi, Babaeian et al. 2017, Babaeian, Sadeghi et al. 2018, Mananze, Pôças et al. 2019).

2.2.2| Thermal infrared sensing of surface soil moisture

The principle behind the SM retrieval through thermal infrared (TIR) region of the electromagnetic spectrum is that land surface temperature is sensitive to surface soil water content due to its influence on surface heating process (heat capacity and thermal conductivity) under bare soil or sparse vegetation cover conditions (Sohrabinia, Rack et al. 2012, Zhao and Li 2013).

Thermal RS approaches that have been used for vegetation studies and SM estimation include the thermal inertia and temperature index methods (Amani, Parsian et al. 2016).These methods are established using the relationship between surface emissivity, temperature and SM, by mainly taking advantage of the water circulation and energy balance principle (Sohrabinia, Rack et al. 2012). Thermal inertia essentially describes a physical property that characterizes the surface resistance to ambient temperature change (Soliman, Heck et al. 2013). Though it is quite effective in bare soils, it demonstrates poor correlation for highly vegetated surfaces (Zhan, Qin et al. 2007). As a result, SM estimation methods based on thermal inertia seem difficult to apply to large-scale SM monitoring. Yet, the use of a new generation of portable thermal infrared sensors that can be installed in airborne platforms provide an opportunity to retrieve land surface temperature with a very high spatial resolution at smaller scales (Petropoulos, Ireland et al. 2015, Amani, Parsian et al. 2016).

The land surface temperature (LST) can be used for assessing physical and biological processes occurring in the land surface, including for monitoring SM. The LST is related with soil surface temperature under conditions of bare soil, while for densely vegetated areas it represents vegetation canopy temperature that can be used as an indicator of vegetation water stress (when canopy temperature increases); in both cases, the LST can be directly or indirectly related with SM, respectively (Zhang and Zhou 2016).

Additionally, there are methodologies coupling visible and TIR data for estimating SM. These methodologies rely on the spatial distribution of the LST-VI feature space. The LST-VI feature space is delimited by the dry and wet edges. The wet edge in the feature space denotes maximum evapotranspiration and suitable SM while the dry edge illustrates situations of vegetation under water stress in which evapotranspiration is minimum and no water is available to be used (SM $\simeq$ 0) (Zhang and Zhou 2016). Different interpretations of the surface temperature in the fully vegetated area gave rise two different methods: the triangular and the trapezoidal methods, whose forms are represented in the figure 4. The triangular method main assumptions include that: (i) the complete underlying vegetation cover from bare soil to full vegetation is known; (ii) the variations in the space are not caused by differences in atmospheric conditions but by variations in the available SM and (iii) the sensitivity of LST to the canopy and soil is different. The triangular method can estimate SM using only remote sensing data, is easy to operate and can monitor drought conditions over large areas. However, it presents many limitations in non-flat regions, requires high number of pixels to cover wide range of soil wetness and vegetation covers, the estimates accuracy is very low with high spatial resolution data, and the space determination is to some extent empirical (Price 1990, Zhang, Tang et al. 2015).

The trapezoidal method, in the other hand, assumes that the surface temperatures of fully vegetated surfaces are much higher than those of the wettest surface. This method does not require high number of pixels; it has a robust physical basis and the space determination from four vertices is close to the true conditions of the land surface (Fig. 4). Nevertheless, the method requires more ground-based parameters to calculate the soil moisture index and the influence of water stress on vegetation cover have a time lag that induces uncertainties for SM estimations (Hazaymeh and Hassan , Moran, Clarke et al. 1994, Zhang and Zhou 2016).

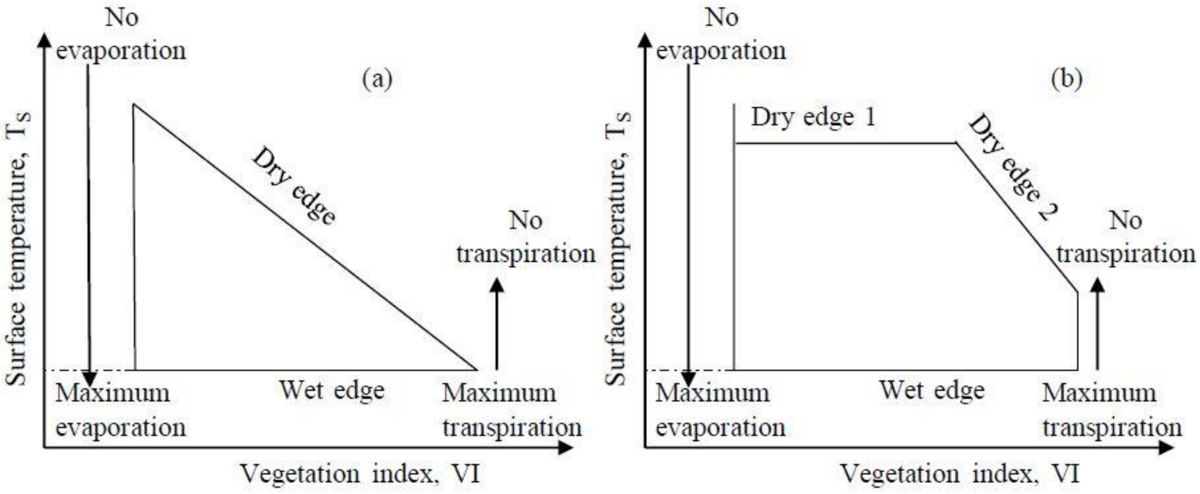


Fig. 4 – Triangular (a) and trapezoidal (b) forms of soil moisture estimation based on the relationship between the surface temperature and vegetation index (Hazaymeh and Hassan).

2.2.3| Microwave sensing of surface soil moisture

The theoretical basis for measuring SM by microwave (MW) techniques is based on the large contrast between the dielectric properties of water and dry soil. Microwave RS has the capability of acquiring data under almost any meteorological conditions and without an external source of illumination.

2.2.3.1| Passive microwave sensing of surface soil moisture

Passive sensors (the so-called microwave radiometers) use very sensitive detectors to measure the naturally emitted intensity of MW emission (at wavelengths k = 1 – 30 cm) from the Earth's surface, expressed as brightness temperature (T_B). The T_B of the surface is related to its emissivity, physical temperature, and contributions from the intervening atmosphere. The interpretation and use of passive MW signatures is made complex by the influence of other surface variables, such as soil roughness and vegetation properties (Petropoulos, Ireland et al. 2015, Amani, Parsian et al. 2016). It is known that the sensitivity to variations of SM is at a maximum for bare and smooth soil, while it is reduced by an increase of soil roughness and/or vegetation biomass. Soil texture and variability in the temperature of both soil and vegetation also affect microwave retrieval (Guerriero, Ferrazzoli et al. 2012).

The L-band passive microwave is the most applied channel to monitor SM (Jackson, Le Vine et al. 1995, Lannoy, Reichle et al. 2013). In fact, two space missions use this technology at the global scale: The Soil Moisture and Ocean Salinity (SMOS, launched by the end of 2009), and the Soil Moisture Active Passive (SMAP, launched in 2014). The SMOS has multi-angular capabilities, which are exploited by the SM retrieval approach: SM and vegetation optical depth are retrieved simultaneously based on SMOS multi-configuration observations, in terms of polarizations and incidence angles (Champagne, Rowlandson et al. 2016, Fascetti, Pierdicca et al. 2016). The SMAP incorporates a radar and a radiometer, both operating at L-band and at the incidence (observation) angle of 40°. The mission concept is to combine the complementary attributes of the radar observations (high spatial resolution but lower SM accuracy) and radiometer observations (higher SM accuracy, but coarse spatial resolution) to retrieve SM at a spatial resolution of 9 km (Nocita, Stevens et al. 2013, Petropoulos, Ireland et al. 2015).

2.2.3.2| Active microwave sensing of surface soil moisture

Active MW instruments supply their own source of illumination and subsequently determine the difference in the energy power between the transmitted and received electromagnetic radiation, often referred to as the backscatter coefficient. Active MW sensors can be divided into imaging (e.g. Synthetic Aperture Radar – SAR) and non-imaging sensors (altimeters and scatterometers).

As in the case of passive MW sensing, the magnitude of the SAR backscatter coefficient is related to surface SM through the contrast of the dielectric constants of bare soil and water. Also, the main perturbing factors affecting the accuracy of SM retrieval include soil surface roughness and vegetation biomass (Moran, Peters-Lidard et al. 2004). However, in contrast to passive MW, active MW sensing corrects for perturbing effects, and the connection between the backscattering coefficient and the surface reflectivity is more complex (Jackson 2002). Thus, several studies have proposed different methods, algorithms, and models relating satellite-based images of SAR backscatter to surface SM retrieval (e.g. (Ahmed, Zhang et al. 2011, Baghdadi, Aubert et al. 2012)). Nevertheless, radar signals relative to low frequency bands (X-band, with wavelength of 3 cm) are not able to penetrate vegetation cover due to the way the dielectric permittivity of the biomass affects radar response, while active MW retrieval utilizing higher frequencies (L-band, with wavelength 23.5 cm) is more applicable (e.g. (Jagdhuber, Hajnsek et al. 2013, Vereecken, Huisman et al. 2014)).

Additionally, with the launch of Sentinel-1 from Copernicus program, various studies have been testing SAR C-band (with wavelength of ~5.6 cm) for the retrieval of SM. Several algorithms and models have been tested for such purpose, including the water cloud model (Li and Wang 2018) and the artificial neural network (ANN) approach (Paloscia, Pettinato et al. 2013).

2.2.4| Synergistic sensing of surface soil moisture

An alternative group of methods employed for deriving surface SM content from EO datasets includes those based on the fusion of spectral information acquired from sensors operating at different parts of the EMR. Those approaches have been developed to primarily address the complementarity and interchangeability between the different types of RS data (Petropoulos, Ireland et al. 2015, Amani, Parsian et al. 2016), and most of such synergistic methods examine the synergy between: (i) Optical and thermal data (Sandholt, Rasmussen et al. 2002, Vicente-Serrano, Pons-Fernández et al. 2004, Petropoulos, Carlson et al. 2009, Holzman, Rivas et al. 2014, Zhang, Tang et al. 2014); (ii) Active and passive microwave data (Bindlish and Barros 2000, Lee and Anagnostou 2004, Liu, Judge et al. 2013, Panciera, Walker et al. 2014, Rüdiger, Doubková et al. 2014) and (iii) Microwave (mainly SAR) and optical/TIR data (Kurucu, Sanli et al. 2009, Temimi, Leconte et al. 2010, Piles, Vall-llossera et al. 2012, Srivastava, Han et al. 2013, Piles, Sánchez et al. 2014, Sánchez-Ruiz, Piles et al. 2014).

Table 2 Summary of remote sensing data and methods used in surface soil moisture estimation. Adapted from (Petropoulos, Ireland et al. 2015).

Sensor type	Methods	Advantages	Disadvantages	Examples
Optical	Reflectance based methods	Good spatial resolution, multiple satellites available, hyperspectral sensors promising, based on mature technology	Weak relationship to SM when the vegetation cover is high, constrained by cloud conditions and nighttime, thin surface layer only	(Zhao and Li 2013), Anne, Abd-Elrahman et al. (2014) Liu, Zhang et al. (2009) Gao, Xu et al. (2013)
	Thermal Infrared-based	Good spatial resolution, multiple satellites available, methods relating SM to thermal inertia are promising	Weak relationship to SM when the vegetation cover is high, constrained by cloud conditions and nighttime, meteorological conditions and surface topography	Minacapilli, Agnese et al. (2009), Qin, Yang et al. (2013), Soliman, Heck et al. (2013), Lei, Bian et al. (2014)
Microwave passive	Various methods proposed	Promising results in SM estimation particularly over bare soil surfaces, not constrained by clouds and/or daytime conditions, high temporal resolution	Coarse spatial resolution, constrained by vegetation cover and surface roughness	Kong and Dorling (2008), de Jeu, Holmes et al. (2014)
Microwave active	Various methods (empirical, semi-empirical, physically - based)	Fine spatial resolution, not constrained by clouds and/or daytime conditions	SM accuracy influenced by surface roughness & vegetation cover amount, coarse temporal resolution	Callens, Verhoest et al. (2006), Al-Bakri, Suleiman et al. (2014), Panciera, Tanase et al. (2014), Panciera, Walker et al. (2014)
Synergistic methods	Optical & Thermal Infrared	High spatial resolution, a range of satellite sensors to choose from, simple & straightforward implementation, based on mature technology	Most methods are of empirical nature (transferability difficult), limited to cloud-free & daytime conditions, poor temporal resolution, low penetration depth	(Wang, Qu et al. 2007, Petropoulos and Carlson 2011, Holzman, Rivas et al. 2014, Zhang, Tang et al. 2014)
	Active & passive MW	Link the high spatial resolution of active MW systems with lower sensitivity to SM with passive MW frequencies, improved temporal resolution & SM retrieval	SM scaling & validation needs caution, different SM measurement depths	Lee and Anagnostou (2004), Narayan, Lakshmi et al. (2006), Liu, Dorigo et al. (2012)
	MW & optical	Vegetation and surface roughness effects can be minimized, promising in SM retrievals	SM scaling & validation needs caution, different SM measurement depths	Zhao and Li (2013)

2.3| Remote Sensing of agricultural drought

There are numerous conceptual and operational drought definitions proposed according to different disciplinary perspectives (Heim Jr 2002). Fundamentally, drought is a temporary water supply deficit relative to some long-term average condition. (Dracup, Lee et al. 1980) and (Wilhite and Glantz 1985) proposed a drought typology based on four distinct types, namely meteorological, agricultural, hydrological and socio-economic.

Agricultural drought occurs when the SM doesn’t support average crop yield or average grass production on range lands (Wilhite and Glantz 1985) and manifests itself through reduction in average crop yield or death of the plant (Llano, Vargas et al. 2012, Martínez-Fernández, González-Zamora et al. 2015, Vyas, Bhattacharya et al. 2015). In very severe drought situations soil loss may occur, which may render agricultural production unviable.

Several RS based drought indices, typically based on a combination of precipitation, temperature and modeled SM, have been derived to assess the effects of agricultural droughts and define its parameters, which include intensity, duration, severity and spatial extent (Vyas, Bhattacharya et al. 2015). These include the Palmer drought index (PDI) (Palmer 1965), the Crop Moisture Index (CMI) (Palmer 1968), the soil moisture deficit index (SMDI), and the evapotranspiration deficit index (ETDI) (Narasimhan and Srinivasan 2005).

Nevertheless, there have been attempts to contemplate site-specific soil proprieties (SM, field capacity and permanent wilting point) given their core role in drought indicators and classification, e.g., soil moisture index, SMI (Sridhar, Hubbard et al. 2008); soil water deficit index, SWDI (Martínez-Fernández, González-Zamora et al. 2015); water deficit index, WDI (Martínez-Fernández, Ceballos et al. 2005); the concept of static vegetation water stress, (Porporato, Laio et al. 2001); the method for using long-term measurements of the soil water deficit to compute site-specific drought probability (Torres, Lollato et al. 2013); soil water deficit, SWD (Torres-Ruiz, Diaz-Espejo et al. 2013); plant available water, PAW (McPherson, Fiebrich et al. 2007); drought severity index, DSI (Cammalleri, Micale et al. 2016), among others.

The SWDI has been recently used in conjunction with remotely sensed SM estimates from the optical region through the OPTRAM (Babaeian, Sadeghi et al. 2018, Mananze, Pôças et al. 2019), as well as from the MW region through SMAP and SMOS (Martínez-Fernández, González-Zamora et al. 2016, Mishra, Vu et al. 2017), for monitoring agricultural drought, Currently, the application of MW-based (e.g., SMAP, SMOS, Sentinel-1) SM retrievals for long-term monitoring of agricultural drought is still limited due to the lack of data for long periods of time. Contrarily, OPTRAM can be used for agricultural drought by using long-term archives of MODIS and Landsat missions, allowing for assessing trends and interannual variability at various spatial scales.

2.4| Remote Sensing for agricultural landscapes mapping

Image classification for land cover land use (LCLU) mapping has gained relevance in the RS research community since its results are the basis for many environmental and socioeconomic applications (Lu and Weng 2007, Tran, Julian et al. 2014, Wu, Zucca et al. 2016). In the agricultural perspective, the most interesting applications of image classification include classifying agricultural landscapes (Pôças, Cunha et al. 2011, Duro, Franklin et al. 2012), determining land use areas (Huang, Fipps et al. 2011), land cover mapping and monitoring (Oetter, Cohen et al. 2001, Azzari and Lobell 2017, Defourny, Bontemps et al. 2019) , tracking land cover change (Xu, Yu et al. 2018), yield mapping and estimating crop production (Lambert, Traoré et al. 2018, Jin, Azzari et al. 2019) , soil moisture estimation (Ahmad, Kalra et al. 2010), revealing crop residues in agricultural fields in liaison to soil erosion evaluation (Daughtry, Doraiswamy et al. 2006), estimating the effect of crop classification errors (Stehman and Milliken 2007).

Efforts have been undertaken in order to develop advanced classification approaches and techniques for improving classification accuracy (Stuckens, Coppin et al. 2000, Franklin and Wulder 2002, Pal and Mather 2003, Gallego 2004, Soe Thwal, Ishikawa et al. 2019). Nevertheless, classifying remotely sensed data into a thematic map remains a challenge because many factors, such as the complexity of the landscape (either temporal or spatial components), selected input data, and the classification algorithms and approaches, may affect the success of the classification (Lu and Weng 2007). The major steps of image classification may include determination of a suitable classification system, selection of training samples, image preprocessing, features extraction, selection of suitable classification approaches, post-classification processing, and accuracy assessment. The user’s needs, scale of the study area, areal extent, data costs, and analyst’s skills are important factors influencing the selection of remotely sensed data, the design of the classification procedure, and the quality of the classification results (Lu and Weng 2007).

Numerous techniques for RS image classification have been described in the literature (Dean and Smith 2003, Ziyu, Wenxia et al. 2004, Mountrakis, Im et al. 2011, Vatsavai, Bright et al. 2011, Liao, Bellens et al. 2012, Gonçalves, Pôças et al. 2019). However, the image classification is broadly divided into two approaches, namely: pixel-based and object-based image analysis (e.g. Duro, Franklin et al. 2012).

2.4.1 Pixel-based classification

In pixel-based classification, individual image pixels are analyzed by the spectral information that they contain (Richards and Jia 2006, Jones and Vaughan 2010). In this classification approach, each pixel in the image is assigned into one of a range of possible land cover types based on its characteristic spectral reflectance, thought other information can also be used. The resulting signature comprises the contributions of all materials present in the training pixels but ignores the impact of the mixed pixels (Lu and Weng 2007, Jones and Vaughan 2010, Duro, Franklin et al. 2012).

Per-pixel classification algorithms can be parametric or non-parametric. The parametric classifiers assume that a normally distributed dataset exists, and that the statistical parameters (e.g. mean vector and covariance matrix) generated from the training samples are representative. However, the two assumptions are often violated, particularly in heterogeneous landscapes and it is difficult to integrate spectral with ancillary data (Lu and Weng 2007). Furthermore, when using parametric classifiers, the information from surrounding pixels, which may improve the accuracy, is not used. Therefore, a class that displays high spectral heterogeneity may have its pixels labelled as different classes. This is more likely to happen for example, in cropping field exhibiting high within-field spectral variability due to variations in soil fertility, soil moisture conditions, pests and diseases, or erratic farm practices (Peña-Barragán, Ngugi et al. 2011, Forkuor, Conrad et al. 2014). Such misclassifications are more pronounced when parametric classifier is applied to highly heterogeneous landscapes like in smallholder agricultural dominated regions (Xiao, Boles et al. 2002, Tran, Julian et al. 2014). The maximum likelihood is the widely used parametric classifier in practice, thanks to its robustness and its availability in almost any image-processing software (Lu and Weng 2007).

The non-parametric classifiers, also known as machine learning classifiers, do not rely on the assumption of normal distribution of the dataset. No statistical parameters are needed to separate image classes. These methods are data driven and they learn the relationship between predictor and response data (e.g. Breiman 2001). Using enough size data set and parameters, the machine learning methods can automatically extract rules and restrictions to find the best model for the new data using decision rules created from input data.

Non-parametric classifiers are thus especially suitable for the incorporation of non-spectral data into a classification procedure. Previous research has indicated that non-parametric classifiers provide better classification results than parametric classifiers in heterogeneous landscapes (Paola and Schowengerdt 1995, Foody 2002, Ok, Akar et al. 2012). The most widely used learning based algorithms for image classification are Random forest, Bagging, Boosting, Decision tree, Artificial neural network, Supported vector machine and K-nearest neighbor (Ok, Akar et al. 2012).

2.4.2 Object-based image analysis

Object-based image analysis (OBIA) is a classification approach often preferred when using high resolution datasets, allowing overcoming some of the limitations of the pixel-based approaches (Räsänen and Virtanen 2019). The OBIA uses image objects instead of pixels for image classification. It allows incorporating different types of input data, including spectral, textural and contextual information to identify thematic classes in an image.

The first step in OBIA is to segment the image into homogeneous objects consisting of contiguous cluster of pixels grouped together in a meaningful way through pre-defined criteria of similarity (Blaschke, Burnett et al. 2004). Each object is composed by multiple pixels that are internally more homogeneous and similar than the pixels from adjacent objects (Gonçalves, Pôças et al. 2019).

There exist many segmentation algorithms including, for example, fuzzy c-means (Bhaskaran, Paramananda et al. 2010), level set (Ball and Bruce 2005), watershed transformation (Dobrin, Viero et al. 1994), mean-shift (Ozdarici Ok and Akyurek 2012) and region-merging (Lassalle, Inglada et al. 2015, Tavares, Beltrão et al. 2019).

In a second step, each object (segment) is classified based on one or more statistical properties of the contained pixels. This means that all pixels within a segment are assigned to one class, eliminating the within-field spectral variability and mixed pixels problems associated with pixel-based approaches. Several studies have confirmed the superiority of OBIA over parametric pixel-based classifications, especially in heterogeneous agricultural landscapes and urban areas (Myint 2006, Blaschke 2010, Peña-Barragán, Ngugi et al. 2011). However, a potential shortcoming of object-based classification is caused by the problem of under-segmentation and over-segmentation. Segmentation of remotely sensed images is a difficult task due to mixed pixels, spectral similarity of many land-cover types (Zerrouki and Bouchaffra 2014).

2.4.3 Feature extraction and selection

Regardless of the classification approach, a crucial task in the classification of spatially heterogeneous landscapes relies on the definition of the kind of spatial resolution and data input combinations for accurately mapping land cover (Räsänen and Virtanen 2019). Inclusion of diversified types of input variables, for example, spectral bands, vegetation indices, transformed images, textural or contextual information, multitemporal images, multi-sensor images, and ancillary data may be useful to improve the classification accuracy (Lu and Weng 2007, Rodriguez-Galiano, Chica-Olmo et al. 2012). However, increasing number of input features may introduce additional complexity related to the increase of computational time and the “curse of dimensionality”. This high dimensionality in the dataset can overcome the expected increase in classification accuracy associated with the inclusion of additional features (Guyon and Elisseeff 2003).

Hence, dimensionality reduction needs to be undertaken in order to reduce irrelevant and redundant input variables. Dimensionality reduction techniques can be categorized mainly into feature extraction and feature selection. In the feature extraction approach, features are projected into a new space with lower dimensionality. A number of different feature extraction techniques have been developed, including: principal component analysis, minimum noise fraction transform, discriminant analysis, decision boundary feature extraction, non-parametric weighted feature extraction, wavelet transform, and spectral mixture analysis (Myint 2001, Okin, Roberts et al. 2001, Rashed, Weeks et al. 2001, Asner and Heidebrecht 2002, Lobell, Asner et al. 2002, Neville, Lévesque et al. 2003, Pal and Mather 2003, Platt and Goetz 2004).

On the other hand, the feature selection approach aims to select a small subset of features that minimize redundancy and maximize the relevance to the target, i.e., it selects features that are capable of better discriminating samples from different classes (Kwak and Choi 2002, Guyon and Elisseeff 2003). Optimal selection of input features for classification has been extensively discussed in previous literature, and can be performed through: Graphic analysis (e.g. bar graph spectral plots, co-spectral mean vector plots, two dimensional feature space plot, and ellipse plots) and statistical methods (e.g. average divergence, transformed divergence, Bhattacharyya distance, Jeffreys– Matusita distance) (Swain and King 1973, Swain and Davis 1978, Mausel, Kramber et al. 1990, Jensen 1996, Mishra and Crews 2014).

Feature selection have been regarded as superior in terms of better readability and interpretability since it preserves the original feature values in the reduced space while feature extraction transforms the data from the original space into a new space with lower dimension, which cannot be linked to the features in the original space (Kwak and Choi 2002, Guyon and Elisseeff 2003).

Section II| Country and study area description

1| Geography of Mozambique

Mozambique occupies a total area of 799,380 km^2, stretching for 2,470 km along Africa's southeast coast, bordering Tanzania to the north; Malawi, Zambia, and Zimbabwe to the west; and South Africa and Swaziland to the south (Fig. 4). The country is characterized by a low-lying plateau broken up by 25 sizable rivers that flow into the Indian Ocean. The population of Mozambique is estimated at 26 million inhabitants and is expected to increase up to 40 million people until 2060. Also, there is a growth of urban population, with estimates pointing to a percentage of 45% in the year 2020 (MINAG 2011). Mozambique has been the fastest growing economy in Sub-Saharan Africa in the last 25 years (8% per year), with the agricultural sector being the second largest contributor (6%) to the Gross Domestic Product (GDP) growth rate (Ackermann, Allgood et al. 2012).

The country administrative organization setting includes 10 provinces grouped into three regions: South (Maputo, Gaza and Inhambane), Center (Sofala, Zambezia, Manica and Tete) and North (Nampula, Cabo Delgado and Niassa) (Fig. 5).

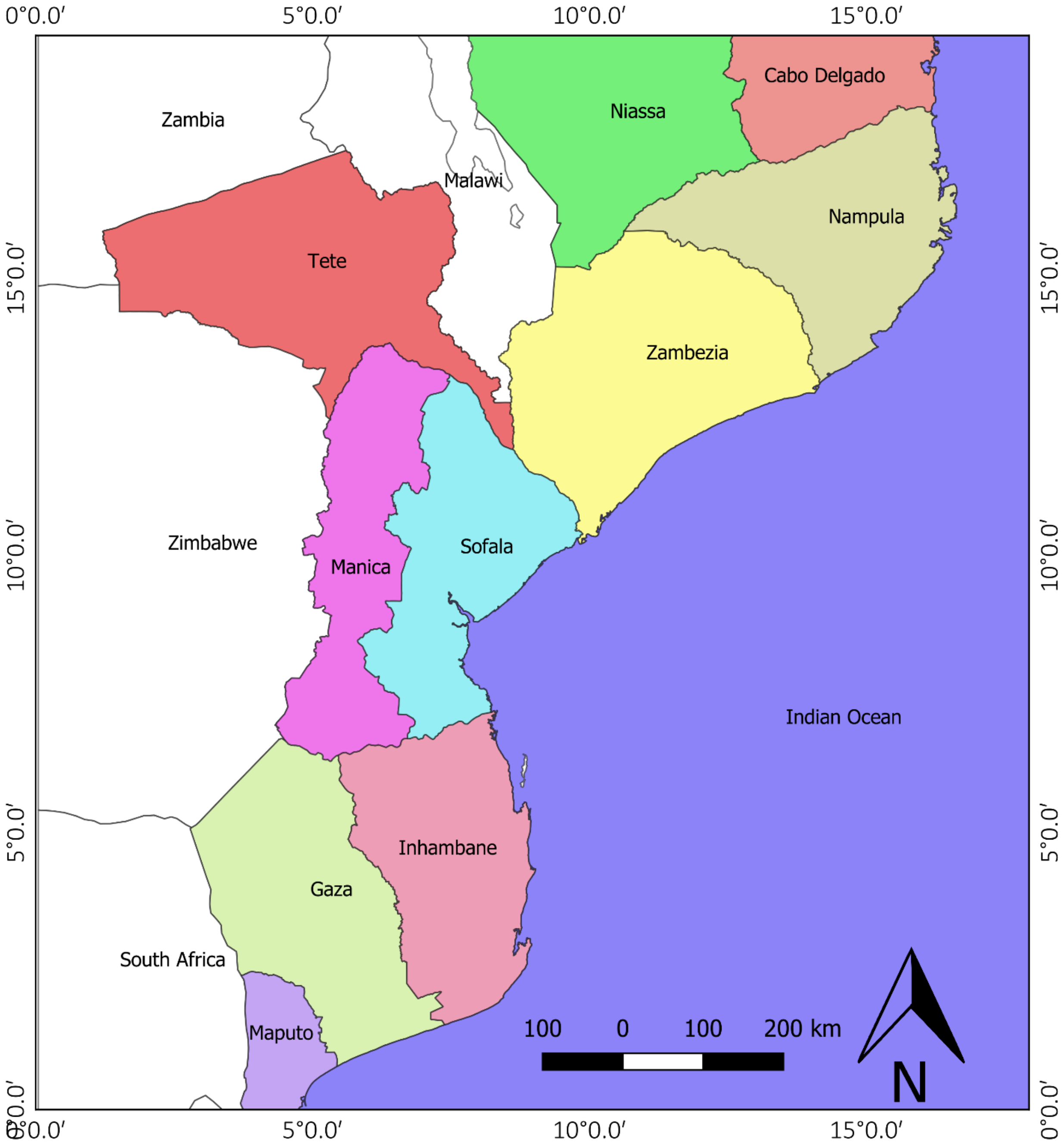


Fig. 5 - Map of Mozambique including the provinces and boundering countries

2| Climate of Mozambique

According to the *Koppen Climate Classification* (Köppen 1900), Mozambique climate varies from mostly tropical (Aw) and subtropical (Cwa and Cwb) conditions in the north and central parts to dry semiarid (BSh) and dry arid desert (Bwh) in the south (Fig. 6). The hottest regions are in the Zambezi basin, the coastline of Cabo Delgado, Nampula, Zambezia and Sofala. The south is the coolest part of the country, with an average maximum and minimum temperature of 30 ºC and 19 ºC respectively. The annual average precipitation for the whole country is 1032 mm and the rainy season lasts from October to April. The precipitation varies widely, decreasing from the coast to the inland areas and from north to south. Average rainfall ranges from 800 to 1000 mm along the coast, with values above 1200 mm occuring between Sofala and Zambezia. The rainfall decreases inland reaching 400 mm at the border with South Africa and Zimbabwe. The north and central part of the country has annual rainfall from 1000 to over 2000 mm because of the northeast monsoon and high mountains. In the southern inland part of the country, the average annual rainfall ranges from 500 to 600 mm. The evapotranspiration varies between 800 and above 1600 mm. Along the coast it varies between 1200 and 1500 mm. Maximum values of above 1600 mm occur in the eastern and middle Zambezi basin. Values around 800 mm occur in central Niassa and on the border with Zimbabwe (FAO 2016).

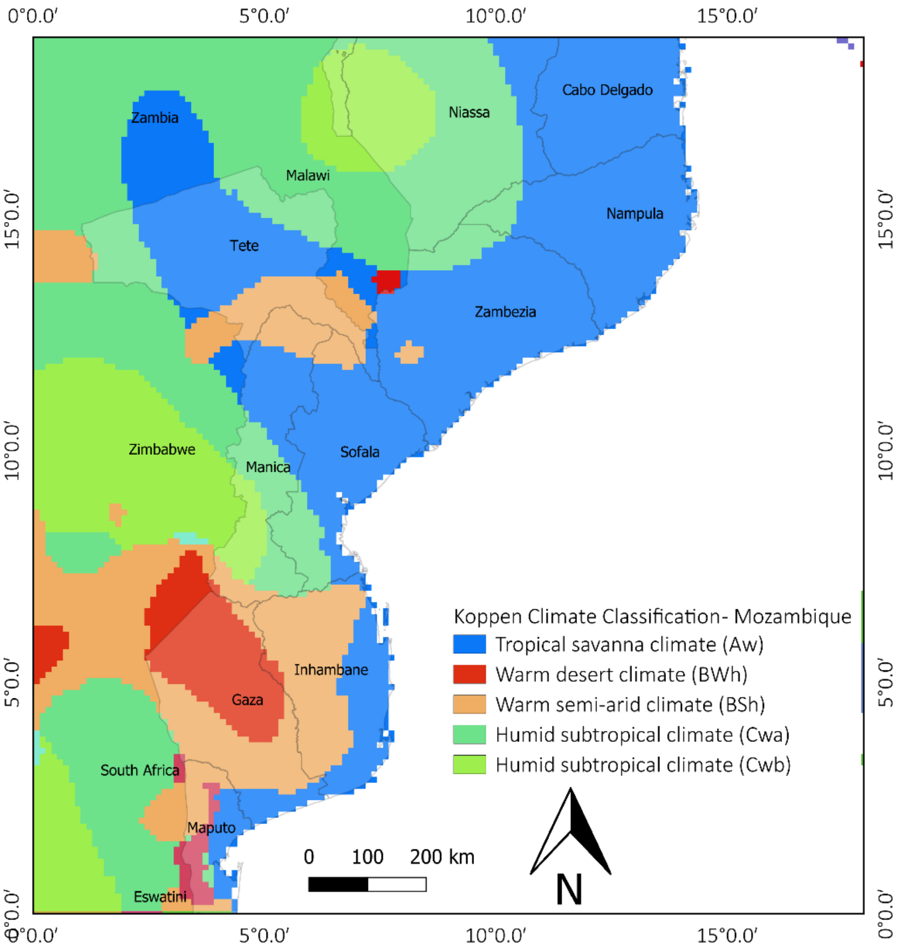


Fig. 6 - Mozambique map of Koppen Climate Classification (FAO 2016).

3| Physiography of Mozambique

Lowlands dominate the southern provinces, narrowing to a mere coastal plain north of the cleft where the Zambezi River cuts through the country’s midsection. The Zambezi valley, the lower section of which is a part of the Eastern (Great) Rift Valley, is Mozambique’s most remarkable geographic feature. Throughout the country the land rises gently from east to west. In the centre and north it slopes steadily into the high plains and ultimately to the mountainous regions on the northwest border with Malawi and Zambia. Four of Mozambique’s five highland regions straddle the west and northwest border areas: the Chimoio Plateau on the border with Zimbabwe, the Marávia highlands bordering Zambia, and the Angónia highlands and Lichinga Plateau, which lie, respectively, west and east of Malawi’s protrusion into Mozambique. Mount Binga, the country’s highest elevation at 2,436 meters, is part of the Chimoio highlands (Dos Muchangos 1999, Sheldon and Penvenne 2019).

4| Soils of Mozambique

There is a high variety of soil types (Figure 7) in Mozambique but generally, the northern and west-central regions are dominated by the Africa’s ancient basement complex of granite rock, while the southern and east-central regions present sedimentary soils. Hence, the northern and central provinces have generally more fertile, water-retention soils than does of the south, where sandy, infertile soils prevail. The quality of soils in the northern and central regions permit agricultural potential to extend outside the river valleys, whereas in the southern region, fertility is largely limited to alluvial soils in the valleys of the Save, Limpopo, Incomáti, Umbelúzi, and Maputo rivers, although several pockets of fertile but heavy soil occur southwest of Inhambane province (Dos Muchangos 1999, Sheldon and Penvenne 2019).

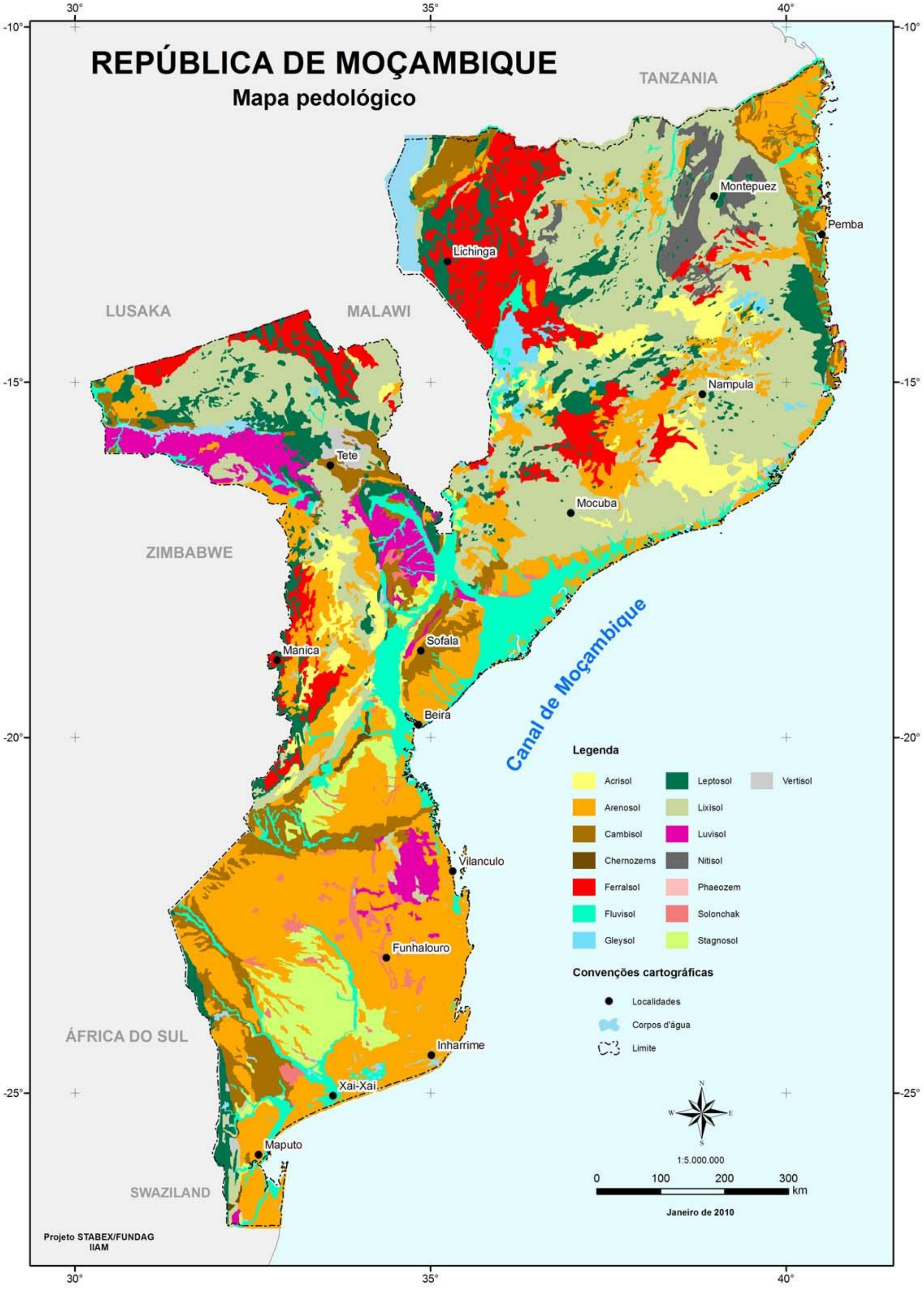


Fig. 7 - Main soil types in Mozambique (Walker, Pitoro et al. 2006)

5| Vegetation of Mozambique

Five phyto-geographical zones are recognized in Mozambique: (i) Swahilian regional centre of endemism (ii) Swahilian-Maputaland regional Transitional zones, (iii) Maputaland-Tongoland regional mosaics, (iv) Zambezian Regional Centre of Endemism and (v) Afromontane centre of endemism (White 1983, Van Wyk and Smith 2001). Within these zones around 5,500 species of plants (including macroalgae) are found, organized in eight communities with miombo, mopane, undifferentiated woodlands and coastal mosaic being the most common (Figure 8). The miombo forests dominate the northern regions while the mopane and undifferentiated woodlands the southern. The highlands are covered with grasses, as well as temperate evergreen forests where rainfall is high. On the coast, swamps are home to mangroves which are part of the Southern Zanzibar-Inhambane Coastal Forest Mosaic ecoregion, which extends from Tanzania to Mozambique.

Two centers of endemism were identified in Mozambique, being the Maputaland coastal forest mosaic ecosystem in the extreme south and the Chimanimani-Nyanga in the central region (White 1983, MICOA. 2009). These ecosystems are home of several commercial timber and non-timber species and about 10% of plant species are used in traditional medicine. About 300 species are included in the IUCN red list of which 120 are threatened. The threats include deforestation for fuelwood production, itinerant agriculture, increasing human settlements and urbanization and uncontrolled fires (MICOA. 2009, Sitoe, Salomão et al. 2012).

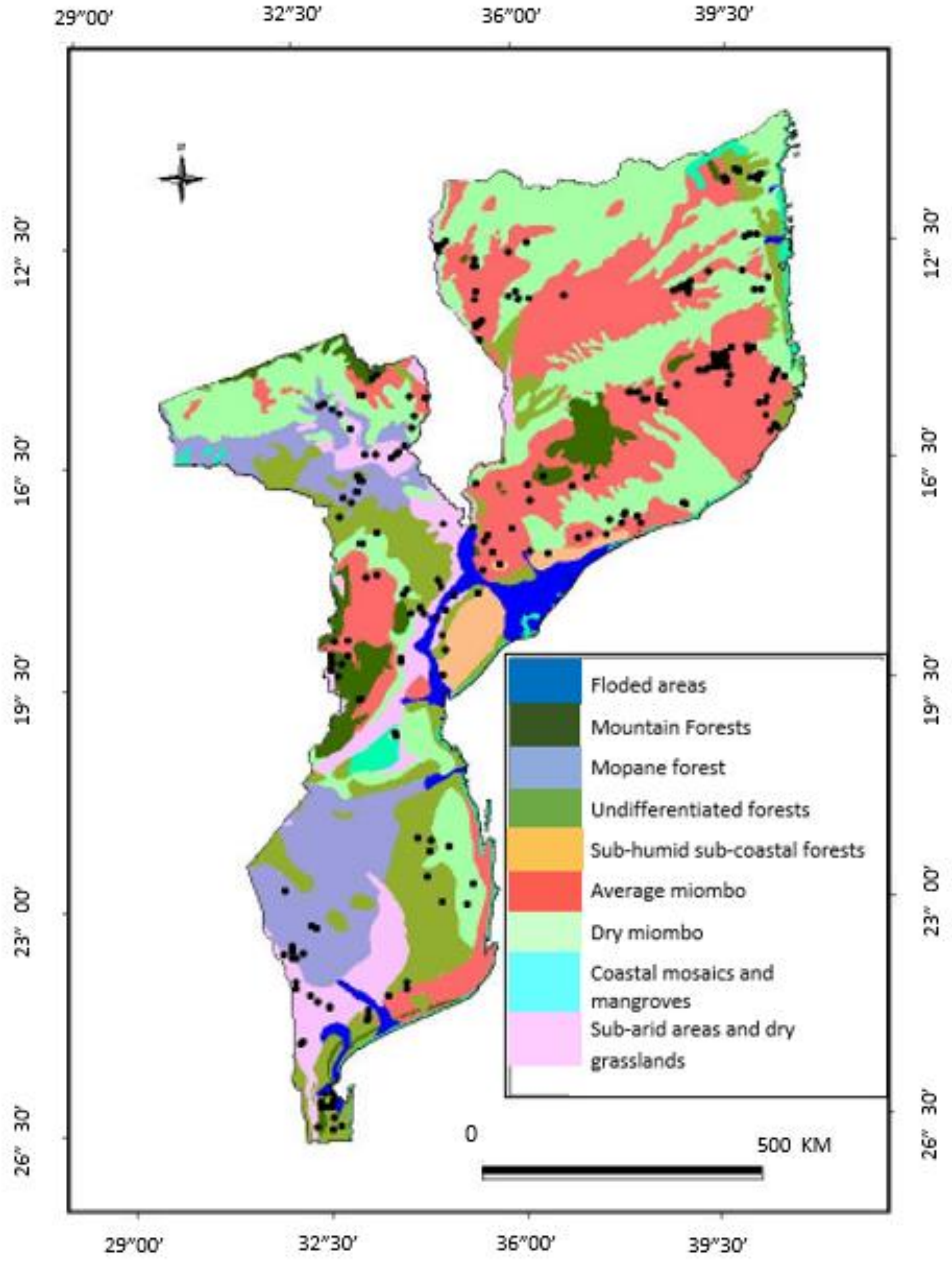


Fig. 8 - Main natural vegetation types in Mozambique (MICOA 2008).

5| Study area description

The study was conducted in the district of Vilankulo, province of Inhambane. The district covers about 5,867 km^2 and has about 193,895 inhabitants. Agriculture, fisheries and tourism are the most important subsistence and economic activities. Lowlands dominate the entire district area and the soils are sandy along the coastal belt, hydromorphic in the river's vicinities, while in the interior loam-sandy soils are predominant. Coastal mosaics, Miombo forests and rangelands are the main natural vegetation types which are human-modified by land use types comprising rain fed and irrigated agriculture.

Figure 9 presents the monthly average precipitation and temperature in the region for the reference period 1934 – 1995. In Vilankulo, the mean annual temperature is 23.0°C. The warmest month, on average, is February with an average temperature of 25.5°C, while the coolest month on average is July, with an average temperature of 19.5°C. The average amount of precipitation for the year in Vilankulo is 125.3 mm. The month with the most precipitation on average is February with 186.6 mm of precipitation, being August, with an average of 23.3 mm, the month with the least precipitation. There is an average of 68 days of precipitation per year, with the most precipitation occurring between January and March and the least precipitation occurring between August and October (Fig. 9).

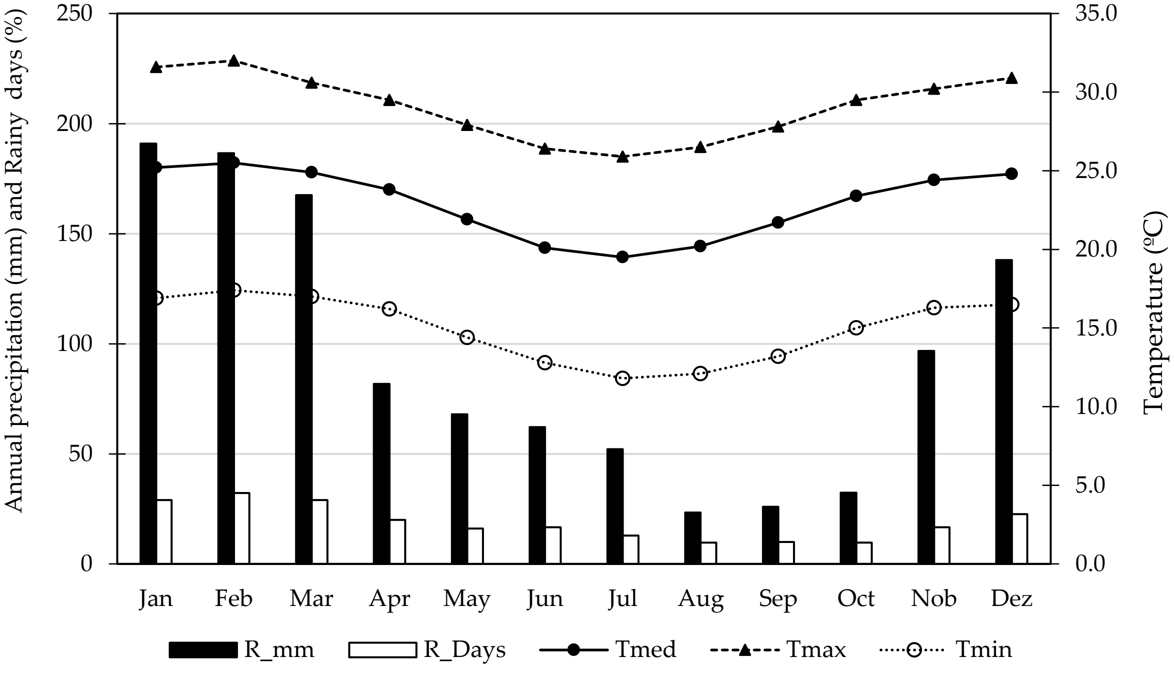


Fig. 9 - Average monthly precipitation (R; mm), Rainy days per month (%) and temperature (Temp.; ºC) temperature in Vilankulo for the reference period (1934 – 1995).

According to Köppen's classification (Köppen 1900), Vilankulo belongs to the climate group Aw - Tropical savannah (Fig. 6). Tropical savanna climates have monthly mean temperature above 18°C in every month of the year and typically a pronounced dry season, with the driest month having precipitation below 60 mm.

According to the Thornthwaite's rational classification of climate, Vilankulo belongs to the climate group C2 B'4 r a', corresponding to Sub humid rainy (Humidity index: 6.17; C2), Mesothermal (Index of thermal efficiency: 1129.0 mm; B'4), with little or no water deficiency (aridity index: 9.53; r) and low variation of temperature throughout the year or low thermal seasonality (summer concentration of thermal efficiency 33.6%; a').

Figure 10 presents the soil water balance for Vilankulo based on (Thornthwaite and Mather 1955), which assumes a 100 mm soil water-holding capacity, highlighting the water deficit and withdrawal from April to November.

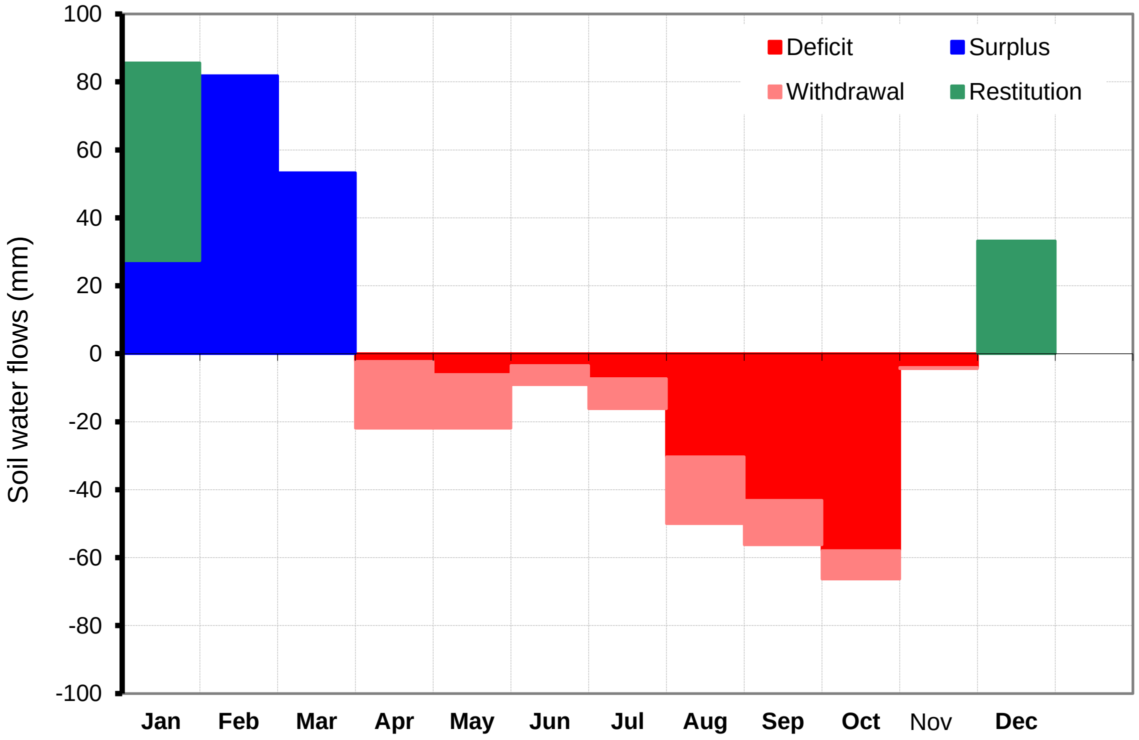


Fig. 10 - Annual distribution of the main components of soil water balance in Vilankulo based on 100 mm soil water-holding capacity (1991 - 2018).

6| Main context of Agricultural sector in Mozambique

The Mozambican Constitution defines agriculture as the base for the national economy development (República 2004) This definition is to some extent motivated by the existing potentialities for agricultural development, including but not limited to: (i) extensive land surface, associated to ten different agro-ecological regions (Figure 11) which are prone to cultivate a large diversity of crop types; (ii) more than 33 million hectares of land classified as good for agriculture, of which only 10 % are currently under use; (ii) more than 3.3 million hectares located in areas with available water for irrigation; (iv) increasing demand for food products as the urban population also tend to increase. Moreover, Mozambique's government signed the African Unions Comprehensive African Agriculture Development Program (CAADP) agreement in late-2010, thereby committing itself to allocating at least 10% of the total budget to agriculture (Benfica, Cunguara et al. 2019).

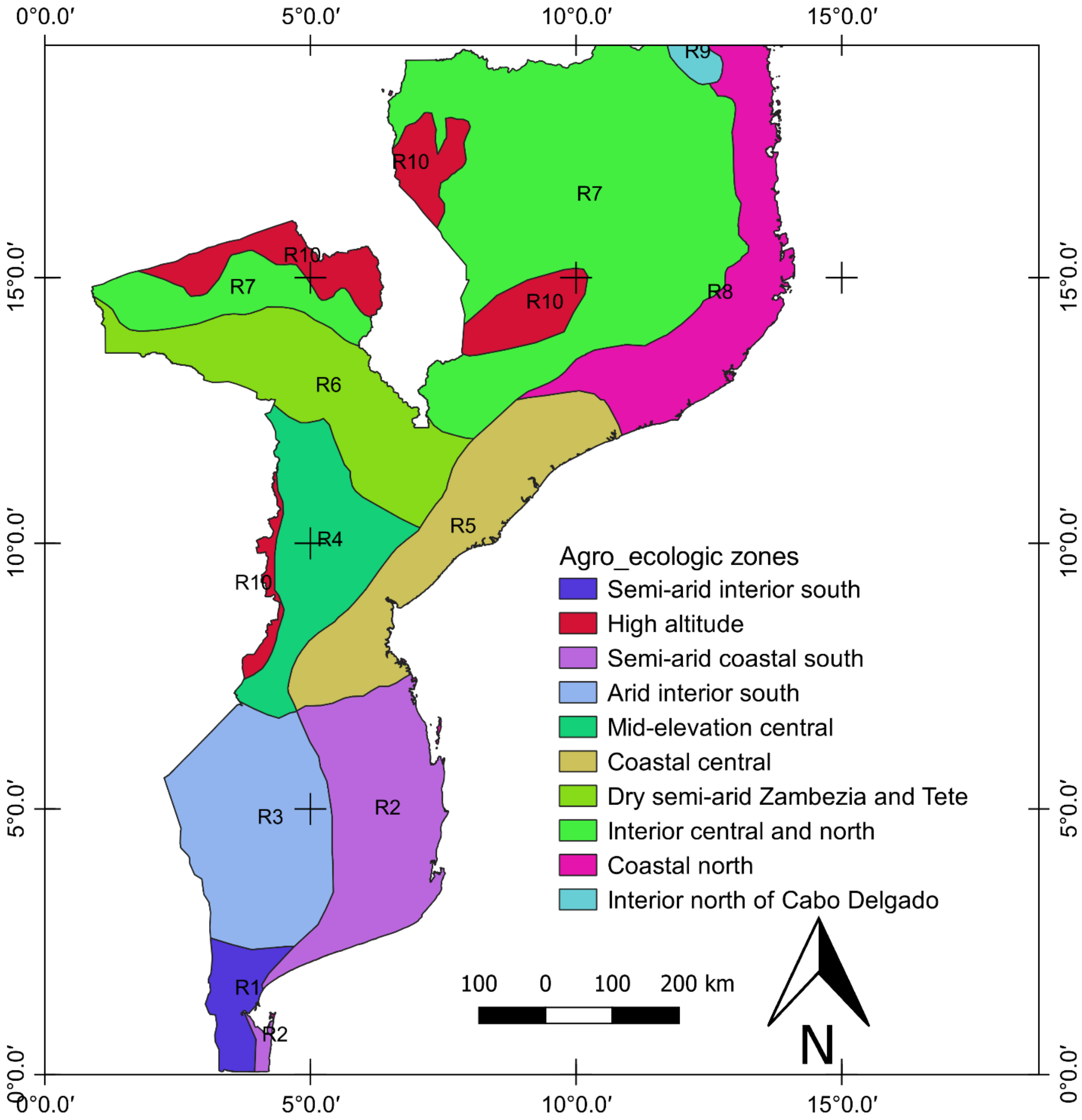


Fig. 11 - Main agri-ecological regions in Mozambique (Walker, Pitoro et al. 2006)

More than 70% of the population in Mozambique relies on agriculture for their livelihoods. Traditional small-holder farming systems dominate the agricultural landscape (Figure 12), growing mainly maize (*Zea mays* L.) or cassava (*Manihot esculenta*) for own consumption. Small-holder farmers, a quarter of them being women, constitute the socio-economic backbone of the country and provide the basis for food security, preserving natural biodiversity and providing environmental protection. These smallholdings farming systems are mostly rain-fed and based on manual and domestic labor. The use of chemical inputs, animal traction, farm machinery, and more sophisticated tools is minimal. Any inputs that are used are mostly acquired through the informal economy (Rutherfood 2010, INE 2011, Achicala, Lampeão et al. 2016, INE 2018). In this context, the crops productivity per hectare is very low.

Figure 12. Prevailing cropland types in the study area. (1) Deforested area for agriculture; 2. Cropland (peanuts and maize); 3. Cropland (maize and banana) and 4. Fallow land.

Over 80% of the total area of cultivated land is used to produce staple food crops. Maize, cassava and cowpeas are the staples produced by most holdings, together comprising 60% of the total cultivated land. Only 5% of the land is cultivated for horticultural crops and 6% for cash crops (sugar cane, cotton, tea, oilseeds, tobacco) (Achicala, Lampeão et al. 2016).

The social division of labour in agriculture involves the whole family. There is usually little use of labour from outside the household. Women are the basis of agricultural production; they are responsible for land preparation, digging, weeding and harvesting. They help to transport, store and market surplus production (Achicala, Lampeão et al. 2016).

As a result of the civil war, the main agricultural infrastructures (e.g., irrigation systems, trade infrastructures) have been destroyed and are under the process of reconstruction.

Therefore, the rural population has "limited" access to markets. Alternative off-farm income sources are, for instance, seasonal labour and, in decreasing order, other forms of wage labour, fishing trade, and sale of farm products.

Commercial agriculture is extremely limited, particularly considering the context of the vast agricultural potential of the country. Exceptions include concession areas that have been made available to expatriate farming initiatives and irrigation projects (particularly in the south on the Maputo River where rice and vegetables are grown on a semi-commercial basis) (Bowen 2000, Rutherfood 2010, Ibrahimo 2017).

This scenario of farming systems imposes physical limitations to the area that can be cultivated and to the achievable yields; therefore, a natural limit is also imposed on the total quantity of food that can be produced for self-consumption and/or for sale (INE 2018). Consequently, food security, agricultural risk assessment and sustainable production are top priorities and challenging societal issues in Mozambique (FAO, Country Programming Framework 2016-2020).

The described scenario is consistent with the lower public and private investment in the agricultural sector. In fact, the overall budget allocated to the sector (for functioning and investment) averaged only 4% of the Total State Budget between 2010 and 2019 (Aiuba and Mosca 2018). In the period 2001- 2010, the private investment was about 27% and was concentrated in cash crops (sugar cane, tobacco and cashew) and decreased to 22% during the period 2010 - 2015. In addition, the amount of credit granted to the agrarian sector represents only 8% of the total granted to the national economy and, more than 50% was destined to cash crops (cotton and sugar cane), which are mostly cultivated by large commercial companies (Mosca 2014).

Recognizing all the concerns in the agricultural sector, the Government of Mozambique adopted the “Plano Nacional de Investimento no Sector Agrario” (PNISA) (GOM (Government of Mozambique) 2012). PNISA targets include doubling the agriculture’s share of public expending and achieving 6% annual growth in agricultural gross domestic product (GDP).

As other countries in East Africa, Mozambique is highly vulnerable to extreme weather events (droughts and floods) and erratic and unpredictable precipitation patterns. These conditions negatively contribute to instability in crop production and often limit access to food for children and adults (Ackermann, Allgood et al. 2012). In this context, analytical tools to timely monitor crop production, risk, and vulnerability are strongly needed to support governmental agencies (national to regional) as well as various actors in the food management and supply chain. Timely and accurate knowledge of land and water use is also essential for understanding how climate variability affects regional crop production and for supporting the allocation of resources on medium and long-term planning.

Section III| Case Studies

Paper 1

Mananze, S.E.; Pôças, I.; Cunha, M. Maize leaf area estimation in different growth stages based on allometric descriptors. *African Journal of Agricultural Research* **2018**, *13(4)*, 202-209, doi:10.5897/AJAR2017.12916.

https://academicjournals.org/journal/AJAR/article-abstract/48FF9D555801.

Paper available in 25 January, 2018

Classification according to journal: research paper

Bibliometric indicators from the Journal Citation Report, Institute for Scientific Investigation
- Journal Impact Factor, Emerging sources citation index

academicJournals

Vol. 13(4), pp. 202-209, 25 January, 2018
DOI: 10.5897/AJAR2017.12916
Article Number: 48FF9D555801
ISSN 1991-637X

http://www.academicjournals.org/AJAR

African Journal of Agricultural Research

Full Length Research Paper

Maize leaf area estimation in different growth stages based on allometric descriptors

Sosdito Estevão Mananze[1,2,4]*, Isabel Pôças[3,4,5] and Mario Cunha[1,4,5]

[1]Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s.n., 4169-007 Porto, Portugal.
[2]Escola Superior de Desenvolvimento Rural, Universidade Eduardo Mondlane, Vilankulo, Moçambique.
[3]Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.
[4]Geo-Space Sciences Research Center, Rua do Campo Alegre s.n., 4169-007 Porto, Portugal.
[5]Institute for Systems and Computer Engineering, Technology and Science (INESC-TEC), CRIIS, Campus da Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias 4200-465 Porto, Portugal.

Received 4 December, 2017; Accepted 20 December, 2017

The literature points out the need for leaf area (LA) calibration models that are suitable for specific varieties (variety-specific). These models should be capable of coping with different crop conditions, growth stages, and agronomic practices. The objective of the current study was to develop a model for estimating the LA of maize (*Zea mays* L.), considering the entire growth cycle, based on non-destructive allometric measurements. The proposed model was derived from a multiple regression analysis of LA data obtained from digital image processing, including the number of leaves per plant (NL) and the product of major leaf length per major leaf width of the greater leaf (MLL × MLW). A high percent of data variability in the LA of maize plants was explained by the model, both in the calibration and validation phases ($R^2 = 0.90$; n = 30). Overall, the selected model presented good performance in the estimation of LA of maize, variety PAN 53, cultivated under the conditions of the present study area. Additionally, the model enabled the estimation of LA at different stages of the crop cycle. The results indicated a positive potential for using the developed model to support several maize cultural practices.

Key words: Allometry, non-destructive measurement, modelling, *Zea mays*.

INTRODUCTION

Leaf area (LA) is a determinant factor in many physiological and agronomic processes, particularly in terms of growth, photosynthesis, transpiration, water and nutrients use and productivity (Gao et al., 2012; Nangju and Wanki, 1980; Pandey and Singh, 2011).

Therefore, implementation of operational and accurate processes for measuring and estimating crop LA has long been a concern for researchers. There are currently several approaches for LA determination, which include direct and indirect methods. Direct methods include

---

*Corresponding author. E-mail: sosdito.mananze@uem.mz.

planimetric or gravimetric analyses of leaves, harvested directly or indirectly (Breda, 2003; Jonckheere et al., 2004). Portable scanning planimeters (e.g., LI-3000, Licor, NE, USA) are often used as a reference method for obtaining the LA.

Direct methods are more accurate but have the disadvantages of being very time-consuming, not user-friendly, and having constraints regarding equipment acquisition, price, and operation (Jonckheere et al., 2004). Moreover, direct methods can be destructive, not allowing successive measurements of LA (Peksen, 2007; Rouphael et al., 2010). One of the most frequently used indirect methods for LA estimation is based on observations and measurements of allometric parameters of the plants, which are used as inputs in mathematical models (Montgomery, 1911; Peksen, 2007). Such mathematical models are based on the correlation between the allometric measures of plants and the area of the leaves. These methods are non-destructive and allow for faster LA determination, eventually being suitable for automation. Nevertheless, an adequate parameterization and calibration of such methods is necessary.

The development of model for maize LA estimate based on alometrics has long been a concern for growers, breeders and researchers. A generalized leaf area equation LA = α × L × W for maize plants was proposed by Montgomery, (1911), based on a rectangle area L × W (L – leaf length, W – leaf width) and on a weighing factor (α) equal to 0.75. However, several authors indicated that the weighting factor may vary depending on the maize variety (Bange et al., 2000; Carvalho and Christoffoleti, 2007; Tivet et al., 2001), plant development stage (Bange et al., 2000), environmental conditions and agronomic practices (Elings, 2000; Sezer et al., 2009). Therefore, application of this classic equation requires a measurement of length and width of all leaves on a plant, which is very labour and time consuming, and can be a source of errors.

An alternative approach for estimating the maize LA based only on the largest leaf allometric measurements was developed for varieties adapted to temperate regions (Valentinuz and Tollenaar, 2006). When it was used on tropical varieties, these equations underestimated LA (Elings, 2000). Mondo et al. (2009) estimated maize LA based on one leaf, but not necessarily the largest. Although these models can perform well in estimating the LA at specific stages of the season, their portability to estimate the LA in different stages of maize development are not yet known.

According to Costa et al. (2016), the flexibility of LA models for use at different crop development stages is an important feature to support, throughout the crop cycle, different agricultural practices of high agronomic, economic and environmental importance, such as management of crop water requirements and dosage parameterization of pesticides applications.

The objective of this study was to develop a non-destructive and expeditious method and mathematical model for estimating TLA in the maize crop, variety PAN 53, at different phenological stages. The specific goals included (i) the development of an estimation methodology based on biometric measurements of specific plant leaf using image processing; and (ii) the development of a dynamic mathematical model that estimates the TLA of the crop stems throughout the cultural cycle of the maize.

MATERIALS AND METHODS

The current study was conducted in a field of 3 ha operated by the Joint Aid Management (JAM), a non-governmental organization, in the district of Vilankulo, within the province of Inhambane in southern Mozambique, latitude: 21° 58`S, longitude: 035° 09`E and altitude of 31 m above sea level (Figure 1).

The district of Vilankulo is characterized by a semi-arid to arid climate, with sandy soils of low fertility, and a high risk of agricultural production failure due to drought. The total annual rainfall is 733.9 mm, while the total annual evapotranspiration is 1135.1 mm, and the average annual temperature is 24.5ºC. The hot and rainy season occurs between November and March, with February being the hottest month (average monthly temperature of 26.9°C), and the average rainfall is about 166 mm. The cold and dry season occurs from April to October. July is the coldest month (average monthly temperature of 19.4°C) and drier, with about 17 mm of monthly rainfall.

Maize seeds of PAN 53 variety (from PANNAR Seeds Company) were used for the present study. Sowing was done on June 9, 2015 in the cold and dry season, and following geometry of 0.50 × 0.20 m. A drip irrigation system was used and fertilization was applied during irrigation. Harvest was done in October 2015. The PAN 53 variety has an average maturity, is resistant to major maize diseases and has a potential yield from 8 to 10 t/ha (PANNAR, s/d).

The alometrics measures took place from June to September 2015 in different phenological stages. The Lancashire et al. (1991) phenological stages description was adopted and data were collected at the following stages: plants with 3 (V3), 6 (V6), 8 (V8), 12 (V12) and 15 (V15) leaves unfolded; flag leaf just visible (VT); inflorescence emergence (R1) and medium milk (RT). Fourteen maize plants were randomly selected and monitored at each phenological stage. The recorded variables in each stage were: i) length and width of the largest leaf, ii) number of leaves per plant and iii) height and diameter of the stem. Additionally, in the stages V8 and R1, the full set of leaves of 30 randomly selected plants was collected, identified, marked and transported to the laboratory for measurements of length and width, using a graduated ruler.

The leaves were also digitized using a camera (Sony - Optical SteadyShot ® DSC - W730; 16.1 megapixels; 8x optical zoom), while keeping constant the distance of the image acquisition. The area of each leaf was determined by digital image processing, using the Image J software 1:48 (Wayne Rasband National Institute of Health, USA) and following the methodology described by Glozer (2008). Previous studies have shown reasonable results of LA estimations using Image J software and other image processing software (Costa et al., 2016). In fact, several authors showed the occurrence of no statistically significant differences between the results provided by this approach and the portable leaf area meter (Liquor Inc., Lincoln, Nebraska, US), which is considered the most accurate equipment for measuring LA (Dombroski et al., 2010; Santos et al., 2014).

A linear regression analysis was performed to assess the relationship between the total leaf area (TLA, which is the sum of

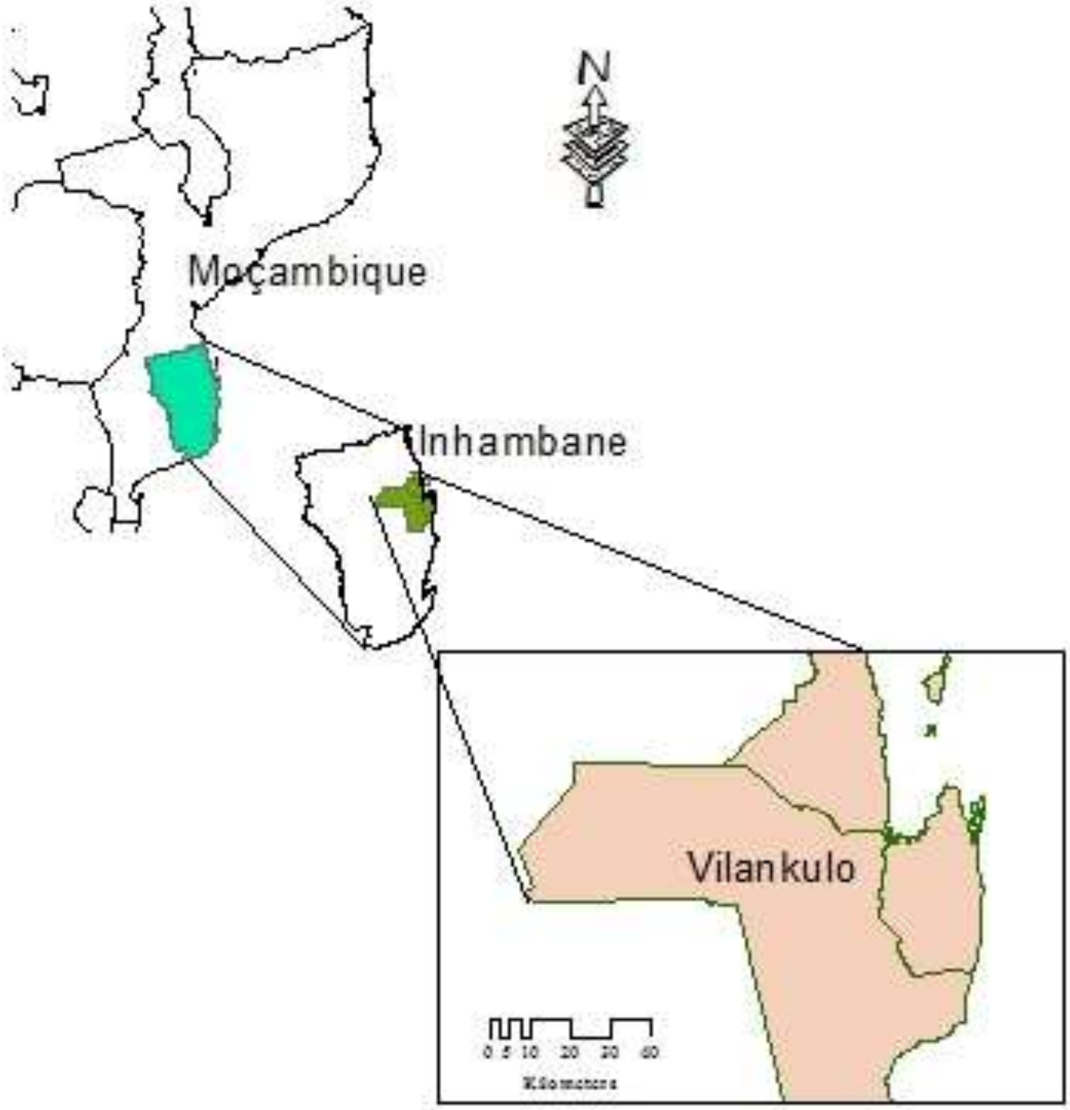


Figure 1. Geographic location of study area in Vilankulo, Mozambique.

LA for all leaves on a plant) and the measured allometric variables.

The dependent variable (TLA) was estimated according to the allometric measurements and their derivatives (transformations), to test the following linear regression models:

$$TLA = \beta 0 + \beta 1 \times NL \times L \times W \quad (1)$$

$$TLA = \beta_0 + \beta_1 \times NL + \beta_2 \times L + \beta_3 \times W \quad (2)$$

$$TLA = \beta_0 + \beta_1 \times NL + \beta_2 \times L \times W \quad (3)$$

$$TLA = \beta_0 + \beta_1 \times NL + \beta_2 \times L + \beta_3 \times W + \beta_4 \times H + \beta_5 \times D \quad (4)$$

where NL is the number of leaves on a plant; L and W are the length and width of the largest leaf; H is the plant height; D is the stem diameter; β_0, β_1, β_2, β_3, β_4 e β_5 are the regression parameters estimated specifically for each model using the ordinary least squares method. For the model calibration, data from the 60 (30 + 30) plants collected at the phenological stages V8 and R1 were aggregated into one sample. The aggregated sample was then divided into two independent random samples, one used for calibration and the other for validation.

Analysis of variance was performed to test statistical differences (F test) for each model. In addition, the standard deviation (SD) was computed for each parameter, and the statistical significance of model parameters were determined using the t-test. For each model, the normality and homoscedasticity assumptions were tested, and the absence of multicollinearity between independent variables assessed. The normal distribution of the residuals was determined through the Jarque-Bera test (Gujarati, 1995). The Breusch-Pagan test (Breusch and Pegan, 1979) was used to identify the homoscedasticity by testing the dependence of the residuals variance on the independent variables. In both tests, the null hypothesis assumed a homogeneous variance of residuals, or a normal distribution of the residuals. The null hypothesis was rejected for p value lower than 5% for the distribution of $X^2_{(2df)}$. The diagnosis of extreme observations or "outliers" was processed through the leverage test, establishing a maximum acceptable value of 1.5 (Montgomery et al., 2012).

The assessment of the model`s goodness-of-fit was done using the coefficient of determination (R^2), the efficiency coefficient – Nash-Sutcliffe – NSE, the linear regression through the origin and the index of agreement (IoA) between simulated and observed values. The NSE is a standard statistics that compares the relative magnitude of the residual variance with the variance of the observed data (Cunha et al., 2016). It has a range of -∞ to 1; the closer to 1, the more accurate is the model. Compared to R^2, the NSE is less sensitive to differences between the means and variances of the observed and predicted values. However, both are sensitive to extreme values, as reported by Legates and McCabe (1999) cited in (Cunha et al., 2016). The IoA has values ranging from 0 – 1, with 0 indicating lack of agreement and 1 perfect agreement.

Analysis of the residuals between observed and estimated values was used to evaluate the model accuracy and precision. Several indicators were considered: i) the absolute average error (AAE), ii) the mean squared error (MSE), iii) the mean root square error (MRSE) and iv) the relative mean root squared error (RMRSE). The Durbin Watson test (DW) was used for evaluating the auto-correlation between residuals assuming that values close to 2 denote the absence of autocorrelation.

For selection of the model with best performance, the Akaike information criteria (AIC, dimensionless), was also used based on the maximum likelihood function that allows generic comparison of models with different number of predictors. The AIC is calculated as follows:

$$AIC = N x ln\left(\frac{SQE}{N}\right) + 2k \qquad (5)$$

where N is the number of observations, SQE is the sum of square error, and K is the number of parameters + 1. Lower values of AIC indicate better models.

Evaluation of the regression assumptions and the model validation are very important for verifying the model suitability as a forecasting tool when using observations of new independent variables. In fact, the regression model can provide a good fit for the calibration sample data, but not when transposed outside the calibration confidence interval. For this reason, the statistical indicators of both calibration and validation phases, were used for model evaluation and selection. In addition to the statistical indicators, the easiness of application and the biophysical meaning were also taken in consideration.

Two validation procedures were applied: cross validation and external validation. The cross validation was applied over the full set of data (*n* = 60) using the “leave-one-out” (LOO) cross-validation method (Cunha et al., 2016). The LOO cross-validation evaluates the model performance for observations not considered in the estimation step, thus providing independent estimates of the predictive capability of the selected models. This technique consists of the removal of one observation from the dataset used, and the estimation of a new regression model with the remaining observations. This new regression model is then used to estimate the stem LA.

For the external validation, about 50% of the observations (30 plants) not used in the model parameter estimation, were used to evaluate the quality of the predictive model for these observations. We assume that the quality of the model validation is greater when the values of the indicators MSE, RMSE, AAE and the RRMSE are similar for the calibration and validation samples. The SPSS 23 software was used for the implementation of all statistical analysis.

RESULTS AND DISCUSSION

The dates for the occurrence of phenology and dynamics of plants growth are presented in Table 1. The maximum height (2325 mm) was recorded when plants presented the flag leaf just visible. The growth rate is in agreement with the expected patterns of crop growth (Table 1). Initially, there was an exponential growth up to the 15th leaf stage and hereafter the growth rate becomes very small.

The mean and standard deviation values for all the allometric descriptors presented in Table 2 were very close for the calibration and validation samples.

Predictors and their corresponding regression coefficients for the proposed models tested for estimating the TLA Equations 1 to 4 are presented in the following equations:

$$TLA = 564.2 + 0.48 x NL x L x W \qquad (1')$$

Int. conf. (95%) (-135.4|1263.8) (0.42|0.55)
t_student (0.11) (0.000)

$$TLA = -10834.3 + 415.6NL + 65.7L + 557.5W \qquad (2')$$

Int. conf. (95%) (-13717.4|-7951.3) (308.4|522.8) (38.7|92.7) (363|751.9)
t_student (0.000) (0.000) (0.000) (0.000)

$$TLA = -4672.8 + 420.3NL + 5.9L x W \qquad (3')$$

Int. conf. (95%) (-6035.4|-3310.2) (315|525.8) (4.4|7.6)
t_student (0.000) (0.000) (0.000)

$$TLA = -8537.9 + 306.1NL + 45.6L + 564.4W + 3.6H + 180.9D \qquad (4')$$

Int. conf. (95%) (-11593.6|-5482.1) (183.9 |428.4) (17.6|73.6) (370.7|758.2) (1.1|6.1) (-79.3|441.1)
t_student (0.000) (0.000) (0.000) (0.000) (0.000) (0.164)

Table 3 presents the statistical indicators for each model, when applied to the calibration and validation samples. Models 2’, 3’ and 4’ show slightly better calibration statistics but there were large differences in the validation statistics, particularly for the residuals indices and AIC (Table 3). The models 2’ and 4’ presented for the validation data-set a value of b in the regression 1:1 much lower than 1 (0.79 and 0.76, respectively) indicating a considerable underestimation when used as predictive tool (Table 3). In the particular case of model 4’, the estimated parameter for diameter and the respective estimation interval are not significant. Instead, model 1’ shows, for the calibration and validation sets, similar results of residual indices and AIC (Table 3), suggesting greater robustness and transferability when compared to the other models. Based upon these findings, the model 1’ was selected for estimating the TLA of maize, variety PAN 53.

The model 1’ explains 90% of the variability of maize TLA at different stages of crop development ($R^2 = 0.90$, n = 30; P <0.000) in both calibration and validation data-sets. The value of the regression coefficient β_1 was significantly different from zero (t_{test} P < 0.000) and the confidence interval for its estimation does not include

Table 1. Day of the year (DOY) for the occurrence of phenological stages and parameters of crop growth dynamics.

Phenological stage	DOY	Crop height (mm)	Growth rate (mm/day)
Sowing	160	---	---
3 leaves unfolded (V3)	177	67	3.9
6 leaves unfolded (V6)	198	207	6.7
8 leaves unfolded (V8)	211	507	23.1
12 leaves unfolded (V12)	224	1345	64.5
15 leaves unfolded (V15)	231	2094	107.0
Flag leaf unfolded (VT)	238	2325	33.0
Inflorescence emergence (R1)	246	2325	0.00
Medium milk (RT)	273	2325	0.00

Table 2. Descriptive statistics for the allometric descriptors used for model calibration and validation.

Allometric descriptors	N	Minimum	Maximum	Mean	Standard deviation
Number of leaves (NL)	30|30	80|90	140|140	111|117	15|13
Leaf length (mm)	30|30	925|860	1145|1170	1043|1043	57|64
Leaf width (mm)	30|30	78|74	115|116	93|98	7|9
Plant height (mm)	30|30	310|310	2306|2640	1105|1560	804|812
Stem diameter (mm)	30|30	150|220	45|45	30|32	6|6
Total leaf area (mm^2)	30|30	36939|41469	79468|85929	58297|65358	11575|11434

The vertical line separates the statistical indicators for the calibration and validation samples, respectively.

Table 3. The assumptions diagnostic and goodness-of-fit indicators for the calibration and the validation of the models proposed for estimating the total leaf area.

Statistics	Model (1’)		Model (2’)		Model (3’)		Model (4’)	
	Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
Model assumptions								
Leverage test	< 0.1	< 0.1	< 0.1	< 0.03	< 0.06	< 0.06	< 0.16	< 0.16
Jarque-Bera test	17.2	12.2	17.2	12.2	17.2	12.2	17.2	12.2
Beuch-Pegan test	0.55	0.16	0.55	0.16	0.55	0.16	0.55	0.16
Association measures								
R^2	0.90	0.90	0.91	0.88	0.91	0.88	0.89	0.89
R^2 [1:1]	0.89	0.89	0.90	0.87	0.91	0.89	0.93	0.85
b [1:1]	0.99	0.90	0.99	0.79	0.90	0.90	0.99	0.76
NSE	0.90	0.91	0.91	0.89	0.90	0.88	0.93	0.88
IoA	0.84	0.84	0.83	0.81	0.84	0.80	0.86	0.88
AIC	357.7	364.1	359.7	375.7	358.3	375.3	354.3	380.0
Durbin Watson	1.45	1.55	1.43	1.74	1.51	1.69	2.05	1.82
Residual indices								
MSE	131560.7	163339	115578.4	197103.7	118106.2	209252.5	84430	198804.5
RMSE	362.7	404.1	339.9	443.9	343.7	457.4	290.5	445.8
AAE	299.3	328.8	291.3	367.9	284.1	384.7	242.7	329.4
RRMSE	0.06	0.06	0.05	0.06	0.05	0.06	0.05	0.06

zero, which proved its statistical significance (model 1’). Additionally, the application of model 1’ was operational throughout the entire crop growth cycle, while the other models showed limitations by presenting negative values of TLA in the initial stages of crop growth (data not shown).

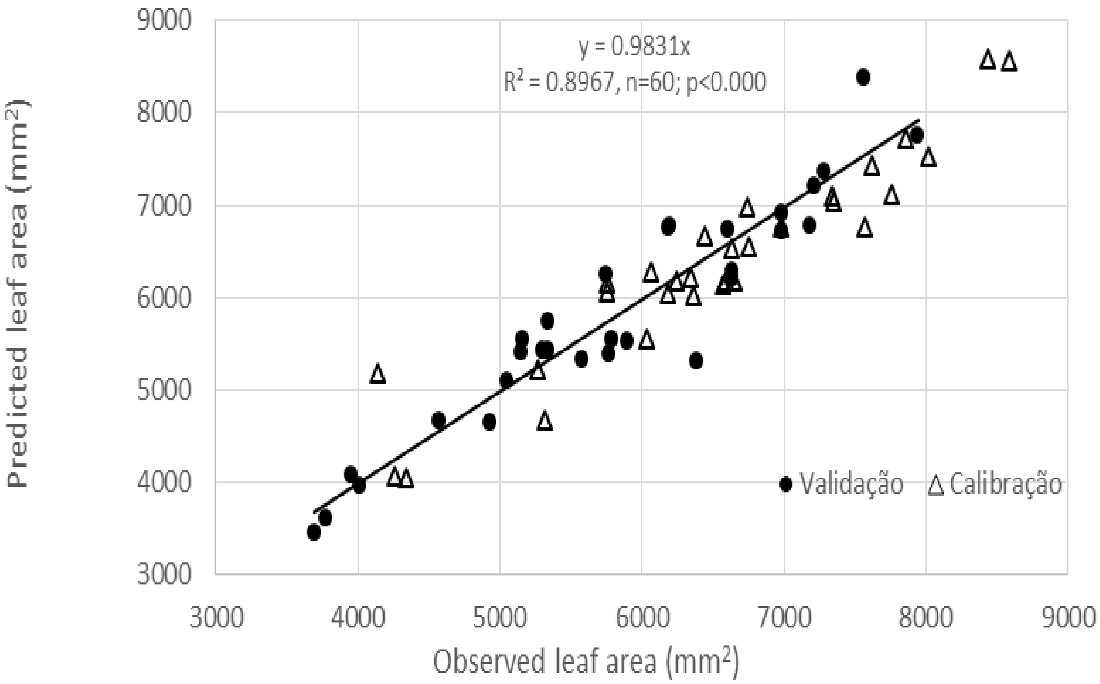


Figure 2. Linear regression through the origin between predicted and observed leaf area for calibration and validation data-sets.

The Jarque–Bera test was statistically significant for both calibration and validation samples, indicating normality of the residual variance, and the homoscedasticity of the variance could be confirmed by the statistical significance of the Breusch–Pegan test (Table 3). The efficiency coefficient for calibration (NSE = 0.90) and validation (NSE = 0.91) are within the range defined for accurate models. Additionally, the model indicates an excellent predictive power, if one considers its high level of agreement (IoA = 0.84). The measures of association suggest strong correlation between observed and predicted TLA, with the coefficient b higher than 1 for both calibration and validation data sets suggesting good accuracy. The slope of the regression through the origin was very close to one (0.99 for calibration and 0.90 for validation) and the coefficient of determination was 89%, showing that the model produced TLA values with high accuracy and precision at different plant development stages (Table 3).

Figure 2 illustrates the relationship between observed and predicted TLA for all data-set (n=60). The slope of the regression line (b) is very close to 1 (0.98), and the value of the coefficient of determination is high (0.89), indicating a good agreement between the observed and predicted values.

The relative difference between observed and predicted leaf area is less than 10% in over 91% of the cases, as shown in Figure 3. The deviations exceeded 10% only in 6.7 and 10% of the cases respectively for the calibration and validation data. The largest deviation (24.6%) was registered in the validation series, with all other cases presenting deviation lower than 20% (Figure 3).

The model relating the product between the number of leaves, length and width of the largest leaf (model 1') proved to be the most suitable for estimating the TLA of maize, variety PAN 53, in the agro-ecological conditions and agronomic practices of the study area. This model performed well when applied to the validation dataset, which suggests its accuracy in forecasting maize TLA. On the other hand, the selected model enabled the estimation of LA at different stages of the crop cycle, unlike other evaluated models which resulted in negative LA values at the initial stages of crop development (data not shown).

Studies using temperate (Valentinuz and Tollenaar, 2006) or tropical (Elings, 2000; Mondo et al., 2009) maize varieties demonstrated that the product of the length and width of the largest leaf is an important descriptor to estimate the total LA. These models were developed for a specific stage of crop development and therefore, do not include the number of leaves as the model developed in the current study. According to Elings (2000), models developed for temperate varieties are not suitable for application in tropical varieties. Likewise, in the current study, an attempt to apply the models developed by Mondo et al. (2009), Sezer et al. (2009) and by Montgomery (1911), resulted in substantially lower fit (R^2 = 0.597, 0.6416 and 0.6453, respectively, data not shown) when compared with the model '1'. As noted by

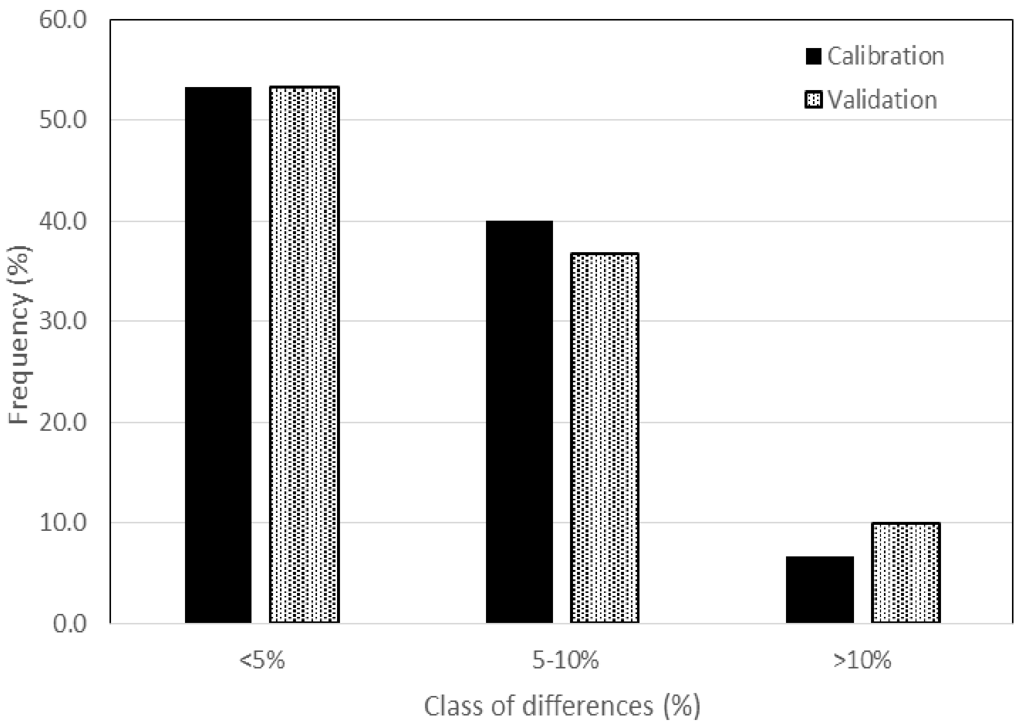


Figure 3. Frequencies (%) of differences between observed and predicted leaf area for calibration and validation sets.

several authors, these differences probably stem from genetic aspects of the studied varieties, agro-ecological conditions and agricultural practices of the study areas (Stoppani et al., 2003; Tsialtas et al., 2008).

Conclusion

The model equation developed from the current study is deemed suitable for estimating the total leaf area of maize plants based on data collected from various stages of the crop cycle. The accuracy of the leaf area estimation results, and the operability of the model developed in the current study are indicators of the model's potential use in different agricultural practices whereby decision-making depends on plant leaf area, such as spraying, fertilization and irrigation as well to support research project.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Bange MP, Hamer GL, Milroy SP, Ricket KG (2000). Improving estimates of individual leaf area of sunflower. Agron. J. 92:5.

Breda NJJ (2003). Ground-based measurements of leaf area index: A review of methods, instruments and current controversies. J. Exp. Bot. 54:15.

Breusch TS, Pegan AR (1979). A simple test for heteroscedasticity and random coefficient variation. Econometrica 47:8.

Carvalho SJP, Christoffoleti PJ (2007). Estimativa da área foliar de cinco espécies do gênero Amaranthus usando dimensoes lineares do limbo foliar. Planta Daninha 25:8.

Costa A, Pôças I, Cunha M (2016). Estimating the leaf area of cut roses in different growth stages using image processing and allometrics. Horticulturae P 2.

Cunha M, Ribeiro H, Abreu I (2016). Pollen-based predictive modelling of wine production: application to an arid region. European J. Agron. 73:42-54.

Dombroski JLD, Rodrigues GSO, Batista TMV, Lopes WAR, Lucena RRM (2010). Análise comparativa de métodos de determinação de área foliar em pinha (*Annona muricata* L.). Rev. Verde Agroecol. Desenvolv. Sust. 5:7.

Elings A (2000). Estimation of leaf area in tropical maize. Agron. J. 92:436-444.

Gao M, Van der Heijdena G, Vosb J, Eveleensc B, Marcelis L (2012). Estimation of leaf area for large scale phenotyping and modeling of rose genotypes. Scient. Hortic. P 138.

Glozer K (2008). Protocol for leaf image analysis – surface area. University of California. Department of Plant Sciences, Davis.

Gujarati D (1995). Basic econometrics, McGraw, New York.

Jonckheere I, Fleck S, Nackaerts K, Muys B, Coppin P, Weiss M, Baret F (2004). Review of methods for in situ leaf area index determination-Part I. Theories, sensors and hemispherical photography. Agric. For. Meteorol. 121:17.

Lancashire PD, Bleiholder H, Van Den Boom T, Langelüddecke P, Stauss R, Weber E, Witzenberger A (1991). A uniform decimal code for growth stages of crops and weeds. Ann. Appl. Biol. 119:41.

Legates DR, McCabe GJ (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35:8.

Mondo VHV, De Carvalho SJP, Labonia VD, Neto DD, Cicero SM (2009). Comparação de Métodos para Estimativa ee Área Foliar em Plantas de Milho. Rev. Bras. Milho Sorgo 8:233-246.

Montgomery D, Peck E, Geoffrey V (2012). Introduction to linear regression analysis, Wiley, Adelaid, Australia.

Montgomery EG (1911). Correlation studies in corn. Agricultural Experiment Station of Nebraska, Lincoln.

Nangju D, Wanki SBC (1980). Estimating leaf area of cowpea and soyabean using dry weights of terminal leaflets. Exp. Agric. 16:3.

Pandey SK, Singh H (2011). A simple, cost-effective method for leaf area estimation. J. Botany P 7.

Peksen E (2007). Non-destructive leaf area estimation model for faba bean (*Vicia faba* L.). Scientia Horticulturae 113: 322–328.

Rouphael Y, Mouneimne A, Ismail A, Mendoza-De Gyves E, Rivera C, Colla G (2010). Modeling individual leaf area of rose (*Rosa hybrida* L.)based on leaf length and width measurement. Photosynthetica 48:9-15.

Santos SND, DIgan RC, Aguilar MG, Souza CAS, Pinto DG, Marinato CS, Arpini TS (2014). Análise comparativa de métodos de determinação de área foliar em genótipos de cacau Biosci. J., Uberlandia 30:9.

Sezer I, Oner F, Mut, Z (2009). Non-destrutive leaf area measurement in maize (*Zea mays* L.). J. Environ. Biol. 30:785-790.

Stoppani MI, Wolf R, Francescangeli N, Marti HRA (2003). Non-destrutive and rapid method for estimating leaf area of broccoli. Adv. Hortic. Sci. 17:3.

Tivet F, Pinheiro BS, Raissac MDE, Dingkuhn M (2001). Leaf blade dimensions of rice (*Oryza sativa* L., *Oryza glaberrima* Steud.). Relationships between tillers and the main stem. Ann. Bot. 88:5.

Tsialtas JT, Koundouras S, Zioziou E (2008). Leaf area estimation by simple measurements and evaluation of leaf área predition models in Cabernet-Sauvignon grapvine leaves. Photosynthetica 46:5.

Valentinuz OR, Tollenaar M (2006). Effect of genotype, nitrogen, plant density, and row spacing on the area-per-leaf profile in maize. Agron. J. 98:6.

Paper 2

Mananze, S., Pôças, I and Cunha, M. 2018. Retrieval of maize leaf area index using hyperspectral and multispectral data. Remote Sensing. 10(12), 1942; https://doi.org/10.3390/rs10121942.

Paper available in: 3 December 2018

Classification according to journal: research paper

Bibliometric indicators from the Journal Citation Report, Web of Science (JCR – WoS), Clarivate Analytics:
- Journal impact factor 4.118
- Journal rank

Remote Sensing: position 7/30; 1st quartile (Q1)

remote sensing

MDPI

Article

Retrieval of Maize Leaf Area Index Using Hyperspectral and Multispectral Data

Sosdito Mananze [1,2,3], Isabel Pôças [2,3,4,5] and Mario Cunha [2,3,5,*]

[1] Escola Superior de Desenvolvimento Rural, Universidade Eduardo Mondlane, Vilankulo PC-257, Mozambique; sosdito.mananze@uem.mz
[2] Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s.n., 4169-007 Porto, Portugal; isabel.pocas@fc.up.pt
[3] Geo-Space Sciences Research Center, Rua do Campo Alegre s.n., 4169-007 Porto, Portugal
[4] Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal
[5] Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Campus da Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
* Correspondence: mccunha@fc.up.pt; Tel.: +351-25-2660-400 or +351-25-2963-631

Received: 2 October 2018; Accepted: 30 November 2018; Published: 3 December 2018

Abstract: Field spectra acquired from a handheld spectroradiometer and Sentinel-2 images spectra were used to investigate the applicability of hyperspectral and multispectral data in retrieving the maize leaf area index in low-input crop systems, with high spatial and intra-annual variability, and low yield, in southern Mozambique, during three years. Seventeen vegetation indices, comprising two and three band indices, and nine machine learning regression algorithms (MLRA) were tested for the statistical approach while five cost functions were tested in the look-up-table (LUT) inversion approach. The three band vegetation indices were selected, specifically the modified difference index ($\mathrm{mDI}_{\mathrm{d}:\,725;\,715;\,565}$) for the hyperspectral dataset and the modified simple ratio ($\mathrm{mSR}_{\mathrm{c}:\,740;\,705;\,865}$) for the multispectral dataset of field spectra and the three band spectral index ($\mathrm{TBSI}_{\mathrm{b}:\,665;\,865;\,783}$) for the Sentinel-2 dataset. The relevant vector machine was the selected MLRA for the two datasets of field spectra (multispectral and hyperspectral) while the support vector machine was selected for the Sentinel-2 data. When using the LUT inversion technique, the minimum contrast estimation and the Bhattacharyya divergence cost functions were the best performing. The vegetation indices outperformed the other two approaches, with the $\mathrm{TBSI}_{\mathrm{b}}$ as the most accurate index (RMSE = 0.35). At the field scale, spectral data from Sentinel-2 can accurately retrieve the maize leaf area index in the study area.

Keywords: field-spectroradiometer; Sentinel-2; hyperspectral; multispectral; leaf area index; vegetation indices; machine learning regression algorithms; PROSAIL; LUT inversion

1. Introduction

The leaf area index (LAI) is a parameter of crop structure, which is key for many agronomic and physiologic studies involving plant growth, light interception, photosynthetic efficiency, evapotranspiration, and plant response to irrigation, fertilization, and other types of agricultural practices [1]. The LAI describes the mass and energy exchange surface between the Earth's surface and the atmosphere, influences the within and bellow canopy microclimate, and determines and controls the canopy water interception, the radiation extinction, and the water and carbon gas exchange [2]. Due to its role as an interface between the ecosystems and the atmosphere, studies involving LAI have applications in a wide range of fields, particularly in agriculture, forests, ecology, hydrology, eco-physiology, and meteorology [3,4].

Remote Sens. **2018**, *10*, 1942; doi:10.3390/rs10121942

Remote sensing data acquired from different types of sensors and processed through various modelling approaches have been revealing a high potential for retrieving and mapping LAI, which is an important contribution to the management of agricultural fields at different spatial and temporal scales [5–7]. The principle inherent to the application of spectral information for LAI (or other biophysical variable) retrieval is related to the changes in crop spectral behavior in response to variations of physiologic and structural conditions of plants as well as to the surrounding environment conditions [8].

Although data from a wide diversity of multispectral sensors is currently available, there are some limitations in their use for LAI retrieval due to the occurrence of saturation in high LAI conditions or high amounts of biomass [9–11]. Thus, there has been an increasing interest in the application of hyperspectral data. The numerous narrow bands of hyperspectral sensors provide continuous spectral measurement across the electromagnetic spectrum, which are more sensitive to subtle variations in reflected energy and, therefore, have a greater potential for detecting differences between surface characteristics [12]. Thus, through hyperspectral data it is possible to use narrow bands specifically suited for quantifying biophysical and/or biochemical variables of vegetation [13].

The estimation of plant biophysical variables using spectral data can be achieved through statistically (or empirical) and physically based approaches. The statistical techniques are used to analyze the relationship between variables of interest measured in situ and the crop spectral reflectance or its transformations in the form of vegetation indices or regression algorithms. Several studies have applied this approach to estimate different biophysical parameters of agricultural crops, including the estimation of LAI [14]. Although this approach is easy to implement and has been successfully applied in several studies, it is often site specific, i.e., the transferability of the resulting statistical relationship usually presents constraints when used under different conditions (locations, stages of crop development, agricultural practices, etc.), or with sensors of different characteristics, namely in their spectral resolution [15]. To cope with this limitation, it is necessary to ensure a broad sampling, considering different stages of crop development, different varieties (within a single crop), and different locations [16].

The statistical relationship between crop biophysical data and their reflectance is generally done through vegetation indices that may assume different formalisms: (i) Indices using single-band reflectance or the difference between two wavelengths; (ii) reflectance ratios between two or more bands; (iii) normalized difference ratios; and (iv) indices based on reflectance derivatives [14]. Alternatively, this relationship can be assessed using machine learning algorithms [17].

In the physical approach, methods of inversion of radiative transfer models (RTMs) are applied. The radiative transference is the physical phenomenon of energy transfer in the form of electromagnetic radiation, which is affected by the absorption, emission, and dispersion processes. The RTMs describe, based on physical laws, the temporal variation of canopy spectral reflectance as a function of its biochemical and structural properties coupled with the soil properties and the illumination conditions of the object [18]. For these reasons, this approach is considered more robust than the statistical. However, the application of RTMs, in addition to reflectance, requires supplemental information relative to canopy structure and reflectance of proximity objects, such as soil, to characterize the physical model and describe the conditions in which the model is valid. This information is not always available and, even when available, the reflectance simulation process is computationally demanding. These are the main limitations for the RTMs' applicability [15]. Nevertheless, the RTMs have been applied in a great diversity of regions and types of vegetation, considering several types of sensors for the retrieval of different vegetation biophysical parameters.

The most commonly used RTMs for the estimation of vegetation biophysical properties are PROSPECT (leaf optical proprieties spectra) and SAIL (scattering by arbitrarily inclined leaves). The PROSPECT model, proposed by Jacquemoud and Baret [19], simulates the reflectance and transmittance of the leaf [20], and it is frequently coupled with the SAIL model to simulate the canopy's bidirectional reflectance [12].

The combination of POSPECT and SAIL (PROSAIL) interconnects the variation of the spectral reflectance of the canopy, which is mainly related to the biochemical attributes of the leaf, with its directional variation, which is mainly related to the canopy architecture and the contrast between soil/vegetation. This linkage is essential for the simultaneous estimation of the biophysical and structural variables of the vegetation [21]. However, the parameterization of PROSAIL seems quite challenging in such a way that most of the studies use nominal values or ranges applied in other studies carried out with the same crops, but in different places. In Table A1 (Appendix A), we present the PROSPECT and SAIL models, including parameters used in several crops and vegetation types.

Overall, although several mentioned studies have achieved good results estimating and mapping crop biophysical properties through different remote sensing data and methods, according to [22], there are still gaps to be explored in terms of the use and operationalization of these data and methods. The diversity of bands selected for estimating the same variable, for the same crop, in different studies, and the variability of the intervals of parameters used in the calibration of the RTMs for the same crop, emphasize the impact of the radiometric characteristics of the sensors, the agricultural practices (which determine the canopy structure), the environmental conditions, and sampling conditions on the models' results. This assumption is even more important when remote sensing data and methods are to be applied in low-input crop systems and highly heterogeneous farming systems, as is the case of the present study area. In this type of farming system, the limitation of the production factors, such as fertilizers and irrigation, leads to crops with very low LAI and high infield heterogeneity, which constitutes an additional challenge in the application of remote sensing techniques.

In this context, the main goal of this work is testing the applicability of different types of remote sensing data (hyperspectral and multispectral) and methods (statistical and physical) for developing models to predict LAI in maize crop systems characterized by low input factors, high spatial and intra-annual variability, and low yield in the south of Mozambique. The specific objectives include (i) to compare the performance of multi and hyperspectral data in estimating the LAI of maize in the study region; and (ii) to test the performance of the RTMs and the statistically-based models in the estimation of LAI of maize.

2. Materials and Methods

2.1. Study Area Description

The current study was conducted in maize crop fields located in the district of Vilankulo, within the province of Inhambane in southern Mozambique (lat.: 21°58′S, long.: 035°09′E, 31.83 m a.s.l) (Figure 1).

The district of Vilankulo is characterized by a semi-arid to arid climate, with sandy soils of low fertility. Figure 2 presents the average precipitation and temperature in the study area for a reference period and for the three years of data collection. For the period studied (2015–2018), the average of total annual rainfall was 403 mm, which is 50% below the 804.4 mm registered in the period (1991–2015). The average annual temperature of 26.9 °C during the study period was warmer than that recorded in the period of 1991–2015 (24.5 °C). The summer and rainy season (January, February, March, October, November, and December) accounted for the highest amount of rainfall for both the reference period (79.8%) and study period (81.2%). The driest conditions were recorded in 2015 with 335.3 mm of total annual rainfall and an average temperature of 29.0 °C.

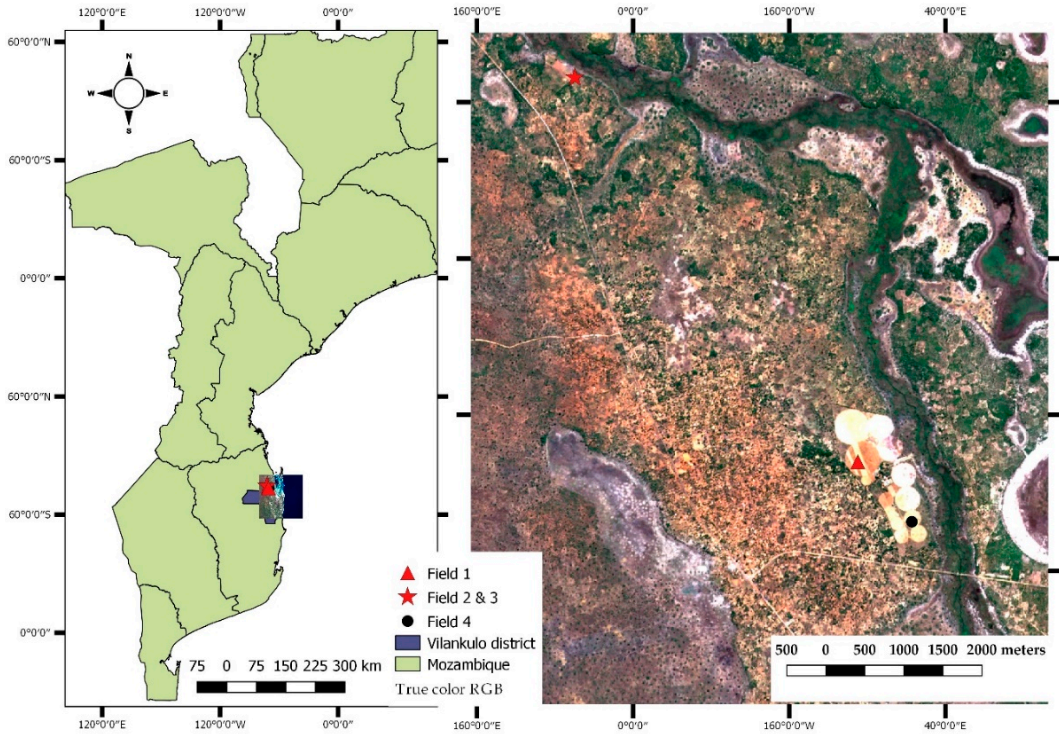


Figure 1. Location of the sampling fields overlaid to a true color composite (Red, Green and Blue) of Sentinel-2 image (**right**), and the boundaries of the country of Mozambique and the district of Vilankulo (**left**).

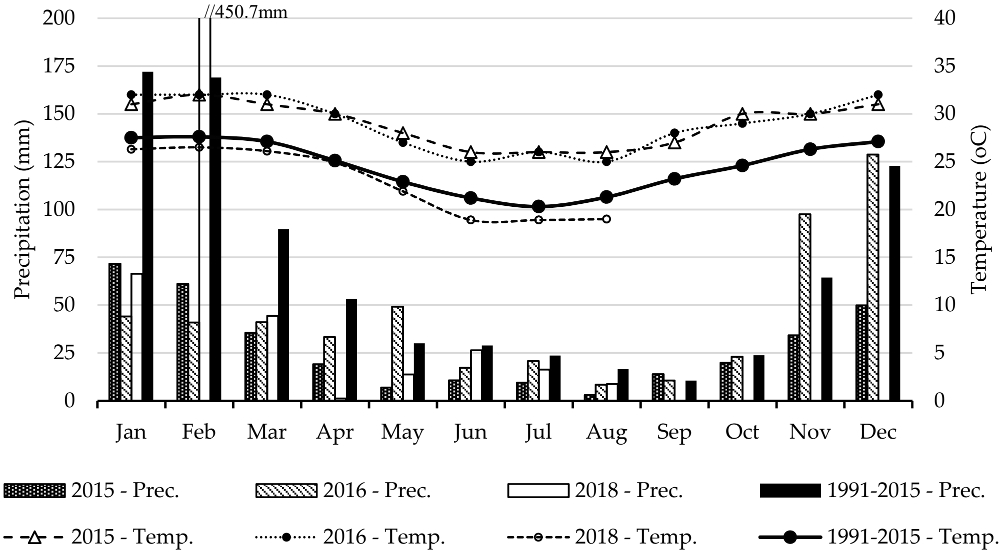


Figure 2. Average monthly precipitation (Prec.) and temperature (Temp.) for the three years of data collection (2015, 2016, and 2018) and comparison with a reference period (1991–2015) [23].

2.2. Characteristics of the Sampled Fields

The data collection took place in four maize fields in 2015 (field 1), 2016 (field 2 and field 3) and 2018 (field 4). In 2015 and 2016, the crop cycle coincided with the cold and dry season (April–September) while in 2018, the crop was grown during hot and rainy season. In each field, two 400 m^2 parcels were established for monitoring and data collection. Table 1 summarizes the cropping practices in each sampled field.

Table 1. Characteristics and main agricultural practices of the sampled fields.

Item	**Field 1 (2015)**	**Field 2 (2016)**	**Field 3 (2016)**	**Field 4 (2018)**
Latitude	21°56′24.19′′	21°59′02.53′′	21°56′20.88′′	21°59′24.62′′
Longitude	35°07′24.19′′	35°09′30.95′′	35°07′18.12′′	35°09′55.53′′
Surface	3 ha	1 ha	1 ha	3 ha
Soils	Sandy-loam soils	Sandy-loam soils	Sandy-loam soils	Sandy-loam soils
Irrigation system and scheduling	Sprinkler: irrigation schedule conditioned by the water pumping availability	Drip irrigation: 3 days interval from V3–V8 and 6 days interval in the following stages	Drip irrigation: 3 days interval from V3–V8 and 6 days interval in the following stages	Central pivot: 5 days irrigation interval (with punctual constraints due power fluctuations)
Variety	PAN 53, a medium maturity variety (125–140 days to harvest)	PAN 53	PAN 53	PAN 67, a medium maturity variety (120–130 days to harvest)
Planting geometry	0.5 × 0.25 cm;	0.75 × 0.2 cm;	0.75 × 0.2 cm;	0.9 × 0.15 cm;
Sowing and harvest dates	9 June/30 October	2 June/2 November	4 July/4 December	10 December 2017/25 April 2018
Crop yield	2.5 Ton/ha	5 Ton/ha	5 Ton/ha	4 Ton/ha
Agricultural practices	Surface fertilization with Urea; manual weed removal; insect control of insect *Sesamia monogriodes* with cipermetrine at stages V8, VT and R	Deep fertilization with Guano (1200 kg/ha) before sowing; 3 applications of Mono-Ammonium Phosphate (MAP) (200 kg/ha) and Ammonium Sulphate (100 kg/ha) throughout the season	Deep fertilization with Guano (1200 kg/ha) before sowing; 3 applications of Mono-Ammonium Phosphate (MAP) (200 kg/ha) and Ammonium Sulphate (100 kg/ha) throughout the season	Deep fertilization with Guano (1200 kg/ha) early before sowing; 3 applications of Mono Amonium-Phosphate (MAP) (200 kg/ha) and Ammonium Sulphate (100 kg/ha) throughout the season; application of insecticides to control the *Spodoptera frugiperda*

2.3. Data Collection

2.3.1. Field LAI Data Collection

The field LAI data was obtained from an allometric model, which estimates the plant total leaf area based on biometric variables (length and width) and the number of leaves. The model variables were collected during the years, 2015, 2016, and 2018, in seven plants randomly selected and marked within each parcel. The maize measurements (allometrics descriptors) were recorded in different crop phenologic stages (Table 2) described in accordance with the leaf collar method [24]. This allometric model was calibrated and validated during the 2015 field campaign using the maize variety, 'PAN 53', and explained 90% of the maize total leaf area variability at different stages of crop development ($R^2 = 0.90$; n = 30; $p < 0.000$) [1]. Additionally, an independent validation of the model was performed using data from the crop variety, 'PAN 67', grown in 2018 ($R^2 = 0.914$, n = 16, $p < 0.000$; data not published).

Table 2. Dates of field leaf area index collection and corresponding crop phenological stages.

Crop Stage	Dates of Data Collection			
	Field 1 (2015)	**Field 2 (2016)**	**Field 3 (2016)**	**Field 4 (2018)**
V3	2 July	28 Jun	30 July	9 January
V6	20 July	19 July	12 August	25/January
V8	31 July	30 July	24 August	5 February
VT	28 August	17 August	13 August	4 March
R1	3 September	28 August	27 September	
RT		16 September	25 October	

V3, V6, V8—stages at which the plants have 3, 6, and 8 leaves with visible leaves, respectively; VT—stage at which the last new leaf is completely developed; R1—first reproductive stage; RT—last reproductive stage.

2.3.2. Field Equivalent Water Thickness and Dry Mater Content

For the estimation of the equivalent water thickness (Cw,) and dry mater content (Cm, g/cm^2), the first full developed leaf, defined from the top to the bottom of each sampled plant, was cut and enclosed in a sealed plastic bag and brought to the laboratory inside a cooler. In the laboratory, a 2 cm diameter disc was cut in each leaf and the wet weight taken before the discs were dried at 65 °C in an oven until constant weight (dry weight). The Cw (cm) was then calculated according to the following equation:

$$\mathrm{Cw} = \frac{\mathrm{F_w} - \mathrm{D_w}}{\mathrm{A}} \times \mathrm{dw} \tag{1}$$

where F_w (g) and D_w (g) are, respectively, the fresh and dry weights of the leaf discs; A (cm^2) is the area of the disc; and dw (1 g cm^{-3}) is the water density. For the estimation of Cm, we used the following equation:

$$\mathrm{Cm} = \frac{\mathrm{D_w}}{\mathrm{A}} \tag{2}$$

2.3.3. Field Spectral Data Collection

The field spectral data were acquired using an Apogee hyperspectral spectroradiometer with sensitivity in the range of 236 nm–1100 nm and a spectral resolution of 1 nm. The instrument is attached to an optic cable with a field of view (FOV) of 30°. The measurements were taken at the nadir position, keeping a constant distance of about 5 cm to the spot, and were carried out around solar noon (between 11 a.m. and 1 p.m.), when the changes in solar zenith are minimal. Standard procedures of the equipment calibration were followed, namely the dark object correction and the acquisition of white reference reflectance of the Spectralon. The reference reflectance was updated every 10 min during the measurements. The spectra readings were done in the same dates (Table 2) and points where the data for LAI determination were collected. In the phenological stages, V3, V6, and V8, while the fraction of canopy cover was still very low, two additional readings were done, one over the interface between two plants and another over bare soil. In each point, 10 replications were made, with each an average of 20 automatic readings.

2.3.4. Spectral Data from Sentinel-2

The multispectral sensors imaging (MSI) on board of the Sentinel-2A and Sentinel-2B, launched, respectively, in June 2015 and March 2017, have a temporal resolution of 10 days and present 4 bands (2, 3, 4, and 8) with 10 m of spatial resolution and 6 bands (5, 6, 7, 8a, 11, and 12) with 20 m of spatial resolution. The high spatial and temporal resolutions and the spectral intervals covered by these bands make Sentinel-2 potentially more interesting than other multispectral sensors for studies related with LAI [25–27].

Atmospherically corrected Sentinel-2A and Sentinel-2B images of the study area, including their value added product bands, were acquired from the data service platform developed by [28]. For all the pixels covering each sampling parcel, the spectral data of all the original Sentinel-2 bands (SP_S2) and of the Sentinel LAI products (LAI_S2) were extracted and then averaged, using a 10 m spatial filter applying the zonal statistics tool of QGIS software, version 2.14.14. The Sentinel-2 images were selected considering the closeness to field spectral data collection dates and low cloud cover. The images from the following dates were then selected and downloaded: 30/06/2016; 20/07/2016; 30/07/2016/; 28/09/2016; 21/01/2018; 05/02/2018, and 02/03/2018. The dates of satellite acquisition and field data collection did not always match due to frequent high cloud cover conditions in the study area at the time of satellite overpass. However, nearby satellite image dates, corresponding to the same crop phenological stage, were always considered.

2.4. Data Analysis

2.4.1. Field Spectral Data Pre-Processing

The reflectance of each sampling point was achieved by averaging the 10 replications. For the stages, V3, V6, and V8, the average included the reflectance of the plant, bare soil, and interface between two plants. A weighted average of the three ground cover components was considered to obtain a reflectance value representative of maize growth conditions at each sampling point and allow the comparability with satellite data. Due to noise in the initial portion of the spectra, data analysis was performed considering data in the interval of 400 nm–1100 nm of the spectra. Afterwards, these data were spectrally aggregated to simulate hyperspectral and multispectral datasets. For the hyperspectral dataset, a 10 nm spectral resolution aggregation (FSP_10) was considered while for the multispectral dataset, the Sentinel-2 band setting was used for aggregation (FSP_S2). Because the field spectra does not match the whole Sentinel-2 spectra, the aggregation was limited to 8 bands (2 A–8 A) acknowledged as being more convenient in LAI estimations [2,29].

2.4.2. Statistical Modelling of LAI Based on Vegetation Indices

Table 3 presents the spectral vegetation indices (VI) assessed as predictors for LAI statistical modelling. The VIs included band combinations in the form of the simple ratio (SR), modified simple ratio (mSR), normalized difference (ND), modified normalized difference (mND), difference between bands (DI), and modified difference between bands (mDI). For each VI, all possible band combinations were tested considering the three types of data: Hyperspectral dataset (FSP_10), the multispectral dataset (FSP_S2), and the Sentinel-2 data (SP_S2). The FSP_10 and FSP_S2 were correlated only to the field leaf area index (Field_LAI) while the SP_S2 was correlated to both the average Field_LAI and Sentinel-2 leaf area index product (LAI_S2). The linear, exponential, and second degree polynomial functions were evaluated for modelling the LAI based on VIs. The analysis of the best band combination per VI and the modelling approach were carried out in the Spectral Index Toolbox of the software Automated Radiative Transfer Models Operator (ARTMO) [15,30].

Table 3. Spectral vegetation indices tested for statistical LAI modelling.

Type of Index	Formulation	Original Index and Source
2 bands index		
ND	$(\rho a - \rho b)/(\rho a + \rho b)$	NDVI; [31]
mND_a	$[(\rho a - \rho b)/(\rho a + \rho b + 0.5)] * 1.5$	SAVI; [32]
SR	$\rho a/\rho b$	SR; [33]
mSR_a	$\rho a/\rho b - 1$	CI Green; [34]
DI	$\rho a - \rho b$	DI; [35]
mDI_a	$(1/\rho a) - (1/\rho b)$	ARI; [36]
3 bands index		
mDI_b	$(\rho a - \rho b) - 0.2 * (\rho a - \rho c)$	CARI; [37]
mND_b	$2.5 * [(\rho a - \rho b)/(\rho a + 6 * \rho b + 7.5 * \rho c + 1)]$	EVI; [38]
mND_c	$(\rho a - \rho b)/(\rho a + \rho b - \rho c)$	VARI; [39]
mSR_b	$(\rho a - \rho b)/(\rho a - \rho c)$	SIPI; [40]
mSR_c	$(\rho a - \rho b)/\rho c$	PSRI; [41]
mDI_c	$[(\rho a - \rho b) - 0.2 * (\rho a - \rho c)] * (\rho a/\rho b)$	mCARI; [42]
mDI_d	$[(1/\rho a) - (1/\rho b)] * \rho c$	mARI; [36]
mDI_e	$[(\rho a + \rho b)/2] - \rho c$	RVSI; [43]
TBSI_a	$(\rho a - \rho c)/(\rho b + \rho a)$	[14]
TBSI_b	$(\rho a - \rho b + 2\rho c)/(\rho a + \rho b + \rho c)$	[44]
TBSI_c	$(\rho a - \rho b - \rho c)/(\rho a + \rho b + \rho c)$	[45]

ρ—Reflectance at given wavelength; SR—simple ratio; ND—normalized difference; mSR—modified simple ratio; mND—modified normalized difference; TBSI—three band spectral index; DI—difference index; EVI—enhanced vegetation index; SAVI—soil adjusted vegetation index; NDVI—normalized difference vegetation index; VARI—visible atmospherically resistant index; CARI—chlorophyll absorption ratio index; mCARI—modified chlorophyll absorption ratio index; CI_{Green}—green chlorophyll index; ARI—anthocyanin reflectance index; mARI—modified anthocyanin reflectance index; SIPI—structure insensitive pigment index; PSRI—plant senescence reflectance index; RVSI—reflectance vegetation stress index. Figure 3 presents the flowchart of the two statistical modelling approaches.

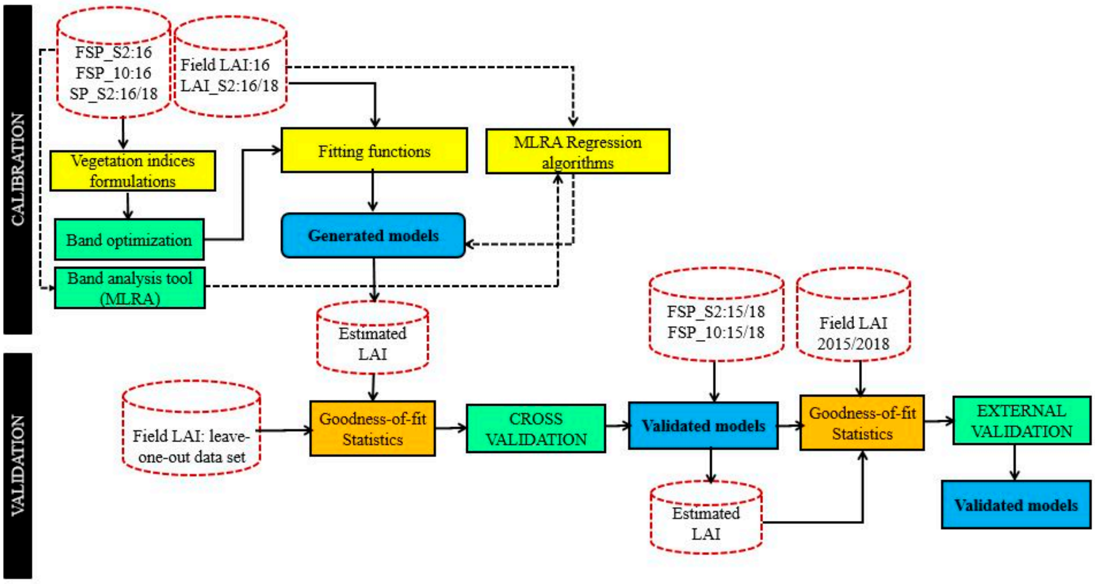


Figure 3. Statistical modelling process with the leaf are index (LAI) and spectral data used in the calibration and validation. In the calibration section, the full line arrows indicate the process for the vegetation indices approach while the dashed arrows indicate the process for the machine learning approach.

2.4.3. Statistical Modelling of LAI Based on Machine Learning Regression Algorithms

The machine learning regression algorithms (MLRA) are nonparametric models that are adjusted to predict a variable of interest using a training dataset of input-output data pairs, derived from synchronised measurements of the parameter and the corresponding reflectance observations. Numerous nonparametric regression algorithms are available in the statistics and machine learning literature, and they are being increasingly applied for vegetation biophysical parameters' retrieval [14,16,17,46,47].

Similarly to the VIs, the MLRA algorithms were evaluated using the three types of data: Hyperspectral dataset (FSP_10), multispectral dataset (FSP_S2), and Sentinel-2 data (SP_S2). The FSP_10 and FSP_S2 were correlated only to the field leaf area index (Field_LAI) and the SP_S2 was correlated to both the Field_LAI and Sentinel-2 LAI product (LAI_S2). Table 4 presents the MLRA evaluated in the present study and the modelling process is presented in Figure 3. Prior to the application of the MLRA algorithms, an analysis of the best number and combination of bands for estimating the LAI using the different datasets was performed based on the band analysis toolbox. The algorithms were also implemented on the MLRA toolbox [17] of the software Automated Radiative Transfer Models Operator (ARTMO).

Table 4. Machine learning regression algorithms tested for LAI modelling.

Algorithm	Source
Regression tree (RT)	[48]
Random Forest (TreeBagger) (RFTB)	[49]
Bagging trees (BaT)	[50]
Relevance vector machine (RVM)	[51]
Kernel ridge regression (KRR)	[52]
Gaussian process regression (GPR)	[53]
Variation Heteroscedastic Gaussian process regression (VH-GPR)	[54]
Support Vector Regression (SVM)	[55]
Random Forest (Fitensemble) (RFF)	[49]

2.4.4. Retrieval of LAI Based on Radiative Transfer Models

To perform an LAI retrieval through RTMs, one need to first build up a look-up-table (LUT), which is a database of simulated canopy reflectance and their corresponding set of input parameters. Afterwards, an inversion technique is applied to search and identify the parameters' combination that yields the best fit between the measured and LUT reflectance [18,56].

Simulation of the Look-Up-Table (LUT)

The PROSAIL model, a widely used RTM model for biophysical parameters' retrieval [14,57], was used to simulate reflectance data that were stored in an LUT. Table 5 presents the input parameters' settings used for the LUT simulations. The input parameters relative to LAI, water equivalent thickness (Cw), dry matter content (Cm), hotspot (hspot), and the reflectance of dry and wet soil were set according to field data. The parameter of leaf structure (N) was set to the 1–1.4 interval as proposed by [2,58,59] for various crops, including maize. The interval of values considered for the chlorophyll a and b content (Ca+b) was defined slightly below the ranges presented in the current literature considering the local conditions, which are prone to stress occurrence due to irrigation deficit conditions and damages by insects. As mentioned by several authors, when plants are exposed to stress (water deficit, insect damage, nutrient deficit, etc.), the pigments' concentration is reduced and/or degraded [60–62], lowering the energy absorption and thus increasing the reflectance in the visible domain. The observer zenith angle (tto), the azimuth angle (psi), and the sun zenital angle (tts) were set according to the geometry of field spectra readings.

A total number of 100,000 random simulations was used for generating the LUT spectral reflectance data. This simulations' number is considered appropriate for good computational efficiency and high accuracy in the parameter estimation [63–65]. The simulated data were aggregated into the eight Sentinel-2 bands (2–8 A) in order to match the FSP_S2 dataset. The PROSAIL simulations were carried out in the software, ARTMO [30].

Table 5. Input parameters of PROSAIL used to generate LUT for maize in the study area.

Model/Parameter	Abbreviation	Unit	Range of Values	Fixed Values
Prospect 5 model				
Equivalent water thickness	Cw	cm	0.001–0.030	-
Leaf chlorophyll content	Cab	$\mu g/cm^2$	5–40	-
Leaf structure coefficient	N	No dimension	1–1.4	-
Dry matter content	Cm	g/cm^2	0.001–0.008	-
Carotenoids content	Car	$\mu g/cm^2$	-	10
Brown pigments content	Cbrown	g/cm^2	-	5
4SAIL model				
Leaf area index	LAI	m^2/m^2	0.01–3.5	-
Average leaf angle	ALA	Degree	20–60	-
Hot-spot size parameter	Hspot	m/m	0.25–1	-
Diffuse/Direct light	Skyl	No dimension	-	10
Soil Coefficient	Psoil	No dimension	-	0.6
Solar Zenith Angle	Tts	Degree	-	10
Observer zenith Angle	Tto	Degree	-	5
Azimuth Angle	Psi	Degree	-	0

LUT Inversion for LAI Retrieval

A look-up table (LUT) based inversion technique was applied for the LAI retrieval. This technique is considered easy-to-use and potentially outperforms the limitations of iterative techniques [63,66]. The LUT inversion consists of a direct comparison of simulated spectral data with the observed spectral data (field or satellite data), through one or various cost functions, aiming to identify the best parameter combinations that yield minimal differences between the observed and simulated data [14,67]. The LUT inversion toolbox [68] of the ARTMO software was used to perform this inversion task. The toolbox enables testing simultaneously a wide diversity of cost functions making it possible to optimize the models for different assumptions on the nature and proprieties of the residuals [68,69]. Several classes of cost functions were considered, including minimum contrast estimation, information measures, and M-estimates.

The minimum contrast estimation considers the bidirectional reflectance as a spectral density of the stochastic process and aims at minimizing the distance between a parametric model and a non-parametric spectral density [69]. Within this class, three cost functions were tested, which are detailed in [69,70]:

$$K(x) = -logx + x,\ then,\ D[E,L] = \sum_{\lambda=1}^{\lambda n}\left[-\log\left(\frac{l(\lambda i)}{e(\lambda i)}\right) + \left(\frac{l(\lambda i)}{e(\lambda i)}\right)\right] \tag{3}$$

$$K(x) = (logx)^2,\ D[E,L] = \sum_{\lambda=1}^{\lambda n}[\log(l(\lambda i) - log(e(\lambda i))]^2 \tag{4}$$

$$K(x) = xlogx - x,\ then,\ D[E,L] = \sum_{\lambda=1}^{\lambda n} l(\lambda i)e(\lambda i)^{-1}\{\log(l(\lambda i)e(\lambda i)^{-1}) - 1\} \tag{5}$$

where, $D[E,L]$ is the distance between two functions, $E = \{e(\lambda i), \ldots, e(\lambda n)\}$ is the FSP_S2 or the SP_S2 datasets, $L = \{l(\lambda i), \ldots, l(\lambda n)\}$ is the PROSAIL simulated reflectance (LUT), and, $\lambda i, \ldots, \lambda n$ is the spectral bands.

The information measures represent measures of distance between two probability distributions. The reflectance is interpreted as a function of probability distribution [68]. In this class, the Bhattacharyya divergence [69] was tested:

$$D[E,L] = -\log+\{\sum_{\lambda i=1}^{\lambda n}\sqrt{e(\lambda i)l(\lambda i)} - \frac{1}{2}(e(\lambda i) + l(\lambda i))\} \tag{6}$$

The M-estimate measures rely on minimizing a dispersion function. Even though the M-estimates are the most commonly used measures, they can be misleading if their basic assumptions of residual normality and absence of extreme values are not met [69]. In this class, we tested the root mean square error (RMSE), which is calculated as:

$$\text{RMSE}: \ D[E,L] = \sqrt{\frac{\sum_{\lambda i=1}^{\lambda i=n}\left(e(\lambda i)-l(\lambda i)\right)^2}{n}} \quad (7)$$

Figure 4 presents the scheme of the LUT generation and the inversion process for LAI retrieval.

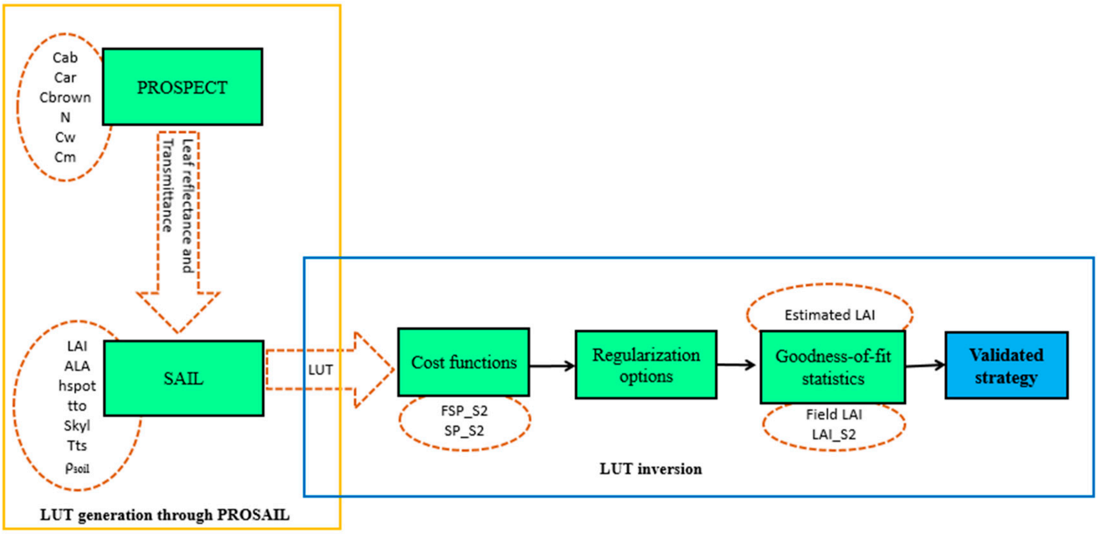


Figure 4. Look-up table (LUT) generation through PROSAIL and LUT inversion process for LAI retrieval.

2.5. LAI Model Calibration, Cross Validation, and External Validation

Data from 2016 were used for the calibration of statistical models (VI and MLRA) because it covered more crop phenological stages and a larger number of observations than the datasets of the other study years (2015 and 2018) (Table 2), resulting in high data variability.

In the VI approach, the calibration consists in a systematic assessment of all possible band combinations, vegetation indices formulations (Table 3), and curve fitting options. This process yields a table of goodness-of-fit measures for the values estimated by the calibrated model against the measured calibration measures [15]. A validation process is simultaneously performed. In this study, we used a k-fold cross validation method, which involves randomly dividing the dataset into k equal-sized sub-datasets. From these k sub-datasets, k − 1 sub-datasets are used as a training dataset and a single k sub-dataset is used as a validation dataset for model testing. The cross-validation process is then repeated k times, with each of the k sub-datasets used as a validation dataset [14]. The results from each of these iterative validation steps are then combined to derive a single estimation value. At the end, all data are used for both training and validation, and each single observation is used for validation exactly once [14,71]. In this study, we used k = 2 and run the calibration/validation process several times until no improvement of the goodness-of-fit statistics were detected in both the calibration and validation processes.

In the MLRA approach, we first used the Gaussian process regression band analysis tool (GPR-BAT) to identify the most sensitive bands to LAI and the minimum number of bands for an acceptable estimate accuracy [17]. Subsequently, we tested nine algorithms using the identified bands for each dataset. We equally used a k = 2 k-fold cross validation process and ran it several times until no improvement of the results were obtained.

For both VI and MLRA, an external validation was performed using a dataset comprising data collected in 2015 and 2018 to assess the model transferability. However, the external validation was not run for the SP_S2 dataset due to the limited number of observations (n = 22).

2.6. Model Assumptions, Accuracy Assessment, and Selection

The assumption of residual normal distribution was assessed by the Jarque-Bera (JB) test [72], and the Breusch-Pegan (BP) test [73] was used to investigate the presence of homoscedasticity, testing the dependence of the residuals' variance on the independent variables. For JB tests, the null hypothesis is that the variance of the residuals is normally distributed while for the BP, the null hypothesis is that the residuals are homogeneous. Thus, for *p*-values less than 5%, we reject the null hypothesis. The models' performance was assessed through three widely used statistical measures: The root mean square error (RMSE), the relative root mean square, and the coefficient of determination (R^2). These measures were compared within the different retrieval approaches (VI, MLRA, and LUT inversion) and between them to decide which approach performed better in estimating LAI in the study area.

3. Results

3.1. LAI Ground Measurements and Derived from Sentinel-2

Figure 5 shows the intra-annual and inter-annual variability of LAI and a comparison between Field_ LAI and LAI_S2 in the study area. The LAI values ranged from 0.23 to 2.68 (averaging 1.34) in 2015, 0.32 to 3.54 (averaging 1.63) in 2016, and 0.49 to 2.28 (averaging 1.38) in 2018.

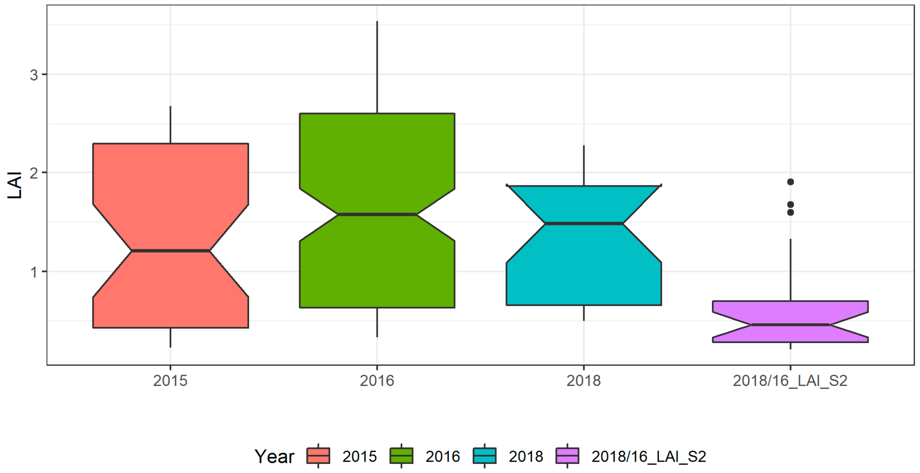


Figure 5. Boxplot of the field LAI for the three years: 2015 (n = 40), 2016 (n = 137), 2018 (n = 23), and derived from sentinel products, 2016/18_LAI_S2 (n = 22).

In 2016 the data collection covered the entire crop growing cycle (Table 2) and thus has an extended range and high average value of LAI, when compared with the data collected in the other two years. Also, the diversity between fields (sowing and irrigation conditions and crop treatments; Table 2) as well as the inter-annual weather conditions (Figure 2) likely contributed for the inter-annual LAI variability. The LAI_S2 that comprises data extracted from Sentinel-2 images acquired in 2016 and in 2018 ranged from 0.21 to 1.90 (averaging 0.61) and is very low if compared to the Field_LAI. A correlation between the Field_LAI and LAI_S2 yielded r = 0.69 (n = 30; *p*-value = 0.0001). This low correlation between measured and Sentinel derived LAI could be explained by several factors, including the scale mismatch and the fact that the LAI_S2 product was not validated in the context of our study area, which has very peculiar characteristics in terms of cropping systems (large in-field heterogeneity) with low LAI and soil background effects (soils with white colour). Also, the temporal mismatch between some LAI_S2 and Field_LAI data, due to cloud cover issues, may have also impacted the correlation between the two LAI products. Nevertheless, the S2 images' acquisition dates are in agreement with the same crop growing stages as those of field data collection dates, minimizing this time effect.

3.2. Vegetation Index Based LAI Models and Validation

3.2.1. Field Spectral Data Resampled to 10 nm (FSP_10)

The model (FSP_10 vs. Field_LAI) goodness-of-fit statistics of the three best performing VIs in the calibration, cross validation, and external validation dataset are shown in Table 6. The mSR_b and mDI_d fitted to a second degree polynomial function outperformed the other VI, showing consistent results for the LAI estimations when applied to the calibration, cross validation, and external validation datasets. For the external validation dataset, the mSR_b presents $R^2 = 0.80$, comparatively higher than the mDI_d ($R^2 = 0.62$), while the RMSE = 0.80 of mSR_b shows lower accuracy than the mDI_d with RMSE = 0.58. The three bands of VI, mSR_b, and mDI_d include bands within the same spectral regions of the green (565 nm) and red edge (715 nm, 725 nm, and 735 nm), suggesting reliability of the band optimization process for LAI estimation.

Figure 6 presents the agreement between the LAI measured and estimated with FSP_10 using the mSR_b and the mDI_d for the cross validation and external validation datasets. For the cross-validation dataset (Figure 6 left panel), the two indices exhibit good matching in the whole range of LAI values, but there is an underestimation for values greater than three. For the independent validation dataset (Figure 6 right panel), the two VIs estimate well the lower values of LAI, but for values above one, the estimation is poorer with an underestimation for the mDI_d and an overestimation for the mSR_b.

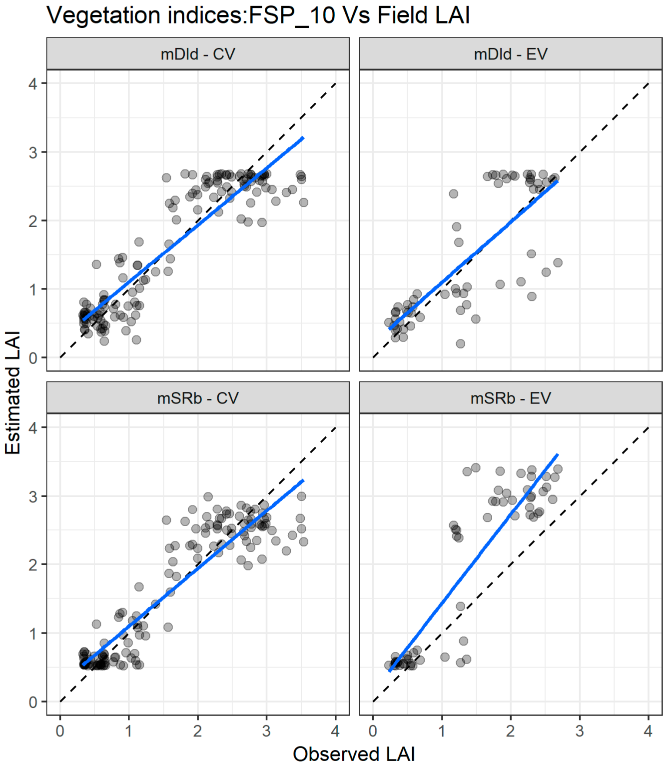


Figure 6. Comparisons between measured and estimated LAI with the two best performing spectral vegetation indices using the FSP_10 dataset for both cross validation (CV) and external validation (EV) datasets. mSR—modified simple ratio; mDI—modified difference index. The full line depicts the linear regression and the dashed line the linear regression through origin.

Table 6. Goodness-of fit statistics of the best performing vegetation indices (VI), using different combinations of spectral and leaf area index data.

VI	Bands	CA			CV			EV			Equation Parameters
		RMSE	R^2	NRMSE	RMSE	R^2	NRMSE	RMSE	R^2	NRMSE	
FSP_10 vs. Field_LAI											
Observations		137			137			63			
mSR_b	735;565;715	0.39	0.85	13.4	0.41	0.83	16.2	0.8	0.8	20.8	$a_2 = 0.07$; $a_1 = -1.056$; $a_0 = 4.65$
mDI_d	725;715;565	0.4	0.84	14.7	0.42	0.82	15.3	0.58	0.62	21.6	$a_2 = -319.73$; $a_1 = -64.78$; $a_0 = -0.59$
mSR_a	705;755	0.47	0.78	15.8	0.49	0.76	17.1	1.74	0.58	19.4	$m = -6.05$; $b = -0.17$
FSP_S2 vs. Field_LAI											
Observations		137			137			63			
mDI_c	705;740;865	0.41	0.83	12.6	0.42	0.83	13.4	0.77	0.71	22.4	$a_2 = -587.33$; $a_1 = -75.51$; $a_0 = 0.37$
mSR_c	740;705;865	0.4	0.84	12.4	0.43	0.82	14.7	0.62	0.82	23.6	$a_2 = -21.63$; $a_1 = 17.03$; $a_0 = -0.63$
mSR_b	842;783;705	0.43	0.82	14.9	0.44	0.8	15.3	0.97	0.75	21.8	$m = -6.87$; $b = 3.33$
SP_S2 vs. Field_LAI											
Observations		22			22						
$TBSI_b$	665;865;783	0.32	0.79	19.2	0.35	0.74	16.1	-	-		$a_2 = 166.6$; $a_1 = 199.5$; $a_0 = 59.9$
mDI_c	865;665;705	0.36	0.74	20.3	0.38	0.71	17.4	-	-		$m = 3.1$; $b = -0.76$
$TBSI_c$	865;665;783	0.36	0.73	20.5	0.38	0.71	19.8	-	-		$m = 9.6$; $b = 2.6$
SP_S2 vs. LAI_S2											
Observations		22			22						
$TBSI_b$	705;842;560	0.14	0.83	11.2	0.18	0.76	11.7	-	-		$k = -2.9$; $n = 0.83$
SR	665;783	0.18	0.73	12.4	0.19	0.72	11.5	-	-		$m = 0.20$; $b = 0.9$
$TBSI_a$	560;705;842	0.17	0.75	12.2	0.19	0.73	10.3	-	-		$m = -1.07$; $b = -0.16$

CV—cross validation; EV—external validation; FSP_10—field spectral data resampled to 10 nm; FSP_S2—field spectral data resampled to Sentinel-2 bands; SP_S2—Sentinel-2 spectral data (SP_S2); Field_LAI—field leaf area index; LAI_S2—Sentinel-2 leaf area index product data; mSR—modified simple ratio; mDI—modified difference index; SR—simple ratio; TBSI—three band spectral index; a_0, a_1, a_2—coefficients of the second degree polynomial function ($y = a_2X^2 + a_1X + a_0$); m, b—coefficients of the linear function ($y = mX + b$); k, n—coefficients of the exponential equation ($y = ne^{kx}$); m, b—coefficients of the linear function ($y = mX + b$).

3.2.2. Field Spectral Data Resampled to Sentinel-2 (FSP_S2)

The goodness-of-fit statistics of the three best performing VIs for LAI estimation in calibration, cross validation, and external validation with the FSP_S2 is also presented in Table 6 (FSP_S2 vs. Field_LAI). The mSR_c, based on the polynomial model with bands located at the red edge (705 nm, 740 nm) and near infrared (865 nm) region of the spectra, presents better LAI estimates with $R^2 \geq 0.82$ and RMSE of 0.40, 0.43, and 0.62, respectively, for the calibration, cross validation, and external validation datasets. The mDI_c presents very similar statistics for the calibration and cross validation datasets, but relatively lower R^2 (0.71) and higher RMSE (0.77) for the independent dataset. It is interesting to note that all the validated VIs derived from FSP_S2 were based on bands within the red edge and near infrared region of the spectra, which may suggest consistency of the band optimization and selection process.

Figure 7 depicts the agreement between measured and estimated LAI for the cross validation and external validation datasets. In the cross validation, there is a good agreement between measured and estimated LAI for the two indices, especially for LAI values below 2. As with the FSP_10 dataset, there is a miss match trend for LAI values higher than 2, which is more pronounced when the external dataset is used (Figure 7).

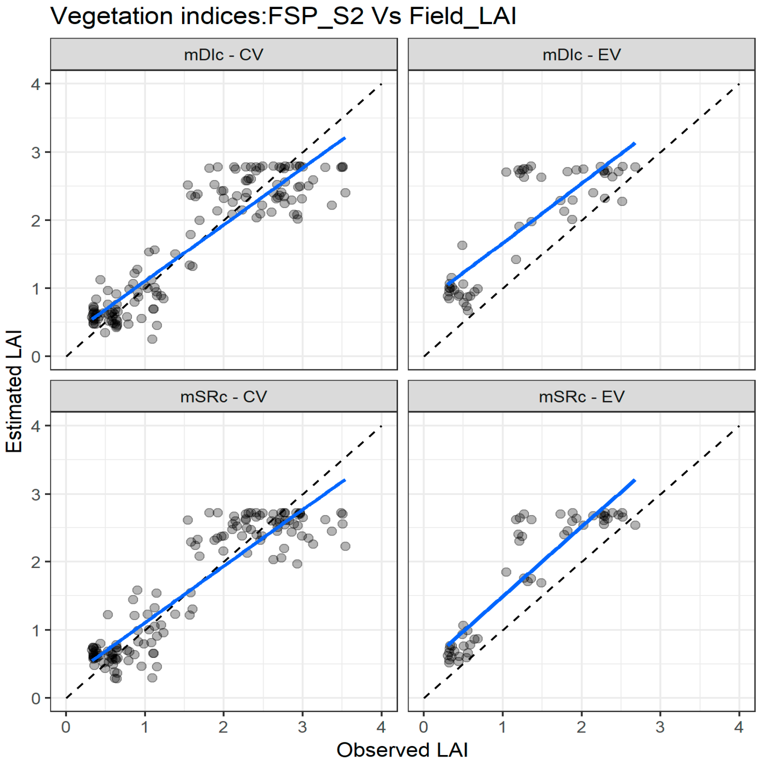


Figure 7. Comparisons of the measured and estimated LAI with the two best performing spectral vegetation indices using the FSP_S2 dataset for both cross validation (CV) and external validation (EV) datasets. mSR—modified simple ratio; mDI—modified difference index. The full line depicts the linear regression and the dashed line the linear regression through the origin.

3.2.3. Sentinel-2 Spectral Data (SP_S2)

The goodness-of-fit statistics of the three best performing VIs for LAI estimation in calibration and cross validation for both SP_S2 vs. Field_LAI and SP_S2 vs. LAI_S2 are also presented in Table 6. For the SP_S2 vs. Field_LAI combination, the selected VIs ($TBSI_b$, mDI_c, $TBSI_c$) present very similar statistics and all were constructed with the same spectral bands, centred at red (665 nm), the red edge (783 nm), and near infrared (865 nm). However, the $TBSI_b$ fitted to a second degree polynomial function outperforms the other VIs with an RMSE = 0.32, R^2 = 0.79 and RMSE = 0.35, R^2 = 0.74 for calibration and cross validation, respectively.

The SP_S2 vs. LAI_S2 combination presents excellent statistics for the three selected VIs ($TBSI_b$, SR, $TBSI_a$), with the $TBSI_b$ the outperforming one with an RMSE = 0.14, R^2 = 0.83 and RMSE = 0.18, R^2 = 0.76, respectively, for calibration and cross validation. This was expected because the SP_S2 was used as input data to derive the LAI_S2 products. However, careful interpretation is recommended given the lower correlation between the Field_LAI and the LAI_S2 (r = 0.69). The $TBSI_b$ was constructed with green (560 nm), red edge (705 nm), and near infrared (865 nm) bands.

Figure 8 depicts the agreement between measured and estimated LAI for the cross validation with the two combinations: SP_S2 vs. Field_LAI (upper panel) and SP_S2 vs. LAI_S2 (lower panel). Clearly, the SP_S2 vs. LAI_S2 combination presents better agreement with the measured and estimated LAI.

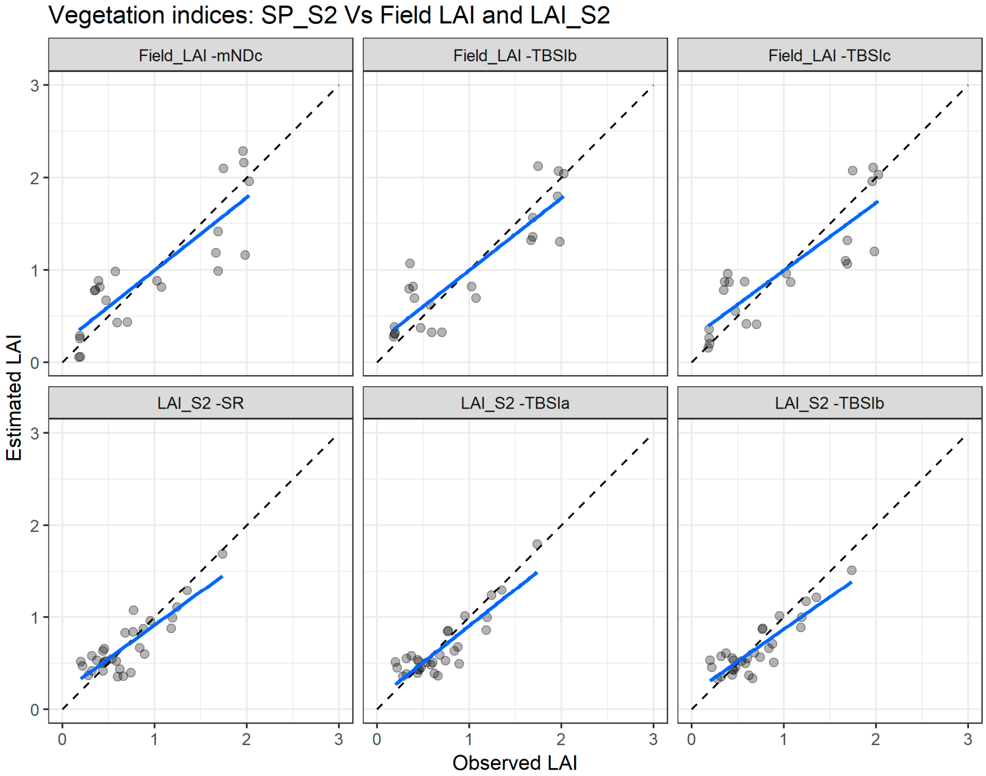


Figure 8. Comparisons of the measured and estimated LAI with the best performing vegetation indices using the Sentinel-2 spectral data (SP_S2) with the field leaf area index (Field_LAI) and with Sentinel-2 leaf area index product (LAI_S2). mNDc—modified difference index; TBSI—three band spectral index; SR—simple ratio. The blue line depicts the linear regression and the black dashed line the linear regression through origin.

3.3. Machine Learning Regression Based LAI Models and Validation

3.3.1. Field Spectral Data Resampled to 10 nm (FSP_10)

The LAI estimation based on MLRA was preceded by a band analysis tool in order to select the most sensitive bands among the initial number of 70 bands included in the FSP_10 dataset. The results indicated that increasing the number of bands in the models reduces the LAI estimation accuracy. Indeed, the maximum average R^2 (0.77) was obtained with three band models while the minimum average R^2 (0.70) was acquired with the total number of bands that comprises the hyperspectral dataset (70 bands) (Figure A1, Appendix A). The three selected bands are centered at 565 nm, 675 nm, and 775 nm, with a 10 nm spectral resolution, corresponding to green, red, and red edge regions of the spectra, respectively.

Table 7 presents the goodness-of-fit statistics of the three best performing MLRA for LAI estimation using the three above mentioned bands. The relevance vector machine (RVM) algorithm shows better consistency for LAI estimation in both cross validation and external validation datasets.

Table 7. Goodness-of-fit statistics for the LAI models of the best performing machine learning regression algorithms (MLRA) applied to different combinations of spectral and leaf area index.

MLRA/Type of Data	**CV**			**EV**		
	RMSE	**R^2**	**NRMSE**	**RMSE**	**R^2**	**NRMSE**
FSP_10 vs. Field_LAI						
Observations	137			63		
Support Vector Regression (SVM)	0.48	0.77	14.9	0.67	0.65	17.5
Variation Heteroscedastic Gaussian Processes Regression (VHGPR)	0.53	0.73	16.5	0.63	0.83	16.3
Relevance vector Machine (RVM)	0.54	0.72	16.7	0.5	0.67	13.9
FSP_S2 vs. Field_LAI						
Observations	137			63		
Support Vector Regression (SVM)	0.48	0.78	14.9	0.9	0.72	24.9
Relevance vector Machine (RVM)	0.52	0.73	26.3	0.53	0.62	15.9
Bagging trees (BaT)	0.63	0.63	19.5	0.64	0.72	19.9
SP_S2 vs. Field_LAI (n = 22)						
Observations	22			-		
Support Vector Regression (SVM)	0.51	0.52	27.6	-	-	
Random Forest (Fitensemble)	0.52	0.51	28.4	-	-	
Gaussian Processes Regression (GPR)	0.5	0.49	27.4	-	-	
SP_S2 vs. LAI_S2 (n = 22)						
Observations	22			-		
Random Forest (Fitensemble)	0.22	0.64	14.1	-	-	
Relevance vector Marchine (RVM)	0.23	0.61	14.8	-	-	
Support Vector Regression (SVM)	0.23	0.6	14.9	-	-	

CV—cross validation; EV—external validation; FSP_10—field spectral data resampled to 10 nm; FSP_S2—field spectral data resampled to Sentinel-2 bands; SP_S2—Sentinel-2 spectral data (SP_S2); Field_LAI—field leaf area index; LAI_S2—Sentinel-2 leaf area index product data.

With the cross validation dataset, the RVM has an RMSE = 0.54, which is higher than the values of RMSE of other MLRA; nevertheless, the RVM strongly outperform the other algorithms when used in the independent dataset with an RMSE = 0.50, which is clearly better compared to, for example, the SVM (RMSE = 0.67). The very high Gaussian process regression (VHGPR) also presented a good performance for the LAI estimation, both with the cross validation and external validation, although with a higher RMSE (0.63), when compared with the RVM algorithm for the external validation dataset.

Figure 9 depicts the comparison between measured and estimated LAI values for cross validation and external validation for the two best performing MLRA with the FSP_10 dataset. For cross validation (Figure 9 left panel), the two algorithms achieved good agreement in the entire LAI range, but there is a slight underestimation for LAI values above two. For the external validation dataset (Figure 9 right panel), the two algorithms exhibit good estimation in the full range of LAI values, although the VHGPR has lower accuracy for values exceeding two.

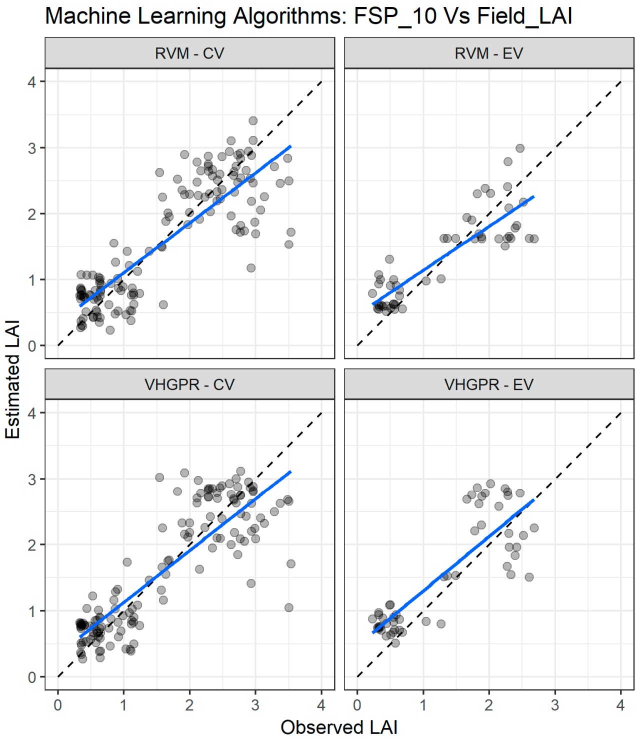


Figure 9. Comparisons of the LAI measured and estimated with the two best performing machine learning regression algorithms (MLRA) using the field spectra resampled to 10 nm (FSP_10). CV—cross validation, EV—external validation. RVM—relevance vector machine; VH-GPR—VH. Gaussian processes regression. The full line depicts the linear regression and the black dashed line the linear regression through origin.

3.3.2. Field Spectral Data Resampled to Sentinel-2 Bands (FSP_S2)

A band analysis was also performed prior to the MLRA assessment to select the best bands of the eight comprising the FSP_S2, for the LAI estimation. The results indicated that the LAI estimation accuracy increases from a minimum average of $R^2 = 0.51$, when only one band is involved, to a maximum average of $R^2 = 0.77$ when three bands are involved in the modelling (Figure A2, Appendix A). There was no accuracy improvement with additional band inclusion in the models. The selected bands match to the red and red edge Sentinel-2 bands, centered at 665 nm, 705 nm, and 783 nm.

The goodness-of-fit statistics of the three best performing MLRA when using the FSP_S2 dataset for LAI estimation are presented in Table 7. As with FSP_10, the RVM algorithm shows better consistency for both cross validation and external validation datasets. In the cross validation, the RVM has an RMSE = 0.52 and $R^2 = 0.73$, which is slightly poorer than the SVM values of RMSE = 0.48 and $R^2 = 0.78$. However, in the independent dataset, the RVM presents higher accuracy with RMSE = 0.53, compared to the SVM values of RMSE = 0.90. The bagging trees (BaT) algorithm is equally consistent, but with relatively higher RMSE (0.63 and 0.64) for the cross validation and external validation datasets, respectively.

Figure 10 presents the comparison between LAI measured and estimated for both cross validation and external validation datasets of the two best performing MLRA for the FSP_S2 dataset. For the cross validation (Figure 10 left panel), the two algorithms reveal good agreement between the measured and estimated LAI throughout the whole range, but the estimation accuracy decreases as the LAI values increase, resulting in a spread-out pattern for LAI values above two. For the external validation, the two algorithms present an overestimation trend in the full LAI range.

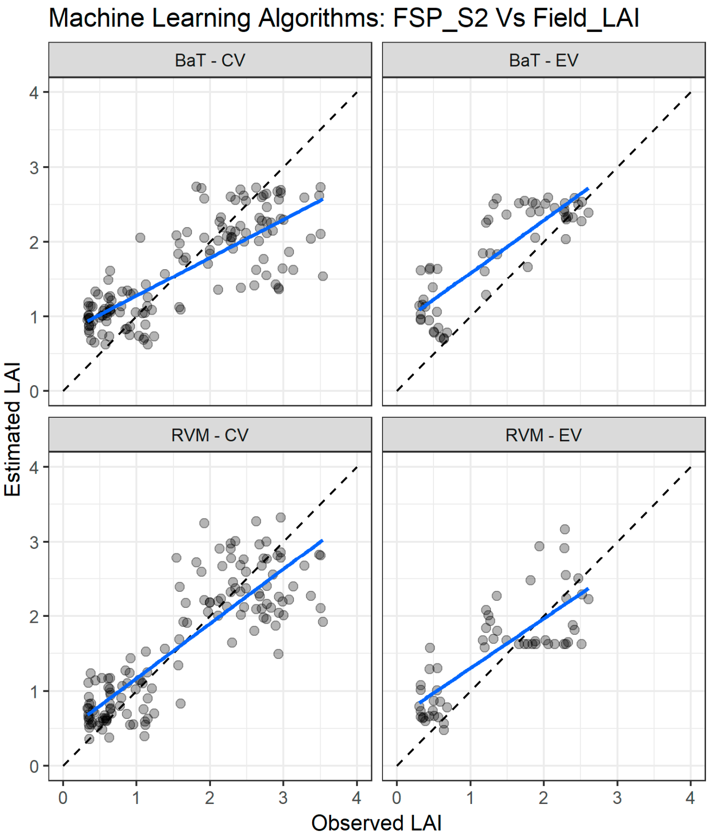


Figure 10. Comparisons of the LAI measured and estimated with the two best performing machine learning regression algorithms (MLRA) using the field spectra resampled to Sentinel-2 bands (FSP_S2). CV—cross validation, EV—external validation. RVM—relevance vector machine; BaT—bagging tree. The blue line depicts the linear regression and the black dashed line the linear regression through origin.

3.3.3. Sentinel-2 Spectral Data (SP_S2)

A band analysis performed preceding the MLRA assessment revealed that two bands centered at blue (490 nm) and red (665 nm) are the most appropriate for LAI estimation using the SP_S2 dataset. Inclusion of more bands decreased the estimation accuracy (Figure A3, Appendix A).

The goodness-of-fit statistics of the three best performing MLRA when using the SP_S2 dataset in combination with both Field_LAI and LAI_S2 are presented in Table 7. For the combination, SP_S2 vs. Field_LAI, the statistics are very poor for the three selected models: SVM (RMSE = 0.51, R^2 = 0.52), RTF (RMSE = 0.52, R^2 = 0.51), and GPR (RMSE = 0.50, R^2 = 0.49). However, as was expected, the combination, SP_S2 vs. LAI_S2, yielded very good statistics: RFF (RMSE = 0.22, R^2 = 0.64), RVM (RMSE = 0.23, R^2 = 0.61), and SVM (RMSE = 0.23, R^2 = 0.60).

Figure 11 presents the comparison between LAI measured and estimated for cross validation for the two best performing MLRA with the combinations, SP_S2 vs. Field_LAI (upper panel) and SP_S2 vs. LAI_S2 (lower panel). For the SP_S2 vs. Field_LAI, the SVM algorithm shows better agreement while the RFF presents the best agreement for the SP_S2 vs. LAI_S2.

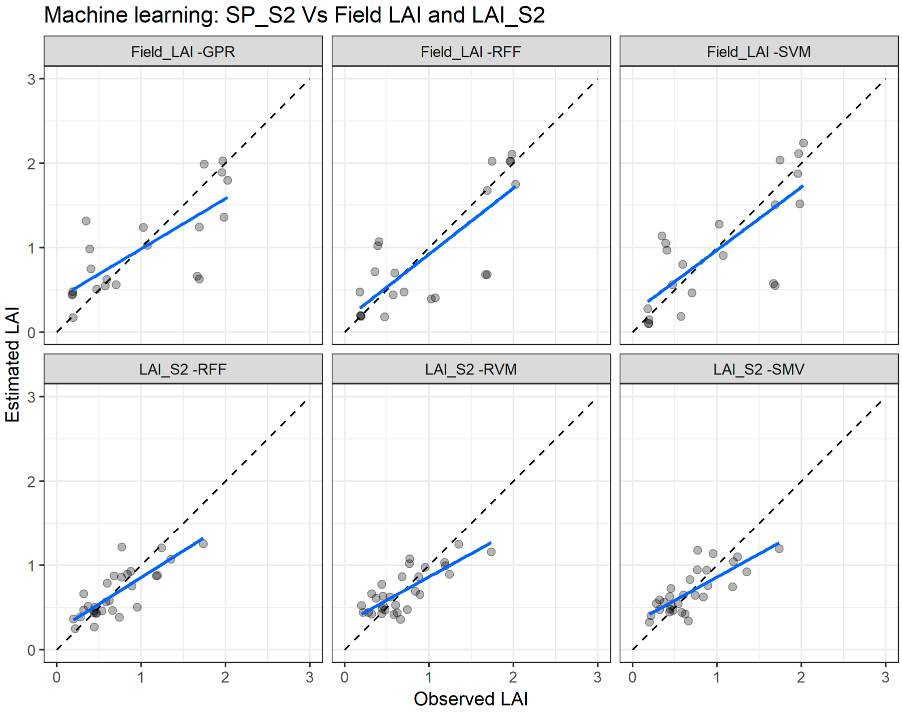


Figure 11. Comparisons of the LAI measured and estimated with the best performing machine learning regression algorithms (MLRA) using the Sentinel-2 spectral data (SP_S2) with the field leaf area index (Field_LAI) and the Sentinel-2 leaf area index product (LAI_S2). GPR—Gaussian process regression; SVM—support vector regression; RFF—random forest Fitensemble; RVM—relevance vector machine. The full line depicts the linear regression and the dashed line the linear regression through the origin.

3.4. LUT Inversion Based LAI Estimation and Validation

In this section, we present the goodness-of-fit statistics of LAI estimation by the LUT inversion approach using three different combinations of spectral and LAI datasets: (i) FSP_S2 and the field LAI (FSP_S2 vs. Field LAI); (ii) spectral data of Sentinel-2 images and the LAI product derived from Sentinel-2 (SP_S2 vs. LAI_S2); and (iii) SP_S2 combined with Field LAI (SP_S2 vs. Field_LAI).

LUT Inversion

Table 8 presents the goodness-of-fit statistics of LAI retrieval through the LUT inversion using the three combinations of datasets. For the three datasets, three cost functions, namely the $K(x) = (\log(x))^2$, the $K(x) = x(\log(x)) - x$, and the Bhattacharyya divergence, evidenced good LAI estimation. Nevertheless, the highest LAI estimation accuracy was achieved applying the Bhattacharyya divergence cost function to the SP_S2 vs. LAI_S2 dataset, resulting in an RMSE = 0.20 and $R^2 = 0.70$. Furthermore, this dataset combination showed better performance than the others with all the evaluated cost function. However, the SP_S2 vs. LAI_S2 dataset produced relatively lower association measure values, $0.65 \leq R^2 \leq 0.72$, compared to the range of $0.82 \leq R^2 \leq 0.86$ for FSP_S2 vs. Field LAI.

Figure 12 illustrates the agreement between the measured and estimated LAI values with the two best performing cost functions and using the three datasets. With the FSP_S2 vs. Field LAI dataset, the two cost functions present a strong underestimation of the LAI for values higher than one, while the estimation for the LAI values below one is more accurate (Figure 11 first two plots). The underestimation is also evident when the SP_S2 vs. Field LAI is used with all the cost functions (Figure 12, third and fourth plots). On the other hand, for the SP_S2 vs. LAI_S2, the two cost functions accurately estimate the LAI throughout the entire range (Figure 12, last two plots).

Table 8. Goodness-of-fit statistics of look-up-table (LUT) inversion using the field spectral data aggregated according to Sentinel-2 bands and the field LAI (FSP_S2 vs. Field LAI); the spectral data of Sentinel-2 images and the LAI product derived from Sentinel-2 (SP_S2 vs. LAI_S2); the spectral data of Sentinel-2 images and the field LAI (SP_S2 vs. Field LAI), while considering different cost function (CF) algorithms.

CF Algorithm	FSP_S2 vs. Field LAI (n = 63)			SP_S2 vs. LAI_S2 (n = 22)			SP_S2 vs. Field LAI (n = 22)		
	RMSE	R^2	NRMSE	RMSE	R^2	NRMSE	RMSE	R^2	NRMSE
$K(x) = (\log(x))^2$	0.43	0.82	18.2	0.28	0.71	17.9	0.53	0.6	28.9
$K(x) = x(\log(x)) - x$	0.49	0.82	20.7	0.24	0.7	15.4	0.56	0.6	30.6
Bhattacharyya divergence	0.61	0.83	25.9	0.2	0.7	12.9	0.62	0.7	33.9
RMSE	0.85	0.86	36.1	0.26	0.65	16.9	0.77	0.7	41.8
$K(x) = -\log(x) + x$	1.09	0.85	46.2	0.41	0.72	26.6	0.92	0.6	49.7

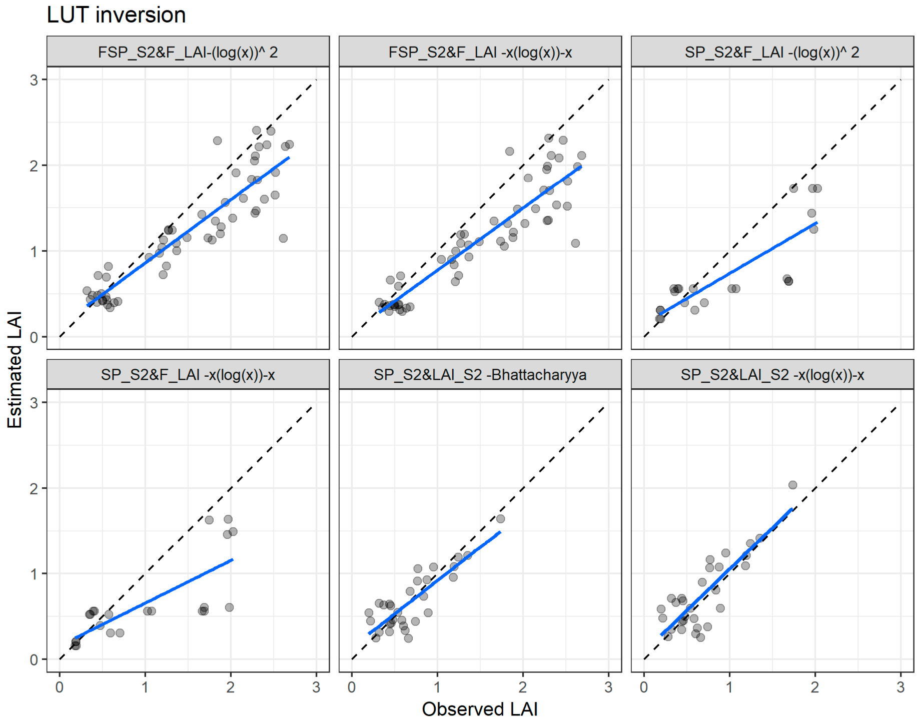


Figure 12. Comparisons of the measured and estimated LAI of the best LUT inversion cost functions using the field spectral data resampled to Sentinel-2 bands with the field leaf area index (FSP_S2 vs. F_LAI—first two plots), Sentinel-2 spectral data with field leaf area index data (SP_S2 vs. F_LAI—third and fourth plots), and Sentinel-2 spectral data with Sentinel-2 leaf area index product data (SP_S2 vs. LAI_S2—last two plots). The full line depicts the linear regression and the dashed line the linear regression through origin.

The Breusch-Pegan (BP) test *p*-value is reasonably high for almost all the datasets and modelling approaches, excluding the MLRA (BaT), for FSP_S2 and VI (mNDc and TBSIc) for SP_S2 vs. Field_LAI, all for cross validation, confirming the presence of homoscedasticity. The residuals' normality was confirmed from the Jarque-Bera (JB) test, showing higher values than *p*-values for all the developed models. The results of BP and JB tests are included in Table A2 (Appendix A).

4. Discussion

The best performing VIs were constructed with wavelengths centered at visible (green—565 nm) and red edge (715, 725, 735 nm) when using the FSP_10 vs. Field_LAI dataset; red edge (705, 740 nm) and near infrared (865 nm) when using the FSP_S2 vs. Field_LAI dataset; red (665 nm), red edge

(783 nm), and near infrared (865 nm) when using the SP_S2 vs. Field_LAI; and green (560 nm), red (665 nm), red edge (705, 783 nm), and near infrared (842 nm) when using the SP_S2 vs. LAI_S2. These findings are similar to those of previous studies for estimating LAI in various crops, e.g., [74] used red edge bands (690–710 nm and 750–900 nm), [75] combined green (580 nm), red edge (700 nm and 710 nm), and near infrared wavelengths (1003 nm); [59] also combined green (550 nm), red (670 nm), and near infrared (800 nm) wavelengths; [76] considered blue (460–480 nm), green (545–565 nm), or red edge (700–710 nm), and red (660–680 nm) wavelengths; and [77] considered red (665 nm), red edge (705, 740 and 783 nm), and near infrared wavelengths (842 nm). Additionally, and as equally found by [14], the three band VIs outperformed the two band VIs formulation for all the datasets.

Regarding the MLRA, our best performing models (RVM, VHGPR, BaT, SVM, and RFF) were equally amongst the best performing models in [14]. Similarly to the VIs, the MLRA integrated bands centered at green (565 nm), red (675 nm), and red edge (775 nm) for the FSP_10 vs. Field_LAI; red (665 nm), red edge (705 and 783 nm) for the FSP_S2 vs. Field_LAI; and blue (490 nm) and red (665 nm) for the SP_S2 vs. Field_LAI and SP_S2 vs. LAI_S2.

Generally, there was no substantial difference in LAI estimation accuracy between the two types of datasets, hyperspectral (FSP_10) and multispectral (FSP_S2) for both VIs and MLRA approaches. In fact, the advantage of hyperspectral data over multispectral data in estimating LAI remains a matter of debate [78]. Theoretically, the high spectral resolution of hyperspectral data disclose the spectral details obscured with multispectral data for LAI estimation [79]. However, the LAI insensitive bands included in hyperspectral data require additional computational time and distort the accuracy of LAI retrieval, with this the reason why there is a need for dimensionality reduction of hyperspectral data [78].

The performance of LUT inversion for LAI retrieval is in accordance with the findings of other studies for maize using the same and other retrieval methods, with RMSE values of 0.40–0.43 [2], 0.46–1.21 [59], 0.73 [80], 0.41–0.76 [81], and 0.63 [82]. The underestimation trend observed while using the FSP_S2 dataset for the inversion was also found in maize by [2]. The reason could be related with the row planting pattern of maize, which diverges from the turbid assumption of the PROSAIL model [83]. In the case of the present study, this could be exacerbated by the low input cropping system of the field area and important differences in the planting geometry (Table 1), which resulted in high heterogeneity within and between the sampled fields.

Concerning the applied cost functions, our findings show that the commonly used RMSE was not the best performing cost function, but instead the contrast function, ($K(x) = (\log(x))^2$), when inverting through the field spectral data (FSP_S2), and the Bhattacharyya divergence cost function, when using the SP_S2 dataset for the inversion. Similar findings were reported by [68] for LAI.

Table 9 summarizes the RMSE and slope (b) of the best models identified in the different combinations of spectral and LAI data and modelling approaches tested.

For all the modelling approaches, the combination, SP_S2 vs. LAI_S2, yielded the best accuracy, which was, in fact, expected because the SP_S2 is an input for the derivation of LAI_S2 products. The comparison of combinations involving Field_LAI shows that the statistical approach based in VIs using the SP_S2 vs. Field_LAI yielded the most accurate LAI estimation (RMSE = 0.35 and b = 0.82), outperforming the physically-based approach of LUT inversion that is often considered more robust. These results may be due to the small geographical scale (field level) of our study and involving a single crop type. According to [78], statistical predictive LAI models are based on numerical relationships, and thus rely greatly on the specific location, including the crop's condition and soil background reflectance, therefore, they are more suitable for small-scale studies. Additionally, the parameterization of PROSAIL for the application of LUT inversion may have been hindered by the large heterogeneity within and between crop fields in the study area.

Table 9. Multi-comparison of LAI measured and estimated based on different combinations of modelling approaches and remote sensing validation data-set.

Spectral Data	Modelling Approach	Field LAI		LAI_S2	
		RMSE	b *	RMSE	b *
FSP_10	VI	0.42	0.99	-	-
FSP_S2	VI	0.43	1.0	-	-
FSP_10	MLRA	0.54	0.95	-	-
FSP_S2	MLRA	0.52	0.99	-	-
FSP_S2	LUT	0.43	1.11	-	-
SP_S2	VI	0.35	0.82	0.18	0.80
SP_S2	MLRA	0.51	0.78	0.22	0.62
SP_S2	LUT	0.53	1.20	0.20	0.88

FSP_10—field spectral data resampled to 10 nm; FSP_S2—Field spectral data resampled to Sentinel-2 bands; Sentinel-2 spectral data (SP_S2); Field_LAI—field leaf area index; LAI_S2—Sentinel-2 leaf area index product data; VI—vegetation index; MLRA—machine learning regression algorithm; * b coefficient of the linear regression between measured and predicted LAI.

5. Conclusions

In this paper, we successfully calibrated and validated models for maize LAI estimation in low-input crop systems based on statistical and physical approaches. Hyperspectral and multispectral data, obtained both from field and satellite sensors, were tested for the LAI retrieval and further compared with field and satellite LAI data. The robustness of the developed models was indicated by the consistency of the selected electromagnetic band regions, whichever the modelling approach or dataset combination was applied. Additionally, the models reasonably copedwith transferability issues by adjusting to relatively different cropping systems (Fields 1–4) and different weather conditions (2015, 2016, 2018), as suggested by the results of the external validation process.

The most accurate model involves the $TBSI_b$ spectral index. This VI is built with three bands centered in the red, red edge, and near infrared regions of the electromagnetic spectrum with widely known biophysical significance for LAI estimation. The hyperspectral data (aggregated to 10 nm) failed to improve the LAI estimation accuracy comparatively to Sentinel-2 multispectral data. This finding is of particular relevance for the operational application of spectral data in crop monitoring, though Sentinel-2 data is freely available and presents good spatial and spectral resolutions. However, future research should consider using field LAI data acquired with high precision equipment, including other crop types and extensive sampling, in order to increase the ground truth data and, as a result, improve the accuracy of LAI retrieval.

Author Contributions: Conceptualization, S.M., I.P. and M.C.; methodology, S.M.; validation, I.P. and M.C.; formal analysis, S.M.; investigation, S.M.; writing—original draft preparation, S.M.; writing—review and editing, I.P. and M.C.; visualization, S.M.; supervision, I.P. and M.C.

Funding: This research received no external funding.

Acknowledgments: Sosdito Mananze acknowledge the Joint Aid Management (JAM-Mozambique) for making available the crop fields for data collection; the Gulbenkian Foundation for the Ph.D. scholarship through the grant number 140227. Isabel Pôças acknowledges Portuguese Foundation for Science and Technology (FCT) for the Post-Doc research grant (SFRH/BPD/79767/2011) and for the Post-Doctoral grant of the project ENGAGE-SKA POCI-01-0145-FEDER-022217, co-funded by FEDER through COMPETE (POCI-01-0145-FEDER-022217).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

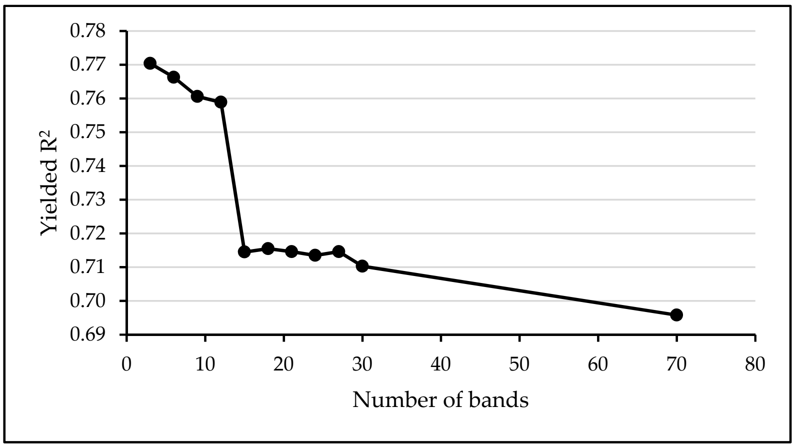


Figure A1. Variation of the models' R^2 as a function of the number of bands involved in machine learning modelling using the field spectral data resampled to 10 nm (FSP_10). The highest R^2 was achieved using three bands centered at 565 nm, 675 nm, and 775 nm.

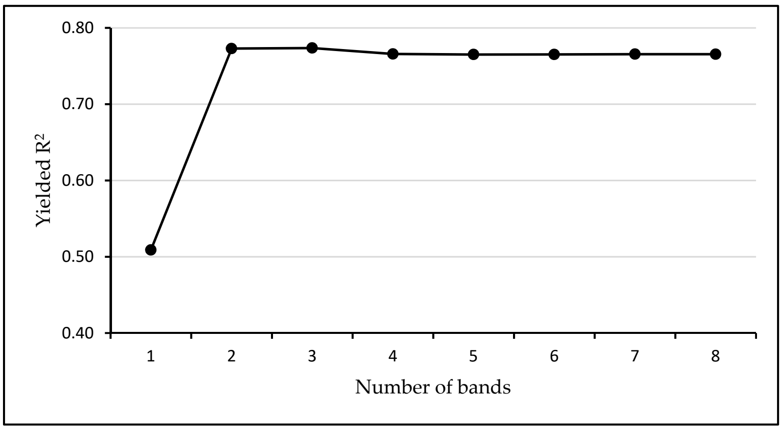


Figure A2. Variation of the models' R^2 as a function of the number of bands involved in machine learning modelling using the field spectral data resampled to Sentinel-2 bands (FSP_S2). The highest R^2 was achieved using three bands: Centered at 665 nm, 705 nm, and 783 nm.

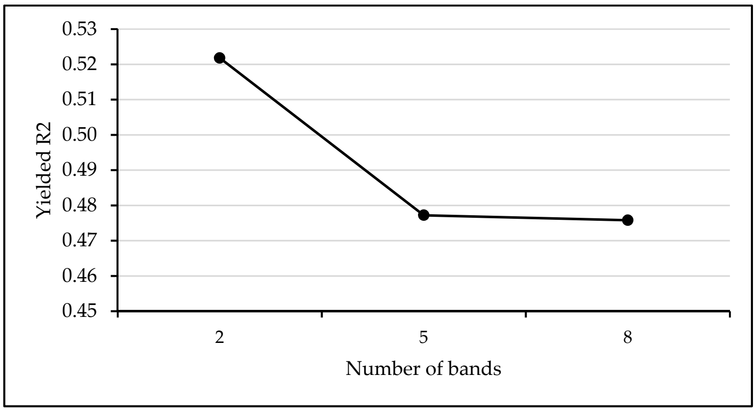


Figure A3. Variation of the models' R^2 as a function of the number of bands involved in machine learning modelling using the spectral data from Sentinel-2 images (SP_S2). The highest R^2 was achieved using two bands: Centered at 490 nm and 665 nm.

Table A1. Summary of PROSPECT and PROSAIL parameters as used in different studies.

Authors	[56]	[67]	[84]	[85]	[2]	[86]	[14]
Model: PROSPECT 5							
Ca+b: Clorofila a+b ($\mu g/cm^2$)	30–70	20–80	15–55	5–70	10–70	10–70	05–75
Equivalent water thickness (g/cm^2)	0.01–0.06	0.01–0.04	0.01–0.02	-	-	0.01–0.03	0.002–0.05
N: Leaf Structural Parameter	1–3	1	1.5–1.9	1.5	1.3–1.7	1–1.6	1.3–2.5
Car: Carotenoids ($\mu g/cm^2$)	-	1	-	-	-	-	-
Cbrown: Brown pigments (g/cm^2)	-	0.05	-	-	-	0–2	-
Cm: Dry matter content (g/cm^2)	0.008–0.025	0.0046	0.005–0.01	0.009	0.004–0.007	0.005–0.021	0.001–0.03
Model: 4SAIL							
LAI: Leaf area Index	1–7	0.1–6	0.3–7.5	0–8	0–6	0–7	0.1–7
ALA: Leaf angle distribution (°)	20–60	70, 57, 45	40–70	35	40–70	40–70	40–70
skyl: Diffuse/Direct light	-	0.1	-	-	-	10	0.05
psoil: Soil Coefficient	-	0.1	0.5–1.5	-	0.7–1.3	0–1	0–1
hspot: Hot spot	0.5/LAI	0.78, 0.40, 0.32	0.05–0.1	0.01	0.05–1	1–1.6	0.05–0.5
tts: Solar Zenit Angle (°)	−20–+80	51, 45, 33	-	30	35	20–50	22.3
tto: Observer zenit Angle (°)	0–55	0	-	10	0	0	20.19
psi: Azimut Angle (°)	−120–+120	0	-	-	0	0	0
Crop/vegetation type	Wheat	Wheat	Rangelands	Maize, vegetables, sunflower, alfafa and vine	Maize and sugar beet	Maize, vegetables and alfafa	Maize, vegetables, alfafa, sunflower, vines

Table A2. Diagnostic of homoscedasticity and residual normality.

Dataset	Modelling Approach		CV		EV	
			BP	JB	BP	JB
FSP_10	VI	mSRb	0.95	13.4	0.12	6.18
		mDId	0.2	3.15	0.28	6.98
	MLRA	RVM	0.17	12.42	0.07	7.29
		VHGPR	0.54	12.42	0.3	7.29
FSP_S2	VI	mDIc	0.08	5.9	0.17	4.9
		mSRc	0.52	13.59	0.81	2.89
	MLRA	BaT	0.02	12.42	0.41	7.29
		RVM	0.13	12.42	0.2	7.29
SP_S2 vs. Field_LAI	VI	mNDc	0.004	0.79		
		TBSIb	0.65	6.76		—
		TBSIc	0.0003	1.89		
	MLRA	GPR	0.41	2.99		
		RFF	0.47	2.99		—
		SVM	0.18	2.99		
SP_S2 vs. LAI_S2	VI	SR	0.77	1.22		
		TBSIa	0.88	5.68		—
		TBSIb	0.88	1.83		
	MLRA	RFF	0.65	10.5		
		RVM	0.51	10.5		—
		SMV	0.39	10.5		

FSP_10—field spectral data resampled to 10 nm; FSP_S2—Field spectral data resampled to Sentinel-2 bands; Sentinel-2 spectral data (SP_S2); Field_LAI—field leaf area index; LAI_S2—Sentinel-2 leaf area index product data; VI—vegetation index; MLRA—machine learning regression algorithm; SR—Simple Ratio; mSR—Modified Simple Ratio; mND—Modified Normalized Difference; TBSI—Three Band Spectral Index; mDI—Modified Difference Index; CV—cross validation; EV—external validation; RVM—Relevance Vector Machine; VHGPR—Variation Heteroscedastic Gaussian process regression; BaT—Bagging Trees; GPR—Gaussian process regression; RFF—Random Forest (Fitensemble); SVM—Support Vector Machine; BP—Breusch-Pegan; JB—Jarque-Bera.

References

1. Mananze, S.E.; Pôças, I.; Cunha, M. Maize leaf area estimation in different growth stages based on allometric descriptors. *Afr. J. Agric. Res.* **2018**, *13*, 202–209. [CrossRef]
2. Richter, K.; Atzberger, C.; Vuolo, F.; Weihs, P.; D'Urso, G. Experimental assessment of the Sentinel-2 band setting for RTM-based LAI retrieval of sugar beet and maize. *Can. J. Remote Sens.* **2009**, *35*, 230–247. [CrossRef]
3. Cattanio, J.H. Leaf area index and root biomass variation at different secondary forest ages in the eastern Amazon. *For. Ecol. Manag.* **2017**, *400*, 1–11. [CrossRef]
4. Fang, H.; Ye, Y.; Liu, W.; Wei, S.; Ma, L. Continuous estimation of canopy leaf area index (LAI) and clumping index over broadleaf crop fields: An investigation of the PASTIS-57 instrument and smartphone applications. *Agric. For. Meteorol.* **2018**, *253–254*, 48–61. [CrossRef]
5. Delegido, J.; Verrelst, J.; Rivera, J.P.; Ruiz-Verdú, A.; Moreno, J. Brown and green LAI mapping through spectral indices. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35*, 350–358. [CrossRef]
6. Houborg, R.; McCabe, M.F.; Gao, F. A Spatio-Temporal Enhancement Method for medium resolution LAI (STEM-LAI). *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *47*, 15–29. [CrossRef]
7. Mokhtari, A.; Noory, H.; Vazifedoust, M. Improving crop yield estimation by assimilating LAI and inputting satellite-based surface incoming solar radiation into SWAP model. *Agric. For. Meteorol.* **2018**, *250–251*, 159–170. [CrossRef]
8. Thenkabail, P.; Lyon, J.; Huete, A. Advances in Hyperspectral Remote Sensing of Vegetation and Agricultural Croplands. In *Hyperspectral Remote Sensing of Vegetation*; CRC Press: Boca Raton, FL, USA, 2011; pp. 3–36. [CrossRef]

9. Thenkabail, P.S.; Enclona, E.A.; Ashton, M.S.; Legg, C.; De Dieu, M.J. Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. *Remote Sens. Environ.* **2004**, *90*, 23–43. [CrossRef]
10. Thenkabail, P.S.; Enclona, E.A.; Ashton, M.S.; Van Der Meer, B. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sens. Environ.* **2004**, *91*, 354–376. [CrossRef]
11. Thenkabail, P.S.; Smith, R.B.; De Pauw, E. Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sens. Environ.* **2000**, *71*, 158–182. [CrossRef]
12. Jones, H.G.; Vaughan, R.A. *Remote Sensing of Vegetation: Principles, Techniques, and Applications*; Oxford University Press: New York, NY, USA, 2010; p. 353.
13. Roberts, D.A.; Roth, K.L.; Perroy, R.L. Hyperspectral vegetation indices. In *Hyperspectral Remote Sensing of Vegetation*; Thenkabail, P.S., Lyon, J.G., Huete, A., Eds.; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA, 2012; pp. 309–327.
14. Verrelst, J.; Rivera, J.P.; Veroustraete, F.; Muñoz-Marí, J.; Clevers, J.G.P.W.; Camps-Valls, G.; Moreno, J. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods—A comparison. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 260–272. [CrossRef]
15. Rivera, J.; Verrelst, J.; Delegido, J.; Veroustraete, F.; Moreno, J. On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization. *Remote Sens.* **2014**, *6*, 4927–4951. [CrossRef]
16. Pôças, I.; Gonçalves, J.; Costa, P.M.; Gonçalves, I.; Pereira, L.S.; Cunha, M. Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 177–190. [CrossRef]
17. Caicedo, J.P.R.; Verrelst, J.; Muñoz-Marí, J.; Moreno, J.; Camps-Valls, G. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1249–1259. [CrossRef]
18. Darvishzadeh, R.; Skidmore, A.K.; Schlerf, M.; Atzberger, C. Inversion of a radiative transfer model for estimating vegetation LAIA and chlorophyll in a heterogeneous grassland. *Remote Sens. Environ.* **2008**, *112*, 2592–2604. [CrossRef]
19. Jacquemoud, S.; Baret, F. PROSPECT: A model of leaf optical properties spectra. *Remote Sens. Environ.* **1990**, *34*, 75–91. [CrossRef]
20. Feret, J.-B.; François, C.; Asner, G.P.; Gitelson, A.A.; Martin, R.E.; Bidel, L.P.R.; Ustin, S.L.; le Maire, G.; Jacquemoud, S. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sens. Environ.* **2008**, *112*, 3030–3043. [CrossRef]
21. Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P.J.; Asner, G.P.; François, C.; Ustin, S.L. PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sens. Environ.* **2009**, *113*, S56–S66. [CrossRef]
22. Mariotto, I.; Thenkabail, P.S.; Huete, A.; Slonecker, E.T.; Platonov, A. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission. *Remote Sens. Environ.* **2013**, *139*, 291–305. [CrossRef]
23. The World Bank. Climate Change Knowledge Portal for Development Practitioners and Policy Makers. Available online: http://sdwebx.worldbank.org/climateportal/index.cfm?page=country_historical_climate&ThisRegion=Africa&ThisCCode=MOZ (accessed on 7 September 2018).
24. Abendroth, L.; Elmore, R.; Boyer, M.; Marlay, S. *Corn Growth and Development*; Iowa State University Extension: Ames, IA, USA, 2011.
25. Clevers, J.G.P.W.; Gitelson, A.A. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 344–351. [CrossRef]
26. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [CrossRef]
27. Malenovský, Z.; Rott, H.; Cihlar, J.; Schaepman, M.E.; García-Santos, G.; Fernandes, R.; Berger, M. Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sens. Environ.* **2012**, *120*, 91–101. [CrossRef]
28. Vuolo, F.; Żółtak, M.; Pipitone, C.; Zappa, L.; Wenng, H.T.; Immitzer, M.; Weiss, M.; Baret, F.; Atzberger, C. Data Service Platform for Sentinel-2 Surface Reflectance and Value-Added Products: System Use and Examples. *Remote Sens.* **2016**, *8*, 938. [CrossRef]

29. Pan, H.; Chen, Z.; Ren, J.; Li, H.; Wu, S. Estimating winter wheat leaf area index and canopy water content with three different approaches using Sentinel-2 multi-spectral instrument (S2 MSI) data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *99*, 1–11.
30. Verrelst, J.; Romijn, E.; Kooistra, L. Mapping Vegetation Density in a Heterogeneous River Floodplain Ecosystem Using Pointable CHRIS/PROBA Data. *Remote Sens.* **2012**, *4*, 2866–2889. [CrossRef]
31. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the Third Earth Resources Technology Satellite Symposium, Washington, DC, USA, 10–14 December 1973; p. 9.
32. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soiladjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 12. [CrossRef]
33. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [CrossRef]
34. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [CrossRef] [PubMed]
35. Roujean, J.-L.; Breon, F.-M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [CrossRef]
36. Gitelson, A.A.; Chivkunova, O.B.; Merzlyak, M.N. Non destructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. *Am. J. Bot.* **2009**, *96*, 7. [CrossRef]
37. Broge, N.H.; Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **2001**, *76*, 156–172. [CrossRef]
38. Huete, A.R.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [CrossRef]
39. Gitelson, A.A.; Stark, R.; Grits, U.; Rundquist, D.; Kaufman, Y.; Derry, D. Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction. *Int. J. Remote Sens.* **2002**, *23*, 2537–2562. [CrossRef]
40. Penuelas, J.; Llusia, J.; Pinol, J.; Filella, I. Photochemical reflectance index and leaf photosynthetic radiation-use-efficiency assessment in Mediterranean trees. *Int. J. Remote Sens.* **1997**, *18*, 2863–2868. [CrossRef]
41. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [CrossRef]
42. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; de Colstoun, E.B.; McMurtrey, J.E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [CrossRef]
43. Merton, R.; Huntington, J. Early simulation results of the ARIES-1 satellite sensor for multi-temporal vegetation research derived from AVIRIS. In Proceedings of the Eighth Annual JPL Airborne Earth Science Workshop, Pasadena, CA, USA, 8–14 February 1999.
44. Wang, W.; Yao, X.; Yao, X.; Tian, Y.; Liu, X.; Ni, J.; Cao, W.; Zhu, Y. Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crop. Res.* **2012**, *129*, 90–98. [CrossRef]
45. Tian, Y.C.; Gu, K.J.; Chu, X.; Yao, X.; Cao, W.X.; Zhu, Y. Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. *Plant Soil* **2013**, *376*, 193–209. [CrossRef]
46. Hastie, T.; Tibshirani, R.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009.
47. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [CrossRef]
48. Breiman, L.; Friedman, J.; Stone, C.; Olshen, R. *Classification and Regression Trees*; Chapman&Hall/CRC: Boca Raton, FL, USA, 1984.
49. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
50. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]

51. Tipping, M.E. Sparse Bayesian Learning and the Relevance Vector Machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244. [CrossRef]
52. Suykens, J.A.K.; Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]
53. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; The MIT Press: New York, NY, USA, 2006.
54. Lazaro-Gredilla, M.; Titsias, M.; Verrelst, J.; Camps-Valls, G. Retrieval of biophysical parameters with heteroscedastic gaussian processes. *IEEE Geosci. Remote Sens. Lett.* **2013**, *99*, 1–5. [CrossRef]
55. Chapelle, O.; Vapnik, V. Model Selection for Support Vector Machines. In Proceedings of the 12th International Conference on Neural Information Processing Systems, Denver, CO, USA, 29 November–4 December 2000; pp. 230–236.
56. Nigam, R.; Bhattacharya, B.K.; Vyas, S.; Oza, M.P. Retrieval of wheat leaf area index from AWiFS multispectral data using canopy radiative transfer simulation. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *32*, 173–185. [CrossRef]
57. Clevers, J.G.P.W.; Kooistra, L.; Schaepman, M.E. Estimating canopy water content using hyperspectral remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 119–125. [CrossRef]
58. Gonzalez-Sanpedro, M.C.; Toan, T.L.; Moreno, J.F.; Kergoat, L.; Rubio, E. Seasonal variations of Leaf Area Index of agricultural fields retrieved from Landsat data. *Remote Sens. Environ.* **2008**, *112*, 810–824. [CrossRef]
59. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejadad, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2003**, *90*, 337–352. [CrossRef]
60. Reddy, G.S.; Rao, C.L.N.; Venkataratnam, L.; Rao, P.V.K. Influence of plant pigments on spectral reflectance of maize, groundnut and soybean grown in semi-arid environments. *Int. J. Remote Sens.* **2001**, *22*, 3373–3380. [CrossRef]
61. Sanchez, R.A.; Hall, A.J.; Trapani, N.; de Hunau, R.C. Effects of water stress on the chlorophyll content, nitrogen level and photosynthesis of leaves of two maize genotypes. *Photosynth. Res.* **1983**, *4*, 35–47. [CrossRef]
62. Schepers, J.S.; Blackmer, T.M.; Wilhelm, W.W.; Resende, M. Transmittance and reflectance measurements of corn leaves from plants with diferent nitrogen and water supply. *J. Plant Physiol.* **1996**, *148*, 523–529. [CrossRef]
63. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Macé, D.; Pragnère, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2003**, *84*, 1–15. [CrossRef]
64. Tang, S.; Chen, J.M.; Zhu, Q.; Li, X.; Chen, M.; Sun, R.; Zhou, Y.; Deng, F.; Xie, D. LAI inversion algorithm based on directional reflectance kernels. *J. Environ. Manag.* **2007**, *85*, 638–648. [CrossRef] [PubMed]
65. Weiss, M.; Baret, F.; Myneni, R.B.; Pragnère, A.; Knyazikhin, Y. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie* **2000**, *20*, 3–22. [CrossRef]
66. Kimes, D.S.; Knyazikhin, Y.; Privette, J.L.; Abuelgasim, A.A.; Gao, F. Inversion methods for physically-based models. *Remote Sens. Rev.* **2000**, *18*, 381–439. [CrossRef]
67. Sehgal, V.K.; Chakraborty, D.; Sahoo, R.N. Inversion of radiative transfer model for retrieval of wheat biophysical parameters from broadband reflectance measurements. *Inf. Process. Agric.* **2016**, *3*, 107–118. [CrossRef]
68. Verrelst, J.; Rivera, J.P.; Leonenko, G.; Alonso, L.; Moreno, J. Optimizing LUT-Based RTM Inversion for Semiautomatic Mapping of Crop Biophysical Parameters from Sentinel-2 and -3 Data: Role of Cost Functions. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 257–269. [CrossRef]
69. Leonenko, G.; Los, O.S.; North, P.R.J. Statistical Distances and Their Applications to Biophysical Parameter Estimation: Information Measures, M-Estimates, and Minimum Contrast Methods. *Remote Sens.* **2013**, *5*, 1355–1388. [CrossRef]
70. Taniguchi, M. On Estimation of Parameters of Gaussian Stationary Processes. *J. Appl. Probab.* **1979**, *16*, 575–591. [CrossRef]
71. Snee, R.D. Validation of Regression Models: Methods and Examples. *Technometrics* **1977**, *19*, 415–428. [CrossRef]
72. Gujarati, D. *Basic Econometrics*; McGraw: New York, NY, USA, 1995.

73. Breusch, T.S.; Pegan, A.R. A simple test for heteroscedasticity and random coefficient variation. *Econometrica* **1979**, *47*, 1287–1294. [CrossRef]
74. Zhao, D.; Huang, L.; Li, J.; Qi, J. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 25–33. [CrossRef]
75. Stagakis, S.; Markos, N.; Sykioti, O.; Kyparissis, A. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sens. Environ.* **2010**, *114*, 977–994. [CrossRef]
76. Gitelson, A.A.; Vina, A.; Arkebauer, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*. [CrossRef]
77. Pan, H.; Chen, Z.; Ren, J.; Li, H.; Wu, S. Estimating winter wheat leaf area index and canopy water content with three different approaches using Sentinel - 2 multi-spectral instrument (S2 MSI) data. *J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *99*, 11.
78. Liu, K.; Zhou, Q.-B.; Wu, W.-B.; Xia, T.; Tang, H.-J. Estimating the crop leaf area index using hyperspectral remote sensing. *J. Integr. Agric.* **2016**, *15*, 475–491. [CrossRef]
79. Schlerf, M.; Atzberger, C.; Hill, J. Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sens. Environ.* **2005**, *95*, 177–194. [CrossRef]
80. Koetz, B.; Kneubühler, M.; Huber, S.; Schopfer, J.; Baret, F. LAI estimation based on multi-temporal CHRIS/PROBA data and radiative transfer modelling. In Proceedings of the Envisat Symposium 2007, Montreux (CH), Switzerland, 23–27 April 2007.
81. Vuolo, F.; Urso, G.D.; Dini, L. Cost-effectiveness of Vegetation Biophysical Parameters Retrieval from Remote Sensing Data. In Proceedings of the 2006 IEEE International Symposium on Geoscience and Remote Sensing, Denver, CO, USA, 31 July–4 August 2006; pp. 1949–1952.
82. Wu, J.D.; Wang, D.; Bauer, M.E. Assessing broadband vegetation indices and QuickBird data in estimating leaf area index of corn and potato canopies. *Field Crop. Res.* **2007**, *102*, 33–42. [CrossRef]
83. Andrieu, B.; Baret, F.; Jacquemoud, S.; Malthus, T.; Steven, M. Evaluation of an improved version of SAIL model for simulating bidirectional reflectance of sugar beet canopies. *Remote Sens. Environ.* **1997**, *60*, 247–257. [CrossRef]
84. Darvishzadeh, R.; Matkan, A.A.; Dashti Ahangar, A. Inversion of a radiative transfer model for estimation of rice canopy chlorophyll content using a lookup-table approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 1222–1230. [CrossRef]
85. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. [CrossRef]
86. Wenng, H.T. Estimation and Validation of the Biophysical Parameter Leaf Area Index for Agricultural Areas from Satellite Sentinel-2a Data. Master's Thesis, University of Natural Resources and Life Science, Vienna, Austria, 2017.

Paper 3

Manaze, S. E., Pôças, I., & Cunha, M. (2019). Agricultural drought monitoring based on soil moisture derived from the optical trapezoid model in Mozambique. *Journal of Applied Remote Sensing, 13*(2), 17.
doi:10.1117/1.JRS.13.024519.

Paper available in: 22 May 2019

Classification according to journal: research paper

Journal of Applied Remote Sensing

Bibliometric indicators from the Journal Citation Report, Web of Science (JCR – WoS), Clarivate Analytics:
- Journal impact factor 1.344
- Journal rank

Remote Sensing: position 26/30; 4th quartile (Q4)
Environmental sciences: position: 198/251; 4th quartile (Q4)

Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Agricultural drought monitoring based on soil moisture derived from the optical trapezoid model in Mozambique

Sosdito Mananze
Isabel Pôças
Mário Cunha

Sosdito Mananze, Isabel Pôças, Mário Cunha, "Agricultural drought monitoring based on soil moisture derived from the optical trapezoid model in Mozambique," *J. Appl. Remote Sens.* **13**(2), 024519 (2019), doi: 10.1117/1.JRS.13.024519.

Agricultural drought monitoring based on soil moisture derived from the optical trapezoid model in Mozambique

Sosdito Mananze,[a,b,c] **Isabel Pôças,**[a,c] **and Mário Cunha**[a,c,d,*]
[a]Universidade do Porto, Faculdade de Ciências, Porto, Portugal
[b]Universidade Eduardo Mondlane, Escola Superior de Desenvolvimento Rural,
Vilankulo, Moçambique
[c]Geo-Space Sciences Research Center, Porto, Portugal
[d]Institute for Systems and Computer Engineering, Technology and Science (INESC TEC),
Campus da Faculdade de Engenharia da Universidade do Porto, Porto, Portugal

Abstract. Soil moisture (SM) at three depths (15, 25, and 30 cm), derived from the optical trapezoidal model (OPTRAM), was used for multiyear, multisite monitoring of agricultural droughts over two agricultural crops (Maize and Soybean) in southern Mozambique. The OPTRAM was implemented using satellite data from Sentinel-2 and was validated against field SM assessed by gravimetric methods and by Watermark Sensors in sandy-soils with very low water holding capacity ($0.13\ \mathrm{cm}^3/\mathrm{cm}^3$). The OPTRAM model estimated the SM at 15 and 25 cm yielding a $R^2 \geq 0.79$ and RMSE $\leq 0.030\ \mathrm{cm}^3/\mathrm{cm}^3$. The OPTRAM-derived SM was successfully used as input to compute and map the soil water deficit index, an indicator of agricultural drought. The results indicate that OPTRAM can provide useful information to improve water productivity in cropland under the specific conditions of Mozambique agricultural systems and for early warning systems development. © *2019 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.13.024519]

Keywords: soil water deficit index; remote sensing; Sentinel-2; shortwave transformed reflectance; Maize; Soya.

Paper 190161 received Mar. 4, 2019; accepted for publication May 2, 2019; published online May 22, 2019.

1 Introduction

In Mozambique, although irrigated croplands represent only 2.1% of the 3 million hectares of the estimated suitable land for irrigation, they account for 73% of national freshwater consumption.[1] These indicators suggest very low water use efficiency in agriculture, which is supported by the fact that, in most cases, irrigation scheduling is based on soil appearance and feel, by hand touching the soil to empirically assess how dry it is, which potentially leads to overirrigation of crops,[1,2] making it unsustainable from both ecological and economic perspectives. In such context, a successful retrieval of soil moisture (SM) would greatly contribute to gathering key information to support sustainable irrigation practices.

The SM is also a key parameter to define, predict, and mitigate agricultural drought.[3] Agricultural drought refers to circumstances when SM availability to plants has dropped to such a level that adversely affects the crop growth and, hence, agricultural profitability. Thus, monitoring agricultural droughts for early detection is essential for managing the evolving crop losses and food safety risks, and for dealing with increased fire risk.[4,5] Therefore, truthful estimation of SM is also essential to develop early warning systems for agricultural droughts and to formulate better water and land use management practices and hazards mitigation strategies.[6]

The systematic acquisition of SM data through direct measurements in sampling points is arduous and expensive,[7] and hardly represents the spatial variability that characterizes this variable, even in small plots.[8,9] In this context, an important contribution has been given by Earth

---

*Address all correspondence to Mário Cunha, E-mail: mccunha@fc.up.pt

observation satellites (EOS), with both passive and active sensors, which provide synoptic and repetitive measurements of key aspects of land surface parameters, and at various spatiotemporal scales. The EOS allow rapidly measuring, monitoring, and developing low-cost indicators for SM-related applications with increasingly more products being made available in recent years.[10] Often, remote sensing (RS) techniques for SM retrieval are based on (active) satellite sensors providing information in the microwave electromagnetic region, mainly in the low-frequency wavelengths (1 to 12 GHz), from where small changes in soil dielectric properties are detectable.[11] Even though the data from the active sensors or radar instruments are not perturbed by clouds and poor weather conditions, it often has low spatial resolution, needs complex image processing to adequately extract information, and achieves SM retrieval accuracy.[12] The data from the optical sensors have also been used for SM retrieval; these sensors can provide enough spatial and temporal resolutions but suffer from the effect of vegetation cover and nonoptimal sensitivity to SM.[10,13] A very promising progress in SM retrieval using EO data is the synergistic approach, combining active and passive microwave observations. Specifically, NASA's soil moisture active passive satellite mission[6] offers the possibility of determining more accurate subpixel variability of SM, providing the best capability for global SM retrieval. However, for local and regional-scale retrievals, data from active microwave (e.g., synthetic aperture radar) and optical or thermal infrared instruments are able to provide surface SM estimates at relatively fine spatial resolutions.[11]

A number of studies explored the relationship between reflectance measured by optical sensors and SM in different soil types and demonstrated that, generally, the reflectance decreases as the SM increases.[14,15] However, other studies reported nonlinear relationships for these variables.[16,17] RS-based vegetation indices related with plant water stress have also been applied to indirectly estimate the SM,[11] but the empirical nature of this approach makes it site-specific and thus, of limited transferability beyond the tested conditions.

The recently developed optical trapezoid model (OPTRAM) is a physically based RS approach for SM estimation grounded merely on optical data.[18] The model is based on the idea of replacing the land surface temperature in the widely known thermal–optical trapezoid model (TOTRAM) by the short-wave infrared transformed reflectance (STR) building a pixels' distribution within the STR-normalized difference vegetation index (NDVI) space.[18,19] The advantage of OPTRAM over TOTRAM is that the STR–NDVI space remains nearly invariant with time because reflectance is a function of the surface properties and not the ambient atmospheric conditions. Therefore, OPTRAM does not need individual parameterization and calibration for each observation date as required in TOTRAM.[18] Additionally, OPTRAM does not depend on satellite images with the thermal band. To date, the OPTRAM has been evaluated only for SM retrieval in a few watersheds in the United States.[18,19] Therefore, examining OPTRAM in other geographical regions of the world and other ecosystems such as agricultural fields will provide insights on the potential generalization, robustness, and applicability of this model.

Although the SM is placed squarely in the core of drought indicators and classification,[20] most of the proposed drought assessment methods are based on long-term atmospheric variables[21] and do not contemplate site-specific soil properties.[4,19] Nevertheless, several studies have demonstrated the applicability of remotely sensed SM data for agricultural drought monitoring.[4,19,22–24] Likewise, SM-based drought indices have been used for agricultural drought monitoring, including the soil moisture index,[25] soil water deficit (SWD),[26] drought severity index,[27] soil moisture drought index,[28] and soil water deficit index (SWDI).[4,19] Compared to the other indices, the SWDI introduces a novelty of fixing the water deficit threshold by the SM at field capacity (θ_{FC}), which is the site-specific upper limit of the available water for plants, rather than using SM estimated from hydrological modeling[21,29] and its formulation does not rely on longer data availability.[24] The idea is that when the soil reaches a water content threshold below θ_{FC}, plants are no longer in optimal conditions and thus, more and more energy is required for water uptake, and water stress may occur.[4] This interpretation is close to the concept of the total available soil water.[30]

The main goal of this study is to evaluate the applicability of OPTRAM model in retrieving SM within croplands in Mozambique. The specific objectives are: (i) to test the performance of OPTRAM over two agricultural crops (maize and soya) and (ii) to establish a multiyear monitoring agricultural drought at crop field level based on OPTRAM estimated SM.

2 Materials and Methods

2.1 *Study Area*

The study took place in croplands around the Vilankulo district in southern Mozambique. Predominant land cover/uses comprise agriculture (rainfed and irrigated), forests, rangelands, and bare soils. Four cropland fields managed by local agricultural associations were sampled during 2015, 2016, 2017, and 2018 (Fig. 1). Two parcels with 400 m^2 were defined for data collection in each field. The main characteristics of each field are presented in Table 1 and a detailed description can be found in Ref. 31.

Figure 2 presents the average precipitation and temperature in the region for a 25-year reference period (1991 to 2015) and for the four years of data collection (2015, 2016, 2017, and 2018). The years 2015 and 2016 were both 4.5°C warmer and, respectively, 469.1 and 289.6 mm dryer than the reference period (804.4 mm). The year 2017 registered relatively high precipitation (929.1 mm) while 2018 was the coolest year with 23.3°C.

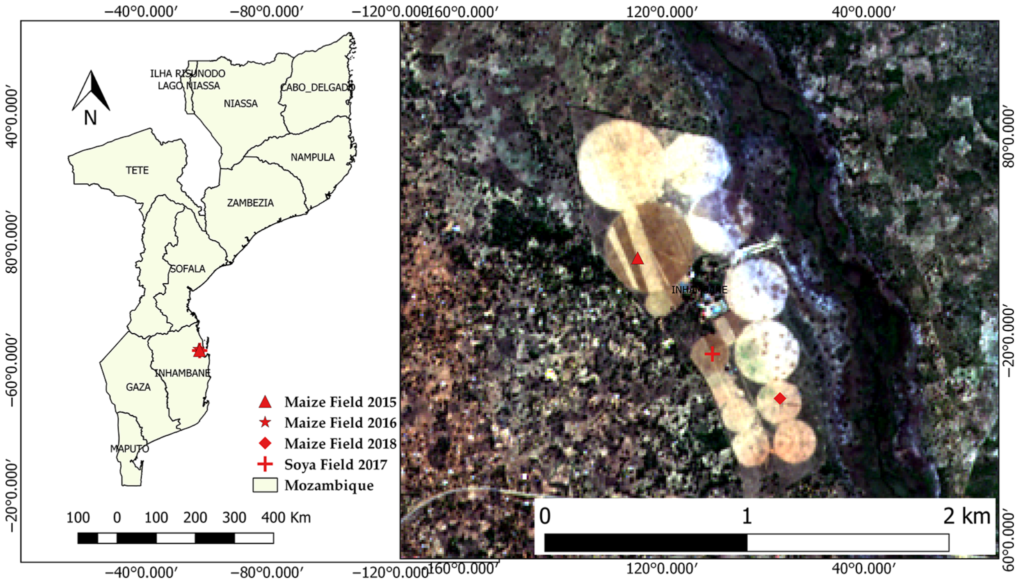


Fig. 1 Location of the sampling fields overlaid to a true color RGB of Sentinel-2 image (right) and the boundaries of Mozambique provinces (left).

Table 1 Characteristics of the sampled fields.

	Field 1 (2015)	Field 2 (2016)	Field 3 (2017)	Field 4 (2018)
Latitude	21°56′24.19″	21°59′02.53″	21°59′17.23″	21°59′23.82″
Longitude	35°07′24.19″	035°09′30.95″	35°09′43.86″	35°09′56.02″
Area (Ha)	3	1	1	3
Irrigation system and scheduling	Sprinkler: 3 days interval from V3–V8 and 6 days interval in the following stages.	Drip irrigation: 3 days interval from V3–V8 and 6 days interval in the following stages.	Central pivot irrigation: twice a week irrigation interval.	Central pivot: 5 days irrigation interval.
Crop/variety	Maize/PAN 53, a medium maturity variety (125–140 days to harvest)	Maize/PAN 53, a medium maturity variety (125–140 days to harvest)	Soya/FALCAO	Maize/PAN 67, a medium maturity variety (120–130 days to harvest).

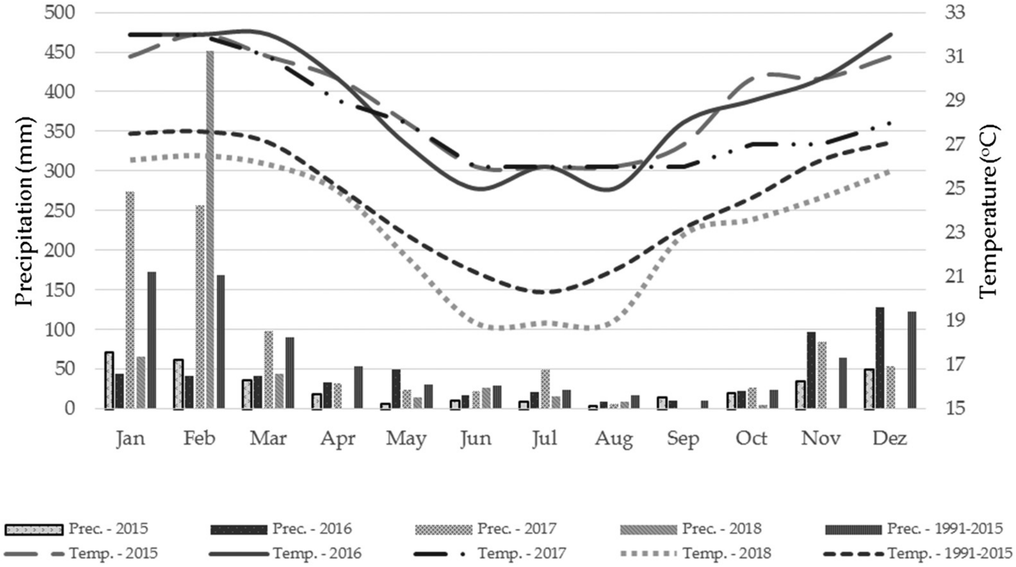


Fig. 2 Average monthly precipitation (prec.) and temperature (temp.) for the 4 years of data collection and comparison with a reference period (1991–2015).[32]

The soil in the sampled fields is classified as sandy loam with a very low water capacity ($0.13\ \mathrm{cm}^3/\mathrm{cm}^3$) within the evaluated soil thickness of 30 cm. The fields were irrigated through different irrigation systems (Table 1), and the irrigation schedule was managed by the farmers without our interference.

2.2 *Data Collection and Preprocessing*

2.2.1 *Ground measurements of soil moisture*

Soil water tension (SWT) data were collected during four consecutive cropping seasons in 2015 (field 1, maize), 2016 (field 2, maize), 2017 (field 3, soya), and 2018 (field 4, maize). In each parcel, seven watermark 200SS sensors were placed at three soil depths, 15, 25, and 30 cm, and the sensors were plugged to a data logger. These sensors are sensitive in the range between 0 (near-saturated soil) to 239 centibars (dry soil),[33] and were set to make readings at an hour interval during the entire crop growing cycle. The measurement depths were chosen according to Ref. 34, who concluded that below a soil depth of 30 cm, the root biomass and the number of roots drastically decrease for annual crops.

Additionally, during the seasons of 2016 (field 2, maize) and 2017 (field 3, soya), soil samples were extracted at equal depths, where the watermark sensors were installed to apply the gravimetric method according to Ref. 35, aiming to determine the SM ($\mathrm{cm}^3/\mathrm{cm}^3$). This gravimetric method was used as a reference.

The field capacity (θ_{FC}) was determined through a field technique, according to Ref. 36. Accordingly, a test plot was irrigated until the soil profile saturation to a depth of about 1 m. Then, the plot was covered to prevent evaporation. The SM was measured every 24 h until the changes were very small, at which point the SM content corresponds to field capacity.[36]

The permanent wilting point (θ_{PWP}) was determined through the procedure described by Ref. 37. Five soil samples were extracted from each of the three soil depths and put in a metal can. Corn seeds were sown, and the plants grown outdoors in adequate SM until the third pair of true leaves was formed. Then, the top of the can was sealed with candle wax to prevent evaporation and the plants were left growing until they wilt. The plants were then transferred to a dark, humid compartment for recovery. In cases when the plants recovered, they were put outside again. The procedure was repeated until the plants remained wilted

Table 2 Gravimetric SM data collection dates in maize and soybean fields.

Field 2, maize in 2016		Field 3, soybean in 2017	
Maize stages	Dates	Soybean Stages	Dates
Three visible leaves V3	June 28	Two visible leaves V2	August 2
Six visible leaves V6	July 19	Four visible leaves V4	August 15
Eight visible leaves V8	July 30	Five visible leaves V5	September 7
Last leaf developed VT	August 17	1 open flower R1	October 3
First reproductive stage R1	August 28	0.5 cm pod R3	November 20
Last reproductive stage RT	September 29	Pod with green seeds R6	December 12

overnight (24 h) in the humid compartment. The SM average for each depth was then determined for obtaining the θ_{PWP}.[37]

The soil porosity (SP, %) calculated as a percentage of the soil volume, was derived from the bulk density (BD; $\mathrm{mg/m^3}$) according to Eq. (1), using a value of soil particle density of $2.65\ \mathrm{mg/m^3}$. The porosity was later used to convert the OPTRAM SM estimates:

$$\mathrm{SP}(\%) = \left(1 - \frac{\mathrm{BD}}{\text{particle density}}\right) \times 100. \tag{1}$$

The SM ($\mathrm{cm^3/cm^3}$) for each depth was calculated according to Eq. (2):

$$\mathrm{SM}(\mathrm{cm^3/cm^3}) = \frac{W_{\mathrm{w}} - D_{\mathrm{w}}}{D_{\mathrm{w}}} \times \mathrm{BD}, \tag{2}$$

where W_{w} is the wet weight of soil samples placed in aluminum tins and then dried to a constant weight in an oven at a temperature of 105°C to obtain the soil dry weight (D_{w}), both weighed in a scale (± 0.001 g).

Table 2 presents the dates and corresponding crop stages at which the gravimetric method was applied.

For each date indicated in Table 2, the average of SWT recorded by watermark sensors and SM at each depth was determined. The relationship between the 36 measurements of gravimetric-based SM and SWT was established by a regression model. The model was then used to standardize all the sensor measurements throughout the study period.

2.2.2 *Satellite data*

The multispectral Sentinel-2A (launched in 2015) and Sentinel-2B (launched in 2017) satellite images were used in the present study. Both Sentinel-2A and Sentinel-2B fully cover Earth every 10 days and jointly, they potentially allow getting data at 5 days interval. The Sentinel-2 spatial resolution is 10 m for four bands (2, 3, 4, and 8), 20 m for six bands (5, 6, 7, 8a, 11, and 12), and 60 for the other two bands (1 and 9).

Cloud-free and atmospherically corrected images (L2A level) of the study area were selected and downloaded from the data service platform developed by Ref. 38 based on the European Space Agency's (ESA) Sen2Cor algorithm. Only four images (September 28, 2016; October 28, 2016; September 13, 2017; and November 22, 2017) completely free of clouds were available for the study area during the studied period. These four images were used for OPTRAM model parameterization. Additionally, nine images (June 30, 2016; July 30, 2016; August 09, 2016; September 28, 2016; November 08, 2016; July 25, 2017; January 01, 2018; January 06, 2018; and March 02, 2018) presenting some cloud coverage percentage were considered for validation purposes.

2.3 *Model Development*

2.3.1 *OPTRAM parameterization*

Because the relationship between the SM and STR is only valid for partially and fully saturated soils but not for oversaturated soils,[18] in all the selected images, the relevant bands of Sentinel-2 (bands 4: $\rho 665$ nm, 8: $\rho 842$ nm, and 12: $\rho 2190$ nm) for OPTRAM parameterization were first masked for oversaturated pixels in permanent or seasonal water bodies (e.g., pounds, lagoons, rivers, etc.). This was achieved through water features extraction using the normalized difference water index [NDWI; Eq. (3)] that produces an image in which the positive values are typically open water areas while negative values comprise nonwater features:[39]

$$\mathrm{NDWI} = \frac{R_{\mathrm{green}} - R_{\mathrm{nir}}}{R_{\mathrm{green}} + R_{\mathrm{nir}}}. \tag{3}$$

The water mask was applied to the Sentinel-2 bands 4, 8, and 12. An additional mask was applied over the course of the main river crossing the study area given that it is densely covered by vegetation (mainly Phragmites sp.) and thus, undetectable as water body through the NDWI. These procedures were applied in QGIS software.

Because the aim of the study is monitoring agricultural droughts at field level, the band 12 was resampled to match the 10-m spatial resolutions of bands 4 and 8, rather than resampling all bands to a coarser resolution, as pointed out by Ref. 18. The resampling was performed within the Sentinel application platform developed by ESA through a bilinear method.

The OPTRAM is based on the linear relationship between SM and STR as follows:[40]

$$\mathrm{SM} = \frac{\mathrm{STR} - \mathrm{STR_d}}{\mathrm{STR_w} - \mathrm{STR_d}}, \tag{4}$$

where $\mathrm{STR_d}$ and $\mathrm{STR_w}$ are STR at dry and wet soil states, respectively. The STR is obtained as follows:

$$\mathrm{STR} = \frac{(1 - \rho_{\mathrm{SWIR}})^2}{2\rho_{\mathrm{SWIR}}}, \tag{5}$$

where ρ_{SWIR} is the surface reflectance for the short-wave infrared (SWIR) domain (Sentinel-2 band 12).

The OPTRAM assumes a linear relationship between soil and vegetation water contents.[18] This relationship was reported in several studies[41–46] applying RS data to relate plant vigor characteristics to the SM at the root zone. Accordingly, the dry and wet edges of the optical trapezoid are defined as follows:[18,19]

$$\mathrm{STR_d} = i_{\mathrm{d}} + s_{\mathrm{d}}\mathrm{NDVI}, \tag{6}$$

$$\mathrm{STR_w} = i_{\mathrm{w}} + s_{\mathrm{w}}\mathrm{NDVI}, \tag{7}$$

where i_d and s_d are the intercept and slope of the dry edge, respectively, and i_{w} and s_{w} are the intercept and slope of the wet edge, respectively. Combining Eqs. (5)–(7), the normalized SM (W) within a given satellite image pixel can be estimated as a function of its STR and NDVI values using Eq. (8):

$$W = \frac{i_{\mathrm{d}} + s_{\mathrm{d}}\mathrm{NDVI} - \mathrm{STR}}{i_{\mathrm{d}} - i_{\mathrm{w}} + (s_{\mathrm{d}} - s_{\mathrm{w}})\mathrm{NDVI}}. \tag{8}$$

The parameters of OPTRAM, dry (i_{d} and s_{d}) and wet (i_{w} and s_{w}) edges, were determined based on the pixels' distribution within the STR-NDVI space, using visual inspection. This approach is preferred over the least-square regression approach because it can successfully neglect oversaturated or shadowed pixels scattered around the main body of the trapezoid.[18,19] Although the user might introduce some bias, the visual inspection does not significantly affect the results.[19]

Values of W in the pixels, where field data of SM were collected, were subsequently converted to volumetric SM multiplying by the SP for comparison with observed field data of SM.

2.3.2 *Model validation and accuracy assessment*

The OPTRAM-derived SM was compared with observed field data of SM through regression model. The accuracy of the estimates was assessed using four widely known statistical measures: root mean square error (RMSE, $\mathrm{cm}^3/\mathrm{cm}^3$), normalized root mean square error (%), mean absolute error ($\mathrm{cm}^3/\mathrm{cm}^3$), and coefficient of determination (R^2).

2.3.3 *Agricultural drought analysis*

The estimation of agricultural drought was performed using the SWDI [Eq. (9)]:

$$\mathrm{SWDI} = 10 \times \frac{\theta - \theta_{\mathrm{FC}}}{\theta_{\mathrm{FC}} - \theta_{\mathrm{PWP}}}, \tag{9}$$

where θ is the actual SM, θ_{FC}, and θ_{PWP} are SM content at field capacity and permanent wilting point, respectively. The OPTRAM-derived SM was used to determine and map the SWDI applying the available Sentinel-2 images while the field SM was used to calculate the SWDI aiming to assess its temporal dynamic throughout the in-season growing period in each sampled point and field.

The SWDI is interpreted as follows: $\mathrm{SWDI} > 0$, excess of water; $\mathrm{SWDI} = 0$, soil is at the field capacity, and $\mathrm{SWDI} < 0$ indicates some level of drought:[4,19] $0 > \mathrm{SWDI} > -2$, mild; $-2 > \mathrm{SWDI} > -5$, moderate; $-5 > \mathrm{SWDI} > -10$, severe; and $-10 > \mathrm{SWDI}$, extreme.

3 Results

3.1 *Soil Moisture Ground Measurements*

Figure 3 depicts the intra- and interannual variability of SM during the entire crop cycles in the depths of 15, 25, and 30 cm, for the four studied cropping seasons. The intra-annual variability

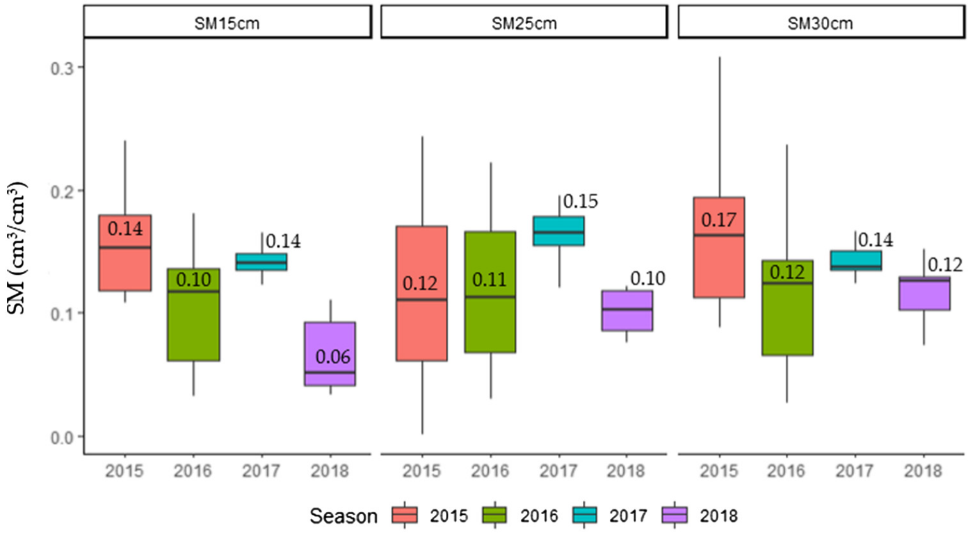


Fig. 3 Boxplot of the field SM at 15-, 25-, and 30-cm depths for the 4 years: 2015 ($n = 14$), 2016 ($n = 12$), 2017 ($n = 16$), and 2018 ($n = 10$). The boxplots are flagged by the average value of SM.

Table 3 Average values of field capacity (θ_{FC}), permanent wilting point (θ_{PWP}), and BD within the sampled fields.

Depth (cm)	θ_{FC} (cm^3/cm^3)	θ_{PWP} (cm^3/cm^3)	BD (g/cm^3)
15	0.204 ± 0.02	0.054 ± 0.008	1.5 ± 0.08
25	0.181 ± 0.07	0.053 ± 0.008	1.5 ± 0.02
30	0.162 ± 0.05	0.049 ± 0.006	1.5 ± 0.03

on the SM content roughly follows the same pattern for the different soil depths. There was a slight increase of SM with the increase of soil depth throughout the study period as indicated by the average of 0.11, 0.12, and 0.14 cm^3/cm^3 at 15-, 25-, and 30-cm depths, respectively. Similar results were also verified by Ref. 47. It is well known that sandy loam soils present low water holding capacity.[48–50] Additionally, other authors observed the same pattern and concluded that the SM at the surface or shallow layers is greatly influenced by infiltration and/or evapotranspiration and is a regular water source for vegetation growth.[51]

For the three depths, the highest values were registered in 2017, while the lower SM values were recorded in the year 2018, likely due to the observed insufficiencies in 2018 irrigation schedule associated to power fluctuation constraints.

The average field capacity (θ_{FC}), permanent wilting point (θ_{PWP}), and soil BD, as assessed in the fields 2 and 3, in the three depths, are presented in Table 3. These soils present very low water holding capacity (~0.13 cm^3/cm^3), being prone to crop water stress when an efficient irrigation schedule is not possible.

3.2 *Relationship Between Soil Moisture and Soil Water Tension*

Figure 4 presents the relationship between gravimetric-based SM (cm^3/cm^3) and SWT (centibars) recorded by watermark sensors over 36 field measurements. Most observations (88.9%) are below the θ_{FC}, including 22.3% observations under the θ_{PWP}. According to the fitted logarithmic regression, about 89% ($R^2 = 0.89$; $n = 36$; $p < 0.000$) of the variability in the gravimetric SM could be explained by the SWT based on watermark measurements. As was expected, the decreasing of SM values corresponded to the increase of SWT indicating that the plants spend relatively more and more energy to extract water from the soil.

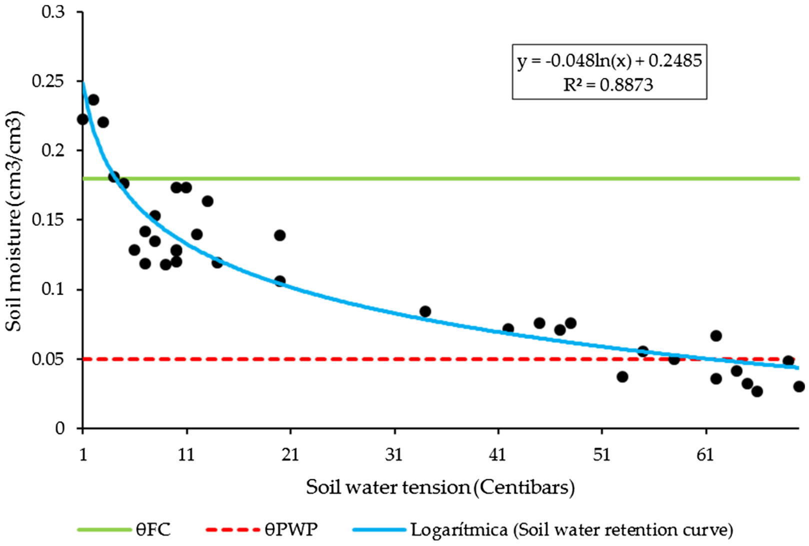


Fig. 4 Relationship between gravimetric SM and soil water tension and the representation of field capacity (θ_{FC}) and permanent wilting point (θ_{PWP}).

The shape of the curve is similar to the soil-water retention curve reported in other studies[33,52] for similar soil texture in other regions. Nevertheless, this relationship is highly site specific and should be experimentally measured for each type of soil under consideration,[33] as done in the present study. The results indicate that, in most cases, the SM stayed below the θ_{FC} and quickly decreased to θ_{PWP} from SWT values of about 30 centibars. This is consistent with Ref. 33, who recommended 30 to 50 centibars as critical values for irrigation in sandy soils, due to their low water holding capacity and limited available water.

3.3 *OPTRAM Parameterization*

The pixels' distribution within the STR-NDVI space for model parameterization is illustrated in Fig. 5. It configures an almost trapezoidal shape indicating a high correlation between SM and STR. The visually estimated parameters are: (i) dry edge: $i_d = 0$, $s_d = 1.1$ and (ii) wet edge: $i_w = 1.2$, $s_w = 4.7$. The parameters captured most of the pixels within the space STR-NDVI, denoting good OPTRAM parameterization for the study area.

According to Ref. 18, coarser spatial resolutions perform better in the oversaturation pixels removal. In this study, even though a fine spatial resolution (10 m) was considered, few oversaturated pixels (above the best-wet edge in Fig. 5) were detected, suggesting good performance of the two procedures of water body masking that were applied to the data prior to OPTRAM parameterization.

3.4 *OPTRAM Soil Moisture Estimates*

The OPTRAM SM estimates for the three soil depths were compared with the corresponding values obtained by *in-situ* measurements. The goodness-of-fit statistics for this comparison showed very good results for the soil depths of 15 and 25 cm (Table 4). Nevertheless, some

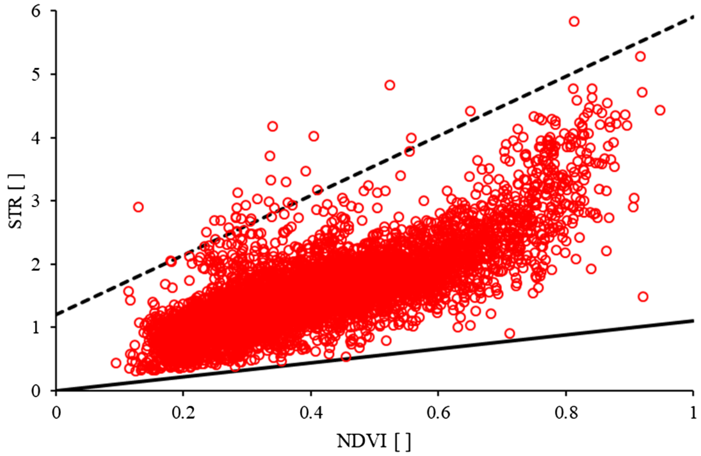


Fig. 5 Pixel distribution within the NDVI-STR space. Full black line, best dry edge; dashed black line, best wet edges optimized for OPTRAM.

Table 4 Goodness-of-fit statistics of the comparison between *in situ* and OPTRAM estimated SM.

Depth	R^2	RMSE (cm^3/cm^3)	Normailzed RMSE (%)	Mean absolute error (cm^3/cm^3)
SM_15 cm	0.87	0.025	10.78	0.018
SM_25 cm	0.79	0.030	13.7	0.024
SM_30 cm	0.43	0.050	16.48	0.036

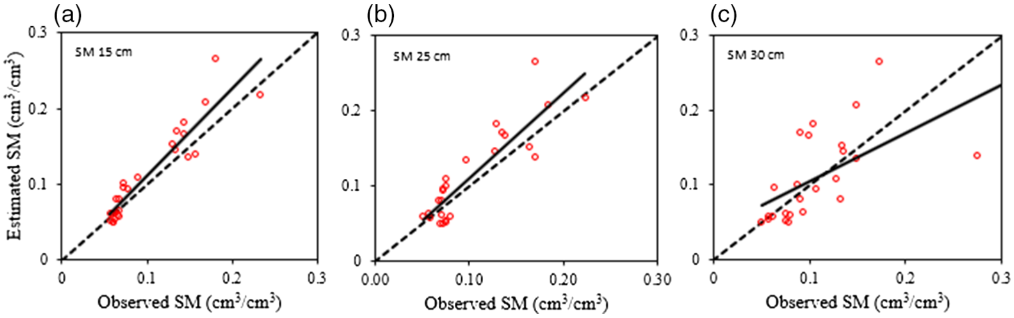


Fig. 6 Comparison between *in situ* and OPTRAM estimated SM (OPTRAM SM). Dashed line, 1:1 relationship; full line, linear regression.

overestimation was observed at all depths, especially for higher values of SM and particularly for the 30-cm depth (Fig. 6).

Figure 6 depicts the comparison between OPTRAM estimated and *in situ* SM. The best accuracy is achieved in the shallow depths and broadly, OPTRAM accuracy decreases with an increase in SM value.

3.5 *Agricultural Drought Analysis*

To examine the agricultural drought evolution throughout the crop growing seasons, the SWDI was computed based on the field SM at 15 cm for each field (Fig. 7). The soil depth of 15 cm was chosen based on the agreement between field and OPTRAM estimated SM. The dynamic of SDWI during each cropping season is illustrated in Fig. 7.

The maize field in 2016 is the one that experienced long periods of extreme drought condition (SWDI < −10) while the maize field in 2015 and the soya field in 2017 registered more periods of moderate and mild droughts. This is consistent with: (i) the irrigation schedule that was many times hampered by water pump malfunction and (ii) very low rainfall and high temperature during the cropping season (Fig. 2).

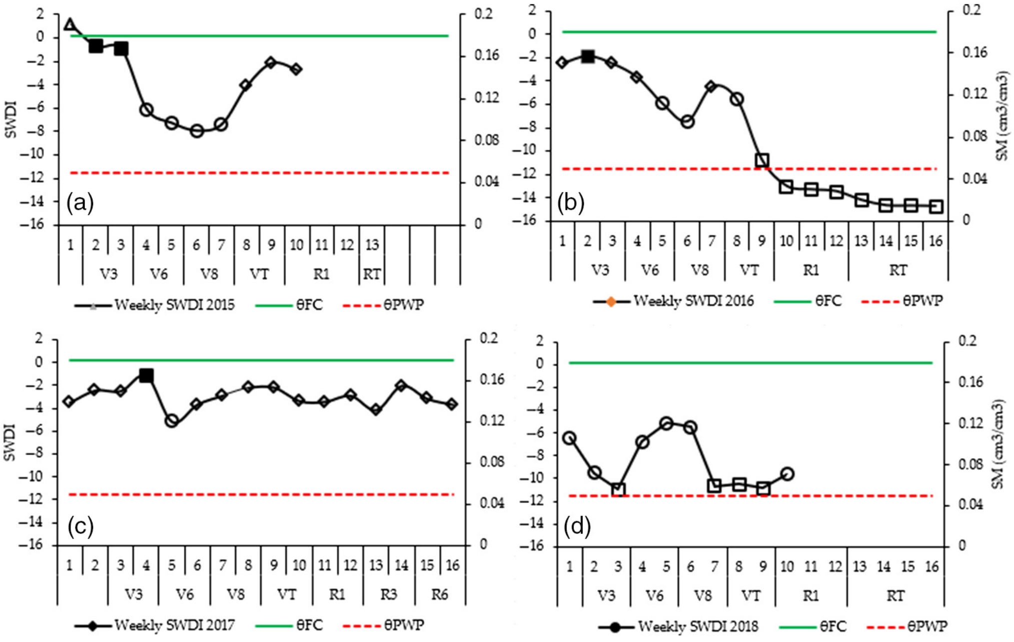


Fig. 7 Weekly evolution of SWDI: (a) – 2015; (b) – 2016; (c) – 2017; (d) – 2018. Interpretation: triangle (no drought); dark square (mild drought); diamond (moderate drought); circle (severe drought); and open square (extreme drought).

3.6 *Agricultural Drought Mapping*

The OPTRAM-derived SM (OPTRAM SM) was used to map the spatial pattern of agricultural drought (from SWDI) within the sampled fields using six Sentinel images temporarily distributed (Fig. 8). Overall, in all 6 days, the fields experienced some degree of drought, ranging from mild to extreme. The results show a great variability of SWDI values between and within image dates. The field sampled in 2016 is the one presenting large area under extreme and severe drought, what is consistent with: (i) the irrigation schedule that was many times hampered by water pump malfunction and (ii) very low rainfall and high temperature during the cropping season (Fig. 2). For 2017 and 2018 fields, large areas are within moderate, mild drought, and no drought being consistent with the relatively high rainfall registered throughout the crop growing periods (Fig. 2). Particularly, the SWDI map for March 2, 2018, shows many pixels with adequate SM (no drought), which is consistent with the fact that the temperature during

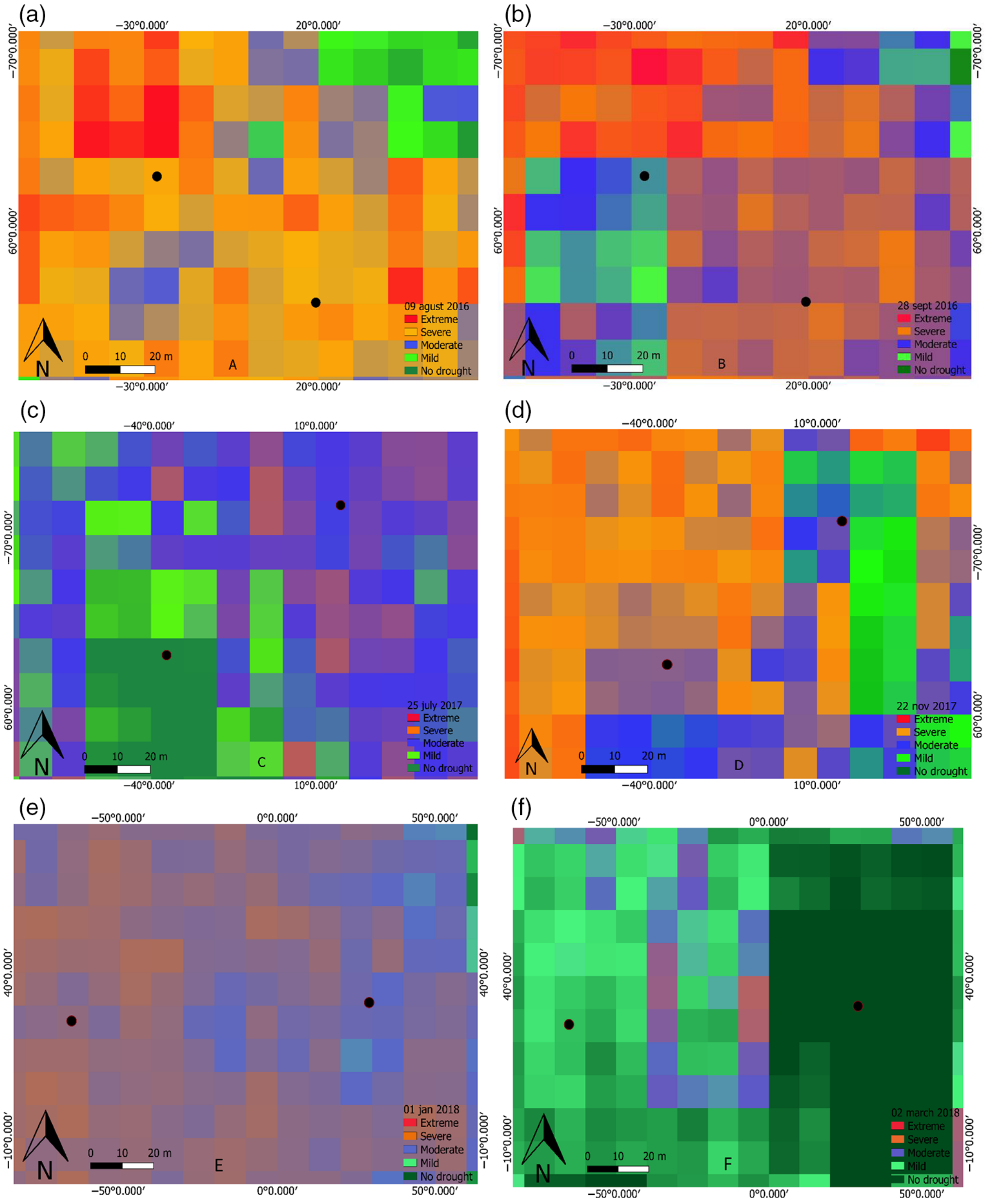


Fig. 8 Spatial map of drought evolution: (a) August 9, 2016; (b) September 28, 2016; (c) July 25, 2017; (d) November 22, 2017; (e) January 1, 2018; and (f) March 2, 2018. Black dots represent centers of the parcels.

January and February of 2018 was lower if compared to the same periods of other years and of the reference period (1991 to 1995) and the rainfall was higher in February. These results highlight the relevant role that SWDI could play as agricultural drought indicator at field level.

4 Discussion

The trapezoidal shape of the pixel's distribution in the space STR-NDVI confirms the good correlation between SM and STR both in soils with sparse and with high vegetation cover, as assessed by the NDVI values, as also reported by Refs. 18 and 19. The STR-NDVI space was built from Sentinel-2 images covering a wide range of SM variability (irrigated and rainfed crop fields) and of vegetation cover types (agriculture, forests, rangelands, bare soils, etc.). Additionally, the larger values of STR and NDVI represent areas with dense vegetation cover as a result of high SM in the root zone and/or high vegetation water content.[19] This was equally verified in our study with high values of STR and NDVI in later stages of crop development and in surrounding forested areas.

Comparing the *in situ* SM with the OPTRAM SM, the estimation errors ($\text{RMSE} \leq 0.050\ \text{cm}^3/\text{cm}^3$) for all depths are within what is described as reasonable agreement accuracy for RS mapping of SM.[53] Previous studies demonstrated good agreement between OPTRAM estimates and *in situ* data measured with electromagnetic sensors installed at 5-cm depth[18] and with *in situ* data acquired with the cosmic-ray neutron providing effective SM within the top soil,[54] achieving similar error values. The higher accuracy of OPTRAM is mostly verified in lower and intermediate values of SM, which is of great interest because it includes SM levels at which the operational decisions for irrigation should be taken.

Also, Roberts et al.[55] observed decreasing correlations with depth when analyzing the relationship between vegetation indices and SM content. This behavior may be explained by the fact that the top soil layer and SM are usually the regular water source for vegetation growth[51] and, thus, more likely to be highly correlated with the vegetation spectral characteristics. Moreover, the assumption is that OPTRAM SM is correlated to root zone SM through the vegetation response to SWD in the root zone layer.[18,19]

Results presented here strengthen previous studies by Refs. 18 and 19 revealing OPTRAM usefulness for SM estimation for a wider variety of soil types, land cover types, climatic conditions, and diversified methods/equipment of *in situ* SM measurements. Furthermore, the fact that OPTRAM requires only observations in the optical electromagnetic domain[18] is a great gain given that freely optical satellite data are available at adequate spatial and temporal resolutions for agricultural-related studies. Consequently, OPTRAM is a truthful tool for SM estimation, a crucial variable to delineate value-added water and land use management practices and hazards mitigation strategies, with great importance in countries like Mozambique, where water is the main limiting resource for crop growing. However, the application of this model may be hampered by satellite data availability because of the frequent and thick cloud coverage that characterizes many regions in Mozambique. Also, it is important to notice that SM estimation accuracy with OPTRAM drops for greater depths (Table 4). This issue must be taken in consideration, particularly for crops with deep roots that explore deeper soil layers, and thus, further work should explore the extent of correlation between surface and root zone SM through vegetation response to SWD in the root zone.[18,19]

Regarding the agricultural drought monitoring, the tested irrigated crops experienced several periods with some level of agricultural drought during the crop cycles ($\text{SDWI} < 0$) and very few periods (weeks) of no agricultural drought ($\text{SDWI} > 0$). However, for most cases, the drought level was above the lower limit of available water for plants ($\text{SWDI} < -10$), meaning that the plants would not be effective under water stress but required more and more energy for water uptake. Previous studies reported similar findings, however, involving other moisture deficit indicators like the atmospheric water deficit, the crop moisture index, and rainfall.[4,19,24]

Broadly, the SWDI derived from OPTRAM showed to be a well-functioning indicator of agricultural droughts providing an opportunity for high spatial and temporal resolution drought monitoring and mapping. In this study, we successfully mapped the within-field variability of drought level (Fig. 8), which could be used to support precision agriculture strategies. The SWDI could also be developed to support strategies of crop irrigation in order to increase the water use efficiency. Thus, SWDI can potentially be used to produce spatially and timely effective information for local farmers, agrarian extension services, entities for sustainable water management strategies and policies as well as for entities for famines and natural hazards management. However, and as equally pointed by Ref. 24, the main challenges are the quantification of soil textures, soil water characteristics and, in particular, for our study area, an extensive quantification of *in situ* SM data including for other crops and land cover types.

5 Conclusions

OPTRAM model was applied to retrieve croplands SM in sandy-soils with very low water-holding capacity in Mozambique. The model accurately estimated SM especially at shallow depths (15 and 25 cm), representing a useful tool for spatial and temporal monitoring of this important agricultural variable. The usefulness of OPTRAM is particularly pertinent because it only requires optical RS data, which are freely available from satellite missions like Landsat and Sentinel-2 and have an appropriate spatial and temporal resolution to support sustainable strategies of water management in irrigation crops. In addition, OPTRAM-derived SM showed good performance when used for agricultural droughts monitoring what potentially enables the availability of this kind of information to local farmers and other stakeholders engaged in water management. This information can contribute to alleviate the impacts of droughts on small scale and subsistence agricultural systems.

Acknowledgments

Sosdito Mananze acknowledges the Joint Aid Management (JAM, Mozambique) for making available the crop fields for data collection; the Gulbenkian Foundation for the PhD scholarship through the Grant No. 140227; the ESUDER technicians for their support during data collection and the remote sensing capacity building project funded by the Austrian Partnership Programme in High Education and Research for Development (Appear) through the Project 176. Isabel Pôças acknowledges the postdoctoral grant supported by the Research Infrastructure Enabling Green E-Science for the SKA (ENGAGE SKA), reference POCI-01-0145-FEDER-022217, funded by COMPETE 2020 and FCT, Portugal.

References

1. M. D. Agricultura, *Estrategia de Irrigação*, A. Nhabetse et al., eds., Ministerio da Agricultura, Maputo, Mozambique, p. 54 (2013).
2. K. Frenken, Irrigation in Africa in figures. AQUASTAT Survey, in *FAO Water Reports*, K. Frenken, Ed., FAO, Rome, Italy, p. 63 (2005).
3. J. Martínez-Fernández et al., “Satellite soil moisture for agricultural drought monitoring: assessment of the SMOS derived soil water deficit index,” *Remote Sens. Environ.* **177**, 277–286 (2016).
4. J. Martínez-Fernández et al., “A soil water based index as a suitable agricultural drought indicator,” *J. Hydrol.* **522**, 265–273 (2015).
5. U. S. Panu and T. C. Sharma, “Challenges in drought research: some perspectives and future directions,” *Hydrol. Sci. J.* **47**(suppl1), S19–S30 (2002).
6. M. E. Brown et al., “NASA’s soil moisture active passive (SMAP) mission and opportunities for applications users,” *Bull. Am. Meteorol. Soc.* **94**(8), 1125–1128 (2013).
7. K. Mallick, B. K. Bhattacharya, and N. K. Patel, “Estimating volumetric surface moisture content for cropped soils using a soil wetness index based on surface temperature and NDVI,” *Agric. Forest Meteorol.* **149**(8), 1327–1342 (2009).
8. D. Fernández-Prieto et al., “Earth observation for land-atmosphere interaction science,” *Biogeosciences* **10**, 261–266 (2013).
9. P. Rahimzadeh-Bajgiran et al., “Estimation of soil moisture using optical/thermal infrared remote sensing in the Canadian prairies,” *ISPRS J. Photogramm. Remote Sens.* **83**, 94–103 (2013).

10. J. A. Sobrino et al., "A method to estimate soil moisture from airborne hyperspectral scanner (AHS) and ASTER data: application to SEN2FLEX and SEN3EXP campaigns," *Remote Sens. Environ.* **117**, 415–428 (2012).
11. G. P. Petropoulos, G. Ireland, and B. Barrett, "Surface soil moisture retrievals from remote sensing: current status, products & future trends," *Phys. Chem. Earth, Parts A/B/C* **83-84**, 36–56 (2015).
12. S. Paloscia et al., "Soil moisture mapping using Sentinel-1 images: algorithm and preliminary validation," *Remote Sens. Environ.* **134**, 234–248 (2013).
13. O. Merlin et al., "Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency," *Remote Sens. Environ.* **112**(10), 3935–3946 (2008).
14. S. Fabre, X. Briottet, and A. Lesaignoux, "Estimation of soil moisture content from the spectral reflectance of bare soils in the $0.4-2.5$ μm domain," *Sensors* **15**, 3262–3281 (2015).
15. Z. Gao et al., "A method of estimating soil moisture based on the linear decomposition of mixture pixels," *Math. Comput. Modell.* **58**(3), 606–613 (2013).
16. M. Nocita et al., "Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy," *Geoderma* **199**, 37–42 (2013).
17. D. B. Lobell and G. P. Asner, "Moisture effects on soil reflectance," *Soil Sci. Soc. Am. J.* **66**(3), 722–727 (2002).
18. M. Sadeghi et al., "The optical trapezoid model: a novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations," *Remote Sens. Environ.* **198**, 52–68 (2017).
19. E. Babaeian et al., "Mapping soil moisture with the optical trapezoid model (OPTRAM) based on long-term MODIS observations," *Remote Sens. Environ.* **211**, 425–440 (2018).
20. T. E. Ochsner et al., "State of the art in large-scale soil moisture monitoring," *Soil Sci. Soc. Am. J.* **77**(6), 1888–1919 (2013).
21. E. H. Hogg, A. G. Barr, and T. A. Black, "A simple soil moisture index for representing multi-year drought impacts on aspen productivity in the western Canadian interior," *Agric. Forest Meteorol.* **178-179**, 173–182 (2013).
22. H. Carrão et al., "An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data," *Int. J. Appl. Earth Obs. Geoinf.* **48**, 74–84 (2016).
23. D. Liu et al., "Performance of SMAP, AMSR-E and LAI for weekly agricultural drought forecasting over continental United States," *J. Hydrol.* **553**, 88–104 (2017).
24. A. Mishra et al., "Drought monitoring with soil moisture active passive (SMAP) measurements," *J. Hydrol.* **552**, 620–632 (2017).
25. V. Sridhar et al., "Development of the soil moisture index to quantify agricultural drought and its" user friendliness" in severity-area duration assessment," *J. Hydrometeorol.* **9**(4), 660–676 (2008).
26. J. M. Torres-Ruiz et al., "Shoot hydraulic characteristics, plant water status and stomatal response in olive trees under different soil water conditions," *Plant Soil* **373**(1-2), 77–87 (2013).
27. C. Cammalleri, F. Micale, and J. Vogt, "A novel soil moisture-based drought severity index (DSI) combining water deficit magnitude and frequency," *Hydrol. Process.* **30**(2), 289–301 (2016).
28. M. M. Sohrabi et al., "Development of soil moisture drought index to characterize droughts," *J. Hydrol. Eng.* **20**(11), 04015025 (2015).
29. E. Dutra, P. Viterbo, and P. M. A. Miranda, "ERA-40 reanalysis hydrological applications in the characterization of regional drought," *Geophys. Res. Lett.* **35**(19), 5 (2008).
30. R. G. Allen et al., *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements, in Irrigation and Drainage*, FAO, Rome (1998).
31. S. E. Mananze, I. Pôças, and M. Cunha, "Retrieval of maize leaf area index using hyperspectral and multispectral data," *Remote Sens.* **10**(12), 1942 (2018).
32. The World Bank, "Climate change knowledge portal for development practitioners and policy makers," 2018, http://sdwebx.worldbank.org/climateportal/index.cfm?page=country_historical_climate&ThisRegion=Africa&ThisCCode=MOZ [cited 2018/09/07].

33. S. Irmak et al., *Principles and Operational Characteristics of Watermark Granular Matrix Sensor to Measure Soil Water Status and Its Practical Applications for Irrigation Management in Various Soil Textures, in Extension*, University of Nebraska-Lincoln Extension, Nebraska, p. 14 (2014).
34. L. Pietola and L. Alakukku, "Root growth dynamics and biomass input by Nordic annual field crops," *Agric. Ecosyst. Environ.* **108**(2), 135–144 (2005).
35. C. A. Black, *Methods of Soil Analysis: Part I Physical and Mineralogical Properties*, American Society of Agronomy, Madison, Wisconsin (1965).
36. W. R. Walker, "Guidelines for designing and evaluating surface irrigation systems," in *FAO Irrigation and Drainage Paper*, FAO, Rome, Italy, vol. 45, (1989).
37. S. A. Taylor and G. L. Ashcroft, "Physical edaphology," in *The Physics of Irrigated and Nonirrigated Soils*, W.H. Freeman and Company, San Francisco, xiii + 533 (1972).
38. F. Vuolo et al., "Data service platform for Sentinel-2 surface reflectance and value-added products: system use and examples," *Remote Sens.* **8**, 938 (2016).
39. S. K. McFeeters, "The use of the normalized difference water index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.* **17**(7), 1425–1432 (1996).
40. M. Sadeghi, S. B. Jones, and W. D. Philpot, "A linear physically-based model for remote sensing of soil moisture using short wave infrared bands," *Remote Sens. Environ.* **164**, 66–76 (2015).
41. W. T. Crow, W. P. Kustas, and J. H. Prueger, "Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model," *Remote Sens. Environ.* **112**(4), 1268–1281 (2008).
42. S. Liu et al., "Spectral responses to plant available soil moisture in a Californian grassland," *Int. J. Appl. Earth Obs. Geoinf.* **19**, 31–44 (2012).
43. C. Peng, M. Deng, and L. Di, "Relationships between remote-sensing-based agricultural drought indicators and root zone soil moisture: a comparative study of Iowa," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.* **7**(11), 4572–4580 (2014).
44. W. J. R. Santos et al., "Soil moisture in the root zone and its relation to plant vigor assessed by remote sensing at management scale," *Geoderma* **221-222**, 91–95 (2014).
45. M. T. Schnur, H. Xie, and X. Wang, "Estimating root zone soil moisture at distant sites using MODIS NDVI and EVI in a semi-arid region of southwestern USA," *Ecol. Inf.* **5**(5), 400–409 (2010).
46. X. Wang et al., "Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions," *J. Hydrol.* **340**(1), 12–24 (2007).
47. Y. Qiu et al., "Spatial variability of soil moisture content and its relation to environmental indices in a semi-arid gully catchment of the Loess Plateau, China," *J. Arid Environ.* **49**, 723–750 (2001).
48. J. G. B. Leenaars et al., "Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa," *Geoderma* **324**, 18–36 (2018).
49. K. T. Morgan, L. R. Parsons, and T. Adair Wheaton, "Comparison of laboratory- and field-derived soil water retention curves for a fine sand soil using tensiometric, resistance and capacitance methods," *Plant Soil* **234**(2), 153–157 (2001).
50. R. Weil and N. C. Brady, *The Nature and Properties of Soils*, 15th ed., Pearson, Columbus (2016).
51. X. Fang et al., "Variations of deep soil moisture under different vegetation types and influencing factors in a watershed of the Loess Plateau, China," *Hydrol. Earth Syst. Sci.* **20**, 3309–3323 (2016).
52. R. Muñoz-Carpena, D. D. Bosch, and A. Ritter, *Field Methods for Monitoring Soil Water Status*, Agricultural and Biological Engineering Department, UF/IFAS Extension, Florida, p. 15 (2015).
53. D. Entekhabi et al., *SMAP Handbook–Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw from Space*, JPL Publication, Pasadena, California, p. 192 (2014).
54. M. Zreda et al., "Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons," *Geophys. Res. Lett.* **35**(21), L21402 (2008).
55. D. Roberts, K. Roth, and R. Perroy, "Hyperspectral vegetation indices," in *Hyperspectral Remote Sensing of Vegetation*, CRC Press, New York, p. 309–328 (2011).

Sosdito Mananze is a PhD candidate at the Faculty of Sciences, Porto University in Portugal, and an associated lecturer at the Eduardo Mondlane University in Mozambique. He received his BSc degree in forestry from Eduardo Mondlane University in 2015 and his MSc degree in natural resources management from Evora University in 2012. He is the author of three journal papers. His current research interest includes remote sensing of agriculture and natural resources.

Isabel Pôças has a PhD in agrarian sciences, a MSc degree in genetic resources and plant breeding of agricultural and forestry species and a first degree in agrarian sciences. She holds a post-doc grant at Centro de Investigação em Ciências Geo-Espaciais—under the project "Enabling Green E-Science for the SKA." She has a teaching experience at the higher degree education level. She published 18 articles in international journals with peer-review and 5 book chapters.

Mário Cunha received his BS, MSc, and PhD degrees in agrarian sciences. He is currently an assistant professor in the Department of Geosciences, Environment and Spatial Planning at Sciences Faculty—University of Porto (FCUP), and a senior researcher at the Institute for Systems and Computer Engineering, Technology and Science (INESC-TEC), CRIIS. His research interests include various topics in remote sensing applications, agricultural engineering, and digital agriculture. In terms of academics, he teaches courses in the first, second, and third cycles, codirector of the MSc in agronomy, and the director of the PhD program in agrarian sciences.

Paper 4

Mananze, S. E., Pôças, I., & Cunha, M. (2020). Mapping and assessing the dynamics of shifting agricultural landscapes using Google Earth Engine cloud computing, a case study in Mozambique. Remote Sensing 12(8), 1279;

doi:10.3390/rs12081279

Paper available in: 17 April 2020
Classification according to journal: research paper
Main scientific areas: Remote Sensing of Agriculture

Classification according to journal: research paper

Bibliometric indicators from the Journal Citation Report, Web of Science (JCR – WoS), Clarivate Analytics:
- Journal impact factor 4.118
- Journal rank

Remote Sensing: position 7/30; 1st quartile (Q1)

remote sensing

Article

Mapping and Assessing the Dynamics of Shifting Agricultural Landscapes Using Google Earth Engine Cloud Computing, a Case Study in Mozambique

Sosdito Mananze [1,2,3,*], Isabel Pôças [1,3] and Mário Cunha [1,3,4]

[1] Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s.n., 4169-007 Porto, Portugal; isabel.pocas@fc.up.pt (I.P.); mccunha@fc.up.pt (M.C.)
[2] Escola Superior de Desenvolvimento Rural, Universidade Eduardo Mondlane, Vilankulo PC-257, Mozambique
[3] Centro de Investigação em Ciências Geo-Espaciais, Observatório Astronómico Prof. Manuel de Barros, Alameda do Monte da Virgem, 4430-146 Vila Nova de Gaia, Portugal
[4] Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Campus da Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
* Correspondence: sosdito.mananze@uem.mz; Tel.: +258-84-2045-113

Received: 16 March 2020; Accepted: 13 April 2020; Published: 17 April 2020

Abstract: Land cover maps obtained at high spatial and temporal resolutions are necessary to support monitoring and management applications in areas with many smallholder and low-input agricultural systems, as those characteristic in Mozambique. Various regional and global land cover products based on Earth Observation data have been developed and made publicly available but their application in regions characterized by a large variety of agro-systems with a dynamic nature is limited by several constraints. Challenges in the classification of spatially heterogeneous landscapes, as in Mozambique, include the definition of the adequate spatial resolution and data input combinations for accurately mapping land cover. Therefore, several combinations of variables were tested for their suitability as input for random forest ensemble classifier aimed at mapping the spatial dynamics of smallholder agricultural landscape in Vilankulo district in Mozambique. The variables comprised spectral bands from Landsat 7 ETM+ and Landsat 8 OLI/TIRS, vegetation indices and textural features and the classification was performed within the Google Earth Engine cloud computing for the years 2012, 2015, and 2018. The study of three different years aimed at evaluating the temporal dynamics of the landscape, typically characterized by high shifting nature. For the three years, the best performing variables included three selected spectral bands and textural features extracted using a window size of 25. The classification overall accuracy was 0.94 for the year 2012, 0.98 for 2015, and 0.89 for 2018, suggesting that the produced maps are reliable. In addition, the areal statistics of the class classified as agriculture were very similar to the ground truth data as reported by the Serviços Distritais de Actividades Económicas (SDAE), with an average percentage deviation below 10%. When comparing the three years studied, the natural vegetation classes are the predominant covers while the agriculture is the most important cause of land cover changes.

Keywords: classification; cropland; shifting agriculture; Landsat; land cover; google earth engine; food security; texture features; landscape dynamics; mapping; random forest

1. Introduction

Food security, sustainable agricultural production, and poverty reduction are top priorities and challenging societal issues in Mozambique. The production and productivity of the main crops is very low due to limited use of improved agricultural inputs and the post-harvest losses are high. Therefore, the access to food for the gross part of Mozambican population relying on agricultural production is only guaranteed during short periods. As such, about 24% of households are food insecure and there is an extremely high level of chronic under-nutrition (43%) which affects almost one in every two children under the age of five years [1].

Mapping the cropland extent, distribution, and characteristics is of great importance to support measures to mitigate these priority problems in Mozambique. In fact, accurate and consistent information on cropland and the dynamic location of major crop types is required to support future policy, investment, and logistical decisions [2,3], as well as production monitoring [4]. Timely cropland information directly feeds early warning systems, such as FAO Global Information and Early Warning System (GIEWS), Global Monitoring for Food Security (GMFS), and Famine Early Warning Systems Network (FEWS NET) [5,6]. Moreover, this information is also relevant for other disciplines, such as environmental impact and mitigation measures of the foreseen climate scenarios [7]. In agriculture monitoring, as well as climate modeling, cropland maps can be used as a mask to isolate agricultural land for (1) time-series analysis for crop condition monitoring and (2) to examine how the cropland responds to different climatic scenarios [3]. Additionally, timely and accurate assessment of land cover and land use is essential in understanding how climate variability affects regional crop production and for supporting the allocation of resources for medium and long-term planning. This issue is particularly relevant considering that Mozambique is highly vulnerable to extreme weather events like droughts and floods (ranks 3rd amongst African countries) and to erratic and unpredictable precipitation patterns. These conditions contribute to instability in crop production and often limit access to food. Traditional smallholder and resources-poor farming systems dominate the agricultural landscape in Mozambique, with 99% of fields smaller than 1.5 hectares and growing mainly maize (*Zea mays* L.) or cassava (*Manihot esculenta* Crantz) for own consumption [8]. These traditional farming systems are characterized by a highly dynamic shifting nature because when the production starts to decline (in a time frame of very few years), the crop fields are left for fallow and new natural areas are open for cultivation.

Smallholder farmers constitute the socio-economic backbone of the country and provide the basis for food security, preserving natural biodiversity, and providing environmental protection [5]. It is therefore imperative to adopt data and tools that enable appropriate monitoring and mapping of these traditional agricultural systems so that they can be included and valued into more global food systems. On the other hand, to meet the rapidly growing demand for food, fueled by a rapid population growth and changing diets, and to enhance economic development, pressure on natural areas (forest and savannahs) is increasing [9,10]. Cultivating more land to increase yield has put natural land and water resources under intense pressure [11,12]. The introduction of agricultural machinery, the use of irrigation systems and fertilizers are contributing to rapidly modify the land surface. Given the rate of agricultural area expansion, the shifting nature of the dominant agricultural system and considering Mozambique's vulnerability to risks from droughts and floods, there is a need to significantly improve the information management system aiming to feed crop statistics as well as to monitor and support the sustainable use of soil, water, and other resources [1]. In such context, the timely, accurate, and cost-efficient mapping of these peculiar and evolving agricultural systems in Mozambique, at fine scales, is paramount.

Remote sensing, through Earth Observation (EO) imagery, has been critical for land cover classification thanks to its wide area and time-repeated coverage. Several global land cover products have been developed and made publicly available. Among others, the following products are the most widespread: (i) the Global Land Cover database for the year 2000 (GLC2000), a SPOT4-based map produced by the European Commission's Joint Research Center (JRC), with a spatial resolution of about 1 km at the equator and larger at higher latitudes [13–15]; (ii) the Globcover, produced by the European Space Agency with ENVISAT-MERIS data, with a 300 m of spatial resolution [16]; (iii) the Moderate-resolution Imaging Spectroradiometer (MODIS) Collection 5 Land Cover Product, with 500 m of spatial resolution [17]; (iv) the Global Land Cover Characterization database (GLCC), resulting from a concerted effort between the U.S. Geological Survey (USGS), University of Nebraska Lincoln (UNL) and the JRC, which is based on the Advanced Very High Resolution Radiometer (AVHRR) and has a nominal spatial resolution of 1 km [18]; (v) the Global Food Security-support Analysis Data Product (GFSAD1000), with a nominal scale of 1 km [19]; (vi) the global Spatial Production Allocation Model (SPAM) dataset, which utilizes several datasets and ancillary information, with a spatial resolution of about 10 km at the equator [13]; (vii) the International Institute for Applied Systems Analysis-International Food Policy Research Institute (IIASA-IFPRI) cropland product, a cropland percentage hybrid map generated using some of the datasets mentioned above, with a spatial resolution of 1 km [2]; the National Geomatic Center of China (NGCC) 30 m global land cover (Globeland 30), based on Landsat and China Environmental Disaster Alleviation Satellite (HJ – 1) [20].

In almost all the above-mentioned products, the coarse spatial resolution hinders the careful assessment of cropland class, which is characterized by a large variety of agro-systems with a dynamic nature. This issue is more pronounced in regions with prevalence of traditional smallholder and resource-poor farming systems, such as in many African countries [3,13,21]. In addition, the authors of [3,12,22] argued that some of these datasets are often not reliable over croplands, as they show important disagreement with each other and with national statistics. Other important limitations of the current global land cover maps are the lack of temporal updates, the limited number of land cover classes, and the types of classes defined, which are of limited use in finer scale applications [13].

There is growing interest in mapping smallholder dominated landscapes [23–31]. In general, smallholder cropland mapping is hindered by characteristics such as small field size, heterogeneity in management practices, landscape fragmentation, and the widespread presence of trees within the fields [24,32]. These characteristics differ widely between countries, agro-ecological zones, socio-economic context, and even region or province, making difficult the definition of the kind of spatial resolution and data input combinations for accurately mapping smallholders, as discussed by [30,32,33]. As such, there are studies reporting accurate land cover classification results within smallholder dominated landscapes using high-resolution imagery [29], very high-resolution imagery [23], or a combination [27]. The classification input data and the features calculated from the data, e.g., textural features at different window sizes, also impact on the classification accuracy, e.g., [33–37]. The textural features represent the spatial variation in contiguous pixel values and thus provide a quantitative measure of spatial context. These texture features are extracted through the application of multiple functions to the image bands at different window sizes, and thus can result in a large volume of input data with potentially relevant information for the classification [37]. Although the use of different sets of features in land cover classification has been widely explored, e.g., [34,35,38], more efforts are needed towards assessing the best combination of features when classifying shifting smallholder agricultural landscapes typical in several African countries as in Mozambique. According to [32] the selection of appropriate imagery and input feature for shifting smallholder classification is probably a context-specific issue.

The development of powerful cloud-based computational frameworks, coupled with the increasing accessibility of imagery resulting from the Landsat Global Archive Consolidation (LGAC) initiative [39] and the European Commission Copernicus programme (Sentinel 2A and 2B) [40] is making custom classification more accessible, and have the potential to help overcome the limitations of existing products. An example of the combined use of these assets is the Google Earth Engine (GEE), a planetary-scale platform for geospatial data science powered by Google's cloud platform [41]. As pointed out by [13], GEE is an invaluable tool by making available a huge data pool of satellite imagery and access to robust algorithms that can be applied for diversified types of uses, including image classification. In fact, numerous studies have been developed to perform classification within GEE at regional [42–45] and global scales [46,47]. Amongst others, the frequently used classifiers include random forest (RF) [48]; support vector machine (SVM) [49], maximum entropy model (MaxEnt) [50]; classification and regression trees (CART) [51]. The RF, which is a classifier based on an ensemble of classification trees, has been demonstrated to provide higher accuracies in image classification applications than other techniques [52,53], independent of the type of data used and of the targeted environment. Moreover, RF is recognized by its ability to cope with a large set of input variables, which results from the use of a random subset of input features in the division of every node [37].

As highlighted by [54], the rising availability of improved RS data and tools offers the opportunity for the development of land cover and land uses maps at finer spatial and temporal resolutions, which are suitable to support applications in an agricultural setting with predominance of smallholder and resource-poor agricultural systems. This is particularly crucial in the face of rising climate variability and the associated uncertainties on smallholder agriculture [40]. Additionally, according to [3], Mozambique is one of the three Southern African countries of high priority for cropland mapping and which would really benefit from a timely land cover/land use update.

Most previous studies on smallholder cropland mapping considered only a single date or a single season time scale. However, making use of multitemporal imagery (from various years) is needed to capture both the spatial dynamics and the temporal shifting characteristic of smallholder agriculture in Mozambique.

In this context, the main objective of the current study is to develop and test an automated land cover classification and mapping for Vilankulo district in Mozambique using GEE cloud computing and based on different types of features derived from Landsat imagery for three years within the period between 2012 and 2018. The types of features tested include Landsat spectral bands, spectral vegetation indices, and textural features derived from the spectral bands. The specific objectives include (i) evaluating the best combination of input variables for the classification process; (ii) assessing the most adequate window size for the texture input features; (iii) assessing the spatial and temporal dynamics of agriculture in Vilankulo; (iv) comparing the statistics of agricultural areas as extracted from the produced maps to those yearly reported by the Local Government entities.

2. Materials and Methods

2.1. Study Area Description

The current study was conducted in the district of Vilankulo, within the province of Inhambane in southern Mozambique (Figure 1), and includes agricultural seasons of the years 2012, 2015, and 2018.

Climatologically, Vilankulo is characterized by a semi-arid to arid climate and presents two distinct seasons, a warm and rainy season running from October to March and a cold and dry season that runs from April to August.

The district of Vilankulo shapes well the complexity and heterogeneity of a landscape characterized by diversified land covers types, including agriculture, open and close canopy natural vegetation, water bodies, grasslands, bare soils, and human settlement areas. The agriculture is characterized by two cropping seasons, one during the rainy season and another during the dry season. The agriculture is mainly rainfed and is practiced only in the rainy season in mixed crop systems (maize, peanuts, beans, cassava, etc.); generally multipurpose trees species are also left within the crop fields for shade, fruits, medicines, and firewood collection. During the dry season, the rainfed croplands are not cultivated. In addition, rainfed croplands may be left for fallow after three or four years of consecutive cropping seasons, when the crop production drops considerably, and new natural areas are open for cultivation. Large scale irrigated crops (mainly maize, beans, soybean, and potatoes) have been introduced more recently and although grown throughout the entire year, it is still restricted to a small number of zones within the study area. Additionally, during the dry season, the cultivation is concentrated in areas neighboring water bodies (lagoons and rivers), where vegetables (e.g., tomatoes, cabbage, onion) are grown in small areas using handheld watering cans or small scales pumping irrigation systems.

The open-canopy natural vegetation (OCNV) comprises areas with vegetation bellow 5 m height and canopy cover below 10%; close canopy vegetation (CCNV) corresponds to areas with vegetation above 5 m height and canopy cover above or equal to 10%; water bodies (WB) comprise areas periodically or permanently covered by water; grasslands areas (GR) are primarily covered by grass with or without scattered shrubs; bare soils (BS) are uncovered areas which in the study area include prepared areas for agriculture, sand bands along water bodies; and agriculture (AGR) consists of cultivated fields.

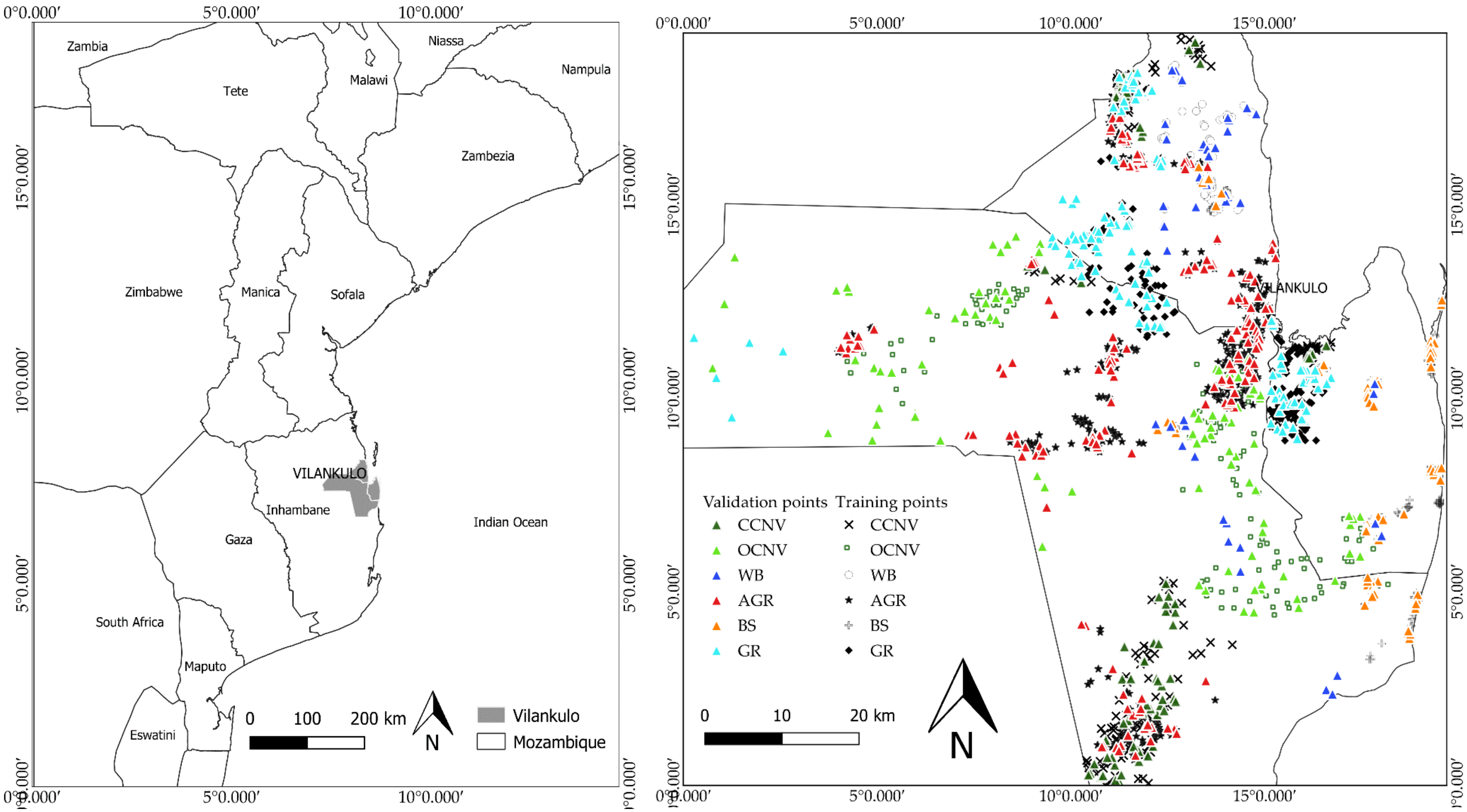


Figure 1. Location of Vilankulo district in Mozambique (**left**) and map of Vilankulo district (**right**) with location of the training and validation datasets. CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland.

2.2. Landsat Image Selection and Pre-Processing

Due to its high number of bands, Landsat 8 OLI/TIRS was taken as the prime source of data for the study and a search for all USGS Surface Reflectance (SR) images from the Landsat 8 OLI/TIRS sensor covering Vilankulo district in the main cropping seasons of 2015 and 2018 (October to March) was performed within the GEE's data pool. The images used for each year are listed in Table 1.

The study area is covered by three Landsat scenes of the path/row 166/75, 166/76, and 167/75.

Pre-processing of Landsat 8 OLI/TIRS images consisted of application of quality bands to remove cloud contamination.

Table 1. Landsat 8 OLI/TIRS used in this study.

Season	Path/Row	Date
2015/2016	166/75	6 October 2015
	166/76	6 October 2015 23 November 2015
	167/75	13 October 2015 14 November 2015
2018/2019	166/75	14 October 2018 1 December 2018 17 December 2018
	166/76	15 November 2018 1 December 2018 17 December 2018
	167/75	5 October 2018 6 November 2018 8 December 2018 24 December 2018

For the year 2012, prior to the Landsat 8 OLI/TIRS launch (February 2013), a search for Landsat 7 ETM+ was done and six atmospherically corrected surface reflectance images and with cloud cover below 5% were identified. In this year, there was an anticipation of cropping season, taking advantage of good meteorological conditions observed in September. As such, an image of September was used for the scene with path/row 167/75. The USGS Gap-Fill algorithm, as implemented within GEE, was applied to fill the SLC-failure gaps in all selected Landsat 7 ETM+ images. The methodology involves using data from other scenes of the same reference (source scenes), to fill the gaps in the scenes of interest (fill scenes). The source scenes were selected within a time interval of less than four months relative to the date of the fill scenes (Table 2).

Table 2. Landsat 7 ETM+ imagery used in the study.

Season	Path/Row	Date (Fill Scenes)	Date (Source Scenes)
2012/2013	166/75	5 October 2012	25 January 2013
	166/76	5 October 2012	6 November 2012
	167/75	10 September 2012	13 November 2012

2.3. Texture Features

Texture is a distinctive property of all surfaces. It contains important information about the spatial distribution, structural arrangement of surfaces, and their relationship to the surrounding environment [55]. Various methods can be used to derive textural information, for example, the first and second-order local statistics [56] and grey level co-occurrence matrix (GLCM) measurements [55]. Local statistics (LS) describe the moments of a neighborhood of individual pixels in an image region, while GLCM-based texture features characterize the distance and angular relationships among pixels. The GLCM is a tabulation of how often different combinations of pixel brightness values (grey levels) occur in an image. It counts the number of times a pixel of value X lies next to a pixel of value Y, in a given direction and distance, and then derives statistics from this tabulation. GLCM can be classified into two groups: (i) occurrence or first order measures, which describe a statistical property, such as pixel variance or skewness within a specified neighborhood of an image [57,58]; and (ii) co-occurrence or second order measures, which characterize the relative frequencies between two pixel brightness values linked by a spatial relation [59,60]. The second order GLCM texture measures have been the most common textural features used in remote sensing studies, including for forest structure modelling [61–63], habitat modelling [64], and land cover classification [65–67].

In this study, the 14 GLCM metrics proposed by [55], namely Angular Second Moment (asm), Contrast (contrast), Correlation (corr), Variance (var), Inverse Difference Moment (idm), Sum Average (savg), Sum Variance (svar), Sum Entropy (sent), Entropy (ent), Difference variance (dvar), Difference entropy (dent), Information Measure of correlation 1 (imcorr1), Information Measure of correlation 2 (imcorr2), and maximum correlation coefficient (maxcorr), and an additional four textural features from [68], specifically, Dissimilarity (diss), Inertia (inertia), Cluster Shade (shade) and Cluster prominence (prom), were computed for each Landsat band. As the value of texture variables depends mostly on window size and spatial resolution [61], six different window sizes (2, 5, 10, 20, 25, and 30) were tested while the 3 × 3 square default kernel size was kept constant.

2.4. Vegetation Indices

Four vegetation indices frequently used in land cover and agricultural studies were computed: normalized difference vegetation index (NDVI) [69]; soil adjusted vegetation index (SAVI) [70]; enhanced vegetation index (EVI) [71]; and the normalized water difference index (NWDI) [72]. NDVI is one of the most used vegetation indices for vegetation phenology characterization thanks to its sensitivity to photosynthetically active biomass and its ability to discriminate vegetation/non-vegetation, as well as wetland/non-wetland cover classes. SAVI was included to make use of its ability to minimize the soil background effect, as the study area presents large areas of sparse vegetation. EVI can be useful in areas of high leaf area index (LAI) where NDVI may saturate and NDWI is useful to discriminate water from land due to its sensitiveness to open water [73].

2.5. Image Composites

For each year studied, when more than one image per scene existed, an aggregation method on a per-pixel, per-band basis was applied by using the median value. The Pearson's correlation coefficient was used to assess the correlation between Landsat bands. Bands of the same spectral region showed correlation coefficient greater than 0.90 for both visible and short-wave infrared regions. As such, only three bands were selected for further analysis, being the band 1, B1 (presenting the least correlation coefficient (0.24) with the B5); band 5, B5 (presenting the least correlation coefficient with all bands (0.36) in average); and band 6, B6 (presenting the least correlation with all bands (0.60), as compared to the B7, of the same spectral region).

The final image composites for the years 2015 and 2018 comprised 61 features: (i) three reflective Landsat 8 bands (Band 1—deep blue; Band 5—NIR; Band 6—SWIR1), (ii) four vegetation indices (NDVI, SAVI, EVI, and NWDI), and (iii) 54 GLCM textural features (18 GLCM × 3 bands). For the year 2012, the final image composites included the same numbers of features, but the textural features were extracted from reflective Landsat 7 bands (Band 1—blue; Band 4—NIR; Band 5—SWIR1).

2.6. Training and Validation Samples Collection

Six land cover types, characteristic of the study area landscape, were considered for image classification: (i) Close canopy natural vegetation (CCNV), (ii) Open canopy natural vegetation (OCNV), (iii) Water Bodies (WB), (iv) Agriculture, (v) Bare soils (BS), and Grassland (GR). Each land cover type is characterized according to the description provided in Section 2.1. The urban settlements were excluded from the classification process by using a mask derived from available cartography.

Global positioning system (GPS) points of each land cover type were collected during two field campaigns in 2017 and in 2019 and used as references during the on-screen collection of training data. During the field campaigns, farmers were asked about cultivation in previous years in order to ensure that only croplands operating for at least two consecutive years were included. The GPS points were then overlaid to the Landsat RGB and to very high-resolution Google Earth imagery of the corresponding cropping season and used as reference to select additional training samples. For the years where no ground data were available, the training data was selected based on visual inspection and interpretation of both Google Earth very high-resolution imagery and Landsat RGB composite. Visual interpretation of reference imagery was based on elements that help to identify land cover features such as location, size, shape, tone, color, shadow, texture, and pattern. A similar procedure was used by [74,75].

The validation data were generated by a stratified random sampling method from the classification results. The method was performed for each map (2012, 2015, and 2018) and, for each cover class, setting the number of validation points to at least 30% of the respective training points. This procedure enabled independence between training and validation data. It was also implemented in previous studies by [34]. Table 3 presents the number of training and validation samples used in each year and Figure 1 shows the spatial distribution of training and validation samples within the study area.

Table 3. Number of training and validation samples per class and per year. CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland.

Cover Class	2012		2015		2018	
	Training	Validation	Training	Validation	Training	Validation
CCNV	319	109	331	110	169	88
OCNV	268	78	205	60	119	93
WB	267	82	166	57	67	42
AGR	551	154	606	177	316	185
BS	181	40	342	172	138	93
GR	330	116	375	108	221	132

2.7. Input Features Selection for Classification

For the selection of optimal input features, the three non-correlated Landsat spectral bands, four vegetation indices (VIs), and the 18 GLCM textural features extracted in the three Landsat bands were initially considered.

Subsequently, combinations involving one or more types of the above features were established, including (Table 4) (i) Landsat bands and 18 GLCM per Landsat band (T1: Angular Second Moment (asm), Contrast (contrast), Correlation (corr), Variance (var), Inverse Difference Moment (idm), Sum Average (savg), Sum Variance (svar), Sum Entropy (sent), Entropy (ent), Difference variance (dvar), Difference entropy (dent), Information Measure of correlation 1 (imcorr1), Information Measure of correlation 2 (imcorr2), Maximum Correlation Coefficient (maxcorr), Dissimilarity (diss), Inertia (inertia), Cluster Shade (shade), Cluster prominence (prom)); (ii) Landsat bands and five GLCM textural features (T2: contrast, corr, variance, sevg, diss); (iii) Landsat bands and five GLCM textural features (T3: asm, shade, ent, prom, inertia); (iv) Landsat bands and six GLCM textural features (T4: asm, contrast, var, corr, ent, and idm); (v) Landsat bands; (vi) Landsat bands and four VIs.

For the combinations involving textural features, six different window sizes (2, 5, 10, 20, 25, and 30) were considered.

The subsets of five and six GLCM textural features selected for the combinations tested (#: 7–20 in Table 4) were identified among the most relevant according to the literature (asm, contrast, variance, correlation, entropy, inverse difference moment, dissimilarity) [60,76–79]. The textural features proposed by [68], were also included. These are suggested to improve the perceptual concepts of uniformity, proximity, and the qualities of texture periodicity and gradient, when combined with the so called most relevant GLCM textural metrics. The inverse divergence moment (idm) indicates the image homogeneity as it assumes larger values for smaller grey tone differences in pair elements. Contrast is a degree of spatial frequency, the difference between the maximum and the minimum values of a neighboring set of pixels; high contrast values denote high coarse texture. The dissimilarity (diss) informs about the heterogeneity of the grey levels; coarser textures present higher values of dissimilarity. Entropy (ent) indicates the disorder of an image; non-texturally uniform image presents very large entropy. Angular second moment (asm) is a measure of textural uniformity (pixel pair repetition); high asm values occur when the grey level distribution has either a contrast or a periodic form. Correlation (corr) indicates the grey tone linear dependencies in the image; high correlation values imply a linear relationship between the grey levels of pixel pairs. More information about the GLCM features can be found in [55,78].

Table 4. Feature combinations as tested and average separability per class of land cover type, per feature combination, and average separability per feature combination.

#	Feature Combinations				Total Features	Average Separability per Classes						Average
	Textures	W	VIs	# Bands		CCNV	OCNV	WB	AGR	BS	GR	
1	T1	2	Nil	3	57	1.66	0.38	0.37	0.37	0.44	0.45	0.61
2	T1	5	Nil	3	57	1.37	0.54	1.30	1.62	0.89	0.60	1.05
3	T1	10	Nil	3	57	1.42	0.72	1.26	1.62	1.10	1.08	1.20
4	T1	20	Nil	3	57	1.52	1.02	1.30	1.62	1.69	1.49	1.44
5	T1	25	Nil	3	57	1.03	1.33	1.23	1.22	1.53	1.50	1.30
6	T1	30	Nil	3	57	0.91	1.33	1.14	1.13	1.35	1.50	1.23
7	T2	2	Nil	3	18	1.67	0.39	0.38	0.38	0.45	0.46	0.62
8	T2	5	Nil	3	18	1.61	0.83	1.61	0.83	0.90	0.88	1.11
9	T2	10	Nil	3	18	1.66	0.73	1.27	1.82	1.11	0.79	1.23
10	T2	20	Nil	3	18	1.23	1.32	1.29	1.37	1.69	1.34	1.37
11	T2	25	Nil	3	18	1.22	1.59	1.54	1.23	1.54	1.60	1.45
12	T2	30	Nil	3	18	1.13	1.60	1.48	1.14	1.36	1.60	1.38
13	T3	2	Nil	3	18	0.32	0.31	1.03	0.27	0.43	0.27	0.44
14	T3	5	Nil	3	18	0.26	0.25	1.03	0.31	0.43	0.32	0.44
15	T3	10	Nil	3	18	0.06	0.07	0.05	0.06	0.20	0.12	0.10
16	T3	20	Nil	3	18	0.09	0.07	0.13	0.06	0.16	0.18	0.11
17	T3	25	Nil	3	18	0.12	0.07	0.13	0.22	0.18	0.30	0.17
18	T3	30	Nil	3	18	0.19	0.13	0.22	0.17	0.38	0.25	0.22
19	T4	2	Nil	3	18	0.40	0.39	1.35	0.10	0.47	0.34	0.51
20	T4	5	Nil	3	18	0.31	0.30	1.27	0.31	0.46	0.32	0.50
21	T4	10	Nil	3	18	0.06	0.05	0.05	0.05	0.17	0.13	0.08
22	T4	20	Nil	3	18	0.08	0.09	0.08	0.42	0.71	0.48	0.31
23	T4	25	Nil	3	18	0.14	0.56	0.34	0.40	0.71	0.32	0.41
24	T4	30	Nil	3	18	0.09	0.11	0.19	0.05	0.08	0.16	0.11
25	Nil		4	3	7	0.21	0.20	0.05	0.22	0.40	0.23	0.22
26	Nil		Nil	3	3	0.43	0.39	1.56	0.42	0.71	0.48	0.66

T1—textural features: Angular Second Moment (asm), Contrast (con*trast), Correlation (corr), Variance (var), Inverse Difference Moment (idm), Sum Average (savg), Sum Variance (svar), Sum Entropy (sent), Entropy (ent), Difference variance (dvar), Difference entropy (dent), Information Measure of correlation 1 (imcorr1), Information Measure of correlation 2 (imcorr2), Maximum Correlation Coefficient (maxcorr), Dissimilarity (diss), Inertia (inertia), Cluster Shade (shade), Cluster prominence (prom); T2—textural features: asm, shade, ent, prom, inertia; T3—textural features: contrast, corr, variance, sevg, diss; T4—textural features: asm, contrast, var, corr, ent, and idm; VIs—vegetation indices; W—window size for texture features; CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland; #—number of order.

The Jeffreys–Matusita distance (JM) was computed for each feature combination and all possible pair of classes. The JM is a pair-wise measure of class separability based on the probability distribution of two classes [80–82]. JM ranges from 0 to 2.0; values below 1.0 indicate that two classes are not separable; values between 1.0 and 1.5 indicate that two classes are separable but with some degree of confusion, and between 1.5 and 1.9, the classes are clearly separable; completely separable classes present JM values above 1.9 [38]. The selection of final feature combination considered the largest JM for the least separable pairs of classes, the number of features in the combination, and the GLCM window size. A similar approach was applied in previous studies [38,83,84].

2.8. Classifier Algorithm

In this study we used the random forest classifier (RF) [48], within the GEE platform. The RF can be described as the collection of tree-structured classifiers. It produces multiple decision trees, using randomly selected subsets of training samples and variables. As such, a classification result is obtained in each tree trained on a subset of the original training sample; subsequently a majority vote is used to create the final classification result. This classifier algorithm is considered very fast, robust against over fitting, allows forming as many trees as the user wants, and each decision tree uses a random subset of training data and a random subset of input predictor variables, which reduces the correlation between decision trees as well as the overall computational complexity [85]. In our study, a RF with 500 trees was created and the number of features per split was set to the square root of the total number of features, which corresponds to the standard settings [86].

2.9. Accuracy Assessment

A confusion matrix and the accuracy assessment of each map were computed using the validation data set for each land cover type. The widely applied overall accuracy (OA), Kappa coefficient (KC), producer accuracy (PA), and consumer accuracy (CA) were computed and used to assess the quality of derived land cover maps. The OA determines the overall efficiency of the algorithm and is measured by dividing the total number of correctly classified samples by the total number of the testing samples. The KC indicates the degree of agreement between the ground truth data and the predicted values. The PA measures how well a pixel has been classified. It includes the error of omission which refers to the proportion of observed features on the ground that are erroneously excluded from a class. The CA measures the reliability of the map, informing how well the map represents what is really on the ground and it includes the error of commission which refers to pixels erroneously included in a class [73,87].

Furthermore, the statistics of cropped areas for each year were computed and compared to those from the Local Government entity responsible for agricultural issues within the district (Serviços Distritais de Actividades Económicas, SDAE). SDAE estimates the agricultural area (ha) and production (ton) within the district based on the share of economically active population as reported by the National Census and share of population in agricultural sector, and the number and size of croplands as reported by the National Agrarian Census.

3. Results

3.1. Input Feature Selection

The separability analysis, based on the JM distance, shows that in 12 feature combinations tested (#: 1, 7, and 13–26), the average separability for all classes is bellow one, denoting that most of the classes are not separable. In the remaining features combinations (#: 2–6 and 8–12) the average separability is between 1 and 1.5 indicating that most of classes are separable but with some level of confusion (Table 4).

Except for the class CCNV, the other classes showed high sensitivity to the window size of textural features (Table 4), increasing the separability with the increase of the window size for the combinations #1–#12. For the combinations #13–#24, the pattern of separability between classes according to the window size was less evident.

The feature combinations #4 and #11 (Table 4) yielded very similar average separability for all classes and for these combinations more homogeneous separability results among classes were obtained (Table 5: #4–#11).

Table 5. Jeffreys–Matusita distance for all possible pairs of classes as assessed for different feature combinations.

		CCNV	**OCNV**	**WB**	**AGR**	**BS**	**GR**		**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**		**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**
		#2: Bands, T1-W 5							#4: Bands, T1-W 20							#6: Bands, T1-W 30					
CCNV	#1: Bands, T1-w 2		1.97	0.17	1.97	1.96	1.98	#3: Bands, T1-w 10		1.98	1.09	0.03	1.47	1.99	#5: Bands, T1-w 25		1.98	1.03	0.04	0.57	1.99
OCNV		1.68		1.96	0.01	0.15	0.04		1.98		1.99	1.99	1.92	0.04		1.98		1.99	1.99	1.97	0.04
WB		1.65	0.01		1.96	1.95	1.97		1	1.99		0.97	1.83	1.99		1.1	1.99		0.89	1.49	1.99
AGR		1.66	0.01	0		0.09	0.08		0.69	1.97	1.52		1.53	1.99		0.07	1.99	0.92		0.75	1.99
BS		1.57	0.15	0.08	0.09		0.31		1.98	0.15	1.99	1.96		1.92		0.95	1.96	1.66	1.12		1.97
GR		1.72	0.04	0.09	0.08	0.31			1.99	0.04	1.99	1.97	0.31			1.99	0.04	1.99	1.99	1.96	
		#8: Bands, T2-W 5							#10: Bands, T2-W 20							#12: Bands, T2-W 30					
CCNV	#7: Bands, T2-w 2		1.97	0.18	1.97	1.97	1.98	#9: Bands, T2-w 10		1.98	1.09	0.35	1.47	1.99	#11: Bands, T2-w 25		1.98	1.05	0.05	0.58	1.99
OCNV		1.69		1.97	0.01	0.16	0.05		1.98		1.99	1.99	1.91	0.05		1.98		1.99	1.99	1.97	0.05
WB		1.68	0.02		1.96	1.95	1.97		1.01	1.99		0.97	1.82	1.99		1.11	1.99		0.89	1.5	1.99
AGR		1.68	0.01	0		0.1	0.09		0.69	1.97	1.53		1.54	1.99		0.07	1.99	0.93		0.77	1.99
BS		1.58	0.16	0.08	0.1		0.33		1.98	0.16	1.99	1.95		1.93		0.97	1.96	1.67	1.15		1.97
GR		1.73	0.05	0.11	0.09	0.33			1.99	0.05	1.99	1.96	0.33			1.99	0.05	1.99	1.99	1.96	
		#14: Bands, T3-W 5							#16: Bands, T3-W 20							#18: Bands, T3-W 30					
CCNV	#13: Bands, T3-w 2		0.03	0.92	0.06	0.29	0.003	#15: Bands, T3-w 10		0.01	0.17	0.02	0.21	0.03	#17: Bands, T3-w 25		0.07	0.18	0.06	0.55	0.09
OCNV		0.03		1.02	0.01	0.16	0.05		0.01		0.13	0.01	0.16	0.05		0.02		0.18	0.08	0.25	0.07
WB		0.91	1.03		1.08	1.27	0.88		0.04	0.06		0.07	0	0.29		0.01	0.05		0.4	0.07	0.29
AGR		0.06	0.01	1.08		0.11	0.09		0.02	0.21	0.004		0.12	0.09		0.04	0.11	0.07		0.27	0.04
BS		0.29	0.16	1.27	0.11		0.33		0.21	0.16	0.08	0.11		0.33		0.31	0.16	0.09	0.22		0.77
GR		0.002	0.05	0.88	0.09	0.33			0.03	0.05	0.12	0.09	0.33			0.24	0.03	0.45	0.66	0.11	

Table 5. *Cont.*

		CCNV	**OCNV**	**WB**	**AGR**	**BS**	**GR**		**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**		**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**
		#20: Bands, T4-W 5							#22: Bands, T4-W 20							#24: Bands, T4-W 30					
CCNV	#19: Bands, T4-w 2		0.03	1.18	0.06	0.28	0.003	#21: Bands, T4-w 10		0.004	0.16	0.02	0.2	0.03	#23: Bands, T4-w 25		0.07	0.19	0.04	0.07	0.08
OCNV		0.03		1.26	0.01	0.16	0.05		0		0.13	0.01	0.16	0.05		0.05		0.26	0.16	0.05	0.02
WB		1.23	1.32		1.3	1.45	1.14		0.04	0.02		0.07	0	0.29		0.01	0.45		0.04	0.01	0.46
AGR		0.06	0.01	1.36		0.1	0.09		0.02	0.01	0.01		0.09	0.09		0.12	0.99	0.17		0.02	0.01
BS		0.29	0.16	1.49	0.1		0.33		0.2	0.16	0.07	0.1		0.33		0.5	0.78	0.85	0.66		0.24
GR		0.002	0.05	1.21	0.09	0.33			0.03	0.05	0.13	0.09	0.33			0.04	0.52	0.23	0.04	0.77	
		#26: Bands																			
CCNV	#25: Bands and VIs		0.04	1.43	0.09	0.53	0.05														
OCNV		0.01		1.55	0.02	0.33	0.01														
WB		0.68	0.75		1.57	1.78	1.49														
AGR		0.05	0.01	0.79		0.21	0.19														
BS		0.28	0.19	1.03	0.12		0.68														
GR		0.02	0.06	0.56	0.11	0.38															

VIs—vegetation indices; T1—textural features: Angular Second Moment (asm), Contrast (contrast), Correlation (corr), Variance (var), Inverse Difference Moment (idm), Sum Average (savg), Sum Variance (svar), Sum Entropy (sent), Entropy (ent), Difference variance (dvar), Difference entropy (dent), Information Measure of correlation 1 (imcorr1), Information Measure of correlation 2 (imcorr2), Maximum Correlation Coefficient (maxcorr), Dissimilarity (diss), Inertia (inertia), Cluster Shade (shade), Cluster prominence (prom); T2—textural features: asm, shade, ent, prom, inertia; T3—textural features: contrast, corr, variance, sevg, diss; T4—textural features: asm, contrast, var, corr, ent, and idm; W—window size for texture features; CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland; # number of order.

Considering the results presented in Tables 4 and 5, combination #11 was selected for the classification in this study, taking over combination #4 mainly due to its relatively low number of features (18). The features in this combination include three Landsat bands and five GLCM textural measures (asm, shade, ent, prom, inertia) extracted from the three Landsat bands using window 25. For this feature combination, the least separable classes were grassland (GR) and open canopy natural vegetation (OCNV) and close canopy natural vegetation (CCNV) and agriculture (AGR). The highest separability was achieved between the following pairs of classes: GR and CCNV water bodies (WB); GR and; GR and agriculture (AGR); GR and bare soils (BS); OCNV and AGR and OCNV and WB (Table 5: #11).

3.2. Spatial and Temporal Dynamics of Agriculture

3.2.1. Classification Results

The spatial and temporal (2012, 2015, and 2018) distribution of each land cover class is mapped in Figure 2. Visual inspection of maps denotes increase of class agriculture (AGR) throughout the study period. Spatially, the cultivated area tends to span from north to the south of district while the western portion is predominately occupied by the class open canopy natural vegetation (OCNV), which is the most predominant class within all the study area.

Table 6 presents the confusion matrices derived from the analysis of the validation dataset for each year. For the three years under consideration, the least accurately classified land cover class was OCNV. Most of OCNV incorrectly classified pixels were attributed to the classes CCNV, AGR, and GR. The class WB was consistently the most accurately classified with all the validation pixels correctly classified.

Table 6. Confusion matrix of classification for the years 2012, 2015, and 2018. CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland; PA—producer accuracy; CA—consumer accuracy.

2012	**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**	**CA**
CCNV	102	7	0	0	0	0	0.94
OCNV	2	67	0	2	0	7	0.86
WB	0	0	82	0	0	0	1.0
AGR	0	1	0	153	0	2	0.99
BS	0	0	0	0	40	0	1.0
GR	0	3	0	3	0	110	0.95
PA	0.98	0.89	1.0	0.97	1.00	0.94	
2015	**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**	**CA**
CCNV	110	0	0	0	0	0	1
OCNV	0	59	0	1	0	0	0.98
WB	0	0	57	0	0	0	1
AGR	1	0	0	176	0	0	0.99
BS	0	0	0	0	172	0	1
GR	0	0	0	0	0	108	1
PA	0.99	1.00	1	0.99	1.00	1.00	
2018	**CCNV**	**OCNV**	**WB**	**AGR**	**BS**	**GR**	**CA**
CCNV	64	8	0	4	0	0	0.85
OCNV	2	64	0	5	0	1	0.82
WB	0	0	66	0	0	0	1.0
AGR	6	1	0	170	2	4	0.91
BS	1	2	0	6	113	7	0.92
GR	5	4	0	4	0	75	0.85
PA	0.85	0.78	1	0.89	0.98	0.79	

The classification overall accuracy (OA) and the kappa coefficient (KC) are presented in Table 7 for the three years. In all years both OA and KC present values consistently above 0.85, suggesting that all the classes were mapped with high accuracy.

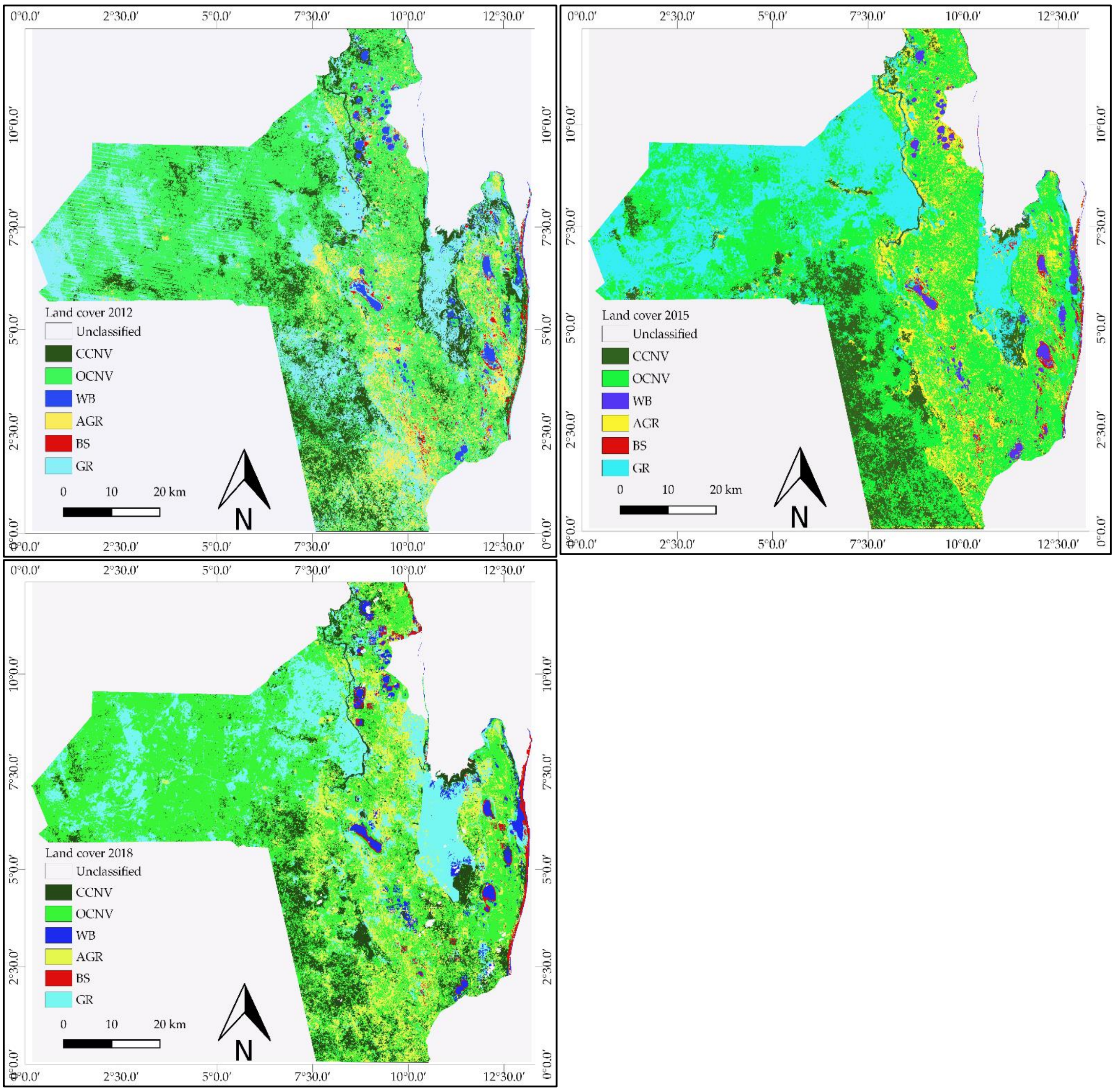


Figure 2. Land cover maps of Vilankulo for the years 2012 (upper left), 2015 (upper right), and 2018 (lower left). CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland.

Table 7. Overall accuracy and Kappa coefficient.

Years	Overall Accuracy	Kappa Coefficient
2012	0.94	0.93
2015	0.98	0.97
2018	0.89	0.87

3.2.2. Land Cover Statistics and Land Cover Changes

Tables 8 and 9 present the land cover changes statistics, respectively, from 2012 to 2015 and from 2015 to 2018, as extracted from classification maps. The results indicate that the landscape of the study area is mostly occupied by natural vegetation classes, namely, close canopy natural vegetation (CCNV), open canopy natural vegetation (OCNV), and grassland (GR). Together the three classes comprise

86%, 85%, and 82% for the years 2012, 2015, and 2018, respectively. During the two time intervals, the changes in these classes occurred mainly within them but also between each of them and the class agriculture. The class agriculture (AGR) showed an increase from 68,417.28 ha in 2012 to 75,564.90 ha in 2015 and 89,442.72 ha in 2018. The water bodies remained almost unchanged from 2012 to 2015 but a larger increase was verified in the time interval between 2015 and 2018. The class bare soils (BS) experienced very small changes throughout the analyzed period (Tables 8 and 9).

Table 8. Land cover change matrix relative to the period 2012–2015.

2012/2015	CCNV	OCNV	WB	AGR	BS	GR	AREA (ha) 2015
CCNV	42,053.31	20,482.56	316.26	187.47	57.87	9113.04	72,210.51
OCNV	32,066.91	181,246.77	460.53	34,703.28	661.05	60,404.49	309,543.03
WB	910.44	952.56	7620.30	5.31	53.91	1506.06	11,048.58
AGR	4030.38	25,792.65	174.42	31,350.96	1717.83	12,498.66	75,564.90
BS	185.22	203.22	410.22	902.97	4362.48	1464.30	7528.41
GR	9851.22	75,631.50	1631.70	1267.29	142.38	57,006.63	145,530.72
AREA (ha) 2012	89,097.48	304,309.26	10,613.43	68,417.28	6995.52	141,993.18	621,426.15
Change (2015–2012)	−16,886.97	5233.77	435.15	7147.62	532.89	3537.54	
	−18.95%	1.72%	4.10%	10.45%	7.62%	2.49%	

CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland.

Table 9. Land cover change matrix relative to the period 2015–2018. CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland.

2015/2018	CCNV	OCNV	WB	AGR	BS	GR	AREA (ha) 2018
CCNV	46,894.14	20,919.24	979.92	1966.68	24.84	3295.35	74,080.17
OCNV	20,724.75	198,877.68	518.31	26,241.84	2151.63	69,778.80	318,293.01
WB	329.31	3091.59	8768.97	397.80	259.56	1812.33	14,659.56
AGR	1383.84	48,959.73	64.35	35,482.32	1251.54	2300.94	89,442.72
BS	10.62	2577.15	428.31	1355.49	3112.65	140.67	7624.89
GR	2867.85	35,117.64	288.72	10,120.77	728.19	68,202.63	117,325.80
AREA (ha) 2015	72,210.51	309,543.03	11,048.58	75,564.90	7528.41	145,530.72	621,426.15
Change (2018–2015)	1869.66	8749.98	3610.98	13,877.82	96.48	−28,204.92	
	2.59%	2.83%	32.68%	18.37%	1.28%	−19.38%	

CCNV—close canopy natural vegetation; OCNV—open canopy natural vegetation; WB—water bodies; AGR—agriculture; BS—bare soils; GR—grassland.

3.3. Comparison between Remote Sensed Classification and Agricultural Statistics

Due to data availability constraints, the comparison between remote sensing derived areas and ground truth data from statistical census was performed only for the class agriculture (AGR), for which annual statistical information is made available by a local Governmental Entity, the 'Serviços Distritais de Actividades Económicas (SDAE)'. Figure 3 depicts the comparison between cultivated area as derived from the Landsat-derived maps and the SDAE data, denoting a good agreement between the two data sources. For the three studied years the mean percentage differences between actual ground truth data and remote sensed agrarian areas were below 10% and always negative, the highest deviation of 15.5% being observed in 2012. These measures of accuracy are in relation to the agrarian census, which also has its own inaccuracies.

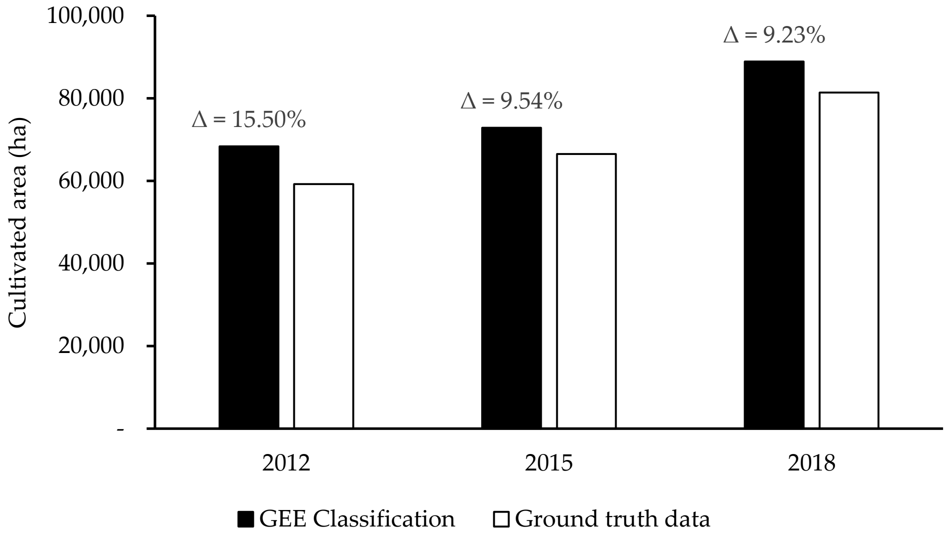


Figure 3. Comparison between cultivated area derived from remote sensing data (classification results) and from 'Serviços Distritais de Actividades Económicas (SDAE)' (Local agrarian data). The percentage of overestimation of Google Earth Engine (GEE) classification relative to ground truth data is provided for each year.

4. Discussion

The selected input features for land cover classification in the study area include five GLCM textural features (angular second moment—asm, entropy—ent, inertia—inertia, cluster shade—shade, cluster prominence—prom), and three Landsat spectral bands (Band 1—deep blue, Band 5—NIR, and Band 6—SWIR1 for Landsat 8; and Band 1—blue, Band 4—NIR, and Band 5—SWIR1 for Landsat 7). Several studies have shown that integrating optical spectral data with texture features representing the spatial variability in pixel values can improve classification accuracy [33]. The textural features asm and ent have been included as optimal features for land cover classification in previous studies, e.g., [76,78,79]. In this study, asm and ent performed poorly when combined together with other so called most relevant textural features (texture features combinations #17–#20; Table 5), but results improved when asm and ent were combined with the textural features inertia, shade, and prom (texture features combinations #1–#12; Table 5). These results suggest that textural features inertia, shade, and prom may have improved the separability of different cover classes as argued by [68]. The Landsat spectral bands selected in this study are widely used as important features for land cover classification in diversified types of landscapes [13,88,89]. The increase of window values for textural feature extraction showed notable gain in land cover class separability up to window equal to 25 for combinations #1–#12 (Table 4). The increase of class separability with the increase of window is plausible for the study area which presents high degree of spatial variability. In a previous study, the authors [35] concluded that for areas displaying a lower degree of spatial variation small window sizes perform better, while for heterogeneous landscapes texture features extracted using large window values highly improve the classification accuracies. The fact that in this study the class close canopy natural vegetation (CCNV) was insensitive to window increase in several combinations of features may be explained by the within homogeneity that characterizes this type of land cover class in the study area.

Overall, the separability between the land cover classes was good. Lower separability occurred between open canopy natural vegetation (OCNV) and grassland (GR) and close canopy natural vegetation (CCNV) and agriculture (AGR) (Tables 4 and 5). The low separability between OCNV and GR could be explained by the fact that the distinction of these classes was based on the tree/shrub cover percentage or density as observed in the field and from high-resolution imagery. Nevertheless, the two classes may present similar spectral behavior as consequence of defoliation of the trees comprising the upper canopy. Defoliation and leaf development have been indicated as the main factor of seasonality of spectral behavior in deciduous forest canopies [90,91]. When the defoliation occurs in the upper canopy trees, the spectral signature of the site is dominated by the understory vegetation. The opposite is verified in case of leaf development [90]. Other authors also reported lower separability performance between land cover classes related with sparse vegetation and low shrublands [88]. The low separability between CCNV and AGR may be explained by the fact that in the study area, forest areas are characterized by high fragmentation, with interstices occupied by other land cover types including AGR. Additionally, AGR is characterized by mixed crop systems, which generally include multipurpose trees within the fields. Additionally, in [92] the authors reported the complexity of classifying mosaic areas comprising different proportions of various homogeneous classes including forests.

The overall accuracy of the classification performed in this study (percentage values: 89%–98%; Table 7) outperformed the results of previous studies equally using the random forest algorithm with varying data types and in diversified types of landscapes and area extensions. For example, 85% overall accuracy in crop type classification using a multi-spectral time series of RapidEye images [93]; more than 80% for a time series of Landsat 7 images in homogeneous regions [94]; around 89% for a time series of Landsat 7 and 8 in heterogeneous regions [13]; up to 71.7 for Worldview 2 time series in heterogeneous regions [95] and 85.9% for crop discrimination using SPOT 5 images [96]. This may be attributed to the fact that in this study, a reduced spatial extent was under consideration, what made it easier to take the local conditions into account and tune the classification accordingly. Nevertheless, the overall accuracy of these studies reports to different land cover classes, including classes relative to different crop types. The lower classification accuracy for the year 2018 (0.89) as compared to the years 2012 (0.94) and 2015 (0.98) may equally be justified by the high fragmentation of forest areas within the study area. In fact, from Figure 2 it is noticeable that the forest fragmentation is higher in 2018, mainly due to agricultural expansion. Due to the early mentioned characteristics of agricultural systems, the confusion between this class and the forested areas is higher in 2018, what is confirmed in the confusion matrices presented in the Table 6.

Most OCNV misclassified pixels were erroneously classified as other vegetated classes (CCNV, AGR, and GR). This may be due to the seasonality in canopy spectral signature, which is characteristic in deciduous vegetation [90]. In fact, within the study area, CCNV pixels may temporally seem OCNV due to defoliation and the same process can make OCNV appear similar to AGR or GR. The inaccurately classified BS pixels as AGR (Table 6) could be due to the effect of soil background during initial stages of crop development, as various satellite images used in the classification were from dates corresponding to the initial phase of crops growth cycle (Table 1).

The land cover analysis showed that the study area landscape is predominantly covered by natural vegetation classes (CCNV, OCNV, and GR), and experiences a strong dynamic from one land cover class to another (Figure 2). Changes between the three natural vegetation classes appear to be largely due to spectral similarities caused by vegetation seasonality characteristics rather than due to conversion from one to another class. Similar findings were reported in previous studies, e.g., [35,38,78]. As was expected, the class agriculture (AGR) ranks first as the driving of effective land cover conversion, specifically for the class OCNV. During the period 2012–2018 a total of 60,945.12 ha of OCNV area was converted into AGR while 74,752.38 ha was recovered from AGR to OCNV. In the same period, a total of 16,203.78 ha was converted from all other cover classes to AGR (Tables 8 and 9). These land cover dynamic of transitions from AGR to other classes and vice-versa confirms the shifting nature of agriculture in the study area, which is characterized by systematically leaving old croplands for fallow and setting new areas for cultivation in natural vegetation areas. As pointed out by [9,11,97], in most African countries natural areas (forests and savannas) are under increasing pressure as the cultivation of more and more land is the means to meet the growing food demand. Additionally, the authors of [11,97–99] pointed out that anthropogenic land conversion is a primary cause of habitat and biodiversity loss.

The statistics of agriculture area as derived from the classification are relatively greater than those reported by the Local Governmental Entity, the SDAE, but the mean difference for the three years is rather low (slightly above 10%; Figure 3). The SDAE's statistics, as the overall agricultural statistics relying on farm surveys, have their own accuracy errors, which may result from the respondent (literacy degree, mistrust), the interviewer, the survey instrument (questionnaire design), the mode of data collection, and the incompleteness of the census list. The mean difference between the satellite-based classification and ground truth statistics from SDAE are within the average error of field surveys based agricultural censuses in developing countries (average error of 11% between the area measured and the area estimated by farmers) [99].

In a study to map cropping intensity of smallholder farms in India, the authors [25] reported a large discrepancy between cropped area estimated through remote sensing, using Landsat data, and the area of agricultural census (percentage difference of 150.8%). Likewise, [100] the authors found that irrigation area derived using MODIS 500 m data was consistently greater than the ground truth data in most of the Indian States but the percentage difference was of about 17.2%.

The good accuracy results obtained for the maps produced in this study indicate their reliability for cropland area estimation, while adding a spatially distribution information. This finding is very important considering that cropland mapping is a pre-requisite for crop monitoring through remote sensing.

The applied methodology may most likely be applied straightforward in other districts of Mozambique, given the similarities of agricultural systems. However, its application in other locations across the globe, should be considered with care as the appropriateness of a method depends on the goals of the study, the scale of analysis, and the characteristics of the agricultural system in question [30,32,33].

5. Conclusions

In this study, Google Earth Engine cloud computing was successfully applied for land cover classification in Vilankulo district in Mozambique. The best performing input variables for the random forest ensemble classifier were identified through systematic testing of different types of features. The selected input features included spectral bands and texture features. Furthermore, the best window size for the texture features computation was selected based on land cover classes separability, allowing the identification of the classes with higher sensitivity to this specific parameter in liaison to the inherent specificities of each class. The classification was applied to various years allowing to evaluate the shifting dynamics of smallholder agriculture in a time frame between 2012 and 2018.

The classes close canopy natural vegetation (CCNV), open canopy natural vegetation (OCNV), and grassland (GR) are the prevalent land cover types within the district while agriculture (AGR) is the most important cause of land cover change either by opening new cropping fields or by leaving the old ones for fallow. The cultivated area consistently increased throughout the entire study period and new spatial patterns were identified, with more concentrated agriculture areas spanning to the south east portion of the district.

The produced maps yielded high overall accuracy and the derived agricultural statistics present reasonable agreement with the ground truth statistics from SDAE.

The operational applications of this study results include timely and spatially distributed monitoring of agricultural areas, which can contribute to the improvement of agricultural statistics, to support local decisions regarding agriculture management and associated inputs management (e.g., irrigation water), to feed crop growth simulation models, among others. Additionally, the results allow the monitoring of natural vegetation dynamics for several years, which can support management decisions related to biodiversity and natural resources conservation. Moreover, on demand products may be derived based upon the methodology developed. Nevertheless, future research should address testing other types of remote sensing data (e.g., Sentinel-1 and Sentinel-2) either paired or not paired with Landsat data.

Author Contributions: Conceptualization, S.M., I.P. and M.C.; methodology, S.M., I.P. and M.C.; validation, I.P. and M.C.; formal analysis, S.M.; writing—original draft preparation, S.M.; supervision, I.P. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Sosdito Mananze acknowledges the Austrian Partnership Programme in Higher Education & Research for Development (APPEAR) for the funding through the project EO4Africa; the Ministério de Ciência e Tecnologia, Ensino Superior e Técnico Profissional, for the funding through the HEST project and Emma Izquierdo-Verdiguier for her support with the Google Earth code writing, and Isabel Pôças acknowledges the Post-Doctoral grant of the Research Infrastructure Enabling Green E-science for the SKA (ENGAGE-SKA) funded by COMPETE2020 and Fundação para a Ciência e a Tecnologia (POCI-01-0145-FEDER-022217).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. FAO. *FAO, Country Programming Framework for Mozambique 2016–2020*; FAO: Maputo, Mozambique, 2016; p. 13.
2. Fritz, S.; See, L.; McCallum, I.; You, L.; Bun, A.; Moltchanova, E.; Duerauer, M.; Albrecht, F.; Schill, C.; Perger, C.; et al. Mapping global cropland and field size. *Glob. Chang. Biol.* **2015**, *21*, 1980–1992. [CrossRef]
3. Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps. *Remote Sens.* **2015**, *7*, 7959–7986. [CrossRef]
4. Justice, C.O.; Becker-Reshef, I. *Report from the Workshop on Developing a Strategy for Global Agricultural Monitoring in the Framework of Group on Earth Observations (GEO)*; UN FAO: Rome, Italy, 2007; p. 67.
5. Hannerz, F.; Lotsch, A. Assessment of remotely sensed and statistical inventories of African agricultural fields. *Int. J. Remote Sens.* **2008**, *29*, 3787–3804. [CrossRef]
6. Vancutsem, C.; Marinho, E.; Kayitakire, F.; See, L.; Fritz, S. Harmonizing and combining existing land cover/land use datasets for cropland area monitoring at the African continental scale. *Remote Sens.* **2012**, *5*, 22. [CrossRef]
7. Lobell, D.B.; Bala, G.; Duffy, P.B. Biogeophysical impacts of cropland management changes on climate. *Geophys. Res. Lett.* **2006**, *33*. [CrossRef]
8. INE. *Censo Agro-Pecuário CAP 2009–2010: Resultados preliminares-Moçambique*; Instituto Nacional de Estatistcas: Madrid, Spain, 2011; p. 125.
9. Morris, M.B.H.; Byerlee, D.; Savanti, P.; Staatz, J. *Awakening Africa's Sleeping Giant: Prospects for Commercial Agriculture in the Guinea Savannah Zone and Beyond*; The World Bank: Washington, DC, USA, 2009.
10. United Nations. *World Population Prospects: The 2012 Revision, Highlights and Advance Tables*; United Nations: New York, NY, USA, 2013.
11. Brink, A.B.; Eva, H.D. Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach. *Appl. Geogr.* **2009**, *29*, 12. [CrossRef]
12. Ramankutty, N.; Evan, A.T.; Monfreda, C.; Foley, J.A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob. Biogeochem. Cycles* **2008**, 22. [CrossRef]
13. Azzari, G.; Lobell, D.B. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sens. Environ.* **2017**, *202*, 64–74. [CrossRef]
14. Bartholomé, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977. [CrossRef]
15. Fritz, S.; Bartholomé, E.; Belward, A.; Hartley, A.; Stibig, H.J.; Eva, H.; Mayaux, P.; Bartalev, S.; Latifovic, R.; Kolmert, S.; et al. *Harmonisation, Mosaicing and Production of the Global Land Cover 2000 Database (Beta Version)*; EC-JRC: Ispra, Italy, 2003.
16. Arino, O.; Gross, D.; Ranera, F.; Leroy, M.; Bicheron, P.; Brockman, C.; Defourny, P.; Vancutsem, C.; Achard, F.; Durieux, L.; et al. GlobCover: ESA service for global land cover from MERIS. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 2412–2415.
17. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [CrossRef]

18. Loveland, T.R.; Reed, B.C.; Brown, J.F.; Ohlen, D.O.; Zhu, Z.; Yang, L.; Merchant, J.W. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens.* **2000**, *21*, 1303–1330. [CrossRef]
19. Thenkabail, P.S.; Knox, J.W.; Ozdogan, M.; Gumma, M.K.; Congalton, R.G.; Wu, Z.; Milesi, C.; Finkral, A.; Marshall, M.; Mariotto, I.; et al. Assessing future risks to agricultural productivity, water resources and food security: How can remote sensing help? *Photogramm. Eng. Remote Sens.* **2012**, *78*, 773–782.
20. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [CrossRef]
21. Anderson, W.; You, L.; Wood, S.; Wood-Sichra, U.; Wu, W. An analysis of methodological and spatial differences in global cropping systems models and maps. *Glob. Ecol. Biogeogr.* **2015**, *24*, 180–191. [CrossRef]
22. Fritz, S.; See, L.; McCallum, I.; Schill, C.; Obersteiner, M.; van der Velde, M.; Boettcher, H.; Havlík, P.; Achard, F. Highlighting continued uncertainty in global land cover maps for the user community. *Environ. Res. Lett.* **2011**, *6*, 044005. [CrossRef]
23. Debats, S.R.; Luo, D.; Estes, L.D.; Fuchs, T.J.; Caylor, K.K. A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes. *Remote Sens. Environ.* **2016**, *179*, 210–221. [CrossRef]
24. Delrue, J.; Bydekerke, L.; Eerens, H.; Gilliams, S.; Piccard, I.; Swinnen, E. Crop mapping in countries with small-scale farming: A case study for West Shewa, Ethiopia. *Int. J. Remote Sens.* **2013**, *34*, 2566–2582. [CrossRef]
25. Jain, M.; Mondal, P.; DeFries, R.S.; Small, C.; Galford, G.L. Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. *Remote Sens. Environ.* **2013**, *134*, 210–223. [CrossRef]
26. Jin, Z.; Azzari, G.; You, C.; Di Tommaso, S.; Aston, S.; Burke, M.; Lobell, D.B. Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sens. Environ.* **2019**, *228*, 115–128. [CrossRef]
27. Lebourgeois, V.; Dupuy, S.; Vintrou, É.; Ameline, M.; Butler, S.; Bégué, A. A Combined Random Forest and OBIA Classification Scheme for Mapping Smallholder Agriculture at Different Nomenclature Levels Using Multisource Data (Simulated Sentinel-2 Time Series, VHRS and DEM). *Remote Sens.* **2017**, *9*, 259. [CrossRef]
28. McCarty, J.L.; Neigh, C.S.R.; Carroll, M.L.; Wooten, M.R. Extracting smallholder cropped area in Tigray, Ethiopia with wall-to-wall sub-meter WorldView and moderate resolution Landsat 8 imagery. *Remote Sens. Environ.* **2017**, *202*, 142–151. [CrossRef]
29. Sweeney, S.; Ruseva, T.; Estes, L.; Evans, T. Mapping Cropland in Smallholder-Dominated Savannas: Integrating Remote Sensing Techniques and Probabilistic Modeling. *Remote Sens.* **2015**, *7*, 15295–15317. [CrossRef]
30. Timmermans, A. *Mapping Cropland in Smallholder Farmer Systems in South-Africa Using Sentinel-2 Imagery*; Université Catholique de Louvain: Louvain, Belgium, 2018.
31. Vogels, M.; De Jong, S.; Sterk, G.; Douma, H.; Addink, E. Spatio-Temporal Patterns of Smallholder Irrigated Agriculture in the Horn of Africa Using GEOBIA and Sentinel-2 Imagery. *Remote Sens.* **2019**, *11*, 143. [CrossRef]
32. Lambert, M.-J.; Pierre, T.; Xavier, B.; Philippe, B.; Pierre, D. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* **2018**, *216*. [CrossRef]
33. Räsänen, A.; Virtanen, T. Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes. *Remote Sens. Environ.* **2019**, *230*, 111207. [CrossRef]
34. Chen, W.; Li, X.; He, H.; Wang, L. Assessing Different Feature Sets' Effects on Land Cover Classification in Complex Surface-Mined Landscapes by ZiYuan-3 Satellite Imagery. *Remote Sensing* **2018**, *10*, 23. [CrossRef]
35. Coburn, C.A.; Roberts, A.C.B. A multiscale texture analysis procedure for improved forest stand classification. *Int. J. Remote Sens.* **2004**, *25*, 4287–4308. [CrossRef]
36. Khatami, R.; Mountrakis, G.; Stehman, S.V. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sens. Environ.* **2016**, *177*, 89–100. [CrossRef]
37. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.M.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107. [CrossRef]
38. Wu, W.; Zucca, C.; Karam, F.; Liu, G. Enhancing the performance of regional land cover mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 422–432. [CrossRef]

39. Wulder, M.A.; White, J.C.; Loveland, T.R.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; Fosnight, E.A.; Shaw, J.; Masek, J.G.; Roy, D.P. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283. [CrossRef]
40. Kpienbaareh, D.; Kansanga, M.; Luginaah, I. Examining the potential of open source remote sensing for building effective decision support systems for precision agriculture in resource-poor settings. *GeoJournal* **2018**. [CrossRef]
41. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
42. Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore, B. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens. Environ.* **2016**, *185*, 142–154. [CrossRef]
43. Padarian, J.; Minasny, B.; McBratney, A.B. Using Google's cloud-based platform for digital soil mapping. *Comput. Geosci.* **2015**, *83*, 80–88. [CrossRef]
44. Patel, N.N.; Angiuli, E.; Gamba, P.; Gaughan, A.; Lisini, G.; Stevens, F.R.; Tatem, A.J.; Trianni, G. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35*, 199–208. [CrossRef]
45. Simonetti, D.; Simonetti, E.; Szantoi, Z.; Lupi, A.; Eva, H.D. First Results from the Phenology-Based Synthesis Classifier Using Landsat 8 Imagery. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1496–1500. [CrossRef]
46. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [CrossRef]
47. Pekel, J.-F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418. [CrossRef]
48. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
49. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
50. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
51. Breiman, L.; Friedman, J.; Stone, C.; Olshen, R. *Classification and Regression Trees*; Chapman&Hall/CRC: Boca Raton, FL, USA, 1984.
52. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
53. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 55–72. [CrossRef]
54. Pretty, J.N.; Noble, A.D.; Bossio, D.; Dixon, J.; Hine, R.E.; Penning de Vries, F.W.T.; Morison, J.I.L. Resource-Conserving Agriculture Increases Yields in Developing Countries. *Environ. Sci. Technol.* **2006**, *40*, 1114–1119. [CrossRef] [PubMed]
55. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. ManCybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
56. Kuplich, T.M.; Curran, P.J.; Atkinson, P.M. Relating SAR image texture to the biomass of regenerating tropical forests. *Int. J. Remote Sens.* **2005**, *26*, 4829–4854. [CrossRef]
57. Asner, G.P.; Palace, M.; Keller, M.; Pereira, R., Jr.; Silva, J.N.M.; Zweede, J.C. Estimating Canopy Structure in an Amazon Forest from Laser Range Finder and IKONOS Satellite Observations1. *Biotropica* **2002**, *34*, 483–492. [CrossRef]
58. Wright, C.; Gallant, A. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sens. Environ.* **2007**, *107*, 582–605. [CrossRef]
59. Augusteijn, M.F.; Clemens, L.E.; Shaw, K.A. Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 616–626. [CrossRef]
60. Baraldi, A.; Parmiggiani, F. An investigation of the textural characteristics associated with Gray Level Co-occurrence Matrix statistical parameters. *Geosci. Remote Sens. IEEE Trans.* **1995**, *33*, 293–304. [CrossRef]
61. Castillo-Santiago, M.; Ricker, M.; Jong, B. Estimation of tropical forest structure from SPOT5 satellite images. *Int. J. Remote Sens.* **2010**, *31*, 2767–2782. [CrossRef]

62. Kayitakire, F.; Hamel, C.; Defourny, P. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sens. Environ.* **2006**, *102*, 390–401. [CrossRef]
63. Ozdemir, I.; Karnieli, A. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 701–710. [CrossRef]
64. Estes, L.D.; Reillo, P.R.; Mwangi, A.G.; Okin, G.S.; Shugart, H.H. Remote sensing of structural complexity indices for habitat and species distribution modeling. *Remote Sens. Environ.* **2010**, *114*, 792–804. [CrossRef]
65. Beekhuizen, J.; Clarke, K.C. Toward accountable land use mapping: Using geocomputation to improve classification accuracy and reveal uncertainty. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 127–137. [CrossRef]
66. Kimothi, M.; Dasari, A. Methodology to map the spread of an invasive plant (Lantana camara L.) in forest ecosystems using Indian remote sensing satellite data. *Int. J. Remote Sens.* **2010**, *31*, 3273–3289. [CrossRef]
67. Pacifici, F.; Chini, M.; Emery, W.J. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sens. Environ.* **2009**, *113*, 1276–1292. [CrossRef]
68. Conners, R.W.; Trivedi, M.M.; Harlow, C.A. Segmentation of a high-resolution urban scene using texture operators. *Comput. Vis. Graph. Image Process.* **1984**, *25*, 273–310. [CrossRef]
69. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of Third Earth Resources Technology Satellite Symposium, Washington, DC, USA, 10–14 December 1973; p. 9.
70. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
71. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
72. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
73. Mahdianpari, M.; Salehi, B.; Mohammadimanesh, F.; Homayouni, S.; Gill, E. The First Wetland Inventory Map of Newfoundland at a Spatial Resolution of 10 m Using Sentinel-1 and Sentinel-2 Data on the Google Earth Engine Cloud Computing Platform. *Remote Sens.* **2018**, *11*, 43. [CrossRef]
74. De Alban, J.D.; Connette, G.; Oswald, P.; Webb, E. Combined Landsat and L-Band SAR Data Improves Land Cover Classification and Change Detection in Dynamic Tropical Landscapes. *Remote Sens.* **2018**, *10*, 306. [CrossRef]
75. Qadir, A.; Mondal, P. Synergistic Use of Radar and Optical Satellite Data for Improved Monsoon Cropland Mapping in India. *Remote Sens.* **2020**, *12*, 522. [CrossRef]
76. Cossu, R. Segmentation by means of textural analysis. *Pixel* **1988**, *1*, 4.
77. Carr, J.R.; Miranda, F.P.d. The semivariogram in comparison to the co-occurrence matrix for classification of image texture. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1945–1952. [CrossRef]
78. Rao, P.; Seshasai, M.; Kandrika, S.; Rao, M.; Rao, B.; Dwivedi, R.; Venkataratnam, L. Textural analysis of IRS-1D panchromatic data for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 3327–3345. [CrossRef]
79. Solberg, A.H.S. Contextual data fusion applied to forest map revision. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1234–1243. [CrossRef]
80. Mishra, N.B.; Crews, K.A. Mapping vegetation morphology types in a dry savanna ecosystem: Integrating hierarchical object-based image analysis with Random Forest. *Int. J. Remote Sens.* **2014**, *35*, 1175–1198. [CrossRef]
81. Richards, J.A.; Jia, X. *Remote Sensing Digital Image Analysis: An Introduction*, 4th ed.; Springer: Berlin, Germany, 2006.
82. Swain, P.H.; King, R.C. Two effective feature selection criteria for multispectral remote sensing. In Proceedings of the International Joint Conference on Pattern Recognition, Washington, DC, USA, 1 January 1973; p. 6.
83. Swain, H.; Davis, S.M. *Remote Sensing: The Quantitative Approach*; McGraw-Hill: New York, NY, USA, 1978.
84. Turner, M.G. Landscape Ecology: The Effect of Pattern on Process. *Annu. Rev. Ecol. Syst.* **1989**, *20*, 171–197. [CrossRef]
85. Breiman, L.; Cutler, A. Random Forests Homepage. Available online: http://www.stat.berkeley.edu/~{}breiman/RandomForests/cc_home.htm (accessed on 29 May 2019).

86. Akar, Ö.; Güngör, O. Integrating multiple texture methods and NDVI to the Random Forest classification algorithm to detect tea and hazelnut plantation areas in northeast Turkey. *Int. J. Remote Sens.* **2015**, *36*, 442–464. [CrossRef]
87. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]
88. Pôças, I.; Cunha, M.; Marçal, A.; Pereira, L. An evaluation of changes in a mountainous rural landscape of Northeast Portugal using remotely sensed data. *Landsc. Urban Plan.* **2011**, *101*, 253–261. [CrossRef]
89. Xu, Y.; Yu, L.; Zhao, F.R.; Cai, X.; Zhao, J.; Lu, H.; Gong, P. Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa. *Remote Sens. Environ.* **2018**, *218*, 13–31. [CrossRef]
90. Samanta, A.; Knyazikhin, Y.; Xu, L.; Dickinson, R.E.; Fu, R.; Costa, M.H.; Saatchi, S.S.; Nemani, R.R.; Myneni, R.B. Seasonal changes in leaf area of Amazon forests from leaf flushing and abscission. *J. Geophys. Res. Biogeosci.* **2012**, *117*. [CrossRef]
91. Wu, Q.; Song, C.; Song, J.; Wang, J.; Chen, S.; Yu, B. Impacts of Leaf Age on Canopy Spectral Signature Variation in Evergreen Chinese Fir Forests. *Remote Sens.* **2018**, *10*, 262. [CrossRef]
92. Carrao, H.; Sarmento, P.; Araújo, A.; Caetano, M. *Separability Aanalysis of Land Cover Classes at Regional Scale: A Comparative Study of MERIS and MODIS Data*; European Space Agency, (Special Publication) ESA SP: Paris, France, 2007.
93. Nitze, I.; Schulthess, U.; Asche, H. Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. Proceedings of 4th Conference on GEographic Object-Based Image Analysis, Rio de Janeiro, Brazil, 7–9 May 2012; p. 35.
94. Tatsumi, K.; Yamashiki, Y.; Canales Torres, M.A.; Taipe, C.L.R. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. *Comput. Electron. Agric.* **2015**, *115*, 171–179. [CrossRef]
95. Aguilar, R.; Zurita-Milla, R.; Izquierdo-Verdiguier, E.; de By, R.A. A Cloud-Based Multi-Temporal Ensemble Classifier to Map Smallholder Farming Systems. *Remote Sens.* **2018**, *10*, 729. [CrossRef]
96. Ok, A.O.; Akar, O.; Gungor, O. Evaluation of random forest method for agricultural crop classification. *Eur. J. Remote Sens.* **2012**, *45*, 421–432. [CrossRef]
97. Jacobson, A.; Dhanota, J.; Godfrey, J.; Jacobson, H.; Rossman, Z.; Stanish, A.; Walker, H.; Riggio, J. A novel approach to mapping land conversion using Google Earth with an application to East Africa. *Environ. Model. Softw.* **2015**, *72*, 1–9. [CrossRef]
98. Pimm, S.L.; Jenkins, C.N.; Abell, R.; Brooks, T.M.; Gittleman, J.L.; Joppa, L.N.; Raven, P.H.; Roberts, C.M.; Sexton, J.O. The biodiversity of species and their rates of extinction, distribution, and protection. *Science* **2014**, *344*, 1246752. [CrossRef] [PubMed]
99. De Groote, H.; Traoré, O. The cost of accuracy in crop area estimation. *Agric. Syst.* **2005**, *84*, 21–38. [CrossRef]
100. Dheeravath, V.; Thenkabail, P.; Chandrakantha, G.; Noojipady, P.; Reddy, G.P.O.; Biradar, C.; Gumma, M.; Velpuri, N.M. Irrigated areas of India derived using MODIS 500 m time series for the years 2001–2003. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 42–59. [CrossRef]

Section IV| Operationalization

1| Applications of research results

The results presented in this PhD thesis relate with remote sensing (RS) applications in agriculture, spanning from case studies focusing the estimation and monitoring of leaf area index (LAI), soil moisture (SM) and agricultural drought, and land cover mapping. These results are of unquestionable relevance as source of information for farm level decision support system, for food aid interventions, as well as for policy makers. Such type of information can be particularly useful in Mozambique conditions, with potential direct impacts on improving the agricultural productivity, the food security, the population livelihood, and the sustainable management of resources.

The information on crop development and condition provided through RS can allow farmers to easily identify and react to inter and intra-field variability in farmlands, allowing them to take appropriate management decisions. Monitoring crop growth and performance during crop growing stages enables the farmer to implement timely interventions ensuring that optimal yield is obtained at the end of the season. Stress factors that often prevent crops from developing (e.g., water and/or nutrient deficiency, fungal, bacterial or viral diseases, attack from insects or other organisms) have great impact in canopy foliage and biomass (Kpienbaareh, Kansanga et al. 2018, ESA. 2019).

The LAI, focused on papers 1 and 2, is a widely used index in agriculture (Richter, Atzberger et al. 2012). It characterizes plant canopy structure and gives an idea of the amount of biomass available in a field. It is also considered a measure of crop growth and productivity. As such, LAI maps of different growth stages can be used to timely monitor crops throughout the cropping season and to estimate crop yield (Atzberger 2013). In other words, RS estimation and mapping of LAI can improve the identification of the above-mentioned stress factors in time (and space) resulting in LAI changes and support the appropriate decisions to be implemented. Additionally, because different crop types have different LAI at different growth stages, due to variation in leaf structure/sizes, knowledge of LAI dynamics can be an important input variable to mapping the spatial distribution of different crop types as well as to feed crop models.

Water is a key resource in agriculture, highly impacting in production; the FAO statistics indicate that irrigated agriculture occupies 20% of the total cultivated land but contributes 40% to the total food produced in the world (FAO 2016). The increasing population (9 billion expected by 2050) and changes in food diets demand an increase in agricultural production, which largely depends on irrigation. However, the inter-sectoral competition for water resources is also increasing and shifts of water allocation between economic sectors are expected, with a pressure over the agricultural sector. Agriculture is recognized as a public good of top priority, but it faces the dramatic challenge of producing more with less or with a more sustainable use of natural resources, taking into account water scarcity and the need to mitigate, as well as to adapt to foreseen climate change scenarios. Therefore, accurately monitoring water in agriculture is paramount for an adequate management of this resource. Within this context, the knowledge about the water available to the crops, including the monitoring of SM on agricultural fields, is thus a crucial component of crop performance monitoring and implementation of appropriate interventions (Justice 2016, ESA. 2019).

The SM is the quantity of water contained in the soil at a given point in time. It controls water and heat energy exchanges between the land surface and atmosphere through evaporation and transpiration. The paper 3 of this thesis provides timely and spatially distributed SM information within agricultural fields. The SM is afterward used to estimate and map the agricultural drought applying an index based on SM thresholds that defines the soil water availability for plants (field capacity and permanent wilting point). This information can assist in: (i) scheduling irrigation, by defining the irrigation timing when the drought index is below the threshold of water availability for plants, (ii) where to irrigate, by identifying the spatial location of areas within the crop field where the drought index is below the threshold of water availability for plants, and (iii) the amount of water to be applied, by indicating the quantity of water to be replenished in order to upgrade the SM to levels of water availability for plants. Moreover, the information about agricultural drought throughout the crops growth cycle, by impacting the crops yield, can be used by regional, national and international agencies for planning the import and export of food products, and mainly to feed food security early warning systems.

The topic of paper 4, relative to land cover mapping and evaluation of spatial and temporal dynamics of smallholder agriculture, enables a better estimation of cropland area, which is an important information for both farmers and agricultural managers (Bontemps, Arino et al. 2012, Gong, Wang et al. 2013, McKenzie, Sparrow et al. 2016, FAO 2017). Farmers and agricultural services entities in Mozambique often use traditional measurement approaches to estimate the area of farms, which normally leads to high under or over-estimation, with consequent impacts on yield forecasting. The land cover maps allow an easier and less costly access to the area of croplands and other land cover types. Moreover, allows a spatially distributed information of land cover types facilitating the analysis of their spatial dynamics throughout a time period (spatial-temporal dynamics component). This type of information can assist decisions related with: (i) food supply planning (ii) land management; (iii) resources management (e.g., water); (iv) biodiversity strategies.

Regardless all the above described opportunities, considering the predominant agricultural systems in Mozambique, there are a number of potential challenges and barriers that farmers face including, but not limited to: (i) lack of resources to guarantee immediate reaction as soon as an anomaly is detected; (ii) lack of scientific understanding; (iii) difficulties associated with extending scientific understanding or technological capability to operational utility; and (iv) lack of human resources to interpret and operationalize the RS data. Therefore, in-farm operational applications are more likely to be currently implemented by large to medium size commercial farming companies, which use relatively sophisticated agricultural practices.

Nevertheless, the results can also be operationally applied by agricultural governmental institutions, entities related to natural hazards management and non-governmental organizations for early detection of crop loss and food security risks. The crop growth and agricultural drought monitoring and cropland mapping are crucial for planning the food security at several scales, ranging from individual farmers up to district or even the whole country (Justice 2016, ESA. 2019). The remotely sensed information about LAI, SM, and cropland mapping can be further used to feed crop models and crop yield models (e.g. Cropping systems simulator; CropSyst or CERES-Maze) leveraging the usefulness of the results of this thesis. Because agriculture is the prime source of food and income for the major part of Mozambican population, the possibilities of obtaining yield estimates or crop performance with reasonable accuracy and prior to harvest is very important, since timely interventions can be made in case of low crop yields. Government and/or aid agencies can adequately plan the import and export of food products based on such information. In fact, this information can feed food security early warning systems enabling prompt interventions in order to alleviate food shortage effects and avoid fatalities.

Furthermore, SM and drought indicators are potentially useful for entities responsible for water management and the crop land maps are important for land planning entities as well as for those related to nature conservation.

Overall, the potential adoption of this thesis results is very high considering that all the field work was conducted in close collaboration with local farmers associations as well as local entities responsible for agriculture planning and management. In addition, dissemination of this thesis results will be done from local to national institutional level.

In the figure 13, a flowchart representing the process from RS data acquisition to decision making at farmer and government level is illustrated.

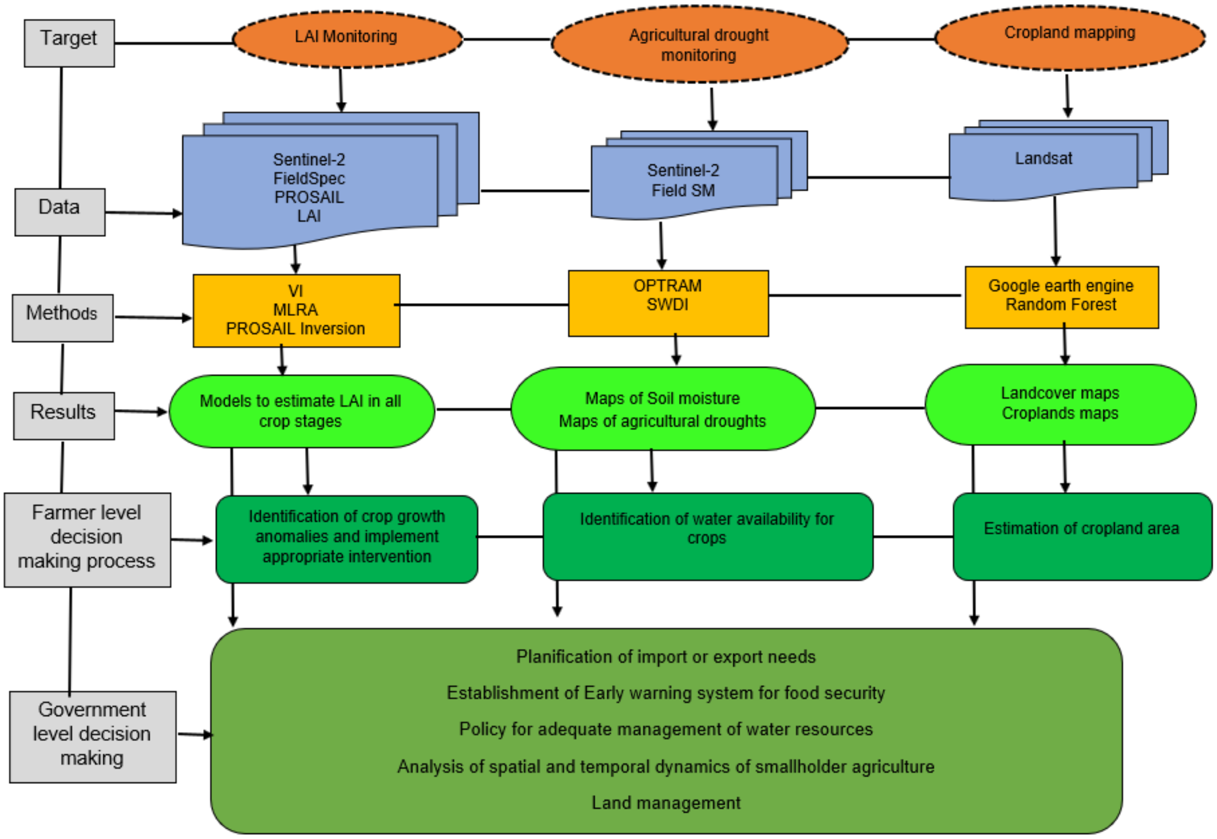


Fig. 13 - Scheme of this thesis results application in support to agricultural decision-making process. LAI – leaf area index, FielddSpec – field espectroradiometer, VI – vegetation index, MLRA – machine learning regression algorithm, OPTRAM – optical trapezoidal model; SWDI - soil water deficit index.

2| Integration with Agenda 2030 Objectives

In September 2015, the United Nations (UN) General Assembly endorsed "Transforming Our World: the 2030 Agenda for Sustainable Development", a global development agenda for all countries and stakeholders to use as a blueprint for progress on economic, social, and environmental sustainability. Seventeen Sustainable Development Goals (SDGs) and associated Targets and Indicators anchor the 2030 Agenda, which specifically calls for new data acquisition and exploitation of a wide range of data sources to support implementation. In fact, Article 76 states, "…We will promote transparent and accountable scaling-up of appropriate public-private cooperation to exploit the contribution to be made by a wide range of data, including Earth observation and geo-spatial information, while ensuring national ownership in supporting and tracking progress" (UN. 2015, Katherine, Barbara et al. 2017).

Earth observations (from satellite, airborne, and *in-situ* sensors) provide accurate and reliable information on the state of the atmosphere, oceans, coasts, rivers, soil, crops, forests, ecosystems, natural resources, ice, snow and built infrastructure, and their change over time, can be directly or indirectly used to assist governments, all economic sectors and almost all day-to-day activities of society (GEO 2017). The 2030 Agenda for Sustainable Development recognizes the significant role that Earth observations and geospatial information could play in making the whole framework feasible through the provision of essential evidence, including the tracking of indicators over time, and supporting the implementation of solutions to reach specific targets.

The effective use of the information in Earth observations can have a transformational impact on many societal challenges, such as those related with better monitor and protect fragile ecosystems, ensure resilient infrastructure, manage climate risks, enhance food security, build more resilient cities, reduce poverty, and improve governance, among others (GEO 2017, Katherine, Barbara et al. 2017). An analysis by the Group on Earth Observations (GEO), an intergovernmental organization working to improve the availability, access and use of Earth observations for the benefit of society and the Committee on Earth Observation Satellites (CEOS), a body for coordination of satellite Earth observation programmes of all the world's civil space agencies, has identified a total of 71 specific Targets and 29 Indicators that can be directly supported by Earth Observations (GEO 2017).

GEO supports the implementation of the 2030 Agenda with two dedicated Initiatives, EO for Sustainable Development (EO4SDG) in Service of the 2030 Agenda and the Global Agricultural Monitoring Initiative (GEOGLAM). The primary purpose of EO4SDG is to organize and help realize the potential of EO to advance the 2030 Agenda and enable societal benefits through achievement of the SDG while GEOGLAM aims to strengthen the international community's capacity to produce and disseminate relevant, timely and accurate forecasts of agricultural production at national, regional and global scales through the use of satellite and ground-based EO. GEOGLAM specifically supports SDG Goal 2, *End hunger, achieve food security and improved nutrition and promote sustainable agriculture, Target 2.c adopt measures to ensure the proper functioning of food commodity markets and their derivatives, and facilitate timely access to market information, including on food reserves, in order to help limit extreme food price volatility* (UN 2015, GEO 2017).

The present study supports the SDG contributing directly to 7 Goals and 12 targets which are presented in the table 3.

Table 3. Agenda 2030 Goals and Targets to which the present study contributes directly. Adaptation from (UN 2015).

Goals	Targets
1: End poverty in all its forms everywhere	**1.4** By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance.
	1.5 By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters.
2: End hunger achieve food security and improved nutrition and promote sustainable agriculture	**2.1** By 2030, end hunger and ensure access by all people, in particular the poor and people in vulnerable situations, including infants, to safe, nutritious and sufficient food all year round.
	2.3 By 2030, double the agricultural productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists and fishers, including through secure and equal access to land, other productive resources and inputs, knowledge, financial services, markets and opportunities for value addition and non-farm employment.
	2.4 By 2030, ensure sustainable food production systems and implement resilient agricultural practices that increase productivity and production, that help maintain ecosystems, that strengthen capacity for adaptation to climate change, extreme weather, drought, flooding and other disasters and that progressively improve land and soil quality.
6: Ensure access to water and sanitation for all	**6.4** By 2030, substantially increase water-use efficiency across all sectors and ensure sustainable withdrawals and supply of freshwater to address water scarcity and substantially reduce the number of people suffering from water scarcity.
8: Promote inclusive and sustainable economic growth, employment and decent work for all	**8.4** Improve progressively, through 2030, global resource efficiency in consumption and production and endeavour to decouple economic growth from environmental degradation, in accordance with the 10-year framework of programmes on sustainable consumption and production, with developed countries taking the lead**.**
12: Ensure sustainable consumption and production patterns	**12.2** By 2030, achieve the sustainable management and efficient use of natural resources**.**
13: Take urgent action to combat climate change and its impacts	**13.1** Strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries**.**
	13.3 Improve education, awareness-raising and human and institutional capacity on climate change mitigation, adaptation, impact reduction and early warning.
15: Sustainably manage forests, combat desertification, halt and reverse land degradation, halt biodiversity loss	**15.1** By 2020, ensure the conservation, restoration and sustainable use of terrestrial and inland freshwater ecosystems and their services, in particular forests, wetlands, mountains and drylands, in line with obligations under international agreements**.**
	15.2 By 2020, promote the implementation of sustainable management of all types of forests, halt deforestation, restore degraded forests and substantially increase afforestation and reforestation globally.

As pointed out by (Blanc 2015), Agenda 2030 goals and targets configures a network system, where goals are linked through targets such that implementation of one target can indirectly contribute to one or more other goals. The figure 14 depicts the existing networking between the goals directly supported by the present study as well as with those to which the study indirectly contributes.

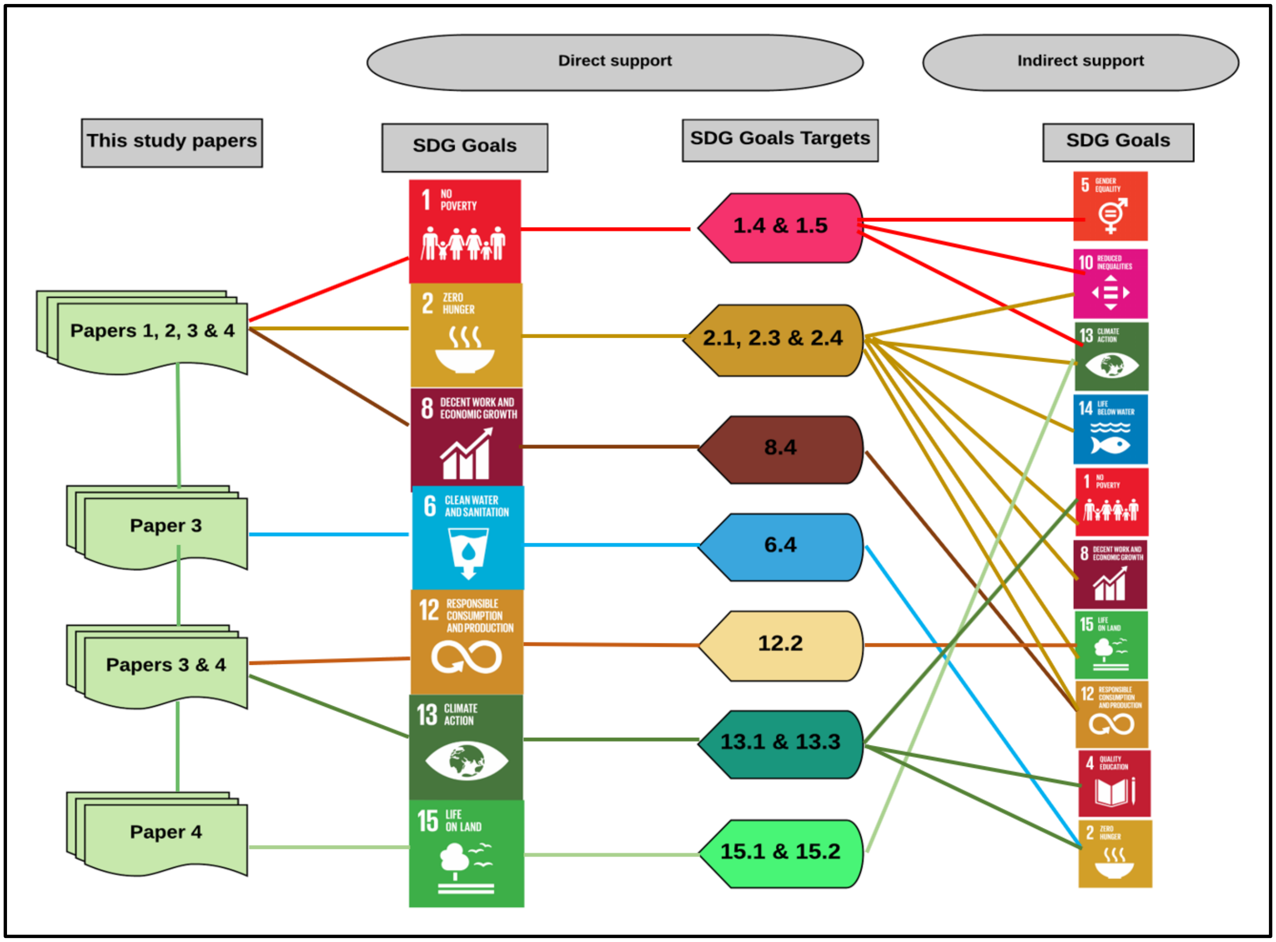


Fig. 14 - Network of goals and targets of Agenda 2030 to which the present study is linked.

Section V| General Discussion

Leaf area retrieval

The reliability of remotely sensed products relative to biophysical variables needs to be assessed, comparing them to in-situ collected data. For leaf area index (LAI), in-situ data can be gathered using direct or indirect methods. Although more accurate, the direct methods are time-consuming, not user-friendly and in many cases imply destroying the plants and thus preventing successive measurements (Jonckheere, Fleck et al. 2004). In paper 1, an indirect method based on measurements of allometric parameters of the plants was used and a mathematical model relating the product between the number of leaves, length and width of the largest leaf was calibrated and validated for the estimation of LAI of maize throughout all growing stages in the agri-ecological conditions and agronomic practices of the study area. Previous studies by (Valentinuz and Tollenaar 2006) and (Elings 2000, Mondo, De Carvalho et al. 2009) developed allometric models respectively for temperate and tropical maize varieties however, for a specific crop growth stage and, as such, do not include the number of leaves and can’t be used to monitor the LAI throughout the entire growing stages as does the model calibrated and validated in the current study.

Following the work implemented in paper 1, three different approaches of LAI retrieval through RS data (parametric, non-parametric and LUT inversion) were tested and compared using hyperspectral data (collected by field spectroradiometer and simulated using the radiative transfer model, PROSAIL) and multispectral data (Sentinel-2 and field spectroradiometer data aggregated to Sentinel-2 band settings) (paper 2).

The parametric approach, involving three vegetation indices (VIs) built with bands located at visible, red edge and near infrared spectral domains performed better in all types of datasets. These bands were equally reported in several studies for retrieval of LAI in various crops, e.g. (Zhao, Huang et al. 2007) (Stagakis, Markos et al. 2010), (Haboudane, R. et al. 2003), (Gitelson, Vina et al. 2003), and (Pan, Chen et al. 2018). Moreover, and as equally found by (Verrelst, Rivera et al. 2015), the three band VIs outperformed the two band VIs formulation for all the datasets.

When using the non-parametric approach, the machine learning algorithms Support Vector Regression (SVM), Variation Heteroscedastic Gaussian Processes Regression (VHGPR), Relevance vector Machine (RVM), Bagging trees (BaT) and Random Forest (Fitensemble) were the best performing and, likewise in the parametric approach, bands within the visible, red edge and infrared were frequently selected, what was also stated by (Verrelst, Rivera et al. 2015).

Generally, there was no considerable difference in LAI estimation accuracy between the hyperspectral and multispectral datasets, for both parametric and non-parametric approaches. Actually, there is still a discussion about the advantage of hyperspectral data as compared to multispectral data for estimating LAI (Liu, Zhou et al. 2016). The high spectral resolution of hyperspectral data potentially unveil the spectral details masked within multispectral data for LAI estimation (Schlerf, Atzberger et al. 2005). However, because the hyperspectral data include LAI insensitive bands, there is a trend to increase the computational time and to distort the accuracy of LAI retrieval due to the excessive number of variables, as such, hyperspectral data need to be coupled with techniques for dimensionality reduction (Liu, Zhou et al. 2016).

The accuracy of LUT inversion for LAI retrieval in paper 2 is within ranges reported by other studies for maize using the same and other retrieval techniques, for example, RMSE [0.40 – 0.43] in (Richter, Atzberger et al. 2009) and [0.46 – 1.21] in (Haboudane, R. et al. 2003). Other studies, such as (Richter, Atzberger et al. 2009), testified a tendency of underestimation in LUT inversion involving data from maize fields, which also occurred in the present case study. This may be explained by the row planting pattern of maize, which diverge from the turbid assumption of the PROSAIL model (Andrieu, Baret et al. 1997). In this study, the underestimation may have been aggravated by the low input cropping system of the field area and high heterogeneity within and between the sampled fields.

For all the modelling approaches, the combination of Sentinel-2 spectral data (SP_S2) and Sentine-2 LAI (LAI_S2) product was the most accurate, an expected finding considering that SP_S2 is an input for the derivation of LAI_S2 products. The parametric approach involving combination of SP_S2 and the in-situ LAI (Field_LAI; derived from paper 1 products) accurately estimated LAI, when compared to the physically based approach of LUT inversion. The lower performance of LUT inversion may be explained by the small size of the sampled area as well as the fact that the study considers only a single crop type. Furthermore, the PROSAIL parameters setting may have been hindered by the large heterogeneity within and between crop fields in the study area. In fact, as stated by (Liu, Zhou et al. 2016), statistical predictive LAI models are the most appropriate for small-scale studies.

This study results highlight the importance of adapting the RS data and methods to the characteristics and conditions of the site and target of interest rather than to merely adopt it. In fact, the area under study presents very peculiar characteristics (low-input and highly heterogeneous farming systems), which constitutes an additional challenge in the application of remote sensing techniques. Thus, testing the three different LAI retrieval approaches using different types of spectral datasets enabled the definition of the best strategy to be applied within the study area.

Soil moisture and agricultural drought

For SM retrieval through the OPTRAM, a physically based model developed by (Sadeghi, Babaeian et al. 2017), was applied using Sentinel-2 images covering a wide range of SM variability (irrigated and rain fed crop fields) and of vegetation cover types (agriculture, forests, rangelands, bare soils, etc.). The in-situ data for validation were collected using both tensiometers sensors and gravimetric methods. The pixels distribution in the space STR – NDVI configures well a trapezoidal shape confirming the good correlation between SM and STR in sparsely and densely vegetated conditions, which was also attested by (Sadeghi, Babaeian et al. 2017, Babaeian, Sadeghi et al. 2018).

For all assessed soil depths (15, 25 and 30 cm), the OPTRAM accurately estimated the SM, achieving error measures reported as reasonable for SM mapping (Entekhabi, Yueh et al. 2014). Nevertheless, and as also verified by (Roberts, Roth et al. 2011), the correlation decreased with depth, which can be explained by the fact that top soil layer SM is usually the regular water source for vegetation growth (Fang, Zhao et al. 2016) and, thus, more likely to be highly correlated with the vegetation spectral characteristics. Moreover, this finding fulfil the assumption that OPTRAM SM is correlated to root zone SM through the vegetation response to soil water in the root zone layer (Sadeghi, Babaeian et al. 2017, Babaeian, Sadeghi et al. 2018).

Another interesting finding is that higher accuracy of OPTRAM was mostly obtained in lower and intermediate values of SM, which includes SM levels at which the operational decisions for irrigation are normally taken.

The OPTRAM derived SM was subsequently used as input to determine an agricultural drought indicator, the soil water deficit index (SWDI). The results demonstrated that the observed crop fields in different years experienced several periods of some degree of agricultural drought (SDWI < 0) during the crop growing cycles and very few periods of no agricultural drought (SDWI > 0). Yet, the drought level was above the minimum threshold of available water for plants (SWDI < -10) in most of the cases, meaning that the plants could extract some water but spending much more energy. Earlier studies described similar findings but involving other moisture deficit indicators like the atmospheric water deficit (AWD), the crop moisture index (CMI) and rainfall (Martínez-Fernández, González-Zamora et al. 2015, Mishra, Vu et al. 2017, Babaeian, Sadeghi et al. 2018).

Broadly, the SWDI derived from OPTRAM SM was effective in monitoring agricultural droughts in the study area, which is an opportunity for high spatial and temporal resolution drought monitoring and mapping. In this study, the within field agricultural drought variability was successfully mapped. Despite this, and as equally pointed by (Mishra, Vu et al. 2017), important challenges such as the quantification of soil textures, soil water characteristics and, in particular for our study area, an extensive collection of in-situ SM data covering diversified crop and land cover types still need to be addressed.

Because spatial information has been pointed out as the major limitation of crop growth models implementation (Kasampalis, Alexandridis et al. 2018) and, as both LAI and SM are important parameters in such models, the results of this study can potentially fill the lack of information being assimilated in crop growth models and thus improving the crop forecasting. For example, (Nearing, Crow et al. 2012) successfully assimilated remotely sensed LAI and SM in the Decision Support System for Agrotechnology Transfer (DSSAT) CropSim-Ceres model.

Land cover mapping

A combination involving two types of input features, namely spectral bands of Landsat and textural features, were identified as the most appropriate for land cover classification and evaluation of spatial and temporal dynamics of the smallholder agriculture in the study area in Vilankulo district, Mozambique. A classification procedure was performed within the Google Earth Engine cloud computing platform applying the Random Forest algorithm aimed at mapping land cover types in three different years between 2012 and 2018. The land cover was classified and mapped with high overall accuracy. Previous studies have shown that including textural features as classification input greatly improves the classification accuracy (Räsänen and Virtanen 2019). Likewise, Random Forest has been recognized as a powerful algorithm for classification, especially in heterogeneous landscapes, as it is the study area (Belgiu and Drăguţ 2016, Gómez, White et al. 2016). The results highlight strong transitions between the natural vegetation land cover classes (close canopy natural vegetation, open canopy natural vegetation and grassland). In our point of view, this is mostly related to similarities in spectral behavior given that the distinction of these classes was defined through the percentage of canopy cover. Previous studies equally reported analogous results (Coburn and Roberts 2004, Wu, Zucca et al. 2016). As was expected, the class agriculture is the prime source of land cover conversion in the study area, which is characteristic of regions of itinerary agriculture predominance but also of agricultural intensification (Brink and Eva 2009, Morris, Byerlee et al. 2009, Pimm, Jenkins et al. 2014, Jacobson, Dhanota et al. 2015). The statistics of cultivated area resulting from the classified maps are consistently higher than those estimated by the local entities. However, the mean difference is small, compared to those stated in other works, e.g. (Dheeravath, Thenkabail et al. 2010, Jain, Mondal et al. 2013).

Section VI| Conclusions and perspectives

Within the context of increasing availability of remote sensing (RS) data, it’s important to assess and define the extent of its applicability in addressing the most challenging societal problems in developing countries. This thesis investigated the usefulness of different types of RS data and methods to retrieve the leaf area index (LAI), the soil moisture (SM) and the agricultural drought in low input agricultural systems, and to map land cover types in a heterogeneous highly dynamic landscape in southern Mozambique. The contribution of RS for both agronomic process inference and agricultural monitoring in resource-poor smallholder farming systems was assessed in different case studies, testing the most suited types of data and methodological approaches.

The thesis results highlight that, although with slightly different accuracy, both multispectral and hyperspectral data and the three tested methods (parametric, non-parametric and radiative transfer models) can be successfully used to estimate and monitor LAI throughout the entire growing season of maize, the most important stable crop type in the study area. Whichever the modelling approach or dataset, the developed models consistently included bands from the same electromagnetic regions. An allometric based model, suitable for estimating the LAI in all crop stages was also calibrated and validated and used to derive the ground truth data.

The SM in the root zone depths was accurately retrieved through the application of Optical Trapezoid Model (OPTRAM) using Sentinel-2 data and subsequently used for spatial and temporal monitoring of agricultural drought. The usefulness of OPTRAM is particularly pertinent because it only requires optical RS data, which are freely available from satellite missions like Landsat and Sentinel-2 and have appropriate spatial and temporal resolution to support sustainable strategies of water management in crops irrigation.

A combination of spectral bands and textural features was appropriate to derive high accurate classification of land cover and to evaluate the spatial and temporal dynamics of smallholder agriculture in the district of Vilankulo. From the classified maps, statistics of cultivated area were calculated which were, to some extent, in agreement to the ground truth information.

Overall, this thesis results present very high potential for agricultural practice-oriented applications, for example: (i) supporting farmers to timely identify different types of stress-induced growth anomalies and/or crop requirements and thus making effective decision about the intervention to be undertaken; (ii) helping policy makers in planning food aids and thus preventing fatalities related to food insecurity and (iii) contributing directly to SDG 7 Goals and 12 targets.

Because the retrieval of LAI and SM was conducted in fields of only two crop types, future work should be carried on developing models involving more crop types grown within the study area.

Moreover, the use of Sentinel-2 data, as well as multi-sensor data (combining different sensors and following harmonized pre-processing procedures), should also be tested for land cover classification, aiming to enhance the regular spatial and temporal analysis of within and between land cover types.

Lisf of References

Achicala, R., S. Lampeão, L. S. Lopes, V. Salustiano, A. Mabota, H. Matlombe, A. Miguel, A. Mazivila, F. Fumo, F. Zezela, F. Camisa and D. Diogo (2016). Anuário de Estatísticas Agrárias 2015, Ministerio da Agricultura e Seguranca Alimentar**:** 66.

Ackermann, N., J. Allgood, S. Aravazhi, S. Ayers, P. Ballantyne, A. Bennett, R. Best, R. Birner, S. Boettiger, A. R. Braun, A. Brett, J. Chambers, A. Clayton, M.-H. Collion, K. Davis, L. D. Rios, K. Dorai, D. Duveskog, J. Ekboir, H. Elliott, W. Ellis and J. Ernestberger (2012). Agricultural Innovation Systems: An investment sourcebook. Agriculture and rural development. Whashington, DC, The World Bank**:** 684.

Ahmad, S., A. Kalra and H. Stephen (2010). "Estimating soil moisture using remote sensing data: A machine learning approach." Advances in Water Resources **33**(1): 69-80.

Ahmed, A., Y. Zhang and S. Nichols (2011). "Review and evaluation of remote sensing methods for soil-moisture estimation." SPIE Reviews **2**: 18.

Aiuba, R. and J. Mosca (2018). Orçamento geral do estado para o sector agrário e desenvolvimento rural. Destaque Rural. Maputo, Observatorio do Meio Rural. **46:** 3.

Al-Bakri, J., A. Suleiman and A. Berg (2014). "A comparison of two models to predict soil moisture from remote sensing data of RADARSAT II." Arabian Journal of Geosciences **7**(11): 4851-4860.

Alchanatis, V. and Y. Cohen (2011). Spectral and Spatial Methods of Hyperspectral Image Analysis for Estimation of Biophysical and Biochemical Properties of Agricultural Crops. Hyperspectral Remote Sensing of Vegetation, CRC Press**:** 289-308.

Amani, M., S. Parsian, S. M. MirMazloumi and O. Aieneh (2016). "Two new soil moisture indices based on the NIR-red triangle space of Landsat-8 data." International Journal of Applied Earth Observation and Geoinformation **50**: 176-186.

Andrieu, B., F. Baret, S. Jacquemoud, T. Malthus and M. Steven (1997). "Evaluation of an improved version of SAIL model for simulating bidirectional reflectance of sugar beet canopies." Remote Sensing of Environment **60**(3): 247-257.

Anne, N. J. P., A. H. Abd-Elrahman, D. B. Lewis and N. A. Hewitt (2014). "Modeling soil parameters using hyperspectral image reflectance in subtropical coastal wetlands." International Journal of Applied Earth Observation and Geoinformation **33**: 47-56.

Asner, G. and K. Heidebrecht (2002). "Spectral Unmixing of Vegetation, Soil and Dry Carbon in Arid Regions: Comparing Multispectral and Hyperspectral Observations." International Journal of Remote Sensing **23**.

Atzberger, C. (2004). "Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models." Remote Sensing of Environment **93**(1–2): 53-67.

Atzberger, C. (2013). "Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs." Remote Sensing **5**(2): 949-981.

Atzberger, C., R. Darvishzadeh, M. Immitzer, M. Schlerf, A. Skidmore and G. le Maire (2015). "Comparative analysis of different retrieval methods for mapping grassland leaf area index using airborne imaging spectroscopy." International Journal of Applied Earth Observation and Geoinformation **43**: 19-31.

Atzberger, C., M. Guérif, F. Baret and W. Werner (2010). "Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat." Computers and Electronics in Agriculture **73**(2): 165-173.

Atzberger, C., T. Jarmer, M. Schlerf, B. Kötz and W. Werner (2003). Retrieval of wheat bio-physical attributes from hyperspectral data and SAILH+PROSPECT radiative transfer model. EARSeL Workshop on Imaging Spectroscopy,, Herrsching

Atzberger, C. and F. Rembold (2013). "Mapping the Spatial Distribution of Winter Crops at Sub-Pixel Level Using AVHRR NDVI Time Series and Neural Nets." Remote Sensing **5**(3): 1335.

Azzari, G. and D. B. Lobell (2017). "Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring." Remote Sensing of Environment **202**: 64-74.

B. Lobell, D., G. Asner, B. Law and R. N. Treuhaft (2002). View angle eVects on canopy re ectance and spectral mixture analysis of coniferous forests using AVIRIS.

Babaeian, E., M. Sadeghi, T. E. Franz, S. Jones and M. Tuller (2018). "Mapping soil moisture with the OPtical TRApezoid Model (OPTRAM) based on long-term MODIS observations." Remote Sensing of Environment **221**: 16.

Bacour, C., F. Baret, D. Béal, M. Weiss and K. Pavageau (2006). "Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation." Remote Sensing of Environment **105**(4): 313-325.

Baghdadi, N., M. Aubert and M. Zribi (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.

Bakhshandeh, E., B. Kamkar and J. Tsialtas (2011). "Aplication of linear models for estimation of leaf area in soybean [*Glycine max* (L.) Merr]." Photosynthetica **49 (3)**: 405-416.

Bakker, W. H., W. Feringa, A. S. M. Gieske, B. G. H. Gorte, K. A. Grabmaier, C. A. Hecker, J. A. Horn, G. C. Huurneman, L. L. Janssen, N. Kerle, F. D. van der Meer, G. N. Parodi, C. Pohl, C. V. Reeves, F. J. van Ruitenbeek, E. M. Schetselaar, K. Tempfli, M. J. C. Weir, E. Westinga and T. Woldai (2009). Principles of Remote Sensing. An introductory textbook. T. K. K. N. G. C. H. J. L. L. F. Enschede, The Netherlands, The International Institute for Geo-Information Science and Earth Observation (ITC)**:** 591.

Ball, J. and L. Bruce (2005). Level set segmentation of remotely sensed hyperspectral images.

Barati, S., B. Rayegani, M. Saati, A. Sharifi and M. Nasri (2011). "Comparison the accuracies of different spectral indices for estimation of vegetation cover fraction in sparse vegetated areas." The Egyptian Journal of Remote Sensing and Space Science **14**(1): 49-56.

Baret, F. and S. Buis (2008). "Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems." Advances in Land Remote Sensing: System, Modelling, Inversion andApplication: 172-301.

Basso, B., D. Cammarano and P. De Vita (2004). "Remotely sensed vegetation indices: Theory and applications for crop management." Rivista Italiana di Agrometeorologia **1**: 17.

Baumgardner, M. F., L. F. Silva, L. L. Biehl and E. R. Stoner (1986). Reflectance Properties of Soils. Advances in Agronomy. N. C. Brady, Academic Press. **38:** 1-44.

Belgiu, M. and L. Drăguţ (2016). "Random forest in remote sensing: A review of applications and future directions." ISPRS Journal of Photogrammetry and Remote Sensing **114**: 24-31.

Beljaars, A., P. Viterbo, M. Miller and A. Betts (1996). "The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil moisture anomalies." Monthly Weather Review **124**: 21.

Benayas, J. M. R. and J. M. Bullock (2012). "Restoration of Biodiversity and Ecosystem Services on Agricultural Land." Ecosystems(15): 16.

Benfica, R., B. Cunguara and J. Thurlow (2019). "Linking agricultural investments to growth and poverty: An economywide approach applied to Mozambique." Agricultural Systems **172**: 91-100.

Berger, K., C. Atzberger, M. Danner, G. D'Urso, W. Mauser, F. Vuolo and T. Hank (2018). "Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study." Remote Sensing **10**(1): 26.

Berjo, A. J., V. E. Cachorro, P. J. Zarco-Tejada and A. de Frutos (2013). "Retrieval of biophysical vegetation parameters using simultaneous inversion of high resolution remote sensing imagery constrained by a vegetation index." Precision Agric **14**: 541–557.

Bernardes, T., M. A. Moreira, M. Adami and B. F. T. Rudorff (2012). Monitoring biennial bearing effect on coffee yield using modis remote sensing imagery. 2012 IEEE International Geoscience and Remote Sensing Symposium.

Betts, A. K. and J. H. Ball (1998). "FIFE surface climate and site-average dataset: 1987–1989." Journal of Atmospheric Sciences **55**: 18.

Bhaskaran, S., S. Paramananda and M. Ramnarayan (2010). "Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data." Applied Geography **30**(4): 650-665.

Bindlish, R. and A. P. Barros (2000). " Multi-frequency soil moisture inversion from SAR measurements with the use of IEM." Remote Sensing of Environment **61**: 27.

Blackburn, G. A. and J. G. Ferwerda (2008). "Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis." Remote Sensing of Environment **112**(4): 1614-1632.

Blanc, D. L. (2015). Towards integration at last? The sustainable development goals as a network of targets. New York, United Nations.

Blaschke, T. (2010). "Object based image analysis for remote sensing." ISPRS Journal of Photogrammetry and Remote Sensing **65**(1): 2-16.

Blaschke, T., C. Burnett and A. Pekkarinen (2004). Image Segmentation Methods for Object-based Analysis and Classification. Remote Sensing Image Analysis: Including The Spatial Domain. S. M. D. Jong and F. D. V. d. Meer. Dordrecht, Springer Netherlands**:** 211-236.

Bonan, G. B. (1995). "Land-Atmosphere interactions for climate system Models: coupling biophysical, biogeochemical, and ecosystem dynamical processes." Remote Sensing of Environment **51**(1): 57-73.

Bontemps, S., O. Arino, P. Bicheron, C. C. Brockman, M. Leroy, C. Vancutsem and P. Defourny (2012). Operational Service Demonstration for Global Land-Cover Mapping: The GlobCover and GlobCorine Experiences for 2005 and 2009. Remote Sensing of Land Use and Land Cover: Principles and Applications. C. P. Giri. Boca Raton, FL, USA., CRC Press**:** 243–265.

Bowen, M. L. (2000). The State Against Peasantry. Rural Struggles in Colonial and Postcolonial Mozambique. Charlottesville, VA, University Press of Virginia.

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

Breiman, L., J. Friedman, C. Stone and R. Olshen (1984). Classification and Regression Trees. Boca, Raton, USA, Chapman&Hall/CRC.

Brink, A. B. and H. D. Eva (2009). "Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach." Applied Geography **29**: 12.

Caicedo, J. P. R., J. Verrelst, J. Muñoz-Marí, J. Moreno and G. Camps-Valls (2014). "Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **7**(4): 1249-1259.

Callens, M., N. E. C. Verhoest and M. W. J. Davidson (2006). "Parameterization of tillage-induced single-scale soil roughness from 4-m profiles." IEEE Transactions on Geoscience and Remote Sensing **44**(4): 878-888.

Cammalleri, C., F. Micale and J. Vogt (2016). "A novel soil moisture-based drought severity index (DSI) combining water deficit magnitude and frequency." Hydrological Processes **30**(2): 289-301.

Casa, R., F. Baret, S. Buis, R. Lopez-Lozano, S. Pascucci, A. Palombo and H. G. Jones (2010). "Estimation of maize canopy properties from remote sensing by inversion of 1-D and 4-D models." Precision Agriculture **11**(4): 319-334.

Cattanio, J. H. (2017). "Leaf area index and root biomass variation at different secondary forest ages in the eastern Amazon." Forest Ecology and Management **400**: 1-11.

CCRS (2017). Fundamentals of Remote Sensing. A Canada Centre for Remote Sensing Remote Sensing Tutorial. N. R. Canada. Canada, Canada Centre for Remote Sensing. **1:** 258.

Champagne, C., T. Rowlandson, A. Berg, T. Burns, J. L'Heureux, E. Tetlock, J. R. Adams, H. McNairn, B. Toth and D. Itenfisu (2016). "Satellite surface soil moisture from SMOS and Aquarius: Assessment for applications in agricultural landscapes." International Journal of Applied Earth Observation and Geoinformation **45**: 143-154.

Chen, J. M. and T. A. Black (1992). "Defining leaf area index for non-flat leaves." Plant, Cell & Environment **15**(4): 421-429.

Chen, J. M. and S. G. Leblanc (1997). "A four-scale bidirectional reflectance model based on canopy architecture." IEEE Transactions on Geoscience and Remote Sensing **35**(5): 1316-1337.

Cheng, Y. B., P. J. Zarco-Tejada, D. Riano, C. A. Rueda and S. L. Ustin (2006). "Estimating vegetation water content with hyperspectral data for different canopy scenarios: relationships between AVIRIS and MODIS indexes." Remote Sensing of Environment **105**: 13.

Clevers, J. G. P. W. and A. A. Gitelson (2013). "Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3." International Journal of Applied Earth Observation and Geoinformation **23**: 344-351.

Clevers, J. G. P. W., L. Kooistra and M. E. Schaepman (2010). "Estimating canopy water content using hyperspectral remote sensing data." International Journal of Applied Earth Observation and Geoinformation **12**(2): 119-125.

Coburn, C. A. and A. C. B. Roberts (2004). "A multiscale texture analysis procedure for improved forest stand classification." International Journal of Remote Sensing **25**(20): 4287-4308.

Cohen, W. B. and S. N. Goward (2004). "Landsat's role in ecological applications of remote sensing." BioScience **54**: 11.

Combal, B., F. Baret, M. Weiss, A. Trubuil, D. Macé, A. Pragnère, R. Myneni, Y. Knyazikhin and L. Wang (2003). "Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem." Remote Sensing of Environment **84**(1): 1-15.

Cong, J. O., Li, N., Xu, Y.J., Gu, W., Le, Z.Y., Huang, S.Q., Xi, B., Lei, Y. (2010). " Relationship between indices of growth, physiology and reflectivity and yield of winter wheat under water stress." Chinese Journal of Eco-Agriculture **18**(1): 67 - 71.

Corbari, C., J. A. Sobrino, M. Mancini and V. Hidalgo (2010). "Land surface temperature representativeness in a heterogeneous area through a distributed energy-water balance model and remote sensing data." Hydrol. Earth Syst. Sci. **14**(10): 2141-2151.

Cornelissen, J., S. Lavorel, E. B. Garnier, S. Diaz, N. Buchmann, D. Gurvich, P. Reich, H. ter Steege, H. D. G. Morgan, M. Van der Heijden, J. G. H. Pausas and H. Poorter (2003). Handbook of protocols for standardised and easy measurement of plant functional traits worldwide.

Croft, H., K. Anderson and N. J. Kuhn (2012). "Reflectance anisotropy for measuring soil surface roughness of multiple soil types." CATENA **93**: 87-96.

Cunha, M., A. R. S. Marçal and L. Silva (2010). "Very early prediction of wine yield based on satellite data from VEGETATION." International Journal of Remote Sensing **31**(12): 3125-3142.

Darvishzadeh, R. (2008). Hyperspectral remote sensing of vegetation parameters using statistical and physical models. PhD, Wageningen University.

Darvishzadeh, R., A. A. Matkan and A. Dashti Ahangar (2012). "Inversion of a radiative transfer model for estimation of rice canopy chlorophyll content using a lookup - table approach." IEEE Journal of selected topics in applied earth observations and remote sensing **5**(4): 9.

Darvishzadeh, R., A. K. Skidmore, M. Schlerf and C. Atzberger (2008). "Inversion of a radiative transfer model for estimating vegetation LAIA and chlorophyll in a heterogeneous grassland." Remote Sensing of Environment **112**(5): 2592 - 2604.

Dash, J. and P. J. Curran (2004). "he MERIS terrestrial chlorophyll index." International Journal of Remote Sensing **25**: 11.

Daughtry, C. S. T., P. C. Doraiswamy, E. R. Hunt, A. J. Stern, J. E. McMurtrey and J. H. Prueger (2006). "Remote sensing of crop residue cover and soil tillage intensity." Soil and Tillage Research **91**(1): 101-108.

de Bello, F., J. Leps and M. T. Sebastia (2006). "Variations in species and functional plant diversity along climatic and grazing gradients. 29, 801–810." Ecography **29**: 11.

de Jeu, R. A. M., T. R. H. Holmes, R. M. Parinussa and M. Owe (2014). "A spatially coherent global soil moisture product with improved temporal resolution." Journal of Hydrology **516**: 284-296.

de Wit, A. J. W. and C. A. van Diepen (2007). "Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts." Agricultural and Forest Meteorology **146**(1): 38-56.

Dean, A. M. and G. M. Smith (2003). "An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities." International Journal of Remote Sensing **24**(14): 2905-2920.

Defourny, P., S. Bontemps, N. Bellemans, C. Cara, G. Dedieu, E. Guzzonato, O. Hagolle, J. Inglada, L. Nicola, T. Rabaute, M. Savinaud, C. Udroiu, S. Valero, A. Bégué, J.-F. Dejoux, A. El Harti, J. Ezzahar, N. Kussul, K. Labbassi, V. Lebourgeois, Z. Miao, T. Newby, A. Nyamugama, N. Salh, A. Shelestov, V. Simonneaux, P. S. Traore, S. S. Traore and B. Koetz (2019). "Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world." Remote Sensing of Environment **221**: 551-568.

Dheeravath, V., P. Thenkabail, G. Chandrakantha, P. Noojipady, G. P. O. Reddy, C. Biradar, M. Gumma and N. M. Velpuri (2010). "Irrigated areas of India derived using MODIS 500 m time series for the years 2001–2003." ISPRS Journal of Photogrammetry and Remote Sensing **65**: 42-59.

Diaz, S., J. Hodgson, K. Thompson, M. Cabido, J. Cornelissen, A. Jalili, G. Montserrat-Martí, J. Grime, F. Zarrinkamar, Y. Asri, S. R. Band, S. Basconcelo, P. Castro-Díez, G. Funes, B. Hamzehee, M. Khoshnevi, N. Pérez-Harguindeguy, C. Pérez-Rontomé, A. Shirvany and M. Zak (2004). The plant traits that drive ecosystems: Evidence from three continents.

Ding, Y., X. Zheng, K. Zhao, X. Xin and H. Liu (2016). "Quantifying the Impact of NDVIsoil Determination Methods and NDVIsoil Variability on the Estimation of Fractional Vegetation Cover in Northeast China." Remote Sensing **8**(1): 29.

Dobrin, B., T. Viero and M. Gabbouj (1994). Fast watershed algorithms: analysis and extensions, SPIE.

Doktor, D., A. Lausch, D. Spengler and M. Thurner (2014). "Extraction of Plant Physiological Status from Hyperspectral Signatures Using Machine Learning Methods." Remote Sensing **6**(12): 12247.

Doraiswamy, P. C., T. R. Sinclair, S. Hollinger, B. Akhmedov, A. Stern and J. Prueger (2005). "Application of MODIS derived parameters for regional crop yield assessment." Remote Sensing of Environment **97**(2): 192-202.

Dorigo, W., R. Richter, F. Baret, R. Bamler and W. Wagner (2009). " Enhanced Automated Canopy Characterization from Hyperspectral Data by a Novel Two Step Radiative Transfer Model Inversion Approach." Remote Sensing **1**: 31.

Dorigo, W. A., R. Zurita-Milla, A. J. W. de Wit, J. Brazile, R. Singh and M. E. Schaepman (2007). "A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling." International Journal of Applied Earth Observation and Geoinformation **9**(2): 165-193.

Dos Muchangos, A. (1999). Moçambique Paisagens e Regiões Naturais Maputo, Tipografia Globo, Lda.

Dracup, J. A., K. S. Lee and E. G. Paulson Jr. (1980). "On the definition of droughts." Water Resources Research **16**(2): 297-302.

Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese and P. Bargellini (2012). "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." Remote Sensing of Environment **120**: 25-36.

Duro, D. C., S. E. Franklin and M. G. Dubé (2012). "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery." Remote Sensing of Environment **118**: 259-272.

Elings, A. (2000). "Estimation of leaf area in tropical maize." Agronomy journal **92**: 436-444.

Engman, T. and G. A. Schultz (2000). Remote Sensing in Hydrology and Water Management**:** 445-457.

Entekhabi, D., S. Yueh, P. E. O'Neill, K. H. Kellogg, A. Allen, R. Bindlish, M. Brown, S. Chan, A. Colliander, W. T. Crow and N. Das (2014). SMAP Handbook–Soil Moisture Active Passive:Mapping Soil Moisture and Freeze/Thaw from Space. Pasadena, CA, Space. JPL Publication.

Eriksson, H., L. Eklundh, A. Kuusk and T. Nilson (2006). Impact of understory vegetation on forest canopy reflectance and remotely sensed LAI estimates.

ESA. (2019). "Earth Observation for Sustainable development: Agriculture and Rural Development knowledge portal." Retrieved 04/07/2019, 2019, from http://eo4idi.eu/eo4sd-knowledge-portal.

Fang, H., Y. Ye, W. Liu, S. Wei and L. Ma (2018). "Continuous estimation of canopy leaf area index (LAI) and clumping index over broadleaf crop fields: An investigation of the PASTIS-57 instrument and smartphone applications." Agricultural and Forest Meteorology **253-254**: 48-61.

Fang, X., W. Zhao, L. Wang, Q. Feng, J. Ding, Y. Liu and X. Zhang (2016). "Variations of deep soil moisture under different vegetation types and influencing factors in a watershed of the Loess Plateau, China." Hydrol. Earth Syst. Sci. **20**: 17.

FAO (2011). Global Strategy to Improve Agricultural and Rural Statistics. Rome, Italy, FAO.

FAO. (2016). "AQUASTAT website. Food and Agriculture Organization of the United Nations (FAO)." Retrieved 2019/07/08, 2019.

FAO (2016). FAO, Country Programming Framework for Mozambique 2016-2020. Maputo, FAO**:** 13.

FAO (2017). Handbook on remote sensing for agricultural statistics. Rome, Italy, FAO**:** 288.

Fascetti, F., N. Pierdicca, L. Pulvirenti, R. Crapolicchio and J. Muñoz-Sabater (2016). "A comparison of ASCAT and SMOS soil moisture retrievals over Europe and Northern Africa from 2010 to 2013." International Journal of Applied Earth Observation and Geoinformation **45**: 135-142.

Féret, J. B., A. A. Gitelson, S. D. Noble and S. Jacquemoud (2017). "PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle." Remote Sensing of Environment **193**: 204-215.

Fonseca, A. D. and J. C. Fernandes (2004). Detecão Remota. Lisbora, Portugal, LIDEL.

Foody, G. M. (2002). "Status of land cover classification accuracy assessment." Remote Sensing of Environment **80**(1): 185-201.

Forkuor, G., C. Conrad, M. Thiel, T. Ullmann and E. Zoungrana (2014). "Integration of Optical and Synthetic Aperture Radar Imagery for Improving Crop Mapping in Northwestern Benin, West Africa." Remote Sensing **6**(7): 6472-6499.

Fourty, T., F. Baret, S. Jacquemoud, G. Schmuck and J. Verdebout (1996). "Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems." Remote Sensing of Environment **56**(2): 104-117.

Franceschini, M. H. D., J. A. M. Demattê, F. da Silva Terra, L. E. Vicente, H. Bartholomeus and C. R. de Souza Filho (2015). "Prediction of soil properties using imaging spectroscopy: Considering fractional vegetation cover to improve accuracy." International Journal of Applied Earth Observation and Geoinformation **38**: 358-370.

Franceschini, M. H. D., J. A. M. Demattê, L. Kooistra, H. Bartholomeus, R. Rizzo, C. T. Fongaro and J. P. Molin (2018). "Effects of external factors on soil reflectance measured on-the-go and assessment of potential spectral correction through orthogonalisation and standardisation procedures." Soil and Tillage Research **177**: 19-36.

Franklin, S. E. and M. A. Wulder (2002). "Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas." Progress in Physical Geography: Earth and Environment **26**(2): 173-205.

Fritz, S., L. See, I. McCallum, L. You, A. Bun, E. Moltchanova, M. Duerauer, F. Albrecht, C. Schill, C. Perger, P. Havlik, A. Mosnier, P. Thornton, U. Wood-Sichra, M. Herrero, I. Becker-Reshef, C. Justice, M. Hansen, P. Gong, S. Abdel Aziz, A. Cipriani, R. Cumani, G. Cecchi, G. Conchedda, S. Ferreira, A. Gomez, M. Haffani, F. Kayitakire, J. Malanding, R. Mueller, T. Newby, A. Nonguierma, A. Olusegun, S. Ortner, D. R. Rajak, J. Rocha, D. Schepaschenko, M. Schepaschenko, A. Terekhov, A. Tiangwa, C. Vancutsem, E. Vintrou, W. Wenbin, M.

van der Velde, A. Dunwoody, F. Kraxner and M. Obersteiner (2015). "Mapping global cropland and field size." Global Change Biology **21**(5): 1980-1992.

Fukshansky, L., N. Fukshansky-Kazarinova and A. M. v. Remisowsky (1991). "Estimation of optical parameters in a living tissue by solving the inverse problem of the multiflux radiative transfer." Applied Optics **30**(22): 3145-3153.

Gago, J., C. Douthe, R. E. Coopman, P. P. Gallego, M. Ribas-Carbo, J. Flexas, J. Escalona and H. Medrano (2015). "UAVs challenge to assess water stress for sustainable agriculture." Agricultural Water Management **153**: 9-19.

Gallego, F. J. (2004). "Remote sensing and land cover area estimation." International Journal of Remote Sensing **25**(15): 3019-3047.

Gao, Z., X. Xu, J. Wang, H. Yang, W. Huang and H. Feng (2013). "A method of estimating soil moisture based on the linear decomposition of mixture pixels." Mathematical and Computer Modelling **58**(3): 606-613.

Gastellu-Etchegorry, J. P., V. Demarez, V. Pinel and F. Zagolski (1996). "Modeling radiative transfer in heterogeneous 3-D vegetation canopies." Remote Sensing of Environment **58**(2): 131-156.

GEO (2017). Earth Observations in Support of the 2030 Agenda for Sustainable Development. Geneva, Switzerland, Group on Earth Observation. **1:** 19.

Ghulam, A., Q. Qin, T. Teyip and Z.-L. Li (2007). "Modified perpendicular drought index (MPDI): a real-time drought monitoring method." ISPRS Journal of Photogrammetry and Remote Sensing **62**(2): 150-164.

Gitelson, A. A., A. Vina, T. J. Arkebauer, D. C. Rundquist, G. Keydan and B. Leavitt (2003). "Remote estimation of leaf area index and green leaf biomass in maize canopies." Geophysical research letters **30**(5): 4.

Glenn, E. P., Huete, A. R., Nagler, P. L., Nelson, S. G. (2008). "Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape." Sensors **8**: 25.

Goel, N. S. (1988). "Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data." Remote Sensing Reviews **4**(1): 1-212.

Goel, N. S. and T. Grier (1988). "Estimation of canopy parameters for inhomogeneous vegetation canopies from reflectance data: III. Trim: A model for radiative transfer in heterogeneous three-dimensional canopies." Remote Sensing of Environment **25**(3): 255-293.

GOM (Government of Mozambique) (2012). Programa Nacional de Investmento do Sector Agrário: PNISA 2013–2017. Maputo, Mozambique, Ministry of Agriculture.

Gómez, C., J. C. White and M. A. Wulder (2016). "Optical remotely sensed time series data for land cover classification: A review." ISPRS Journal of Photogrammetry and Remote Sensing **116**: 55-72.

Gonçalves, J., I. Pôças, B. Marcos, C. A. Mücher and J. P. Honrado (2019). "SegOptim—A new R package for optimizing object-based image analyses of high-spatial resolution remotely-sensed data." International Journal of Applied Earth Observation and Geoinformation **76**: 218-230.

Gong, P., J. Wang, L. Yu, Y. Zhao, Y. Zhao, L. Liang, Z. Niu, X. Huang, H. Fu, S. Liu, C. Li, X. Li, W. Fu, C. Liu, Y. Xu, X. Wang, Q. Cheng, L. Hu, W. Yao, H. Zhang, P. Zhu, Z. Zhao, H. Zhang, Y. Zheng, L. Ji, Y. Zhang, H. Chen, A. Yan, J. Guo, L. Yu, L. Wang, X. Liu, T. Shi, M. Zhu, Y. Chen, G. Yang, P. Tang, B. Xu, C. Giri, N. Clinton, Z. Zhu, J. Chen and J. Chen (2013). "Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data." International Journal of Remote Sensing **34**(7): 2607-2654.

Govaerts, Y., S. Jacquemoud, M. Verstraete and S. Ustin (1996). Three-dimensional radiation transfer modeling in a dicotyledon leaf.

Govaerts, Y., M. Verstraete, B. Pinty and N. Gobron (2014). Designing optimal spectral indices: A feasibility and proof of concept study. **20**.

Govender, M., K. Chetty and H. Bulcock (2007). "A review of hyperspectral remote sensing and its application in vegetation and water resource studies." Water SA **33**(2): 7.

Gu, Y., J. F. Brown, J. P. Verdin and B. Wardlow (2007). "A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States." Geophysical Research Letters **34**(6).

Guérif, M. and C. L. Duke (2000). "Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation." Agriculture, Ecosystems & Environment **81**(1): 57-69.

Guerriero, L., P. Ferrazzoli and R. Rahmoune (2012). A synergyc view of L-band active and passive remote sensing of vegetated soil. 2012 12th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad).

Guyon, I. and A. Elisseeff (2003). An introduction to variable and feature selection, JMLR.org.

Haboudane, D., J. R. Miller, N. Tremblay, P. J. Zarco-Tejada and L. Dextraze (2002). "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture." Remote Sensing of Environment **81**(2): 416-426.

Haboudane, D., M. J. R., E. Pattey, P. J. Zarco-Tejadad and I. B. Strachan (2003). "Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture." Remote Sensing of Environment **90**: 15.

Hannerz, F. and A. Lotsch (2008). "Assessment of remotely sensed and statistical inventories of African agricultural fields." International Journal of Remote Sensing **29**(13): 3787-3804.

Hapke, B. (2008). "Bidirectional reflectance spectroscopy: 6. Effects of porosity." Icarus **195**(2): 918-926.

Hastie, T., R. Tibshirani and J. H. Friedman (2009). The elements of statistical learning: data mining, inference, and prediction. New York, Springer-Verlag.

Hazaymeh, K. and Q. K. Hassan.

Heim Jr, R. R. (2002). "A review of twentieth-century drought indices used in the United States." Bulletin of the American Meteorological Society **83**(8): 1149-1165.

Holzman, M. E., R. Rivas and M. C. Piccolo (2014). "Estimating soil moisture and the relationship with crop yield using surface temperature and vegetation index." International Journal of Applied Earth Observation and Geoinformation **28**: 181-192.

Homolová, L., Z. Malenovský, J. G. P. W. Clevers, G. García-Santos and M. E. Schaepman (2013). "Review of optical-based remote sensing for plant trait mapping." Ecological Complexity **15**: 1-16.

Houborg, R., J. B. Fisher and A. K. Skidmore (2015). "Advances in remote sensing of vegetation function and traits." International Journal of Applied Earth Observation and Geoinformation **43**: 1-6.

Huan-Jun, L., Z. Yuan-Zhi, Z. Xin-Le, Z. Bai, S. Kai-Shan, W. Zong-Ming and T. Na (2009). "Quantiative analysis of smoisture effect on black soil reflectance." Pedosphere **19**(4): 9.

Huang, Y., G. Fipps, E. L. Ronald and J. T. Steven (2011). "Landsat Satellite Multi-Spectral Image Classification of Land Cover and Land Use Changes for GIS-Based Urbanization Analysis in Irrigation Districts of Lower Rio Grande Valley of Texas." Appl Remote Sens J **2**: 27-36.

Huemmrich, K. F. (2001). "The GeoSail model: a simple addition to the SAIL model to describe discontinuous canopy reflectance." Remote Sensing of Environment **75**(3): 423-431.

Ibrahimo, M. (2017). Perfil socio-económico dos pequenos agricultores do sul de moçambique: Realidades de chokwé, guijá e kamavota. Maputo, Observatorio do Meio Rural**:** 35.

INE (2011). Censo Agro-Pecuário CAP 2009-2010: Resultados preliminares - Moçambique. I. N. d. E. M. d. Agricultura, Instituto Nacional de Estatistcas**:** 125.

INE (2018). Statistic Yearbook 2017 - Mozambique. N. S. Institute. Maputo, INE**:** 134.

Jackson, R. D., S. B. Idso, R. J. Reginato and P. J. Pinter Jr. (1981). "Canopy temperature as a crop water stress indicator." Water Resources Research **17**(4): 1133-1138.

Jackson, T. J. (2002). "Remote sensing of soil moisture: implications for groundwater recharge." Hydrogeology Journal **10**(1): 40-51.

Jackson, T. J., D. M. Le Vine, C. T. Swift, T. J. Schmugge and F. R. Schiebe (1995). "Large area mapping of soil moisture using the ESTAR passive microwave radiometer in Washita'92." Remote Sensing of Environment **54**(1): 27-37.

Jacobson, A., J. Dhanota, J. Godfrey, H. Jacobson, Z. Rossman, A. Stanish, H. Walker and J. Riggio (2015). "A novel approach to mapping land conversion using Google Earth with an application to East Africa." Environmental Modelling & Software **72**: 1-9.

Jacquemoud, S., C. Bacour, H. Poilvé and J. P. Frangi (2000). "Comparison of Four Radiative Transfer Models to Simulate Plant Canopies Reflectance: Direct and Inverse Mode." Remote Sensing of Environment **74**(3): 471-481.

Jagdhuber, T., I. Hajnsek, A. Bronstert and K. P. Papathanassiou (2013). "Soil Moisture Estimation Under Low Vegetation Cover Using a Multi-Angular Polarimetric Decomposition." IEEE Transactions on Geoscience and Remote Sensing **51**(4): 2201-2215.

Jain, M., P. Mondal, R. S. DeFries, C. Small and G. L. Galford (2013). "Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors." Remote Sensing of Environment **134**: 210-223.

Jensen, J. R. (1996). Introduction to Digital Image Processing: A remote sensing perspective. Upper Saddle River, NJ., Piscataway, NJ: Prentice Hall.

Jensen, J. R. (2000). Remote sensing of environment. An earth resource perspective. New Jersey, Prentice Hall, Inc.

Jin, Z., G. Azzari, C. You, S. Di Tommaso, S. Aston, M. Burke and D. B. Lobell (2019). "Smallholder maize area and yield mapping at national scales with Google Earth Engine." Remote Sensing of Environment **228**: 115-128.

Jonckheere, I., S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss and F. Baret (2004). "Review of methods for in situ leaf area index determination—Part I. Theories, sensors and hemispherical photography." Agricultural and Forest Meteorology **121**: 17.

Jones, H. G. and R. A. Vaughan (2010). Remote Sensing of Vegetation: Principles, Techniques, and Applications. New York, NY, USA, Oxford Oxford University Press.

Justice, C. (2016). Earth observation for agricultural monitoring and food security. EARSEL. Prague, University of Maryland, USA**:** 72.

Kasampalis, D. A., T. K. Alexandridis, C. Deva, A. Challinor, D. Moshou and G. Zalidis (2018). "Contribution of Remote Sensing on Crop Models: A Review." Journal of Imaging **4**(4): 52.

Katherine, A., R. Barbara, S. William, K. Argyro and F. Lawrence (2017). " Earth observation in service of the 2030 Agenda for Sustainable Development." Geo-spatial Information Science **20**(2): 21.

Katherine, A., R. Barbara, S. William, K. Argyro and F. Lawrence (2017). "Earth observation in service of the 2030 Agenda for Sustainable Development." Geo-spatial Information Science **20**(2): 21.

Kattge, J., S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, G. Bönisch, E. Garnier, M. Westoby, P. B. Reich, I. J. Wright, J. H. C. Cornelissen, C. Violle, S. P. Harrison, P. M. Van Bodegom, M. Reichstein, B. J. Enquist, N. A. Soudzilovskaia, D. D. Ackerly, M. Anand, O. Atkin, M. Bahn, T. R. Baker, D. Baldocchi, R. Bekker, C. C. Blanco, B. Blonder, W. J. Bond, R. Bradstock, D. E. Bunker, F. Casanoves, J. Cavender-Bares, J. Q. Chambers, F. S. Chapin Iii, J. Chave, D. Coomes, W. K. Cornwell, J. M. Craine, B. H. Dobrin, L. Duarte, W. Durka, J. Elser, G. Esser, M. Estiarte, W. F. Fagan, J. Fang, F. Fernández-Méndez, A. Fidelis, B. Finegan, O. Flores, H. Ford, D. Frank, G. T. Freschet, N. M. Fyllas, R. V. Gallagher, W. A. Green, A. G. Gutierrez, T. Hickler, S. I. Higgins, J. G. Hodgson, A. Jalili, S. Jansen, C. A. Joly, A. J. Kerkhoff, D. Kirkup, K. Kitajima, M. Kleyer, S. Klotz, J. M. H. Knops, K. Kramer, I. Kühn, H. Kurokawa, D. Laughlin, T. D. Lee, M. Leishman, F. Lens, T. Lenz, S. L. Lewis, J. Lloyd, J. Llusià, F. Louault, S. Ma, M. D. Mahecha, P. Manning, T. Massad, B. E. Medlyn, J. Messier, A. T. Moles, S. C. Müller, K. Nadrowski, S. Naeem, Ü. Niinemets, S. Nöllert, A. Nüske, R. Ogaya, J. Oleksyn, V. G.

Onipchenko, Y. Onoda, J. Ordoñez, G. Overbeck, W. A. Ozinga, S. Patiño, S. Paula, J. G. Pausas, J. Peñuelas, O. L. Phillips, V. Pillar, H. Poorter, L. Poorter, P. Poschlod, A. Prinzing, R. Proulx, A. Rammig, S. Reinsch, B. Reu, L. Sack, B. Salgado-Negret, J. Sardans, S. Shiodera, B. Shipley, A. Siefert, E. Sosinski, J.-F. Soussana, E. Swaine, N. Swenson, K. Thompson, P. Thornton, M. Waldram, E. Weiher, M. White, S. White, S. J. Wright, B. Yguel, S. Zaehle, A. E. Zanne and C. Wirth (2011). "TRY – a global database of plant traits." Global Change Biology **17**(9): 2905-2935.

Kerr, A., H. Rafuse, G. Sparkes, J. Hinchey and H. Sandeman (2011). Visible/infrared spectroscopy (VIRS) as a research tool in economic geology; background and pilot studies from New Foundland and Labrador.

Kerr, Y. (2007). "Soil moisture from space: Where we are?" Hydrogeology Journal **15**: 4.

Kimes, D. S., Y. Knyazikhin, J. L. Privette, A. A. Abuelgasim and F. Gao (2000). "Inversion methods for physically-based models." Remote Sensing Reviews **18**(2-4): 381-439.

Kleyer, M., R. M. Bekker, I. C. Knevel, J. P. Bakker, K. Thompson, M. Sonnenschein, P. Poschlod, J. M. van Groenendael, L. Klimes and e. a. Klimesˇova´ (2008). " The LEDA traitbase: a database of life-history traits of the Northwest European flora." Journal of Ecology **96**: 9.

Knyazikhin, Y., J. Glassy, J. Privette, Y. Tian, A. Lotsch, Y. Zhang, Y. Wang, J. Morisette, P. Votava, R. Myneni, R. Nemani and S. Running (1999). MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document.

Knyazikhin, Y., M. A. Schull, P. Stenberg, M. Mõttus, M. Rautiainen, Y. Yang, A. Marshak, P. Latorre Carmona, R. K. Kaufmann, P. Lewis, M. I. Disney, V. Vanderbilt, A. B. Davis, F. Baret, S. Jacquemoud, A. Lyapustin and R. B. Myneni (2013). "Hyperspectral remote sensing of foliar nitrogen content." Proceedings of the National Academy of Sciences **110**(3): E185-E192.

Koetz, B., F. Baret, H. Poilvé and J. Hill (2005). "Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics." Remote Sensing of Environment **95**(1): 115-124.

Kokaly, R. F., G. P. Asner, S. V. Ollinger, M. E. Martin and C. A. Wessman (2009). " Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies." Remote Sensing of Environment **113**: 14.

Kong, X. and S. R. Dorling (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.

Köppen, W. (1900). " Versuch einer Klassifikation der Klimate, Vorzugsweise nachihren Beziehungen zur Pflanzenwelt [Attempted climate classification in relationto plant distributions]." Geographische Zeitschrift **6**: 18.

Kpienbaareh, D., M. Kansanga and I. Luginaah (2018). "Examining the potential of open source remote sensing for building effective decision support systems for precision agriculture in resource-poor settings." GeoJournal.

Kraft, N. J. B., R. Valencia and D. D. Ackerly (2008). " Functional traits and niche-based tree community assembly in an Amazonian forest." Science **322**(580): 3.

Kuhn, M. and K. Johnson (2013). Applied Predictive Modeling. New York, Springer Science +Business Media.

Kurucu, Y., F. B. Sanli, M. T. Esetlili, M. Bolca and C. Goksel (2009). "Contribution of SAR images to determination of surface moisture on the Menemen Plain, Turkey." International Journal of Remote Sensing **30**(7): 1805-1817.

Kwak, N. and C.-H. Choi (2002). Input feature selection for classification problems, IEEE Press.

Lambert, M.-J., P. C. S. Traoré, X. Blaes, P. Baret and P. Defourny (2018). "Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt." Remote Sensing of Environment **216**: 647-657.

Lannoy, G. J. M. D., R. H. Reichle and V. R. N. Pauwels (2013). "Global Calibration of the GEOS-5 L-Band Microwave Radiative Transfer Model over Nonfrozen Land Using SMOS Observations." Journal of Hydrometeorology **14**(3): 765-785.

Lassalle, P., J. Inglada, J. Michel, M. Grizonnet and J. Malik (2015). "A Scalable Tile-Based Framework for Region-Merging Segmentation." IEEE Transactions on Geoscience and Remote Sensing **53**(10): 5473-5485.

Lavorel, S., K. Grigulis, P. Lamarque, M. Colace, D. Garden, J. Girel, G. Pellet and R. Douzet (2011). " Using plant functional traits to understand the landscape distribution of multiple ecosystem services." Journal of Ecology **99**: 12.

le Maire, G., C. François and E. Dufrêne (2004). "Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements." Remote Sensing of Environment **89**(1): 1-28.

Lee, K.-H. and E. N. Anagnostou (2004). "A combined passive/active microwave remote sensing approach for surface variable retrieval using Tropical Rainfall Measuring Mission observations." Remote Sensing of Environment **92**(1): 112-125.

Lei, S.-g., Z.-f. Bian, J. L. Daniels and D.-l. Liu (2014). "Improved spatial resolution in soil moisture retrieval at arid mining area using apparent thermal inertia." Transactions of Nonferrous Metals Society of China **24**(6): 1866-1873.

Li, J. and S. Wang (2018). "Using SAR-Derived Vegetation Descriptors in a Water Cloud Model to Improve Soil Moisture Retrieval." Remote Sensing **10**(9): 1370.

Li, L., Y. B. Cheng, S. Ustin, X. T. Hu and D. Riaño (2008). "Retrieval of vegetation equivalent water thickness from reflectance using genetic algorithm (GA)-partial least squares (PLS) regression." Advances in Space Research **41**(11): 1755-1763.

Liang, S. (2004). Quantitative Remote Sensing of Land Surfaces. Hoboken, NJ, John Wiley & Sons.

Liao, W., R. Bellens, A. Pizurica, W. Philips and Y. Pi (2012). "Classification of Hyperspectral Data Over Urban Areas Using Directional Morphological Profiles and Semi-Supervised Feature Extraction." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5**(4): 1177-1190.

Liu, H.-J., Y.-Z. Zhang, X.-L. Zhang, B. Zhang, K.-S. Song, Z.-M. Wang and N. Tang (2009). "Quantitative Analysis of Moisture Effect on Black Soil Reflectance." Pedosphere **19**(4): 532-540.

Liu, K., Q.-b. Zhou, W.-b. Wu, T. Xia and H.-j. Tang (2016). "Estimating the crop leaf area index using hyperspectral remote sensing." Journal of Integrative Agriculture **15**(2): 475-491.

Liu, K., Q. Zhou, W. Wu, T. Xia and H. Tang (2016). "Estimating the crop leaf area index using hyperspectral remote sensing." Journal of Integrative Agriculture **15**(2): 475–491.

Liu, P. W., J. Judge, R. DeRoo, A. W. England and A. Luke (2013). "Utilizing complementarity of active/passive microwave observations at L-band for soil moisture studies in sandy soils." In IGARSS: 4.

Liu, Y. Y., W. A. Dorigo, R. M. Parinussa, R. A. M. de Jeu, W. Wagner, M. F. McCabe, J. P. Evans and A. I. J. M. van Dijk (2012). "Trend-preserving blending of passive and active microwave soil moisture retrievals." Remote Sensing of Environment **123**: 280-297.

Llano, M. P., W. Vargas and G. Naumann (2012). "Climate variability in areas of the world with high production of soya beans and corn: its relationship to crop yields." Meteorological Applications **19**(4): 385-396.

Lobell, D. B. and G. P. Asner (2002). "Moisture effects on soil reflectance." Soil Science Society American Journal **66**(3): 6.

Lobell, D. B., G. P. Asner, B. E. Law and R. N. Treuhaft (2002). "View angle effects on canopy reflectance and spectral mixture analysis of coniferous forests using AVIRIS." International Journal of Remote Sensing **23**(11): 2247-2262.

Lu, D. and Q. Weng (2007). "A survey of image classification methods and techniques for improving classification performance." International Journal of Remote Sensing **28**(5): 823-870.

Mahlein, A.-K., E.-C. Oerke, U. Steiner and H.-W. Dehne (2012). "Recent advances in sensing plant diseases for precision crop protection." European Journal of Plant Pathology **133**(1): 197-209.

Main, R., M. A. Cho, R. Mathieu, M. M. O'Kennedy, A. Ramoelo and S. Koch (2011). "An investigation into robust spectral indices for leaf chlorophyll estimation." ISPRS Journal of Photogrammetry and Remote Sensing **66**(6): 751-761.

Mananze, S. E., I. Pôças and M. Cunha (2018). "Maize leaf area estimation in different growth stages based on allometric descriptors." African Journal of Agricultural Research **13(4)**: 202-209.

Mananze, S. E., I. Pôças and M. Cunha (2018). "Retrieval of Maize Leaf Area Index Using Hyperspectral and Multispectral Data." Remote Sensing **10**(12): 30.

Mananze, S. E., I. Pôças and M. Cunha (2019). "Agricultural drought monitoring based on soil moisture derived from the optical trapezoid model in Mozambique." Journal of Applied Remote Sensing **13**(2): 17.

Martínez-Fernández, J., A. Ceballos, S. Casado, C. Morán and V. Hernández (2005). Runoff and soil moisture relationships in a small forested basin in the Sistema Central Ranges (Spain). IAHS-AISH Publication.

Martínez-Fernández, J., A. González-Zamora, N. Sánchez and A. Gumuzzio (2015). "A soil water based index as a suitable agricultural drought indicator." Journal of Hydrology **522**: 265-273.

Martínez-Fernández, J., A. González-Zamora, N. Sánchez, A. Gumuzzio and C. M. Herrero-Jiménez (2016). "Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index." Remote Sensing of Environment **177**: 12.

Martínez, B., F. Camacho, A. Verger, F. J. García-Haro and M. A. Gilabert (2013). "Intercomparison and quality assessment of MERIS, MODIS and SEVIRI FAPAR products over the Iberian Peninsula." International Journal of Applied Earth Observation and Geoinformation **21**: 463-476.

Mausel, P. W., W. J. Kramber and J. K. Lee (1990). "Optimum band selection for supervised classification of multispectral data." Photogrammetric Engineering and Remote Sensing **56**: 6.

McKenzie, N. J., A. D. Sparrow and J. P. Guerschman (2016). The role of remote sensing in agricultural development and poverty alleviation – the STARS Landscaping Study. . Cambera, Australia, CSIRO Australia / ITC at the University of Twente**:** 172.

McPherson, R. A., C. A. Fiebrich, K. C. Crawford, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. Kloesel, A. D. Melvin, H. Shrivastava, J. M. Wolfinbarger, J. P. Bostic, D. B. Demko, R. L. Elliott, S. J. Stadler, J. D. Carlson and A. J. Sutherland (2007). "Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet." Journal of Atmospheric and Oceanic Technology **24**(3): 301-321.

Meroni, M., E. Marinho, N. Sghaier, M. M. Verstrate and O. Leo (2013). "Remote Sensing Based Yield Estimation in a Stochastic Framework — Case Study of Durum Wheat in Tunisia." Remote Sensing **5**(2): 539-557.

MICOA (2008). Relatório do estudo de avaliação da interacção entre a biodiversidade e pobreza em Mozambique. Maputo, Mozambique, Ministry of Coordination of Environmental Affairs**:** 138.

MICOA. (2009). The National Report on Implementation of the Convention on Biological Diversity in Mozambique. Maputo, Mozambique, Ministry for the Coordination of Environmental Affairs**:** 95.

Milton, E. J., M. E. Schaepman, K. Anderson, M. Kneubuhler and N. Fox (2009). " Progress in field spectroscopy." Remote Sensing of Environment **113**: 17.

Minacapilli, M., C. Agnese, F. Blanda, C. Cammalleri, G. Ciraolo, G. D'Urso, M. Iovino, D. Pumo, G. Provenzano and G. Rallo (2009). "Estimation of actual evapotranspiration of Mediterranean perennial crops by means of remote-sensing based surface energy balance models." Hydrol. Earth Syst. Sci. **13**(7): 1061-1074.

MINAG (2011). Plano Estrategico para o Desenvolvimento do Sector Agrario: 2011 - 2020. Maputo, Ministero da Agricultura**:** 76.

Mishra, A., T. Vu, A. V. Veetti and D. Entekhabi (2017). " Drought monitoring with soil moisture active passive (SMAP) measurements." Journal of Hydrology **552**: 13.

Mishra, N. and K. Crews (2014). "Mapping vegetation morphology types in a dry savanna ecosystem: Integrating hierarchical object-based image analysis with Random Forest." International Journal of Remote Sensing **35**: 1175-1198.

Mondo, V. H. V., S. J. P. De Carvalho, V. D. Labonia, D. D. Neto and S. M. Cicero (2009). "Comparação de Métodos para Estimativa ee Área Foliar em Plantas de Milho." Revista Brasileira de Milho e Sorgo **8**(3): 233-246.

Monteith, J. L. and M. H. Unsworth (1990). Principles of environmental physics. London, Edward Arnold.

Moran, M. S., T. R. Clarke, Y. Inoue and A. Vidal (1994). "Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index." Remote Sensing of Environment **49**(3): 246-263.

Moran, M. S., C. D. Peters-Lidard, J. M. Watts and S. McElroy (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.

Morris, M. B. H., D. Byerlee, P. Savanti and J. Staatz (2009). Awakening Africa's Sleeping Giant: Prospects for Commercial Agriculture in the Guinea Savannah Zone and Beyond. Washington, DC, USA, The World Bank.

Mosca, J. (2014). Agricultura familiar em moçambique: Ideologias e políticas. Lisboa, Portugal, School of Economics & Management**:** 34.

Mountrakis, G., J. Im and C. Ogole (2011). "Support vector machines in remote sensing: A review." ISPRS Journal of Photogrammetry and Remote Sensing **66**(3): 247-259.

Myint, S. W. (2001). "A Robust Texture Analysis and Classification Approach for Urban Land-Use and Land-Cover Feature Discrimination." Geocarto International **16**(4): 29-40.

Myint, S. W. (2006). "A New Framework for Effective Urban Land Use and Land Cover Classification: A Wavelet Approach." GIScience & Remote Sensing **43**(2): 155-178.

Myneni, R. B., S. Hoffman, Y. Knyazikhin, J. L. Privette, J. Glassy, Y. Tian, Y. Wang, X. Song, Y. Zhang, G. R. Smith, A. Lotsch, M. Friedl, J. T. Morisette, P. Votava, R. R. Nemani and S. W. Running (2002). "Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data." Remote Sensing of Environment **83**: 17.

Narasimhan, B. and R. Srinivasan (2005). "Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring." Agricultural and Forest Meteorology **133**(1): 69-88.

Narayan, U., V. Lakshmi and T. J. Jackson (2006). "High-resolution change estimation of soil moisture using L-band radiometer and Radar observations made during the SMEX02 experiments." IEEE Transactions on Geoscience and Remote Sensing **44**(6): 1545-1554.

Nearing, G. S., W. T. Crow, K. R. Thorp, M. S. Moran, R. H. Reichle and H. V. Gupta (2012). "Assimilating remote sensing observations of leaf area index and soil moisture for wheat yield estimates: An observing system simulation experiment." Water Resources Research **48**(5).

Neville, R. A., J. Lévesque, K. Staenz, C. Nadeau, P. Hauff and G. A. Borstad (2003). "Spectral unmixing of hyperspectral imagery for mineral exploration: comparison of results from SFSI and AVIRIS." Canadian Journal of Remote Sensing **29**(1): 99-110.

Nocita, M., A. Stevens, C. Noon and B. van Wesemael (2013). "Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy." Geoderma **199**: 37-42.

North, P. R. J. (1996). "Three-dimensional forest light interaction model using a Monte Carlo method." IEEE Transactions on Geoscience and Remote Sensing **34**(4): 946-956.

Oetter, D. R., W. B. Cohen, M. Berterretche, T. K. Maiersperger and R. E. Kennedy (2001). "Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data." Remote Sensing of Environment **76**(2): 139-155.

Ok, A. O., O. Akar and O. Gungor (2012). "Evaluation of random forest method for agricultural crop classification." European Journal of Remote Sensing **45**(1): 421-432.

Okin, G. S., D. A. Roberts, B. Murray and W. J. Okin (2001). "Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments." Remote Sensing of Environment **77**(2): 212-225.

Orwin, K. H., S. M. Buckland, D. Johnson, B. L. Turner, S. Smart, S. Oakley and R. D. Bardgett (2010). "Linkages of plant traits to soil properties and the functioning of temperate grassland." Journal of Ecology **98**: 9.

Ozdarici Ok, A. and Z. Akyurek (2012). "A segment-based approach to classify agricultural lands by using multi-temporal optical and microwave data." International Journal of Remote Sensing **33**(22): 7184-7204.

Pal, M. and P. M. Mather (2003). "An assessment of the effectiveness of decision tree methods for land cover classification." Remote Sensing of Environment **86**(4): 554-565.

Palmer, W. (1965). Meteorological Drought. Washington, DC,, U.S. Weather Bureau.

Palmer, W. C. (1968). "Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index." Weatherwise **21**(4): 156-161.

Paloscia, S., S. Pettinato, E. Santi, C. Notarnicola, L. Pasolli and A. Reppucci (2013). "Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation." Remote Sensing of Environment **134**: 15.

Pan, H., Z. Chen, J. Ren, H. Li and S. Wu (2018). "Estimating winter wheat leaf area index and canopy water content with three different approaches using Sentinel - 2 multi-spectral instrument (S2 MSI) data." Journal of Sellected Topics in Applied Earth Observations and Remote Sensing **99**: 11.

Panciera, R., M. A. Tanase, K. Lowell and J. P. Walker (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.

Panciera, R., J. P. Walker, T. J. Jackson, D. A. Gray, M. A. Tanase, D. Ryu, A. Monerris, H. Yardley, C. Rüdiger, X. Wu, Y. Gao and J. M. Hacker (2014). "The Soil Moisture Active Passive Experiments (SMAPEx): Toward Soil Moisture Retrieval From the SMAP Mission." IEEE Transactions on Geoscience and Remote Sensing **52**(1): 490-507.

Panigada, C., M. Rossini, M. Meroni, C. Cilia, L. Busetto, S. Amaducci, M. Boschetti, S. Cogliati, V. Picchi, F. Pinto, A. Marchesi and R. Colombo (2014). "Fluorescence, PRI and canopy temperature for water stress detection in cereal crops." International Journal of Applied Earth Observation and Geoinformation **30**: 167-178.

Panu, U. S. and T. C. Sharma (2002). "Challenges in drought research: some perspectives and future directions." Hydrological Sciences Journal **47**(sup1): S19-S30.

Paola, J. D. and R. A. Schowengerdt (1995). "A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery." International Journal of Remote Sensing **16**(16): 3033-3058.

Paula, S., M. Arianoutsou, D. Kazanis, C. Tavsanoglu, F. Lloret, C. Buhk, F. Ojeda, B. Luna, J. M. Moreno, A. Rodrigo, J. M. Espelta, S. Palacio, B. Ferna´ ndez-Santos and P. M. Fernandes, Pausas, J.G. (2009). " Fire-related traits for plant species of the Mediterranean Basin." Ecology **90**(1420).

Peña-Barragán, J. M., M. K. Ngugi, R. E. Plant and J. Six (2011). "Object-based crop identification using multiple vegetation indices, textural features and crop phenology." Remote Sensing of Environment **115**(6): 1301-1316.

Petropoulos, G. P. and T. N. Carlson (2011). Retrievals of turbulent heat fluxes and soil moisture content by Remote Sensing. Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications. T. a. Francis**:** 556.

Petropoulos, G. P., T. N. Carlson, M. J. Wooster and S. Islam (2009). "A review of Ts/VI remote sensing based methods for the retrieval of land surface fluxes and soil surface moisture content." Prog. Phys. Geogr. **33**(2): 28.

Petropoulos, G. P., G. Ireland and B. Barrett (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.

Piles, M., N. Sánchez, M. Vall-llossera, A. Camps, J. Martínez-Fernández, J. Martínez and V. González-Gambau (2014). "A Downscaling Approach for SMOS Land Observations: Evaluation of High-Resolution Soil Moisture Maps Over the Iberian Peninsula." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **7**(9): 3845-3857.

Piles, M., M. Vall-llossera, L. Laguna and A. Camps (2012). A downscaling approach to combine SMOS multi-angular and full-polarimetric observations with MODIS VIS/IR data into high resolution soil moisture maps. 2012 IEEE International Geoscience and Remote Sensing Symposium.

Pimm, S. L., C. N. Jenkins, R. Abell, T. M. Brooks, J. L. Gittleman, L. N. Joppa, P. H. Raven, C. M. Roberts and J. O. Sexton (2014). "The biodiversity of species and their rates of extinction, distribution, and protection." Science **344**(6187): 1246752.

Platt, R. V. and A. Goetz (2004). "A comparison of AVIRIS and Landsat for land use classification at the urban fringe." Photogrammetric Engineering and Remote Sensing **70**: 813-819.

Pôças, I., M. Cunha, A. Marçal and L. Pereira (2011). "An evaluation of changes in a mountainous rural landscape of Northeast Portugal using remotely sensed data." Landscape and Urban Planning **101**: 253-261.

Pôças, I., J. Gonçalves, P. M. Costa, I. Gonçalves, L. S. Pereira and M. Cunha (2017). "Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region." International Journal of Applied Earth Observation and Geoinformation **58**: 177-190.

Pôças, I., T. Paço, P. Paredes, M. Cunha and L. S. Pereira (2015). "Estimation of Actual Crop Coefficients Using Remotely Sensed Vegetation Indices and Soil Water Balance Modelled Data." Remote sensing **7**: 2373-2400.

Pôças, I., T. A. Paço, M. Cunha, J. A. Andrade, J. Silvestre, A. Sousa, F. L. Santos, L. S. Pereira and R. G. Allen (2014). "Satellite based evapotranspiration of a super-intensive olive orchard: application of METRIC algorithms." Biosystems Engineering **128**: 69-81.

Pôças, I., A. Rodrigues, S. Gonçalves, P. Costa, I. Gonçalves, L. S. Pereira and M. Cunha (2015). "Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices." Remote Sensing **7**: 16460–16479.

Porporato, A., F. Laio, L. Ridolfi and I. Rodriguez-Iturbe (2001). "Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: III. Vegetation water stress." Advances in Water Resources **24**(7): 725-744.

Price, J. C. (1990). "Using spatial context in satellite data to infer regional scale evapotranspiration." IEEE Transactions on Geoscience and Remote Sensing **28**(5): 940-948.

Qin, J., K. Yang, N. Lu, Y. Chen, L. Zhao and M. Han (2013). "Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia." Remote Sensing of Environment **138**: 1-9.

Qiu, J., W. Gao and B. M. Lesht (1998). "Inverting optical reflectance to estimate surface properties of vegetation canopies." International Journal of Remote Sensing **19**(4): 641-656.

Räsänen, A. and T. Virtanen (2019). "Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes." Remote Sensing of Environment **230**: 111207.

Rashed, T., J. R. Weeks, M. S. Gadalla and A. G. Hill (2001). "Revealing the Anatomy of Cities through Spectral Mixture Analysis of Multispectral Satellite Imagery: A Case Study of the Greater Cairo Region, Egypt." Geocarto International **16**(4): 7-18.

Rembold, F., C. Atzberger, I. Savin and O. Rojas (2013). "Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection." Remote Sensing **5**(4): 1704-1733.

República, A. d. (2004). Constituição da República de Moçambique. A. d. República. Maputo, Imprensa Nacional**:** 97.

Richards, J. A. and X. Jia (2006). Remote Sensing Digital Image Analysis: An Introduction. Berlin, Springer-Verlag.

Richter, K., C. Atzberger, T. B. Hank and W. Mauser (2012). "Derivation of biophysical variables from Earth observation data: validation and statistical measures." Journal of Applied Remote Sensing **6**(1): 063557-063551-063557-063523.

Richter, K., C. Atzberger, F. Vuolo, P. Weihs and G. D'Urso (2009). "Experimental assessment of the Sentinel-2 band setting for RTM-based LAI retrieval of sugar beet and maize." Canadian Journal of Remote Sensing **35**: 230-247.

Richter, T. and L. Fukshansky (1996). "Optics of a bifacial leaf. 1: A novel combined procedure for deriving the optical parameters." Journal of Photochemistry and Photobiology **63**: 10.

Ripley, B. (1996). Pattern Recognition and Neural Networks. Cambridge, Cambridge UniversityPress.

Rivera, J., J. Verrelst, J. Delegido, F. Veroustraete and J. Moreno (2014). "On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization." Remote Sensing **6**: 4927-4951.

Roberts, D., K. Roth and R. Perroy (2011). Hyperspectral Vegetation Indices. Hyperspectral Remote Sensing of Vegetation. New York, USA, CRC Press**:** 309-328.

Rodrigues, A., A. R. S. Marcal and M. Cunha (2013). "Monitoring Vegetation Dynamics Inferred by Satellite Data Using the PhenoSat Tool." IEEE Transactions on Geoscience and Remote Sensing **51**(4): 2096-2104.

Rodriguez-Galiano, V. F., M. Chica-Olmo, F. Abarca-Hernandez, P. M. Atkinson and C. Jeganathan (2012). "Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture." Remote Sensing of Environment **121**: 93-107.

Roumet, C., C. Fortunel, E. Kazakou, E. Garnier, D. Vile, M. Papadimitriou, M. P. Zarovali, V. P. Papanastasis, C. Skarpe, G. Rusch, M. Sternberg, A. Thébault, F. Quétier, K. Grigulis, M. Robson, S. Lavorel, C. Jouany, J.-P. Theau, P. Cruz, P. Ansquer, H. Castro, H. Freitas, J. Lepš, J. Dolezal, H. Quested, O. Eriksson, C. Golodets, J. Kigel, M. Kleyer, T. Meier, V. Lehsten and R. Pakeman (2006). "Assessing the Effects of Land-use Change on Plant Traits, Communities and Ecosystem Functioning in Grasslands: A Standardized Methodology and Lessons from an Application to 11 European SitesGarnier et al. — Methodology to Assess Effects of Land-use ChangeGarnier et al. — Methodology to Assess Effects of Land-use Change." Annals of Botany **99**(5): 967-985.

Rüdiger, C., M. Doubková, J. R. Larsen, W. Wagner and J. P. Walker (2014). "Similarities Between Spaceborne Active and Airborne Passive Microwave Observations at 1 km Resolution." IEEE Geoscience and Remote Sensing Letters **11**(12): 2178-2182.

Rutherfood, R. J. (2010). An Assessment of Rai-Fed Crop Production Potential in South Africa's Neighbouring Countries. South Africa, Department of Water Affairs**:** 41.

Sadeghi, M., E. Babaeian, M. Tuller and S. B. Jones (2017). "The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations." Remote Sensing of Environment **198**: 52-68.

Sadeghi, M., S. B. Jones and W. D. Philpot (2015). "A linear physically-based model for remote sensing of soil moisture using short wave infrared bands." Remote Sensing of Environment **164**: 66-76.

Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka and H. Ohno (2005). "A crop phenology detection method using time-series MODIS data." Remote Sensing of Environment **96**(3): 366-374.

Sánchez-Ruiz, S., M. Piles, N. Sánchez, J. Martínez-Fernández, M. Vall-llossera and A. Camps (2014). "Combining SMOS with visible and near/shortwave/thermal infrared satellite data for high resolution soil moisture estimates." Journal of Hydrology **516**: 273-283.

Sandholt, I., K. Rasmussen and J. Andersen (2002). "A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status." Remote Sensing of Environment **79**(2): 213-224.

Santos, W. J. R., B. M. Silva, G. C. Oliveira, M. M. L. Volpato, J. M. Lima, N. Curi and J. J. Marques (2014). "Soil moisture in the root zone and its relation to plant vigor assessed by remote sensing at management scale." Geoderma **221**: 91-95.

Say, S. M., M. Keskin, M. Sehri and Y. E. Sekerli (2018). "Adoption of precision agriculture technologies in developed and developing countries." The Online Journal of Science and Technology **8**(1): 8.

Schimmelpfennig, D. (2016) "Precision Agriculture Technologies and Factors Affecting Their Adoption."

Schlerf, M. and C. Atzberger (2006). "Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data." Remote Sensing of Environment **100**(3): 281-294.

Schlerf, M., C. Atzberger and J. Hill (2005). "Remote sensing of forest biophysical variables using HyMap imaging spectrometer data." Remote Sensing of Environment **95**(2): 177-194.

Schwilch, G., B. Bestelmeyer, S. Bunning, W. Critchley, J. Herrick, K. Kellner, H. P. Liniger, F. Nachtergaele, C. J. Ritsema, B. Schuster, R. Tabo, G. van Lynden and M. Winslow (2011). "Experiences in monitoring and assessment of sustainable land management." Land Degradation & Development **22**(2): 214-225.

Sehgal, V. K., D. Chakraborty and R. N. Sahoo (2016). "Inversion of radiative transfer model for retrieval of wheat biophysical parameters from broadband reflectance measurements." Information Processing in Agriculture **3**(2): 107-118.

Sheldon, K. E. and J. M. Penvenne (2019) "Mozambique."

Shen, Y., L. Wu, L. Di, G. Yu, H. Tang, G. Yu and Y. Shao (2013). "Hidden Markov Models for Real-Time Estimation of Corn Progress Stages Using MODIS and Meteorological Data." Remote Sensing **5**(4): 1734-1753.

Sims, D. A. and J. A. Gamon (2002). "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages." Remote Sensing of Environment **81**(2–3): 337-354.

Sitoe, A., A. Salomão and S. Wertz-Kanounnikoff (2012). The context of REDD+ in Mozambique: Drivers, agents and institutions. Bogor, Indonesia, CIFOR.

Soe Thwal, N., T. Ishikawa and H. Watanabe (2019). Land cover classification and change detection analysis of multispectral satellite images using machine learning.

Sohrabi, M. M., J. H. Ryu, J. Abatzoglou and J. Tracy (2015). "Development of soil moisture drought index to characterize droughts." Journal of Hydrologic Engineering **20**(11).

Sohrabinia, M., W. Rack and P. Zawar-Reza (2012). "GEOSTATISTICAL ANALYSIS OF SURFACE TEMPERATURE AND IN-SITU SOIL MOISTURE USING LST TIME-SERIES FROM MODIS." Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. **XXXIX-B7**: 17-21.

Soliman, A., J. R. Heck, A. Brenning, R. Brown and S. Miller (2013). "Remote Sensing of Soil Moisture in Vineyards Using Airborne and Ground-Based Thermal Inertia Data." Remote Sensing **5**(8).

Soliman, A., R. J. Heck, A. Brenning, R. Brown and S. Miller (2013). "Remote sensing of soil moisture in vineyards using airborne and ground-based thermal inertia Data. ." Remote Sensing **5**(8): 20.

Sridhar, V., K. G. Hubbard, J. You and E. D. Hunt (2008). "Development of the Soil Moisture Index to Quantify Agricultural Drought and Its “User Friendliness” in Severity-Area-Duration Assessment." Journal of Hydrometeorology **9**(4): 660-676.

Srivastava, P. K., D. Han, M. R. Ramirez and T. Islam (2013). "Machine Learning Techniques for Downscaling SMOS Satellite Soil Moisture Using MODIS Land Surface Temperature for Hydrological Application." Water Resources Management **27**(8): 3127-3144.

Stagakis, S., N. Markos, O. Sykioti and A. Kyparissis (2010). "Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: an application on a

Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations." Remote Sensing of Environment **114**: 17.

Stehman, S. V. and J. A. Milliken (2007). "Estimating the effect of crop classification error on evapotranspiration derived from remote sensing in the lower Colorado River basin, USA." Remote Sensing of Environment **106**(2): 217-227.

Stuckens, J., P. R. Coppin and M. E. Bauer (2000). "Integrating Contextual Information with per-Pixel Classification for Improved Land Cover Classification." Remote Sensing of Environment **71**(3): 282-296.

Swain, H. and S. M. Davis (1978). Remote Sensing: The Quantitative Approach. New York, McGraw- Hill.

Swain, P. H. and R. C. King (1973). Two effective feature selection criteria for multispectral remote sensing. International Joint Conference on Pattern Recognition, Washington, D.C, Purdue University.

Tavares, P., N. Beltrão, U. Guimarães and A. Teodoro (2019). "Integration of Sentinel-1 and Sentinel-2 for Classification and LULC Mapping in the Urban Area of Belém, Eastern Brazilian Amazon." Sensors **19**.

Temimi, M., R. Leconte, N. Chaouch, P. Sukumal, R. Khanbilvardi and F. Brissette (2010). "A combination of remote sensing data and topographic attributes for the spatial and temporal monitoring of soil wetness." Journal of Hydrology **388**(1): 28-40.

Thenkabail, P. S., R. B. Smith and E. De Pauw (2000). "Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics." Remote Sensing of Environment **71**(2): 158-182.

Thornthwaite, C. W. and J. R. Mather (1955). "The water balance." Publications inClimatology **8**: 105.

Tipping, M. E. (2001). "Sparse Bayesian Learning and the Relevance Vector Machine." Journal of Machine Learning Research **1**(3): 211-244.

Torres-Ruiz, J. M., A. Diaz-Espejo, A. Morales-Sillero, M. J. Martín-Palomo, S. Mayr, B. Beikircher and J. E. Fernández (2013). "Shoot hydraulic characteristics, plant water status and stomatal response in olive trees under different soil water conditions." Plant and Soil **373**(1): 77-87.

Torres, G. M., R. P. Lollato and T. E. Ochsner (2013). "Comparison of Drought Probability Assessments Based on Atmospheric Water Deficit and Soil Water Deficit." Agronomy Journal **105**(2): 428-436.

Tran, T., J. Julian and K. de Beurs (2014). "Land Cover Heterogeneity Effects on Sub-Pixel and Per-Pixel Classifications." ISPRS International Journal of Geo-Information **3**: 540-553.

Turner, D. P., W. B. Cohen, R. E. Kennedy, K. S. Fassnacht and J. M. Briggs (1999). " Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites." Remote Sensing of Environment **70**: 16.

UN (2015). Addis Ababa Action Agenda of the Third International Conference on Financing for Development (Addis Ababa Action Agenda). Addis Ababa, Ethiope, United Nations**:** 37.

UN. (2015). Transforming our world: The 2030 agenda for sustainable development. New York, USA, United Nations**:** 41.

Ustin, S. L. and J. A. Gamon (2010). "Remote sensing of plant functional types." New Phytologist **186**: 21.

Valentinuz, O. R. and M. Tollenaar (2006). "Effect of genotype, nitrogen, plant density, and row spacing on the area-per-leaf profile in maize." Agronomy Journal **98**: 6.

Van Wyk, A. E. and G. F. Smith (2001). Regions of fl oristic endemism in southern Africa: a review with emphasis on succulents. Hatfi eld, Pretoria. Pretoria, South Africa, UmdausPress.

Vatsavai, R. R., E. Bright, C. Varun, B. Budhendra, A. Cheriyadat and J. Grasser (2011). Machine learning approaches for high-resolution urban land cover classification: a comparative study. Washington, DC, USA, Association for Computing Machinery.

Verbesselt, J., R. Hyndman, G. Newnham and D. Culvenor (2010). "Detecting trend and seasonal changes in satellite image time series." Remote Sensing of Environment **114**(1): 106-115.

Verbrugghe, M. M. and J. Cierniewski (1997). Influence of soil surface roughness on soil bidirectional reflectance.

Vereecken, H., J. A. Huisman, Y. Pachepsky, C. Montzka, J. van der Kruk, H. Bogena, L. Weihermüller, M. Herbst, G. Martinez and J. Vanderborght (2014). "On the spatio-temporal dynamics of soil moisture at the field scale." Journal of Hydrology **516**: 76-96.

Verger, A., F. Baret and F. Camacho (2011). "Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with CHRIS/PROBA observations." Remote Sensing of Environment **115**(2): 415-426.

Verhoef, W. (1984). "Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model." Remote Sensing of Environment **16**(2): 125-141.

Verrelst, J., J. Muñoz, L. Alonso, J. Delegido, J. P. Rivera, G. Camps-Valls and J. Moreno (2012). "Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3." Remote Sensing of Environment **118**: 127-139.

Verrelst, J., J. P. Rivera, F. Veroustraete, J. Muñoz-Marí, J. G. P. W. Clevers, G. Camps-Valls and J. Moreno (2015). "Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – A comparison." ISPRS Journal of Photogrammetry and Remote Sensing **108**: 260-272.

Verrelst, J., M. E. Schaepman, Z. Malenovský and J. G. P. W. Clevers (2010). "Effects of woody elements on simulated canopy reflectance: Implications for forest chlorophyll content retrieval." Remote Sensing of Environment **114**(3): 647-656.

Vicente-Serrano, S. M., X. Pons-Fernández and J. M. Cuadrat-Prats (2004). "Mapping soil moisture in the central Ebro river valley (northeast Spain) with Landsat and NOAA satellite imagery: a comparison with meteorological data." International Journal of Remote Sensing **25**(20): 4325-4350.

Violle, C., M. L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel and E. Garnier (2007). "Let the concept of trait be functional." Oikos **116**: 11.

Vyas, S. S., B. K. Bhattacharya, R. Nigam, P. Guhathakurta, K. Ghosh, N. Chattopadhyay and R. M. Gairola (2015). "A combined deficit index for regional agricultural drought assessment over semi-arid tract of India using geostationary meteorological satellite data." International Journal of Applied Earth Observation and Geoinformation **39**: 28-39.

Waldner, F., S. Fritz, A. Di Gregorio and P. Defourny (2015). "Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps." Remote Sensing **7**(6): 27.

Walker, T., R. Pitoro, A. Tomo, I. Sitoe, C. Salência, R. Mahanzule, C. Donovan and F. Mazuze (2006). Priority Setting for Public-Sector Agricultural Research in Mozambique with the National Agricultural Survey Data. Research Report Series. Maputo, Institute of agricultural research of mozambique.

Wang, L., J. J. Qu, S. Zhang, X. Hao and S. Dasgupta (2007). "Soil moisture estimation using MODIS and ground measurements in eastern China." International Journal of Remote Sensing **28**(6): 1413-1418.

Wang, W., X. Yao, Y.-c. Tian, X.-j. Liu, J. Ni, W.-x. Cao and Y. Zhu (2012). "Common Spectral Bands and Optimum Vegetation Indices for Monitoring Leaf Nitrogen Accumulation in Rice and Wheat." Journal of Integrative Agriculture **11**(12): 2001-2012.

Wang, Z., A. K. Skidmore, T. Wang, R. Darvishzadeh and J. Hearne (2015). "Applicability of the PROSPECT model for estimating protein and cellulose+lignin in fresh leaves." Remote Sensing of Environment **168**: 205-218.

Weidong, L., F. Baret, G. Xingfa, T. Qingxi, Z. Lanfen and Z. Bing (2002). "Relating soil surface moisture to reflectance." Remote Sensing of Environment **81**(2): 238-246.

Weiss, M., F. Baret, R. B. Myneni, A. Pragnère and Y. Knyazikhin (2000). "Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data." Agronomie **20**(1): 3-22.

Wenng, H. T. (2017). Estimation and validation of the biophysical parameter leaf area index for agricultural areas from satellite sentinel-2a data. Master of Science, University of Natural Resources and Life Science, Vienna.

Wetzel, P. J. and J. T. Chang (1988). " Evaporation from non-uniform surface: A first approach for short-term numerical weather prediction." Monthly Weather Review **116**: 22.

White, F. (1983). The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. Paris, UNESCO.

Wilhite, D. A. and M. H. Glantz (1985). "Understanding: the Drought Phenomenon: The Role of Definitions." Water International **10**(3): 111-120.

Wold, S., H. Martens and H. Wold (1983). The multivariate calibration problem in chemistry solved by the PLS method, Berlin, Heidelberg, Springer Berlin Heidelberg.

Wright, I. J., P. B. Reich, M. Westoby, D. D. Ackerly, Z. Baruch, F. Bongers, J. Cavender-Bares, T. Chapin and e. a. Cornelissen (2004). "The worldwide leaf economics spectrum." Nature **428**: 7.

Wu, W., C. Zucca, F. Karam and G. Liu (2016). "Enhancing the performance of regional land cover mapping." International Journal of Applied Earth Observation and Geoinformation **52**: 422-432.

Xiao, Q. G., W. Y. Chen and Y. W. Shen (1994). "A study on soil moisture monitoring using NOAA satellite." Quart. J. Appl. Meteorol **5**(2): 6.

Xiao, X., S. Boles, S. Frolking, W. Salas, B. Moore, C. Li, L. He and R. Zhao (2002). "Landscape-scale characterization of cropland in China using Vegetation and Landsat TM images." International Journal of Remote Sensing **23**(18): 3579-3594.

Xu, Y., L. Yu, F. R. Zhao, X. Cai, J. Zhao, H. Lu and P. Gong (2018). "Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa." Remote Sensing of Environment **218**: 13-31.

Yao, X., Y. Huang, G. Shang, C. Zhou, T. Cheng, Y. Tian, W. Cao and Y. Zhu (2015). "Evaluation of Six Algorithms to Monitor Wheat Leaf Nitrogen Concentration." Remote Sensing **7**(11): 14939.

Yin, Z., S. C. Dekker, B. J. J. M. van den Hurk and H. A. Dijkstra (2014). "Effects of vegetation structure on biomass accumulation in a Balanced Optimality Structure Vegetation Model (BOSVM v1.0)." Geoscientific Model Development **7**: 24.

Zarco-Tejada, P. J., V. González-Dugo, L. E. Williams, L. Suárez, J. A. J. Berni, D. Goldhamer and E. Fereres (2013). "A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index." Remote Sensing of Environment **138**: 38-50.

Zerrouki, N. and D. Bouchaffra (2014). Pixel-based or Object-based: Which approach is more appropriate for remote sensing image classification? 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC).

Zhan, Z., Q. Qin, A. Ghulan and D. Wang (2007). "NIR-red spectral space based new method for soil moisture monitoring." Science in China Series D: Earth Sciences **50**(2): 283-289.

Zhang, C., E. Pattey, J. Liu, H. Cai, J. Shang and T. Dong (2017). "Retrieving Leaf and Canopy Water Content of Winter Wheat Using Vegetation Water Indices." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing: 16.

Zhang, D., R. Tang, B. Tang, H. Wu and Z. Li (2015). "A Simple Method for Soil Moisture Determination From LST–VI Feature Space Using Nonlinear Interpolation Based on Thermal Infrared Remotely Sensed Data." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **8**(2): 638-648.

Zhang, D., R. Tang, W. Zhao, B. Tang, H. Wu, K. Shao and Z. L. Li (2014). "Surface Soil Water Content Estimation from Thermal Remote Sensing based on the Temporal Variation of Land Surface Temperature." Remote Sensing **6**(4): 18.

Zhang, D., R. Tang, W. Zhao, B. Tang, H. Wu, K. Shao and Z. L. Li (2014). " Surface soil water content estimation from thermal remote sensing based on the temporal variation of land surface temperature." Remote Sensing **6**(4): 18.

Zhang, D. and G. Zhou (2016). "Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review." Sensors **16**(8): 1308.

Zhang, J., J. Yang, H. Lu, W. Wu, J. Huang and S. Chang (2015). "Subwavelength TE/TM grating coupler based on silicon-on-insulator." Infrared Physics & Technology **71**: 542-546.
Zhang, L., H. Lin and D. Li (2002). Influences of the illumination and viewing geometry on the vegetation indices.
Zhao, D., L. Huang, J. Li and J. Qi (2007). "A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy." ISPRS Journal of Photogrammetry and Remote Sensing **62**(1): 8.
Zhao, W. and Z.-L. Li (2013). "Sensitivity study of soil moisture on the temporal evolution of surface temperature over bare surfaces." International Journal of Remote Sensing **34**(9-10): 3314-3331.
Ziyu, W., W. Wenxia, Z. Shuhe and C. Xiuwan (2004). Object-oriented classification and application in land use classification using SPOT-5 PAN imagery. IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium.