Université de Liège
Faculté des Sciences
Département des Sciences et Gestion de l'Environnement

Advancing agricultural monitoring for improved yield estimations using SPOT-VGT and PROBA-V type remote sensing data

Yetkin Özüm Durgun

Thèse présentée en vue de l'obtention
du grade de docteur en sciences

juin 2018

Composition du jury:

Président:	Pr Philippe ANDRE (ULiège)
Promoteur:	Pr Bernard TYCHON (ULiège)
Lecteurs:	Dr Riad BALAGHI (INRA-Morocco)
	Dr Anne GOBIN (VITO)
	Dr Mirco BOSCHETTI (CNR-Italy)
	Dr Grégory DUVEILLER (JRC-Ispra)
	Ir Sven GILLIAMS (VITO)
	Dr Joost WELLENS (ULiège) (Secrétaire)

Avec l'appui de:

Hayat kısa,
Kuşlar uçuyor.
Cemal Süreya

Summary

Accurate and timely crop condition monitoring is crucial for food management and the economic development of any nation. However, accurately estimating crop yield from the field to global scales is a challenge. According to the global strategy of the World Bank, in order to improve national agricultural statistics, crop area, crop production, and crop yield are key variables that all countries should be able to provide. Crop yield assessment requires that both an estimation of the quantity of a product and the area provided for that product should be available. The definition seems simple; however, these measurements are time consuming and subject to error in many circumstances. Remote sensing is one of several methods used for crop yield estimation. The yield results from a combination of environmental factors, such as soil, weather, and farm management, which are responsible for the unique spectral signature of a crop captured by satellite images. Additionally, yield is an expression of the state, structure, and composition of the plant. Various indices, crop masks, and land observation sensors have been developed to remotely observe and control crops in different regions.

This thesis focuses on how much low spatial resolution satellites, such as Project for On-Board Autonomy-Vegetation (PROBA-V), can contribute to global crop monitoring by aiding the search for improved methods and datasets for better crop yield estimation. This thesis contains three chapters. The first chapter explores how an existing product, Dry Matter Productivity (DMP), that has been developed for Satellites Pour l'Observation de la Terre or Earth-observing Satellites-VeGeTation (SPOT-VGT), and transferred to PROBA-V, can be improved to more closely relate to yield anomalies across selected regions. This chapter also covers the testing of the contribution of stress factors to improve wheat and maize yield estimations. According to Monteith's theory, crop biomass linearly correlates with the amount of Absorbed Photosynthetically Active Radiation (APAR) and constant Radiation Use Efficiency (RUE) downregulated by stress factors such as CO_2, fertilization, temperature, and water stress. The objective of this chapter is to investigate the relative importance of these stress factors in relation to the regional biomass production and yield. The production efficiency model Copernicus Global Land Service-Dry Matter Productivity (CGLS-DMP), which follows Monteith's theory, is modified and evaluated for common wheat and silage maize in France, Belgium, and Morocco using SPOT-VGT for the 1999–2012 period. The correlations between the crop yield data and the cumulative modified DMP, CGLS-DMP, Fraction of APAR (fAPAR), and Normalized Difference Vegetation Index (NDVI) values are analyzed for different crop growth stages. The best results are obtained when combinations of the most appropriate stress factors are included for each selected region, and the modified DMP during the reproductive stage is accumulated. Though no single solution can demonstrate an improvement of the global product, the findings support an extension of the methodology to other regions of the world.

The second chapter demonstrates how PROBA-V can be used effectively for crop identification mapping by utilizing spectral matching techniques and phenological characteristics of different crop types. The study sites are agricultural areas spread across the globe, located in Flanders (Belgium), Sria (Russia), Kyiv (Ukraine), and Sao Paulo (Brazil). The data are collected for the 2014–2015 season. For each pure pixel within a field, the NDVI profile of the crop type for its growing season is matched with the reference NDVI profile. Three temporal windows are tested within the growing season: green-up to senescence, green-up to dormancy, and minimum NDVI at the beginning of the growing season to minimum NDVI at the end of the growing season. In order of importance, the crop phenological development period, parcel size, shorter time window, number of ground-truth parcels, and crop calendar similarity are the main reasons behind the differences between the results. The methodology described in this chapter demonstrates the potentials and limitations of using 100 m PROBA-V with revisiting frequency every 5 days in crop identification across different regions of the world.

The final chapter explores the trade-off between the different spatial resolutions provided by PROBA-V products versus the temporal frequency and, additionally, explores the use of thermal time to improve statistical yield estimations. The ground data are winter wheat yields at the field level for 39 fields across Northern France during one growing season 2014–2015. An asymmetric double sigmoid function is fitted, and the NDVI values are integrated over thermal time and over calendar time for the central pixel of the field, exploring different thresholds to mark the start and end of the cropping season. The integrated NDVI values with different NDVI thresholds are used as a proxy for yield. In addition, a pixel purity analysis is performed for different purity thresholds at the 100 m, 300 m, and 1 km resolutions. The findings demonstrate that while estimating winter wheat yields at the field level with pure pixels from PROBA-V products, the best correlation is obtained with a 100 m resolution product. However, several fields must be omitted due to the lack of observations throughout the growing season with the 100 m resolution dataset, as this product has a lower temporal resolution compared to 300 m and 1 km.

This thesis is a modest contribution to the remote sensing and data analysis field with its own merits, in particular with respect to PROBA-V. The experiments provide interesting insight into the PROBA-V dataset at 1 km, 300 m, and 100 m resolutions. Specifically, the results show that 100 m spatial resolution imagery could be used effectively and advantageously in agricultural crop monitoring and crop identification at local – field level – and regional – the administrative regions defined by the national governments – levels. Furthermore, this thesis discusses the limitations of using a low-resolution satellite, such as the PROBA-V 100 m dataset, in crop monitoring and identification. Also, several recommendations are made for space agencies that can be used when designing the new generation of satellites.

Résumé

Un monitoring précis et à temps des conditions culturales est crucial pour le développement économique d'une nation et de sa gestion alimentaire. Cependant, cette estimation précise des rendements culturaux reste un défi de l'échelle locale à l'échelle globale. Selon la stratégie globale de la Banque Mondiale pour l'amélioration des statistiques agricoles nationales, les surfaces, les rendements et les productions des cultures sont des variables clés que chaque pays devrait pouvoir fournir. L'estimation du rendement d'une culture exige de connaître tant la quantité produite que celle de la surface pour cette production. La définition est simple; cependant leurs mesures sont couteuses en temps et elles sont toutes les deux sources d'erreur dans beaucoup de circonstances. La télédétection est une méthode utilisée parmi d'autres pour l'estimation des rendements des cultures. Le rendement d'une culture est lié à un ensemble de facteurs comme le sol, le climat et la gestion de l'agriculteur. Ces facteurs sont à la base de la signature spectrale unique de chaque culture observée dans une image satellite, signature qui exprime l'état, la structure et la composition de la plante. Divers indices, masques culturaux et senseurs d'observation de la terre ont été développés pour observer à distance et contrôler les cultures dans différentes régions.

Cette thèse s'intéresse à l'apport de satellites à basse résolution du type de Project for On-Board Autonomy-Vegetation (PROBA-V) dans le suivi global des cultures par l'amélioration de méthodes et l'apport de nouvelles données pour une meilleure estimation des rendements. Elle a 3 chapitres. Le premier explore comment un produit existant, la productivité de matière sèche (DMP), qui a été développé pour Satellites Pour l'Observation de la Terre or Earth-observing Satellites-VeGeTation (SPOT-VGT) (et transféré à PROBA-V) peut être amélioré pour mieux exprimer les anomalies de rendements dans des régions sélectionnées. On s'est concentré sur la contribution de facteurs de stress dans l'amélioration de l'estimation du rendement du blé et du maïs. Selon la théorie de Monteith, la biomasse culturale est corrélée linéairement avec la quantité de rayonnement photosynthétiquement actif absorbé par la plante (APAR) et une constante d'efficience d'utilisation du rayonnement (RUE) ajustée en fonction de facteurs de stress tels que le niveau de CO2, de fertilisation, de température et d'eau disponible dans le sol. L'objectif était d'investiguer l'importance relative de ces facteurs de stress en relation avec les rendements et productions de biomasse à échelle régionale. Le modèle d'efficience de production Copernicus Global Land Service – Dry Matter Productivity (CGLS-DMP), qui suit la théorie de Monteith a été modifié et évalué pour le blé tendre et le maïs ensilage en France, Belgique et Maroc sur base d'images SPOT-VGT pour la période 1999-2012. La corrélation entre les données officielles de rendements et les valeurs cumulées de DMP modifiée, CGLS-DMP, fAPAR et NDVI ont été analysées pour différents stades de développement de la culture. Les meilleurs résultats ont été obtenus lorsque l'on inclut des combinaisons des facteurs de stress les plus appropriés pour chaque région sélectionnée et en cumulant le DMP modifié pendant le stade de développement

correspondant à la phase de reproduction. Bien qu'il n'existe pas une solution unique à l'amélioration du produit global, nos découvertes nous encouragent à développer la méthode dans d'autres régions du monde.

Le second chapitre montre comment PROBA-V peut être utilisé pour l'identification cartographique des cultures en utilisant les techniques de correspondances spectrales et les caractéristiques phénologiques de différents types de cultures. Les sites d'études étaient répartis globalement dans des zones agricoles de Flandre (Belgique), de la région de Sria (Russie), de la région de Kiev (Ukraine) et de la région de Sao Paulo (Brésil). Le suivi a été réalisé au cours de la saison 2014-2015. Pour chaque pixel pur à l'intérieur d'un champ, le profil de NDVI au cours de la croissance culturale a été mis en correspondance avec un profil NDVI de référence correspondant aux différentes cultures de chaque région. Trois fenêtres temporelles ont été testées pendant la période de croissance : reprise de végétation à senescence, reprise de végétation à dormance, et minimum de NDVI au début de la croissance au minimum de NDVI à la fin du cycle de la culture. Par ordre d'importance, la période de développement de la phénologie de la culture, la taille de la parcelle, la fenêtre temporelle plus réduite, le nombre de parcelles de contrôle de terrain et la similarité du calendrier cultural étaient les principales raisons expliquant les différences dans les résultats. La méthodologie décrite dans ce chapitre démontre le potentiel mais aussi les limites liées à l'utilisation d'image PROBA-V à 100 m avec une fréquence de revisite de 5 jours dans l'identification des cultures dans différentes régions du monde.

Le dernier chapitre analyse le compromis entre les différentes résolutions spatiales disponibles sur PROBA-V et sa fréquence temporelle de revisite. On y étudie aussi l'apport du temps thermique dans l'amélioration des estimations statistiques de rendement. Les données de terrain comprennent les rendements de 39 champs de blé pendant une campagne agricole (2014-2015) dans le Nord de la France. Une fonction sigmoïde asymétrique double a été ajustée et les valeurs de NDVI ont été intégrées en fonction du temps thermique et du calendrier classique pour le pixel central du champ en testant différents seuils pour définir le début et la fin du cycle de la culture. Les valeurs intégrées de NDVI avec différents seuils de NDVI ont été utilisées comme des proxy du rendement. En plus, l'analyse de la pureté du pixel a été réalisée pour différents seuils de pureté à 100, 300 et 1000 m de résolution. On montre que pour l'estimation du rendement du blé d'hiver au niveau des parcelles avec des pixels purs de PROBA-V, la meilleure corrélation est obtenue avec les produits à 100 m de résolution. Cependant, plusieurs champs ont dû être éliminés en raison du manque d'observations pendant la saison culturale avec le senseur à 100 m à cause d'une résolution temporelle plus faible que pour les produits à 300 m et 1 km.

Cette thèse est une contribution modeste au domaine de la télédétection et de l'analyse de données avec ses propres mérites, en particulier en ce qui concerne PROBA-V. Les expériences ont fourni des informations intéressantes sur les données PROBA-V à des résolutions de 1 km, 300 m et 100 m. En particulier, les résultats montrent qu'une imagerie à résolution spatiale de 100 m pourrait être utilisée efficacement et avantageusement dans l'identification et la surveillance des cultures agricoles au niveau local (champ) et régional (régions administratives définies par les gouvernements nationaux). En outre, cette thèse traite des limites de l'utilisation de données satellitaires à basse résolution du type PROBA-V 100 m dans la surveillance et l'identification des cultures. Plusieurs recommandations sont faites pour les agences spatiales en vue de la conception de satellites de nouvelle génération.

Acknowledgements

This thesis could not be realized without priceless support and help of many people. I would like to thank everyone who offered their supervision throughout my PhD period at VITO and ULg over the past five years.

This thesis has been realized in the framework of VEGETATION project collaboration and support of VITO and BELSPO (Belgian Federal Science Policy Office), without which this doctoral study could not be accomplished.

I would like to express my sincerest thanks and gratitude to my thesis committee members. Bernard Tychon guided, supervised and encouraged me thoroughly throughout this thesis. I have learnt a lot from him. My special thanks go to Grégory Duveiller who advised me, supported me and always gave challenging ideas. His significant criticisms and expertise during this thesis period added a different perspective to my scientific point of view. A special thanks is extended to Anne Gobin who has joined later to the committee but has always been understanding, supporting and advising me. I am particularly thankful to her for her editing skills in the completion of our scientific papers. I convey my thanks to Sven Gilliams for his support and guidance during my PhD period. I would also like to thank Riad Balaghi and Mirco Boschetti for having accepted to be part of the jury, to Philippe Andre to be the president of the jury and to Joost Wellens to be the secretary of the jury.

I would like to thank all data providers, from which ground data and satellite data was received. Without their support, this thesis would have never been completed. Thanks to E-AGRI project, MARS-JRC for SPOT-VGT, PROBA-V imageries and meteorological data, NASA for the MODIS imagery, INRA Meknès, GDI - Belgium, Drone Agricole - France, FP7 SIGMA project partners: IKI - Russia, SRI - Ukraine and Guerric le Maire - Brazil.

I also thank to all colleagues at VITO and ULg for their support and friendship during the period of this work.

My special thanks go to my family. I deeply thank my parents for their unconditional love, trust and encouragement. I would also like to thank our sunshine, Enar Atlas, who joined us towards the end of this thesis. Tough this thesis could be finished far before, without him it would not be that fun and same as it was. Finally, I thank to Cihan for sharing life together, caring me every day, still surprising me and being love of my life. I am particularly grateful to him for his understanding, support and time for ‘scientific’ discussions throughout this PhD journey.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AC	Agreement Coefficient
ADSF	Asymmetric Double Sigmoid Function
AET	Actual EvapoTranspiration
APAR	Absorbed Photosynthetically Active Radiation
AR	Autotrophic Respiration
AVHRR	Advanced Very High Resolution Radiometer
BE	Belgium
BPs	Biophysical Proxies
CGLS-DMP	Copernicus Global Land Service-Dry Matter Productivity
DMP	Dry Matter Productivity
EC	Eddy Covariance
ED	Euclidian Distance
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization of the United Nations
fAPAR	fraction of Absorbed Photosynthetically Active Radiation
fCOVER	fractional cover
FR	France
GAI	Green Area Index
GDD	Growing Degree Days
GPP	Gross Primary Productivity
JRC-MARS	Joint Research Centre Monitoring Agriculture by Remote Sensing
LAI	Leaf Area Index
LAI_g	green Leaf Area Index
MAR	Morocco
MARSOP	Monitoring Agriculture by Remote Sensing OPerational
MAE	Mean Absolute Error
MBE	Mean Bias Error
MODIS	MODerate resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
NPP	Net primary Productivity
PET	Potential EvapoTranspiration
PROBA-V	Project for On-Board Autonomy-Vegetation
RMSE	Root Mean Square Error
RRMSE	Relative RMSE
RUE	Radiation Use Efficiency
SIGMA	Stimulating Innovation for Global Monitoring of Agriculture
SMTs	Spectral Matching Techniques
SPIRITS	Software for the Processing and Interpretation of Remotely sensed Image Time Series

SPOT-VGT	Satellites Pour l'Observation de la Terre or Earth-observing Satellites - VeGeTation
SSV	Spectral Similarity Value
VIs	Vegetation Indices

List of Symbols

ε_{AR}	Fraction kept after Autotrophic Respiration
μ_r	Mean pure pixel NDVI profile from validation set
μ_{ref}	Mean reference NDVI profile
$C_{CO_2}^{ref}$	Concentration in the reference year 1833
C_{CO_2}	Actual CO_2 concentration
D_d	Middle dates of the decreasing segments
D_i	Middle dates of the increasing segments
E_1	Index of model performance
ED_{normal}^2	Normalized Euclidean distance between the different candidate reference NDVI profiles and the pure pixel profiles
$\varepsilon_{RUE_{MAX}}$	Maximum Radiation Use Efficiency
F_{CO_2}	CO_2 Assimilation rate in year *x*
$F_{CO_2}^{ref}$	CO_2 Assimilation rate in the reference year 1833
Kp	Michaelis–Menten constant for CO_2
M	Historical maximum NDVI of the reference profile for a logical comparison
m	Historical minimum NDVI of the reference profile for a logical comparison
p	Overall changing rates of the increasing slope
q	Overall changing rates of the decreasing slope
R	Total shortwave incoming solar radiation (0.2–3.0 µm) (GJ/ha/day)
R^2	Coefficient of determination
ref_i	Reference NDVI profile at time *i* from 1 to *n*
r_i	Pure pixel NDVI profile from validation set at time *i* from 1 to *n*
T_{12}	Day time temperature
T_{24}	Day and night temperatures
T_{base}	Base temperature
T_M	Ten daily mean composite of mean air temperature
T_{max}	Maximum daily temperature
T_{min}	Minimum daily temperature
$V(t)$	NDVI at the time *t*
V_a	Amplitude of NDVI variation within the current growing cycle
V_b	Background NDVI value corresponding to non-growing season
V_{pmax}	Maximum phosphoenolpyruvate carboxylase activity
W	Above ground biomass
ε_{CO2}	Normalized CO_2 fertilization effect
ε_{H2O}	Water stress
ε_T	Normalized temperature effect
ρ^2	Correlation coefficient
$\rho_{min}^{-2.6}$	Minimum reflectance in the red channel of the satellite

σ_r	Standard deviation of the pure pixel NDVI profile from the validation set
σ_{ref}	Standard deviation of the reference NDVI profile

Introduction

Crop Yield Estimation

Article 25 of the United Nations Universal Declaration of Human Rights (1948) states that all people have the right to a standard of living that is adequate for the health and well-being of themselves and their families, including food [1]. According to projections, as displayed in Figure 1, the world population will increase by 51%, cereal production will grow by 54%, and arable land area will expand 2% by 2080 [2].

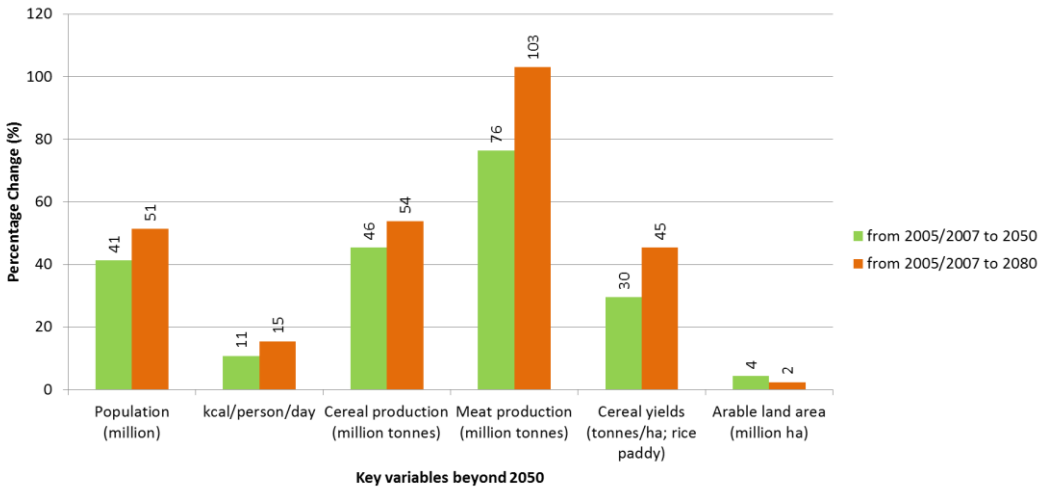


Figure 1. Key variable percentage changes from 2005/2007 to 2050 and 2080, adapted from Alexandratos and Bruinsma [2]

However, this production increase will probably not be sufficient to ensure food security for everyone [3]. Almost a third of the world's population could remain chronically undernourished due to unequal distribution of and access to food [4]. Increasing agricultural productivity could contribute toward increased food security and allow countries to produce a surplus, particularly of cereals, for export to food-deficient regions [4].

Governments and organizations monitor crop growth and estimate yields in order to plan the national food production and tackle any unexpected weather events throughout the growing season. There is, therefore, an increasing demand for methods that provide timely, accurate, efficient, and affordable information about crop yield and cultivated area estimates in near-real time at the regional, national, and global scales. Through information about the current situation of global and local agricultural production, expectations could be shaped for future food prices which would allow markets to

function more efficiently. Furthermore, timely information about potential and observed harvest shortfalls could hasten the early identification of problem areas and help to organize and optimize food supplies on a regional scale [5].

Vegetation behavior depends on the nature of the vegetation itself; its interactions with solar radiation and other climatic factors, such as precipitation, temperature, etc.; and the availability of chemical nutrients and water in the soil [6]. Therefore, crop yield depends on the amounts of the following: energy, mainly incoming solar radiation; water, including precipitation and water retention; and nutrients, including soil chemistry, provided to the plant. Crop yield also depends on the capacity of the crop to use all of these growing factors, as expressed in the crop phenotype, and, of course, on the cropping areas [6].

Remote sensing is a key tool to analyze short-term and long-term changes in vegetation growth. Remote sensing data can provide accurate and timely information on crop yield estimates and crop identification mapping. It enables the acquisition of remote information from the electromagnetic spectrum at different wavelengths, including visible, near infrared, thermal infrared, and microwave. When the sun's radiation reaches the Earth's surface, some of the energy at specific wavelengths is absorbed, and the rest of the energy is reflected. Green leaves absorb the blue and red wavelengths and reflect the green wavelengths, which are, therefore, detected by human eyes. Plant chlorophyll pigment molecules absorb electromagnetic radiation in the wavelength range of 400–700 nm – visible light – which is called the Photosynthetically Active Radiation (PAR) spectral region. The reflectance of solar radiation differs for soil, green vegetation, and dead vegetation over the range of wavelengths, 300–2500 nm, in the electromagnetic spectrum. This spectral signature of the surface material differs by what is useful to detect differences in the crop growth during the growing season. The reflectance for green vegetation is low in both the blue and red regions of the spectrum because of the absorption by chlorophyll pigments. The reflectance has a small peak in the green region and is much higher in the near infrared region than in the visible band because of the cellular structure of the leaves. The shape of the reflectance spectrum, known as the spectral signature, can be used for identification of the vegetation type and vegetation stress [6]. For instance, dry vegetation has a higher reflectance in the visible region but a lower reflectance in the near infrared region. Therefore, remotely sensed images enable the monitoring of vegetation status which can change depending on factors such as the leaf moisture content and health of the plants.

The ability of remote sensing satellites to obtain extensive coverage of the Earth has made this tool an asset for the monitoring of crop production. The first civilian satellite for Earth observation, called Landsat-1, was launched in 1972, and in 1974 the North-American Large Area Crop Inventory Experiment demonstrated the improvement in United States Department of Agriculture (USDA) wheat production predictions that were achieved by using satellite imagery in the United States, Canada, and the former Soviet Union. The methodology was extended to other crops and regions by the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing program which used satellite imagery to evaluate crop condition throughout the growing season and to estimate yield [7]. The Monitoring Agriculture by Remote Sensing (MARS) program was established by the European Union (EU) in 1988 to provide agricultural statistics on yield and crop areas for EU countries and is operationally based on the combined use of remote sensing observations and agrometeorological modeling [5]. There are currently several operational agricultural monitoring systems that use remote sensing to observe crop growth conditions and operate at national and international scales. Several examples of these systems include the United Nations Food and Agricultural Organization global information and early warning system, the United States agency for international

development Famine Early Warning Systems NETwork (FEWS-NET), the USDA Foreign Agricultural Service (FAS) GLobal Agriculture Monitoring (GLAM) System, Group on Earth Observations Global Agricultural Monitoring Initiative, the Chinese CropWatch system, the Indian Forecasting Agricultural output using Space, Agro-meteorology and Land based observations system, and the Belgian crop growth monitoring system [5,8].

Various approaches have been developed for monitoring crop growth and estimating crop yields, and they can be broadly divided into two categories. The first category includes empirical-statistical models that relate regional crop yields to climatic variables and remote sensing-based vegetation indices using methods such as multiple regression analysis [9]. Crops respond differently to changes in weather conditions, including drought, water excess, and low and high temperatures, depending on the crop type and management practices. Remote sensing images for the current growing season can be compared with the corresponding long-term statistics or statistics from the previous year to derive anomaly images in order to observe the interannual changes [10]. Therefore, this approach allows anomaly images to be analyzed for qualitative crop growth monitoring and for useful information from temporal, or seasonal, profiles of remotely derived vegetation indices to be obtained [11]. One of the main shortcomings of these models is that their application is valid only for the areas for which they have been calibrated [12]. Other drawbacks include problems of co-linearity between predictor variables, for example, temperature and precipitation; assumptions of stationarity, for example, that past relationships will hold in the future, even if management events evolve;, and low signal-to-noise ratios in yield in many locations [13]. An additional flaw is that crop masks and reference information, such as agricultural statistics and crop yield, are required for the application and calibration of the approach [11]. Although the reference information can be accessible on a regional level through the national statistics of each country, crop masks are, mostly, not readily available. As a crop mask, one of the following could be used: a cropland mask that covers all crops, a crop-specific mask that is prepared for a specific study area, or a dynamic crop mask. For this reason, lack of a crop mask limits the applicability of the method in many regions of the world.

The second category includes process-based crop growth models and is based on simulation models to create detailed information about plant physiological processes in order to explain how crops function as a whole [8]. These models try to understand crop behavior by explaining crop growth and development through underlying physiological mechanisms. The main drawback of these models is that they have many parameters which all need to be calibrated. The calibration of all of these parameters is frequently not possible, leading to poor performance in the yield assessment when upscaling. Another drawback of these models is their extensive data input requirements, including data on cultivar, management, and soil conditions. This data is mostly limited to the study regions and is unavailable in the majority of the world. A soil-vegetation-atmosphere model is one model that combines the biological and physiological knowledge of plants and the interactions between plants and their environments [14]. Remotely sensed canopy state variables have been assimilated into these soil-vegetation-atmosphere models. There are a number of shortcomings of these models, such as the model input requirements that limit the use of cropping system models and practical problems including cloud cover and satellite revisit time [8]. Crop models require many site-specific inputs for detailed calibration in order to be able to perform realistic simulations. Additionally, optical remote sensing observations could be restricted by cloud cover throughout a growing season, particularly when combined with the revisit time of the satellite and when frequent imaging is required.

Remote Sensing and Agricultural Monitoring

Remote sensing provides objective observations over both time and space by various instruments observing the Earth. One major group of instruments is optical sensors. These record solar energy reflected back to space across various wavelengths including visible light and invisible infrared bands [6]. The number of bands varies depending on the instrument, and this variation defines the spectral resolution of the sensor. The higher the spectral resolution, the more accurate the characterization of different materials is [15]. Another type of instrument resolution is spatial resolution, which is the minimum size of detail observable – a pixel – in an image, depending on the orbit and characteristics of the instrument [16]. A third distinction can be made for temporal resolution, which is the amount of time needed to revisit and acquire data for the exact same location depending on the orbital characteristics of the sensor platform and its sensor characteristics [17]. Radiometric resolution is the fourth resolution type. The sensitivity of the image acquisition of a sensor to the magnitude of the electronic energy determines the radiometric resolution [18]. Finer radiometric resolution is more sensitive for the detection of small differences in reflected or emitted energy [18]. Each application using remote sensing data searches for the best trade-off between spectral, spatial, temporal, and radiometric resolution. For instance, in the context of agricultural monitoring with remote sensing, the spatial resolution must be high enough to monitor an individual crop, the temporal resolution must be high enough to collect crop-specific information throughout the growing season, and the differences in agricultural land cover must be distinguishable by their responses over distinct wavelength ranges.

As a result of technological developments, the definition of high and low spatial resolution has changed. Traditionally, a pixel size of 10 m to 30 m has been referred to as high spatial resolution. The multispectral scanner, thematic mapper, and thematic mapper plus instruments on board the Landsat satellites and the high resolution visible, visible and infrared high resolution, and high resolution geometric instruments on board the Satellites Pour l'Observation de la Terre (SPOT) satellite are some examples. Low spatial resolution instruments have pixels of about 1 km, such as the Advanced Very High Resolution Radiometer (AVHRR) on-board the national oceanic and atmospheric administration satellite. Moderate Resolution Imaging Spectroradiometer (MODIS), Project for On-Board Autonomy-Vegetation (PROBA-V), and Sentinel-3, with a pixel resolution from 100 m to 300 m, were previously considered to be medium resolution satellites, but now they are considered to have a low resolution. Today, satellites such as Sentinel-2 with 10 m spatial resolution are considered to be medium resolution. There is a new generation of instruments providing images below 5 m, such as FORMOSAT-2, WorldView-2, and QuickBird, which are considered to be high-resolution satellites. The spatial and temporal resolutions of satellites were inversely proportional until early 2000s [8]. Images from low spatial resolution sensors are mostly available on a daily basis, such as from AVHRR, MODIS, PROBA-V, and Sentinel-3. The revisit time for medium resolution sensors varies. For instance, the revisit period of Landsat7/ETM+ is 16 days, but Sentinel-2 images are available within 3 to 5 days, depending on the latitude. Currently, some high spatial resolution instruments, such as RapidEye and FORMOSAT-2, have daily revisit capabilities over a limited geographic extent. Besides the technological developments to provide better temporal and spatial resolution, the availability of historical time records for these satellites is also valuable, and these records are currently available for low and medium resolution satellites. Also, to assure the retrieval of reliable crop-specific information for agricultural monitoring with remote sensing at regional to global scales, it is necessary to have high temporal resolution of the satellite for a fine diagnostic of crop growth and sufficiently high spatial resolution for monitoring the crop fields [5].

Green plants use visible wavelengths as a source of energy in the process of photosynthesis. Conversely, pigments in plant leaves reflect electromagnetic radiation with wavelengths between 700 to 1000 nm, near-infrared light. Information on plant canopies can be retrieved from the electromagnetic signal measured by satellite remote sensing due to the diverse spectral properties of green vegetation [5]. For vegetation monitoring by remote sensing, there has been a tradition of using vegetation indices (VIs), which are a combination of spectral bands to highlight the spectral properties of green plants in relation to vegetation status, such as leaf water content, leaf density, and distribution. Examples of VIs include the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). An optimal VI is sensitive to the desired information and is as insensitive as possible to atmospheric and illumination conditions, soil properties, and the viewing geometry of the imaging instrument [5]. While such VIs at a global scale have limits in the agricultural monitoring context, VIs have been used as a proxy for general crop growing conditions and linked to crop yields in a statistical way [5,12,19–23]. Although VIs can contain spectral information relating to vegetation behavior, they cannot directly describe the functioning of a canopy as biophysical variables can. Examples of biophysical variables involved in important physical and physiological processes retrieved from remote sensing are the fraction of Absorbed Photosynthetically Active Radiation (fAPAR) and the Leaf Area Index (LAI).

Other biophysical variables that have originated from Monteith's theory and measure vegetation productivity are the Gross Primary Productivity (GPP), Net Primary Productivity (NPP), and Dry Matter Productivity (DMP). According to the radiation use efficiency model originating from Monteith's theory, carbon flux is a function of the photosynthetically active radiation absorbed by green vegetation, and the efficiency of light absorption is used as the reference for carbon fixation [69]. Gross primary productivity is the rate at which plants capture and store atmospheric carbon dioxide (CO_2) to generate oxygen and energy as biomass [24]. Terrestrial GPP constitutes the largest flux component in the global carbon budget, however, significant uncertainties remain in GPP estimates and its seasonality [25]. Recent methods of carbon flux assessment integrate FLUXNET site-level observations, satellite remote sensing, and meteorological data using upscaling approaches based on machine learning methods [26–28]. Furthermore, several studies reveal that solar-induced chlorophyll fluorescence from spaceborne observations exhibits a strong correlation with GPP, including cropland ecosystems even at a higher spatial resolution [25,29,30]. Wang et al. [31] proposed a simple equation to generate functional relationships between the ratio of leaf-internal to ambient carbon dioxide partial pressure and growth temperature, vapor pressure deficit, and elevation. This equation is a single global equation embodying these relationships and unifying the empirical radiation use efficiency model with the standard model of C_3 photosynthesis [31]. This proposed model is used in the TerrA-P project to estimate GPP and aboveground biomass production from MEdium Resolution Imaging Spectrometer (MERIS) and Sentinel-3 data [32].

The difference between GPP and the energy lost during plant autotrophic respiration is the NPP. Furthermore, DMP is analogous to NPP but expressed in different units – kgDM/ha/day instead of $gC/m^2/day$ – for agro-statistical purposes [33]. For this reason, the DMP is a particularly interesting variable because it estimates photosynthetic carbon uptake and represents the daily accumulation of standing biomass. Currently, the Copernicus Global Land Service-Dry Matter Productivity (CDLS-DMP) is available as a part of the operational processing chain of SPOT-VGT and PROBA-V at the Flemish Institute for Technological Research, also known as the Vlaamse Instelling Voor Technologisch Onderzoek. Although studies have proven the usefulness of DMP in grasslands and forests [34–36], there are few studies related to DMP in crop yields [37,38]. Additionally, there are several areas identified for improvement in DMP modeling, such as introducing a

water stress index and re-parameterization of stress factors. Hereafter, VIs and biophysical variables are called biomass proxies (BPs).

Biomass proxies should be combined with crop maps for better crop yield estimates. Remote sensing data are a promising source of information for deriving crop maps at the regional to continental scale due to the synoptic capabilities of spaceborne instruments to acquire timely images in different spectral bands and provide repeatable, continuous, human-independent measurements for large territories [39]. Land use and land cover datasets, which have been the primary input in estimating crop yield, provide variable quality and reliability in representing croplands [40]. Regardless of crop type, cropland maps from previously mentioned studies have been proven to improve crop yield forecasting [41]. Therefore, timely crop-specific maps could potentially further enhance crop yield estimates. Cropland areas are intensively managed and modified through a variety of human activities; therefore, timely crop mapping to characterize the cropping patterns on a repetitive basis is valuable to scientists and policy-makers [42]. Although, it is challenging, discriminating croplands from non-croplands and identifying different crop types can be achieved with remote sensing-based crop growth monitoring and, in particular, with indices that quantify the distinct green-up and senescence stages of the crop cycle [43]. Furthermore, multiband pixel and object-based image analysis at a chosen time during the crop season or at different times during the season have also proven their high efficiency in crop identification [44,45].

PROBA-V Low-Resolution Satellite

New technological improvements have supported the development of satellites that could solve the problems mentioned in the previous section. Regarding crop monitoring, these improvements provide remotely sensed data with larger geographic coverage, higher temporal resolution, adequate spatial resolution relative to the typical field size, and minimal cost [42].

PROBA-V is a small satellite constructed to be a successor of the VEGETATION instruments onboard the French SPOT-4 and SPOT-5 Earth observation missions. It was initiated by the Belgian Science Policy Office and has been operated by the European Space Agency. It was designed to support land use, vegetation classification, crop monitoring, famine prediction, food security, climate, and coastal applications. The original mission was to fill the gap between Satellites Pour l'Observation de la Terre or Earth-observing Satellites -VeGeTation (SPOT-VGT), which reached the end of its life in May 2014, and Sentinel-3 satellites, which were launched in February 2016 and April 2018. PROBA-V was launched in May 2013 with a 5-year expected operational lifetime and has four spectral bands: blue, centered at 0.463 µm; red, 0.655 µm; Near InfraRed (NIR), 0.845 µm; and Short-Wavelength InfraRed (SWIR), 1.600 µm. The spectral characteristics of the instrument are nearly identical to the SPOT-VGT instrument. The Vegetation-PROBA instrument has a swath width of 2250 km across four bands. Compared to SPOT-VGT, which provided daily global images with spatial resolution at 1 km, the central camera of the PROBA-V satellite provides a 100 m data product when delivering global coverage every 5 days, whereas global daily images are acquired at 300 m and 1 km resolutions. Radiometric performances in terms of signal-to-noise ratio are also increased with respect to SPOT-VGT [46]. The PROBA-V data is available to the CGLS (formerly "Global Monitoring for Environment and Security") users through https://land.copernicus.eu/global/themes/vegetation.

Crop growth monitoring and yield forecasting using space observations often rely on the analysis of long-term data archives in order to identify anomalies against a long-term average that reflects adverse crop growth conditions or to define local empirical regression models between remote sensing products and observed yield [46]. Therefore, PROBA-V is particularly important to ensure the continuity of the operationally used biophysical

products in such an analysis. Additionally, its particular configuration providing a 100 m spatial resolution combined with 5-day repetition could be very useful for agricultural monitoring, particularly for monitoring the continuity of crop yield estimation studies and crop identification. Over, the unique characteristics of PROBA-V open a new set of opportunities for scientists to explore various Earth surface phenomena at finer spatial and temporal resolutions [47].

Satellite Combinations for Operational Continuity and Better Crop Monitoring

PROBA-V images are interesting when viewed alone, but they are even more interesting when combined with information from other satellites to either complete time series or to study areas with different pixel sizes. Currently, the Sentinel-2 – launched in 2015 and available at 10 m and 20 m resolutions with 3 to 5 days of revisit time and Sentinel-3 – launched in 2016 and available at 300 m resolution with 1 to 2 days of revisit time – satellites can both be used for crop monitoring. Smets et al. [48] assumed that regular monitoring of the globe would continue at 300 m and 100 m resolutions, and hot spot monitoring at 30 m and 10 m resolutions. For instance, Sentinel-2 with 10 m spatial resolution and 5-day temporal resolution could be used for emergency management and security for the present time until the historical archive is ready to conduct global operational crop monitoring. Sentinel-2 for crop yield forecasting may also become rapidly interesting for small-size plots.

Use of Sentinel-2 data for global crop monitoring is questionable and subject to discussion at the present time. In general, Sentinel-2 provides much greater data volume than is necessary for yield estimation studies. With the current technology, in operational conditions, greater data volume would usually take excessive space and processing time. Traditionally, time series of the satellite data have been used in crop monitoring studies. The factors impacting the yield vary greatly over space and time. Therefore, the historical archive of low spatial resolution satellite images may be in use, similar to SPOT-VGT type images. In the future in 10 to 15 years, when the archive has been built for Sentinel-2 images, and when the processing time and storage problems have been overcome by the technological improvements, the Sentinel-2 archive will probably be used in global crop monitoring studies. However, until then, it is important to explore how much low-resolution images, such as those obtained by PROBA-V, can contribute to global crop monitoring.

Additionally, there has been a need for temporal continuity assessment between sensors, SPOT-VGT, PROBA-V, and Sentinel-3. For instance, the spectral continuity between the global daily datasets of PROBA-V and SPOT-VGT has already been validated [49].

The technological developments can ensure free images with better spatial and temporal resolution. Thus, the popularity of using high spatial resolution data will probably increase in different vegetation monitoring studies. However, since low-resolution datasets will still have a longer historical time series, they may be used for crop yield estimation until the high-resolution datasets have established a longer time series and have proven that they are more efficient.

Scope and Objectives

According to the World Bank's global strategy to improve agricultural statistics, crop area, crop production, and crop yield are key variables that all countries should be able to provide [50]. The present research focuses on the crop yield. Estimation of crop yield at a global scale through surveys requires measurements that are time consuming and are subject to error in many circumstances [51]. Remote sensing is one of several methods

used for crop yield estimation. The yield is produced by a combination of environmental factors, such as soil, weather, and farm management. Green plants have a unique spectral signature that is captured in satellite images and is defined by the state, structure, and composition of the plant. There are several limitations of the use of satellite images for crop yield estimations.

First, an appropriate BP derived from remote sensing can better explain variation in irradiation conditions, short-term environmental stresses, and respiration costs compared to other simpler satellite-derived BPs such as NDVI and fAPAR. The improved CGLS-DMP model could be a better option to use in crop yield estimation studies because it explains the variation mentioned previously and represents the standing biomass computed based on Monteith's approach. According to Monteith's theory, crop biomass is linearly correlated with the amount of absorbed photosynthetically active radiation (APAR) and constant radiation use efficiency (RUE), downregulated by stress factors such as CO_2 fertilization, temperature, and water stress. The current CGLS-DMP model is open for improvements to its parameters to be used in agricultural applications, such as stress factors.

Second, the spatial resolution of available satellite imagery should allow the crop identification in fields to be defined. As a result of satellites, such as PROBA-V and Sentinel-2, the spatial resolution of satellite imagery has been improved, and the usefulness of these images can be studied in terms of crop identification improvement. Furthermore, better spatial resolution could improve the crop yield estimation results by avoiding, or at least reducing, the problem of mixed pixels. Traditionally, images used for global and regional crop production studies usually have 1 km spatial resolution. PROBA-V with 100 m, 300 m, and 1 km resolution datasets could be further studied to determine the impact of three different resolutions in improving crop yield estimation.

Third, the temporal resolution of satellite imagery is particularly important for crop monitoring, because vegetation status changes rapidly during the growing season. Therefore, frequent cloud cover during this season can be a major problem for crop monitoring. In most cases, when the spatial resolution is increased, the temporal resolution is decreased because of the trade-off between the spatial and temporal characteristics of the satellite. For this reason, satellite images with better spatial resolution should be studied for their impact on crop yield estimation studies regarding their temporal resolution. For instance, PROBA-V provides a 100 m data product with a 5-day revisit time and daily images at 300 m and 1 km resolutions. The greater difficulty in cleaning 100 m products from cloud cover compared to the 300 m and 1 km products could affect the crop yield estimation results.

This thesis aims to explore how much PROBA-V-like imagery, such as low spatial resolution data, can contribute to global crop monitoring through possible improvements in crop yield estimation. In this context, the main objective of this thesis is to advance agricultural monitoring for improved yield estimation using SPOT-VGT and PROBA-V-type remote sensing data. More specifically, three research questions are listed below.

- *Is a modified DMP a more accurate proxy for crop yield estimations compared to CGLS-DMP, NDVI, and fAPAR?*
- *Is it possible to retrieve crop identification with a 100 m PROBA-V product that could possibly be used in crop yield estimations?*
- *What is the optimal spatial resolution between 100 m, 300 m, and 1 km PROBA-V products to estimate crop yield at field scale?*

Although the research questions are handled individually, they can be understood in combination with each other. Due to the ground data and satellite data availability per topic – for instance, the thesis was started in 2011, but the PROBA-V dataset was only available in late 2013 – different datasets are used in each chapter while investigating the questions.

Outline of This Thesis

The chapters are based on scientific papers published in peer-reviewed journals and a manuscript currently in progress to be submitted. Each chapter introduced below states its research purposes. The main findings and future perspectives are given in the general conclusion and perspectives. Additionally, each chapter can be considered as addressing an independent research question.

Chapter 1 explores how an existing product, CGLS-DMP, that has been developed for SPOT-VGT, and transferred to PROBA-V, can be improved to better relate to yield anomalies over selected regions. This chapter investigates the relative importance of stress factors in relation to regional biomass production and yield. The stress factors are CO_2 fertilization, temperature, and water stress, which could downregulate APAR and RUE according to Monteith's theory. The CGLS-DMP, which follows Monteith's theory, is modified and evaluated for common wheat and silage maize. The results reveal that each study region has its own climatic characteristics that affect the impact of these stress factors.

Chapter 2 demonstrates how PROBA-V can be used effectively for crop identification mapping using spectral matching techniques (SMTs). This chapter provides a method developed for crop identification inspired by SMTs and based on the phenological characteristics of different crop types. For each pure pixel within the field, the 100 m PROBA-V NDVI profile of the crop type for its growing season was matched with the reference NDVI profile based on the training set extracted from the study site where the crop type originated. The resulting maps can be used as crop-specific maps in crop yield estimation studies.

Chapter 3 explores the trade-off between the different spatial resolutions provided by PROBA-V products versus the temporal frequency, and, additionally, this chapter explores the use of thermal time to improve statistical yield estimations. This chapter compares the suitability of three different spatial resolutions – 1 km, 300 m, 100 m – of PROBA-V NDVI datasets for wheat yield forecasting. The NDVI-integrated values as a proxy for yield are computed using daily NDVI and cumulative growing degree day images over thermal time and over calendar time. The results suggest that the greater the spatial resolution, the greater the correlation between the estimated and real yields.

1 Testing the Contribution of Stress Factors to Improve Wheat and Maize Yield Estimations Derived from Remotely-Sensed Dry Matter Productivity[1]

Abstract

According to Monteith's theory, crop biomass is linearly correlated with the amount of absorbed photosynthetically active radiation (APAR) and a constant radiation use efficiency (RUE) down-regulated by stress factors such as CO_2 fertilization, temperature and water stress. The objective is to investigate the relative importance of these stress factors in relation to regional biomass production and yield. The production efficiency model Copernicus Global Land Service-Dry Matter Productivity (CGLS-DMP), which follows Monteith's theory, is modified and evaluated for common wheat and silage maize in France, Belgium and Morocco using SPOT-VGT for the period 1999–2012. For each study site the stress factor that has the highest correlation with crop yield is retained. The correlation between crop yield data and cumulative modified DMP, CGLS-DMP, fAPAR (Fraction of Absorbed Photosynthetically Active Radiation), and NDVI (Normalized Difference Vegetation Index) values are analyzed for different crop growth stages. A leave-one-year-out cross validation is used to test the robustness of the model. On average, R^2 values increase from 0.49 for CGLS-DMP to 0.68 for modified DMP, RMSE (t/ha) values decrease from 0.84–0.61, RRMSE (%) values reduce from 13.1–8.9, MBE (t/ha) values decrease from 0.05–0.03 and the index of model performance (E1) increases from 0.08–0.28 for the selected sites and crops. The best results are obtained when combinations of the most appropriate stress factors are included for each selected region and the modified DMP during part of the growing season that includes the reproductive stage are cumulated. Though no single solution to an improvement of a global product can be demonstrated, the findings support an extension of the methodology to other regions of the world.

1.1. Introduction

Regional to global scale crop monitoring and yield forecasting are important for agricultural management and food security [52,53]. Satellite remote sensing enables assessment of agricultural crop growth and yield across large territories [12,54–57]. Various Biophysical Proxies (hereafter BPs) have been developed from remote sensing

[1] Adapted from Durgun, Y.Ö.; Gobin, A.; Gilliams, S.; Duveiller, G.; Tychon, B. Testing the Contribution of Stress Factors to Improve Wheat and Maize Yield Estimations Derived from Remotely-Sensed Dry Matter Productivity. *Remote Sens.* **2016**, *8*, 170.

imagery to be used in empirical regressive models that monitor agricultural crop growth and estimate crop yield [58–60]. NDVI, fAPAR, LAI, GAI and EVI (Enhanced Vegetation Index) are examples of BPs that have been derived from remote sensing and that are used in vegetation monitoring and crop yield forecasting [55,57,61–66]. Though relationships between BPs and yield have been established, few studies relate Dry Matter Productivity (DMP) to crop yield [37,38].

Vegetation productivity can be defined in several ways. Gross primary productivity (GPP) is the rate at which plants capture and store atmospheric carbon dioxide (CO_2) to generate oxygen and energy as biomass [24]. Net primary productivity (NPP) is the difference between GPP and the energy lost during plant autotrophic respiration. NPP is thus the rate of atmospheric carbon uptake through the process of photosynthesis and represents the daily accumulation of standing biomass. DMP is analogous to NPP, but expressed in different units (kgDM/ha/day instead of $gC/m^2/day$), for agro-statistical purposes [33]. The efficiency of the conversion between carbon and dry matter is on average 0.45 gC/gDM [67].

Three types of models are used to estimate the photosynthetic carbon uptake and understand its spatio-temporal variability [68–70]. The first group consists of empirical models often with a limited applicability outside the area where they have been calibrated. The second group of models is based on major biophysical and biochemical processes of photosynthesis and respiration measured under laboratory conditions [71]. These models have a high computational demand. The third group of models are parametric models driven by remote-sensing-derived variables and weather data, calibrated with data derived from flux measurement sites. Constant parameters are used to link measurements to biophysical processes rendering these models particularly suitable for the local scale. While such assumptions may be difficult to hold at a global scale, parametric models may offer a balance between simplicity and process description [69].

Production efficiency models, such as Monteith parametric models, have been developed to monitor the primary production of vegetation [72,73]. Monteith suggested that crop growth under non-stressed conditions linearly correlates with their Radiation Use Efficiency (RUE) times the amount of Absorbed Photosynthetically Active Radiation (APAR) [74,75]. According to a review of experimental studies, RUE values for C_3 species range from 1.32–3.50 gDM/MJ of intercepted PAR and for C_4 species from 1.48 to 4.32 gDM/MJ of intercepted PAR during different crop development stages [76]. The seasonal variability of photosynthetic activity, however, depends on environmental constraints [68]. For example, RUE is negatively related to water stress and positively related to temperature for annual crops [68]. Water stress has been estimated as a function of soil moisture [77,78], water deficits [79] or satellite-derived land surface water index [80].

Production efficiency models have been widely used to estimate regional or global carbon balances in crops, grasslands and forests due to the simplicity of the RUE concept and the availability of remotely sensed data [37,77,78,81–87]. However, their performance has been shown to vary in describing the carbon budget [88]. A comparison of modelled gross and net primary productivity [72] of six different models, CASA [89], GLO-PEM [90], TURC [91], MOD17 [92], BEAMS [93] and C-Fix [94], illustrates how the various methodologies used to calculate vegetation productivity can be generalized in the following two equations:

$$GPP = PAR \times fAPAR \times RUE_{MAX} \times Stress\ factors \quad (1.1)$$

$$NPP = GPP - AR \quad (1.2)$$

where PAR is Photosynthetically Active Radiation (MJ/m^2), fAPAR is the Fraction of Absorbed PAR (dimensionless), RUE_{MAX} is the maximum Radiation Use Efficiency (gC/MJ) (*i.e.* RUE under no stress), which is downregulated by a coefficient that encompasses the effects of all stress factors such as temperature stress or water stress, and AR is Autotrophic Respiration (gC/m^2/day). In general, variation among the models is caused by the differences in RUE, the incorporation of stress factors, and AR [72]. The initial model conditions, model parameters, model structures and accuracy of input data also play an important role [88,95]. Uncertainties in global GPP/NPP monitoring also relate to the determination of RUE, AR and the quality of meteorological and the biophysical data [72,92,96,97].

In this study, the Copernicus Global Land Service-Dry Matter Productivity (CGLS-DMP) product is analyzed. This model is part of the operational processing chain of SPOT-VGT and the following PROBA-V at the Vlaamse Instelling Voor Technologisch Onderzoek. The parametric CGLS-DMP model estimates carbon mass fluxes at local, regional and global scales [94], and has proved usefulness in vegetation productivity studies such as grasslands and forests [34–36]. For a given location, carbon fluxes are estimated on a decadal basis using SPOT-VGT and resulting in a DMP product. Different DMP versions have evolved since the early 1990s [98]. The C-Fix approach (a variant of Monteith's approach) was applied, to monitor the overall carbon balance over Europe in terms of NPP [83]. In the late 1990s, an operational chain was established to retrieve dekadal DMP images over Belgium based on C-Fix using SPOT-VGT for crop monitoring. Around 2000, a consortium was established for Joint Research Centre - Monitoring Agricultural ResourceS (JRC-MARS). There have been different Monitoring Agricultural ResourceS OPerational (MARSOP) projects in operation (MARSOP1 from 2001–2003, MARSOP2 from 2004–2007, MARSOP3 from 2008–2014, MARSOP4 from 2015–2019). The DMP coverage has been global except for MARSOP1 which was the Mediterranean Basin, part of the Commonwealth of Independent States (former Soviet Union), Mercosur in South-America, and the Horn of Africa. The procedure for MARSOP1 was the same as C-Fix. The RUE in optimal conditions has been taken as 1.10 kg·C/GJ for NPP and 2.45 kg·DM/GJ for DMP. It is considered constant for all land cover types, while differences exist between biomes, causing the operational DMP product to over/underestimate reality [68,99]. The RUE value is subsequently reduced by normalized temperature and CO_2 fertilization dependency factors. With the start of MARSOP2, the temperature errors were corrected using day time temperature and CO_2 fertilization efficiencies, and day and night temperatures for autotrophic respiration instead of simply using mean temperature (Equations (1.3) and (1.4)).

$$T_{12} = 0.75 * (T_{min} + T_{max}) \quad (1.3)$$

$$T_{24} = 0.50 * (T_{min} + T_{max}) \quad (1.4)$$

where T_{12} stands for the day time temperature, T_{24} stands for the day and night temperatures, T_{min} is the minimum daily temperature and T_{max} is the maximum daily temperature. The grid of the meteorological data has been changed from 1° to 0.25° with MARSOP3. For all projects, while calculating CO_2 fertilization efficiency, the CO_2 value used was fixed at 355.61ppmv, the global mean level of the year 1994. For MARSOP1 and 2, in

order to estimate fAPAR, the linear relationship between NDVI and fAPAR was used. For MARSOP3 and 4, the global fAPAR is computed based on Swinnen et al. [33]. The DMP has been produced in the CGLS since 2014 using the same meteo data, algorithm and constants as the MARSOP4 DMP [100]. As mentioned in Smets et al. [100], some remarks should be taken into consideration when interpreting the product. Firstly, since no direct water stress factor is implemented in the DMP algorithm, the retrieved values should be considered as optimal values. The CGLS-DMP model is only partially water limited through the sensitivity of fAPAR to vegetation water stress [83]. For the C-Fix model, Verstraeten et al. [94] therefore proposed to limit RUE by using estimates of soil moisture content and water vapor deficit which required several empirical coefficients tuned to local vegetation conditions [77]. Maselli et al. [36] successfully introduced a water stress index to the C-Fix model, which in turn enabled simulations of the gross and net carbon fluxes of Mediterranean forest ecosystems. Secondly, although RUE values differ between biome types, it is considered as a global constant for all land cover types. Thirdly, the temperature stress factor was parameterized for European forests, and reflects neither the difference between C_3 and C_4 plants nor the differences within one plant type. Fourthly, although the CO_2 level increases each year, it is considered as a global constant. Finally, the AR is a linear function of daily mean temperature and is assumed biomass-independent [102]. The average value of the original AR fraction is around 0.7 for Europe and is an overestimation compared to other values found in the literature [52,88,103–105].

We hypothesize that the relative importance of the stress factors strongly contributes to explaining regional differences in biomass production and yield. An accordingly modified DMP product could therefore serve as a better proxy for crop yield than the currently available CGLS-DMP, and perhaps also outperform other simpler satellite-derived biophysical proxies such as fAPAR and NDVI. The present study aims to parameterize the dry matter productivity approach specifically for wheat and maize and to compare this modified DMP model with the original CGLS-DMP and with other BPs derived from satellite imagery. The value of this improvement is evaluated in the context of arable productivity at the regional administrative or agro-ecological level. The model performance is assessed using statistical metrics based on the comparison between regional cumulated BPs and regional crop yield statistics of silage maize and common wheat for selected sites in Belgium, France and Morocco during the period 1999–2012.

1.2. Materials

1.2.1. Study Areas and Crops

The selected study sites are located in Belgium, France and Morocco (Figure 1.1). Silage maize and common wheat are the most dominant crops in the selected sites.

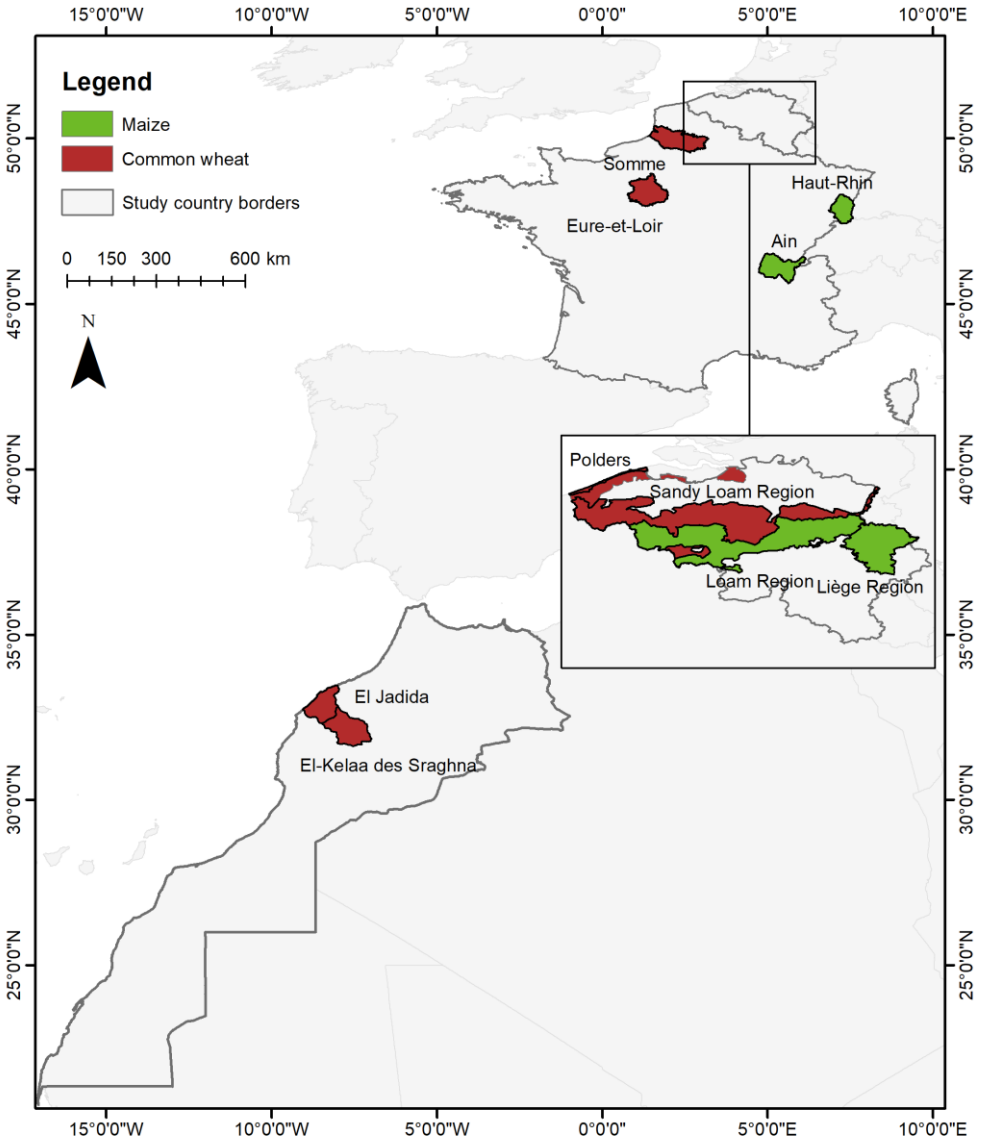


Figure 1.1. Study areas with study crops in Belgium, France and Morocco.

From Northern Europe to Northern Africa, the average daily temperature and potential evapotranspiration increases, and the average cumulative rainfall decreases. The regions were selected to capture this trend in climate regimes. Figure 1.2 shows the long term averages of temperature, precipitation, Penman-Monteith potential evapotranspiration (PET) and water balance computed per region based on meteorological data obtained from the JRC-MARSOP project [106]. The water balance was calculated using the Thornthwaite-Mather method as precipitation minus PET [107]. Water and heat stress are the major factors that influence arable yields in Belgium and France [108,109]. The total annual water used for irrigation across France can amount to 80% of the crop water use during dry periods, while on average 10%–20% is used during a typical growing season [110]. In Morocco, most arable crops are rain fed, having frequent dry periods with high temperatures and irregular rainfall [111].

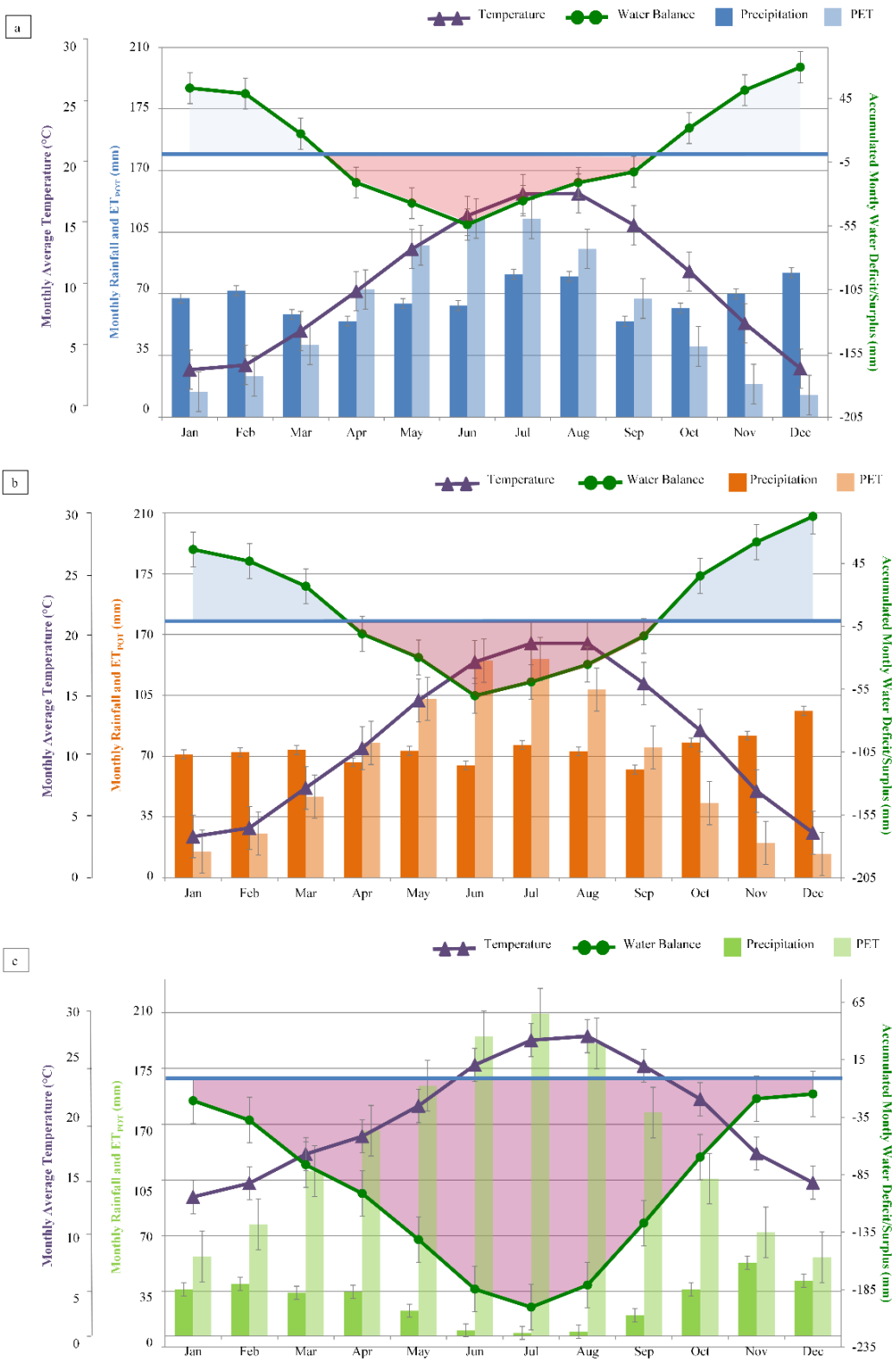


Figure 1.2. Cumulative monthly rainfall, PET (potential evapotranspiration), water balance (accumulated monthly water deficit/surplus) (mm) and average monthly temperature (°C) during the 1999–2012 period for study sites in Belgium (**a**), France (**b**) and Morocco (**c**). Error bars show the standard deviation for the meteorological indicators.

1.2.2. Data Description

Daily meteorological data were obtained from JRC-MARSOP [106] and 10-daily fAPAR, NDVI and CGLS-DMP from SPOT-VGT (1 km) [112] for the period 10/1998–10/2012. Temperature and solar radiation are at a 25 km grid, while rainfall, PET and AET are at a 0.25° grid.

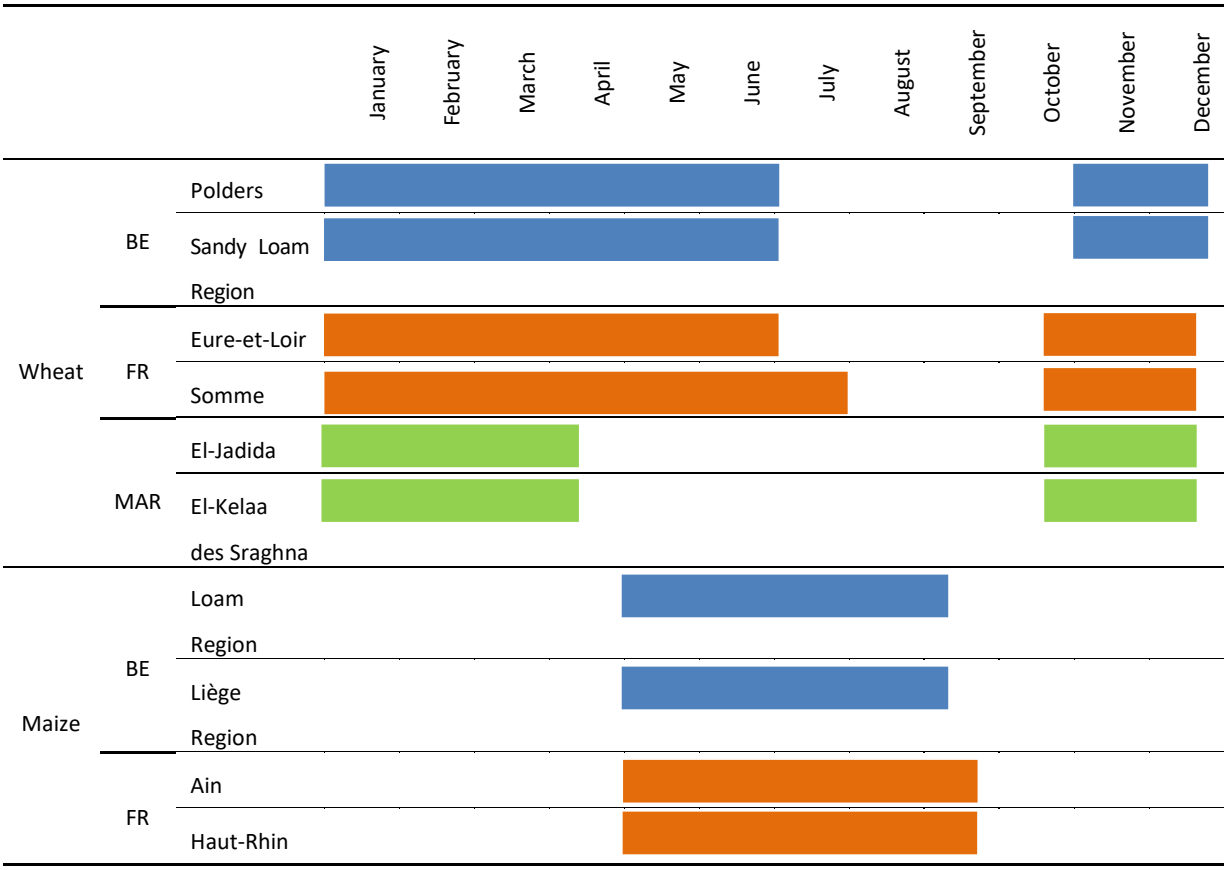
A cropland mask was retrieved from ESA GlobCover 2009 [113]. The classes used were "Post-flooding or irrigated croplands (or aquatic)", "Rainfed croplands", "Mosaic cropland (50%–70%)/vegetation (grassland/shrubland/forest) (20%–50%)" and "Mosaic vegetation (grassland/shrubland/forest) (50%–70%)/cropland (20%–50%)". The soil water holding capacity was retrieved from the European soil map [114] and Digital Soil Map of the World [115] and used in AgroMetShell [116]. AgroMetShell provides a toolbox for agrometeorological crop monitoring and forecasting developed by the FAO. This toolbox includes a database with the weather, climate and crop data used to analyze the impact of weather on crops [117]. AgroMetShell also contains crop-specific water balance parameters for different crop growth stages. In this study, using the meteorological from JRC-MARSOP, AgroMetShell was utilized to compute the actual evapotranspiration (AET) for every season during the period 1999–2012. The soil types and textures of the different regions show a high suitability for arable agriculture due to the presence of silt (Table 1.1).

Table 1.1. The soil types and textures of the study sites.

			Soil Type	**Soil Texture**
Wheat	BE	Polders	Calcaric Regosols + Calcaric Fluvisols	Loam + Silt loam
		Sandy Loam Region	Dystric Podzoluvisols + Orthic Luvisols	Sandy loam + Loam
	FR	Eure-et-Loir	Gleyic Luvisols + Orthic Luvisols	Clay Loam + Loam
		Somme	Orthic Luvisols	Loam
	MAR	El Jadida	Calcic Kastanozems + Vertisols + Eutric Fluvisols	Loam + Silty Clay + Silt Loam
		El Kelaa	Eutric Gleysols + Calcic Xerosols	Clay Loam + Loam
Maize	BE	Loam Region	Orthic Luvisol + Dystric Podzoluvisol	Loam + Sandy Loam
		Liège Region	Stagno-Gleyic Luvisol + Orthic Luvisols + Dystric Cambisol	Clay Loam + Loam + Silt Loam
	FR	Ain	Gleyic Luvisols + Orthic Luvisols + Eutric Cambisols	Clay Loam + Loam + Silt Loam
		Haut Rhin	Gleyic Luvisols + Orthic Luvisols	Clay Loam + Loam

The information for the planting dekad (period of 10 days) and the length of the growing season as required in AgroMetShell were extracted from the MARS database [106] for Belgium and France; and obtained from Balaghi et al. [111] for Morocco (Table 1.2).

Table 1.2. Crop growing periods for maize and common wheat in the case study sites.



Official crop yield statistics for common wheat and silage maize for the period 1999–2012 were obtained from the national statistical services. The coefficient of variation is 54% for Morocco, 9% for France and 7% for Belgium for wheat, and 16% for France and 7% for Belgium for maize.

Annual global MOD17A3 GPP and NPP data for 14 years (2000–2013) with a spatial resolution of 30-arcsec were downloaded from the Numerical Terradynamic Simulation Group at the University of Montana [118]. This dataset was used to compute NPP/GPP ratio which was compared with the autotrophic respiration fraction computed by the CGLS-DMP and the modified DMP.

1.3. Methods

1.3.1. Algorithm Description of the DMP Model

The CGLS-DMP model [83] uses the following equation for calculating Dry Matter Productivity:

$$DMP = R \times \varepsilon_p \times fAPAR \times \varepsilon_{RUE_{MAX}} \times \varepsilon_T \times \varepsilon_{CO2} \times \varepsilon_{AR} \quad (1.5)$$

where DMP is daily dry matter productivity (kgDM/ha/day). R is total shortwave incoming solar radiation (0.2–3.0 µm) (GJ/ha/day). ε_p is the ratio of effective photosynthetically active radiation (PAR) (0.4–0.7 µm) to the total incident radiation. Because the energy in the PAR band at the surface of the earth is approximately 48% of global radiation, a value of 0.48 has been used in the model [83]. fAPAR is the fraction of PAR absorbed by green vegetation and is estimated based on remote sensing [119]. $\varepsilon_{RUE_{MAX}}$ is the maximum

Radiation Use Efficiency (RUE). In order to make a distinction between C_3 and C_4 crops, values of 2.75 kgDM/GJ for wheat and 3.5 kgDM/GJ for maize are used [76]. ε_T is the normalized temperature effect [120] and indicates the role of air temperature in the photosynthesis efficiency [121]. ε_{CO2} is the normalized CO_2 fertilization effect and takes into account the thermodynamic properties of the carboxylation/oxygenation reactions during photosynthesis [83]. ε_{AR} is the fraction kept after Autotrophic Respiration (AR).

$$DMP = R \times \varepsilon_p \times fAPAR \times \varepsilon_{RUE_{MAX}} \times \varepsilon_T \times \varepsilon_{CO2} \times \varepsilon_{AR} \times [\varepsilon_{H2O}] \quad (1.6)$$

The reduction factor "ε_{H2O}" was introduced in Equation (1.6) to account for water stress. The details of the changed parameters are explained in the following sections.

1.3.1.1. Temperature Stress Factor

Plant species show characteristic variations in the way photosynthetic processes respond to temperature [122]. For instance, wheat, as a C_3 plant, has an optimum temperature around 20 °C and maize, as a C_4 plant, has an optimum temperature around 30 °C [123–128]. Different crop varieties may have different temperature responses to dry matter accumulation (e.g., [129]). The focus in this study is on crop type rather than variety. The temperature stress factor (ε_T) in CGLS-DMP is parameterized for C_3 plants (blue line in Figure 1.3) [83]. The temperature response function by Yan & Hunt [128] is used for maize (red line in Figure 1.3).

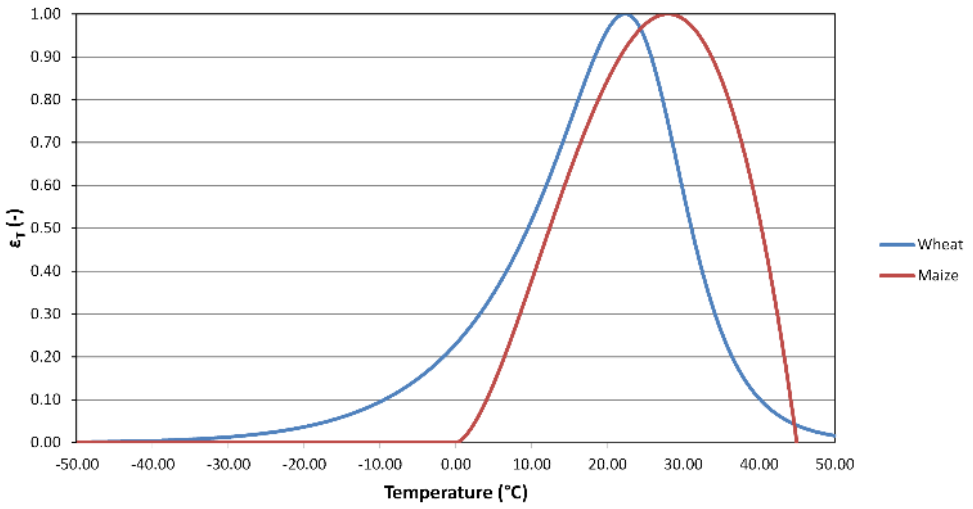


Figure 1.3. The temperature functions of ε_T used in this study for wheat and maize.

1.3.1.2. CO_2 Fertilization Effect

The CO_2 fertilization effect is the increase in carbon assimilation due to CO_2 concentrations above the atmospheric reference level [83]. The CO_2 concentration level in CGLS-DMP was fixed to a constant value, although the globally averaged records show a tendency to increase from 1980–2012 and beyond (Figure 1.4). In Figure 1.5a, the dots represent the CO_2 fertilization effect of CGDL-DMP at a fixed CO_2 level with variable temperature (from −20 °C–40 °C). In order to simulate the measured CO_2 increase, the yearly variable CO_2 values were used in the modified DMP.

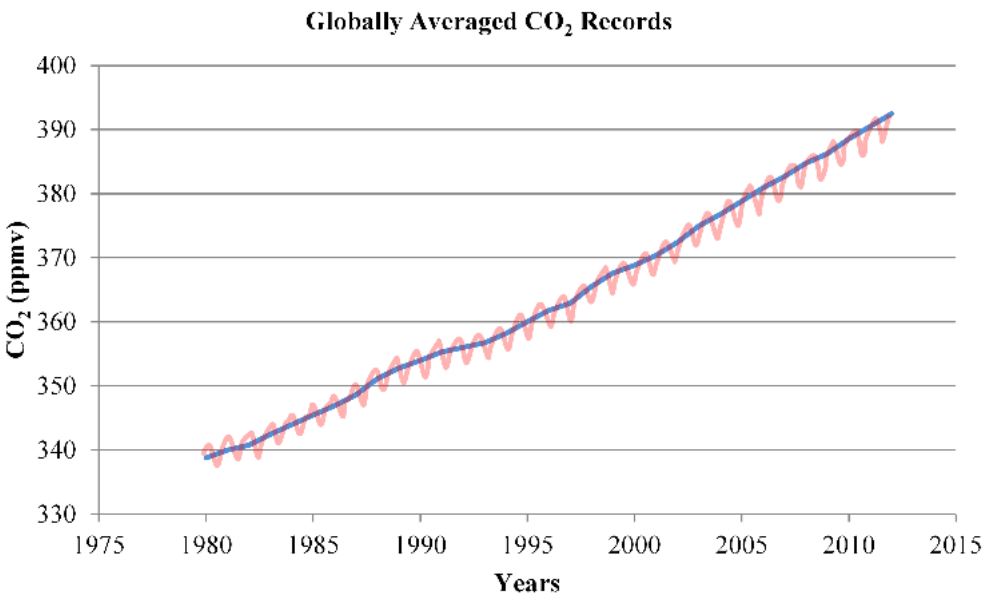


Figure 1.4. Yearly (blue line) and monthly (red line) globally averaged CO_2 records (source: [130]).

In the CGLS-DMP model, the CO_2 fertilization effect was parametrized for C_3 plants with a fixed CO_2 level to year 1994 (See Veroustraete et al. [83] for details on the equation). Above-ground dry matter increases with elevated CO_2 for C_3 plants [131–133]. The blue to red lines in Figure 1.5b show the increase of the CO_2 fertilization effect for C_3 plants from 1999–2012.

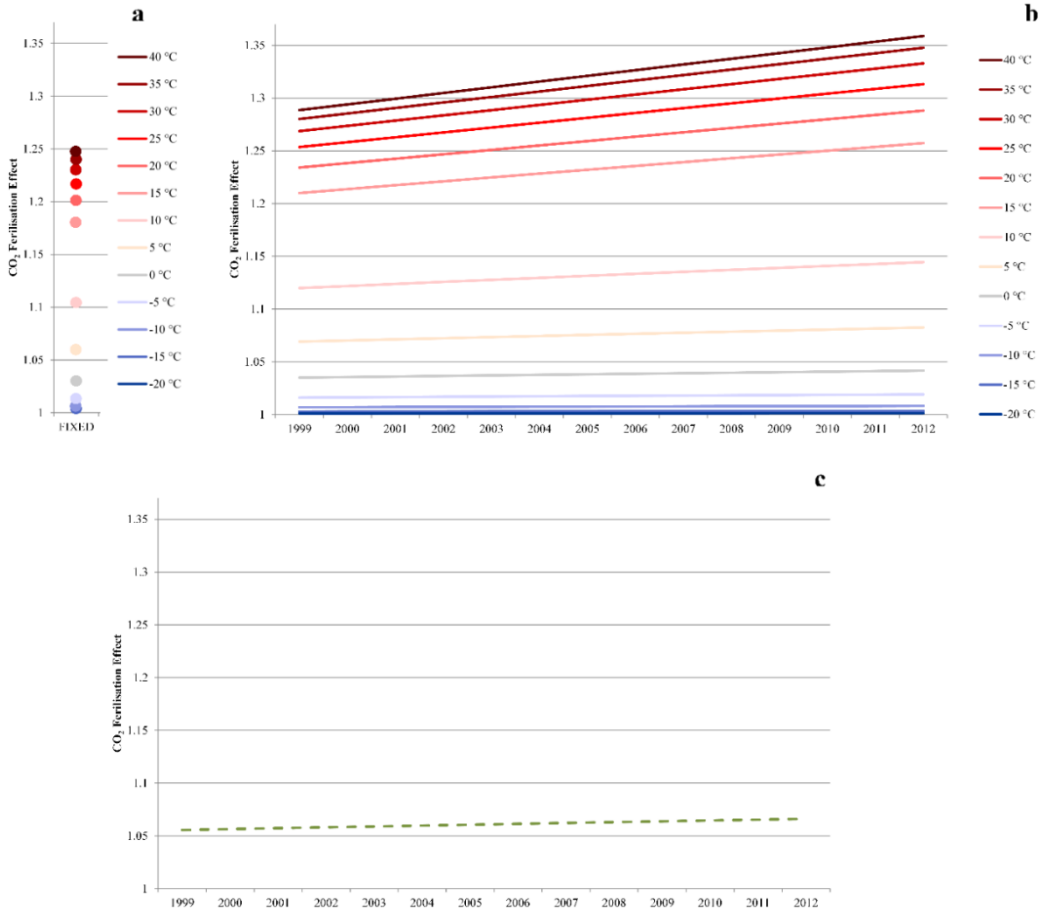


Figure 1.5. Evolution of ε_{CO2} both for CGLS-DMP (Copernicus Global Land Service-Dry Matter Productivity) at the fixed rate of CO_2 value (dots) (a); modified DMP variable CO_2 rates for C_3 (blue to red lines represent the change in temperature from −20 °C–40 °C) (b); and C_4 plants (dashed green line) (c); for the years 1999–2012.

For C_4 plants little evidence of biomass accumulation in response to CO_2 enrichment was observed over a wide range of temperatures [131,134]. Based on the CO_2 assimilation rate of C_4 at lower CO_2 partial pressures [135], the evolution of ε_{CO2} for C_4 plants can be rewritten (Equation (1.7)).

$$\varepsilon_{CO2} = \frac{F_{CO_2}}{F_{CO_2}^{ref}} = \frac{C_{CO_2} * V_{pmax}}{C_{CO_2} + K_p} * \frac{C_{CO_2}^{ref} + K_p}{C_{CO_2}^{ref} * V_{pmax}} \quad (1.7)$$

where F_{CO_2} is the CO_2 assimilation rate in year x, $F_{CO_2}^{ref}$ is the CO_2 assimilation rate in the reference year 1833, C_{CO_2} is the actual CO_2 concentration, V_{pmax} is the maximum phosphoenolpyruvate carboxylase activity, $C_{CO_2}^{ref}$ is the CO_2 concentration in the reference year 1833 (281 ppmv (or µbar)) and Kp is the Michaelis–Menten constant for CO_2 (80 µbar according to [135]). Since the variability in CO_2 assimilation rates of C_4 plants is dependent on CO_2 concentration, the changes from one year to another are very small (dotted green line in Figure 1.5c).

1.3.1.3. Water Stress Factor

The water stress factor (ε_{H2O}, Equation (1.6)(1.5)) is included in the algorithm to account for the immediate effect of water stress on the vegetation at each dekad. The model uses a simple water stress factor computed on the basis of PET and AET estimates [101]. AgroMetShell was used to compute AET from total evaporation, actual rainfall and crop coefficients between planting and harvesting dates [117]. The theoretical base of this index is the same as the water scalar in the CASA model [77,81]. According to the CASA model, water availability up to a maximum of 50% can limit the photosynthesis which may be complemented by a subsequent fAPAR decrease to account for long-term water stress [77,101]. There is no direct water limitation in the CGLS-DMP and water stress is incorporated indirectly through the impact on fAPAR corresponding to a visible impact on the vegetation [33]. Maselli et al. [35,101] demonstrated that this modification is effective in improving the C-Fix model simulations for periods affected by significant water stress in Mediterranean tree ecosystems.

$$\varepsilon_{H2O} = 0.5 + (0.5 \times \frac{AET}{PET}) \tag{1.8}$$

ε_{H2O} can vary between 0.5, when strong water shortage reduces photosynthesis to half of its potential value, and 1, when there is no water shortage and photosynthetic reduction [101].

1.3.1.4. Autotrophic Respiration Factor

The autotrophic respiration factor in CGLS-DMP (ε_{AR}, Equation (1.9)) is dependent on daily temperature only [83].

$$\varepsilon_{AR} = 1 - (-3.049 + (0.01145 \times T_{24})) \tag{1.9}$$

where T_{24} is the daily mean air temperature (in °K). The model performance is improved by introducing biomass information in the modified DMP model, as in the GLOPEM2 model [78]. A semi-empirical relationship accounts for above-ground biomass (Equation (1.10)).

$$\varepsilon_{AR} = 1 - (\,(-3.049 + (0.01145 \times T_M)) \times (0.53 \times (\frac{W}{W + 50}))) \tag{1.10}$$

where T_M is ten daily mean composite of mean air temperature and W is above ground biomass, according to

$$W = 7166.1 \times (\rho_{min}^{-2.6}) \tag{1.11}$$

The variable $\rho_{min}^{-2.6}$ is the minimum reflectance in the red channel of AVHRR in the GLOPEM2 model. The lower $\rho_{min}^{-2.6}$, the higher the biomass [136]. Prince and Goward [137] argue that visible reflectance is positively related to standing biomass and canopy closure. The ten daily minimum composite values of the red channel of SPOT-VGT were used assuming both instruments reach comparable minimum values in similar conditions.

The long-term average autotrophic respiration fraction computed from both the original and modified formula was compared to the annual MOD17A3 NPP/GPP ratio in order to detect similarity and differences. The long-term average values of MODIS NPP and GPP were calculated separately, whereafter the NPP/GPP ratio was computed.

Table 1.3 presents a summary of unchanged and changed parameters in the DMP equation for both the CGLS and the modified version.

Table 1.3. Summary of unchanged and changed parameters in the DMP equation from CGLS as compared to the modified version.

	CGLS-DMP	Modified DMP	
R	Obtained from JRC-MARSOP on a daily basis at a 0.25° grid		
ε_p	$\varepsilon_p = 0.48$		
fAPAR	Derived from 10-daily SPOT VGT imagery at 1km² resolution		
ε_{H_2O}	No water stress factor	Water stress factor based on AET	
		C_3 plants (wheat)	C_4 plants (maize)
$\varepsilon_{RUE_{MAX}}$	2.54 kgDM/GJ for all C_3 plants	2.75 kgDM/GJ for wheat	3.5 kgDM/GJ for maize
ε_T	Blue curve in Figure 1.3	Blue curve in Figure 1.3	Red curve in Figure 1.3
ε_{CO2}	Blue to red dots in Figure 1.5a	Blue to red lines in Figure 1.5b	Dashed green line in Figure 1.5c
ε_{AR}	Figure 1.7a, range: 0.65–0.85 for study regions	Figure 1.7b, range: 0.5–0.7 for study regions	

1.3.2. Regression Analysis

The relative importance of CO_2 fertilization, temperature effect and water stress was determined through different combinations of these stress factors per site and crop type. Previous studies showed that NDVI and fAPAR were used to define site specific relations with crop yield [12,55,56,63,138,139]. Therefore, NDVI and fAPAR were compared with both CGLS derived and modified DMP.

The regression analysis was performed to relate remote sensing Biomass Proxies (BPs) and crop yields of silage maize and common wheat for the period 1999–2012 in Belgium, France and Morocco. A linear regression was calculated between official yield statistics and a regional BP value cumulated over the different periods during the growing season for each crop type and region [56]. Different dekadal combinations of BPs were explored, e.g., the maximum BPs, maximum BPs plus number of dekads, end of season BPs and sum of BPs between flowering and ripening. Although the optimal temporal window varies with crop type and region, the cumulated period of the BPs in the majority of cases includes the reproductive stage. The model performance was assessed using the coefficient of determination (R^2). The *p*-value, calculated by Pearson correlation, was also given to assess whether the relations between the crop yields and calibrated BPs were statistically significant.

A leave-one-year-out cross validation enabled testing the model robustness and goodness of fit using the best combination of variables in the regression analysis. The root mean square error (RMSE), the relative RMSE (RRMSE), mean bias error (MBE) and the index of model performance (E_1) were presented to evaluate the model performance. E_1 is a dimensionless index of model-observation agreement [140]. E_1 takes the value of 1 for perfect agreement. Although not negatively bounded, the value of 0 indicates that such a model has no more ability to predict the observed values than the observed mean (*i.e.*, a null model). All values below 0 reflect a model that performs worse than the null model.

1.4. Results

An overview of spatio-temporal information containing a comparison of modified DMP for C_3 and C_4 versions with CGLS-DMP is provided for Liège Region (BE) (Figure 1.6). These absolute values are estimated from the modified DMP for C_3 and C_4, and CGLS-DMP versions for different time periods during the growing season. The figure presents a clear difference between three versions of DMP.

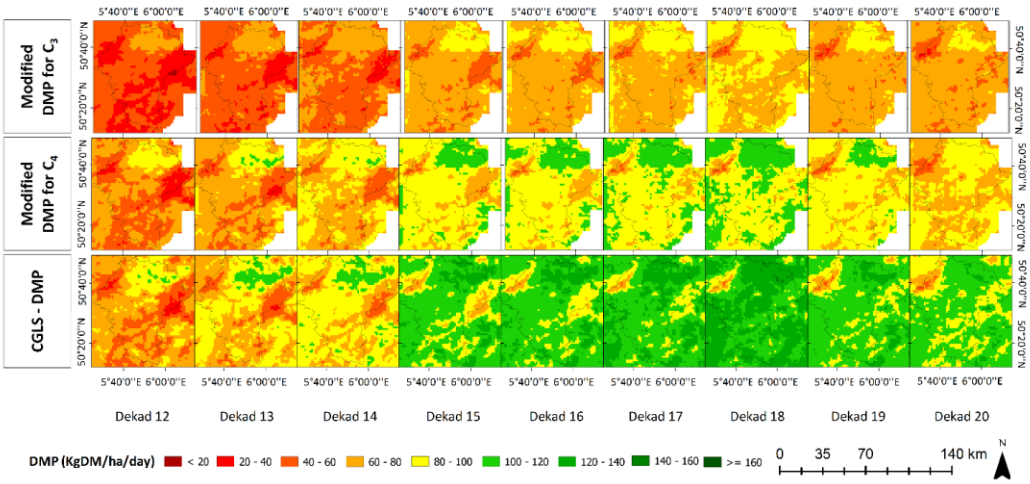


Figure 1.6. Comparison of modified DMP for C_3 and C_4 versions with CGLS-DMP for Liège Region (BE).

1.4.1. Autotrophic Respiration Fraction (ε_{AR})

The range for the ratio of NPP to GPP (which represents the ε_{AR} term in the DMP model) varies with land use type. The values of this ratio are 0.5 at the global scale [88,103,141]; 0.4 for corn and soybeans [52]; 0.4–0.6 for forest, 0.55 for cropland and 0.6 for grassland [104]; and 0.3–0.6 for maize, rice and wheat [105].

The modified ε_{AR} showed a better agreement with the values found in the literature and with the MODIS NPP/GPP ratio as compared to the original CGLS-DMP (Figure 1.7). Although it is expected to have higher ε_{AR} values in high altitudes and in places with higher temperatures, the differences between the modified DMP and MODIS NPP/GPP ratio are relatively large, particularly in mountainous and desertified regions. The main reason could be that the standing biomass was parameterized on the growing season reflectance value. The areas with very high ε_{AR} values in Figure 1.7b are sparsely vegetated [137] and fall outside cropland area (the scope of this study). Figure 1.7 presents the ε_{AR} versions and the MODIS NPP/GPP ratio for a visual interpretation of difference.

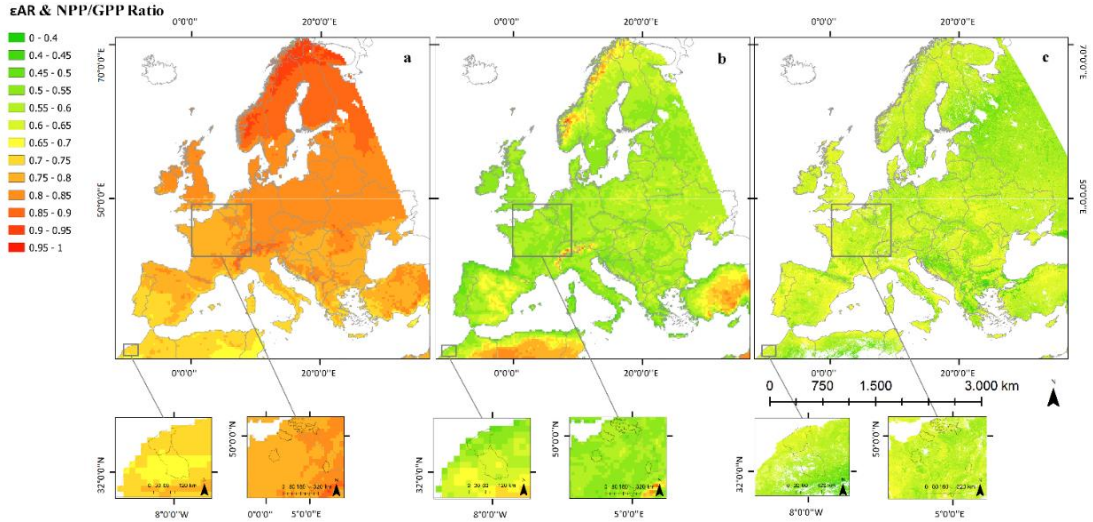


Figure 1.7. Average ε_{AR} calculated with CGLS-DMP (a), modified DMP (this study) (b) for the period 1999–2012, and MODIS NPP/GPP ratio from 2000–2013 (c) with the study sites extent.

In order to support the difference/similarity between the raster data presented (Figure 1.7), a numerical comparison was performed and difference maps created (Figure 1.8 and Figure 1.9). Agreement Coefficient (AC) is a measure for comparison of different spatial datasets e.g. two images from different algorithms [142]. AC ranges from 0 to 1 indicating the degree of agreement from complete disagreement to complete agreement [143]. As a measure of agreement, AC was calculated. AC values computed between ε_{AR} calculated with modified DMP (this study) and CGLS-DMP is 0.53, between MODIS NPP/GPP ratio from 2000 – 2013 and ε_{AR} calculated with CGLS-DMP is 0.57 and between MODIS NPP/GPP ratio from 2000 – 2013 and ε_{AR} calculated with modified DMP (this study) is 0.91.

Figure 1.8 displays the difference maps. On average, the difference between ε_{AR} calculated with modified DMP (this study) and CGLS-DMP is -0.23, MODIS NPP/GPP ratio from 2000 – 2013 and ε_{AR} calculated with CGLS-DMP is -0.18 and MODIS NPP/GPP ratio from 2000 – 2013 and ε_{AR} calculated with modified DMP (this study) is 0.05.

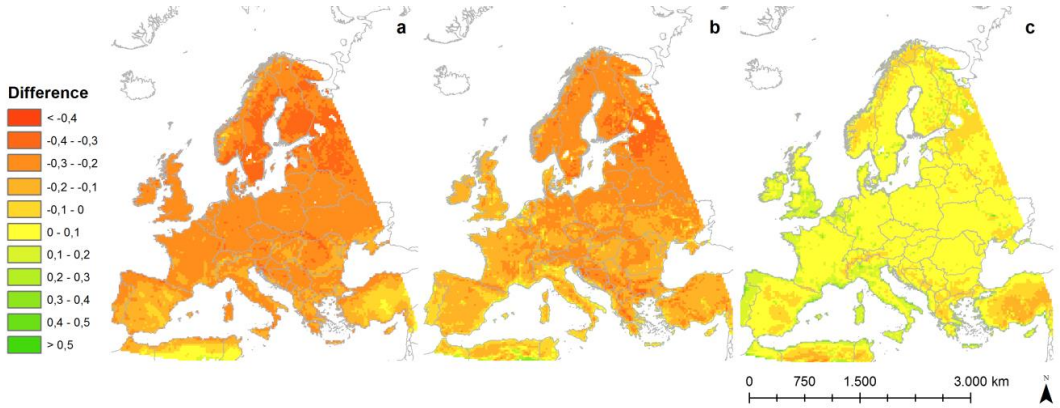


Figure 1.8. Difference maps of εAR calculated with modified DMP (this study) & CGLS-DMP (a), MODIS NPP/GPP ratio from 2000 – 2013 & εAR calculated with CGLS-DMP (b) and MODIS NPP/GPP ratio from 2000 – 2013 & εAR calculated with modified DMP (this study) (c).

The scatterplots were computed and presented in Figure 1.9. Comparison of numerical simulation results show that ε_{AR} calculated with modified DMP is closely related to the MODIS NPP/GPP ratio.

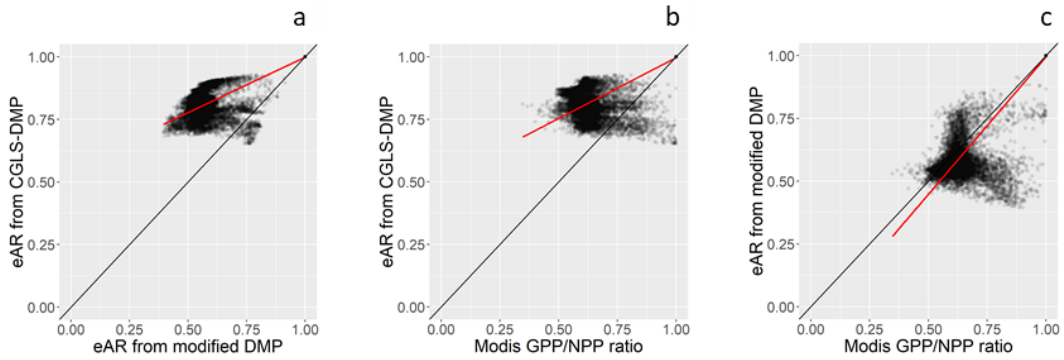


Figure 1.9. Scatterplots of ε_{AR} calculated with modified DMP (this study) & CGLS-DMP (a), MODIS NPP/GPP ratio from 2000 – 2013 & ε_{AR} calculated with CGLS-DMP (b) and MODIS NPP/GPP ratio from 2000 – 2013 & ε_{AR} calculated with modified DMP (this study) (c). The dotted lines are the 45° reference lines and the red lines are trend lines.

1.4.2. Linear Regression Analysis

Figure 1.10 shows the highest R^2 values for the relationship between the official yield statistics and each regional cumulative modified DMP computed with different stress factor combinations. In Belgium common wheat has the highest correlation with DMP, without any stress factors in the Polders and with DMP including CO_2 fertilization effect in the Sandy Loam Region. In France common wheat has the highest correlation with DMP, with water stress in Eure-et-Loir and DMP with CO_2 fertilization effect and temperature stress in Somme. In Morocco common wheat has the highest correlation with DMP, with all three stress factors in both El Jadida and El Kelaa des Sraghna. In Belgium silage maize has the highest correlation with DMP, with temperature and water stress factors in both the Loam and Liège regions. In France silage maize has the highest correlation with DMP, with all three stress factors in Ain and Haut-Rhin.

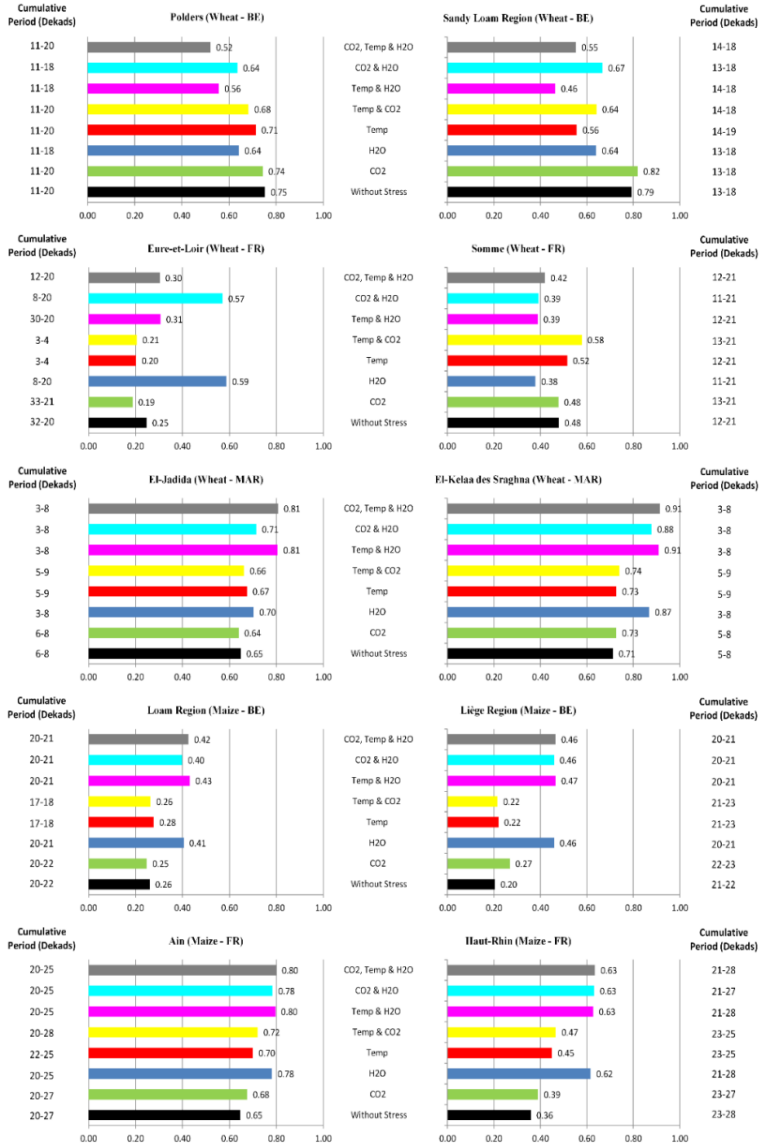


Figure 1.10. R^2 for a linear regression between official yield statistics and regional cumulative modified DMP computed with different stress factors or combinations. CO_2: with CO_2 fertilization effect; H_2O: with water stress factor; Temp: with temperature stress factor and combinations of these stress factors. The cumulative periods of modified DMP for each study region are presented in dekads.

The highest correlations for modified DMP per region and crop type (Figure 1.10) are compared with CGLS-DMP, fAPAR and NDVI in Figure 1.11. In each study region the modified DMP has a higher correlation than the CGLS-DMP, and in general performs better than fAPAR or NDVI. The correlations are significant for all study sites and BPs, except for fAPAR in the Sandy Loam Region (BE), NDVI in Somme (FR) and CGLS-DMP in Liège Region (BE). The correlations between observed yield and calibrated BPs were significant at the 0.01 level for modified DMP in all regions, but displayed a mixed picture in the case of CGLS-DMP, fAPAR and NDVI.

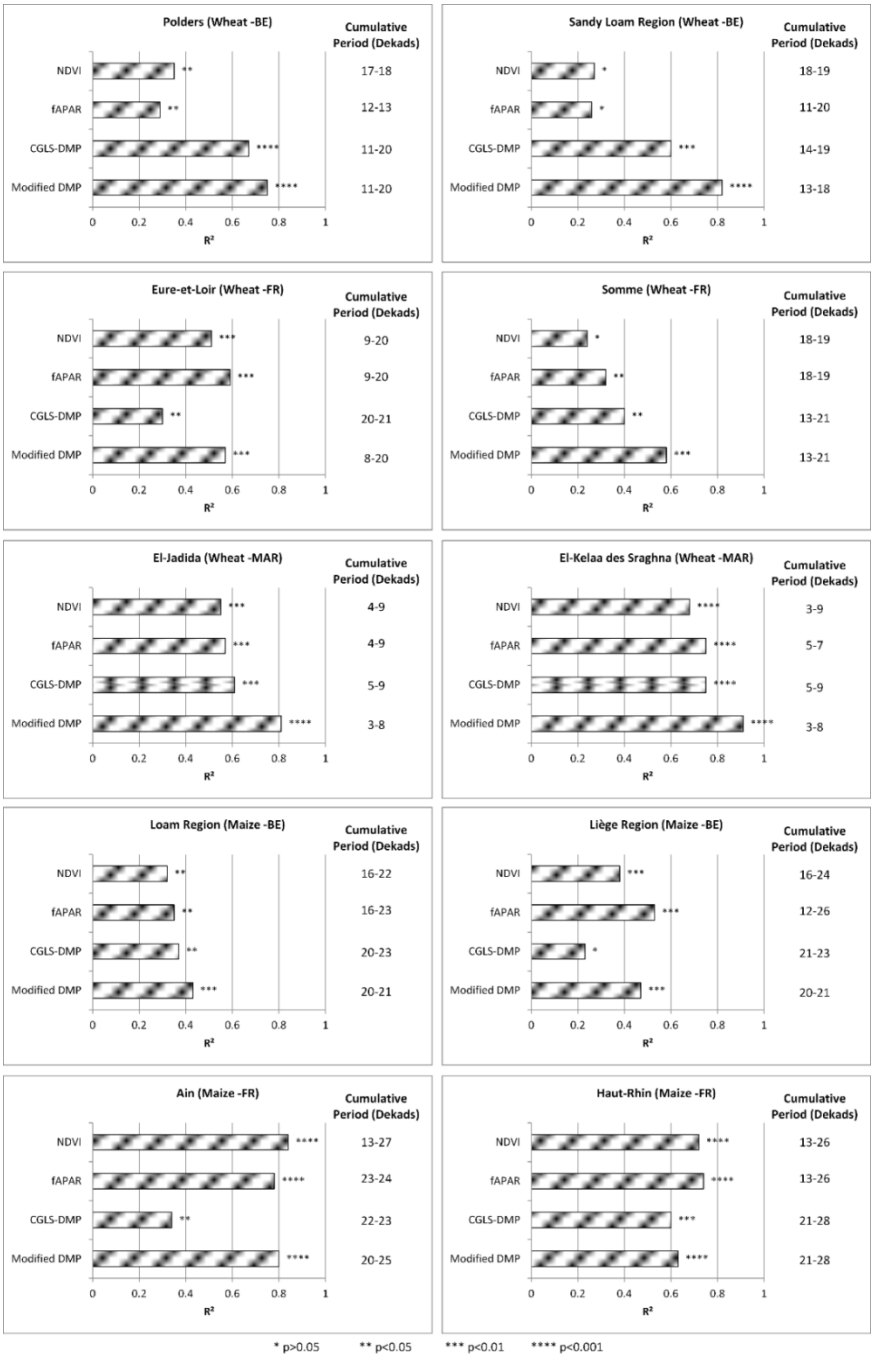


Figure 1.11. Cumulative period with highest R^2 of BPs (modified DMP, CGLS-DMP, fAPAR, NDVI), *p*-value (calculated by Pearson's correlation for r^2) and coefficient of determination (R^2 for the linear model) for wheat and maize per region.

The best combination of BPs from the regression analysis were calibrated with a leave-one-out cross validation. The RMSE (t/ha), the RRMSE (%), MBE (t/ha) and E_1 values calculated in the validation analysis are presented in Figure 1.12. The modified DMP has lower RMSE values than the other BPs which range from 0.25 t/ha–0.44 t/ha for common wheat and 0.68 t/ha–1.41 t/ha for silage maize. In general, the RRMSE values are lowest for the modified DMP. The RMSE values followed a similar trend as RRMSE. The MBE values are between −0.05 t/ha–0.05 t/ha except for the Liège Region (BE), Ain (FR) and Haut Rhin (FR) for all BPs. Overall, higher E_1 values were recorded for modified DMP compared to the other BPs. In addition, E_1 values show that the model is not performant for maize in Belgium and France (except Ain).

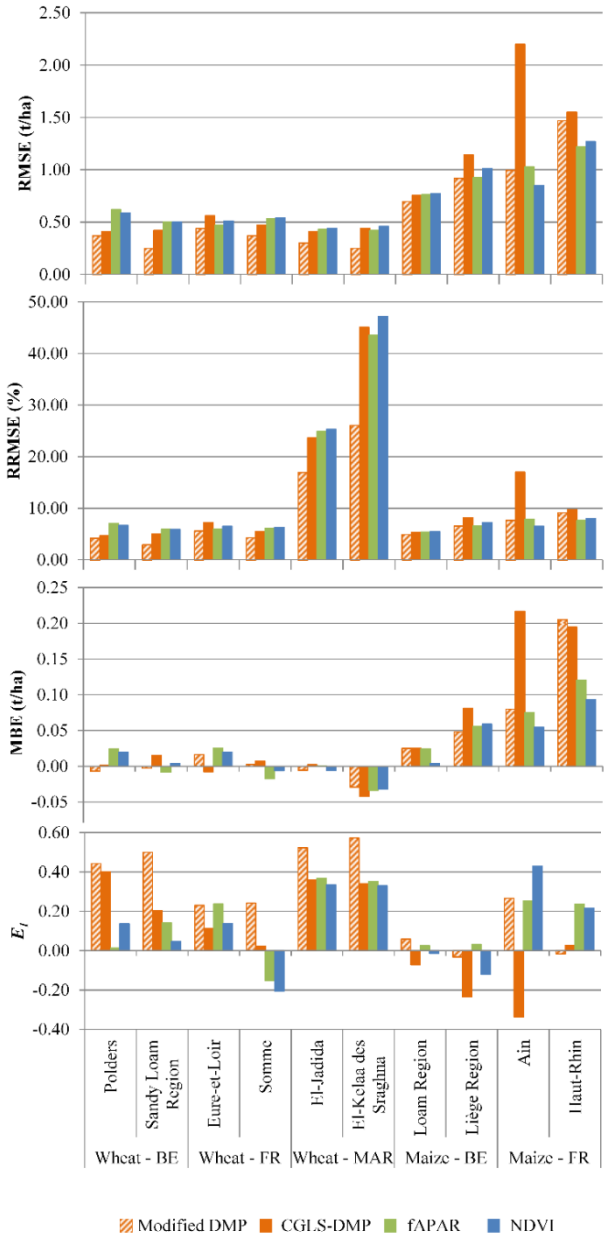


Figure 1.12. RMSE (t/ha), RRMSE (%), MBE (t/ha) and E_1 based on the correlation between the calibrated BP (modified DMP, CGLS-DMP, fAPAR, NDVI) and yield statistics per region for wheat and maize using a leave-one-out cross validation. The index accumulation period according to different BPs is not the same.

Figure 1.13 presents the temporal trends of the actual crop yield and predicted yield of BPs calibrated by the leave-one-out cross validation throughout the study period from 1999–2012. In general, the modified DMP followed a similar trend as the observed yield although there are some discrepancies for maize in the Loam Region (BE), Liège Region (BE) and Haut Rhin (FR).

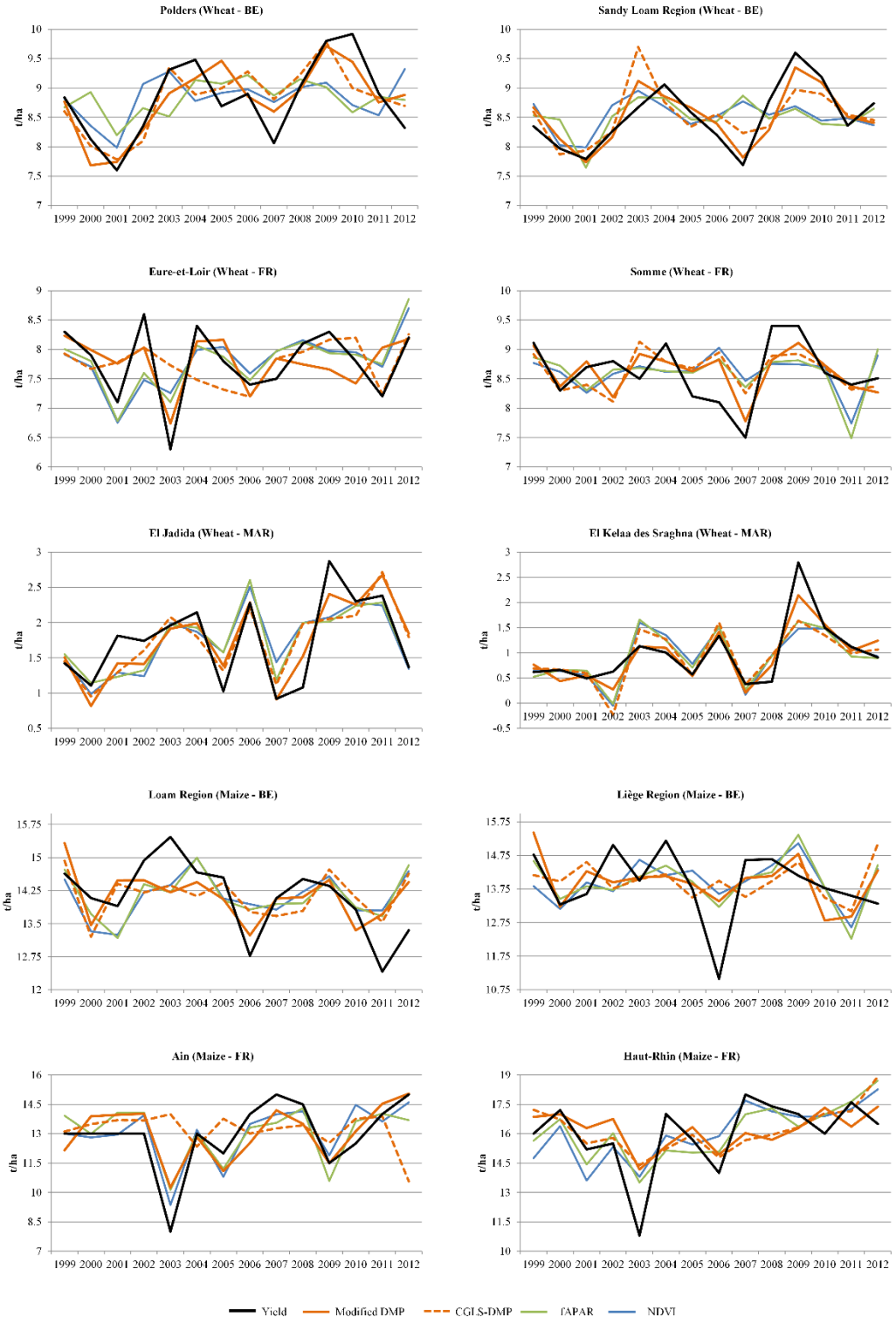


Figure 1.13. Temporal trends of predicted yield calibrated by leave-one-out cross validation technique confronted with the actual yield for 1999–2012 period per region.

1.5. Discussion

The modified DMP, the empirically identified model, is in general more efficient in describing the yield variability for maize and wheat yield across the study sites when different combinations of stress factors are included in the BPs regression analysis together with official statistics. Overall, the modified DMP correlates better with crop yield statistics than the CGLS-DMP, fAPAR and NDVI. On average, R^2 values increased from 0.49 for CGLS-DMP to 0.68 for the modified DMP, RMSE (t/ha) decreased from 0.84–0.61, RRMSE (%) reduced from 13.1–8.9, MBE (t/ha) decreased from 0.05–0.03 and E_1 increased from 0.08–0.28. In general, the cumulative BPs with the highest correlations coincide with periods that include the reproductive stage, which is in agreement with previous studies [23,64,144,145]. These periods are the most critical stages where any water stress may result in reduced crop yields [64]. In Canada, Mkhabela et al. [64] reported RMSE values ranging from 104 kg/ha–697 kg/ha from arid zones to sub-humid zones when relating MODIS-NDVI and spring wheat yield. In China, Ren et al. [22] reported a RMSE value of 214 kg/ha when relating MODIS-NDVI and wheat yield. The values for maize in Figure 1.12 are similar to those recorded by Sakamoto et al. [146] when relating MODIS Wide Dynamic Range Vegetation Index and maize yield in the US. However, E_1 values demonstrated that for maize the modified DMP model is performant only in Ain, which is also evident for the other BPs. In general, E_1 values for the modified DMP are higher than the CGLS-DMP. Hansen and Indeje [147] reported RMSE values ranging from 0.962 t/ha–1.195 t/ha when predicting field-scale maize yields simulated by CERES-maize in Kenya. In agro-ecological zones of Belgium, Klein [148] reported RMSE values fluctuating from 550 kg/ha–1430 kg/ha when relating the simulations of the B-CGMS crop growth model and winter wheat, and from 2150 kg/ha–7730 kg/ha for maize.

The modified DMP includes plant specific parameterizations for C_3 and C_4 plants, introduces potential water limitation and incorporates the impact of different stress factors during the most sensitive cropping periods. According to the results, the inclusion of different stress factors can improve local empirical models. The results indicate that the stress factors play different roles in different climate regimes and for different C_3 and C_4 plants. For instance, the water stress index in CGLS-DMP proved to be a limiting factor, particularly in Morocco. AgroMetShell was used to generate AET estimations for each crop during different stages of the growth period of 1999–2012; this water balance model requires time for collecting the necessary data and computation. An alternative to AgroMetShell is LSA-SAF MSG AET which has been available at the continental scale since 2009 and provides estimations closer to eddy covariance (EC) measurements [149,150]. Since the available years from MSG derived AET are too few for this study, AgroMetShell was used to compute AET. Another alternative could be the Global Land Evaporation Amsterdam Model. The model is a set of algorithms separately estimate different components of land evapotranspiration and has two datasets which are different from each other regarding their forcing and temporal coverage [151]. Another stress factor is the CO_2 fertilization effect which is highly dependent on plant species, soil properties and soil nutrient status [152–154]. Study sites in Belgium were more responsive to this stress factor compared to the soils with different fertility level in Morocco. Irrigation can alleviate the effect of stress factors such as drought and temperature, while fertilization alleviates plant nutrition stress [155]. However, the model approach used in this study did not consider either irrigation or fertilization. The autotrophic respiration fraction is overestimated in the CGLS-DMP model. Therefore, it was proposed in this study a new ε_{AR} which followed a similar trend as the findings in Zhang et al. [141]. For example, densely vegetated areas in Europe have lower ε_{AR} values. However, the values in sparsely

vegetated and high altitude areas are different than the values in Zhang et al. [141]. These areas are beyond the scope of this study and should be disregarded.

The modified DMP could explain more variation in irradiation conditions, short term environmental stresses and respiration costs compared to NDVI and fAPAR. Thus, it could be a more robust choice, although it has several modelling assumptions. Similar to the findings in Gitelson et al. [156,157] and Peng and Gitelson [158], Phillips et al. [159] found that NDVI underestimated productivity due to backscatter effects in the lower values and saturation in the higher values. Additionally, the frequency and intensity of extreme weather events could have significant impacts on biomass production and crop yields [160,161]. For example, floods and droughts can harm crops, reduce yields and increase crop prices. The amplitude of extreme events can be much larger at regional scales than at global scales [162]. For example, the 2003 heat wave affected the local natural environment, society and economy in many parts of Europe. Van der Velde et al. [110] reported that 2003 maize yield loss in France equaled 1.5 t/ha compared to the 2000–2006 average. This trend can also be seen in the study sites in France within BPs used in this study. The modified DMP is therefore better suited than CGLS-DMP to the extreme event in 2003 (e.g., for Ain, the predicted yield from the BPs calibrated for maize yields are 10 t/ha for modified DMP, 14t/ha for CGLS-DMP, 10 t/ha for fAPAR and 9 t/ha for NDVI, while the actual yield is 8 t/ha).

Although the proposed method shows promising results in crop yield forecasting for the study sites, there are a number of limitations. Cereal yields depend on the harvest index which is the fraction of the total aboveground biomass allocated to the grains [63]. Different yields for the same amount of aboveground biomass can be obtained because the harvest index varies with crop varieties, management practices, and water and nitrogen availability, all of which play at the field scale [63]. For this reason, the harvest index was not in the scope of the regional study. A limitation to predicting yield is the amplitude of the inter-annual crop yield variability. López-Lozano et al. [56] showed that the regression analysis performs better when the inter-annual variability is high, which is confirmed in this study: higher performance was seen in Morocco where the inter-annual variability is higher. In addition, many weeds, pests and fungi flourish under warmer temperatures, wetter climates and increased CO_2 levels, and affect yields [160]. Additionally, the quality of official crop statistics is arguable because they are collected based on the farmer's declarations which could be biased.

Products derived from SPOT-VGT, a sensor with a coarse 1 km spatial resolution, were used, and a mask was applied to these products using the GlobCover 2009 crop map in order to minimize the influence of non-agricultural land cover types. DMP is subsequently calculated per pixel located in cropland assuming wheat (C_3) or maize (C_4) cover. The two resulting maps could be combined using weights corresponding to the proportion of each crop, provided that the share per crop is known in a particular year. However, this proportion is not available before the end of the season, and it is beyond the scope of this paper to explore the feasibility of such extension. Furthermore, crop-specific maps, even dynamic crop maps which take into account crop rotations from one year to another, could minimize the effects of mixed signals through unmixing techniques or by selecting subsamples of purer pixels [163], thus improving the yield estimations [164]. The impact of stress is important for estimating dry matter productivity of cereal crops. The research could be extended to include other agricultural regions and arable crops. The incorporation of new generation remote sensing products with higher spatio-temporal resolutions into the DMP model is likely to improve yield estimates.

1.6. Conclusions

A modified DMP model was developed to include water stress based on actual evapotranspiration calculated with AgroMetShell and to adapt the existing factors (CO_2 fertilization effect, temperature stress and autotrophic respiration). The best results were obtained when different combinations of stress factors were included for each selected region and the modified DMP was cumulated between the flowering and ripening period. A linear regression between the modified DMP and crop yield statistics showed an increased model performance as compared to the CGLS-DMP. On average, for all sites and crops studied, RMSE (t/ha) decreased from 0.84 for CGLS-DMP to 0.61 for modified DMP, RRMSE (%) reduced from 13.1–8.9, MBE (t/ha) decreased from 0.05–0.03 and E_1 increased from 0.08–0.28. Although results did not differ much from what can be obtained using simpler BPs such as fAPAR and NDVI, obtaining similar results remains encouraging as the DMP approach may lend itself better to extrapolation in more extreme conditions. The results showed the potential of using the modified DMP for estimating crop yield at the regional scale. A combination of different stress factors produces better local yield estimates, but no single solution to an improvement of a global product could be demonstrated. The inter-annual variability of official yield statistics is low in Belgium and France, and high in Morocco. The latter enabled a better exploration of stress factors. Including appropriate stress factors and their impact during sensitive stages improves dry matter productivity estimates of cereal crops. For improving real RUE estimations, different remote sensing methods will be available such as chlorophyll-related vegetation indices, passive measurement of solar-induced chlorophyll fluorescence, and the photochemical reflectance index. With more accurate crop masks becoming available than the currently used GlobCover 2009, crop yield estimations might potentially further improve the results.

Chapter 1 explored how an existing product, CGLS-DMP, that was developed for SPOT-VGT, and transferred to PROBA-V, can be improved to more closely relate to yield anomalies over selected regions. The modified DMP could potentially be transferred to PROBA-V because PROBA-V was constructed as a successor of SPOT-VGT with similar bands, but with additional improved spatial resolutions. Besides PROBA-V's 1 km product compatible to SPOT-VGT, PROBA-V has a product at 100 m spatial resolution which has a potential to be used in crop identification mapping. In crop yield estimation studies, it is important to differentiate the cropland from non-cropland. Instead of a cropland mask, using crop specific masks would be an ideal approach for crop yield forecasting. Spatial resolution is the primary drawback of low spatial resolution satellite images because the pixel size is generally bigger than the typical field sizes. Therefore, the signal from these sensors has mostly mixed information from different surface types. Additionally, agricultural statistics are generally aggregated at regional levels for analysis with low and medium resolution satellite images [11]. Thus, availability of crop mask is particularly important for accurate yield estimations and to overcome mixed information problem. In Chapter 2, SMTs will be used to retrieve crop identification maps with PROBA-V 100 m time series at a global scale across different countries. Instead of a general cropland mask, a crop identification map as an outcome of the chapter could be possibly used in improving crop yield estimations. When the analysis was done in Chapter 2, the only available operational PROBA-V product was the NDVI. Since there was no fAPAR product at that time, the DMP could not be calculated and used in Chapter 2.

2 Crop Identification Mapping Using 100 m PROBA-V Time Series[2]

Abstract

A method is developed for crop identification inspired by spectral matching techniques (SMTs) and based on phenological characteristics of different crop types applied using 100 m PROBA-V NDVI data for the season 2014–2015. Ten-daily maximum value NDVI composites are created and smoothed. The study sites are globally spread agricultural areas located in Flanders (Belgium), Sria (Russia), Kyiv (Ukraine) and Sao Paulo (Brazil). For each pure pixel within the field, the NDVI profile of the crop type for its growing season is matched with the reference NDVI profile based on the training set extracted from the study site where the crop type originated. Three temporal windows are tested within the growing season with the order of importance from least to most: green-up to senescence, green-up to dormancy and minimum NDVI at the beginning of the growing season to minimum NDVI at the end of the growing season. Post classification rules are applied to the results to aggregate the crop type at the plot level. The overall accuracy (%) ranges between 65 and 86, and the kappa coefficient changes from 0.43–0.84 according to the site and the temporal window. In order of importance, the crop phenological development period, parcel size, shorter time window, number of ground-truth parcels and crop calendar similarity are the main reasons behind the differences between the results. The methodology described in this study demonstrates that 100 m PROBA-V has the potential to be used in crop identification mapping across different regions in the world.

2.1. Introduction

Accurate and timely information on the cropping area and crop type obtained from remote sensing data either or not in combination with ground surveys is key for estimating crop production. This information has significant environmental, policy, agricultural and economic implications for most national governments, since crop production figures are used for determining the amount of food to import or export at the end of the growing season [165,166]. The error introduced to crop production estimation from general agricultural land cover maps is minimized with accurate crop extent maps [41,43,165,167]. For remote sensing-based crop production estimates, the ideal approach would be to combine biomass proxies and crop maps. Biomass proxies have been available for decades and from different sensors at different spatial resolutions and studies are currently going on to improve these proxies. However, creating crop-specific

[2] Adapted from Durgun, Y.Ö.; Gobin, A.; Van De Kerchove, R.; Tychon, B. Crop Area Mapping Using 100-m PROBA-V Time Series. *Remote Sens.* 2016, *8*, 585.

maps has remained a challenge. In general, cropland maps, regardless of crop type, have proved to improve crop production forecasting [41].

Discriminating croplands from non-croplands and identifying different crop types can be achieved with remote sensing-based crop growth monitoring and in particular with indices that quantify the distinct green-up and senescence of the crop cycle [43]. Since different crops show different spectral responses depending on their maturity stage, the temporal dimension of remote sensing data is most useful for identifying major crop types and their phenology [165,168,169]. However, using remote sensing data in an operational context for crop area assessment requires a wide geographic coverage and high spatio-temporal resolution at a minimal cost [168].

Each vegetated land cover class represents a distinctive phenology (i.e., green-up, maturity, senescence and dormancy). Different datasets have been used to monitor the crop signature in remote sensing. The temporal resolution of high spatial resolution data is too low to derive crop phenology directly, whereas medium/low resolution data do not have sufficient spatial resolution to capture the crop-specific signature [170,171]. Despite these limitations, several studies successfully used low to high spatial resolution data or a combination of different resolutions for arable crop identification in the Great Plains by using 1 km NOAA-AVHRR and 500 m MODIS time series [168,172], for paddy rice identification in Japan by using 500 m MODIS time series [173] and in northeast China by using a Landsat-based phenology algorithm [174]. Liu et al. [165] developed a method for combining high spatial resolution data (Landsat, 30 m) with high temporal resolution data (MODIS, 500 m) to achieve a superior classification of crops in the Mississippi River Basin. At the present time, crop identification mapping is possible with Sentinel-2 with a high spatial resolution (10 m) and a three to five-day revisiting time. However, Sentinel-2 dataset would cause a problem of managing a high volume of data in terms of storage and processing time. Therefore, the idea in this study is to study whether 100 m dataset is already good enough for crop identification mapping. Since the final goal is to use the outcome map for crop production estimates, it is hypothesized that 100 m PROBA-V data can fulfil the requirements of both revisiting time and spatial resolution for crop identification mapping at the regional scale.

Different methods for discriminating cropland and mapping different crop types based on vegetation phenology exist. The work in Wardlow et al. [168] investigated the class separability between specific crop types in time series vegetation index data using the Jeffries–Matusita distance. In another study, Xavier et al. [175] applied a cluster analysis and used the Euclidean distance to compute the temporal distance of enhanced vegetation index values among samples. Support vector machines were used to map abandoned agriculture at large scales with coarse-resolution MODIS imagery and phenology metrics calculated with TIMESAT [176] (For the details of TIMESAT, please refer to [177]). The work in Foerster et al. [178] used agro-meteorological data containing information on times of crop growth stages, which were utilized to obtain the average phenological pattern for each individual crop type.

Several studies were done on crop identification mapping using different remote sensing techniques and analyzing time series data at different resolutions. Machine learning algorithms, such as random forest, artificial neural network and support vector machine, perform significantly better compared to the traditional supervised classification methods [179,180]. Dynamic time warping has emerged as a promising new technique for time series data mining applications that include land cover mapping [181]. The major disadvantage of these methods is their computational complexity [182]. Spectral matching techniques (SMTs) are an innovative method of identifying and labelling information classes in historical time series data [183]. Originally, SMTs were used in hyperspectral analysis of minerals [183]. Time series data, however, can be treated in a similar manner

as hyperspectral data where hundreds of bands stack a single instance of a hyperspectral image [184]. According to the method, two time series are matched with the target 'spectra', which are acquired from ideal end-member classes known through census data, ground truth or maps of the study area. The work in Thenkabail et al. [183] tested the method using monthly AVHRR data in the Krishna River Basin in India and demonstrated that spectral similarity was the best method.

A number of factors affect the accuracy of crop classification. Differences in crop phenology [185], agricultural field size [186,187] and the observation period length are all known to have a significant effect on accuracies. In addition, each crop has unique phenological features, which are affected by regional variations in climate, management practices and the variability among cultivars of a same species which might be more different than across species [168,173,188]. Furthermore, varying spectral responses from the soil can change the ability to discriminate the crop type throughout the growing season [189]. Accordingly, image acquisition during those periods when crop separability is the highest is crucial to increase crop classification accuracy [42,189,190].

The objective of this study is to develop a crop identification mapping approach applicable at a global level inspired by the SMT method [183] on a seasonal basis using 100 m PROBA-V NDVI data. PROBA-V data at a 100 m spatial and five-day temporal resolution likely improve land monitoring studies compared to the 250 m spatial and eight-day temporal resolution of MODIS data, or the 300 m spatial and one-day temporal resolution of PROBA-V, or the 10 km spatial and one-day temporal resolution of NOAA-AVHRR [25]. Although the 100 m PROBA-V data tend to be more advantageous, the time series data are currently limited, as they became available in May 2013. SMTs were used for seasonal crop identification with time series data by using different temporal windows throughout the growing season: from green-up to senescence, from green-up to dormancy and from minimum NDVI at the beginning of the growing season to minimum NDVI at the end of the growing season. The method aims to facilitate crop production estimates by developing crop-specific maps.

2.2. Materials

2.2.1. Study Areas and Ground Data

The study sites are globally-spread agricultural areas, which include Flanders (Belgium), Sria (Russia), Kyiv (Ukraine) and Sao Paulo (Brazil) (Figure 2.1). The areas are characterized by different climatic conditions, agricultural management, soil types and topography (Table 2.1).

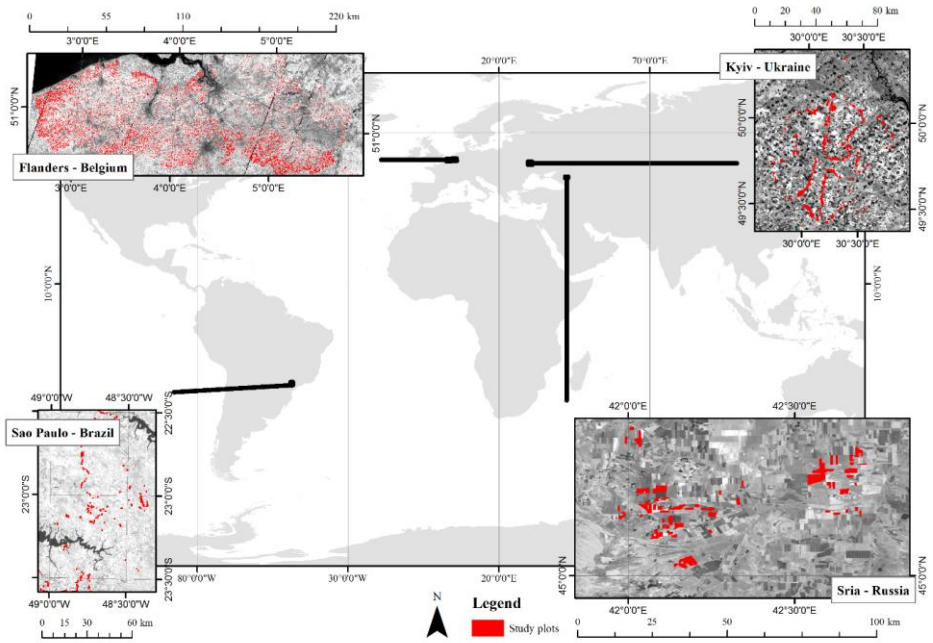


Figure 2.1. Study sites overlaid with field boundaries. The background images were extracted from the 100 m PROBA-V red band.

The extent and characteristics varied between the study areas (Table 2.2). The number of fields of Flanders (Belgium) is relatively high compared to the other study sites, since the database covers the entire Flanders region. Sria (Russia) has the largest field sizes followed by Kyiv (Ukraine), Sao Paulo (Brazil) and Flanders (Belgium).

Table 2.1. Site characteristics.

Characteristics	Flanders (Belgium)	Sria (Russia)	Kyiv (Ukraine)	Sao Paulo (Brazil)
Surface Area	20,000 km^2	3700 km^2	11,000 km^2	9000 km^2
Climatic conditions	Moderate maritime climate [191]	Temperate-continental climate with cold winters and hot dry summers [192]	Humid continental [193]	Humid tropical [194]
Soil types	Albeluvisols, Luvisols, Podzols and Fluvisols [195]	Chestnut soils and chernozems [192]	Chernozems [196]	Ferralsols, 20% clay [194]
Topography	The topography is flat to hilly [197]	The topography is mostly flat with slopes ranging from 0–2%; and nearly 15% of the territory is hilly with slopes from more than 2%. [192]	The topography is mostly flat with slopes ranging from 0–2%. Near 10% of the territory is hilly with slopes about 2–5% [193]	The local topography is hilly, with elevations ranging from 500 m–650 m.
Crop calendar	Maize: April–November [198]	Flax: April–July [192]	Winter barley: September–July [193]	Maize: September–April [198]
	Potato: March–July [198]	Maize: May–November [192]	Winter wheat: September–August [193]	Soybean: October–May [198]
	Sugar beet: April–October [198]	Peas: April–August [192]	Spring wheat: May–September [193]	Sugarcane: September–March [198]
	Winter barley: September–July [198]	Soybean: April–November [192]	Maize: May–October [193]	
	Winter wheat: October–August [198]	Spring barley: April–August [192]	Rape: September–August [193]	
		Sugar beet: April–October [192]	Spring barley: April–August [193]	
		Sunflower: May–October [192]	Soybean: April–September [193]	
		Winter barley: October–July [192]	Sugar beet: April–October [193]	
		Winter rape: September–August [192]	Sunflower: May–October [193]	

Table 2.2. Crop cover characteristics of the study areas.

Study Area	Crop Type	Number of Fields	Acreage (ha)	Field Size Range (ha)	Mean Area of Fields (ha)	Ratio of Pure to Non-Pure Pixels
Flanders (Belgium)	Grain maize	42,517	36,000	1–26	1	0.01
	Potato	16,941	35,000	1–45	2	0.03
	Sugar beet	7697	19,000	1–43	2	0.04
	Winter barley	6818	11,000	1–24	2	0.02
	Winter wheat	29,910	54,000	1–37	2	0.03
Sria (Russia)	Flax	29	2098	23–298	83	1.34
	Maize	18	1755	65–167	76	1.58
	Peas	6	663	49–217	72	1.94
	Soybean	8	370	27–78	27	1.33
	Spring barley	3	165	25–82	25	1.26
	Sugar beet	1	110	110	110	1.59
	Sunflower	11	1259	49–409	73	2.15
	Winter barley	29	2276	36–172	64	1.49
	Winter rape	17	1561	53–305	91	1.72
Kyiv (Ukraine)	Winter barley	2	628	22–30	26	0.53
	Winter wheat	186	12,498	1–193	67	1.31
	Spring wheat	23	791	3–101	34	1.14
	Maize	83	4385	2–162	53	1.20
	Winter rape	49	2389	2–161	49	0.99
	Spring barley	21	628	1–143	30	0.89
	Soybean	110	3000	1–123	27	0.72
	Sugar beet	18	1623	3–270	90	1.98
	Sunflower	34	1503	3–160	44	1.20
Sao Paulo (Brazil)	Maize	30	478	2–81	16	0.39
	Soybean	91	2211	1–101	24	0.42
	Sugarcane	154	3481	1–122	23	0.41

2.2.2. Ground Data

Ground data, containing crop type and parcel information, were obtained from the FP7 Stimulating Innovation for Global Monitoring of Agriculture (SIGMA) project and the digital map parcels dataset 'Geo-Data Infrastructure-Flanders' [199]. The crop information in the parcels dataset was declared by farmers in Flanders-Belgium and collected through field campaigns in other study sites. The dataset provided a good approximation of the actual agricultural land use [200], though it cannot be regarded as 100% correct because deviations can occur due to differences in planting and declaration [201]. For Flanders (Belgium), five main crops from the parcel information database were selected: grain maize, potato, sugar beet, winter barley and winter wheat. For other study sites, the parcel size and crop type information for the 2014–2015 growing season were obtained from the SIGMA project (geoglam-sigma.info).

2.2.3. NDVI Data Description

PROBA-V was launched in May 2013 to fill the gap between SPOT-VGT and Sentinel-3 satellites. PROBA-V has 4 spectral bands: blue (centered at 0.463 µm), red (0.655 µm), NIR (0.845 µm) and SWIR (1.600 µm). The central camera of the PROBA-V satellite provides a 100 m data product with a 5–8 days revisiting time and daily images at 300 m and 1 km resolution. Non-composited atmospherically-corrected NDVI images from 100 m PROBA-V were obtained from http://www.vito-eodata.be. Ten-daily maximum value NDVI composites were created and smoothed in Software for the Processing and Interpretation of Remotely sensed Image Time Series (SPIRITS) with the algorithm of Swets et al. [202] for the growing season 2014–2015. SPIRITS is a free software used to analyze

satellite-derived image time series in crop and vegetation monitoring that can be downloaded from http://spirits.jrc.ec.europa.eu/. The smoothing algorithm was used to remove higher frequency noise [203].

2.3. Methods

The methodology applied consisted of 5 different steps: (i) collecting training/validation samples; (ii) deriving reference NDVI profiles and phenological stages; (iii) classification using SMTs; (iv) post-classification; and (v) accuracy assessment. More details on each step is listed below, and a flowchart is presented in Figure 2.2:

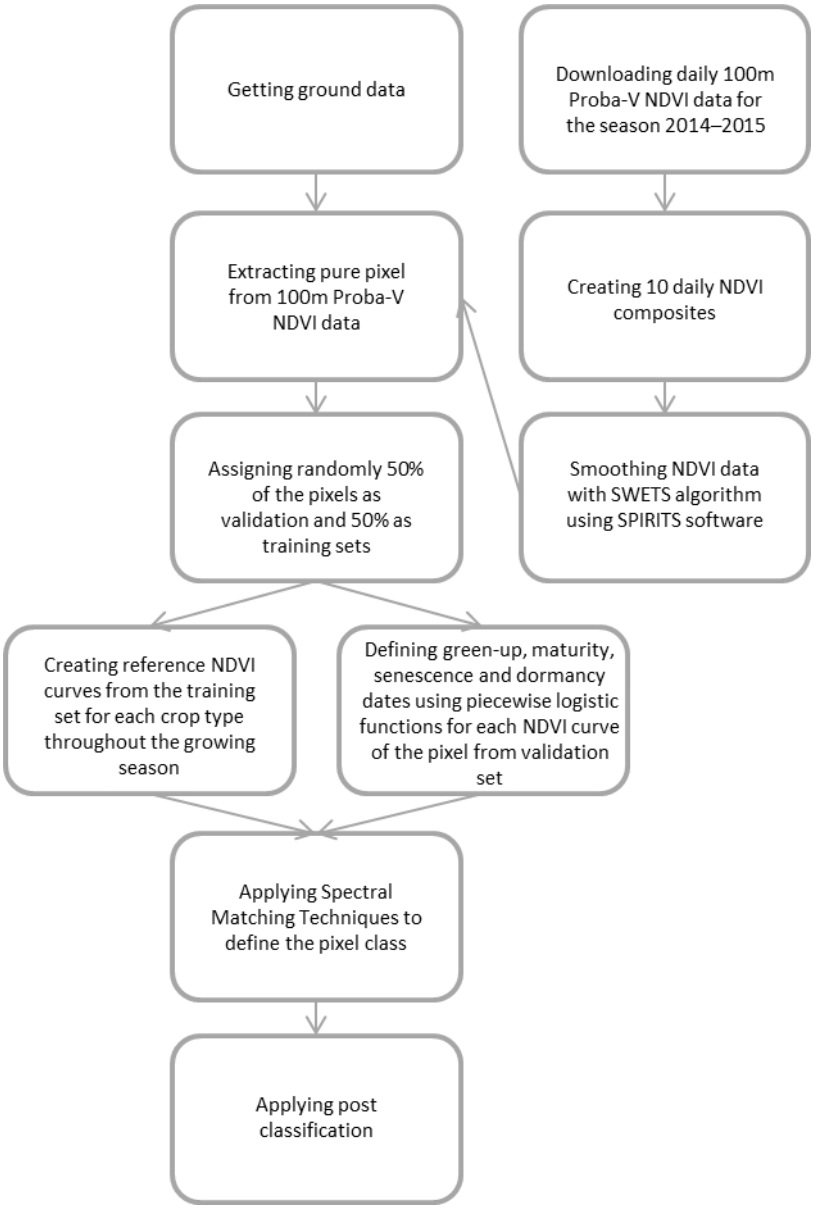


Figure 2.2. Flowchart of the crop identification mapping methodology.

2.3.1. Collecting Training/Validation Samples

For each study area, the PROBA-V 100 m NDVI images were overlaid with the crop field boundaries, and both pure (i.e., homogenous pixels with a 100 m resolution) and

mixed pixels were derived. In this study, only pure pixels were used. They were randomly divided into two equal groups for each crop type, one for training and one for validation. A random sampling scheme was preferred, as this is likely to prevent bias to the accuracy assessment [204].

The first PROBA-V 100 m image was available during the second dekad of March 2014. The analysis was performed from the first available image onwards, thereby leaving out the planting period of winter crops.

2.3.2. Deriving Reference NDVI Profiles and Phenological Stages

For each study area and all the different crop types, a reference or 'ideal' NDVI profile was calculated by taking the average NDVI using all of the pure pixels from the training set. For Sao Paulo where double cropping occurs, the time of the year ground data were collected was considered to decide on the growing season. The crop calendars, which were obtained from AGIV [199] and Homayouni and Roux [205], were used to compare with the reference NDVI profiles and establish similarity.

We used piecewise logistic functions (similar to Zhang et al. [206]) to define the four transition dates in the reference NDVI profiles: green-up (onset of photosynthetic activity, (a) in Figure 2.3), maturity (maximum plant green leaf area, (b) in Figure 2.3), senescence (rapid decrease of photosynthetic activity and green leaf area, (c) in Figure 2.3) and dormancy (zero physiological activity, (d) in Figure 2.3) [206] (For a more detailed description of the algorithm, see Zhang et al. [206]). The four transition points defined the boundaries of different time intervals that corresponded to distinctly different crop stages. Subsequently, the classification results were compared for three time windows: from green-up to senescence ((a–c) in Figure 2.3), from green-up to dormancy ((a–d) in Figure 2.3) and from minimum NDVI at the beginning of the growing season to minimum NDVI at the end of the growing season. The different time windows were chosen to explore possibilities for early crop detection, which in turn enables crop identification mapping as early as possible during the growing season.

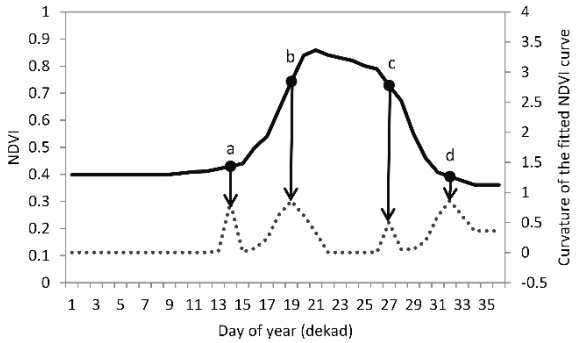


Figure 2.3. A schematic presentation of the annual cycle of crop phenology characterized by four key transition dates ((a) green-up, (b) maturity, (c) senescence and (d) dormancy) calculated using values in the rate of change in the curvature (adapted from Zhang et al. [206]).

2.3.3. Classification Using Spectral Matching Techniques

The crop type of each pure pixel in the validation set was identified by 'matching' the pixel profile with different reference NDVI profiles during the specific time window as defined in the previous section using piecewise logistic functions within the growing periods of the reference crop type. To determine the actual crop type, the spectral

similarity value (*SSV*) [183] was calculated between each pure pixel from the validation set and each candidate reference NDVI profile. To calculate *SSV*, the following formula was used:

$$SSV = \sqrt{ED_{normal}^2 + (1-\rho)^2} \quad (2.1)$$

where ρ and ED_{normal}^2 are the correlation coefficient and the normalized Euclidean distance between the different candidate reference NDVI profiles and the pure pixel profiles, respectively. These parameters are calculated as follows:

$$\rho = \frac{1}{n-1}\left[\frac{\sum_{i=1}^{n}(ref_i - \mu_{ref})(r_i - \mu_a)}{\sigma_{ref}\sigma_a}\right] \quad (2.2)$$

and:

$$ED_{normal} = (ED - m)/(M - m) \quad (2.3)$$

which is the normalized version of:

$$ED = \sqrt{\sum_{i=1}^{n}(ref_i - a_i)^2} \quad (2.4)$$

where ref_i is the reference NDVI profile at time *i* from 1 to *n*; μ_{ref} is the mean reference NDVI profile; a_i is the pure pixel NDVI profile from validation set at time *i* from 1 to *n*; μ_a is the mean pure pixel NDVI profile from validation set; σ_{ref} is the standard deviation of the reference NDVI profile; and σ_a is the standard deviation of the pure pixel NDVI profile from the validation set. ρ values vary between 0 and 1 and measure the shape of the temporal NDVI profile over time. The higher the ρ, the higher the similarity in the shape of the temporal NDVI profiles. The Euclidian distance (*ED*), normalized by using the historical minimum (*m*) and historical maximum (*M*) NDVI of the reference profile for a logical comparison, represents the closeness between the two profiles. ED_{normal} values vary between 0 and 1. The lower the ED_{normal}, the closer the profiles are.

Accordingly, *SSV* is a similarity measure, which combines both the shape (ρ) and distance (ED_{normal}) measures [207]. *SSV* values vary between 0 and a maximum of the square root of the two measures [207]. The smaller the *SSV*, the more similar the profiles. The pure pixel was assigned from the validation data with the label of the reference NDVI, which has the smallest SSV.

2.3.4. Post-Classification Filtering

In a final step, a post-classification rule was applied based on the mode value per crop type where the maximum frequency in one parcel was used to remove outliers in the classified parcel. The parcel was subsequently labelled with the crop type that had the majority of the pixels. Crop identification maps were created after applying the post-classification.

2.3.5. Accuracy Assessment

Confusion matrices at the parcel level were constructed to compare predicted and actual class membership. Based on the confusion matrices, classification accuracy statistics included overall accuracy, producer's accuracy, user's accuracy and kappa coefficients. Kappa analysis provided a measure of the magnitude of agreement between the predicted and actual class membership. A kappa value of 0 represents a total random classification, while a kappa value of 1 corresponds to a perfect agreement between the reference and classification data.

2.4. Results

Figure 2.4 presents particular time windows for two selected crops in two test sites. Depending on the crop type and region, the minimum and maximum NDVI differ from each other. For instance, the maximum NDVI value of soybean is close to 0.8 in Russia and 0.9 in Ukraine. Additionally, also the length of the growing season is different in different regions for the same crop, e.g., the planting and harvesting period for maize in Belgium is longer than in Brazil. It should be noted that the phenological shift due to the late start of the growing season was not studied.

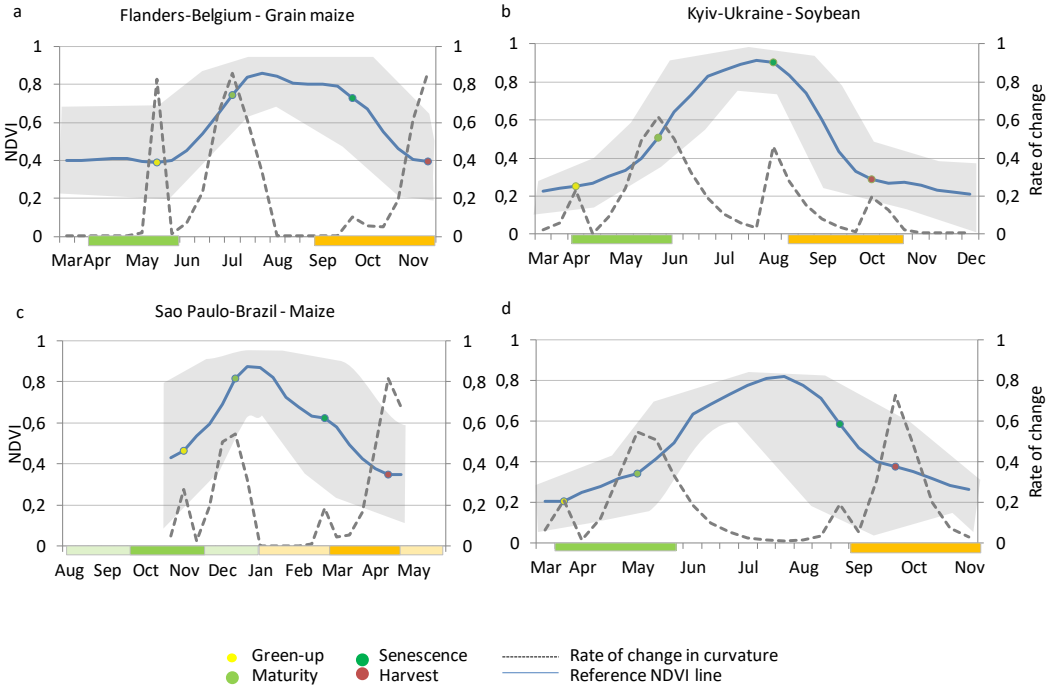


Figure 2.4. Crop time windows for maize in Flanders-Belgium (a) and Sao Paulo, Brazil (c); and for soybean in Kyiv, Ukraine (b), and Sria, Russia (d). The four phenological transition dates were calculated from piecewise logistic functions. The grey zone represents the minimum and maximum NDVI values in the training dataset. The crop calendar is presented below each graph, where green represents the planting time and orange the harvesting time. Light green and orange colors represent low activity for maize in Brazil.

2.4.1. Accuracy Assessment

Overall, the proposed method using 100 m PROBA-V data was effective in crop type classification with relatively high accuracies. The accuracy ranged between 75% and 80% in Flanders-Belgium, 72% and 86% in Sria, Russia, 71% and 86% in Kyiv, Ukraine, and 65% and 77% in Sao Paulo, Brazil (Table 2.3). The kappa coefficient ranged between 0.67 and 0.74 in Flanders-Belgium, 0.67 and 0.84 in Sria, Russia, 0.63 and 0.82 in Kyiv, Ukraine, and 0.43 and 0.61 in Sao Paulo, Brazil (Table 2.3).

In all four study sites, accuracies and kappa coefficient values increased with respect to when a longer time window was considered. The results were considerably better when the time window covered the entire growing season compared to the window from green-up to senescence. Crops with similar phenological profiles were sometimes incorrectly classified particularly when only part of the growing season was considered, e.g., for summer and winter crops. When the crop growth profile had a distinctive feature compared to other crops, it was easier to differentiate it from the other crops. For

instance, sugar beet in Ukraine has a longer period between maturity and senescence compared to the other summer crops. Another important outcome of the results is that post-classification improved the accuracy and kappa results for all sites, except for Belgium (see Table 2.3 for classification analysis and Table 2.4 for post-classification analysis). The overall poorest result was obtained for producer accuracy in Brazil, due to the mixing of soybean with maize and sugarcane pixels.

Table 2.3. Confusion matrix of classification analysis for green-up to senescence, green-up to dormancy and minimum NDVI at the beginning of the growing season to minimum NDVI at the end of the growing season assessment of Flanders-Belgium (a), Sria-Russia (b), Kyiv-Ukraine (c) and Sao Paulo-Brazil (d). The number of correctly-classified crops, the producer accuracy, the user accuracy, the overall accuracy and the kappa coefficient are presented.

a

		Ground truth																			
		Green-up to senescence							Green-up to dormancy							minimum NDVI to minimum NDVI						
Classified in satellite image as:		Grain maize	Potato	Sugar beet	Winter barley	Winter wheat	Total	User accuracy	Grain maize	Potato	Sugar beet	Winter barley	Winter wheat	Total	User accuracy	Grain maize	Potato	Sugar beet	Winter barley	Winter wheat	Total	User accuracy
	Grain maize	**1342**	310	207	1	3	1863	72%	**1391**	289	171	3	7	1861	75%	**1382**	292	150	4	6	1834	75%
	Potato	85	**1285**	107	5	15	1497	86%	69	**1745**	168	1	31	2014	87%	69	**1733**	146	1	34	1983	87%
	Sugar beet	107	239	**846**	0	4	1196	71%	97	144	**848**	1	13	1103	77%	92	108	**882**	0	20	1102	80%
	Winter barley	9	22	4	**333**	296	664	50%	3	0	0	**470**	485	958	49%	3	0	1	**467**	297	768	61%
	Winter wheat	22	86	13	143	**2743**	3007	91%	6	11	7	41	**2738**	2803	98%	2	19	5	41	**3034**	3101	98%
	Not classified	18	312	34	36	392	792		17	65	17	2	179	280		35	102	27	5	62	231	
	Total	1583	2254	1211	518	3453	9019		1583	2254	1211	518	3453	9019		1583	2254	1211	518	3453	9019	
	Producer accuracy	85%	57%	70%	64%	79%			88%	77%	70%	91%	79%			87%	77%	73%	90%	88%		
Overall classification accuracy					73%							80%							83%			
Kappa coefficient					0.65							0.74							0.78			

b

		Ground truth																									
		Green-up to senescence											Green-up to dormancy											minimum NDVI to minimum NDVI										
Classified in satellite image as:		Flax	Maize	Peas	Soybean	Spring barley	Sugar beet	Sunflower	Winter barley	Winter rape	Total	User accuracy	Flax	Maize	Peas	Soybean	Spring barley	Sugar beet	Sunflower	Winter barley	Winter rape	Total	User accuracy	Flax	Maize	Peas	Soybean	Spring barley	Sugar beet	Sunflower	Winter barley	Winter rape	Total	User accuracy
	Flax	**454**	8	124	0	1	0	19	0	0	606	75%	**379**	1	6	13	8	0	5	0	0	412	92%	**539**	1	1	22	3	0	6	0	0	572	94%
	Maize	2	**298**	0	1	0	0	105	0	0	406	73%	1	**297**	0	0	0	0	109	0	0	407	73%	3	**332**	0	0	0	0	125	0	0	460	72%
	Peas	158	2	**146**	29	3	0	3	0	0	341	43%	329	0	**296**	0	1	0	1	0	0	627	47%	176	0	**294**	0	0	0	0	0	0	470	63%
	Soybean	66	4	0	**116**	0	0	86	0	0	272	43%	111	5	0	**121**	0	0	74	0	0	311	39%	115	6	0	**121**	0	0	82	0	0	324	37%
	Spring barley	26	4	43	3	**67**	0	5	7	111	266	25%	16	0	12	1	**61**	0	3	9	29	131	47%	18	0	21	1	**65**	0	0	12	8	125	52%
	Sugar beet	46	0	0	8	0	**49**	165	0	0	268	18%	16	1	0	7	1	**49**	48	0	0	122	40%	21	1	0	10	1	**49**	48	0	0	130	38%
	Sunflower	46	59	0	2	0	0	**205**	0	0	312	66%	16	76	0	3	0	0	**335**	0	0	430	78%	4	43	0	3	0	0	**340**	0	0	390	87%
	Winter barley	0	2	0	0	0	0	0	**840**	86	928	91%	0	0	0	0	0	0	0	**875**	33	908	96%	0	0	0	0	0	0	0	**909**	18	927	98%
	Winter rape	2	0	0	0	0	0	1	127	**423**	553	76%	0	0	0	0	0	0	0	118	**640**	758	84%	0	0	0	0	0	0	3	78	**684**	765	89%
	Not classified	107	7	3	0	0	0	20	33	95	265		39	4	2	14	0	0	34	5	13	111		31	1	0	2	2	0	5	8	5	54	
	Total	907	384	316	159	71	49	609	1007	715	4217		907	384	316	159	71	49	609	1007	715	4217		907	384	316	159	71	49	609	1007	715	4217	
	Producer accuracy	50%	78%	46%	73%	94%	100%	34%	83%	59%			42%	77%	94%	76%	86%	100%	55%	87%	90%			59%	86%	93%	76%	92%	100%	56%	90%	96%		
Overall classification accuracy					**62%**											**72%**											**79%**							
Kappa coefficient					0.56											0.68											0.75							

c

Classified in satellite image as:	Ground truth: Green-up to senescence										
	Winter barley	Winter wheat	Spring wheat	Maize	Winter rape	Spring barley	Soybean	Sugar beet	Sunflower	Total	User accuracy
Winter barley	**14**	606	0	2	264	0	10	0	0	896	2%
Winter wheat	4	**2792**	40	0	131	1	2	0	0	2970	94%
Spring wheat	0	12	**120**	17	20	8	1	2	12	192	63%
Maize	0	0	1	**1163**	0	0	221	0	10	1395	83%
Winter rape	0	8	0	0	**195**	0	0	1	0	204	96%
Spring barley	0	122	152	7	15	**223**	10	0	0	529	42%
Soybean	0	1	0	99	0	0	**180**	19	28	327	55%
Sugar beet	0	0	0	19	0	0	0	**554**	0	573	97%
Sunflower	0	0	0	65	38	0	203	26	**531**	863	62%
Not classified	0	46	9	186	78	25	395	1	61	801	
Total	18	3587	322	1558	741	257	1022	603	642	8750	
Producer accuracy	78%	78%	37%	75%	26%	87%	18%	92%	83%		
Overall classification accuracy				**66%**							
Kappa coefficient				0.58							

Classified in satellite image as:	Ground truth: Green-up to dormancy										
	Winter barley	Winter wheat	Spring wheat	Maize	Winter rape	Spring barley	Soybean	Sugar beet	Sunflower	Total	User accuracy
Winter barley	**16**	572	0	0	78	0	0	0	0	666	2%
Winter wheat	2	**2495**	24	0	53	1	0	0	0	2575	97%
Spring wheat	0	54	**237**	2	15	14	0	1	1	324	73%
Maize	0	0	1	**1265**	0	0	321	0	10	1597	79%
Winter rape	0	317	0	0	**555**	0	0	1	0	873	64%
Spring barley	0	118	41	0	1	**239**	0	1	0	400	60%
Soybean	0	1	0	255	0	0	**567**	0	44	867	65%
Sugar beet	0	0	0	4	0	0	0	**600**	0	604	99%
Sunflower	0	0	5	28	20	0	131	0	**587**	771	76%
Not classified	0	30	14	4	19	3	3	0	0	73	
Total	18	3587	322	1558	741	257	1022	603	642	8750	
Producer accuracy	89%	70%	74%	81%	75%	93%	55%	100%	91%		
Overall classification accuracy				**75%**							
Kappa coefficient				0.69							

Classified in satellite image as:	Ground truth: minimum NDVI to minimum NDVI										
	Winter barley	Winter wheat	Spring wheat	Maize	Winter rape	Spring barley	Soybean	Sugar beet	Sunflower	Total	User accuracy
Winter barley	**15**	190	0	0	66	0	0	0	0	271	6%
Winter wheat	3	**3105**	21	0	78	1	0	0	0	3208	97%
Spring wheat	0	55	**232**	2	23	4	0	0	1	317	73%
Maize	0	0	1	**1283**	0	0	310	0	11	1605	80%
Winter rape	0	94	0	0	**548**	0	0	1	0	643	85%
Spring barley	0	138	59	0	5	**241**	0	1	0	444	54%
Soybean	0	1	0	236	0	0	**558**	0	43	838	67%
Sugar beet	0	0	0	4	0	0	0	**601**	0	605	99%
Sunflower	0	0	5	25	20	5	127	0	**587**	769	76%
Not classified	0	4	4	8	1	6	27	0	0	50	
Total	18	3587	322	1558	741	257	1022	603	642	8750	
Producer accuracy	83%	87%	72%	82%	74%	94%	55%	100%	91%		
Overall classification accuracy				**82%**							
Kappa coefficient				0.77							

d

Classified in satellite image as:	Ground truth: Green-up to senescence				
	Maize	Soybean	Sugarcane	Total	User accuracy
Maize	**64**	145	38	247	26%
Soybean	17	**137**	74	228	60%
Sugarcane	8	52	**510**	570	89%
Not classified	16	51	28	95	
Total	105	385	650	1140	
Producer accuracy	61%	36%	78%		
Overall classification accuracy			**62%**		
Kappa coefficient			0.40		

Classified in satellite image as:	Ground truth: Green-up to dormancy				
	Maize	Soybean	Sugarcane	Total	User accuracy
Maize	**73**	67	7	147	50%
Soybean	17	**167**	6	190	88%
Sugarcane	8	66	**596**	670	89%
Not classified	7	85	41	133	
Total	105	385	650	1140	
Producer accuracy	70%	43%	92%		
Overall classification accuracy			**73%**		
Kappa coefficient			0.55		

Classified in satellite image as:	Ground truth: minimum NDVI to minimum NDVI				
	Maize	Soybean	Sugarcane	Total	User accuracy
Maize	**82**	59	12	153	54%
Soybean	7	**174**	10	191	91%
Sugarcane	5	57	**584**	646	90%
Not classified	11	95	44	150	
Total	105	385	650	1140	
Producer accuracy	78%	45%	90%		
Overall classification accuracy			**74%**		
Kappa coefficient			0.57		

Table 2.4. Confusion matrix of the post-classification analysis for green-up to senescence, green-up to dormancy and growing season of Flanders-Belgium (**a**), Sria, Russia (**b**), Kyiv, Ukraine (**c**), and Sao Paulo, Brazil (**d**). The number of correctly-classified crops, the producer accuracy, the overall accuracy and the kappa coefficient are presented.

a

		Ground truth																			
		Green-up to senescence							Green-up to dormancy							minimum NDVI to minimum NDVI						
		Grain maize	Potato	Sugar beet	Winter barley	Winter wheat	Total	User accuracy	Grain maize	Potato	Sugar beet	Winter barley	Winter wheat	Total	User accuracy	Grain maize	Potato	Sugar beet	Winter barley	Winter wheat	Total	User accuracy
Classified in satellite image	Grain maize	**1329**	292	189	1	2	1813	73%	**1384**	260	156	3	5	1808	77%	**1376**	205	135	4	5	1725	80%
	Potato	69	**1345**	93	3	16	1526	88%	59	**1738**	162	1	33	1993	87%	59	**1720**	142	1	31	1953	88%
	Sugar beet	91	227	**843**	0	6	1167	72%	84	132	**831**	0	13	1060	78%	81	72	**880**	0	17	1050	84%
	Winter barley	8	23	2	**340**	304	677	50%	3	0	0	**473**	421	897	53%	3	0	1	**470**	260	734	64%
	Winter wheat	17	82	12	130	**2870**	3111	92%	6	10	7	39	**2748**	2810	98%	1	13	5	38	**3023**	3080	98%
	Not classified	69	285	72	44	255	725		47	114	55	2	233	451		63	244	48	5	117	477	
	Total	1583	2254	1211	518	3453	9019		1583	2254	1211	518	3453	9019		1583	2254	1211	518	3453	9019	
	Producer accuracy	84%	60%	70%	66%	83%			87%	77%	69%	91%	80%			87%	76%	73%	91%	88%		
Overall classification accuracy					75%							80%							83%			
Kappa coefficient					0.67							0.73							0.77			

b

		Ground truth																									
		Green-up to senescence											Green-up to dormancy											minimum NDVI to minimum NDVI										
		Flax	Maize	Peas	Soybean	Spring barley	Sugar beet	Sunflower	Winter barley	Winter rape	Total	User accuracy	Flax	Maize	Peas	Soybean	Spring barley	Sugar beet	Sunflower	Winter barley	Winter rape	Total	User accuracy	Flax	Maize	Peas	Soybean	Spring barley	Sugar beet	Sunflower	Winter barley	Winter rape	Total	User accuracy
Classified in satellite image as:	Flax	**593**	0	140	0	0	0	0	0	0	733	81%	**461**	0	0	38	10	0	0	0	0	509	91%	**684**	0	0	38	0	0	0	0	0	722	95%
	Maize	0	**321**	0	0	0	0	115	0	0	436	74%	0	**321**	0	0	0	0	115	0	0	436	74%	0	**350**	0	0	0	0	102	0	0	452	77%
	Peas	58	0	**155**	38	0	0	0	0	0	251	62%	338	0	**316**	0	0	0	0	0	0	654	48%	132	0	**316**	0	0	0	0	0	0	448	71%
	Soybean	175	0	0	**121**	0	0	65	0	0	361	34%	74	0	0	**121**	0	0	65	0	0	260	47%	74	0	0	**121**	0	0	41	0	0	236	51%
	Spring barley	0	0	21	0	**71**	0	0	0	98	190	37%	0	0	0	0	**61**	0	0	0	30	91	67%	10	0	0	0	**71**	0	0	0	0	81	88%
	Sugar beet	35	0	0	0	0	**49**	77	0	0	161	30%	0	0	0	0	0	**49**	48	0	0	97	51%	0	0	0	0	0	**49**	48	0	0	97	51%
	Sunflower	34	63	0	0	0	0	**352**	0	0	449	78%	24	63	0	0	0	0	**381**	0	0	468	81%	0	34	0	0	0	0	**381**	0	0	415	92%
	Winter barley	0	0	0	0	0	0	0	**846**	76	922	92%	0	0	0	0	0	0	0	**883**	0	883	100%	0	0	0	0	0	0	0	**943**	0	943	100%
	Winter rape	0	0	0	0	0	0	0	161	**541**	702	77%	0	0	0	0	0	0	0	124	**685**	809	85%	0	0	0	0	0	0	0	64	**715**	779	92%
	Not classified	12	0	0	0	0	0	0	0	0	12		10	0	0	0	0	0	0	0	0	10		7	0	0	0	0	0	37	0	0	44	
	Total	907	384	316	159	71	49	609	1007	715	4217		907	384	316	159	71	49	609	1007	715	4217		907	384	316	159	71	49	609	1007	715	4217	
	Producer accuracy	65%	84%	49%	76%	100%	100%	58%	84%	76%			51%	84%	100%	76%	86%	100%	63%	88%	96%			75%	91%	100%	76%	100%	100%	63%	94%	100%		
Overall classification accuracy					**72%**											**78%**											**86%**							
Kappa coefficient					0.67											0.74											0.84							

c

Classified in satellite image as:	Ground truth: Green-up to senescence										
	Winter barley	Winter wheat	Spring wheat	Maize	Winter rape	Spring barley	Soybean	Sugar beet	Sunflower	Total	User accuracy
Winter barley	**12**	607	0	0	405	0	27	0	0	1051	1%
Winter wheat	6	**2871**	49	0	84	0	0	0	0	3010	95%
Spring wheat	0	0	**91**	1	1	15	0	0	0	108	84%
Maize	0	0	1	**1282**	0	0	284	0	9	1576	81%
Winter rape	0	0	0	0	**195**	0	0	0	0	195	100%
Spring barley	0	107	181	0	0	**242**	0	0	0	530	46%
Soybean	0	0	0	72	0	0	**332**	0	49	453	73%
Sugar beet	0	0	0	34	0	0	0	**603**	0	637	95%
Sunflower	0	0	0	108	42	0	252	0	**554**	956	58%
Not classified	0	2	0	61	14	0	127	0	30	234	
Total	18	3587	322	1558	741	257	1022	603	642	8750	
Producer accuracy	67%	80%	28%	82%	26%	94%	32%	100%	86%		
Overall classification accuracy				**71%**							
Kappa coefficient				0.63							

Classified in satellite image as:	Ground truth: Green-up to dormancy										
	Winter barley	Winter wheat	Spring wheat	Maize	Winter rape	Spring barley	Soybean	Sugar beet	Sunflower	Total	User accuracy
Winter barley	**18**	494	0	0	72	0	0	0	0	584	3%
Winter wheat	0	**2684**	15	0	2	0	0	0	0	2701	99%
Spring wheat	0	0	**221**	0	1	15	0	0	0	237	93%
Maize	0	0	1	**1305**	0	0	338	0	9	1653	79%
Winter rape	0	242	0	0	**602**	0	0	0	0	844	71%
Spring barley	0	159	70	0	0	**242**	0	0	0	471	51%
Soybean	0	0	0	216	0	0	**542**	0	35	793	68%
Sugar beet	0	0	0	0	0	0	0	**603**	0	603	100%
Sunflower	0	0	4	33	42	0	126	0	**598**	803	74%
Not classified	0	8	11	4	22	0	16	0	0	61	
Total	18	3587	322	1558	741	257	1022	603	642	8750	
Producer accuracy	100%	75%	69%	84%	81%	94%	53%	100%	93%		
Overall classification accuracy				**78%**							
Kappa coefficient				0.73							

Classified in satellite image as:	Ground truth: minimum NDVI to minimum NDVI										
	Winter barley	Winter wheat	Spring wheat	Maize	Winter rape	Spring barley	Soybean	Sugar beet	Sunflower	Total	User accuracy
Winter barley	**12**	73	0	0	26	0	0	0	0	111	11%
Winter wheat	0	**3342**	14	0	27	0	0	0	0	3383	99%
Spring wheat	0	0	**221**	0	43	0	0	0	0	264	84%
Maize	0	0	1	**1339**	0	0	364	0	9	1713	78%
Winter rape	0	170	78	0	**635**	0	0	0	0	883	72%
Spring barley	0	0	0	0	0	**242**	0	0	0	242	100%
Soybean	0	0	0	182	0	0	**529**	0	35	746	71%
Sugar beet	0	0	0	0	0	0	0	**603**	0	603	100%
Sunflower	0	0	0	33	0	12	123	0	**598**	766	78%
Not classified	6	2	8	4	10	3	6	0	0	39	
Total	18	3587	322	1558	741	257	1022	603	642	8750	
Producer accuracy	67%	93%	69%	86%	86%	94%	52%	100%	93%		
Overall classification accuracy				**86%**							
Kappa coefficient				0.82							

d

Classified in satellite image as:	Ground truth: Green-up to senescence				
	Maize	Soybean	Sugarcane	Total	User accuracy
Maize	**69**	153	36	258	27%
Soybean	24	**145**	42	211	69%
Sugarcane	8	74	**525**	607	86%
Not classified	4	13	47	64	
Total	105	385	650	1140	
Producer accuracy	66%	38%	81%		
Overall classification accuracy			**65%**		
Kappa coefficient			0.43		

Classified in satellite image as:	Ground truth: Green-up to dormancy				
	Maize	Soybean	Sugarcane	Total	User accuracy
Maize	**75**	72	8	155	48%
Soybean	21	**200**	13	234	85%
Sugarcane	8	92	**625**	725	86%
Not classified	1	21	4	26	
Total	105	385	650	1140	
Producer accuracy	71%	52%	96%		
Overall classification accuracy			**79%**		
Kappa coefficient			0.62		

Classified in satellite image as:	Ground truth: minimum NDVI to minimum NDVI				
	Maize	Soybean	Sugarcane	Total	User accuracy
Maize	**88**	50	7	145	61%
Soybean	4	**200**	7	211	95%
Sugarcane	5	91	**594**	690	86%
Not classified	8	44	42	94	
Total	105	385	650	1140	
Producer accuracy	84%	52%	91%		
Overall classification accuracy			**77%**		
Kappa coefficient			0.61		

Figure 2.5 illustrates the potential of PROBA-V 100 m data for crop identification mapping in the study areas for selected sites. The results are shown for the pure pixels from the beginning to the end of the growing season. In general, the fields have been classified correctly for both sites. However, in Flanders (Belgium), a in Figure 2.5, some sugar beet fields have been classified as potato fields and winter wheat fields as winter barley. In Sria (Russia), b in Figure 2.5, some maize fields have been classified as sunflower and vice versa. Likewise, sunflower fields have been classified as sugar beet and winter barley as winter rape.

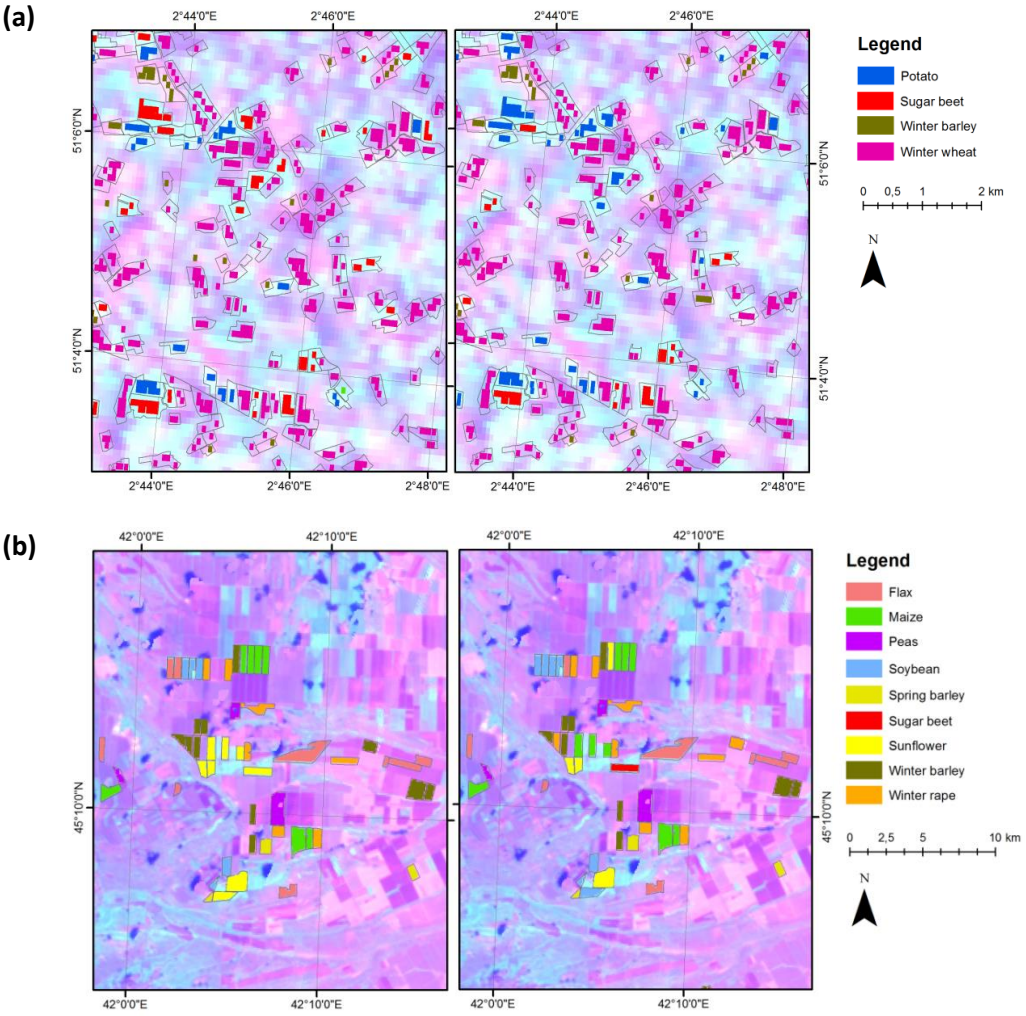


Figure 2.5. Comparison of the classification results based on pure pixels during the entire growing season for a selected area in Flanders-Belgium (a) and Sria, Russia (b). (left) The overlay of PROBA-V and ground data; (right) the overlay with post-classification results.

2.5. Discussion

This study demonstrated the suitability of spectral matching techniques (SMTs) for mapping crop types using 100 m PROBA-V data for the 2014–2015 season. The methodology integrated multi-temporal satellite imagery and parcel boundaries retrieved from the SIGMA project and 'Geo-Data Infrastructure-Flanders' databases. The SMTs were ideal for analyzing remote sensing time series data during the crop growth period. Spectral similarity values (SSV) were calculated. SSVs are measures of the shape and

magnitude similarities of the time series spectra and found the most useful SMTs, similar to Thenkabail et al. [183]. Subsequently, SMTs were applied to match the ideal spectra, i.e., the reference NDVI profiles, to the class spectra, i.e., the individual pure pixel NDVI profiles.

The methodology demonstrated that 100 m PROBA-V has the potential to be used in crop identification across different regions in the world. As a next step, a regular classification method could be applied to PROBA-V data to show whether a regular classification method could give similar results as provided in this study. This would show whether the results were obtained due to the applied SSVs, the sensor itself or both. PROBA-V is a relatively new satellite, and therefore, there are limited studies available for crop identification mapping. The work in Roumenina et al. [47] reported crop identification accuracies in the range of 72.4%–86.2% for 100 m PROBA-V data for mapping summer and winter crops in Bulgaria. In another study, Lambert et al. [208] achieved an overall accuracy of 84% using the 100 m PROBA-V sensor for cropland mapping of Sahelian and Sudanian agro-ecosystems. These reported ranges are in line with the results presented in this study. When using post-classification, the overall accuracy (%) ranged between 65 and 86, and the kappa coefficient changed from 0.43–0.84. In general, post-classification improved the overall accuracy results around an additional 6%–7% for Ukraine and Brazil and 11% for Russia compared to the initial classification results. For Belgium, the post-classification technique did not improve the classification results (see Table 2.3 and Table 2.4). The results are best in Sria, Russia, followed by Kyiv, Ukraine, Flanders-Belgium and Sao Paulo, Brazil. A couple of reasons could explain the differences between accuracies across the different study areas. Firstly, better results were observed in the areas where the crop phenological development was not spread over a long time period. For instance, the planting time for maize in Brazil stretched from August–December with a period of highest activity in October and November. This prevented extracting the distinctive characteristic of the reference NDVI profiles. Secondly, the parcel sizes played an important role. When parcels covered a small number of satellite pixels, the results were less accurate, as was the case for Belgium. Thirdly, classification errors of crop types increased when the time window covered only part of the cropping period. Another reason behind the classification errors is related to the number of ground-truth parcels available from the study site, as is the case for the winter barley fields in Kyiv, Ukraine, compared to other crop types in the same site. Finally, crops with similar growing periods might cause classification errors, such as sunflower and maize in Sria, Russia. In addition, the extent of the study area played a role. Accuracies potentially improved when region specific NDVI reference profiles were included from different agro-ecological regions. Based on these results, crop identification mapping was challenging, but the use of 100 m PROBA-V proved a valid option even when mapping at the field level.

Our results were in close agreement with other studies that used different classification methods and/or other higher resolution satellite images. A multi-temporal sequence of 100 m PROBA-V images was used covering one to two growing seasons. The work in Aurdal et al. [170] reported an overall accuracy of 63% for vegetation mapping in southern Norway using 25 m resolution Landsat images. Another similar study reported an overall accuracy of 62.7% using the NDVI temporal profiles approach and 72.8% using a maximum likelihood classifier in the northeast of Germany with phenological information and spectral-temporal profiles from Landsat TM/ETM [178]. The use of multiple sensors seemed to increase the accuracy. For instance, De Wit and Clevers [201] updated the crop classification in the land cover database of The Netherlands by combining Landsat TM, Indian Remote Sensing Satellite - Linear Imaging Self Scanner and European remote sensing satellite 2 - synthetic aperture radar and reported an overall accuracy value of

90%. Almost one million pixels were used at the national level covering not only the different types of cereals, but also grassland and flower bulbs. The use of homogeneous pixels improved the classification accuracy. The overall accuracy ranged from 73% for very heterogeneous pixels to 89% for homogeneous pixels in North Carolina and Virginia with 250 m MODIS NDVI [209]. The number of homogenous pixels used in their study was 1014, which included 475 pixels for agriculture. Specific crop identification mapping results per-field were presented. In another study, both per-field and per-area results were presented. The work in Zhong et al. [210] reported a maximum overall accuracy of 66% and a kappa coefficient of 0.60 per field and a maximum overall accuracy of 70% and a kappa coefficient of 0.64 per area for mapping specific crop types in Central Valley of California based on the time series of Landsat TM/ETM+.

Although the method showed promising results in crop identification, a number of limitations were identified. The reference NDVI profiles for the growing season of each crop type had to be defined in advance, either based on ground data, on user knowledge of the field or on a literature review. Another limiting factor occurred when the parcel size was smaller than the pixel size. Having larger parcel sizes than pixel sizes was an advantage, particularly because pure pixels tremendously improved the classification results.

The maps based on the methodology could be extended to regional or national-level crop production estimations and all crop types of interest. It is shown that the within-field spectral variability could be reduced with accurate field boundaries. These boundaries eliminated classification errors due to mixed pixels [201]. Object-based image analysis could enable the detection of field boundaries in regions without parcel information. To this extent, Zhong et al. [210] used image segmentation to delineate the field borders prior to classification.

2.6. Conclusions

This study demonstrated the potential of phenology-based crop identification mapping at the global level using adapted spectral matching techniques (SMTs) applied to multi-temporal 100 m PROBA-V images for the 2014–2015 season. Phenological metrics were extracted from NDVI time series using piecewise logistic functions. These metrics represented the crop growing seasons and identified the specific calendar of each crop type. The spectral matching was applied on all NDVI values within each window. A distinct advantage of the SMTs was their simplicity and ease of application. In addition, the method can be extended to other areas based on the reference NDVI profiles, which are predefined either by ground data, field knowledge or literature review. The crop classification accuracies obtained could be compared favorably to the results derived from classifications with higher resolution data. The overall accuracy ranged between 65% and 86%, and the kappa coefficient varied between 0.43 and 0.84 depending on the site and the temporal window used.

Chapter 1 explored how CGLS-DMP that was developed for SPOT-VGT, and transferred to PROBA-V, can be improved to more closely relate to yield anomalies. In Chapter 1, details of the modified DMP and its potential in agricultural applications were presented. DMP can be a better BP compared to simpler BPs because it is a more robust model as being more responsive to changes in climatic conditions and being sensitive to variations in respiration costs. Presence of a crop specific map was one of the limiting factors for yield estimation analyses in Chapter 1. Chapter 2 showed how PROBA-V can be used effectively for crop identification mapping using SMTs. New technological improvements have added another dimension to crop yield estimation studies: finer spatial resolution data with more frequent

revisit times. Rembold et al. [11] discuss that there is a general link between higher spatial resolution and the quality of yield forecasts although the relationship is not strictly linear, particularly due to the field sizes. There is another reason of nonlinearity in the relationship. When the spatial resolution of the satellite images increases, the temporal resolution usually decreases due to the more limited revisit time. Chapter 3 explores the trade-off between the different spatial resolutions provided by PROBA-V products versus the temporal frequency and, additionally, explores the use of thermal time to improve statistical yield estimation. The study was done at the field level; however, the field level data can be aggregated to regional level. When the analysis was done in Chapter 3, the only available operational PROBA-V product was NDVI. Since there was no fAPAR product at that time, DMP could not be calculated and the comparison of NDVI, fAPAR and DMP could not be done in Chapter 3.

3 Comparison of PROBA-V 100 m, 300 m, and 1 km NDVI datasets for yield forecasting at the field level[3]

Abstract

PROBA-V has delivered images for retrieving vegetation and land surface characteristics at a spatial resolution of 100 m, 300 m, and 1 km since 2013. This study explores the suitability of PROBA-V NDVI products at 5-daily 100 m resolution, daily 300 m resolution and daily 1 km resolution to estimate winter wheat yields at the field level for 39 fields across Northern France during one growing season from 2014 to 2015. An asymmetric double sigmoid function is fitted, and the NDVI values are integrated over thermal time and over calendar time for the central pixel of the field, exploring different NDVI thresholds, to mark the start and end of the cropping season. The integrated values with different NDVI thresholds are used as a proxy for yield. In addition, a pixel purity analysis is performed for different purity thresholds at 100 m, 300 m, and 1 km resolutions. A simple linear regression is calculated between the integrated NDVI values and yield data collected at the field level. Depending on the resolution and the NDVI threshold used, the adjusted R^2 ranges from 0.20 to 0.74; jackknifed – leave-one-field-out cross validation – Root Mean Squared Error (RMSE) ranges from 0.6 to 1.07 t/ha and Mean Absolute Error (MAE) ranges between 0.46 and 0.90 t/ha for thermal time analysis. The best results (adjusted R^2 = 0.74, RMSE = 0.6, and MAE = 0.46) are obtained with a 100 m resolution at 0.2 NDVI threshold integrated over thermal time with no purity threshold. The findings demonstrate that winter wheat yields can be estimated at the field level with pure pixels from PROBA-V products.

3.1. Introduction

Accurate and timely yield estimates are important due to their large impact on strategic planning and world markets [211]. Extensive research has been done over the past decades to apply remote sensing for predicting yields at different scales, from field to national levels. Therefore, these studies are crucial for individual farmers, national governments, and international organizations. Satellite images have different spectral, temporal, and spatial resolutions. The images most commonly used for regional studies have a coarse resolution of 1 km and a daily revisit time. Time-series data from the AVHRR, SPOT-VGT, MODIS, and MERIS are examples of low-resolution data. Low-resolution images have been used to establish statistical regressions between VIs and crop yield statistics [43,211]. The advantage of low-resolution data is their long historical time series, high temporal resolution, low cost,

[3] Manuscript is in preparation for publication.

and global coverage which allows following vegetation development even when overcast conditions frequently occur [11,41,212,213]. Despite their valuable input for assessing crop production, the most relevant limitation of coarse-resolution time-series data is that their spectral reflectance measurements contain mixed information from several surfaces and vegetation types, making it difficult to interpret the signal and directly relate a spectral or temporal signature to a specific crop condition [213]. Images with higher spatial resolutions, every 5 days and daily, are available from PROBA-V 100 m and 300 m data. The advantage of the 100 m dataset compared to higher-resolution satellite data is that its spatial resolution is sufficient for yield estimate studies, thereby avoiding larger processing times of heavier data volumes, as is the case for Sentinel-2 and Landsat [214]. The PROBA-V time series became available in May 2013, and, as a result, few studies exist that relate PROBA-V data to crop yield [46,215]. The comparison of PROBA-V data at different resolutions – 100 m, 300 m, and 1 km – for crop yield estimation has not been studied to date.

From satellite observations, the vegetation signal can be captured through mathematical combinations of reflectance at different spectral bands into VIs, such as the NDVI [43]. Several studies have demonstrated that seasonal VIs are significantly correlated with official crop yield statistics [21,211,216–219]. Consequently, remote sensing is an attractive tool for monitoring spatio-temporal patterns of vegetation growth [43]. Measures of green vegetation provided by VIs can be used to estimate yield. Using SPOT-VGT imagery at 1 km, Meroni et al. [220] computed the integral between the start of the growing season and the beginning of the descending stage for winter wheat yield estimates on a national level and discovered a significant correlation.

Vegetation indices can be expressed as a function of thermal time instead of calendar time to derive more crop physiologically sound relations. Crop physiological development in relation to thermal time is a well-known modeling method used in climate impact research [16]. Plants require certain accumulated heat energy over the growing season to develop. Thermal time for organ growth and development depends on the temperature [139,220,221] and can be defined as the cumulative daily average temperature throughout the growing period, expressed as the Growing Degree Day (GDD) [222,223]. The GDD has been used to obtain smoother and more temporally consistent time series of biophysical variables [221]. Lobell et al. [224] have used GDD rather than calendar days to account for changes in crop development due to accumulated temperature. Skakun et al. [39] have used GDD to account for discrepancies which are temporally and spatially non-uniformities of the winter crop development due to the presence of different agro-climatic zones.

Several studies have demonstrated that pixel purity can lead to better results for characterizing crop growth [66,225–227]. Pixel purity is the degree of homogeneity of the signal with respect to the target crop, and it allows explicit control over the relationship between the remote sensing observation and the target [5]. Pixel purity represents the relative contribution of the surface of interest to the signal detected by the remote sensing instrument [225]. This approach has been used to restrict the analysis to a subset of a region's pixels and to identify the maximum tolerable pixel size for both crop growth monitoring and crop area estimation [214,227].

Comparing PROBA-V datasets at 100 m, 300 m, and 1 km resolutions leads to the hypothesis that the higher the spatial resolution, the greater the correlation between NDVI time series and observed yields at the field level. In addition to spatial resolution, temporal resolution of the data obtained from satellite products could affect the accuracy of crop yield estimates. The main objective of this study is to compare the 5-daily and daily PROBA-V NDVI datasets at 100 m, 300 m, and 1 km resolutions to estimate wheat yield at the field level. The results are investigated in terms of pixel sizes and purity. The model performance is evaluated using statistical metrics based on the comparison between wheat yields

measured at the field scale across different regions of Northern France and the time integral of PROBA-V NDVI time series at different resolutions during the 2014–2015 growing season.

3.2. Materials

3.2.1. Study Area and Ground Data

The study area encompassed Nord-Pas-de Calais, Picardie and Champagne-Ardenne in Northern France (Figure 3.1). These regions are amongst the highest wheat-producing regions of France.

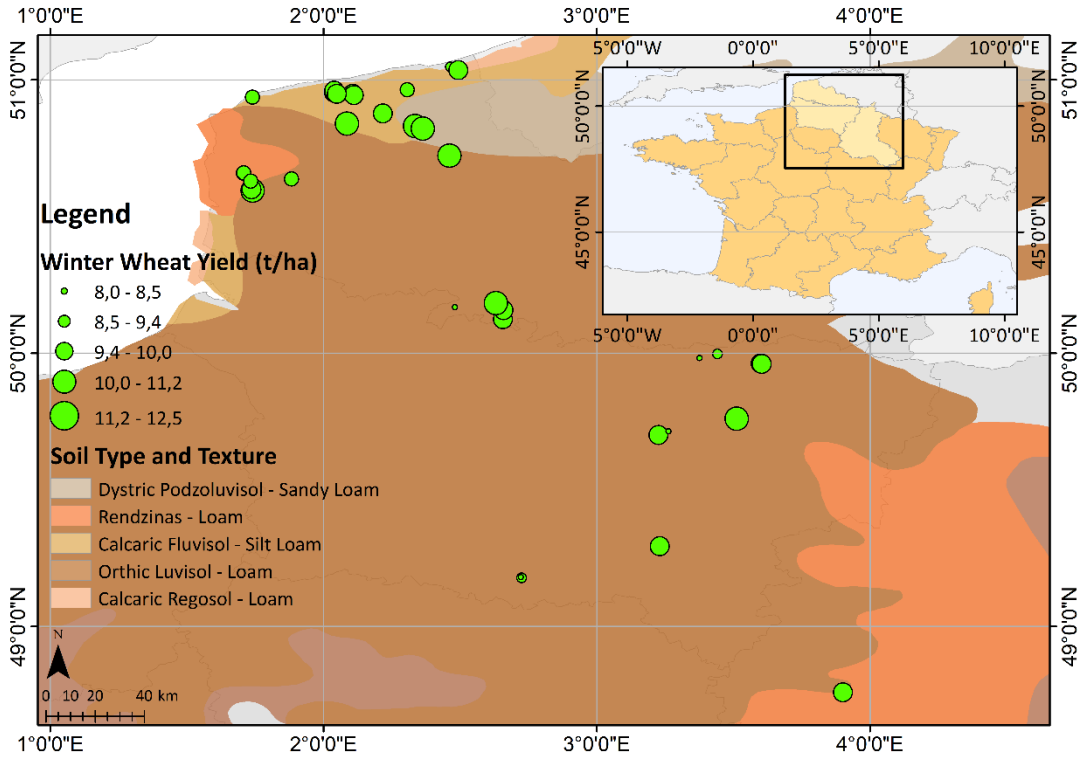


Figure 3.1. Location of winter wheat fields, corresponding yields (t/ha), and soil type of the study area.

Ground data, field boundaries and winter wheat yields were obtained from farmers' declarations during the 2015 crop campaign by a private company called Drone Agricole. The sowing dates were in October and November, and the harvest took place in July and August. After applying field selection thresholds further explained in the Methods section, 39 winter wheat fields from a total of 56 fields were selected with sizes ranging from 8 ha to 12.5 ha. Average yield for the area was around 10 t/ha with a standard deviation of 1.2 t/ha. Neighboring fields can have different yields which can be explained by different biophysical environments, such as soil type and different management strategies relating to cultivar, sowing time, and agri-chemical application rates.

3.2.2. NDVI and Meteorological Data

PROBA-V was launched in May 2013 to fill the gap between the SPOT-VGT and Sentinel-3 satellites. PROBA-V has four spectral bands: blue, centered at 0.463 µm; red, 0.655 µm; NIR, 0.845 µm; and SWIR, 1.600 µm. The Vegetation-PROBA instrument has a swath width of 2250 km across four bands. The central camera of the PROBA-V satellite

provides a 100 m data product delivering global coverage every 5 days, whereas global daily images are acquired at 300 m and 1 km resolutions. Non-composited (S1) atmospherically corrected PROBA-V NDVI images and status maps at 100 m, 300 m, and 1 km resolutions were obtained from http://www.vito-eodata.be. Status maps, containing information about snow, ice, shadow, clouds, and land or sea for every pixel, were used to extract high quality pixels and NDVI series.

Meteorological input, daily minimum and maximum temperatures, were available at a 0.25 x 0.25 degree grid from the JRC-MARSOP project [106]. The meteorological data acquired through that project are regularly collected from ground stations in near real time and provided directly by national meteorological institutes or regional authorities.

3.3. Methods

NDVI values at 100 m, 300 m, and 1 km resolutions were extracted for the central pixel of each field where yield data were available. The number of NDVI observations during the growing season can severely affect the integral calculated after curve fitting. Because of its lower revisit time, the 100 m dataset had a poorer temporal resolution compared to the 300 m and 1 km resolution datasets. Two field selection thresholds were, therefore, applied to retain the fields for which sufficient data were available for the three datasets. The selection criteria were: (1) a minimum of 10 NDVI observations from PROBA-V during the growing season and (2) less than two-month gaps between two consecutive observations. Based on these criteria, 39 winter wheat fields out of 56 fields were selected for further analysis.

Pixel purity percentages were investigated for the three resolution datasets to study the effects. The same purity thresholds were defined for all datasets, although the differences between the resolution and associated sample sizes were present. For instance, for the 1 km dataset, the number of samples within the fields was too low for a 20% purity threshold, whereas for the 100 m dataset, almost all of the samples had more than 95% purity.

Temperature values were extracted from the JRC-MARSOP database [106], and thermal time was computed in cumulative GDD (Equation (3.1)).

$$GDD = \frac{T_{max}+T_{min}}{2} - T_{base} \quad (3.1)$$

where T_{max} is the maximum temperature, T_{min} is the minimum temperature, and T_{base} is the base temperature. The base temperature for wheat was 0 °C, as used in Duveiller's study [228]. All fields in the study site belonged to the same climatic region. The period used for the calculation of cumulative GDD for all fields was from the beginning of January until the end of July.

To monitor vegetation growth, remotely sensed imagery time series should be gap-free ideally. However, in reality, the time series are irregular, and relationships are sought to explain growth during the season. An asymmetric double sigmoid function (ADSF) was fitted to the NDVI time series of the central pixel of each field for thermal time and for calendar time. As applied in the study of Zhong et al. [229], the ADSF was used for curve fitting and expressed as follows in Equation (3.2):

$$V(t) = V_b + \frac{1}{2}V_a\left[tanh\left(p(t-D_i)\right) - tanh(q(t-D_d))\right] \quad (3.2)$$

where $V(t)$ is the NDVI at the time t. V_b is the background NDVI value corresponding to the non-growing season. V_a is the amplitude of NDVI variation within the current growing cycle.

The overall changing rates of the slopes were characterized by p and q. Middle dates of the increasing and the decreasing segments were represented by D_i and D_d.

The fitted curves for each field were integrated at the 100 m, 300 m, and 1 km resolutions using seven NDVI thresholds in the range 0–0.6. A linear regression was calculated between yield data and the integral values of the fitted ADSF for each field. The statistical metrics, namely the adjusted coefficient of determination (R^2), Pearson's correlation's (r^2) associated p-value, jackknifed (leave-one-field-out cross validation) RMSE (t/ha), and MAE (t/ha), were used to evaluate the goodness of fit.

3.4. Results

The fitted ADSF integral areas demonstrated a higher goodness of fit with an increase in pixel purity and resolution of the PROBA-V NDVI datasets at 100 m, 300 m, and 1 km resolutions for thermal time and calendar time as shown for one field with a 0.3 NDVI threshold in Figure 3.2.

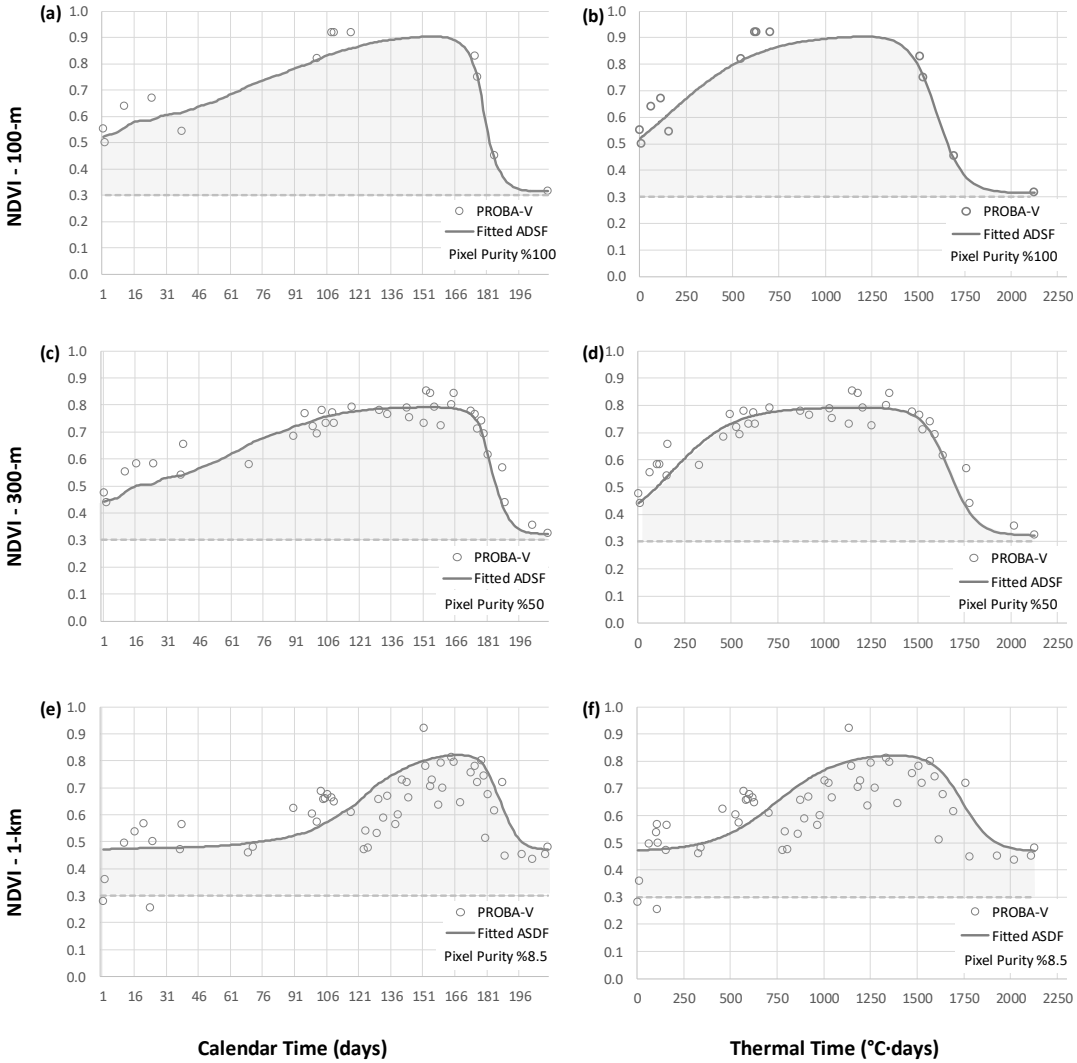


Figure 3.2. The depiction is an example of fitted ADSF and integral area of PROBA-V NDVI datasets at 100 m, 300 m, and 1 km resolutions for calendar time (days) and thermal time (°C·days). The circles refer to NDVI values for the field. The grey line represents the fitted ADSF curve, and the shaded area represents the integral above the NDVI threshold 0.3. The example field was located in Nord-Pas-de Calais (50.879°N, 2.218°E).

For all sampled fields, the results proved that the 1 km and 300 m datasets exhibited a lower performance due to their spectral reflectance, which contains mixed information from several surface types, as compared to the 100 m dataset. Figure 3.3 indicates that the 100 m dataset performed better than the 300 m NDVI which, in turn, was performed better than the 1 km for both the thermal time and calendar time analyses. The thermal time performed systematically better, except for threshold – 0.5 and 0.6 for 300 m. The optimal NDVI threshold was around 0.2, and a higher threshold value reduced the performance.

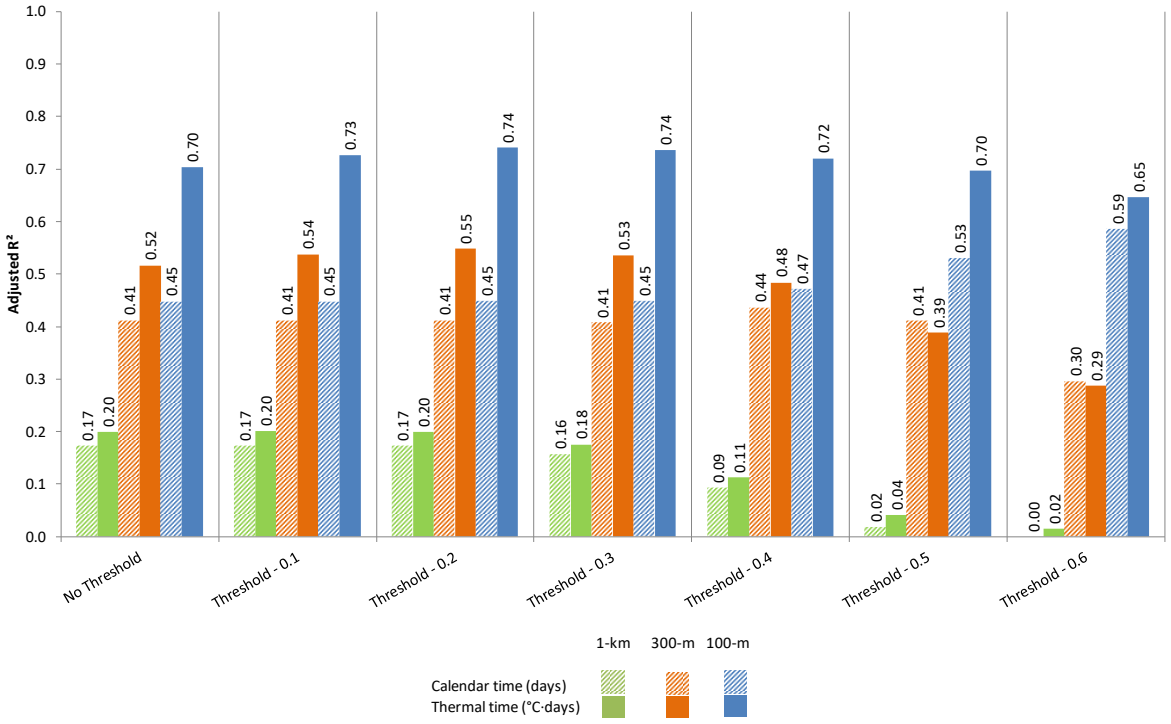


Figure 3.3. The comparison of calendar time (days) and thermal time (°C·days) using adjusted R^2 results at 100 m, 300 m, and 1 km resolutions for different NDVI thresholds for all fields, with no purity threshold applied.

Overall, for all fields, the 100 m dataset had a greater performance by 32% in RMSE compared to the 300 m dataset which had a greater performance by 36% compared to the 1 km NDVI dataset with an NDVI threshold of 0.2 and integrated over thermal time. The 100 m dataset was better because it included higher pixel purities. Compared to calendar time, integration over thermal time exhibited a greater performance by 15% for 1 km, 33% for 300 m, and 65% for 100 m for an NDVI threshold of 0.2 in adjusted R^2 (Figure 3.3).

The retrieval results were presented against both the actual yield and the integral values at 100 m, 300 m, and 1 km resolutions for different NDVI thresholds in the scatterplots in Figure 3.4. There were noticeable differences between the three resolution results. In addition, deviation from the 1:1 line was less pronounced for thermal time as compared to calendar time.

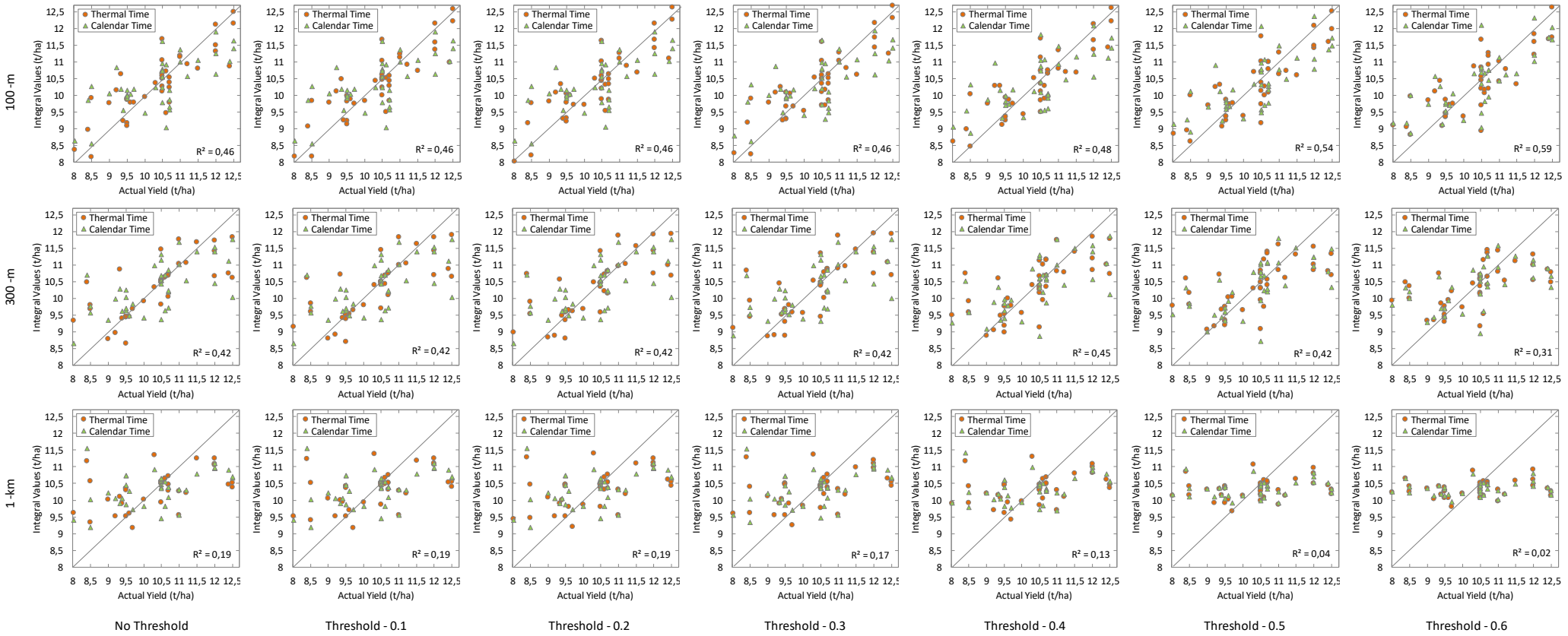


Figure 3.4. Scatterplots comparing the actual yield with integral values at 100 m, 300 m, and 1 km resolutions for different NDVI thresholds.

The relationships between yield data and the integrated NDVI based on ADSF was statistically significant. In all cases, *p*-values were smaller than 0.001, except for the *p*-values for NDVI thresholds of 0.5 and 0.6 at the 1 km resolution, which were larger than 0.05 (Figure 3.5.). Integration over thermal time with a 0.2 NDVI threshold gave the strongest results at the 100 m, 300 m, and 1 km resolutions.

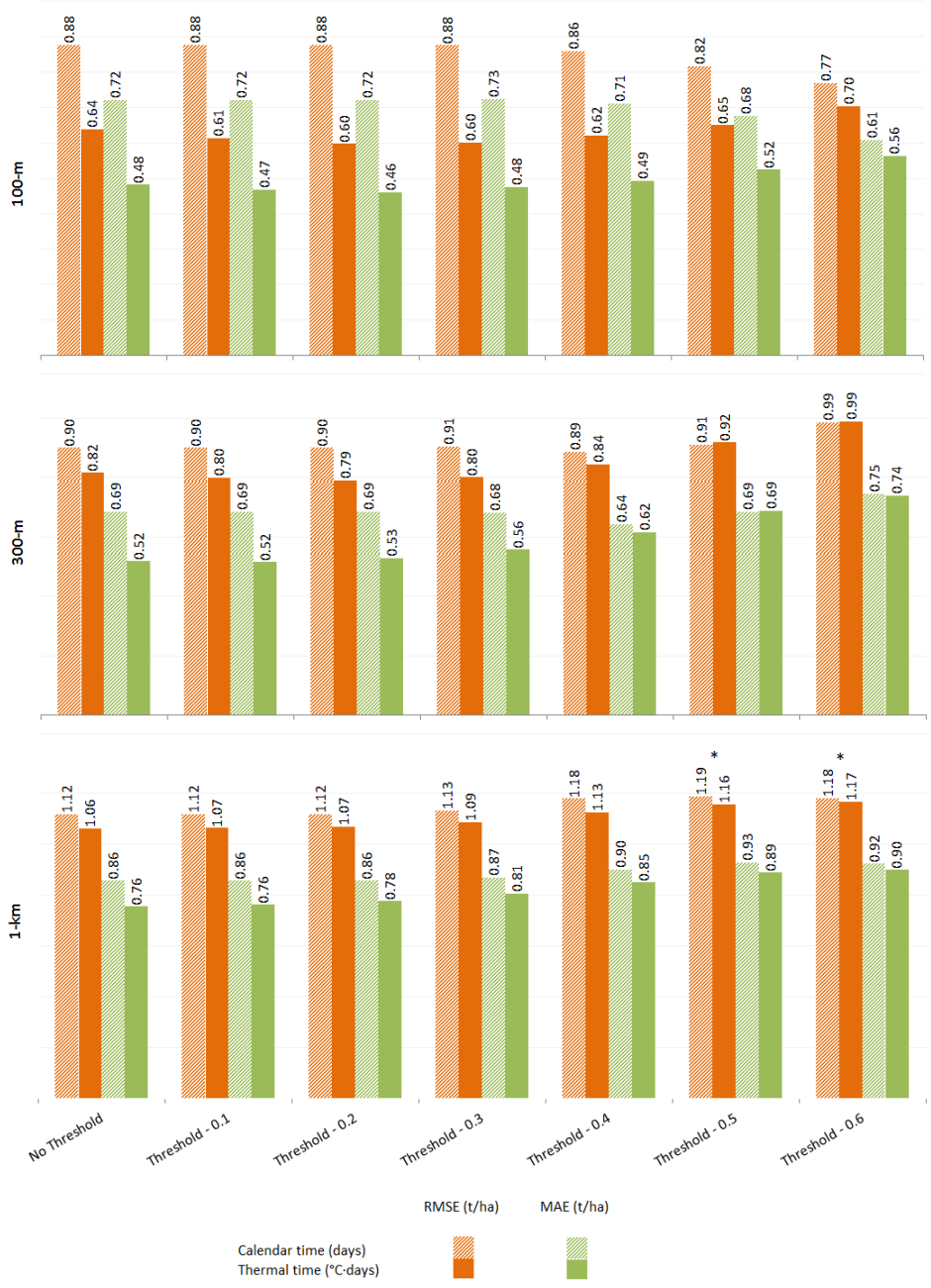


Figure 3.5. Jackknifed RMSE (t/ha) and MAE (t/ha) for calendar time (in days) and thermal time (in °C·days) analyses at 100 m, 300 m, and 1 km resolutions for different NDVI thresholds with *p*-values smaller than 0.001 in all cases, except for the ones labeled with a star symbol, which were larger than 0.05.

Pixel purity percentages increased with resolution. The pixel purity percentage ranged between 61% to 100% for the 100 m dataset, between 14% to 100% for the 300 m dataset, and between 2% to 51% for the 1 km dataset. The same purity thresholds were used for pixel purity comparisons of calendar time (days) and thermal time (°C·days) at 100 m, 300 m, and 1 km resolutions for different NDVI thresholds (Table 3.1).

Table 3.1 Adjusted R^2 results for pixel purity comparison for calendar time (days) and thermal time (°C·days) analyses at 100 m, 300 m, and 1 km resolutions for different NDVI thresholds (ranging from 0 to 0.3) and pixel purity thresholds (ranging from 0% to 90%).

		100 m										300 m										1 km									
		Calendar Time					Thermal Time					Calendar Time					Thermal Time					Calendar Time					Thermal Time				
Pixel purity (%)		0	15	40	65	90	0	15	40	65	90	0	15	40	65	90	0	15	40	65	90	0	15	40	65	90	0	15	40	65	90
Number of fields		39	39	39	37	34	39	39	39	37	34	39	37	27	12	5	39	37	27	12	5	39	9	1	0	0	39	9	1	0	0
NDVI threshold	No	0.4	0.4	0.4	0.5	0.6	0.7	0.7	0.7	0.7	0.7	0.4	0.4	0.5	0.8	0.9	0.5	0.5	0.6	0.6	0.6	0.2	0.3	N/A	N/A	N/A	0.2	0.3	N/A	N/A	N/A
	0.1	0.4	0.4	0.4	0.5	0.6	0.7	0.7	0.7	0.7	0.8	0.4	0.4	0.5	0.8	0.9	0.5	0.5	0.6	0.6	0.6	0.2	0.3	N/A	N/A	N/A	0.2	0.4	N/A	N/A	N/A
	0.2	0.4	0.4	0.4	0.5	0.5	0.7	0.7	0.7	0.7	0.8	0.4	0.4	0.5	0.8	0.9	0.5	0.5	0.6	0.7	0.7	0.2	0.3	N/A	N/A	N/A	0.2	0.4	N/A	N/A	N/A
	0.3	0.4	0.4	0.4	0.5	0.5	0.7	0.7	0.7	0.7	0.8	0.4	0.4	0.5	0.8	0.9	0.5	0.5	0.6	0.6	0.6	0.2	0.2	N/A	N/A	N/A	0.2	0.3	N/A	N/A	N/A

The adjusted R^2 values (Table 3.1) illustrated that with higher purity, there was a higher correlation between the integral value and the yield. This result demonstrated that high quality yield estimates were possible to obtain even at lower resolutions that had a higher pixel purity. However, while pixel purity percentages increased, the number of fields retained for analysis decreased. The rate of reduction in sample size was higher for the 300 m dataset and highest for the 1 km dataset. The bigger sample sizes resulted in more accurate results. For instance, the values increased comparably more for purer pixels for calendar time (days) in the 300 m resolution dataset. The reason behind this could be that the number of available sample sizes in the correlation increased the margin of error.

3.5. Discussion

This study reveals that the fitted ADSF integral area of NDVI values using thermal time provides a better estimate for yield than calendar time on the field scale. Therefore, meteorological maps with appropriate accuracy are necessary to make accurate thermal time calculations. The results of this study demonstrate that the PROBA-V 100 m resolution provided more accurate estimates of wheat yields as compared to PROBA-V 300 m and 1 km resolutions. Furthermore, for all resolutions and thermal time, the highest correlation was obtained by using an NDVI threshold of 0.2.

The methodology demonstrates that the 100 m and 300 m PROBA-V datasets have the potential to be used for estimating winter wheat yield. Since PROBA-V is a relatively new satellite, there are limited studies available for crop yield estimation. This study's results align with reported ranges [8, 9]. Zheng et al. [215] generated 100 m land surface reflectances by fusing the PROBA-V 100 m and 300 m S1 products. They achieved 0.663 for R^2, 0.63 for RMSE (t/ha), and 7.27% for RRMSE at Yucheng County, and they achieved 0.631 for R^2, 0.42 for RMSE (t/ha), and 4.88% for RRMSE at Guantao County for winter wheat yield estimates in China for one growing season in 2014–2015 [215]. In another study, Meroni et al. [46] evaluated the NDVI data continuity between SPOT-VGT and PROBA-V missions for operational yield forecasting in North African countries for yield estimates of barley, soft wheat, and durum wheat. They obtained 0.78 for R^2 and 4.51 for RMSE difference in percentage in Morocco, 0.94 for R^2 and 3.27 for RMSE difference in percentage in Algeria, and 0.91 for R^2 and -31.76 for RMSE difference in percentage in Tunisia. Within this study, statistical metrics were highest for thermal time, high pixel purity, and a 0.2 NDVI threshold; these were improved with increasing resolution. An NDVI threshold of 0.2 draws a border on the lower limit of the curve for most of the fields because the minimum NDVI within the

growing season is around the 0.2 NDVI threshold. Therefore, this NDVI threshold level gave the most accurate result. The adjusted coefficients of determination (R^2) were 0.2 at 1 km, 0.55 at 300 m, and 0.74 at 100 m resolutions. The jackknifed RMSE amounted to 1.07 t/ha for 1 km resolution, 0.79 t/ha for 300 m resolution, and 0.60 t/ha for 100 m resolution, while MAE was 0.78 t/ha for 1 km resolution, 0.53 t/ha for 300 m, and 0.46 t/ha for 100 m resolutions.

The results of this study closely agree with other studies that compared the performance of the spatial resolutions of satellites. Guindin-Garcia evaluated MODIS 8-day and 16-day composite products for monitoring maize using the green leaf area index (LAI_g) [230]. The results revealed that 250 m MODIS products provided more accurate estimates of maize LAI_g compared to 500 m products during the entire growing season [230]. Chen et al. [231] assessed the potential use of the MODIS Enhanced Vegetation Index (EVI) at 250 m, 500 m, and 1 km resolutions for monitoring maize. According to their results, MODIS EVI at 1 km data had lower EVI values compared to 250 m and 500 m data in the vegetative stage. The reason was that the higher resolution pixels contained a larger proportion of agricultural coverage [231].

Though this study showed promising results in comparing different PROBA-V resolution data for crop yield estimates, there were a number of limitations. One of these limitations was the cloud cover which caused missing data during the growing season. The 100 m resolution dataset was particularly affected because number of observations made throughout the growing season presented a poorer temporal resolution compared to 300 m and 1 km resolution datasets. For this reason, field selection thresholds were applied, and 39 winter wheat fields out of 56 fields were selected. The 300 m and 1 km resolution datasets would allow fields to be sampled with a higher revisit frequency. Therefore, it might be preferable to use the 300 m NDVI despite the fact that the 100 m dataset would, in principal, result in a stronger performance. Another limitation was the limited availability of the field-level ground data. Collecting ground data is time consuming, labor intensive, and expensive. Instead of collecting field data, crop identification maps, as explained in Chapter 2, could be used. These maps could be computed based on VI values [232] and subsequently compared with regional yield statistics, a common practice in comparing yield statistics with remote sensing indices.

This study could be extended to the national level and to other crop types across different regions of the world. Kempeneers et al. [233] implemented a Kalman filter recursive algorithm that integrated 100 m and 300 m resolution images to generate the assimilated imagery at 100 m resolution. This simulated 100 m data based on data assimilation methods could be used to increase the temporal resolution of the PROBA-V 100 m product.

3.6. Conclusions

This study demonstrated the potential of the 100 m and 300 m PROBA-V NDVI datasets for yield estimation compared to the 1 km dataset for the 2014–2015 growing season of winter wheat in Northern France. After applying an ADSF, the integrated values were computed using daily or 5-daily NDVIs for thermal time and calendar time. A linear regression was subsequently calculated between the yield observations at the field level and the integral values. Depending on the resolution and the NDVI threshold used for the thermal time analysis, the adjusted R^2 ranged between 0.20 and 0.74, jackknifed – leave-one-field-out cross validation – RMSE (t/ha) ranged between 0.6 and 1.07 t/ha, and MAE (t/ha) ranged between 0.46 and 0.90 t/ha. The most accurate results were obtained with a 100 m resolution at the 0.2 NDVI threshold integrated over thermal time with no purity threshold.

Overall, thermal time allowed for better yield estimates as compared to calendar time. Despite the higher accuracy of the 100 m dataset, the reduction of data in time made the 100 m dataset suboptimal; therefore, the 300 m dataset with thermal time could offer a good compromise. Additionally, the 300 m dataset with calendar time and high purity of 65% and 90% produced the highest adjusted R^2 values. However, the most probable reason for this result would be the low sample size which increased the margin of error. Furthermore, the 300 m dataset compared to the 100 m dataset might allow more fields to be sampled that are representative of the landscape. The method can be extended to larger datasets, other crops, and regions in the world. This approach might be very useful for estimating yield at the parcel and regional levels across the world.

General conclusions and perspectives

Conclusions

The objective of this thesis is to advance agricultural monitoring for improved yield estimations at regional to global scales using low spatial resolution SPOT-VGT and PROBA-V-type remote sensing data. In order to reach this goal, the research questions were studied in three independent analyses, which have been realized using different datasets and ground data.

Is a modified DMP a more accurate proxy for crop yield estimations compared to CGLS-DMP, NDVI, and fAPAR?

Remote sensing plays an important role in regional crop yield estimation studies by providing near real-time data of crop surface status. Satellite imagery provides not only BPs, but also crop identification mapping. Both are the primary inputs for developing crop yield estimates. Biomass proxies are diverse, and each one has its own advantages and disadvantages. One of these BPs is CGLS-DMP, which is a parametric growth model. This model can provide information on arable productivity and can be used as a proxy for crop yield. There are several limitations in the current version of CGLS-DMP; for instance, it was parameterized for European forests. Consequently, the model is open for improvement.

The first research question explores how CGLS-DMP, which was developed for SPOT-VGT, can be improved to more closely relate to yield anomalies across selected regions. Within this thesis, a modified DMP model was developed that included a water stress factor based on the actual evapotranspiration calculated with AgroMetShell and an adaptation of the existing factors: the CO_2 fertilization effect, temperature stress, and autotrophic respiration. The strongest results were provided by using different combinations of stress factors for each selected region and by dispersing the modified DMP between the flowering and ripening period. Model performance of the modified DMP increased when compared to the CGLS-DMP. The variation of results between the modified DMP and simpler BPs, such as fAPAR and NDVI, did not fluctuate.

The DMP approach is encouraging because it may be more applicable to extrapolation in more extreme conditions. According to the results, the modified DMP could potentially be used in crop yield estimates on a regional scale. Although no single solution to the improvement of a global product could be demonstrated, a combination of different stress factors produced more precise local yield estimates. Because the interannual variability of official yield statistics is higher in Morocco compared to Belgium and France, the model in Morocco enabled a more thorough exploration of stress factors. DMP estimates of cereal crops were improved by including appropriate stress factors and their impact during sensitive stages. Although there are few studies available that relate crop yield forecasting to radiation or light use efficiency [37,38], the CGLS-DMP has been

studied for the first time for crop production estimation. This study can give insight into the opportunities and limitations of the model for any possible future research related to agricultural applications.

Is it possible to retrieve crop identification with a 100 m PROBA-V product that could possibly be used in crop yield estimations?

Although low-resolution images are used extensively to estimate crop yield, their spatial resolution is usually too coarse for crop area mapping. However, the 100 m PROBA-V product could potentially be used in crop identification mapping, due to its higher spatial resolution. Crops have different spectral responses during their growing seasons, and their temporal profiles can be identified by remote sensing data. Therefore, it is possible to identify major crop types and their phenology.

The second research question explores how multi-temporal 100 m PROBA-V images can be used effectively for phenology-based crop identification mapping at a global scale using adapted SMTs for the 2014–2015 season. Phenological metrics were extracted from the NDVI time series using piecewise logistic functions. These metrics represented the crop growing seasons and identified the unique calendar of each crop type. The crop classification accuracies obtained were comparable to the results derived from classification with higher resolution data. With this study, PROBA-V 100 m data has been used for the first time at a global scale for crop identification mapping. Although this study shows the potential of using the 100 m dataset to retrieve crop maps, this type of study performed at a higher spatial resolution could give superior results. For instance, the Sentinel-2 dataset could be a potentially more logical option with its 10 m spatial and 5-day temporal resolution, which is provided on an open and free basis at a global scale.

What is the optimal spatial resolution between 100 m, 300 m, and 1 km PROBA-V products to estimate crop yield at the field scale?

The spatial pattern of the study areas defines the optimal spatial resolution requirement for crop growth monitoring. This research question was contemplated at the field level in terms of pixel size and partly by pixel purity. The time series of PROBA-V images is not long enough to convey a traditional crop yield estimation study, yet the datasets at different resolutions can be compared at the field level for one growing season to observe whether an increase in spatial resolution benefits crop yield forecasting. The third research question explores the trade-off between the different spatial resolutions provided by PROBA-V products versus the temporal frequency and, additionally, explores the use of thermal time to improve statistical yield estimations. In this thesis, the potential for using 100 m and 300 m resolution PROBA-V NDVI images in yield estimation was demonstrated and compared to the 1 km dataset for the 2014–2015 growing season in Northern France. As a proxy for yield, the integrated values were computed using the daily NDVI and cumulative GGD images. The results showed that better spatial resolution obtained the highest correlation between the real yield data and the proxy for yield. In addition, in the event of cloud cover, the temporal resolution was even lower. This problem was resolved by applying field selection thresholds to retain sufficient data during the growing season. However, it must be mentioned that almost 30% of the fields had to be discarded. This study is the first to compare PROBA-V NDVI datasets at 100 m, 300 m, and 1 km resolutions to estimate crop yield at the field level. The potential of the PROBA-V 100 m dataset, which has been demonstrated within this research, is particularly important when its global coverage is considered.

Perspectives

Many technological developments have been achieved in the course of this thesis that might challenge the choices made in the beginning of the research. The general approach followed in this thesis could be improved in several ways. First, PROBA-V's life expectancy ends in 2019; its gap-filling mission is finished now that Sentinel-3 sensors are in orbit. Today, the length of the time series plays an important role for the majority of yield prediction methods, which are based on the use of these series for comparisons with previous years or with the average situation [11]. Therefore, the continuity of existing systems remains crucial to ensure the availability of these series, particularly in large operational systems that cannot profit immediately from the availability of higher spatial resolution sensors [11]. Since PROBA-V can ensure this continuity of the crop monitoring studies based on a long historical archive, it was essential to study the applicability of PROBA-V products in crop monitoring studies. As a future recommendation, it is suggested to apply the methodologies used in this thesis to Sentinel-3 images. Further research efforts on sensor intercalibration are necessary to simplify the access to longer time series for remotely sensed data from different sensors [11]. For instance, SPOT-VGT halted its operation in May 2014, and PROBA-V, which is the direct successor mission to SPOT-VGT, has been operational since October 2013. Therefore, the spectral continuity between SPOT-VGT and PROBA-V daily global datasets has been validated [49].

Second, although crop yield estimation studies are traditionally based on long-term archives, there have also been recent crop yield estimation studies combining remote sensing data, such as Sentinel-1 and Sentinel-2. Battude et al. [234] estimate maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2-like remote sensing data. Veloso et al. [235] try to understand the temporal behavior of crops using Sentinel-1-like and Sentinel-2-like data for agricultural applications. Skakun et al. [236] combine the use of Landsat-8 and Sentinel-2A images for winter crop mapping and winter wheat yield assessment on a regional scale. Moreover, crowdsourced data is provided by farmers, including the date of sowing, crop varieties, amount of fertilizer used, and previous actual yield data. This data could be used in new methodologies for regional crop forecasting combined with Sentinel-2-type remote sensing data. The transition to using higher spatial resolution in crop monitoring studies would occur sooner if new methodologies became developed enough to conduct operational applications at a global scale.

Third, since Sentinel-2 has been providing 10 m imagery across the globe every 5 days, it should potentially and probably provide better results for crop identification mapping. However, it should be noted that the amount of Sentinel-2 data is quite large compared to PROBA-V 100 m data. Within this thesis, a potential for using a PROBA-V product for crop identification mapping has been illustrated. Therefore, for today, to avoid handling big data and processing, PROBA-V imagery or an equivalent product would be a more useful option. In the future, however, technological improvements could make the processing and handling of data much easier.

Fourth, there are new possibilities to estimate GPP and NPP from satellite images using sun-induced fluorescence, machine-learning from flux-towers, or a universal model for carbon dioxide uptake by Wang et al. [31]. As a future recommendation, these new models could be investigated and combined with a modified DMP. For instance, instead of calculating RUE, real RUE estimations could be obtained through remote sensing methods, such as chlorophyll-related VIs, passive measurement of solar-induced chlorophyll fluorescence, and the photochemical reflectance index.

For another improvement or investigation, a number of study sites on the global level could be selected, and fieldwork campaigns could be conducted to collect yield and field

information for the main crops of each specific region within growing seasons. The methods used in Chapter 1 and Chapter 2 could be applied to simulate similar studies at a global scale. If the results would show similar outcomes to this thesis, the resulting products could be used in operational systems for monitoring global agriculture.

There has been a continuous development in technology, including in the design of satellites. Space agencies continue designing instruments dedicated to agricultural monitoring. Several recommendations for space agencies and governments emerge from this thesis regarding the requirements for satellites for agricultural remote sensing. Within this research, the PROBA-V 100 m dataset has been studied and its potentials and limitations in agricultural monitoring have been depicted. First, the 100 m pixel size fills a shortcoming between high-resolution sensors with a lower revisit time and low-resolution sensors with daily global coverage. Sentinel-3 was designed as a successor of Environmental Satellite; specifically, the Ocean and Land Color Instrument was built on the heritage of MERIS with a 300 m spatial resolution [237]. Sentinel-3's mission is to continue the legacy of 300–1000 m resolution optical measurements – for example, MERIS, Advanced Along Track Scanning Radiometer, SPOT VGT and Meteorological Operational AVHRR – and to provide daily coverage of the Earth for atmospheric land services, including regional-scale and continental-scale land cover mapping and bio-geophysical parameters [238]. Additionally, it was designed to support previous Global Monitoring for Environment and Security services – currently CGLS which are mostly situated in the framework of global services [238]. Satellite requirements focused on the provision of high-resolution measurements over land are beyond the optical instrument capability proposed for Sentinel-3 [238]. It is not possible to satisfy all user requirements set out based on the spatial resolution requirements, depending on the service and product required. On the other hand, the user requirements of Sentinel-2 are different compared to Sentinel-3; these include land cover and land use mapping and emergency services, which require a higher spatial resolution compared to Sentinel-3 [239]. While considering the variety of terrestrial application-related services, user requirements, and constraints – such as time, orbit selection, and avoiding the duplication of the design of the missions – one finds the choice of spatial resolution and spectral bands difficult to make [239]. In short, neither Sentinel-3 nor Sentinel-2 has been designed with a 100 m spatial resolution. However, because several studies, including this thesis, point out the potentials of using a 100 m spatial resolution in both crop monitoring and crop identification mapping, the next generation's satellites might consider a spatial resolution of 100 m.

Second, one of the main limitations of using a PROBA-V dataset was its temporal resolution for the 100 m resolution. One area of improvement could be the introduction of new satellites in constellations for higher temporal resolution. Over the last decade, significant progress has been made in developing and launching satellites in constellations that can provide daily revisits across the globe [240]. However, there are still missions that cannot fulfill the need for daily revisits, such as PROBA-V and Sentinel-2. Future remote sensing missions with a similar or higher spatial resolution than 100 m and improved temporal resolution, such as daily instead of every 5 days could particularly enhance crop production estimates.

Third, most of the current crop growth monitoring studies depend on the long-term image archive. In 10–15 years, Sentinel-2 will have enough archives to be used in crop estimation studies at regional and global scales. Additionally, due to fast technological improvements, in the near future, storage and data processing with high spatial and temporal resolution images may no longer a problem. Therefore, there will probably be no need to use low-resolution data such as PROBA-V and Sentinel-3. However, at the present

time, it is still important to discover and study the images that are currently available for crop monitoring because they have a sufficiently long historical archive.

This thesis aimed to explore how much low spatial resolution data, for instance PROBA-V imagery, could contribute to global crop monitoring by possible improvements in crop yield estimation. The approach of monitoring crop growth and estimating yield can support governments and organizations to plan national food production and tackle any unexpected weather events throughout the growing season. The FAO discusses the challenges that agriculture will face in the 21st century: producing more food and fiber to feed a growing population with a smaller rural labor force, contributing to the overall development in the many agriculture-dependent developing countries, and adapting to climate change [3]. Therefore, food production will be more vulnerable in the future due to climate change and the increase in the world's population. Expected production shortages and decreasing short-term price instability can be projected by improved monitoring of crop production. Market efficiency can be increased through accurate commodity price forecasts. Recording persistent shortfalls can stimulate policies to prioritize efforts toward ameliorating vulnerable agricultural systems [5]. Identification of production fluctuations and the evaluation of agricultural system flexibility can be supported by long-term data records for agricultural monitoring. Thus, it is fundamentally important to monitor crop growth and estimate yields for the crop growing season, particularly for the major crops on a global level, such as wheat, rice, and maize.

References

1. UN General Assembly. *Universal Declaration of Human Rights*; New York, UN, 1948.
2. Alexandratos, N.; Bruinsma, J. *World Agriculture: Towards 2030/2050: The 2012 Revision*; ESA Working paper No. 12-03; Rome, FAO, 2012.
3. FAO. Global Agriculture towards 2050. In *High Level Expert Forum on How to feed the world 2050*; FAO: Rome, 2009; pp 1–4.
4. FAO; IFAD; WFP. *The State of Food Insecurity in the World. The Multiple Dimensions of Food Security*; FAO: Rome, FAO, 2013.
5. Duveiller, G. Crop Specific Green Area Index Retrieval from Multi-Scale Remote Sensing for Agricultural Monitoring, Université Catholique de Louvain (UCL), 2011.
6. Ashraf, M.; Maah, J.; Yusoff, J. Introduction to Remote Sensing of Biomass. In *Biomass and remote sensing of biomass*; Islam, A., Ed.; InTech: China, 2011; pp 129–170.
7. Biatwright, G. O.; Whitehead, V. S. Early Warning and Crop Condition Assessment Research. *Geosci. Remote Sensing, IEEE Trans.* 1986, *GE-24* (1), 54–64.
8. Curnel, Y. Satellite Remote Sensing Priorities for a Better Assimilation in Crop Growth Models, Université Catholique de Louvain (UCL), 2014.
9. Mo, X.; Liu, S.; Lin, Z.; Xu, Y.; Xiang, Y.; McVicar, T. R. Prediction of Crop Yield, Water Consumption and Water Use Efficiency with a SVAT-Crop Growth Model Using Remotely Sensed Data on the North China Plain. *Ecol. Modell.* 2005, *183* (2–3), 301–322.
10. Eerens, H.; Haesen, D.; Rembold, F.; Urbano, F.; Tote, C.; Bydekerke, L. Image Time Series Processing for Agriculture Monitoring. *Environ. Model. Softw.* 2014, *53*, 154–162.
11. Rembold, F.; Atzberger, C.; Savin, I.; Rojas, O. Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection. *Remote Sens.* 2013, *5* (4), 1704–1733.
12. Becker-Reshef, I.; Vermote, E.; Lindeman, M.; Justice, C. A Generalized Regression-Based Model for Forecasting Winter Wheat Yields in Kansas and Ukraine Using MODIS Data. *Remote Sens. Environ.* 2010, *114* (6), 1312–1323.
13. Lobell, D. B.; Burke, M. B. On the Use of Statistical Models to Predict Crop Yield Responses to Climate Change. *Agric. For. Meteorol.* 2010, *150* (11), 1443–1452.

14. Dorigo, W. A.; Zurita-Milla, R.; de Wit, A. J. W.; Brazile, J.; Singh, R.; Schaepman, M. E. A Review on Reflective Remote Sensing and Data Assimilation Techniques for Enhanced Agroecosystem Modeling. *Int. J. Appl. Earth Obs. Geoinf.* 2007, *9* (2), 165–193.

15. Aspinall, R. J.; Marcus, W. a; Boardman, J. W. Considerations in Collecting , Processing, and Anlaysing High Spational Resolution Hyperspectral Data for Environmental Investigations. *J. Geogr. Syst.* 2002, *4*, 15–29.

16. Chuvieco, E.; Huete, A. *Fundamentals of Satellite Remote Sensing*; CRC Press: FL, USA, 2010.

17. Théau, J. Temporal Resolution. In *Encyclopedia of GIS*; Springer US: Boston, MA, 2008; pp 1150–1151.

18. Natural Resources Canada. Radiometric Resolution http://www.nrcan.gc.ca/node/9379 (accessed Nov 9, 2017).

19. Tucker, C. J.; Holben, B. N.; Elgin, J. H. Remote Sensing of Total Dry-Matter Accumulation in Winter Wheat. 1981, *189*.

20. Maselli, F.; Chiesi, M.; Brilli, L.; Moriondo, M. Simulation of Olive Fruit Yield in Tuscany through the Integration of Remote Sensing and Ground Data. *Ecol. Modell.* 2012, *244*, 1–12.

21. Ferencz, C.; Bognár, P.; Lichtenberger, J.; Hamar, D.; Tarcsai†, G.; Timár, G.; Molnár, G.; Pásztor, S. Z.; Steinbach, P.; Székely, B.; Ferencz, O. E.; Ferencz-Árkos, I. Crop Yield Estimation by Satellite Remote Sensing. *Int. J. Remote Sens.* 2004, *25* (20), 4113–4149.

22. Ren, J.; Chen, Z.; Zhou, Q.; Tang, H. Regional Yield Estimation for Winter Wheat with MODIS-NDVI Data in Shandong, China. *Int. J. Appl. Earth Obs. Geoinf.* 2008, *10*, 403–413.

23. Funk, C.; Budde, M. E. Phenologically-Tuned MODIS NDVI-Based Production Anomaly Estimates for Zimbabwe. *Remote Sens. Environ.* 2009, *113* (1), 115–125.

24. Pinker, R. T.; Zhao, M.; Wang, H.; Wood, E. F. Impact of Satellite Based PAR on Estimates of Terrestrial Net Primary Productivity. *Int. J. Remote Sens.* 2010, *31* (19), 5221–5237.

25. Frankenberg, C.; Fisher, J. B.; Worden, J.; Badgley, G.; Saatchi, S. S.; Lee, J. E.; Toon, G. C.; Butz, A.; Jung, M.; Kuze, A.; Yokota, T. New Global Observations of the Terrestrial Carbon Cycle from GOSAT: Patterns of Plant Fluorescence with Gross Primary Productivity. *Geophys. Res. Lett.* 2011, *38* (17), 1–6.

26. Tramontana, G.; Jung, M.; Schwalm, C. R.; Ichii, K.; Camps-Valls, G.; Ráduly, B.; Reichstein, M.; Arain, M. A.; Cescatti, A.; Kiely, G.; Merbold, L.; Serrano-Ortiz, P.; Sickert, S.; Wolf, S.; Papale, D. Predicting Carbon Dioxide and Energy Fluxes across Global FLUXNET Sites with Regression Algorithms. *Biogeosciences* 2016, *13* (14), 4291–4313.

27. Jung, M.; Reichstein, M.; Schwalm, C. R.; Huntingford, C.; Sitch, S.; Ahlström, A.; Arneth, A.; Camps-Valls, G.; Ciais, P.; Friedlingstein, P.; Gans, F.; Ichii, K.; Jain, A. K.; Kato, E.; Papale, D.; Poulter, B.; Raduly, B.; Rödenbeck, C.; Tramontana, G.; Viovy, N.; Wang, Y. P.; Weber, U.; Zaehle, S.; Zeng, N. Compensatory Water Effects Link Yearly Global Land CO 2 Sink Changes to Temperature. *Nature* 2017, *541* (7638), 516–520.

28. Fluxcom. FLUXCOM http://www.fluxcom.org/ (accessed May 30, 2018).

29. Guanter, L.; Zhang, Y.; Jung, M.; Joiner, J.; Voigt, M.; Berry, J. A.; Frankenberg, C.; Huete, A. R.; Zarco-Tejada, P.; Lee, J.-E.; Moran, M. S.; Ponce-Campos, G.; Beer, C.; Camps-Valls, G.; Buchmann, N.; Gianelle, D.; Klumpp, K.; Cescatti, A.; Baker, J. M.; Griffis, T. J. Global and Time-Resolved Monitoring of Crop Photosynthesis with Chlorophyll Fluorescence. *Proc. Natl. Acad. Sci.* 2014, *111* (14), E1327–E1333.

30. Sun, Y.; Frankenberg, C.; Wood, J. D.; Schimel, D. S.; Jung, M.; Guanter, L.; Drewry, D. T.; Verma, M.; Porcar-Castell, A.; Griffis, T. J.; Gu, L.; Magney, T. S.; Köhler, P.; Evans, B.; Yuen, K. OCO-2 Advances Photosynthesis Observation from Space via Solar-Induced Chlorophyll Fluorescence. *Science (80-.).* 2017, *358* (6360).

31. Wang, H.; Prentice, I. C.; Keenan, T. F.; Davis, T. W.; Wright, I. J.; Cornwell, W. K.; Evans, B. J.; Peng, C. Towards a Universal Model for Carbon Dioxide Uptake by Plants. *Nat. Plants* 2017, *3* (9), 734–741.

32. VITO; Imperial College London; University of Antwerp. OVERALL GOAL OF TERRA-P https://terra-p.vito.be/about/overall-goal-terra-p (accessed May 30, 2018).

33. Swinnen, E.; Van Hools, R.; Eerens, H. *GIO-GL Lot 1, Algorithm Theorethical Basis Document, Dry Matter Productivity (DMP)*; Mol, 2015.

34. Chirici, G.; Anna, B.; Maselli, F. Modelling of Italian Forest Net Primary Productivity by the Integration of Remotely Sensed and GIS Data. *For. Ecol. Manage.* 2007, *246* (2–3), 285–295.

35. Maselli, F.; Argenti, G.; Chiesi, M.; Angeli, L.; Papale, D. Simulation of Grassland Productivity by the Combination of Ground and Satellite Data. *Agric. Ecosyst. Environ.* 2013, *165*, 163–172.

36. Chiesi, M.; Fibbi, L.; Genesio, L.; Gioli, B.; Magno, R.; Maselli, F.; Moriondo, M.; Vaccari, F. P. Integration of Ground and Satellite Data to Model Mediterranean Forest Processes. *Int. J. Appl. Earth Obs. Geoinf.* 2011, *13* (3), 504–515.

37. Bastiaanssen, W. G. M.; Ali, S. A New Crop Yield Forecasting Model Based on Satellite Measurements Applied across the Indus Basin, Pakistan. *Agric. Ecosyst. Environ.* 2003, *94* (3), 321–340.

38. Liu, J.; Pattey, E.; Miller, J. R.; McNairn, H.; Smith, A.; Hu, B. Estimating Crop Stresses, Aboveground Dry Biomass and Yield of Corn Using Multi-Temporal Optical Data Combined with a Radiation Use Efficiency Model. *Remote Sens. Environ.* 2010, *114* (6), 1167–1177.

39. Skakun, S.; Franch, B.; Vermote, E.; Roger, J. C.; Becker-Reshef, I.; Justice, C.; Kussul, N. Early Season Large-Area Winter Crop Mapping Using MODIS NDVI Data, Growing Degree Days Information and a Gaussian Mixture Model. *Remote Sens. Environ.* 2017, *195*, 244–258.

40. Tsendbazar, N. E.; de Bruin, S.; Herold, M. Assessing Global Land Cover Reference Datasets for Different User Communities. *ISPRS J. Photogramm. Remote Sens.* 2015, *103*, 93–114.

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

42. Wardlow, B. D.; Egbert, S. L. Large-Area Crop Mapping Using Time-Series MODIS

250 M NDVI Data: An Assessment for the U.S. Central Great Plains. *Remote Sens. Environ.* 2008, *112* (3), 1096–1116.

43. Duncan, J. M. A.; Dash, J.; Atkinson, P. M. The Potential of Satellite-Observed Crop Phenology to Enhance Yield Gap Assessments in Smallholder Landscapes. *Front. Environ. Sci.* 2015, *3* (56), 1–16.

44. Peña-Barragán, J. M.; Ngugi, M. K.; Plant, R. E.; Six, J. Object-Based Crop Identification Using Multiple Vegetation Indices, Textural Features and Crop Phenology. *Remote Sens. Environ.* 2011, *115* (6), 1301–1316.

45. Li, M.; Ma, L.; Blaschke, T.; Cheng, L.; Tiede, D. A Systematic Comparison of Different Object-Based Classification Techniques Using High Spatial Resolution Imagery in Agricultural Environments. *Int. J. Appl. Earth Obs. Geoinf.* 2016, *49*, 87–98.

46. Meroni, M.; Fasbender, D.; Balaghi, R.; Dali, M.; Haffani, M.; Haythem, I.; Hooker, J.; Lahlou, M.; Lopez-Lozano, R.; Mahyou, H.; Ben Moussa, M.; Sghaier, N.; Wafa, T.; Leo, O. Evaluating NDVI Data Continuity Between SPOT-VEGETATION and PROBA-V Missions for Operational Yield Forecasting in North African Countries. *IEEE Trans. Geosci. Remote Sens.* 2016, *54* (2), 795–804.

47. Roumenina, E.; Atzberger, C.; Vassilev, V.; Dimitrov, P.; Kamenova, I.; Banov, M.; Filchev, L.; Jelev, G. Single- and Multi-Date Crop Identification Using PROBA-V 100 and 300 M S1 Products on Zlatia Test Site, Bulgaria. *Remote Sens.* 2015, *7* (10), 13843–13862.

48. Smets, B.; Lacaze, R.; Freitas, S. C.; Jann, A.; Calvet, J. C.; Camacho, F.; Baret, F.; Paulik, C.; D'Andrimont, R.; Tansey, K.; others. Operating The Copernicus Global Land Service. In *ESA Living Planet Symposium*; 2013; Vol. 722, p 66.

49. Dierckx, W.; Swinnen, E.; Kempeneers, P. Validation of Spectral Continuity between PROBA-V and SPOT-VEGETATION Global Daily Datasets. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 2015, *XL-7/W3* (7W3), 1155–1162.

50. World Bank. *Global Strategy to Improve Agricultural and Rural Statistics*; Report Number 56719-GLB. Washington, DC: World Bank, United Nations, and Food and Agricultural Organization of the United Nations, 2010.

51. Fermont, A.; Benson, T. *Estimating Yield of Food Crops Grown by Smallholder Farmers: A Review in the Uganda Context*; IFPRI Discussion Paper Washington D.C., Kampala, 2011.

52. Xin, Q.; Gong, P.; Yu, C.; Yu, L.; Broich, M.; Suyker, A.; Myneni, R. A Production Efficiency Model-Based Method for Satellite Estimates of Corn and Soybean Yields in the Midwestern US. *Remote Sens.* 2013, *5* (11), 5926–5943.

53. Tao, F.; Yokozawa, M.; Zhang, Z.; Xu, Y.; Hayashi, Y. Remote Sensing of Crop Production in China by Production Efficiency Models : Models Comparisons, Estimates and Uncertainties. *Ecol. Modell.* 2005, *183* (4), 385–396.

54. Wall, L.; Larocque, D.; Léger, P. M. The Early Explanatory Power of NDVI in Crop Yield Modelling. *Int. J. Remote Sens.* 2008, *29* (8), 2211–2225.

55. Kogan, F.; Kussul, N.; Adamenko, T.; Skakun, S.; Kravchenko, O.; Kryvobok, O.; Shelestov, A.; Kolotii, A.; Kussul, O.; Lavrenyuk, A. Winter Wheat Yield Forecasting in Ukraine Based on Earth Observation, Meteorological Data and Biophysical Models. *Int. J. Appl. Earth Obs. Geoinf.* 2013, *23*, 192–203.

56. López-Lozano, R.; Duveiller, G.; Seguini, L.; Meroni, M.; Garcia-Condado, S.; Hooker, J. D.; Leo, O.; Baruth, B. Regional Grain Yield Prediction with 1km-Resolution EO Biophysical Products: Strengths and Limitations at Pan-European Level. *Agric. For. Meteorol. (In Press.* 2015.

57. Doraiswamy, P. C.; Moulin, S.; Cook, P. W.; Stern, A. Crop Yield Assessment from Remote Sensing. *Photogramm. Eng. Remote Sensing* 2003, *69* (6), 665–674.

58. Fensholt, R.; Sandholt, I.; Rasmussen, M. S. Evaluation of MODIS LAI, fAPAR and the Relation between fAPAR and NDVI in a Semi-Arid Environment Using in Situ Measurements. *Remote Sens. Environ.* 2004, *91* (3–4), 490–507.

59. Balaghi, R.; Tychon, B.; Eerens, H.; Jlibene, M. Empirical Regression Models Using NDVI, Rainfall and Temperature Data for the Early Prediction of Wheat Grain Yields in Morocco. *Int. J. Appl. Earth Obs. Geoinf.* 2008, *10*, 438–452.

60. Kuri, F.; Murwira, A.; Murwira, K. S.; Masocha, M. Predicting Maize Yield in Zimbabwe Using Dry Dekads Derived from Remotely Sensed Vegetation Condition Index. *Int. J. Appl. Earth Obs. Geoinf.* 2014, *33*, 39–46.

61. Prasad, A. K.; Chai, L.; Singh, R. P.; Kafatos, M. Crop Yield Estimation Model for Iowa Using Remote Sensing and Surface Parameters. *Int. J. Appl. Earth Obs. Geoinf.* 2006, *8*, 26–33.

62. Gusso, A.; Ducati, J. R.; Veronez, M. R.; Arvor, D.; Jr, L. G. D. S. Spectral Model for Soybean Yield Estimate Using. *Int. J. Geosci.* 2013, *2013* (November), 1233–1241.

63. Meroni, M.; Marinho, E.; Sghaier, N.; Verstrate, M.; Leo, O. Remote Sensing Based Yield Estimation in a Stochastic Framework — Case Study of Durum Wheat in Tunisia. *Remote Sens.* 2013, *5* (2), 539–557.

64. Mkhabela, M. S.; Bullock, P.; Raj, S.; Wang, S.; Yang, Y. Crop Yield Forecasting on the Canadian Prairies Using MODIS NDVI Data. *Agric. For. Meteorol.* 2011, *151* (3), 385–393.

65. Vicente-Serrano, S.; Cuadrat-Prats, J.; Romo, A. Early Prediction of Crop Production Using Drought Indices at Different Time Scales and Remote Sensing Data : Application in the Ebro Valley (North - East Spain). *Int. J. Remote Sens.* 2006, *27* (3), 511–518.

66. Duveiller, G.; Baret, F.; Defourny, P. Remotely Sensed Green Area Index for Winter Wheat Crop Monitoring: 10-Year Assessment at Regional Scale over a Fragmented Landscape. *Agric. For. Meteorol.* 2012, *166–167*, 156–168.

67. Ajtay, G. L.; Ketner, P.; Duvigneaud, P. *Terrestrial Primary Production and Phytomass*, The Global.; Bolin, B., Degens, E. T., Ketner, S., Kempe, P., Eds.; Wiley: New York, 1979.

68. Garbulsky, M. F.; Peñuelas, J.; Papale, D.; Ardö, J.; Goulden, M. L.; Kiely, G.; Richardson, A. D.; Rotenberg, E.; Veenendaal, E. M.; Filella, I. Patterns and Controls of the Variability of Radiation Use Efficiency and Primary Productivity across Terrestrial Ecosystems. *Glob. Ecol. Biogeogr.* 2010, *19*, 253–267.

69. Ruimy, A.; Saugier, B.; Dedieu, G. Methodology for the Estimation of Terrestrial Net Primary Production from Remotely Sensed Data. *J. Geophys. Res.* 1994, *99* (D3), 5263.

70. Matsushita, B.; Xu, M.; Chen, J.; Kameyama, S.; Tamura, M. Estimation of Regional

Net Primary Productivity (NPP) Using a Process-Based Ecosystem Model: How Important Is the Accuracy of Climate Data? *Ecol. Modell.* 2004, *178*, 371–388.

71. Horn, J. E.; Schulz, K. Identification of a General Light Use Efficiency Model for Gross Primary Production. *Biogeosciences* 2011, *8*, 999–1021.

72. McCallum, I.; Wagner, W.; Schmullius, C.; Shvidenko, A.; Obersteiner, M.; Fritz, S.; Nilsson, S. Satellite-Based Terrestrial Production Efficiency Modeling. *Carbon Balance Manag.* 2009, *4*, 8.

73. Ruimy, A.; Kergoat, L.; Bondeau, A.; Intercomparison, T. P. O. T. P. Comparing Global Models of Terrestrial Net Primary Productivity (NPP): Analysis of Differences in Light Absorption and Light-Use Efficiency. *Glob. Chang. Biol.* 1999, *5* (S1), 56–64.

74. Monteith, J. L. Solar Radiation and Productivity in Tropical Ecosystems. *J. Appl. Ecol.* 1972, *9* (3), 747.

75. Monteith, J. L.; Moss, C. J. Climate and the Efficiency of Crop Production in Britain [and Discussion]. *Philos. Trans. R. Soc. B Biol. Sci.* 1977, *281* (980), 277–294.

76. Sinclair, T. R.; Muchow, R. C. Radiation Use Efficiency. *Adv. Agron.* 1998, *65*, 215–265.

77. Field, C. B.; Randerson, J. T.; Malmström, C. M. Global Net Primary Production: Combining Ecology and Remote Sensing. *Remote Sens. Environ.* 1995, *51* (1), 74–88.

78. Goetz, S. J.; Prince, S. D.; Goward, S. N.; Thawley, M.; Small, J. Satellite Remote Sensing of Primary Production: An Improved Production Efficiency Modeling Approach. *Ecol. Model.* 1999, *122*, 239–255.

79. Running, S. W.; Thornton, P. E.; Nemani, R. R.; Glassy, J. M. Global Terrestrial Gross and Net Primary Productivity from the Earth Observing System. In *Methods in Ecosystem Science*; Sala, O., Jackson, R., Mooney, H., Eds.; Springer: New York, 2000; pp 44–57.

80. Xiao, X.; Hollinger, D.; Aber, J.; Goltz, M.; Davidson, E. a.; Zhang, Q.; Moore, B. Satellite-Based Modeling of Gross Primary Production in an Evergreen Needleleaf Forest. *Remote Sens. Environ.* 2004, *89*, 519–534.

81. Potter, C. S.; Randerson, J. T.; Field, C. B.; Matson, P. A.; Vitousek, P. M.; Mooney, H. A.; Klooster, S. A. Terrestrial Ecosystem Production: A Process Model Based on Global Satellite and Surface Data. *Global Biogeochem. Cycles* 1993, *7* (4), 811–841.

82. Ruimy, A.; Dedieu, G.; Saugier, B. TURC: A Diagnostic Model of Continental Gross Primary Productivity and Net Primary Productivity. *Global Biogeochem. Cycles* 1996, *10* (2), 269–285.

83. Veroustraete, F.; Sabbe, H.; Eerens, H. Estimation of Carbon Mass Fluxes over Europe Using the C-Fix Model and Euroflux Data. *Remote Sens. Environ.* 2002, *83* (3), 376–399.

84. Running, S. W.; Nemani, R. R.; Heinsch, F. A.; Zhao, M.; Reeves, M.; Hashimoto, H. A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. *Bioscience* 2004, *54* (6), 547.

85. Xiao, X.; Zhang, Q.; Braswell, B.; Urbanski, S.; Boles, S.; Wofsy, S.; Moore, B.; Ojima, D. Modeling Gross Primary Production of Temperate Deciduous Broadleaf Forest Using Satellite Images and Climate Data. *Remote Sens. Environ.* 2004, *91*, 256–270.

86. Wu, C.; Niu, Z.; Gao, S. Gross Primary Production Estimation from MODIS Data with Vegetation Index and Photosynthetically Active Radiation in Maize. *J. Geophys. Res. Atmos.* 2010, *115*, 1–11.

87. Rossini, M.; Migliavacca, M.; Galvagno, M.; Meroni, M.; Cogliati, S.; Cremonese, E.; Fava, F.; Gitelson, A.; Julitta, T.; di Cella, U. M.; Siniscalco, C.; Colombo, R. Remote Estimation of Grassland Gross Primary Production during Extreme Meteorological Seasons. *Int. J. Appl. Earth Obs. Geoinf.* 2014, *29*, 1–10.

88. Ogutu, B. O.; Dash, J. Assessing the Capacity of Three Production Efficiency Models in Simulating Gross Carbon Uptake across Multiple Biomes in Conterminous USA. *Agric. For. Meteorol.* 2013, *174–175*, 158–169.

89. Potter, C.; Klooster, S.; Tan, P.; Steinbach, M.; Kumar, V.; Genovese, V. Variability in Terrestrial Carbon Sinks over Two Decades: Part 2 - Eurasia. *Glob. Planet. Change* 2005, *49*, 177–186.

90. Cao, M.; Prince, S. D.; Small, J.; Goetz, S. J. Remotely Sensed Interannual Variations and Trends in Terrestrial Net Primary Productivity 1981?2000. *Ecosystems* 2004, *7*, 233–242.

91. Lafont, S.; Kergoat, L.; Erard, G.; Chevillard, A.; Karstens, U.; Kolle, O. Spatial and Temporal Variability of Land CO 2 Fluxes Estimated with Remote Sensing and Analysis Data. *Tellus* 2002, *54B*, 820–833.

92. Heinsch, F. A.; Zhao, M.; Running, S. W.; Kimball, J. S.; Nemani, R. R.; Davis, K. J.; Bolstad, P. V.; Cook, B. D.; Desai, A. R.; Ricciuto, D. M.; Law, B. E.; Oechel, W. C.; Kwon, H.; Luo, H.; Wofsy, S. C.; Dunn, A. L.; Munger, J. W.; Baldocchi, D. D.; Xu, L.; Hollinger, D. Y.; Richardson, A. D.; Stoy, P. C.; Siqueira, M. B. S.; Monson, R. K.; Burns, S. P.; Flanagan, L. B. Evaluation of Remote Sensing Based Terrestrial Productivity from MODIS Using Regional Tower Eddy Flux Network Observations. *IEEE Trans. Geosci. Remote Sens.* 2006, *44* (7), 1908–1923.

93. Sasai, T.; Ichii, K.; Yamaguchi, Y.; Nemani, R. Simulating Terrestrial Carbon Fluxes Using the New Biosphere Model "biosphere Model Integrating Eco-Physiological and Mechanistic Approaches Using Satellite Data" (BEAMS). *J. Geophys. Res.* 2005, *110*, 1–18.

94. Verstraeten, W. W.; Veroustraete, F.; Feyen, J. On Temperature and Water Limitation of Net Ecosystem Productivity: Implementation in the C-Fix Model. *Ecol. Modell.* 2006, *199* (1), 4–22.

95. Wang, W.; Dungan, J.; Hashimoto, H.; Michaelis, A. R.; Milesi, C.; Ichii, K.; Nemani, R. R. Diagnosing and Assessing Uncertainties of Terrestrial Ecosystem Models in a Multimodel Ensemble Experiment: 1. Primary Production. *Glob. Chang. Biol.* 2011, *17*, 1350–1366.

96. Cramer, W.; Kicklighter, D. W.; Bondeau, A.; Iii, B. M.; Churkina, G.; Nemry, B.; Ruimy, A.; Schloss, A. L.; Intercomparison, T. P. O. T. P. Comparing Global Models of Terrestrial Net Primary Productivity (NPP): Overview and Key Results. *Glob. Chang. Biol.* 1999, *5* (S1), 1–15.

97. Turner, D. P.; Ritts, W. D.; Cohen, W. B.; Maeirsperger, T. K.; Gower, S. T.; Kirschbaum, A. a.; Running, S. W.; Zhao, M.; Wofsy, S. C.; Dunn, A. L.; Law, B. E.; Campbell, J. L.; Oechel, W. C.; Kwon, H. J.; Meyers, T. P.; Small, E. E.; Kurc, S. a.; Gamon, J. a. Site-Level Evaluation of Satellite-Based Global Terrestrial Gross

Primary Production and Net Primary Production Monitoring. *Glob. Chang. Biol.* 2005, *11* (August 2004), 666–684.

98. Veroustraete, F.; Patyn, J.; Myneni, R. B. Estimating Net Ecosystem Exchange of Carbon Using the Normalized Difference Vegetation Index and an Ecosystem Model. *Remote Sens. Environ.* 1996, *58*, 115–130.

99. Swinnen, E.; Van Hoolst, R.; Toté, C. *GIO Global Land Component - Lot I "Operation of the Global Land Component", Framework Service Contract N° 388533 (JRC), Quality Assessment Report, Dry Matter Productivity (DMP)*; 2014.

100. Smets, B.; Swinnen, E.; Van Hools, R. *GIO-GL Lot 1, Product USer Manual, Dry Matter Productivity (DMP)*; 2015.

101. Maselli, F.; Papale, D.; Puletti, N.; Chirici, G.; Corona, P. Combining Remote Sensing and Ancillary Data to Monitor the Gross Productivity of Water-Limited Forest Ecosystems. *Remote Sens. Environ.* 2009, *113* (3), 657–667.

102. Veroustraete, F.; Sabbe, H.; Rasse, D. P.; Bertels, L. Carbon Mass Fluxes of Forests in Belgium Determined with Low Resolution Optical Sensors. *Int. J. Remote Sens.* 2004, *25* (Woodwell 1984), 769–792.

103. Yuan, W.; Liu, S.; Yu, G.; Bonnefond, J.; Chen, J.; Davis, K.; Desai, A. R.; Goldstein, A. H.; Gianelle, D.; Rossi, F.; Suyker, A. E.; Verma, S. B. Remote Sensing of Environment Global Estimates of Evapotranspiration and Gross Primary Production Based on MODIS and Global Meteorology Data. *Remote Sens. Environ.* 2010, *114* (7), 1416–1431.

104. Beer, C.; Reichstein, M.; Ciais, P.; Farquhar, G. D.; Papale, D. Mean Annual GPP of Europe Derived from Its Water Balance. *Geophys. Res. Lett.* 2007, *34* (5), L05401.

105. Amthor, J. The McCree–de Wit–Penning de Vries–Thornley Respiration Paradigms: 30 Years Later. *Ann. Bot.* 2000, *86* (1), 1–20.

106. MARSOP-3. AGRI4CAST Interpolated Meteorological Data http://mars.jrc.ec.europa.eu/mars/About-us/AGRI4CAST/Data-distribution/AGRI4CAST-Interpolated-Meteorological-Data (accessed May 18, 2015).

107. Westenbroek, S. M.; Kelson, V. A.; Dripps, W. R.; Hunt, R. J.; Bradbury, K. R. *SWB — A Modified Thornthwaite-Mather Soil-Water- Balance Code for Estimating Groundwater Recharge*; 2010.

108. Gobin, A. Modelling Climate Impacts on Crop Yields in Belgium. *Clim. Res.* 2010, *44*, 55–68.

109. Gobin, A. Impact of Heat and Drought Stress on Arable Crop Production in Belgium. *Nat. Hazards Earth Syst. Sci.* 2012, *12*, 1911–1922.

110. Van der Velde, M.; Wriedt, G.; Bouraoui, F. Estimating Irrigation Use and Effects on Maize Yield during the 2003 Heatwave in France☆. *Agric. Ecosyst. Environ.* 2010, *135* (1–2), 90–97.

111. Balaghi, R.; Jlibene, M.; Tychon, B.; Eerens, H. *Agrometeorological Cereal Yield Forecasting in Morocco*; INRA: Morocco, 2013; Vol. 26.

112. VITO. Product Distribution Portal http://www.vito-eodata.be/PDF/portal/Application.html#Home (accessed May 18, 2015).

113. ESA. GlobCover http://due.esrin.esa.int/page_globcover.php (accessed May 18, 2015).

114. JRC. European Soil Database http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB/index.htm (accessed May 18, 2015).

115. FAO. Digital Soil Map of the World http://data.fao.org/map?entryId=446ed430-8383-11db-b9b2-000d939bc5d8 (accessed May 18, 2015).

116. FAO SDRN. AgroMetShell http://www.hoefsloot.com/agrometshell.htm (accessed May 18, 2015).

117. Yildiz, H.; Mermer, A.; Aydogdu, M. Forecasting of Winter Wheat Yield for Turkey Using Water Balance Model. In *Fourth International Conference on Agro-Geoinformatics*; IEEE: Istanbul, 2015; pp 352–356.

118. NTSG. MODIS NPP/GPP Project http://www.ntsg.umt.edu/project/mod17 (accessed Aug 11, 2015).

119. Gobron, N.; Pinty, B.; Aussedat, O.; Chen, J. M.; Cohen, W. B.; Fensholt, R.; Gond, V.; Huemmrich, K. F.; Lavergne, T.; Mélin, F.; Privette, J. L.; Sandholt, I.; Taberner, M.; Turner, D. P.; Verstraete, M. M.; Widlowski, J.-L. Evaluation of Fraction of Absorbed Photosynthetically Active Radiation Products for Different Canopy Radiation Transfer Regimes: Methodology and Results Using Joint Research Center Products Derived from SeaWiFS against Ground-Based Estimations. *J. Geophys. Res.* 2006, *111* (D13), D13110.

120. Wang, K. Acclimation of Photosynthetic Parameters in Scots Pine after Three Years Exposure to Elevated Temperature and CO2. *Agric. For. Meteorol.* 1996, *82* (1–4), 195–217.

121. Kang, E.; Lu, L.; Xu, Z. Vegetation and Carbon Sequestration and Their Relation to Water Resources in an Inland River Basin of Northwest China. *J. Environ. Manage.* 2007, *85* (3), 702–710.

122. Atwell, B. J.; Kriedemann, P. E.; Turnbull, C. G. . *Plants in Action: Adaptation in Nature, Performance in Cultivation*; Macmillan Education Australia: Melbourne, 1999; Vol. 86.

123. Bonhomme, R. Bases and Limits to Using “Degree.day” Units. *Eur. J. Agron.* 2000, *13* (1), 1–10.

124. Brown, D. Crop Heat Units for Corn and Other Warm Season Crops in Ontario. In *Factsheet Ministry of agriculture and food Ontario*; 1969; pp 1–6.

125. Warrington, I. J.; Kanemasu, E. T. Corn Growth Response to Temperature and Photoperiod I. Seedling Emergence, Tassel Initiation, and Anthesis. *Agron. J.* 1983, *75*, 749–754.

126. Warrington, I. J.; Kanemasu, E. T. Corn Growth Response to Temperature and Photoperiod II. Leaf-Initiation and Leaf- Appearance Rates. *Agron. J.* 1983, *75*, 755–761.

127. White, J. W. Modeling Temperature Response in Wheat and Maize. In *Proceedings of a Workshop, CIMMYT, El Bata´n, Mexico, 23–25 April, 2001. NRG-GIS Series 03-01*; Mexico, D.F, CIMMYT, 2001.

128. Yan, W.; Hunt, L. A. An Equation for Modelling the Temperature Response of Plants

Using Only the Cardinal Temperatures. *Ann. Bot.* 1999, *84*, 607–614.

129. Tollenaar, M.; Mihajlovic, M.; Aguilera, A. Temperature Response of Dry Matter Accumulation, Leaf Photosynthesis, and Chlorophyll Fluorescence in an Old and a New Maize Hybrid during Early Development. *Can. J. Plant Sci.* 1991, *71* (2), 353–359.

130. NOAA. CO2 Records http://www.esrl.noaa.gov/gmd/ccgg/trends (accessed Aug 11, 2015).

131. Ainsworth, E. a.; Long, S. P. What Have We Learned from 15 Years of Free-Air CO2 Enrichment (FACE)? A Meta-Analytic Review of the Responses of Photosynthesis, Canopy Properties and Plant Production to Rising CO2. *New Phytol.* 2005, *165* (2), 351–372.

132. Wang, C.; Guo, L.; Li, Y.; Wang, Z. Systematic Comparison of C3 and C4 Plants Based on Metabolic Network Analysis. *BMC Syst. Biol.* 2012, *6*.

133. Yin, X.; Struik, P. C. C3 and C4 Photosynthesis Models: An Overview from the Perspective of Crop Modelling. *NJAS - Wageningen J. Life Sci.* 2009, *57* (1), 27–38.

134. Kim, S.; Gitz, D.; Sicher, R.; Baker, J.; Timlin, D.; Reddy, V. Temperature Dependence of Growth, Development, and Photosynthesis in Maize under Elevated CO2. *Environ. Exp. Bot.* 2007, *61* (3), 224–236.

135. von Caemmerer, S. *Biochemical Models of Leaf Photosynthesis*; CSIRO: Collingwood, 2000.

136. Goetz, S. J.; Prince, S. D.; Goward, S. N.; Thawley, M. M.; Small, J.; Johnston, A. Mapping Net Primary Production and Related Biophysical Variables with Remote Sensing: Application to the BOREAS Region. *J. Geophys. Res.* 1999, *104* (D22), 27719.

137. Prince, S. D.; Goward, S. N. Global Primary Production: A Remote Sensing Approach. *J. Biogeogr.* 1995, *22*, 815–835.

138. Dabrowska-Zielinska, K.; Kowalik, W.; Budzynska, M.; Malek, I. Assessment of Evapotranspiration and Soil Moisture for Biebrza Wetlands Using Thermal Remote Sensing and In- Situ Data . In; Halounová, L., Ed.; EARSel, 2011; pp 281–288.

139. Duveiller, G.; López-Lozano, R.; Baruth, B. Enhanced Processing of 1-Km Spatial Resolution fAPAR Time Series for Sugarcane Yield Forecasting and Monitoring. *Remote Sens.* 2013, *5* (3), 1091–1116.

140. Legates, D. R.; Mccabe, G. J. A Refined Index of Model Performance: A Rejoinder. *Int. J. Climatol.* 2013, *33* (4), 1053–1056.

141. Zhang, Y.; Xu, M.; Chen, H.; Adams, J. Global Pattern of NPP to GPP Ratio Derived from MODIS Data: Effects of Ecosystem Type, Geographical Location and Climate. *Glob. Ecol. Biogeogr.* 2009, *18* (3), 280–290.

142. Richter, K.; Atzberger, C.; Hank, T. B.; Mauser, W. Derivation of Biophysical Variables from Earth Observation Data: Validation and Statistical Measures. *J. Appl. Remote Sens.* 2012, *6* (SEPTEMBER), 063557–1.

143. Yuan, W.; Liu, S.; Yu, G.; Bonnefond, J. M.; Chen, J.; Davis, K.; Desai, A. R.; Goldstein, A. H.; Gianelle, D.; Rossi, F.; Suyker, A. E.; Verma, S. B. Global Estimates of Evapotranspiration and Gross Primary Production Based on MODIS and Global Meteorology Data. *Remote Sens. Environ.* 2010, *114* (7), 1416–1431.

144. Salazar, L.; Kogan, F.; Roytman, L. Use of Remote Sensing Data for Estimation of Winter Wheat Yield in the United States. *Int. J. Remote Sens.* 2007, *28* (November 2013), 3795–3811.

145. Mkhabela, M. S.; Mkhabela, M. S.; Mashinini, N. N. Early Maize Yield Forecasting in the Four Agro-Ecological Regions of Swaziland Using NDVI Data Derived from NOAA's-AVHRR. *Agric. For. Meteorol.* 2005, *129*, 1–9.

146. Sakamoto, T.; Gitelson, A. a.; Arkebauer, T. J. MODIS-Based Corn Grain Yield Estimation Model Incorporating Crop Phenology Information. *Remote Sens. Environ.* 2013, *131*, 215–231.

147. Hansen, J. W.; Indeje, M. Linking Dynamic Seasonal Climate Forecasts with Crop Simulation for Maize Yield Prediction in Semi-Arid Kenya. *Agric. For. Meteorol.* 2004, *125*, 143–157.

148. Klein, A. Bulletin Agro-Météorologique : Ajustement Des Dates de Semis, Université de Liège, 2014.

149. Hu, G.; Jia, L.; Menenti, M. Comparison of MOD16 and LSA-SAF MSG Evapotranspiration Products over Europe for 2011. *Remote Sens. Environ.* 2015, *156*, 510–526.

150. LSA SAF. Evapotranspiration http://landsaf.meteo.pt/algorithms.jsp?seltab=7&starttab=7 (accessed Aug 11, 2015).

151. Martens, B.; Miralles, D. Gleam https://www.gleam.eu/ (accessed Jun 1, 2018).

152. Ise, T.; Moorcroft, P. R. Simulating Boreal Forest Dynamics from Perspectives of Ecophysiology, Resource Availability, and Climate Change. *Ecol. Res.* 2010, *25* (3), 501–511.

153. Procter, A. C. Effects of Past and Future CO2 on Grassland Soil Carbon and Microbial Ecology, 2012.

154. Polley, H. W.; Derner, J. D.; Jackson, R. B.; Wilsey, B. J.; Fay, P. A. Impacts of Climate Change Drivers on C4 Grassland Productivity: Scaling Driver Effects through the Plant Community. *J. Exp. Bot.* 2014, *65* (13), 3415–3424.

155. Waraich, E. A.; Ahmad, R.; Halim, A.; Aziz, T. Alleviation of Temperature Stress by Nutrient Management in Crop Plants : A Review. 2012, *12* (2), 221–244.

156. Gitelson, A. a.; Peng, Y.; Masek, J. G.; Rundquist, D. C.; Verma, S.; Suyker, A.; Baker, J. M.; Hatfield, J. L.; Meyers, T. Remote Estimation of Crop Gross Primary Production with Landsat Data. *Remote Sens. Environ.* 2012, *121*, 404–414.

157. Gitelson, A. a.; Viña, A.; Masek, J. G.; Verma, S. B.; Suyker, A. E. Synoptic Monitoring of Gross Primary Productivity of Maize Using Landsat Data. *IEEE Geosci. Remote Sens. Lett.* 2008, *5* (2), 133–137.

158. Peng, Y.; Gitelson, A. a. Application of Chlorophyll-Related Vegetation Indices for Remote Estimation of Maize Productivity. *Agric. For. Meteorol.* 2011, *151* (9), 1267–1276.

159. Phillips, L.; Hansen, A.; Flather, C. Evaluating the Species Energy Relationship with the Newest Measures of Ecosystem Energy: NDVI versus MODIS Primary Production. *Remote Sens. Environ.* 2008, *112* (12), 4381–4392.

160. EPA. Climate Impacts on Agriculture and Food Supply http://www.epa.gov/climatechange/impacts-adaptation/agriculture.html (accessed Feb 10, 2015).

161. Deryng, D.; Conway, D.; Ramankutty, N.; Price, J.; Warren, R. Global Crop Yield Response to Extreme Heat Stress under Multiple Climate Change Futures. *Environ. Res. Lett.* 2014, *9*, 34011.

162. Luterbacher, J.; Dietrich, D.; Xoplaki, E.; Grosjean, M.; Wanner, H. European Seasonal and Annual Temperature Variability, Trends, and Extremes since 1500. *Science* 2004, *303* (March), 1499–1503.

163. Duveiller, G.; Lopez-Lozano, R.; Cescatti, A. Exploiting the Multi-Angularity of the MODIS Temporal Signal to Identify Spatially Homogeneous Vegetation Cover: A Demonstration for Agricultural Monitoring Applications. *Remote Sens. Environ.* 2015, *166*, 61–77.

164. Baruth, B.; Kucera, L. Crop Masking – Needs for the MARS Crop Yield Forecasting System. In *ISPRS Archives XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates*; Baruth, B., Royer, A., Genovese, G., Eds.; ISPRS: Ispra, 2007; pp 3–6.

165. Liu, M. W.; Ozdogan, M.; Zhu, X. Crop Type Classification by Simultaneous Use of Satellite Images of Different Resolutions. *Geosci. Remote Sensing, IEEE Trans.* 2014, *52* (6), 3637–3649.

166. Reynolds, C. A.; Yitayew, M.; Slack, D. C.; Hutchinson, C. F.; Huete, A.; Petersen, M. S. Estimating Crop Yields and Production by Integrating the FAO Crop Specific Water Balance Model with Real-Time Satellite Data and Ground-Based Ancillary Data. *Int. J. Remote Sens.* 2000, *21* (18), 3487–3508.

167. Bolton, D. K.; Friedl, M. A. Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics. *Agric. For. Meteorol.* 2013, *173*, 74–84.

168. Wardlow, B.; Egbert, S.; Kastens, J. Analysis of Time-Series MODIS 250 M Vegetation Index Data for Crop Classification in the U.S. Central Great Plains. *Remote Sens. Environ.* 2007, *108* (3), 290–310.

169. Ozdogan, M.; Yang, Y.; Allez, G.; Cervantes, C. Remote Sensing of Irrigated Agriculture: Opportunities and Challenges. *Remote Sens.* 2010, *2* (9), 2274–2304.

170. Aurdal, L.; Ragnar Bang Huseby; Eikvil, L.; Solberg, R.; Vikhamar, D.; Solberg, A. Use of Hidden Markov Models and Phenology for Multitemporal Satellite Image Classification: Applications to Mountain Vegetation Classification. In *Proceedings of the IEEE third international workshop on the analysis of multitemporal remote sensing images*; IEEE, 2005; pp 220–224.

171. Carrão, H.; Gonçalves, P.; Caetano, M. Contribution of Multispectral and Multitemporal Information from MODIS Images to Land Cover Classification. *Remote Sens. Environ.* 2008, *112*, 986–997.

172. Jakubauskas, M. E.; Legates, D. R.; Kastens, J. H. Crop Identification Using Harmonic Analysis Oftime-Series AVHRR NDVI Data. *Comput. Electron. Agric.* 2002, *37*, 127–139.

173. Sakamoto, T.; Yokozawa, M.; Toritani, H.; Shibayama, M.; Ishitsuka, N.; Ohno, H. A Crop Phenology Detection Method Using Time-Series MODIS Data. *Remote Sens.*

Environ. 2005, *96* (3–4), 366–374.

174. Dong, J.; Xiao, X.; Kou, W.; Qin, Y.; Zhang, G.; Li, L.; Jin, C.; Zhou, Y.; Wang, J.; Biradar, C.; Liu, J.; Moore, B. Tracking the Dynamics of Paddy Rice Planting Area in 1986–2010 through Time Series Landsat Images and Phenology-Based Algorithms. *Remote Sens. Environ.* 2015, *160*, 99–113.

175. Xavier, A. C.; Rudorff, B. F. T.; Shimabukuro, Y. E.; Berka, L. M. S.; Moreira, M. A. Multi-temporal Analysis of MODIS Data to Classify Sugarcane Crop. *Int. J. Remote Sens.* 2006, *27* (4), 755–768.

176. Alcantara, C.; Kuemmerle, T.; Prishchepov, A. V.; Radeloff, V. C. Mapping Abandoned Agriculture with Multi-Temporal MODIS Satellite Data. *Remote Sens. Environ.* 2012, *124*, 334–347.

177. Jönsson, P.; Eklundh, L. TIMESAT - A Program for Analyzing Time-Series of Satellite Sensor Data. *Comput. Geosci.* 2004, *30* (8), 833–845.

178. Foerster, S.; Kaden, K.; Foerster, M.; Itzerott, S. Crop Type Mapping Using Spectral-Temporal Profiles and Phenological Information. *Comput. Electron. Agric.* 2012, *89*, 30–40.

179. Dixon, B.; Candade, N. Multispectral Landuse Classification Using Neural Networks and Support Vector Machines: One or the Other, or Both? *Int. J. Remote Sens.* 2008, *29* (4), 1185–1206.

180. Yang, C.; Everitt, J. H.; Murden, D. Evaluating High Resolution SPOT 5 Satellite Imagery for Crop Identification. *Comput. Electron. Agric.* 2011, *75* (2), 347–354.

181. Petitjean, F.; Inglada, J.; Gançarski, P. Satellite Image Time Series Analysis Under Time Warping. *IEEE Trans. Geosci. Remote Sens.* 2012, *50* (8), 3081–3095.

182. 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. In *Proceedings of the 4th GEOBIA*; 2012; pp 35–40.

183. Thenkabail, P. S.; Gangadhararao, P.; Biggs, T. W.; Krishna, M.; Turral, H. Spectral Matching Techniques to Determine Historical Land-Use / Land-Cover (LULC) and Irrigated Areas Using Time-Series 0 . 1-Degree AVHRR Pathfinder Datasets. *Photogramm. Eng. Remote Sens.* 2007, *73* (9), 1029–1040.

184. Gumma, M. K. Mapping Rice Areas of South Asia Using MODIS Multitemporal Data. *J. Appl. Remote Sens.* 2011, *5* (1), 53547.

185. Siachalou, S.; Mallinis, G.; Tsakiri-Strati, M. A Hidden Markov Models Approach for Crop Classification: Linking Crop Phenology to Time Series of Multi-Sensor Remote Sensing Data. *Remote Sens.* 2015, *7* (4), 3633–3650.

186. Medhavy, T. T.; Sharma, T.; Dubey, R. P.; Hooda, R. S.; Mothikumar, K. E.; Yadav, M.; Manchanda, M. L.; Ruhal, D. S.; Khera, A. P.; Jarwal, S. D. Crop Classification Accuracy as Influenced by Training Strategy, Data Transformation and Spatial Resolution of Data. *J. Indian Soc. Remote Sens.* 1993, *21* (1), 21–28.

187. Dadhwal, V. K.; Ruhal, D. S.; Medhavy, T. T.; Jarwal, S. D.; Khera, A. P.; Singh, J.; Sharma, T.; Parihar, J. S. Wheat Acreage Estimation for Haryana Using Satellite Digital Data. *J. Indian Soc. Remote Sens.* 1991, *19* (1), 1–15.

188. Chmielewski, F.-M.; Müller, A.; Bruns, E. Climate Changes and Trends in Phenology

of Fruit Trees and Field Crops in Germany, 1961–2000. *Agric. For. Meteorol.* 2004, *121* (1–2), 69–78.

189. Van Niel, T. G.; McVicar, T. R.; Fang, H.; Liang, S. Calculating Environmental Moisture for per-Field Discrimination of Rice Crops. *Int. J. Remote Sens.* 2003, *24*, 885–890.

190. Lu, D.; Weng, Q. A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *Int. J. Remote Sens.* 2007, *28* (5), 823–870.

191. Cools, M.; Moons, E.; Creemers, L.; Wets, G. Changes in Travel Behavior in Response to Weather Conditions. In *CD Proceedings of the 89th Annual Meeting of the Transportation Research Board*; Washington D.C., 2010.

192. JECAM. Russia -- Stavropol Kray http://www.jecam.org/?/site-description/russia-stavropol-kray (accessed Jun 14, 2016).

193. JECAM. Ukraine - Kyiv http://www.jecam.org/?/site-description/ukraine (accessed Jun 14, 2016).

194. JECAM. Brazil - São Paulo http://www.jecam.org/?/site-description/Brazil-Sao-Paulo-Itatinga (accessed Jun 14, 2016).

195. Peeters, A. Country pasture/forage resource profile for Belgium http://www.fao.org/ag/AGP/AGPC/doc/Counprof/Belgium/belgium.htm (accessed Jun 14, 2016).

196. Słowińska-Jurkiewicz, A.; Bryk, M.; Medvedev, V. V. Long-Term Organic Fertilization Effect on Chernozem Structure. 2013, No. 2000, 81–87.

197. JECAM. Belgium http://www.jecam.org/?/site-description/belgium (accessed Jun 14, 2016).

198. AMIS. *AMIS Crop Calendar*; Rome, 2012.

199. AGIV. GDI-Vlaanderen Landbouwgebruikspercelen ALV, 2014 http://www.geopunt.be/download?container=landbouwgebruikspercelen&title=Landbouwgebruikspercelen 2014 (accessed Feb 24, 2016).

200. Verhoeve, A.; Dewaelheyns, V.; Kerselaers, E.; Rogge, E.; Gulinck, H. Virtual Farmland: Grasping the Occupation of Agricultural Land by Non-Agricultural Land Uses. *Land use policy* 2015, *42*, 547–556.

201. De Wit, A. J. W.; Clevers, J. G. P. W. Efficiency and Accuracy of per-Field Classification for Operational Crop Mapping. *Int. J. Remote Sens.* 2004, *25* (20), 4091–4112.

202. Rembold, F.; Meroni, M.; Urbano, F.; Royer, A.; Atzberger, C.; Lemoine, G.; Eerens, H.; Haesen, D.; Aidco, D. G.; Klisch, A. Remote Sensing Time Series Analysis for Crop Monitoring with the SPIRITS Software: New Functionalities and Use Examples. *Front. Environ. Sci.* 2015, *3* (July), 129–134.

203. Swets, D.; Reed, B.; Rowland, J.; Marko, S. A Weighted Least-Squares Approach to Temporal NDVI Smoothing. *Proc. 1999 ASPRS Annu. Conf. From Image to Information, Portland, Oregon, May 17-21 1999* 1999, No. Figure 1.

204. Geerken, R. A. An Algorithm to Classify and Monitor Seasonal Variations in Vegetation Phenologies and Their Inter-Annual Change. *ISPRS J. Photogramm. Remote Sens.* 2009, *64* (4), 422–431.

205. Homayouni, S.; Roux, M. Hyperspectral Image Analysis for Material Mapping Using Spectral Matching. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 2004, *35*, 1682–1750.

206. Zhang, X.; Friedl, M.; Schaaf, C.; Strahler, A. Monitoring Vegetation Phenology Using MODIS. *Remote Sens. Environ.* 2003, *84*, 471–475.

207. Cheema, M. J. M.; Bastiaanssen, W. G. M. Land Use and Land Cover Classification in the Irrigated Indus Basin Using Growth Phenology Information from Satellite Data to Support Water Management Analysis. *Agric. Water Manag.* 2010, *97* (10), 1541–1552.

208. Lambert, M.; Waldner, F.; Defourny, P. Cropland Mapping over Sahelian and Sudanian Agrosystems : A Knowledge-Based Approach Using PROBA-V Time Series at 100-M. 2016.

209. Knight, J. F.; Lunetta, R. S.; Ediriwickrema, J.; Khorram, S. Regional Scale Land Cover Characterization Using MODIS-NDVI 250 M Multi-Temporal Imagery: A Phenology-Based Approach. *GIScience Remote Sens.* 2007, *43* (1), 1–23.

210. Zhong, L.; Gong, P.; Biging, G. S. Phenology-Based Crop Classification Algorithm and Its Implications on Agricultural Water Use Assessments in California's Central Valley. *Photogramm. Eng. Remote Sens.* 2012, *78* (8), 799–813.

211. Doraiswamy, P. C.; Akhmedov, B.; Beard, L.; Stern, A.; Mueller, R. Operational Prediction of Crop Yields Using Modis Data and Products. In *ISPRS Archives XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates*; Baruth, B., Royer, A., Genovese, G., Eds.; ISPRS: Ispra, 2007; pp 45–49.

212. Junior, J. Z.; Coltri, P. P.; do Valle Gonçalves, R. R.; Romani, L. A. S. Multi-Resolution in Remote Sensing for Agricultural Monitoring: A Review. *Rev. Bras. Cartogr. a* 2014, *66/7*, 1517–1529.

213. Kibet, S. Assessing Field-Level Maize Yield Variability in Tanzania Using Multi-Temporal Very Fine Resolution Imagery. 2016.

214. Duveiller, G.; Defourny, P. A Conceptual Framework to Define the Spatial Resolution Requirements for Agricultural Monitoring Using Remote Sensing. *Remote Sens. Environ.* 2010, *114* (11), 2637–2650.

215. Zheng, Y.; Zhang, M.; Zhang, X.; Zeng, H.; Wu, B. Mapping Winter Wheat Biomass and Yield Using Time Series Data Blended from PROBA-V 100- and 300-M S1 Products. *Remote Sens.* 2016, *8* (10), 824.

216. Bégué, A.; Lebourgeois, V.; Bappel, E.; Todoroff, P.; Pellegrino, A.; Baillarin, F.; Siegmund, B. Spatio-Temporal Variability of Sugarcane Fields and Recommendations for Yield Forecast Using NDVI. *Int. J. Remote Sens.* 2010, *31* (20), 5391–5407.

217. Rasmussen, M. S. Assessment of Millet Yields and Production in Northern Burkina Faso Using Integrated NDVI from the AVHRR. *Int. J. Remote Sens.* 1992, *13* (18), 3431–3442.

218. Meroni, M.; Marinho, E.; Sghaier, N.; Verstrate, M.; Leo, O. Remote Sensing Based Yield Estimation in a Stochastic Framework — Case Study of Durum Wheat in Tunisia. *Remote Sens.* 2013, *5*, 539–557.

219. PINTER, P. J.; JACKSON, R. D.; IDSO, S. B.; REGINATO, R. J. Multidate Spectral Reflectance as Predictors of Yield in Water Stressed Wheat and Barley. *Int. J. Remote Sens.* 1981, *2* (1), 43–48.

220. Meroni, M.; Marinho, E.; Sghaier, N.; Verstrate, M. M.; Leo, O. Remote Sensing Based Yield Estimation in a Stochastic Framework - Case Study of Durum Wheat in Tunisia. *Remote Sens.* 2013, *5* (2), 539–557.

221. Duveiller, G.; Baret, F.; Defourny, P. Using Thermal Time and Pixel Purity for Enhancing Biophysical Variable Time Series: An Interproduct Comparison. *IEEE Trans. Geosci. Remote Sens.* 2013, *51* (4), 2119–2127.

222. Franch, B.; Vermote, E. F.; Becker-Reshef, I.; Claverie, M.; Huang, J.; Zhang, J.; Justice, C.; Sobrino, J. A. Improving the Timeliness of Winter Wheat Production Forecast in the United States of America, Ukraine and China Using MODIS Data and NCAR Growing Degree Day Information. *Remote Sens. Environ.* 2015, *161* (May), 131–148.

223. Kouadio, L.; Duveiller, G.; Djaby, B.; El Jarroudi, M.; Defourny, P.; Tychon, B. Estimating Regional Wheat Yield from the Shape of Decreasing Curves of Green Area Index Temporal Profiles Retrieved from MODIS Data. *Int. J. Appl. Earth Obs. Geoinf.* 2012, *18* (1), 111–118.

224. Lobell, D. B.; Asner, G. P.; Ortiz-Monasterio, J. I.; Benning, T. L. Remote Sensing of Regional Crop Production in the Yaqui Valley, Mexico: Estimates and Uncertainties. *Agric. Ecosyst. Environ.* 2003, *94* (2), 205–220.

225. Duveiller, G. Caveats in Calculating Crop Specific Pixel Purity for Agricultural Monitoring Using MODIS Time Series. *Proc. SPIE Remote Sens. Agric. Ecosyst. Hydrol. XIV, 24-27 Sept. 2012, Edinburgh, UK.* 2012, *8531* (September 2012), 85310J–85310J–10.

226. Kastens, J. H.; Kastens, T. L.; Kastens, D. L. A.; Price, K. P.; Martinko, E. A.; Lee, R. Y. Image Masking for Crop Yield Forecasting Using AVHRR NDVI Time Series Imagery. *Remote Sens. Environ.* 2005, *99* (3), 341–356.

227. Löw, F.; Duveiller, G. Defining the Spatial Resolution Requirements for Crop Identification Using Optical Remote Sensing. *Remote Sens.* 2014, *6* (9), 9034–9063.

228. Duveiller, G.; Baret, F.; Defourny, P. Crop Specific Green Area Index Retrieval from MODIS Data at Regional Scale by Controlling Pixel-Target Adequacy. *Remote Sens. Environ.* 2011, *115* (10), 2686–2701.

229. Zhong, L.; Gong, P.; Biging, G. S. Efficient Corn and Soybean Mapping with Temporal Extendability: A Multi-Year Experiment Using Landsat Imagery. *Remote Sens. Environ.* 2014, *140*, 1–13.

230. Guindin-Garcia, N. Estimating Maize Grain Yield from Crop Biophysical Parameters Using Remote Sensing. 2010, 115.

231. Chen, P.-Y.; Fedosejevs, G.; Tiscareño-LóPez, M.; Arnold, J. G. Assessment of MODIS-EVI, MODIS-NDVI and VEGETATION-NDVI Composite Data Using Agricultural Measurements: An Example at Corn Fields in Western Mexico. *Environ. Monit. Assess.* 2006, *119* (1–3), 69–82.

232. Durgun, Y. Ö.; Gobin, A.; Van De Kerchove, R.; Tychon, B. Crop Area Mapping Using 100-M Proba-V Time Series. *Remote Sens.* 2016, *8* (7), 585.

233. Kempeneers, P.; Sedano, F.; Piccard, I.; Eerens, H. Data Assimilation of PROBA-V 100 and 300 M. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2016, *9* (7), 3314–3325.

234. Battude, M.; Al Bitar, A.; Morin, D.; Cros, J.; Huc, M.; Marais Sicre, C.; Le Dantec, V.; Demarez, V. Estimating Maize Biomass and Yield over Large Areas Using High Spatial and Temporal Resolution Sentinel-2 like Remote Sensing Data. *Remote Sens. Environ.* 2016, *184*, 668–681.

235. Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J. F.; Ceschia, E. Understanding the Temporal Behavior of Crops Using Sentinel-1 and Sentinel-2-like Data for Agricultural Applications. *Remote Sens. Environ.* 2017, *199*, 415–426.

236. Skakun, S.; Vermote, E.; Roger, J.-C.; Franch, B. Combined Use of Landsat-8 and Sentinel-2A Images for Winter Crop Mapping and Winter Wheat Yield Assessment at Regional Scale. *AIMS Geosci.* 2017, *3* (2), 163–186.

237. Donlon, C.; Berruti, B.; Buongiorno, A.; Ferreira, M.-H.; Féménias, P.; Frerick, J.; Goryl, P.; Klein, U.; Laur, H.; Mavrocordatos, C.; Nieke, J.; Rebhan, H.; Seitz, B.; Stroede, J.; Sciarra, R. The Global Monitoring for Environment and Security (GMES) Sentinel-3 Mission. *Remote Sens. Environ.* 2012, *120* (2012), 37–57.

238. Donlon, C. *Sentinel-3 Mission Requirements Traceability Document (MRTD)*; 2011.

239. ESA Sentinel-2 Team. *GMES Sentinel-2 Mission Requirements Document*; 2007.

240. Toulios, L. Remote Sensing As a Service Tool for the Detection of Water Resources and Crop Conditions Monitoring. In *Towards Climatic Services. International scientific conference*; Šiška, B., Nejedlík, P., Eliášová, M., Eds.; Slovak University of Agriculture, Nitra: Nitra, Slovakia, 2015.