Use of remote sensing approaches for agricultural applications

by

Luciana Nieto

B.S., National University of Mar del Plata, 2017

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Abstract

Remote sensing is a technology that has been utilized extensively in agriculture due to its capacity to provide precise and detailed information on various aspects of agricultural production. Farmers and researchers have utilized this technology to gain valuable insights regarding, among other things, crop phenology, yield prediction, land classification, soil quality, water management, and environmental monitoring. The present dissertation is structured into six chapters, with the first serving as an introduction to remote sensing technology in agriculture and the last chapter offering concluding remarks. The remaining chapters delve into various applications of remote sensing technology in agriculture.

The second and third chapters examine the potential of remote sensing to classify maize phenology in Kansas, utilizing three distinct image resolutions. The second chapter identifies the optimal combination of spectral bands, vegetation indices, and weather data for phenology classification using Landsat 8 as a source of spectral information. In chapter three, greater temporal and spatial resolution was tested using Sentinel-2 and Planet Fusion, and the classification performance of both sources was compared. The model was tested in different areas, and the results emphasized the significance of temporal and spatial resolution for traits like phenology that can change rapidly.

Chapter four explores the use of remote sensing technology to identify areas in Cambodia with traditional management practices where conservation agriculture could play a critical role. The study employs 3-meter daily imagery from Planet Fusion and image segmentation tools to distinguish between burned patches and bare soil after ploughing. The results demonstrate that these images, in conjunction with image segmentation tools, have the potential to identify management practices in areas where obtaining ground-truth data could be challenging.

Finally, chapter five discusses the current state of the art and the necessary changes to integrate soil science methods and remote sensing for determining soil organic carbon. This chapter examines the challenges and opportunities associated with using remote sensing to monitor soil properties and offers viable solutions for bridging these two domains.

Use of remote sensing approaches for agricultural applications

by

Luciana Nieto

B.S., National University of Mar del Plata, 2017

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Approved by:

Major Professor
Dr. Ignacio A. Ciampitti

Copyright

© Luciana Nieto 2023.

Abstract

Remote sensing is a technology that has been utilized extensively in agriculture due to its capacity to provide precise and detailed information on various aspects of agricultural production. Farmers and researchers have utilized this technology to gain valuable insights regarding, among other things, crop phenology, yield prediction, land classification, soil quality, water management, and environmental monitoring. The present dissertation is structured into six chapters, with the first serving as an introduction to remote sensing technology in agriculture and the last chapter offering concluding remarks. The remaining chapters delve into various applications of remote sensing technology in agriculture.

The second and third chapters examine the potential of remote sensing to classify maize phenology in Kansas, utilizing three distinct image resolutions. The second chapter identifies the optimal combination of spectral bands, vegetation indices, and weather data for phenology classification using Landsat 8 as a source of spectral information. In chapter three, greater temporal and spatial resolution was tested using Sentinel-2 and Planet Fusion, and the classification performance of both sources was compared. The model was tested in different areas, and the results emphasized the significance of temporal and spatial resolution for traits like phenology that can change rapidly.

Chapter four explores the use of remote sensing technology to identify areas in Cambodia with traditional management practices where conservation agriculture could play a critical role. The study employs 3-meter daily imagery from Planet Fusion and image segmentation tools to distinguish between burned patches and bare soil after ploughing. The results demonstrate that these images, in conjunction with image segmentation tools, have the potential to identify management practices in areas where obtaining ground-truth data could be challenging.

Finally, chapter five discusses the current state of the art and the necessary changes to integrate soil science methods and remote sensing for determining soil organic carbon. This chapter examines the challenges and opportunities associated with using remote sensing to monitor soil properties and offers viable solutions for bridging these two domains.

Table of Contents

List of Figures xi

List of Tables xv

Acknowledgements xvi

Dedication xvii

1 Remote sensing and agronomy 1

1.1 Fundamentals of remote sensing 1

1.2 Understanding satellite images 3

1.2.1 Earth Observation satellites and fusion products 4

1.3 Spectral indices 7

1.4 A glimpse into the future 8

1.5 Dissertation objectives 9

2 An integrated approach of field, weather, and satellite data for monitoring maize phenology 14

2.1 Introduction 15

2.2 Materials and Methods 18

2.2.1 Study area and environmental conditions 18

2.2.2 Data gathering and feature engineering 19

2.2.3 Model training and validation, random forest classification model . . 21

2.2.4 Model testing 23

2.3 Results 26

2.3.1 Best combination of features and model performance 26
2.3.2 Classification using VIVPD model 27
2.3.3 Classification using VIVPD model across years 27
2.3.4 Classification using VIVPD model across space 28
2.4 Discussion 28
2.5 Conclusion 31

3 Impact of High-Cadence Earth Observation in Maize Crop Phenology Classification 38
3.1 Introduction 40
3.2 Materials and Methods 43
3.2.1 Study area 43
3.2.2 Data characteristics 44
3.2.3 Data preparation 48
3.2.4 Random forest 49
3.3 Results 52
3.3.1 Best combination of variables 52
3.3.2 Model performance using PF data in both regions 53
3.3.3 Model performance using Sentinel-2 data 53
3.4 Discussion 54
3.5 Conclusions 57

4 Using EfficientNet Pretrained Model for Segmentation-Based Classification of Agricultural Land Features: a case study in Battambang, Cambodia 63
4.1 Introduction 64
4.2 Materials and methods 67
4.2.1 Study area 67
4.2.2 Satellite imagery and vegetation indices 68
4.2.3 Analysis 69

4.2.4 Metrics . . . 73
4.3 Results . . . 74
4.3.1 Change detection . . . 74
4.3.2 Segmentation results . . . 74
4.4 Discussion . . . 75
4.5 Conclusion . . . 77

5 Limitations and future perspectives for satellite-based soil carbon monitoring . . 83
5.1 Introduction . . . 84
5.2 Current scenario: progress and global footprint . . . 85
5.3 Limitations of current approaches from ground collection to remote sensing . 87
5.4 Future perspectives . . . 93
5.5 Conclusions . . . 95

6 Conclusions . . . 97
6.1 Best combination (Chapter 2) . . . 98
6.2 Resolution matters (Chapter 3) . . . 99
6.3 Finding patterns (Chapter 4) . . . 100
6.4 Why do we need protocols? (Chapter 5) . . . 101
6.5 Looking ahead . . . 102

Bibliography . . . 103

A Appendix A, Chapter 2 . . . 137

B Appendix B Chapter 4 . . . 145

List of Figures

1.1 Interaction of the incident radiation with the target object, plant, showing the proportion absorbed, transmitted, and reflected. Right panel show the electromagnetic spectrum with the different proportions of interest. 11

1.2 representing the different types of sensors based on the source of energy. Left side passive sensor where the energy comes from the sun; right side shows the counterpart, where the energy is emitted by the sensors. 12

1.3 Example with the most extended satellites (and constellations) used for agricultural purposes. Temporal and Spatial refers to the resolutions and bands was used as a short form for the spectral resolution provided by each satellite. This list is only for explanatory purposes, it is by no means an extensive one. 13

2.1 a) area of study highlighting, the USA, in green and Kansas in blue (b) Area of study. The purple points correspond to phenology measurements in Kansas's Southwest Agricultural District; the orange area corresponds to Landsat 8 tile 30/34, which covers the area under study. (c) Zoomed-in highlight of the Southwest Agricultural District (light blue area) and the location of each phenology measurement. Maps were created using ArcGIS software version 10.7 by Esri. ArcGIS and ArcMap are the intellectual property of Esri and are used herein under license. Copyright Esri. All rights reserved. For more information about Esri software, please visit www.esri.com. 33

2.2 Accuracy values for each model during the season from 2013 to 2018. The black horizontal line shows the lowest value recorded. B = spectral bands 2, 3, 4, 5, 6, 7, 10, and 11 for Landsat 8; BW = spectral bands in addition to precipitation, maximum and minimum temperature, VPD, and GDU (weather parameters); VIB = spectral bands in addition to NDVI, EVI, GCVI, GVMI, and NDWI (vegetation indices); VIVPD = vegetation indices and vapor pressure deficit (VPD); W = weather parameters; VI = vegetation indices; FULL = spectral bands, weather parameters, and vegetation indices; LLD = latitude, longitude and doy. 34

2.3 For the years 2013 to 2018, a classification matrix was created that included all classified elements, their magnitude, as well as precision and recall scores. The true classes are on the Y axe, while the model's predictions are on the X axe. The number of fields classified is represented visually by the size of the bubbles; the numbers along the main diagonal represent those successfully classified. 35

3.1 (a) Area of study (a.1) state of Kansas (US), (a.2) Southwest (SW) region; the large orange rectangle represents the Sentinel-2 tile covering the area; the green rectangle corresponds with the Planet Fusion (PF) tiles over the area; pink dots represent ground truth field data collection sites, (a.3) Central KS (CK) region; green rectangle corresponds with the PF tile over the area; yellow dots represent ground truth field data collection sites. (b) Maize phenology stages, with the phenology scale used during ground truth data collection, following Ciampitti et al. (2016). (c) Sentinel-2 and PF data imagery and GRIDMET data from the Google Earth Engine (GEE) repository, data availability differs temporally with 8 images collected for Sentinel-2, and daily data for both PF and weather data during the 2017 growing season, varying with the day of the year (DOY). 59

3.2 EVI temporal series for the 2017 maize growing season for SW fields using Sentinel-2 (solid green line) and PF (dash blue line). The shaded area around both lines represent the standard deviation of the EVI values for the fields under study (pink dots 3.1 (a.2)). Maize phenology progress and development stages is depicted on the x-axis. 60

3.3 (a) Confusion matrices with PF-based classification results for the Southwest region (SW) and for (b) Central KS region (CK), including recall and precision for each crop phenology stage, and F1-score and overall accuracy (OA). . . . 61

3.4 Confusion matrices with S-2 based classification results for the Southwest (SW) region, (a) including imagery on the date of data collection, (b) including data within 3 days before and after, (c) including data within 10 days before and after. All matrices include recall and precision for each crop phenology stage, and F1-score and overall accuracy (OA). 62

4.1 a) Cambodia (purple) and its neighboring countries. b) zoomed-in view of Cambodia and the Battambang province, shaded in a dark green color. White dotted box shows the area of study, presented in section c. c) Detailed view of the study area with the NDVI map for March 28th 2020. The NDVI map highlights the vegetation density in the area, with forest in the East and bare soil covering the rest of the image. The map is an example of the high spatial resolution captured with the use of Planet Fusion images. 78

4.2 a) shows an image over the selected territory, with both burned (darker values) and bare soil (brighter values) areas. b) corresponds to the classification after doodling over the left image. Magenta maps over the burned areas while dark cyan maps over the bare soil. 79

4.3 change detection over the area of study, after subtracting permanent forest. The darker the image less or no change was detected, brighter pixels indicate stronger changes. This image corresponds to the second pair February 13th and February 17th, 2020. Dotted squares highlight the areas selected for further analysis. 80

4.4 a) and b) panels showing an example image from the test batch and the corresponding prediction in purple and yellow, corresponding to two classes, burned and bare. c) shows the intersection over the union with the outline from the predicted mask. d) image shows the confusion matrix with the accuracy values for each class. Elements along the main diagonal correspond to the correctly classified ones, while those outside the diagonal are misclassified. 81

5.1 Current limitations and future perspectives for soil organic carbon (SOC) assessments from soil sample collection and analysis to remote sensing. . . . 96

A.1 Example of geographical distribution of the phenology samples for 2016. Panel (a) presents both the training and test datasets. Panel (b) white points present the distribution for the test dataset. Panel (c) blue points present the distribution for the training dataset. 138

A.2 Feature importance analysis for all the variables present in the dataset. The variables are described in Appendix A, Table 1. 139

List of Tables

2.1 Phenology classes (IV1-H) after re-grouping the phenological stages (VE - R6). IV1 = initial vegetative 1, IV2 = initial vegetative 2, EV = early vegetative, LV = late vegetative, ER = early reproductive, LR1 = late reproductive 1, LR2 = late reproductive 2, LR3 = late reproductive 3, H = harvest. 36

2.2 Classification metrics when 2014, 2015, 2017, and 2018 years were used to train the model, and 2013 and 2016 as validation dataset. A total of 14431 elements (corresponding to the phenological measurements for the years 2013 and 2016) were used for validation and a total of 26088 elements (corresponding to the p phenological measurements for the years 2013, 2015, 2017 and 2018) were used for training. 36

2.3 Classification metrics when testing for spatial transferability using all the years. 37

4.1 Classes and corresponding description. 82

A.1 Summary of variables used for random forest classification models. All variables were extracted from GEE archives with the exception of DOY, latitude, longitude, and growth classes, which were extracted from the ground truth dataset provided by Crop Quest Inc. 140

A.2 Summary of metrics presenting the results of overall accuracy, OOB, kappa value, precision, recall, F-score, and supports for each class (IV1, IV2, EV, LV, ER, LR1, LR2, LR3, H) from 2013 to 2018. Acronyms defined in Table 1 description 141

B.1 Before and after dates, and the distance between both 145

Acknowledgments

I would like to express my heartfelt gratitude to my husband Hernán. He has not only supported me throughout my journey, but also made my dreams his own. For six long years, he held my hand and reassured me that everything would work out. Without his constant support, I would not have made it this far. To my family, my dad Horacio, my mom Claudia, and my two amazing brothers Agustin and Ramiro. They have always been there for me, providing unconditional support through thick and thin, as well as to my grandparents Carlos and Susana. To my old friends Flavia and Renata, as well as the countless new ones I have made here at Kansas State University, such as Rachel, Cote, Debo, Coni, Nacho, and so many more. Their unwavering support and encouragement have been instrumental in my success.

I would like to express my profound appreciation to Dr. Ignacio Ciampitti for his absolute trust and for his dedication to open all doors and provide me with outstanding opportunities throughout my academic career. In addition, I would like to express my gratitude to Drs. Vara Prasad, Rasmus Houborg, and Bradley Olson, who served as committee members. Their extensive knowledge and insightful comments have been instrumental in moulding my research and assisting me in achieving my academic objectives.

This dissertation was made possible by the generous funding provided by the Kansas State University Department of Agronomy, the Kansas Corn Commission, and the USAID-SIIL.

Dedication

To my beloved husband, Hernán. His guidance, patience, support, and constant love have been the main source of strength and inspiration for me.

To my family my dad -Horacio-, my mom -Claudia-, and my two wonderful brothers -Agustin and Ramiro- for always being there for me.

To my grandparents who never missed a key event.

To all the people who shaped me into who I am today.

Chapter 1

Remote sensing and agronomy

1.1 Fundamentals of remote sensing

The sun is the source that fuels processes that take place on earth. It does this by supplying the light and heat that are essential to maintaining life. The energy from the sun comes in the form of electromagnetic radiation, where different wavelengths and frequencies (electromagnetic spectrum) will interact with matter (Teillet, 2015). This energy will be absorbed, reflected, or transmitted, depending on the properties of the material (Fig. 1.1). If the object is a photosynthetic organ, the proportion absorbed will play a crucial role in photosynthesis. Chlorophyll is the primary pigment responsible for absorbing electromagnetic radiation in the blue and red wavelength regions, which then drives the photosynthesis process. On the other hand, transmittance occurs when the radiation passes through the plant without being absorbed or reflected, with the proportion depending on several factors, such as plant species, leaves angle, season, among some of the most relevant factors (Woolley, 1971). The reflected proportion of electromagnetic radiation depends on the complex interactions of the

light with the plant's surface and the pigments absorbed in different wavelengths (Knipling, 1970; Woolley, 1971).

Taking advantage of these fundamentals, remote sensing is the art and science of collecting the reflected proportion, from afar (Teillet, 2015). If a sensor is capturing the energy available provided by an external source (i.e., the sun) we call it passive sensor (Teillet, 2015), when the opposite occurs and the sensor is the one emitting the energy and capturing the proportion after the interaction, we call it active sensor (Teillet, 2015) (Fig. 1.2). The main difference is that active sensors (with a source of energy) can help to collect information during the day or night, and usually can examine wavelengths not sufficiently provided by the sun. However, the requirement of carrying a source of energy to adequately impact a target could present a disadvantage in comparison with their passive counterpart. In general, the passive sensors used for Earth Observation collect data on the visible, near infrared, thermal, and microwaves, while the active sensors operate in the microwave band. Although these sensors can be further classified depending on how far they are from the target, for the scope of this dissertation we will center only on passive sensors on board satellites, orbiting the Earth. Lastly, we can additionally define sensors as multispectral or hyperspectral based on their ability to capture discrete segments of the electromagnetic spectrum (bands) or as a continuum. Access to these fractions would enable us to observe and comprehend what is unfolding on Earth, estimate vegetation cover, health, floods, among other features.

1.2 Understanding satellite images

Satellites for Earth observation (EO) play a critical role in providing information about the environment, climate, and natural resources (Dubovik et al., 2021). To understand how this data can be used, first we need to describe the four main characteristics (resolutions), that will define these sensors. These are spatial, temporal, spectral, and radiometric resolution (Fig. 1.3).

The spatial resolution refers to the smallest area on the Earth's surface that can be captured by the sensor. Spatial resolution is determined by the size of the individual pixel in the image, with smaller pixel sizes translating in higher spatial resolution, and subsequent more details can be observed in the image. A wide array of spatial resolutions is available nowadays, with sensors able to capture at sub-meter level all the way up to half a kilometer (Teillet, 2015).

The temporal resolution is another important characteristic of these sensors. We can define temporal resolution as the time that it takes the satellite to complete an orbit and capture the data over the same point on Earth. Hence, the temporal resolution is determined by the satellite's orbit, and swath width. In this sense, we can have satellites that cover the same area every day, while others capture the same territory every 16 days, for example (Teillet, 2015).

The spectral resolution refers to the range of wavelengths that a sensor can detect, typically measured in nanometers, where the narrower the range the greater is the resolution. As previously mentioned, multispectral sensors capture discrete bands, usually between 4 to 36

bands, while hyperspectral sensors can divide the same spectrum into hundreds of narrower bands, creating something close to a continuum (Teillet, 2015).

Finally, the radiometric resolution refers to the ability of the sensor to detect small differences in the intensity of the electromagnetic radiation. It is measured in bits and determines the number of discrete levels of brightness that can be detected by the sensor. Higher radiometric resolution results in more accurate and detailed images but requires more storage and processing capacity (Teillet, 2015).

1.2.1 Earth Observation satellites and fusion products

According to the database provided by the Union of Concerned Scientists (n.d.), more than 1000 EO satellites, between commercial and governmental, are currently active and orbiting the Earth. More than half of those are equipped with optical sensors (multispectral and hyperspectral). Although providing an exhaustive list is beyond the scope of this dissertation, below are some examples of the most extended for agricultural purposes.

MODIS (Moderate Resolution Imaging Spectroradiometer) is an instrument on board the Terra and Aqua spacecraft, that provides information about Earth's surface in almost real-time. It has 36 spectral bands with different spatial resolution, where most of the bands (8-36) have 1000 meters, while others range between 250 meters (bands 1-2) and 500 meters (bands 3-7). This allows for regional analysis, such as land cover, vegetation indices, leaf area index, gross primary production, and net primary production (Robinson et al., 2018), etc.

Landsat, the Landsat program dates back to 1972 when Landsat 1 was launched, and spans until today with the newly operational Landsat 9 and the Landsat Next mission planned for the coming years (NASA, n.d.). The first 3 missions, associated to the Landsat program, collected information on 4 bands (green, red, and two infrared), while missions 4 and 5 had 7 bands (6 in the visual, near infrared and short ware infrared, and one in the thermal). In 1991 Landsat 7 (still operational) was successfully launched and in addition to the 7 bands already collected by its predecessors this sensor was also equipped with a panchromatic band (15 meters). In 2013 Landsat 8 observing in 11 bands, producing 12-bit data, at 30, 15 and 100 meters became available. Landsat 9, available from 2022, collects data at 14-bit improving the quality of the images and securing the continuity of the program. The future looks even brighter with the Landsat next trio of smaller satellites, proposed for late-2030. This mission will collect data in 26 bands, with finer spatial resolution (10 to 20 meters), and a more frequent revisit time (compared with the other missions' 16 days). The Landsat database is one of the most impactful programs (Wulder et al., 2022), that helped in the advancement of science over the past 50 years, and even more importantly since 2009 when it became freely available for the public (Wulder et al., 2012). More than 700.000 publications are listed using Landsat images, and of those 18 % belong to agricultural studies (Wulder et al., 2022). Among the many uses related with agriculture (Leslie et al., 2017) we can highlight its use for mapping crops (S. Wang et al., 2020), evaluating land tenure and field size (E. V. White & Roy, 2015), forecasting yields (dos Santos Luciano et al., 2021), crop phenology (Nieto et al., 2021), among many others.

Another well-extended mission serving EO is the Sentinel program, particularly for the scope of this dissertation, Sentinel-2A and 2B. Equipped with a multispectral sensor capable of retrieving 13 bands at 10 to 60 meters, with a cadence of 5 days (Drusch et al., 2012), its use was rapidly extended thanks in part to the no cost availability. The application of these images for agricultural purposes is similar to the ones outlined previously for Landsat; however, is its finer resolution, the feature that standout this product for applications that requires for example a better temporal revisit and, or higher spatial resolution (Misra et al., 2020; Nieto et al., 2022; Segarra et al., 2020).

In a distinct category, we can locate satellites with very high resolution (less than 10 meters), primarily utilized for commercial purposes. Some examples of such satellites, and constellations, include SkySat, PlanetScope, Pléiades Neo, SPOT, and ÑuSat, among others (European Space Agency, n.d.). These satellites are distinguished by their ability to capture data ranging from 10 meters to sub-meter, along with a high revisit time and tasking capabilities. Although their spectral resolution is typically restricted to the visual and near infrared, and occasionally short-wave infrared, the applications for these satellites are extensive.

To take advantage of the differences between different acquisition patterns, in recent years different researchers, both in academia and private sector, have developed fusion products. This is the combination of two or more sensors to provide a more complete product, improving temporal, spectral and spatial resolution. A good example for this is the Planet Fusion data, which integrates PlanetScope, MODIS/VIIRS, Landsat 8 and Sentinel-2 into an analysis ready, harmonized, free of clouds and shadows, with daily cadence, 3 meters resolution with

4 spectral bands (blue, green, read, near infrared) product (Planet, 2022). This product has been proven useful for a variety of applications, from rice flooding date (Brinkhoff et al., 2022), monitoring small water bodies (Perin et al., 2021), and to crop phenology (Nieto et al., 2022).

In a similar note the Harmonized Landsat Sentine-2 (HLS) product combines Sentinel-2 images with Landsat 8 and 9 to obtain a 30-meter product with higher temporal resolution (2-3 days), improving land monitoring capabilities (Claverie et al., 2018).

1.3 Spectral indices

Spectral indices derive information about the targeted object from the specific combinations of bands. One of the most well know vegetation indices is the normalized difference vegetation index (NDVI) developed by (Rouse Jr et al., 1974), combining the reflectance in the red and near infrared (NIR) bands. This index allows the study photosynthetically active vegetation to derive conclusions about its status and patterns. The NDVI is widely used to monitor crop growth and health and assess changes in land cover and vegetation density over time. Other vegetation indices include the Enhanced Vegetation Index (EVI) (A. Huete et al., 2002), computed to reduce the impact of atmospheric conditions and soil reflectance on vegetation signal, and the Soil-Adjusted Vegetation Index (SAVI), which is similar to NDVI but is adjusted for varying levels of vegetation cover and soil background reflectance. In a similar fashion, other vegetation indices to retrieve information about bare soil (C. T. Nguyen et al., 2021), water content, both free water and leaf content (Ceccato et al., 2002), burned

areas (Parks et al., 2014), and even minerals (Ninomiya, 2002) have been developed over the years. Hundreds of indices are available across the literature (Montero et al., 2022), but this dissertation will cover only a tiny fraction of those, based on the available bands in the tested products.

1.4 A glimpse into the future

Throughout the previous half-century, there has been extraordinary development in satellite technology (Wulder et al., 2022). Not only has the number of operational satellites expanded, but so too has the quality of equipment, the scale of satellites, the democratization of data access, and most importantly, the expansion in the frontiers of science (Dubovik et al., 2021). The advancement of spectral technologies including multispectral, hyperspectral, optical, and radar has made a substantial contribution in this regard. The next decade will see a significant increase in the number of launched satellite missions all around the world, with several missions collecting data on the hyperspectral range, and active sensing (Dubovik et al., 2021). In addition, the availability of vast computational resources and archives, such as Google Earth Engine (Gorelick et al., 2017), in combination with the growing use of artificial intelligence (Yang et al., 2022), and the sustained reduction in the cost of commercial imagery, gives an immense potential for achievements previously unimaginable.

Although many studies exploring pipelines to integrate hyperspectral and multispectral, as well as optical and radar are available, the coming years very likely will see these convert into products easily deployed to the study of agriculture, helping us to predict in season

quality traits such as protein content, oil, etc., allowing not only to set commodity prices based on yields but also segmenting the market based on the quality of the agricultural product. In a similar way, studies on leaf area index (Kimm et al., 2020), soil moisture (Gruber et al., 2020), heat stress (Gerhards et al., 2019) will become commercially available products in the near future, such as the case of Planetary variables from Planet (Planet, 2020).

1.5 Dissertation objectives

Chapter 1 describes the primary research setting, the technology needed for understanding the topics, and the relevance of this dissertation within the context of agriculture. This chapter gives the reader a clear knowledge of the research's aim and scope and the primary tools that will be used throughout the dissertation.

Chapter 2 presents the potential of Landsat 8 images, tested to address maize phenology classes in southwest Kansas. The purpose of this chapter was to examine the performance of Landsat 8 imagery data to effectively monitor growth and development of maize (*Zea mays* L.) crop in Kansas, United States (US). Hence, the study provides useful insights on the use of remote sensing for agricultural system monitoring.

Chapter 3 investigates the performance of high temporal and spatial resolution images (Sentinel-2 and Planet Fusion) to determine each crop phenological stage for maize crop at on-farm scale. This chapter also aims to determine the optimal combination of spectral and meteorological variables for correctly tracking maize development and growth. The main

findings of this chapter provide critical information for enhancing the use of remote sensing for crop monitoring.

Chapter 4 focuses on evaluating the use Planet Fusion data to classify burned and bare soil areas in Cambodia. The study investigates changes in agricultural patterns and management techniques in this region and provides better assessments for conservation agriculture. This chapter sheds light on the potential of remote sensing to monitor agricultural systems in developing countries, where traditional field-based monitoring methods may be restrictive, labor- and cost-intensive. Chapter 5 discuss the current state of the art and future changes needed to combine soil science methods and remote sensing for determining soil organic carbon. This chapter examines the problems and opportunities connected with employing remote sensing to monitor soil properties, particularly soil organic carbon and provides viable solutions for linking these two domains.

Lastly, Chapter 6 provides a summary of key findings and a discussion on future steps to continue strengthening earth observation while preserving sustainable goals and guaranteeing food security.

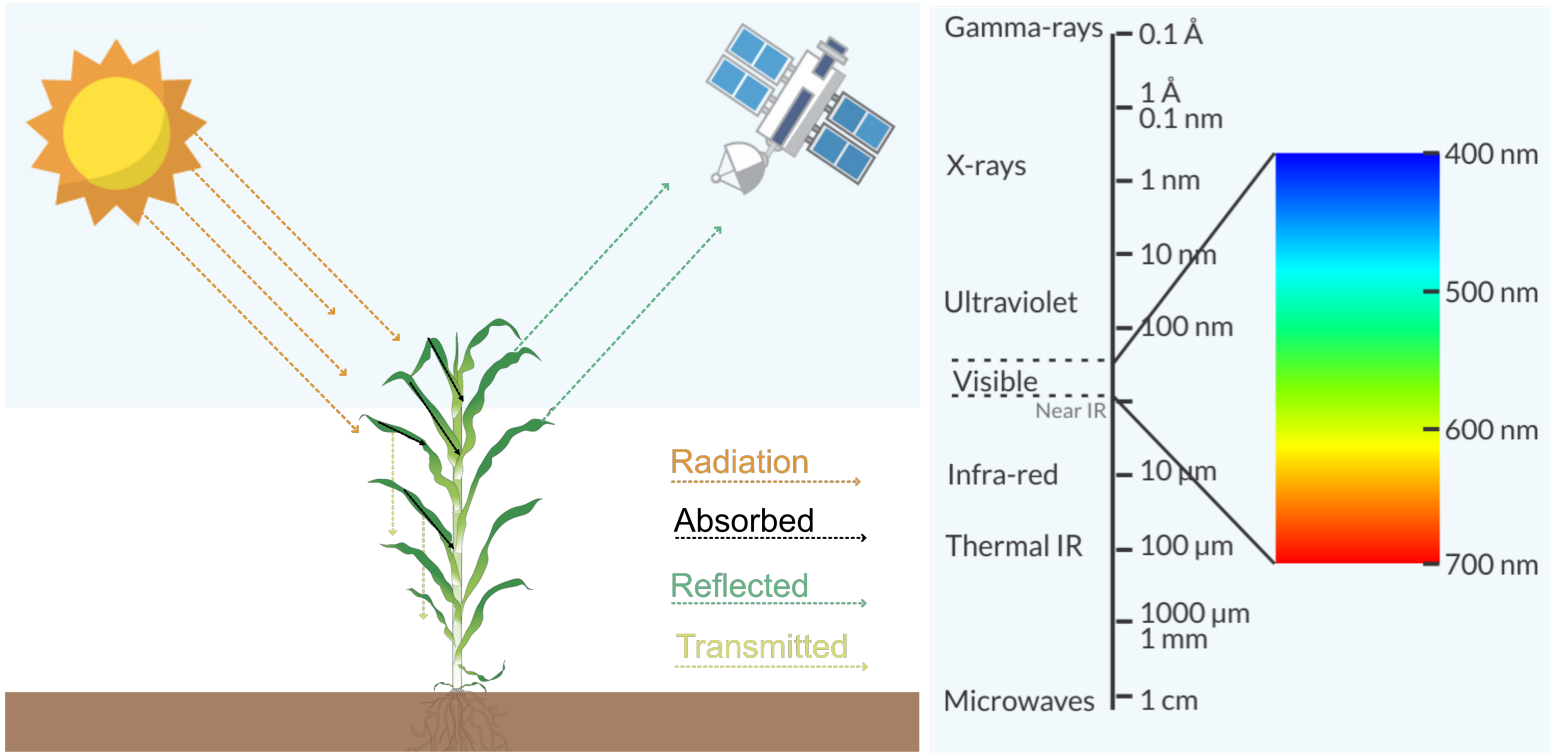

Figure 1.1: Interaction of the incident radiation with the target object, plant, showing the proportion absorbed, transmitted, and reflected. Right panel show the electromagnetic spectrum with the different proportions of interest.

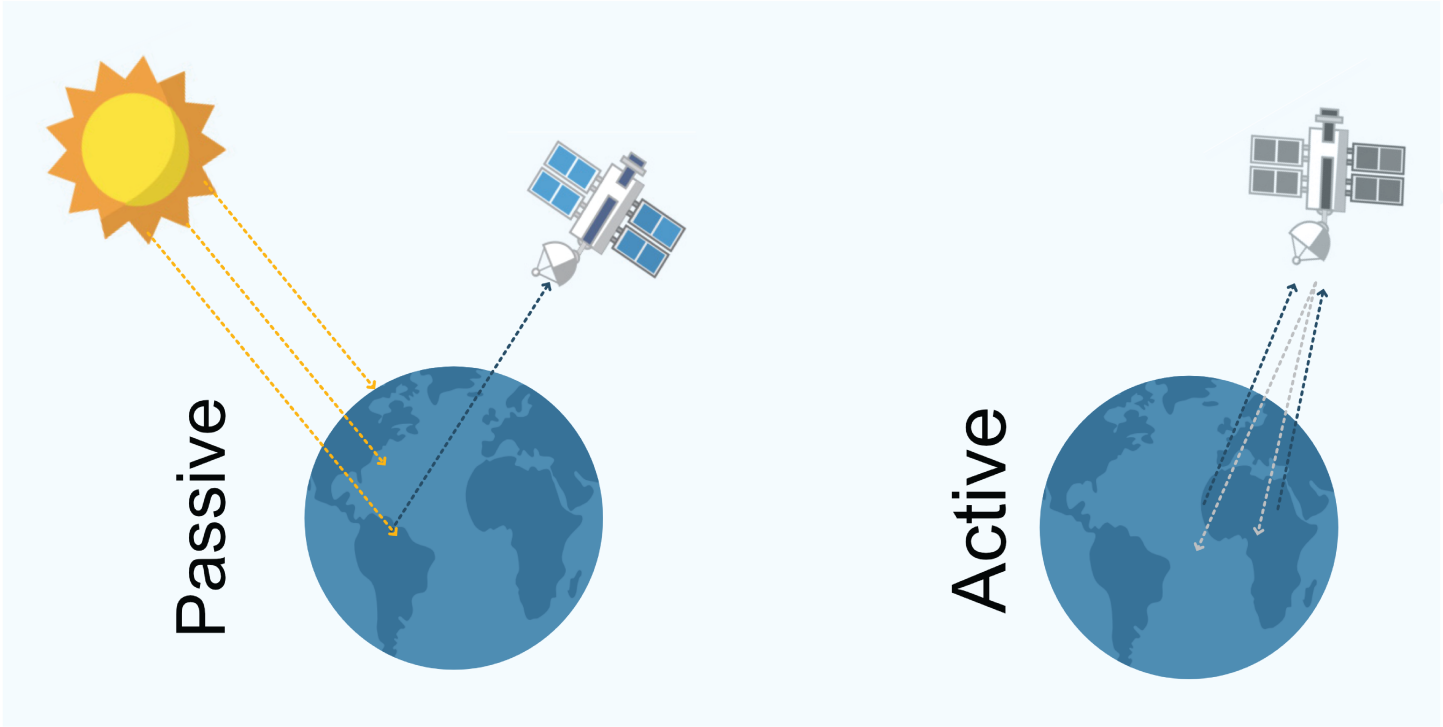

Figure 1.2: representing the different types of sensors based on the source of energy. Left side passive sensor where the energy comes from the sun; right side shows the counterpart, where the energy is emitted by the sensors.

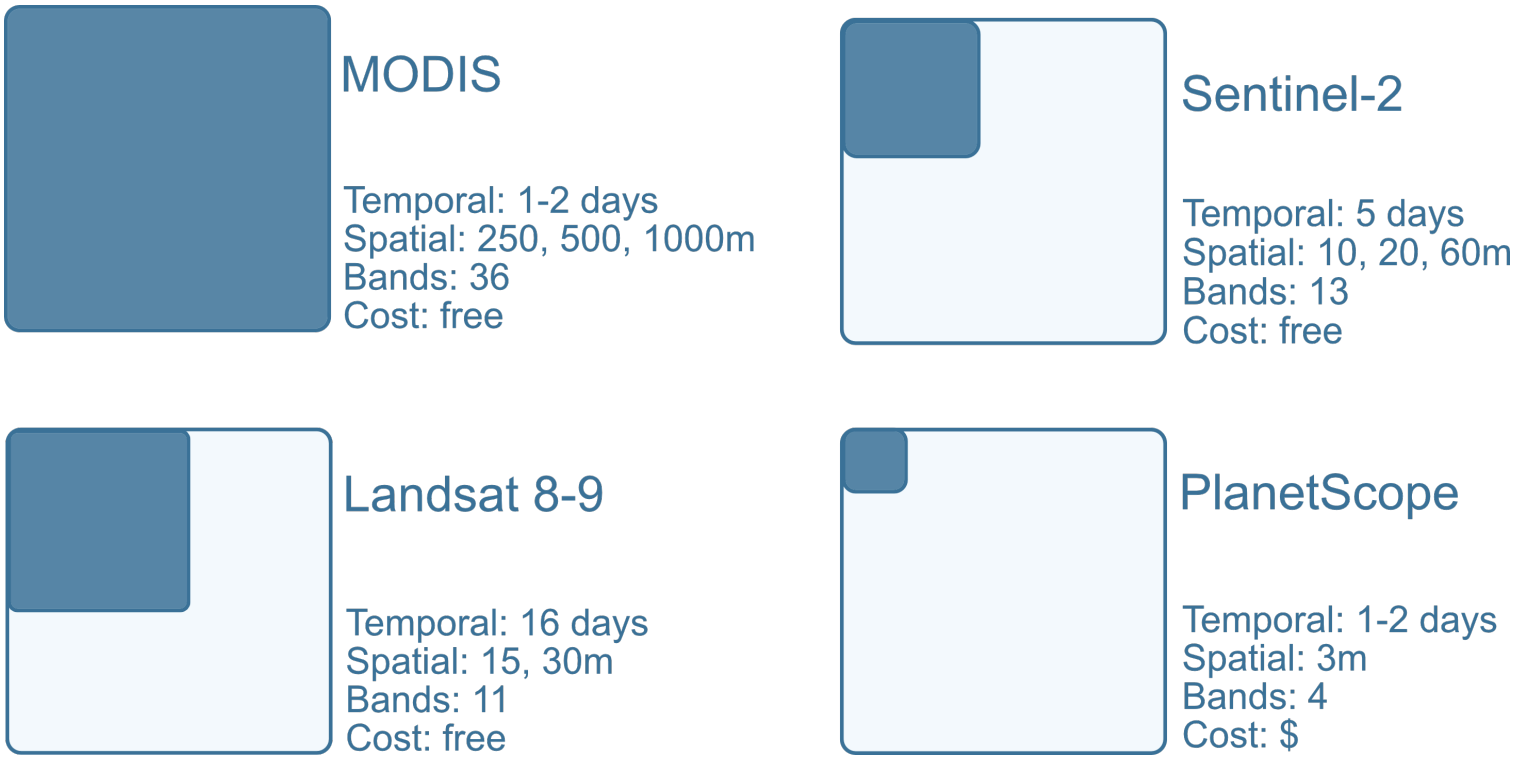

Figure 1.3: Example with the most extended satellites (and constellations) used for agricultural purposes. Temporal and Spatial refers to the resolutions and bands was used as a short form for the spectral resolution provided by each satellite. This list is only for explanatory purposes, it is by no means an extensive one.

Chapter 2

An integrated approach of field, weather, and satellite data for monitoring maize phenology

*Published in Scientific Reports

Nieto, L., Schwalbert, R., Prasad, P.V.V. et al. An integrated approach of field, weather, and satellite data for monitoring maize phenology. Sci Rep 11, 15711 (2021).

https://doi.org/10.1038/s41598-021-95253-7

Abstract

Efficient, more accurate reporting of maize (*Zea mays* L.) phenology, crop condition, and progress is crucial for agronomists and policy makers. Integration of satellite imagery with

machine learning models has shown great potential to improve crop classification and facilitate in-season phenological reports. However, crop phenology classification precision must be substantially improved to transform data into actionable management decisions for farmers and agronomists. An integrated approach utilizing ground truth field data for maize crop phenology (2013–2018 seasons), satellite imagery (Landsat 8), and weather data was explored with the following objectives: (i) model training and validation—identify the best combination of spectral bands, vegetation indices (VIs), weather parameters, geolocation, and ground truth data, resulting in a model with the highest accuracy across years at each season segment (step one) and (ii) model testing—post-selection model performance evaluation for each phenology class with unseen data (hold-out cross-validation) (step two). The best model performance for classifying maize phenology was documented when VIs (NDVI, EVI, GCVI, NDWI, GVMI) and vapor pressure deficit (VPD) were used as input variables. This study supports the integration of field ground truth, satellite imagery, and weather data to classify maize crop phenology, thereby facilitating foundational decision making and agricultural interventions for the different members of the agricultural chain.

2.1 Introduction

Maize (*Zea mays* L.) is one of the leading grain crops, with 193 M ha harvested globally and more than 33 M ha in the United States (US) during the 2018 growing season. US maize production is largely concentrated in the central region, accounting for roughly 85 % of total US production (USDA-NASS, 2019) The US Department of Agriculture (USDA) via

the National Agricultural Statistics Service (NASS) releases a weekly report for in-season crop progress, termed Crop Progress and Report Conditions (CPRC), which provides a subjective estimate of crop phenology and condition in major US producing states (USDA-NASS, 2019). Estimates from the CPRC are based on survey data gathered on a weekly basis from an extensive network of regional agricultural agents based on their field observations (Zhong, Hu, et al., 2016).

Although the data is informative, the collection process is labor intensive, time consuming, financially ineffective, and subject to bias (Carletto et al., 2015; Gourlay et al., 2017). With challenges eminent due to funding restrictions and travel limitations, new approaches should be pursued, tested, and quickly implemented. An extensive record of scientific literature portrays the relevance of implementing well-timed management practices to improve yields and input use efficiency, such as timely irrigation, fertilization, crop protection, and harvest (Cai, Guan, Nafziger, et al., 2019; Morris et al., 2018; Ruml & Vulić, 2005). Remotely sensed satellite data presents a significant opportunity to improve timely agricultural interventions and monitoring of crop vegetation. With various spectral, radiometric, temporal and spatial resolutions, satellites function as a critical source of data to aid tracking of field crop phenology progress (Bolton & Friedl, 2013; Duncan, n.d.; Sakamoto et al., 2005). Combining satellite data features permit the generation of different vegetation indices (VIs), such as the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979), Enhanced Vegetation Index (EVI) (A. Huete et al., 2002), Green Chlorophyll Vegetation Index (GCVI) (Gitelson et al., 2003), Global Vegetation Moisture Index (GVMI) (Ceccato et al., 2002) and

Normalized Difference Water Index (NDWI) (B.-C. Gao, 1995), among many others. These VIs describe changes in vegetation dynamics and are correlated with plant traits such as leaf area index, leaf chlorophyll concentration, and canopy water content (Ceccato et al., 2002; B.-C. Gao, 1995; Gitelson et al., 2003; A. Huete et al., 1994; Nguy-Robertson et al., 2012; Schlemmer et al., 2013). Integrating remote sensing with ground truth and weather data has great potential to advance science and improve overall prediction of crop yield, in-season progress, and crop phenology (Cai, Guan, Lobell, et al., 2019; Sakamoto et al., 2014; Zhong, Hu, et al., 2016).

The development of programs such as Google Earth Engine (Gorelick et al., 2017) facilitates dataset manipulation and analysis by integrating all assets in one place. Directly working in the cloud without downloading large sets of data and using a parallel processing approach allows computation across a large number of machines (Kennedy et al., 2018). This new tool accelerates analysis of remote sensing projects over large areas and facilitates real-time crop behavior and progress exploration (Gorelick et al., 2017). Of the many tools used to perform classifications, the random forest (RF) algorithm (Breiman, 2001) presents steady performance with large (Zhong, Yu, et al., 2016) and unbalanced datasets (Pal, 2005) in addition to modeling non-linearity correlations among the feature space and the dependent variable.

Recent studies focused on predicting field crop phenology, particularly in maize and other field crops (Vina et al., 2004; Zhong, Hu, et al., 2016), concentrated on providing estimates for dormancy, green-up, mid-season growth and development, day of senescence, and end of

the season. Although these metrics are convenient from a remote sensing analytical viewpoint, more detailed phenology descriptions (Ghamghami et al., 2020; Liao et al., 2019) are typically desirable to provide timely actionable agricultural decisions and interventions. Due to field data availability, this study is focused on the Southwest (SW) Agricultural District in Kansas, US. Ground truth data was provided by an industry partner, comprising a large dataset on crop phenology for maize fields during the 2013-2018 growing seasons. Therefore, the aim of this research study was to evaluate a classification of satellite-derived maize crop phenology and integrate in-season weather information to develop a classification model benchmarked with field survey data. To achieve this overarching goal, we established the following objectives: i) model training and validation – understand how different variables affect the model performance and identify the best combination of spectral features, weather parameters, geolocation and ground truth data, resulting in a model with the highest accuracy across years at each season segment (step one); and ii) model testing – post-selection model performance evaluation for a) each phenology class with unseen data (hold-out cross-validation); b) temporal transferability; c) spatial transferability (step two).

2.2 Materials and Methods

2.2.1 Study area and environmental conditions

This research was conducted in the SW Agricultural District, comprising a total of 10 counties, Kansas, US. Annual normal precipitation in this area ranges from 381 to 635 mm from west to east, with the southwest corner (bordering the states of Oklahoma and Colorado)

presenting the lowest precipitation (Fig. 2.1). Average mean daily temperature fluctuates from 12 to 14 °C, with the average minimum ranging from 7.5 to 15 °C and the average maximum between 19 and 22 °C (Goodin et al., 1995). The maize growing season spans from late April (sowing) to late October (harvest), with yields usually enhanced by groundwater irrigation from the Ogallala Aquifer (Masialeti et al., 2010), which currently has more than half million of hectares under irrigation (Lanning-Rush, 2016).

2.2.2 Data gathering and feature engineering

Surface reflectance data

Surface reflectance data was accessed via Google Earth Engine (GEE). Landsat 8, 30 meters spatial resolution, Surface Reflectance (SR) collection, tier 1, was retrieved from the GEE data repository.

Images from 2013 to 2018 from May to September (season segment, May = 1, June = 2, July = 3, August = 4, September = 5) each year were collected for the path/row 30/34. Starting the season in May and extending it until September were supported by the Start of Season Time (SOST) and End of Season Time (EOST) analyses of the NDVI each year (Hao et al., 2015). Images with less than 30% land cloud cover were selected, and remaining clouds and shadows were masked by taking advantage of the pixel quality band in the SR product. Five indices were calculated (NDVI, EVI, GCVI, NDWI and GVMI) over each of these images, and the mean value per index at each field was stored by season segment.

Meteorological data

Gridded surface meteorological data from GEE archives was used to extract weather variables (Abatzoglou, 2013) This dataset integrates high-resolution spatial data (2.5 arc minute, 4km) from PRISM (Daly et al., 2008) with daily temporal resolution data from the North American Land Data Assimilation System (NLDAS). The main weather variables gathered were precipitation, mean minimum and maximum temperature, and vapor pressure deficit (VPD). The Growing Degree Units (GDU) were calculated using the following equation: GDU = [(Max. temperature (C) – Min. temperature (C))/2] – 10 °C (Base temperature) [35].

Field data collection and Feature engineering

The initial dataset provided by Crop Quest Inc. included more than 70.000 fields distributed across the states of Kansas, Oklahoma, Colorado, and Nebraska (USA). The first step was to select only the maize data points within the SW District in Kansas, presenting more than 20.000 fields (Fig. 2.1). This dataset contained geolocated maize phenology measurements (latitude and longitude), associated with the day of the year (DOY), for crop data collected from the 2013 to 2018 growing seasons. To not only avoid potential classifier issues due to the specificity level of the crop growth stages, but also provide informative data more readily translated into actionable decisions for potential agricultural interventions, the stages were re-grouped into nine classes (IV1, IV2, EV, LV, ER, LR1, LR2, LR3 and H), as presented in Table. 2.1.

Remaining data points were geographically aligned and merged with polygon shapefiles

containing the field boundaries. This new file was uploaded as an asset into GEE and used to summarize the satellite and weather information within each field. Additionally, as an approximation, a filter was applied to separate the center-pivot irrigation fields from rainfed agricultural farms. This filter was built using the Canny Edge Detector (Canny, 1986). Although this technique returned nearly all fields under irrigation, some omission errors were present due to the different types of irrigation systems; only circular pivots of 500 m radius were extracted.

The final dataset included 21 independent variables and Growth Class as the dependent variable used to train and validate the models (full description in Table. A.1).

2.2.3 Model training and validation, random forest classification model

Random forest (RF) has been extensively used in the remote sensing field to solve classification problems. This method has been proven to be easy to train with a little or non-computational cost while also dealing with complex interactions and highly correlated variables (Belgiu & Drăguţ, 2016; L. Zhang et al., 2020). Furthermore, RF is extensively documented in the scientific literature to outperform other methods. In the research conducted by (Fernández-Delgado et al., 2014), testing 178 classifiers, RF resulted in the classifier with the best performance. Analysis conducted by (Belgiu & Drăguţ, 2016; Gómez et al., 2016; Talukdar et al., 2020) among others, presented the same outcome.

Following the abovementioned rationale, all models were trained using a RF classification

algorithm from scikit learn software (Pedregosa et al., 2011) in a Python 3.8.0 environment (van Rossum, 2007) This algorithm was adjusted by fine-tuning hyperparameters, executing a grid search to find the best data combinations, once the best features were obtained. The main hyperparameters explored in our analysis were the following: bootstrap, maximum depth of the tree, maximum number of features for the best split, minimum samples per leaf, minimum samples to perform the split, and the number of estimators (Breiman, 2001). The best model considered the following: bootstrap = true, maximum depth of the tree = 50, maximum number of features for the best split = 9, minimum samples per leaf = 2, minimum samples to perform the split = 4, and the number of estimators = 50.

To evaluate the classifier, datasets were split into training, validation, and test sections. The proportion retained for training and validation was 80% of the total (of this 80%, 80% were used as training and 20% as validation), and 20% was used as a test for model evaluation, the numbers for each class each year can be find in Table A.2. An example of the geographical distribution for field observations is presented as supplementary material (Fig. A.1) The selected proportions resulted from a sensitivity analysis using the entire dataset spanning from 2013 to 2018 period with all variables (later referred to as the Full model). In this analysis, the overall accuracy was used as a metric to measure model performance, with data proportions adjusted at increments of 10% (from 10 to 90%, to find the optimal combination for training and validation (80/20%).

2.2.4 Model testing

Performance measurements

For this study, model classification performance was evaluated using five different metrics calculated using the test portion of the dataset. These metrics were as follows: precision, the number of true positives divided by the sum of true and false positives; recall, the number of true positives divided by the sum of true positives and false negatives; overall accuracy metric, the sum of all correctly classified elements divided by total elements; kappa coefficient of agreement, a measure describing how well the model is performing compared to the randomness (taking into account the possibility of agreement occurring by chance); and finally, F-score, the harmonic mean of the precision and recall (Congalton, 1991) The harmonic mean is biased towards observations with lower values. When all values have the same magnitude, the harmonic and arithmetic means are equivalent. The F-score has a maximum value of 1, representing perfect precision. Therefore, the F-score is maximized by simultaneously maximizing precision and recall.

The out of bag score (OOB) was computed in addition to evaluating the performance of the model. This measure computes the average misclassification ratio of non-training samples. The importance of this metric relies on offering an unbiased assessment of classification performance, with values closer to 1 indicating good performance (Belgiu & Drăguţ, 2016).

Effects of feature selection on model performance

The openly exploration of different combinations of variables to address model performance was motivated by two main reasons. First to understand how different features affect the performance of the model (Cai et al., 2018) second to overcome any potential bias resulting from a feature importance analysis. This technique has been proven to favor variables presenting high cardinality (Strobl et al., 2008, 2007), condition present on the dataset for geographical features as well as growth DOY. Nonetheless a feature importance was calculated for one of the models to observe the behavior.

To test this abovementioned point, the following variable combinations were used for each year and for each season segment: all bands (B) including spectral bands 2, 3, 4, 5, 6, 7, 10, and 11 for Landsat 8; all bands and all weather parameters (BW), including the mentioned spectral bands, precipitation, maximum and minimum temperature, VPD, and GDU; all indices and all bands (VIB), including the mentioned spectral bands and NDVI, EVI, GCVI, GVMI, and NDWI as vegetation indices; all indices and weather parameters (VIW), including the five above mentioned weather parameters and five vegetation indices; only weather parameters (W), testing five weather parameters; only indices (VI), including only five vegetation indices; indices, weather parameters, and bands together (FULL), and finally a model only containing latitude, longitude and day of the year (LLD). The models most stable across years and within season (VIW; W; VI) were trained again analyzing all possible interactions between the variables presented. Thus, we explored 31 combinations for the W models, 31 combinations for the VI models, and 1023 combinations for the VIW mod-

els. The number of models actually fitted was greater than the previous sum (31+31+1023) due to the grid search approach for fine-tuning the hyperparameters. In order to account for spatio-temporal correlation in the response variable induced by other features that may not be captured by remote sensing (surface reflectance) and weather data, all model combinations included field geolocation and DOY (Moeyersoms & Martens, 2015).

Because not all crop phenology classes were present between season segments due to the characteristics of this variable, the accuracy values for each model during each season segment for each year were retrieved to better understand the performance within and between years.

Model performance under temporal and spatial transferability

Following the completion of the first stage, the subsequent analyses were established to better understand the performance and stability of the final model. To learn about the stability of the model across years, the dataset was divided into training and validation based on year separation. For this evaluation, we grouped the 6 years (2013-2018) according to the weather trend combining average years (2014, 2017) with dry (2013, 2018) and wet ones (2015, 2016). The same metrics (Overall accuracy, OOB, precision, recall and F1-score) were utilized to test this analysis.

To explore the ability of the model while leading with spatial transferability (Belgiu & Drăguţ, 2016), the dataset was divided into training and validation based on spatial attributes, with a total of 40.519 elements (sum of all the phenological measurements across the 6 years). This resulted on two datasets, a larger one comprised by the fields located in the western part of the area under study (training) where fields are more aggregated. The

validation dataset was smaller and composed by the fields in the east side of the geographic area.

2.3 Results

2.3.1 Best combination of features and model performance

All models explored in this study classified crop phenology classes with an accuracy above 70%, except for the model combining all 21 variables (accuracy below 60%)(Fig. 2.2). For this model, we also conducted a feature importance analysis (Fig. A.2). Although most of the variables shared a similar importance, some were more relevant, such as latitude, longitude and growth doy, associated to each phenology measurement (variables with high cardinality), while others such as precipitation were deleterious. For remaining models, accuracy ranged from 70 to 100% within a season and between years. The models with the best performance between years and within season were I) combination between vegetation indices and weather, II) weather III) vegetation indices, and IV) latitude, longitude and doy. For these four models, all possible combinations between variables were tested. This resulted in two models with the best performances, one composed by 5 vegetation indices and only one weather parameter, VPD, where the accuracy assessment ranged from 86% to 98%; and the other model composed only by latitude, longitude and doy, with an accuracy of 100% for almost all the season segments – years (Fig. 2.2). Further discussion on this model is presented in in the discussion, later on this manuscript.

2.3.2 Classification using VIVPD model

The combination of features optimizing model performance was the following: DOY, EVI, NDVI, NDWI, GCVI, GVMI, VPD, latitude, and longitude. A more detailed view of model performance, spanning all growing seasons from 2013 to 2018 period, is presented as a confusion matrix (Fig. 2.3). The overall accuracy for this classifier was 94%, (OOB 0.94), and the Kappa coefficient was 0.93. Further details for each year and class are presented in Table. A.2. Overall, model behavior was similar across classes and years except when the number of elements (supports) was small (10). The latter scenario can be visualized with class H for the 2018 year, where only 6 elements were present, and the metrics were considerably smaller compared to other classes with more elements.

2.3.3 Classification using VIVPD model across years

Table. 2.2 presents the results when the model was transferred across years. In this analysis the years 2014, 2015, 2017 and 2018 were part of the training dataset, and 2013 and 2016 (dry and wet years respectively) used as validation dataset. The overall accuracy was 93 percent, which was close to the hold-out cross-validation accuracy (Fig. 2.3). Furthermore, all of the parameters for each class were similar, if not better, with precision values ranging from 0.85 to 1, recall values ranging from 0.85 to 0.99, and f1 scores ranging from 0.86 to 0.97.

In addition to this split of years, other combinations were tested to observe the behavior when using only wet years for the validation (2015 and 2016) and not present in the training

set, and the opposite, allowing only dry years for the validation (2013 and 2018). In both scenarios the overall accuracy was 0.94 for the first approach and 0.95 for the second, and OOB score of 0.95 for both.

2.3.4 Classification using VIVPD model across space

The model tested for the spatial transferability of RF resulted on lower values of overall accuracy and low metrics for some classes, such as H, LR2, ER, intermediate for IV1, EV,LR1 and better for IV2,LV and LH3. In all the cases the values were lower when compared against the models previously mentioned (Table. 2.3).

2.4 Discussion

The proposed approach to characterize maize phenology presented the following main advantages: i) use of publicly available data as model input, ii) prediction of phenology using high-spatial resolution (30m), and iii) use of random forest for classification, allowing the trained model to be easily deployed in GEE for efficient computing processing. These aspects, associated with the stable model performance of phenology prediction when transferred across years, ensures high generalization power. Future uses can be easily adapted for using higher spatial resolution satellite data (e.g., Sentinel 2, 10 m), providing better opportunities for small-scale farming in developing countries. In exploring a wide feature space for model training and validation (5 weather variables, 5 VIs and 8 surface reflectance bands, latitude, longitude, and DOY for image collection), we found that the full model (considering all 21

predictors) did not perform well. The inferior results for the full model were potentially related to the high correlation among the surface reflectance bands. Predictive models suffer from multicollinearity issues when the independent variables are correlated, causing unpredictable variance in the model outputs (overfitting) (Dormann et al., 2013; Schwalbert et al., 2018). In the opposite scenario, a model based solely on geographical features and dates created a potentially spurious ideal classification. When dealing with classification problems, this situation is often avoided by dropping or transforming the features with high cardinality (Moeyersoms & Martens, 2015). Even yet, this aspect requires further consideration, particularly when it comes to classification studies integrating agronomical data.

Alternatively, we documented superior performance when VIs and VPD were combined. In addition to VPD (Basso & Ritchie, 2018; Fletcher et al., 2007; Messina et al., 2015), VIs NDVI, EVI, and GCVI have been reported as useful in predicting crop development during the growing season, especially for maize and soybean (Glycine max L.) (Masialeti et al., 2010; Nguy-Robertson et al., 2012; Zhong, Hu, et al., 2016). Furthermore, latitude, longitude, and DOY are strong indicators of how weather and solar radiation patterns influence crop development (Deines et al., 2019; Z. Jin, Azzari, & Lobell, 2017) and dropping them for the rest of the models could result on loosing relevant information.

From a testing standpoint, the model comprised of VI and VPD was stable across all phenology classes, resulting in high classification metrics for the vast majority of them, even in years when weather dynamics substantially impacted crop progress (F. Gao et al., 2020; M. A. White et al., 1997), and when tested explicitly for temporal transferability. This

latter becomes even more relevant if we consider that the classes correctly classified in all the scenarios correspond with the key developmental stages of the crop, such as emergence, effective density, and yield definition.

These findings align with what was reported by Cai et al. (2018); S. Wang et al. (2019). It is worth mentioning that at least one quality image (with less than 30% cloud cover) was retrieved each month between May and September in every year. Although images were available, data retrieved from certain fields was not sufficient to balance the dataset with regard to phenology classes due in part to the characteristics of the ground truth dataset. This resulted in good precision for classes with more observations but poor performance for the underrepresented classes with a smaller number of data points (Japkowicz & Stephen, 2002; Shan et al., 2006). The underperformance of the model transferability across space is also congruent with findings from Belgiu & Drăguţ (2016); Cai et al. (2018); S. Jin et al. (2018). According to Pelletier et al. (2016); S. Wang et al. (2019) this can be overcome when the territory remains similar as the training dataset, or they share similar weather parameters. Although this could be true for land classification problems, our findings suggest that phenology classification can be more susceptible to these changes.

The lack of quality satellite imagery data, the resulting decrease in the number of classified fields, and the changing environment with the consequent data shift could be remediated with at least three possible solutions not currently addressed by this study. First, thanks to the ability to obtain regular intervals of data for time series analysis, a fusion of satellites (Bhowan et al., 2009; Liao et al., 2019) would facilitate retrieval of higher-quality data, with

improved spatio-temporal resolution preventing data loss, and a better scenario for smoothing techniques such as Savitzky-Golay fitting, locally weighed regression, spline smoothing, and others. Second, because not all classes contained a balanced number of elements, different methods can be applied to managed unbalanced data (F. Gao et al., 2017; Kitchenham, 1998), and ultimately result in a more robust classifier. Lastly, integration of remote sensing data and crop model outputs (i.e., mechanistic- or process-based models) could aid in maximizing predictability power and spatio-temporal limits (Azzari et al., 2017; Clevers, 1997; Dal Pozzolo et al., 2015; Z. Jin, Azzari, Burke, et al., 2017; Lobell et al., 2015; Sibley et al., 2014).

2.5 Conclusion

Crop phenology monitoring is crucial for agricultural management since it enables growers, stakeholders, policymakers, and government agencies to determine when the most critical stages are occurring particularly during non-ideal conditions, due to biotic or abiotic stress, where the theory departs from the reality. Is particularly in those cases where models like the one described on this paper can facilitate the report of phenology, and to take more inform decisions. Improving crop phenology classification is becoming more prevalent as current and future research focuses on satellite data fusion and the use of mechanistic or process-based models to enhance spatio-temporal resolution. The findings presented on this paper showed strong classification metrics across years and proven the ability of spectral and weather features to assist in phenology classification. The proposed model also has

been tested for spatial transferability and although promising future work should be put on understand the mechanism behind the model behavior.

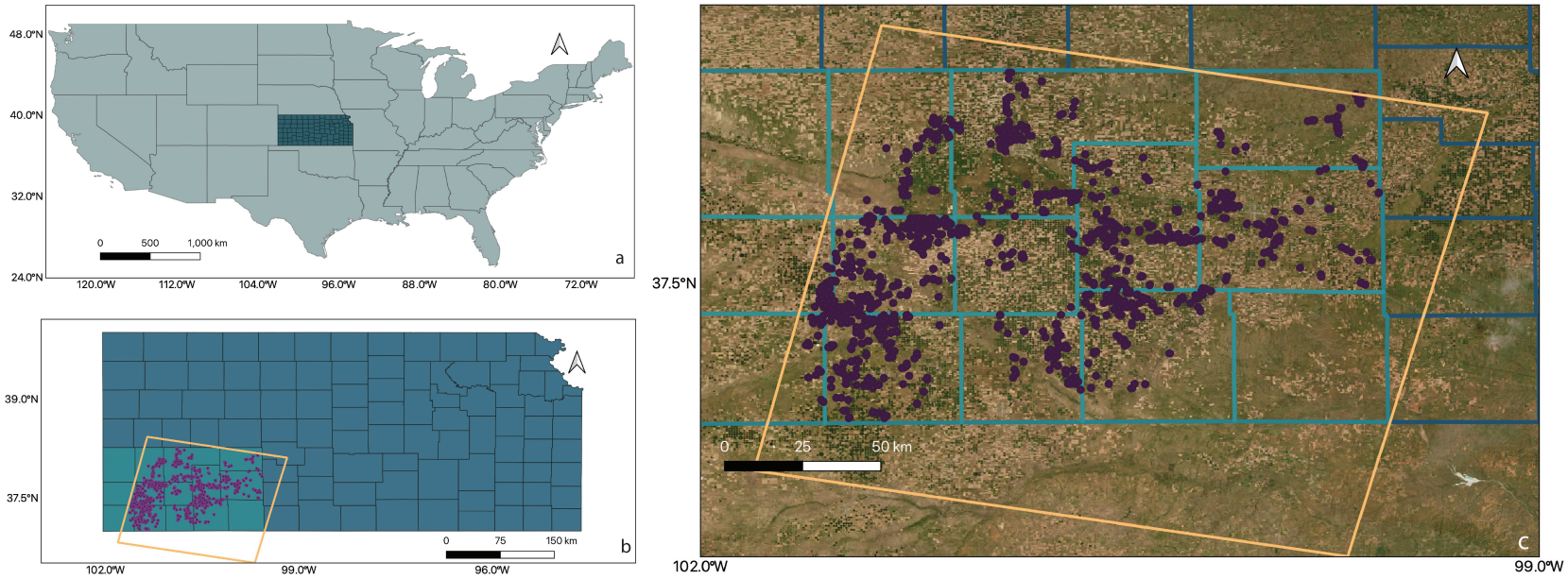

Figure 2.1: a) area of study highlighting, the USA, in green and Kansas in blue (b) Area of study. The purple points correspond to phenology measurements in Kansas's Southwest Agricultural District; the orange area corresponds to Landsat 8 tile 30/34, which covers the area under study. (c) Zoomed-in highlight of the Southwest Agricultural District (light blue area) and the location of each phenology measurement. Maps were created using ArcGIS software version 10.7 by Esri. ArcGIS and ArcMap are the intellectual property of Esri and are used herein under license. Copyright Esri. All rights reserved. For more information about Esri software, please visit www.esri.com.

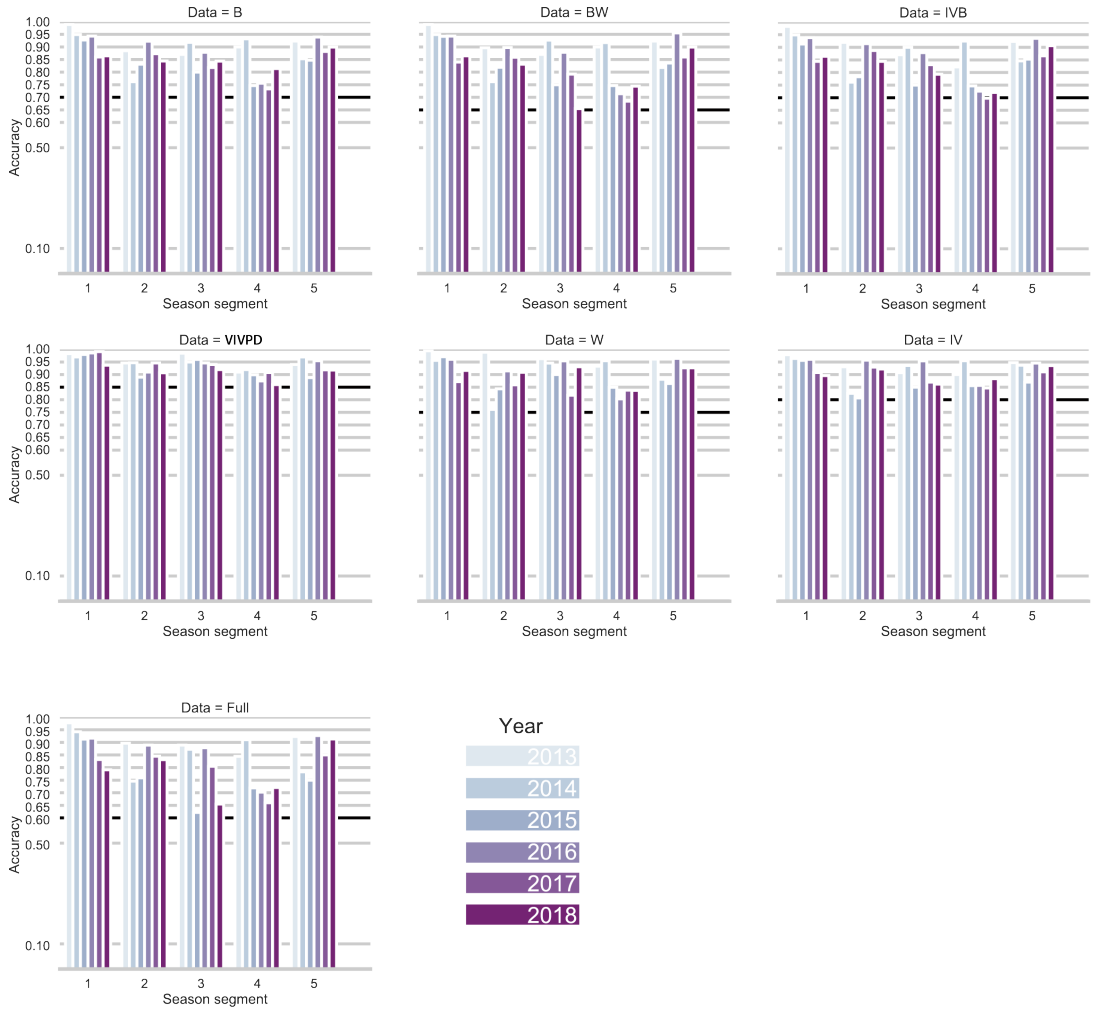

Figure 2.2: Accuracy values for each model during the season from 2013 to 2018. The black horizontal line shows the lowest value recorded. B = spectral bands 2, 3, 4, 5, 6, 7, 10, and 11 for Landsat 8; BW = spectral bands in addition to precipitation, maximum and minimum temperature, VPD, and GDU (weather parameters); VIB = spectral bands in addition to NDVI, EVI, GCVI, GVMI, and NDWI (vegetation indices); VIVPD = vegetation indices and vapor pressure deficit (VPD); W = weather parameters; VI = vegetation indices; FULL = spectral bands, weather parameters, and vegetation indices; LLD = latitude, longitude and doy.

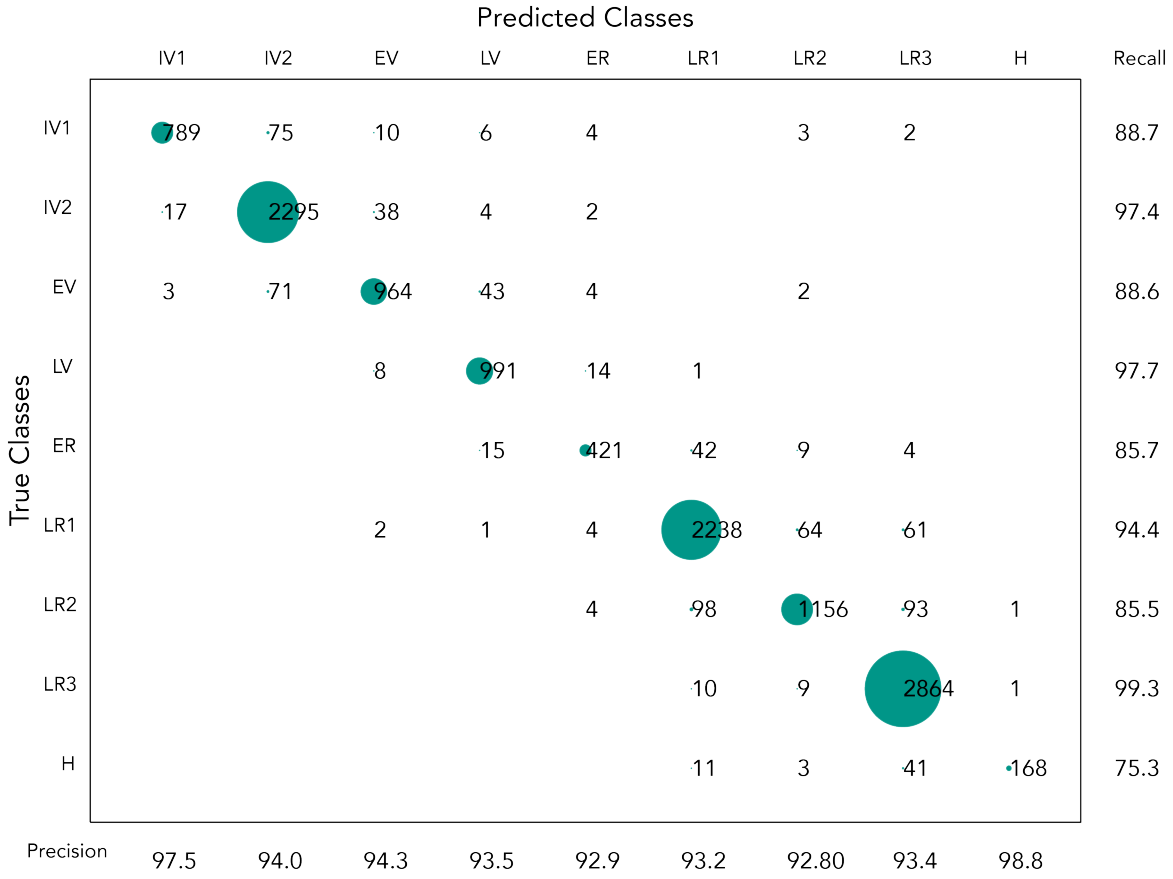

Figure 2.3: For the years 2013 to 2018, a classification matrix was created that included all classified elements, their magnitude, as well as precision and recall scores. The true classes are on the Y axe, while the model's predictions are on the X axe. The number of fields classified is represented visually by the size of the bubbles; the numbers along the main diagonal represent those successfully classified.

Table 2.1: Phenology classes (IV1-H) after re-grouping the phenological stages (VE - R6). IV1 = initial vegetative 1, IV2 = initial vegetative 2, EV = early vegetative, LV = late vegetative, ER = early reproductive, LR1 = late reproductive 1, LR2 = late reproductive 2, LR3 = late reproductive 3, H = harvest.

Class	**Stages on the dataset**
IV1	VE-V1
IV2	V2-V4
EV	V5-V8
LV	V9-Vnth
ER	VT-R3
LR1	R4
LR2	R5
LR3	R6
H	Harvest

Table 2.2: Classification metrics when 2014, 2015, 2017, and 2018 years were used to train the model, and 2013 and 2016 as validation dataset. A total of 14431 elements (corresponding to the phenological measurements for the years 2013 and 2016) were used for validation and a total of 26088 elements (corresponding to the p phenological measurements for the years 2013, 2015, 2017 and 2018) were used for training.

Class	**f1-score**	**Precision**	**Recall**
IV1	0.91	0.98	0.85
IV2	0.94	0.94	0.94
EV	0.86	0.85	0.86
LV	0.91	0.85	0.96
ER	0.92	0.94	0.91
LR1	0.96	0.98	0.94
LR2	0.88	0.88	0.88
LR3	0.97	0.95	0.99
H	0.95	1.0	0.90
Accuracy	0.93		
OOB	0.94		

Table 2.3: Classification metrics when testing for spatial transferability using all the years.

Class	**F1-Score**	**Precision**	**Recall**
IV1	0.60	0.97	0.43
IV2	0.83	0.77	0.90
EV	0.67	0.73	0.62
LV	0.85	0.80	0.91
ER	0.40	0.60	0.30
LR1	0.65	0.55	0.78
LR2	0.33	0.68	0.22
LR3	0.85	0.76	0.96
H	0.23	0.40	0.15
Accuracy	0.72		
OOB	0.75		

Chapter 3

Impact of High-Cadence Earth Observation in Maize Crop Phenology Classification

*Published in remote Sensing

Nieto, L.; Houborg, R.; Zajdband, A.; Jumpasut, A.; Prasad, P.V.V.; Olson, B.J.S.C.; Ciampitti, I.A. Impact of High-Cadence Earth Observation in Maize Crop Phenology Classification. Remote Sens. 2022, 14, 469.

https://doi.org/10.3390/rs14030469

Abstract

For farmers, policymakers, and government agencies, it is critical to accurately define agricultural crop phenology and its spatial-temporal variability. At the moment, two approaches are utilized to report crop phenology. On one hand, land surface phenology provides information about the overall trend, whereas weekly reports from USDA-NASS provide information about the development of particular crops at the regional level. High-cadence earth observations might help to improve the accuracy of these estimations and bring more precise crop phenology classifications closer to what farmers demand. The second component of the proposed solution requires the use of robust classifiers (e.g., random forest, RF) capable of successfully managing large data sets. To evaluate this solution, this study compared the output of a RF classifier model using weather, two different satellite sources (Planet Fusion; PF and Sentinel-2; S-2), and ground truth data to improve maize (*Zea mays* L.) crop phenology classification using two regions in Kansas (Southwest and Central) as a testbed during the 2017 growing season. Our findings suggests that high temporal resolution (PF) data can significantly improve crop classification metrics (F1-score = 0.94) relative to S-2 (F1-score = 0.86). Additionally, a decline in the F1-score between 0.74 and 0.60 was obtained when we assessed the ability of S-2 to extend the temporal forecast for crop phenology. This research highlights the critical nature of very high temporal resolution (daily) earth observation data for crop monitoring and decision making in agriculture.

3.1 Introduction

Crop phenology is the study of biological processes such as emergence, flowering, and senescence that are linked to and in response to environmental growing conditions (Liang et al., 2011). Understanding changes in crop phenology and their response to climatic conditions is critical for production-beneficial management practices (Ruml & Vulić, 2005). While the literature describes several initiatives to address crop phenology via earth observations (remote sensing) (Henebry & de Beurs, 2013; Sakamoto et al., 2010; Shihua et al., 2014; H. Wang et al., 2020), or mathematical models supported on weather information (Rezaei et al., 2018; Sharifi et al., 2017), the metrics reported from those studies are not easily translated into effective agronomic management strategies. For instance, Land Surface Phenology (LSP) determines the change in green vegetation condition by identifying changes in the vegetation seasonal pattern using remote sensing technologies (Henebry & de Beurs, 2013), but these changes do not directly reflect the true phenological stage of a distinctive crop (F. Gao & Zhang, 2021). Crop phenology shows more relevant stages from an agronomic perspective, which are closely connected to the critical stages when management approaches may and should be handled.

In the United States, official phenology crop progress estimates are based on weekly survey data obtained from a network of regional extension agricultural agents (F. Gao & Zhang, 2021). These reports summarize crop development by agricultural district in percentage terms, combining some growth stages into broader categories. While this is a valuable source of knowledge, the data gathering procedure is time-consuming and labor-intensive, and may

not adequately represent a county or district, or the reality of a single farm, with a high degree of fidelity (F. Gao et al., 2017). Additionally, a marked decrease in the reporting of certain crops such as soybean (*Glycine max L.*), sorghum (*Sorghum bicolor L.*), or sunflower (*Helianthus annus L.*), has been documented in recent years (United States Department of Agriculture (USDA), 2018a,b), based on the information from the weekly reports, creating an information gap for critical crops.

Whereas remote sensing technologies have historically assisted in the execution of a variety of Earth observation activities (Vina et al., 2004; M. A. White et al., 1997; X. Zhang et al., 2003) current sensors and future missions are aiding in addressing tasks that need high cadence observation (Houborg et al., 2015; Rast & Painter, 2019; Wulder et al., 2019). These features are one of the most crucial factors to address crop phenology progress and monitor changes in vegetation dynamics over space and time domains (F. Gao & Zhang, 2021). Sensors with daily or near-daily coverage have not traditionally been available at the spatial resolution needed for field analysis, while those sensors with high spatial resolution have been characterized by a comparatively low revisit period. Several fusion efforts are underway to merge those characteristics, scattered through different spaceborne and airborne sensors, into one product to retrieve higher temporal, spatial, and spectral resolutions (Arun et al., 2022; Claverie et al., 2018; F. Gao et al., 2017; Liao et al., 2019). Data fusion, leaning on the use of deep learning and multimodality, is currently the best path to achieve within season analysis capable of thoroughly characterizing crop growth in both large- and small-scale environments or landscapes (Arun et al., 2022; F. Gao et al., 2017; Schramowski et al.,

2020).

Along with high-cadence data, robust classifiers capable of managing large data sets with acceptable trade-offs and stability while staying computationally feasible are required for classifying the numerous phenological phases successfully (Gislason et al., 2006). For decades, Random Forest (RF) (Breiman, 2001) has proved to be useful for remote sensing applications (Belgiu & Drăguţ, 2016; Pal, 2005; V. F. Rodriguez-Galiano et al., 2012). In a study comparing more than 170 classifiers, Fernández-Delgado et al. (2014) reported a superior performance for RF mainly when dealing with large and unbalanced data sets. These characteristics, along with their simplicity and processing speed, make RF the classifier of choice in a wide variety of scenarios, where data sets are not large enough to test more complex machine learning methodologies such as neural networks (V. Rodriguez-Galiano et al., 2015). A significant question addressed in this publication is whether high-cadence satellite data influence crop phenology classification for agronomic applications. The focus of this research is to compare the output of an RF classifier model, testing two different sources of satellite data varying in their temporal and spatial resolution in combination with weather data. The specific objectives for this study were: (i) to utilize surface reflectance data from Planet's data fusion product (PF-SR) and ESA's Sentinel-2 sensors to quantify performance for crop phenology classification; and (ii) to evaluate the spatial transferability of the generated model by testing its performance in two distinct maize production regions, Southwest: mainly irrigated with average yields above 11 Mg ha-1; and Central: rainfed, yields below 9 Mg ha-1, within the state of Kansas, in the Great Plains region of the United

States (US) during the 2017 growing season.

3.2 Materials and Methods

3.2.1 Study area

The present study was conducted in the state of Kansas (US) within the Southwest (SW) region comprising the counties of Morton, Stanton, Grant, and Stevens and the Central (CK) region constituted by the counties of Stafford and Pratt (Fig. 3.1,a). The precipitation pattern in Kansas is highly uneven in both quantity and distribution, ranging from over 1000 mm from the southeast to less than 300 mm in the west region of the state. For the two regions selected in this study, these levels ranged from 500 to 800 mm for the CK (rainfed region) and from 300 to 500 mm for the SW region (Lin et al., 2017). The size of the fields, as well as standard management practices such as ground water irrigation, differ in these two regions. The SW region is mostly dominated by larger fields relative to the CK, and a higher number of irrigation pivots (fed from the Ogallala Aquifer), resulting in higher maize yields averaging 11 Mg ha-1 for year 2019 (United States Department of Agriculture (USDA), 2018b). In contrast, the CK region is characterized by smaller fields and mainly dominated by rainfed agriculture with an average maize yield of 9 Mg ha-1 for the year 2019 (United States Department of Agriculture (USDA), 2018b).

3.2.2 Data characteristics

Reference ground field data

Acquiring quality environmental data may be challenging and expensive, even more when dealing with extensive geographic areas (Hanks et al., 2011; Ruiz-Gutierrez et al., 2016). This requires not just data collection and subsequent storage, but also the training of a person capable of delivering the information with minimal to no bias or errors. Despite these substantial obstacles, ground-truth field data continue to be the most reliable and closest to the true condition of the system (Dickinson et al., 2012; Hooten et al., 2007). For this study, we were able to acquire a valuable source of ground-truth data from Crop Quest Inc., which included geolocated fields spanning many states in the Central region of the United States. This data set was developed during a five-year period, from 2013 to 2017, by conducting repeated trips to each farmer field to document maize crop phenology changes during the growing season. Only the growing season 2017 was analyzed in the present manuscript due to a conjunction of events, related to the sources of information.

On average, each field was visited 5 times during the season; although, this varied within all field observation collected and summarized in this large database. The crop phenology stages present in the data set followed the scale described on Fig. 3.1,b , based on Ciampitti et al. (2016). The final database included geolocated fields for both regions, crop phenology measurements per field, and the date of data collection.

Remote sensing data and weather variables

Google Earth Engine (GEE) was used to retrieve and analyze spectral remote sensing data, as well as for the weather information (Fig. 3.1, c). This cloud-based platform enabled the processing of large data sets and provided access to a wide library of pre-processed satellite images, climate, and land cover data, as well as built-in algorithms and vector data (Gorelick et al., 2017; Kumar & Mutanga, 2018).

Planet Fusion Product

For the 2017 season, our industry partner provided daily, 3 m resolution, Planet Fusion (PF) Surface Reflectance (SR) data (Planet Fusion Team, 2021). Planet Fusion constitutes a comprehensive fusion methodology based on the CESTEM algorithm (Houborg & McCabe, 2018a,b) to inter-calibrate, harmonize, enhance, and fuse multi-sensor data streams leveraging rigorously calibrated public mission (i.e., Sentinel-2, Landsat 8, MODIS, and VIIRS) products in concert with the higher spatial and temporal resolution CubeSat images from the PlanetScope constellations. PF uses the Framework for Operational Radiometric Correction for Environmental Monitoring (FORCE) (Frantz, 2019) to generate a harmonized Sentinel-2 and Landsat 8 BRDF adjusted SR product to be used as the calibration target during the CESTEM-based harmonization step. Additional PF features include: (1) sub-pixel fine geometric alignment of source imagery; (2) rigorous, temporally driven, cloud and cloud shadow detection; (3) fusion of Sentinel-2 and Landsat 8 data to help fill gaps in PlanetScope coverage; and (4) advanced gap-filling. PF delivers daily, gap-filled, 4-band (0.45–0.51, 0.53–0.59,

0.64–0.67, and 0.85–0.88 µm) sensor agnostic SR data characterized by enhanced radiometric stability and consistency across space and time to support advanced analytics. The PF data were provided with a 3 m pixel size as regularly gridded raster tiles (24 by 24 km) in UTM projection.

Due to the high temporal frequency, spatial resolution, and quality of data, this product could prove to be an excellent resource for dealing with detailed crop phenology and its temporal variation. For the SW region, tiles 11E-172 N and 11E-173 N were utilized, as well as tile 21E-174 N for the CK region. Both spectral and associated metadata from each tile were uploaded to GEE via the Google Cloud Repository and organized into a collection for convenient access during feature extraction.

Sentinel-2

Sentinel-2 (S-2) is an optical satellite constellation developed by the European Space Agency (ESA) that is composed of two complementary spacecraft (S-2A and S-2B). This constellation was chosen to compare the PF product to a publicly available, well-studied, and widely reported satellite with a high temporal resolution (median average revisit intervals of 3.7 days, (J. Li & Roy, 2017)) and spatial resolution (10, 20, and 60 m, (Drusch et al., 2012)).

Tile T14SKG covering southwest Kansas, was used to generate level-2A surface reflectance data using the Sen2Cor processor (Main-Knorn et al., 2017). This operation was carried out manually using the Sentinel Application Platform (SNAP) and the Sen2Cor plugin (version 2.8), as not all tiles for level-2A are available from the GEE archives for the year 2017 over the state of Kansas. Once the images were converted to bottom of atmosphere

reflectance, they were uploaded to GEE for the purpose of calculating the various vegetation indices.

Even though 12 images were produced for the 2017 growing season, only eight of them showed less than 20% cloud cover, and only five of them were time corresponding with observations made during the ground truth field data gathering.

Spectral bands and vegetation indices

A suite of Vegetation indices (VI) including the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Green Chlorophyll Vegetation Index (GCVI), Chlorophyll Vegetation Index (CVI), and Normalized Green Index (NGI) were calculated based on the spectral bands from both products (PF and S-2). Extensive literature is associated with these indices and their capability to monitor vegetation dynamics (Gitelson et al., 2003; Tucker, 1977, 1979; Turner et al., 1999; Viña et al., 2011).

NDVI is supported by the fact that green/healthy vegetation will have a higher reflectance in the near infrared (NIR) while absorbing in the red band due to the chlorophyll content in the leaf (Tucker, 1979). While extensively utilized, it is well established that this indicator saturates at intermediate biomass levels, owing mostly to its limited sensitivity to elevated chlorophyll concentrations, especially problematic in maize fields. The EVI integrates the blue band to limit the impact of certain atmospheric phenomena and more precisely capture changes in plant canopy, particularly under high productivity areas (A. Huete et al., 2002; H. Q. Liu & Huete, 1995), making this vegetation index as an effective alternative to the problem from using NDVI. GCVI has been proven to present a greater linear connection

with the leaf area index of maize than other indices and, unlike NDVI, does not saturate as quickly at high biomass levels (Gitelson et al., 2003). Lastly, both CVI and NGI have been widely used due to their responsiveness to changes in the chlorophyll content of the crop (Sripada et al., 2006; Vincini & Frazzi, 2011).

Weather variables

Weather variables were derived from the GEE archives based on gridded surface meteorological data (GRIDMET), (Abatzoglou, 2013). This data set integrates high-resolution spatial data from PRISM (2.5 arc minute, 4 km) with regular temporal resolution data from the North American Land Data Assimilation System (NLDAS) (Abatzoglou et al., 2014). The main meteorological variables retrieved for the analysis were mean, minimum, and maximum temperatures, precipitation, and vapor pressure deficit (VPD), all on a daily temporal scale.

3.2.3 Data preparation

To extract circular fields linked to irrigated pivot operations, the Canny edge detector (Canny, 1986) method was implemented on the GEE platform. Using this mask, in conjunction with the fields under study, we were able to construct a layer that included information about the geolocation and phenology of each irrigation field in the data set. As a result, the mean values of the spectral bands (red, green, blue, and NIR), VIs (NDVI, EVI, GCVI, CVI, and NGI), and the weather parameters (precipitation, minimum temperature, maximum temperature, and VPD) associated with each field for each given date were all gathered to construct the main database with both sensors.

Despite PF daily availability, only images day-coincident with the ground-based phenology collections (a total of 91 dates) were utilized (Fig. 3.2). These dates were distributed between 1 May and 30 October. The same method was applied to S-2 data after filtering based on a cloud mask with a cloud threshold less or equal to 20%. During that process, S-2 images with less than 20% cloud cover were retained for the analyses. With this filtering protocol, only five S-2 images (dates: 15 June, 30 June, 5 July, 25 July, and 14 August) were available on the same date as the phenology collections (Fig. 3.2). Due to this temporal limitation, additional studies were conducted to also consider phenology data acquired within 3 and 10 days of the image collection. As a result, this stage of the study analyzed three data sets: (i) containing the spectral band values, VIs, and weather parameters for the exact phenology date; (ii) same parameters as in (i) spanning three days before and after the crop phenology collection date (7 days total); and (iii) same parameters as in (i) ranging ten days before and after the crop phenology collection date (21 days total).

3.2.4 Random forest

The Random Forest (RF) classifier was implemented with the Scikit-Learn library (Pedregosa et al., 2011), in a Python 3.8 environment. RF represents a supervised classifier, learning from training samples; thus, the quality of training instances is a key component in the classification process. The RF classifier, in addition, is an ensemble classifier, in which multiple decision trees are utilized to improve generalizability and, therefore, resilience by using a randomly selected subset of training samples and variables (Dietterich, 2000). All these as-

pects contribute to the creation of accurate classifications, as well as the processing of large amounts of data. It has been utilized in remote sensing for more than two decades, proving valuable in the handling of high-dimensional data, unbalanced data sets, and multi-target problems, while remaining simple and adaptable using multiple hyperparameters (Belgiu & Drăguţ, 2016; Gómez et al., 2016; Pal, 2005; Pelletier et al., 2016; V. F. Rodriguez-Galiano et al., 2012; Tatsumi et al., 2015).

To produce a more robust model in this study, the hyperparameters were derived based on an exhaustive grid search checking for various values at each hyperparameter and retaining the ones that resulted in the best performance. The data set for this test was created by combining the variables in their final configuration (further details presented in the section below). This resulted in the following hyperparameter values: bootstrap: True; maximum tree depth: 10; maximum number of features required to do the best split: 'auto'; minimum sample size per leaf: 2; minimum sample size required to execute the split: 2; number of trees: 1000. To evaluate performance, the data sets (PF and S-2) were split into training, validation, and test sections for each experiment. From the total of the data, 70% was used for training purposes and validation (fraction A) and 30% for model assessment (fraction B). Fraction A was further divided into 70% for training (fraction Aa) and 30% for validation (fraction Ab).

Performance metrics

When evaluating the performance of RF, the most commonly used metrics are overall accuracy, precision, recall, and F1-score (Congalton, 1991). Considering the examples in the

literature we can obtain these metrics following the Equations 3.1, 3.2, 3.3, 3.4 (Goutte & Gaussier, 2005). Most of the literature studying the performance of these metrics is centered on balanced binary classifications.

In this paper, the ground-truth data set presents more than 18 classes, where each class does not contain the same number of samples. According to (Branco et al., 2016), F1-score, defined as the harmonic mean of precision and recall, can be a useful metric to use with respect to unbalanced data sets. In contrast, the use of overall accuracy, the total of correctly identified elements divided by total number of elements to be classified, can be misleading since it will favor the most represented class (G. H. Nguyen et al., 2009). Precision represents the fraction of samples assigned to the class that are in fact part of the class, and recall represents how well the class was predicted.

$$accuracy = \frac{True\ positives\ +\ True\ negatives}{All\ samples} \tag{3.1}$$

$$precision = \frac{True\ positives}{True\ positives\ +\ False\ positives} \tag{3.2}$$

$$recall = \frac{True\ positives}{True\ positives\ +\ False\ negatives} \tag{3.3}$$

$$F1\ score = \frac{Precision\ x\ Recall}{Precision\ +\ Recall} \tag{3.4}$$

Feature importance

In classification tasks, feature selection is commonly used to reduce the dimensionality of the data set (Gislason et al., 2004; V. F. Rodriguez-Galiano et al., 2012) and to inform on the relative influence of each variable in the model. To narrow down the possible combinations of variables to produce the best classifier, a feature importance analysis was conducted based on the PF data set, for both the SW and CK regions. For this purpose, all the variables (VIs, spectral bands, weather data, geolocation of the fields, and day of the year of field collections) were utilized.

3.3 Results

3.3.1 Best combination of variables

The analysis of the feature importance from the SW region data set revealed that NIR band (B4) and EVI, as well as the minimum and maximum temperatures (Tmin and Tmax), VPD, and DOY, were associated with the highest scores. When all the variables were combined, we found that the best F1-scores were 0.94 for the SW region and 0.93 for the CK region. Additionally, and to investigate other model possibilities, several combinations were attempted (only spectral data; only weather parameters; EVI and weather; B4 and weather, etc.). The lowest F1-scores were derived when the model was only driven by weather variables, with 0.79 for SW and 0.76 for CK regions.

3.3.2 Model performance using PF data in both regions

The model including NIR, EVI, Tmin, Tmax, VPD, and DOY was utilized to classify regions, yielding F1 values of 0.94 for the SW and 0.93 for the CK regions (Fig. 3.3, a, b). The classification results for each crop phenological stage are presented as a confusion matrix, where all the components that were correctly identified are along the main diagonal, as well as the resulting precision and recall along the X and Y axes.

When examining the SW area in detail, most of stages produced values between 1.0 and 0.75 (recall) and 1.0 to 0.80 (precision), whereas certain stages, such as R1, R3, and V15, showed poor metrics ranging from 0 to 0.66. Similarly, the findings for the CK region showed that most stages had recall values ranging from 1.0 to 0.87, but some (R2, V10, V11, V12, and V13) exhibited recall values below 0.67. When the precision is considered, all stages ranged from 0.76 to 1.0.

3.3.3 Model performance using Sentinel-2 data

When the same model was applied to S-2 data, the findings were inferior to those previously discussed in this paper. The first data set included just three crop phenological stages, due to the season timeframe captured by the images. The classification resulted in an overall high F1-score (0.86), but inconsistent metrics for each stage, for example, for VT stage both recall and precision were zero (Fig. 3.4, a). As a second step, the performance of the extended data set, including data from 3 days before and after the time of field data collection, was tested. In this experiment, the overall F1-score was reduced to 0.74 but more phenological

stages were part of the data set. Lastly, the test was repeated by expanding the search from 3 to 10 days before and after the field data collection time. This resulted in an F1-score of 0.58 with several crop phenology stages still not classified (null metrics). When the S-2 data set was employed, the overall performance was consistently inferior to when the PF product was utilized (Fig. 3.4).

3.4 Discussion

This study demonstrates the benefits of having access to high-cadence earth observation at both spatio-temporal dimensions. This analysis employs a simple model trained on spectral characteristics, meteorological variables, and high-quality ground truth labels to achieve optimal crop phenology classification, using maize crop in this US region as testbed for this case study. The findings presented in this study are in accordance with those observed when tracking crop phenology (Bandaru et al., 2020; Gislason et al., 2004; Liang et al., 2011; B. Peng et al., 2018; M. A. White et al., 1997), most notably the importance of enhanced temporal resolution (F. Gao et al., 2017; F. Gao & Zhang, 2021; Nieto et al., 2021).

This combination of variables resulted in overall classification metrics higher than 0.90, with each phenological stage identified accurately as described in sections above. Despite the disparities in the characterization of both regions under analysis (SW and CK) in terms of climate, field size, and average yield, the model performance was similar for both areas using the same combination of variables. Previous studies have highlighted the importance of VIs when monitoring changes in crop phenology, and particularly EVI has been found

to outperform a broad variety of VIs in maize fields (F. Gao et al., 2020; Tatsumi et al., 2015; Zhong, Hu, et al., 2016; Zhong, Yu, et al., 2016). In addition, promising results were achieved by Cai, Guan, Lobell, et al. (2019) relying on the NIR band. On the other hand, inclusion of relevant weather variables is crucial to describing changes in crop phenology and development, as demonstrated in modeling scenarios (Bai et al., 2010; Cai et al., 2018), with air temperature among the most relevant variables (Joshi et al., 2021). In addition, VPD has been strongly associated with water stress and the consequent impact on crop development (Azzari et al., 2017; S. Zhang et al., 2017).

This study also highlights the relevancy of high-cadence satellite data for classifying crop phenology more accurately. Due to cloud cover only a few S-2 images were available during the 2017 growing season. This significantly reduced the likelihood of finding an image acquired on the same day as the field data collection, which affected the ability to classify the crop phenology stages. In order to emphasize the importance of temporal resolution for this purpose, a sensitivity analysis using PF data only was conducted evaluating the performance implications of excluding data from within 1, 2, and 3 days of the date of the field data collection, respectively. The ability of the classifier to achieve high prediction metrics was substantially reduced by 12% with the 2-day, and by roughly more than 40%, with the 3-day departure, especially during crop developmental phases around the time of flowering (e.g., V12, V13, and V19). These results emphasize the importance of having access to low-latency spectral data, particularly during periods of rapid crop growth.

Potential caveats and uncertainties from the analysis are related to problems associated

with multiclass classification and imbalanced data sets. Most of the methods to treat imbalanced data sets rely on binary problems (Hsiao et al., 2019) and solutions are not extensively proven. Having access to daily spectral data as well as increasing the number of ground truth samples represent key factors in the attempt to build more reliable predictions. In this study, this was partially overcome using the daily PF product that resulted in very robust and consistent classification results. More efforts are being conducted towards the development and availability of fusion products, such as the Harmonized Land-sat-Sentinel (HLS); although, it was not implemented on this analysis since, at the time, it was suggested to not be yet used for scientific purposes (Hoens et al., 2012).

Further research should focus on verifying the existing model on smaller scales (e.g., smallholder farms) to determine its transferability to more complex farming systems and management approaches, as well as its contribution to precision agriculture initiatives (LP DAAC, n.d.; Mulla, 2013). Additionally, since this study focused only on maize, a crop with well-defined stages, further research incorporating soybeans, sorghum, and sunflower, among others, should be explored. Extending the research period to include seasonal weather scenarios representative of current and future climatic conditions (e.g., heat, drought, and changes in rain patterns) in the context of climate change is another critical avenue to pursue, as is examining the influence of hyperspectral data from DECIS, HISUI, PRISMA, and future CHIME, and SBG missions (Seeley & Asner, 2021).

3.5 Conclusions

A proposed random forest model combining satellite-based (NIR band surface reflectance, and EVI), weather data (maximum and minimum temperatures, and VPD), and day of the year enabled a high-quality crop phenology classification for maize crops under two contrasting production regions. The study highlights the critical importance of having access to high-cadence earth observation, as well as high-quality ground truth data. Classification accuracy was shown to decrease significantly when comparing performance metrics from a scenario with limited temporal availability (S-2 clear-sky images) with a daily data scenario (Planet Fusion product), particularly when the same image was used to make inferences about the future or the past. In summary, this study successfully portrayed the spatial transferability of this satellite- and weather-based crop phenology model applied to contrasting maize producing regions while preserving performance. Therefore, this simple crop phenology model can become a promising tool that still needs to be globally tested.

Most of the existing crop phenology solutions rely on broad metrics as well as large geographic regions. While these are important tools for policymakers and governments to use when making global or regional decisions, there is still room for development in terms of farmer solutions. Additionally, these measurements do not reflect the crop's real status but rather the vegetation's spectral information. This discrepancy between what is reported and what occurs in the field might result in misconceptions and incorrect conclusions. Improving crop phenology classification and reporting has a substantial influence on every step of the food production chain, enabling accurate, proactive, rather than reactive, strategies at the

regional and field level.

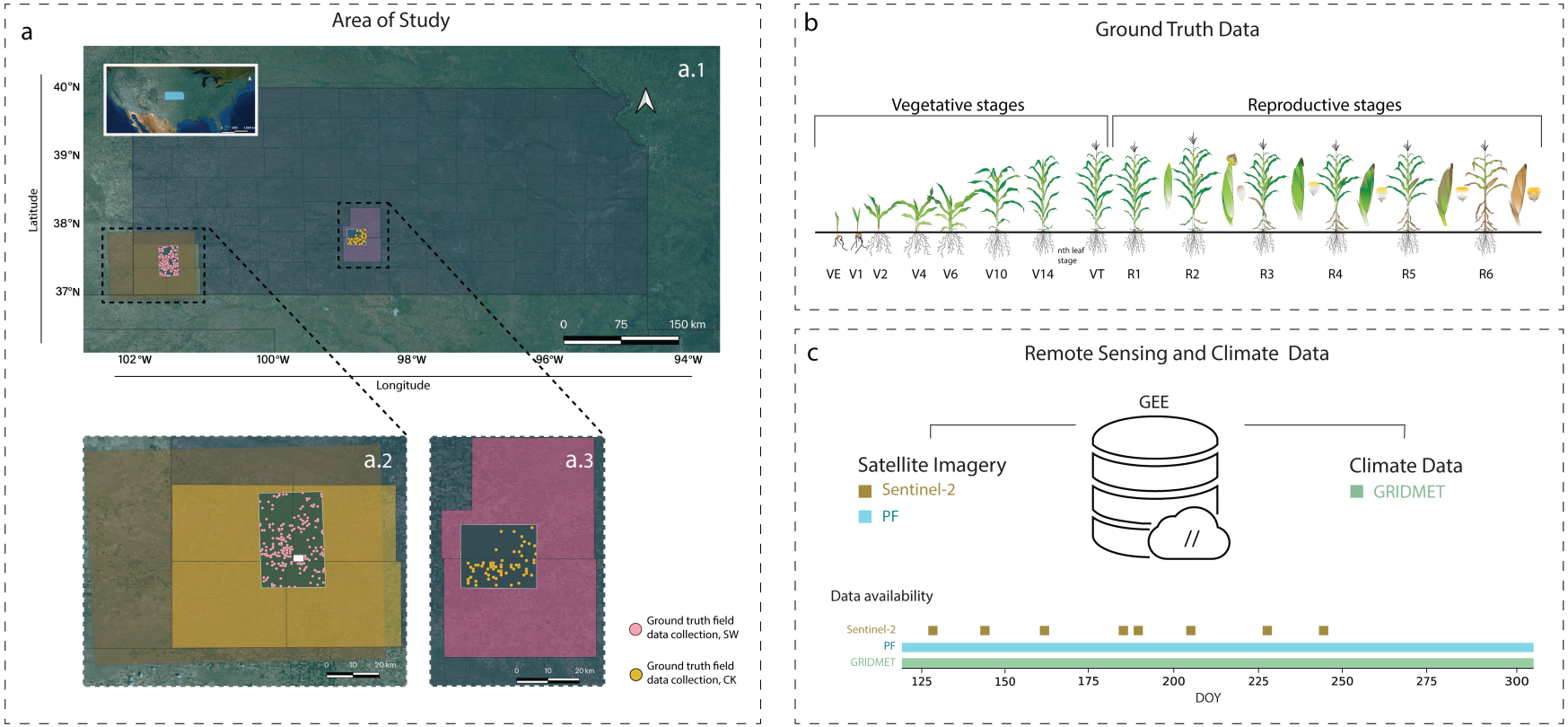

Figure 3.1: (a) Area of study (a.1) state of Kansas (US), (a.2) Southwest (SW) region; the large orange rectangle represents the Sentinel-2 tile covering the area; the green rectangle corresponds with the Planet Fusion (PF) tiles over the area; pink dots represent ground truth field data collection sites, (a.3) Central KS (CK) region; green rectangle corresponds with the PF tile over the area; yellow dots represent ground truth field data collection sites. (b) Maize phenology stages, with the phenology scale used during ground truth data collection, following Ciampitti et al. (2016). (c) Sentinel-2 and PF data imagery and GRIDMET data from the Google Earth Engine (GEE) repository, data availability differs temporally with 8 images collected for Sentinel-2, and daily data for both PF and weather data during the 2017 growing season, varying with the day of the year (DOY).

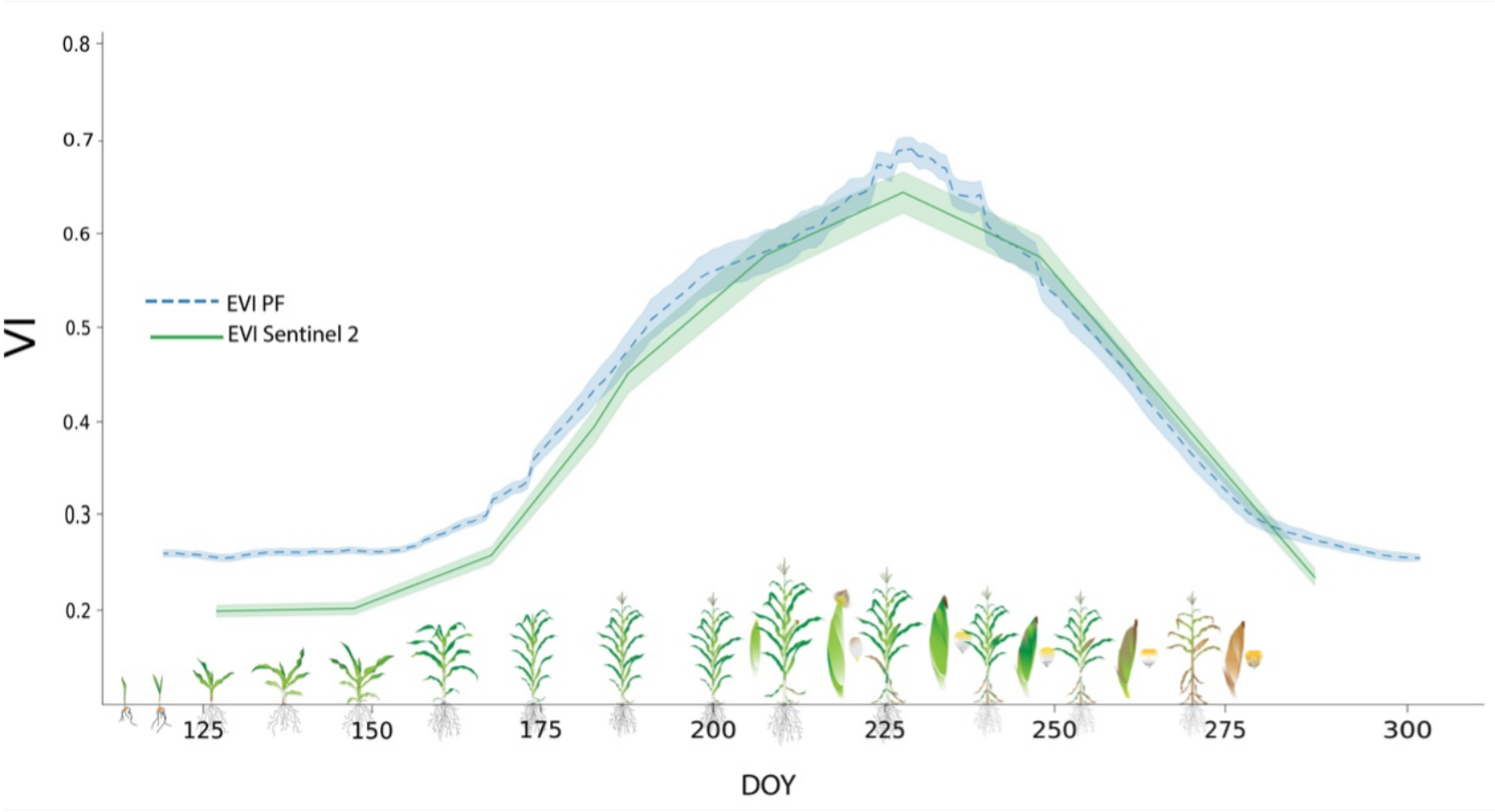

Figure 3.2: EVI temporal series for the 2017 maize growing season for SW fields using Sentinel-2 (solid green line) and PF (dash blue line). The shaded area around both lines represent the standard deviation of the EVI values for the fields under study (pink dots 3.1 (a.2)). Maize phenology progress and development stages is depicted on the x-axis.

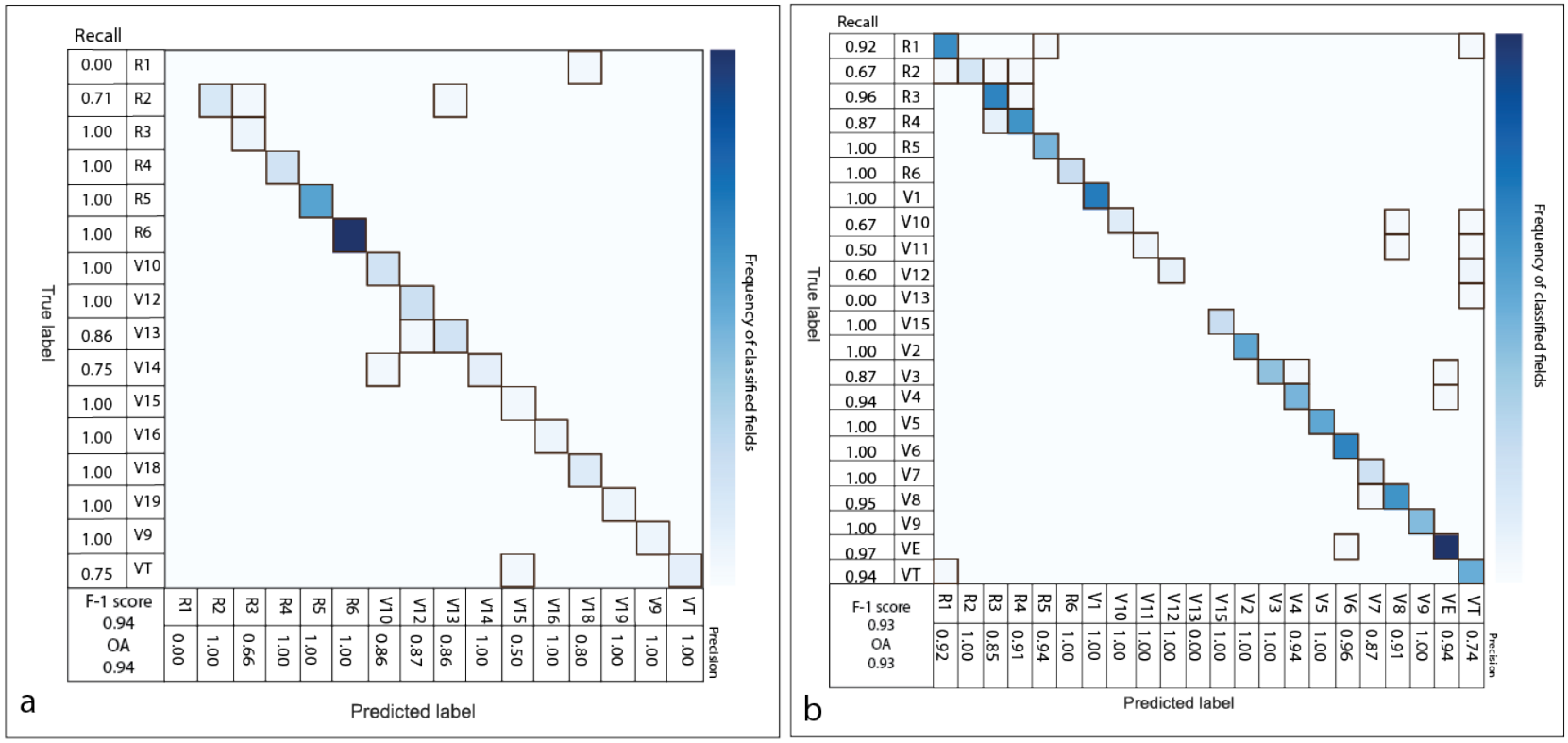

Figure 3.3: (a) Confusion matrices with PF-based classification results for the Southwest region (SW) and for (b) Central KS region (CK), including recall and precision for each crop phenology stage, and F1-score and overall accuracy (OA).

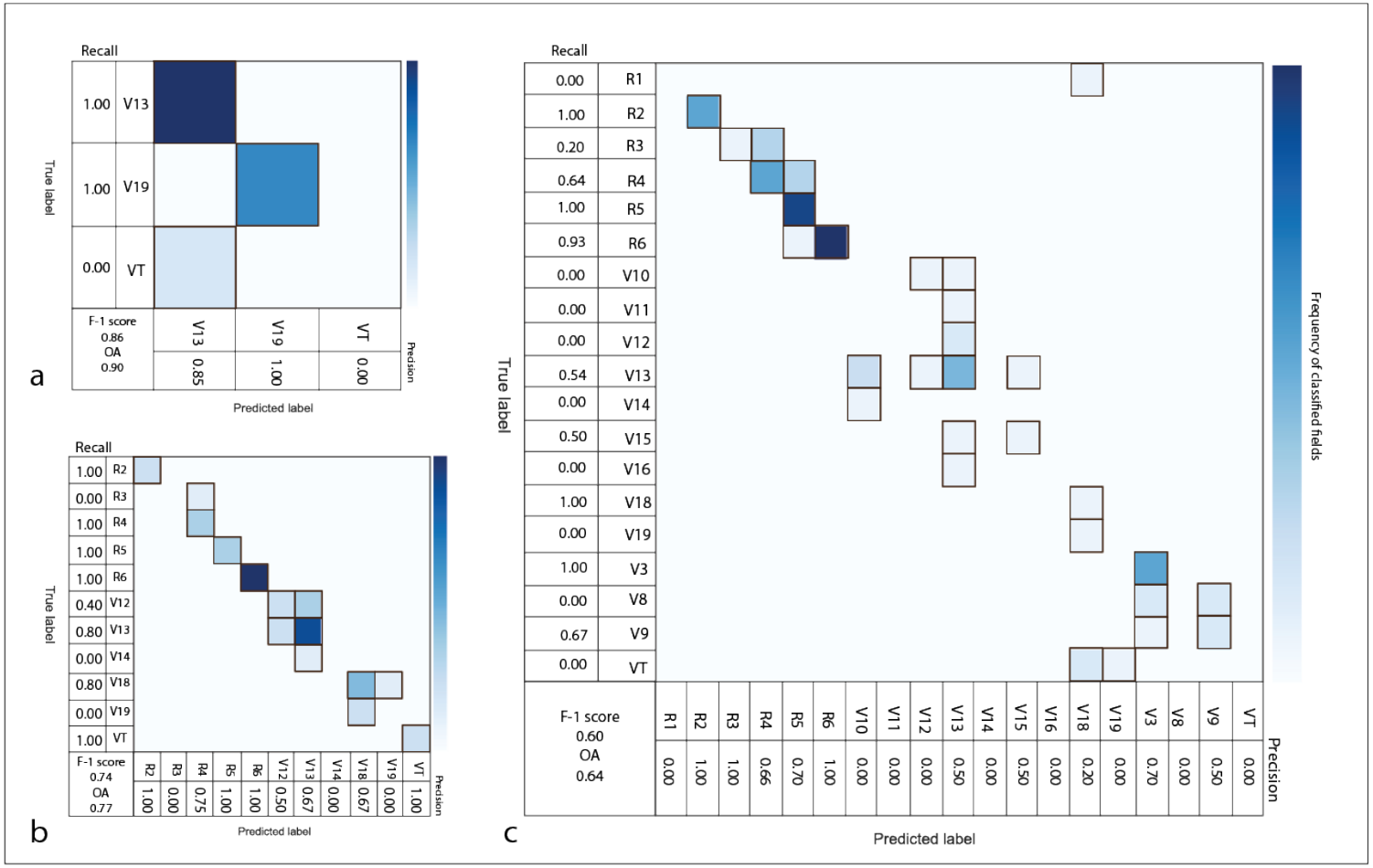

Figure 3.4: Confusion matrices with S-2 based classification results for the Southwest (SW) region, (a) including imagery on the date of data collection, (b) including data within 3 days before and after, (c) including data within 10 days before and after. All matrices include recall and precision for each crop phenology stage, and F1-score and overall accuracy (OA).

Chapter 4

Using EfficientNet Pretrained Model for Segmentation-Based Classification of Agricultural Land Features: a case study in Battambang, Cambodia

Abstract

Conservation agriculture is an enticing approach for improving soil health, increasing crop yields, and offsetting the negative effects of conventional agriculture. Recent changes in land use and agricultural expansion in Cambodia have brought to rapid changes in soil degradation, threatening long-term productivity of current agricultural systems. To address this issue, a systematic method for monitoring changes in agricultural land use is required, and

remote sensing can be a useful tool for achieving this objective. Therefore, the aims of this study are to implement a segmentation-based method for classifying land features linked to changes in agricultural land use, placed as a case study in the province of Battambang, Cambodia. To accomplish the image segmentation task, a deep learning model EfficientNetB2 was applied to a dataset that comprised a total of 6 classes (road, vegetation, forest, burned, bare, and unclassified). Accuracy, Jaccard coefficient, and F1-score were used to evaluate the performance of the model. The model used to classify all 6 classes resulted on an average accuracy of 0.88, a Jaccard coefficient of 0.72, and an F1-score of 0.83. The outcomes of this study demonstrate the potential of combining remote sensing and deep learning models to effectively classify land features related to agricultural use. This can aid in monitoring agricultural land changes over time and establish a foundation for building a strategy to assist farmers in preserving their most valuable asset, their land.

4.1 Introduction

Conservation agriculture has evolved as a promising strategy for food production, emphasizing the preservation of natural resources such as soil and water (FAOa, n.d.). A set of practices linked to conservation agriculture has gained substantial traction in the last decades due to its potential to improve soil health, increase crop yields, and offset the negative environmental implications of conventional farming practices (Palm et al., 2014). By adopting those practices such as minimal (or reduced) tillage, crop rotation, and cover crops, conservation agriculture aims to develop a robust and sustainable food production system over the

long term (Hobbs et al., 2008; Lal, 2015). With an overgrowing population and increasing pressure on natural resources, efficient and sustainable agricultural production is essential (Lal, 2015).

As in many regions around the globe, new efficient and conservation approaches are needed to advance towards the Sustainable Development Goals of the United Nations, and the South Asia region is not exempt from this complex and challenging task. Within this region, nations such as Cambodia, characterized by the small-scale family farms (FAOb, n.d.) dominating the agricultural landscape is a relevant hot spot for advancing this critical aspect for sustainable agricultural intensification.

According to Samkol et al. (2015) most of Cambodian farmers list rice (*Oryza sativa* L.) as the main crop, with the average households owning 0.86 hectares (FAOb, n.d.). As a result of the lack of crop diversification, land tenure, size of farm, and other economic restrictions, farmers are frequently reluctant to take risks (Haushofer & Fehr, 2014), such as the one posed by changing the traditional practices, even when on-farm studies have already documented clear benefits in adoption of conservation practices. In addition, in recent decades, the conversion of native forested regions to agricultural lands has resulted in rapid changes in land use (Sourn et al., 2021). This process ends with negative consequences, impacting natural resources and promoting soil degradation, including soil erosion and decreasing soil fertility (Lemenih et al., 2005). In the short term, a clear effect on reduction in crop yield is expected, but in the long-term, this process compromises the overall productivity of the land (Homma et al., 2003).

Assessing the status of degraded lands and overall area is relevant to better understanding the nature of the problem and present solutions. Recent efforts on implementing strategies avoiding new deforestation (Peter & Pheap, 2015) and investing in expanded protected areas will help to mitigate the speed of land degradation over time. However, there is a trade-off between expanding protected areas versus improving the overall management of the newly established ones (Adams et al., 2019). It is evident that establishing a more systematic approach for monitoring land changes over time is needed at the country-scale. Following this rationale, implementation of remote sensing to assist in identifying land degradation areas (Dubovyk, 2017; Sourn et al., 2022), and identifying regions where soil has been left exposed (Viña et al., 2003) is a potential path forward. In recent years, access to powerful computers as well as cloud computing allowed exploration of more complex classification models, capable of working with larger datasets, and performing end-to-end training (Xie et al., 2019). The use of deep convolutional neural networks (CNN) started to become more ubiquitous to perform tasks for land use and classification (Ji et al., 2018; Zhong et al., 2019). In the case of smallholders in Cambodia, when agricultural fields are small and closely spaced, evaluating entire images without first segmenting them may result in misclassification due to the intersection with trees or bare land. In addition to reducing this risk, segmentation can minimize image noise and eliminate irrelevant characteristics that may hinder the model performance of the classification process (L. Chen et al., 2021; Shirvani et al., 2023).

Developing a more robust data pipeline for segmentation of remote sensing satellite imagery can assist with the classification of land features and provide a foundation for es-

tablishing a method to track changes in land use and degradation for smallholder farmers in Cambodia. Therefore, the aims of this study were to implement a segmentation-based method for classifying features associated with the agricultural management practices, using high spatial resolution images, with focus in the Battambang province (as a case study) in Cambodia.

4.2 Materials and methods

4.2.1 Study area

The present research was carried out in Cambodia, more specifically in the province of Battambang, over the Sangkae District (Fig. 4.1). The district is located between the coordinates 13.1°N and 103.1°E and encompasses approximately 421 square kilometers. The climate of the region is tropical, with distinct wet and dry seasons.

The average length of the rainy season is from May to October, with the wettest months from August to September (Sourn et al., 2021). The dry season typically runs from November to April, with daytime temperatures between 25 and 35 degrees Celsius and nighttime lows between 20 and 25 degrees Celsius (Sourn et al., 2021). Most of the soils of the region are composed of lacustrine alluvial, and their high fertility makes them ideal for agricultural use (Crocker, 1962).

4.2.2 Satellite imagery and vegetation indices

The analysis of satellite images was temporally concentrated during the dry season for this region of interest, spanning from November to April, coincident with the time when farmers prepare the soil for the next planting season.

Planet Fusion (PF) combines the best attributes from public and private satellite sensor resources to create a sensor-agnostic dataset. This is done by combining data from several sensors (Sentinel-2, Landsat 8, MODIS, VIIRS, and CubeSat) into a daily, gap free, 3 meters resolution image. This Analysis Ready Data (ARD) product has been proven in a variety of scenarios, where both temporal and spatial resolutions are needed. Examples of this are inter-day change detection, disturbance monitoring, dynamic land use and land cover classification, crop phenology, vegetation growth monitoring, and biophysical property and process retrievals (Planet, 2022). The PF data was provided as regularly gridded raster tiles (24 by 24 kilometers) in UTM projection with a pixel size of 3 meters. The spectral data and metadata from tile 13E-60N were uploaded to GEE via the Google Cloud Repository and grouped into a collection to facilitate its later manipulation.

From the satellite imagery data, blue, green, read, and near infrared spectral bands along with the calculation of the following vegetation indices Normalized Difference Vegetation Index (NDVI), the Soil-Adjusted Vegetation Index (SAVI), and the Normalized Difference Water Index (NDWI) were employed to acquire information on vegetation patterns, particularly since paddy rice is the predominant crop in the region. The NDVI (Rouse Jr et al., 1974) is computed from reflectance measurements in the red and near-infrared (NIR) bands

and it has been extensively utilized for monitoring vegetation, mapping land cover, and assessing ecosystems. On the other hand, SAVI is a modification of NDVI that implements a soil adjustment factor that mitigates the impact of soil reflectance on NDVI measurements (A. R. Huete, 1988), making this index more suitable with sparse vegetation or strong soil background reflectance. Lastly, the NDWI measures the water content of vegetation and the existence of water bodies using the reflectance values in the green and NIR bands (McFeeters, 1996), with positive values indicating the presence of water, and negative values the absence of this element.

4.2.3 Analysis

Change detection - Spectral distance

Change detection is an essential component of remote sensing analysis, identifying and measuring variations in land cover and land use across time (Asokan & Anitha, 2019). Spectral distance, a frequently employed technique for change detection, evaluates the distance between two or more spectral signatures and can be determined using several metrics, such as the Euclidean distance (Carvalho Júnior et al., 2011). In this study, change detection was performed using spectral distance on PF imagery, employing all the bands and indices described in the previous section. The purpose was to identify the most appropriate period during which major changes occurred. This step also assisted on selected relevant areas generating enough variability for subsequent image segmentation tasks and to avoid relying solely on subjective human observation.

To accomplish this, we calculated the Euclidean distance between two consecutive daily images taken before and after the "event" date, which was specified as a list of dates from January 3rd to April 28th, 2020, and 2021. The resulting distances were then sorted, and the three greatest distances were identified for additional visual assessment with RGB composites (Table. B.1). It is worth mentioning that only areas with evident agricultural patterns were used, and areas with permanent forest were masked. After verifying the changes, two major areas with the greatest distances were identified in the image, and smaller images were taken from these regions for use in the image segmentation task.

Image segmentation

In general terms, we can define image segmentation as the process of partitioning an image into homogeneous groups. This process in remote sensing is very well tight to classification tasks, and in this way, objects are considered as the basic unit for analysis instead of pixels, as it is the case for other popular classification tool such as Random Forest, k-means, maximum likelihood (M. Li et al., 2014).

Once the targeted areas and timing were defined during the change detection step, smaller images (chips) were generated. For this step, 50 random points were assigned within each area and for each selected time an image of 512 by 512 pixels was retrieved for each point. This resulted in 200 chips that later were uploaded for mask generation with 3 channel depth (red, green, and blue).

The labeling step, probably one of the most time-consuming steps in any machine learning project, was carried out using the tool Doodler Labeler (Buscombe & Goldstein, 2022). This

free tool allows the user to doodle over the image with a color-coded pen, based on the user pre-defined classes. This creates as semi-interactive, one-by-one, image labeling assuring the user though visual inspection if the image was correctly segmented or not (Buscombe & Goldstein, 2022). For our experiment, 6 classes were defined: road, vegetation, forest, burned, bare, and unclassified (Table. 4.1). An example for the classification highlighting the burned areas (crop residue burned before planting next crop) and bare soil is presented in figure 4.2, where the magenta color belongs to the burned areas and cyan areas is the label color for the bare soil.

Image augmentation

To increase the variability of the dataset and reduce overfitting, while also increasing the number of images, data augmentation was applied to the images. The following augmentations were applied: Random crop, this augmentation randomly crops the input image to the specified "width" and "height". Horizontal and vertical flip, these augmentations flip the input image horizontally and vertically respectively. Rotation, this augmentation rotates the input image by a random angle within the specified range of 60 to 300 degrees. Random brightness and contrast adjustments, this augmentation randomly adjusts the brightness and contrast of the input image. The brightness was adjusted by a random factor within the range of -0.2 to 0.3, and the contrast was adjusted by a random factor within the range of -0.2 to 0.2. CLAHE, grid distortion, and optical distortion, this augmentation randomly applies one of these three types of distortion to the input image. The CLAHE augmentation applies Contrast Limited Adaptive Histogram Equalization to enhance details in low-contrast areas

of the image. The Grid Distortion augmentation applies a random grid distortion to the image to introduce more variation. The Optical Distortion augmentation applies a random optical distortion to the image to simulate changes in perspective. The final augmented dataset was used to train the deep learning model.

Convolutional Neural Network – U-Net and Efficient Net

For this analysis, we employed a U-Net architecture from the segmentation models library (Iakubovskii, 2019). The U-Net is a convolutional neural network architecture that was initially developed for biomedical image segmentation but has since been adapted for different types of image segmentation tasks. A contracting path (consisting of convolutional and pooling layers that reduce the spatial resolution of the input) is followed by an expanding path in the U-Net architecture (which consists of transposed convolutions and concatenation operations that increase the spatial resolution of the output). Important for accurate segmentation, this architecture enables the network to capture both high-level and low-level features of the input image. Efficient Nets is a family of convolutional neural network architecture designed to achieve state-of-the-art image classification performance while keeping a small model size (Tan & Le, 2019). It employs a compound scaling strategy that optimizes the performance of the model by balancing the number of layers width, and resolution of the network. For this study, we used EfficientNet B2 (Tan & Le, 2019) as the backbone to initialize an encoder-decoder segmentation model in the U-Net architecture. This network can take advantage of the pre-trained features learned by EfficientNet B2 to improve the accuracy of segmentation. The overall model was then compiled with the categorical

cross-entropy loss, trained using the Adam optimizer, and early stopping was used as a regularization technique. To evaluate the performance of the model, we implemented K-fold cross-validation with k=4 folds. The model was trained on the training set, and the accuracy, Jaccard coefficient, confusion matrix, and F1 score were calculated on the test set.

4.2.4 Metrics

Several metrics commonly used in image segmentation tasks were used to evaluate the performance of the model, including accuracy, confusion matrix, F1 score and Jaccard coefficient. We can define accuracy as the ratio between the correctly classified elements (or pixels) over the total number; F1 score is calculated as the harmonic mean between precision and recall, hence these other two metrics were also calculated for the dataset. Precision can be defined as proportion of true positive predictions among all positive predictions, whereas recall is the proportion of true positive estimates among all true positive samples (Congalton, 1991). Jaccard coefficient is the metric most associated to image segmentation, which calculates the similarity between two sets of data, usually between the true label and the predicted (Wu et al., 2019).

4.3 Results

4.3.1 Change detection

The greatest distance between two consecutive daily images occurred on March 18th and March 22nd, 2021, with a distance of 118. The second-greatest distance was observed between February 13th and February 17th, 2020, with a distance of 111. The third-greatest distance was observed between February 18th and February 22nd, 2020, with a distance of 104. Finally, the fourth-greatest distance was observed between April 26th and April 30th, 2020, with a distance of 101. Each pair of images for the four dates were visually inspected to corroborate the change, and to identify the areas where larger distances were present due to a sharper change. Those areas were later transformed into polygons to extract random smaller sections for the segmentation analysis. This process was repeated for each image, brighter areas displaying a more pronounced change and darker areas where no changes were detected during that timeframe (Fig. 4.3).

4.3.2 Segmentation results

The deep learning model EfficientNetB2 was applied to the dataset consisting of 6 distinct classes in order to accomplish the image segmentation task. The model has an adequate accuracy of 0.88 and it was able to produce segmentation masks with an average Jaccard coefficient of 0.72. The Jaccard coefficient quantifies the degree to which the predicted and actual masks resemble one another (Fig. 4.4, c). A good overall performance was indicated by the fact that the average F1 score was 0.83, which is the harmonic mean of the precision

and recall scores. The matrix providing class-specific performance (Fig. 4.4, d) displays along the main diagonal those classes that have been correctly classified. According to the matrix, the model achieved a higher level of accuracy for the majority of the classes, with the exception of the road class that it was classified as an element of each of the other classes (top row).

4.4 Discussion

This work illustrates the potential of combining remote sensing and deep learning models to effectively classify land features relevant to monitoring changes in agricultural land use in Battambang province. This approach can be helpful in identifying patterns of burning and the intensity of tillage, as well as in informing management plans that encourage sustainable agricultural practices. In addition, the utilization of images with a high spatial resolution, and scene reduction by prior identification of the areas where larger changes were present, for proper image segmentation was proven effective. Although the performance of the model was high (Tan & Le, 2019), there is an opportunity for further improving the current model.

Regarding the performance related to the intersection between the generated masks and the actual values, the model achieved significantly high values compared to similar analysis found in the literature. For instance, Huerta et al. (2021) reported 0.57 while mapping urban settings, while Wu et al. (2019) testing 5 different models over 10 classes were able to obtain a wide array of values ranging from 0 to 0.96. Shirvani et al. (2023) on the other hand, sustained high values for semantic segmentation of fires using Residual Attention UNet

trained with a combination of spectral indices.

Alternative backbones were tested, with results that were significantly below the ones that were obtained; however, due to limitations in the running environment, the larger model that we could test was EfficientNetB2. Further investigation should explore the performance of more complex ones like EfficientNet B7; findings reported by (Tan & Le, 2019) highlight this difference with a substantial increase in the overall metrics.

Despite the encouraging findings and high model performance (Tan & Le, 2019), there is an opportunity for further improving the current model. For example, the sample size is considerable small, and even though image augmentation techniques are a reliable alternative, it would be preferable to have large, labeled datasets to achieve better results. In addition, this research was only carried out over a limited area featuring very specific timeframes and landscapes. It is recommended that additional heterogeneous areas be included in order to test the performance of the model across a wider range of landscapes.

In this preliminary study, spectral features were not investigated, since the band (short wave-infrared) that better correlates with burned patches and bare soil was not available in the dataset. An interesting approach was carried out by Illarionova et al. (2021), using a generative adversarial network (GAN) approach to generate the NIR band from RGB channels from high-resolution satellite imagery and use this to segment forest in the images. This could be a relevant step forward to further improve this classification and for more precisely monitoring agricultural land changes for smallholder farmers.

4.5 Conclusion

This study offers the basis for the application of remote sensing and deep learning models in recognizing patterns of burning and tillage intensity with the goal of monitoring agricultural land use that promotes sustainable agriculture practices. For image segmentation, it has been demonstrated that high spatial resolution images and scene reduction approaches, together with the detection of areas with significant changes in advance, is an effective strategy. However, there are drawbacks to this study, such as the relatively small sample size and the need to include more diverse areas to test the efficacy of the model across a broader range of landscapes. In addition, future study may investigate the use of more complex models and spectral characteristics for segmentation.

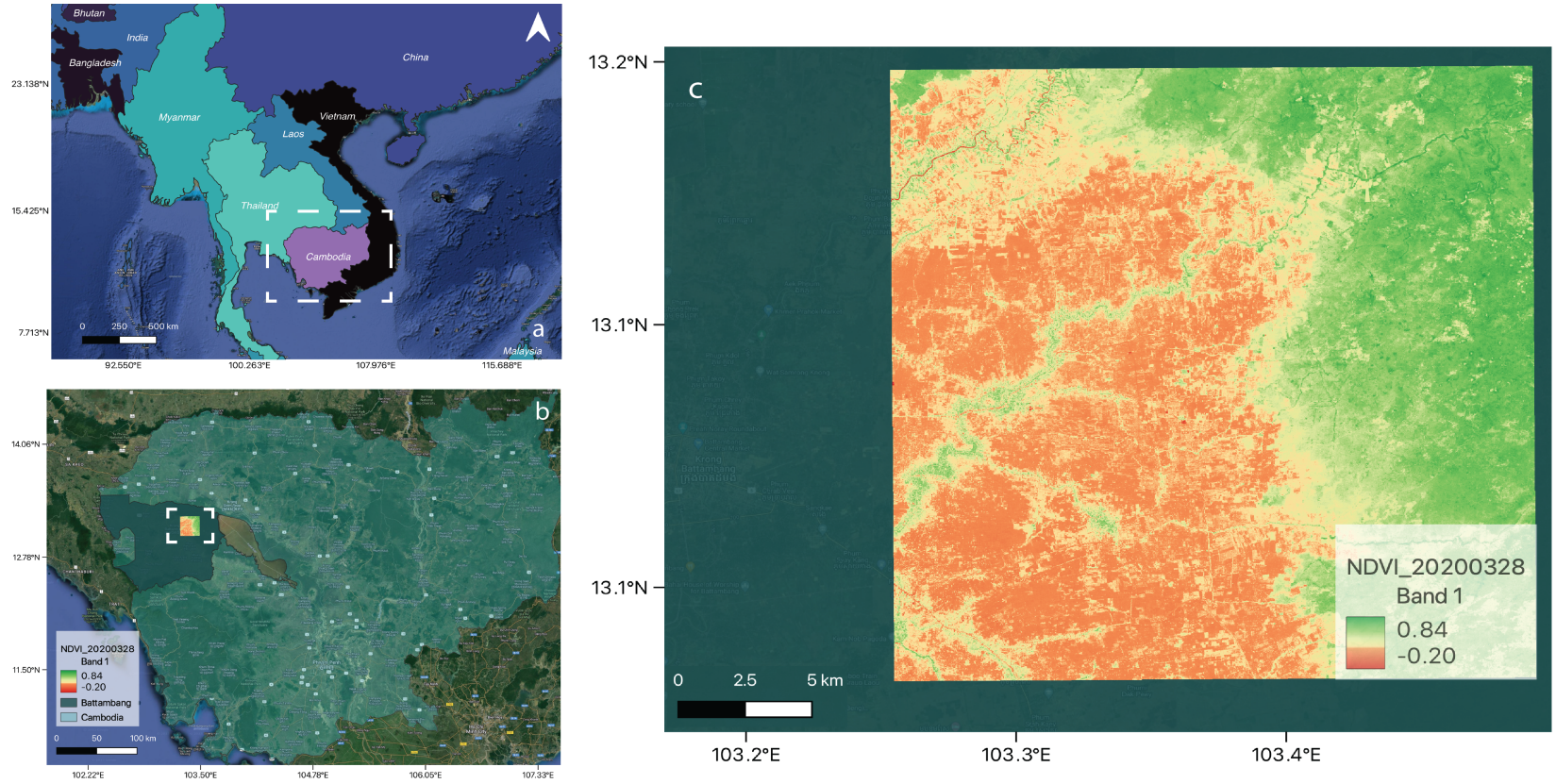

Figure 4.1: a) Cambodia (purple) and its neighboring countries. b) zoomed-in view of Cambodia and the Battambang province, shaded in a dark green color. White dotted box shows the area of study, presented in section c. c) Detailed view of the study area with the NDVI map for March 28th 2020. The NDVI map highlights the vegetation density in the area, with forest in the East and bare soil covering the rest of the image. The map is an example of the high spatial resolution captured with the use of Planet Fusion images.

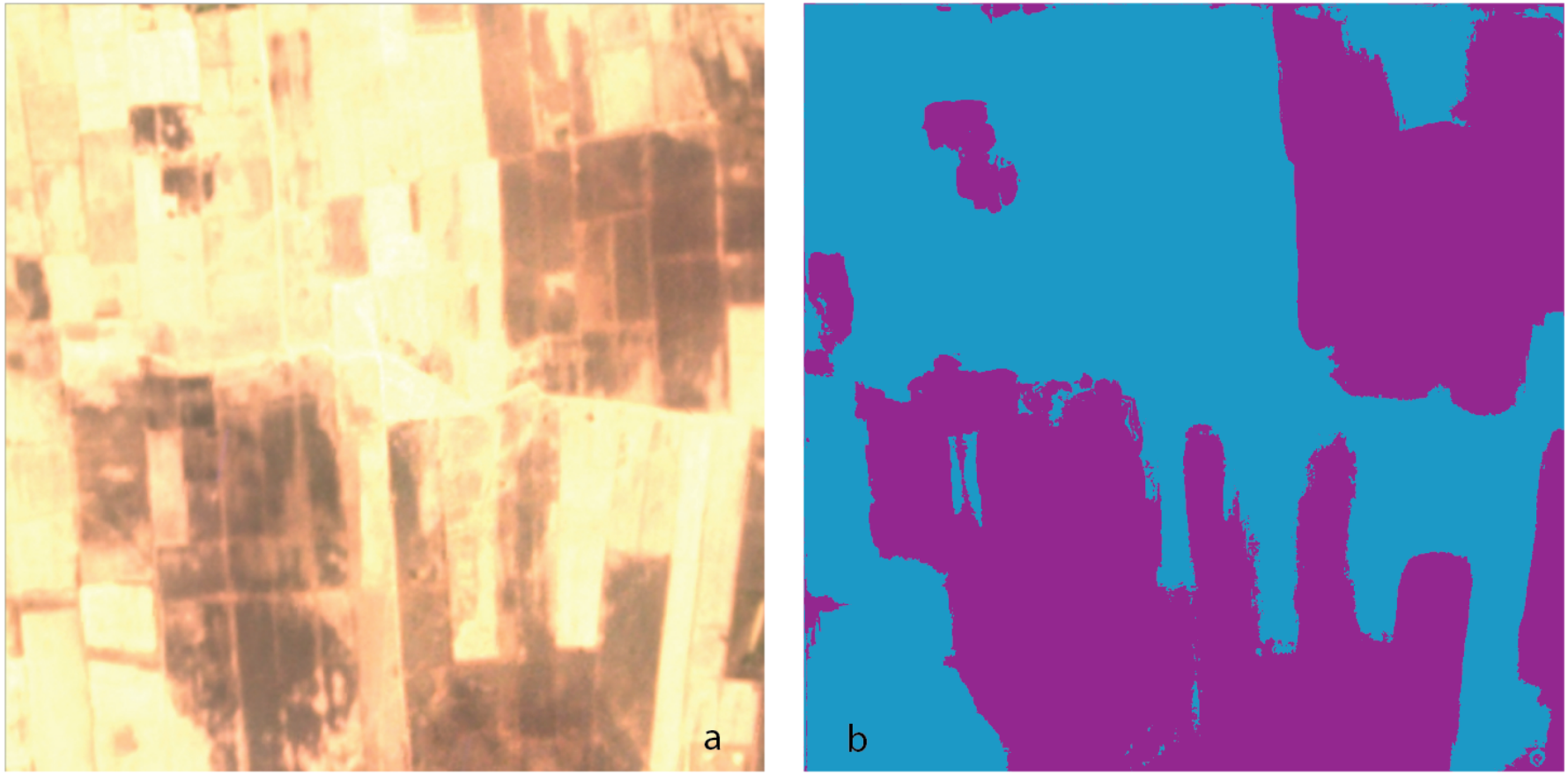

Figure 4.2: a) shows an image over the selected territory, with both burned (darker values) and bare soil (brighter values) areas. b) corresponds to the classification after doodling over the left image. Magenta maps over the burned areas while dark cyan maps over the bare soil.

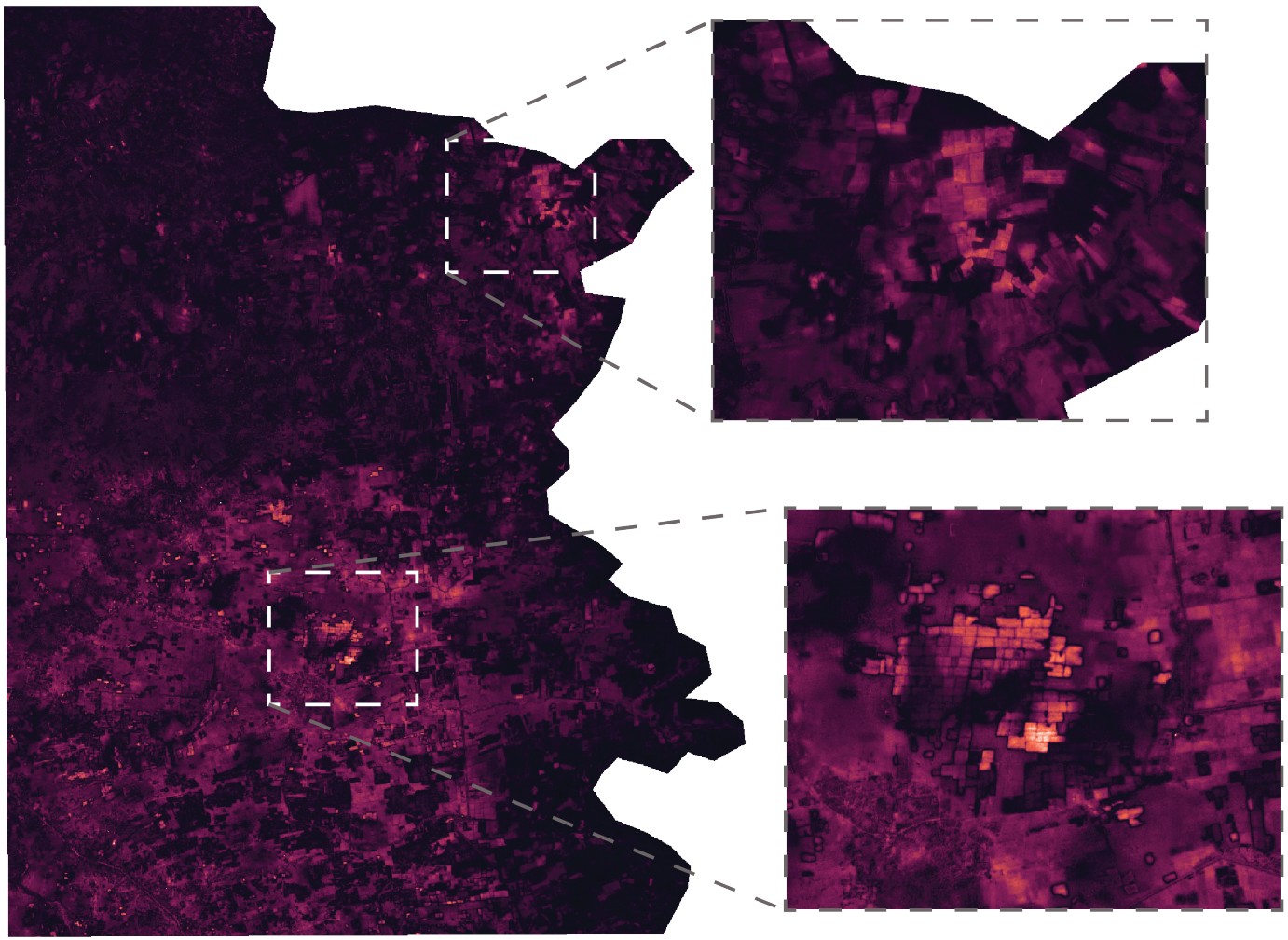

Figure 4.3: change detection over the area of study, after subtracting permanent forest. The darker the image less or no change was detected, brighter pixels indicate stronger changes. This image corresponds to the second pair February 13th and February 17th, 2020. Dotted squares highlight the areas selected for further analysis.

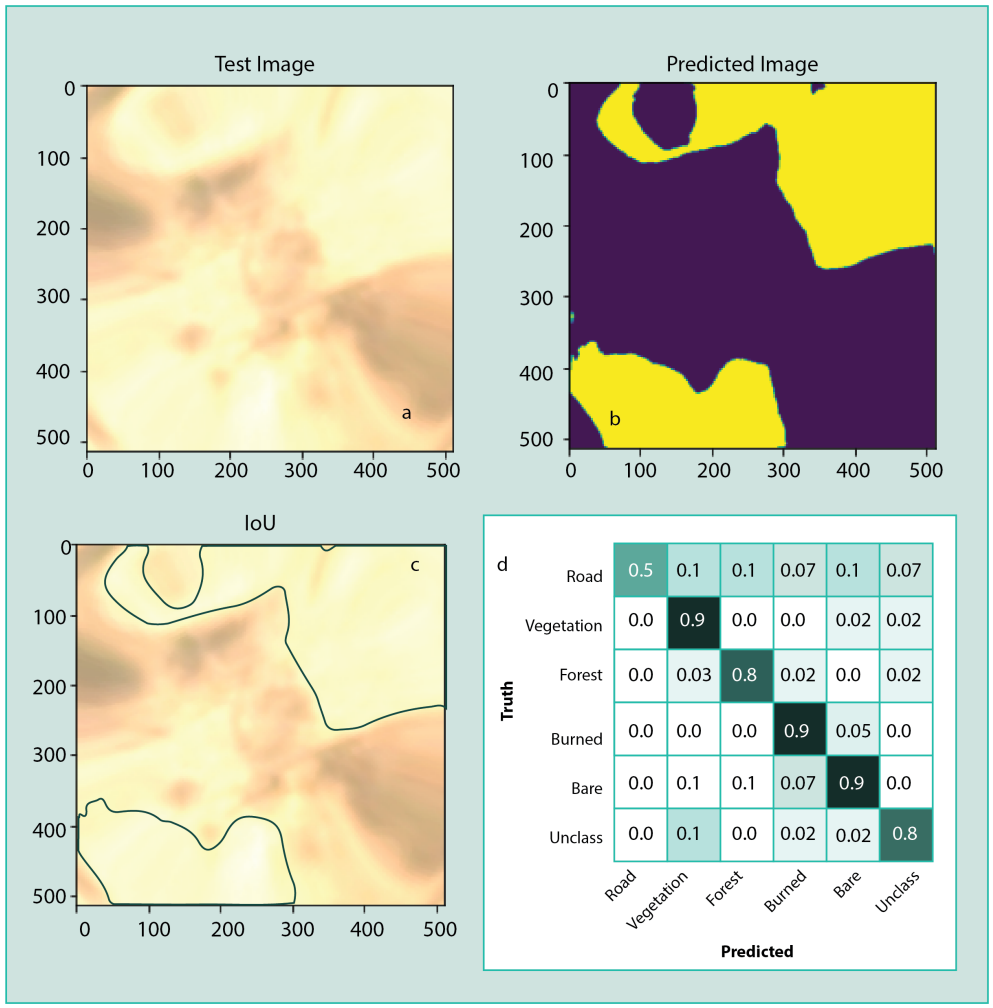

Figure 4.4: a) and b) panels showing an example image from the test batch and the corresponding prediction in purple and yellow, corresponding to two classes, burned and bare. c) shows the intersection over the union with the outline from the predicted mask. d) image shows the confusion matrix with the accuracy values for each class. Elements along the main diagonal correspond to the correctly classified ones, while those outside the diagonal are misclassified.

Table 4.1: Classes and corresponding description.

Class	**Description**
Road	Linear segments with clear bare soil
Vegetation	annual vegetation
Forest	Trees groups around fields
Burned	Dark brown areas
Bare	Bare soil
Unclassified	Elements unable to be classified within the 5 classes

Chapter 5

Limitations and future perspectives for satellite-based soil carbon monitoring

* Under review in Science of the Total Environment

Nieto, L.; Houborg, R.; Tivet, F.; Prasad, P.V.V.; Olson, B.J.S.C.; Ciampitti, I.A.

Abstract

Soil organic carbon (SOC) plays a crucial role in terrestrial C storage and ecosystem services. Agricultural management practices have the potential to increase C inputs and reduce its losses. However, uniform standard protocols for measuring, monitoring, and assessing changes using remote sensing is lacking for SOC in the scientific literature. In this discussion

paper, we present techniques for collecting and analyzing ground samples and employing remote sensing to quantify SOC, along with its limitations and future perspectives. Our analysis identified a number of key limitations to advancing the science for remotely sensed terrestrial C in croplands including i) lack of consensus in sampling depth and density, ii) the absence of a standard (or universally accepted) laboratory procedure and statistical methodology, and iii) lack of details on imagery pre-processing or information on the spectral properties of the targeted soils. Establishing standard protocols for ground-truth data collection and remote sensing approaches, as well as a knowledge of the impacts of diverse soil types, land uses, and landscapes on C assessment, are all required to enhance the accuracy and reliability of future SOC assessments.

5.1 Introduction

Terrestrial carbon (C) comprises more than 3,000 petagrams (Pg) of C (Lal, 2004), with the potential to store significantly more by absorbing CO2 from the atmosphere. The terrestrial C component influences several ecosystem services, including soil water cycle, soil structure, plant nutrition, and plant growth (Rawls et al., 2003). The largest proportion of this pool, around 1500 Pg C, exists as soil organic C (SOC) in an equilibrium between sequestration and emissions as a sink or source. Initiatives such as "4 per 1000" explore the potential that soil has to capture more atmospheric C (Minasny et al., 2017), helping to tackle global challenges such as climate change, land degradation, biodiversity loss, and food security (Lal, 2006).

Agricultural management practices have the potential to help to achieve this goal in two fundamental ways: i) by increasing C inputs, and ii) by lowering C losses, both of which can be accomplished via implementation of best management practices such as optimal fertilization rates, use of cover crops, and no-tillage agroforestry, crop – livestock integration, amongst others (Tiefenbacher et al., 2021). However, to fully understand these dynamics and use them to contribute to the improvement of C storage, we must first be able to precisely measure and monitor changes of soil C over time, along with an accurate description of land use and management practices that are among the main drivers of SOC dynamics. This discussion paper presents a brief summary on (i) the limitations associated with soil data collection and remote sensing approaches for SOC assessments; (ii) the importance and need for standard protocols for measurements and associated issues; and (iii) future directions necessary to advance towards a more homogeneous implementation of remote sensing for SOC assessment (Fig. 5.1)

5.2 Current scenario: progress and global footprint

Globally, several attempts have been made to advance the understanding of SOC dynamics from the sky (Angelopoulou et al., 2019). This becomes particularly important when looking for solutions that: (i) allow for efficient global deployment, (ii) are non-destructive, (iii) cost effective, and (iv) can be easily updated. These studies may involve optical and/or radar sensors, soil spectral libraries, and proxies such as digital elevation models or other co-variables. In recent review studies Angelopoulou et al. (2019) and Vaudour et al. (2022)

described and discussed the role of different sensors, spectral ranges, methodologies, and spatial distribution over the past decades.

From a SOC ground sampling perspective, soil data collection presents a broad range of sampling depths (0-10 cm, 0-20 cm, 0-30 cm or 0-100 cm), number of samples per unit area (0.1 samples km-2 to 2.7 samples km-2, to 201 per field), and laboratory methodologies (wet oxidation and dry combustion) (Casa et al., 2013; Castaldi et al., 2016; D. Chen et al., 2019; Vaudour et al., 2022). While these studies were distributed across the world, the geographical scope has been limited to small regions and many regions/countries remain heavily under-represented or not represented at all.

From a remote sensing standpoint, sensors mounted on planes or drones tend to be costly. Add to that the requirement for highly qualified personnel to conduct the flights, the additional time for pre-processing of the images, calibration and understanding of images, and the limited scalability of the findings. On the other hand, spaceborne platforms can retrieve data from anywhere across the world, with significantly reduced cost. Spaceborne sensors deliver data with varying spatial, temporal, and spectral resolutions, such as MODIS for regional scale studies or PlanetScope images for field-level analyses. Increasing availability of analysis-ready data (ARD) and free access via government or educational initiatives are additional benefits of satellite-based sources. Significant advancements in cloud computing infrastructure (e.g., Google Earth Engine, AWS) (Gorelick et al., 2017) have also contributed to a greater uptake of satellite-based data by enabling an increasing number of images to be analyzed/processed in very little time. The past years have seen an explosion

in satellite-based studies (Vaudour et al., 2022), coincident with the launch of Sentinel-2. Sentinel-2 represents a significant advance over previous sensors with improvements in both spatial, temporal, and spectral resolution, allowing for enhanced super-spectral analyses at 10 meters with a 5-day revisit (Drusch et al., 2012). Add to that, studies based on other multi-spectral missions such as Landsat and MODIS (Y. Peng et al., 2015; Sayão & Demattê, 2018) and hyperspectral systems such as HyMap, Hyperion, PRISMA, and CHRIS (Hbirkou et al., 2012; Hong et al., 2020). While multi-spectral sensors collect data in discrete bands (channels) across the electromagnetic spectrum, hyperspectral imaging (HSI) typically measures in 100's of continuous bands. HSI has been shown to deliver more precise retrievals given access to high spectral resolution information across specific portions of the electromagnetic spectrum (Angelopoulou et al., 2019; Chabrillat et al., 2019). The disadvantage, however, is in relation to the cost per image, spatial coverage and trained personnel to process the data. The utility of spaceborne HSI has traditionally been hindered by a lack of global monitoring missions in addition to more complicated image pre-processing tasks (atmospheric rectification, radiometric interferences) (Angelopoulou et al., 2019).

5.3 Limitations of current approaches from ground collection to remote sensing

Multiple attempts have been made on a global scale for quantifying SOC, as presented by Vaudour et al. (2022), but still large uncertainties are associated with its estimation

(Stockmann et al., 2015). For the purpose of this review and discussion, these uncertainties may be divided into three major categories: those connected with i) characterizing land use and management practices, ii) soil sample collection and analysis, and those linked to iii) remote sensing tools and image processing.

Land use and management practices are among the main drivers of SOC dynamics governing biomass-C inputs and outputs, diversity of organic inputs and their turnover rate associated with soil aggregation processes, abundance and diversity of soil biota that will shape the different SOC pools. Land use and management practices are usually not fully described in soil database or simply refer to generic terms such as conventional and best management practices with few descriptive elements that do not always allow an external reader to assess the nature of the practices and/or the historical path of the land use (Fujisaki et al., 2022).

Spatial distribution, sampling density (number of samples per unit of area), sample depth, bulk density, laboratory methodologies, and statistical methods are among the sources of uncertainty in the second group. The importance of sampling density for SOC prediction was studied by Long et al. (2018) who tested 20 different densities and concluded that sampling density affects the accuracy while presenting a correlation with the landscape, and soil orders ultimately affecting C stocks estimations (de Moraes Sá et al., 2013). This emphasizes the complexity associated with conducting research on a worldwide scale, as the required sample density would be substantially larger in order to capture the variability. Thereby making it even more challenging, costly, and laborious to collect and analyze the

data. In addition, if the propagation of error is not properly considered, scaling up field level data to regional or larger scales is likely to introduce substantial uncertainties. In this context, the spatial auto-correlation between samples plays a key role in combination with the methodology used to interpolate the values to move from sampling level to field or regional level. Few studies have been conducted covering extensive geographical areas across the globe (Vaudour et al., 2022). It is impossible for such a limited evaluation with such a low sample density to accurately portray the intricacy of this topic in its entirety. Regarding soil sample depths, (Vaudour et al., 2022) highlighted the wide range of sample depths used in literature studies in addition to the lack of a complete characterization of other relevant management conditions. The importance of sample depth and its effect on SOC stocks has been highlighted (Olson & Al-Kaisi, 2015), concluding that optimal depth should correspond with the root zone (e.g., 1-2 meters). However, this contributes to the challenge related with the cost of the analysis as well as the Intensity and laboriousness of the task.

Bulk density not only affects the flow of water and oxygen within a soil profile and the availability of nutrients but could be also source of large uncertainties when estimating SOC stocks. In overall, bulk density differed among land management uses and over time, with usually greater bulk densities under no-till cropping systems during the first years of implementation when compared with conventional management (Y. Li et al., 2020). SOC stocks should always be estimated as an equivalent mass in order to reduce uncertainties in the results analysis (Wendt & Hauser, 2013). Unfortunately, many studies do not correct soil thickness based on difference of bulk density across management practices, resulting on

one of the main reasons for high uncertainty in SOC storage estimates (Walter et al., 2016), specifically when comparing the effect of management practices (Wendt & Hauser, 2013). Attempts are also made to substitute missing bulk density data on soil database covering large scale. (Xu et al., 2016) compared several methods and reported that statistical methods involving the mean and median tend to overestimate the SOC storage by 50%, while methods using pedo-transfer functions underestimated the SOC by 8%, with a decreasing accuracy with sampling depth.

In terms of laboratory procedures for soil C determination, wet oxidation and dry combustion are the two most often stated procedures (Davis et al., 2017; Meersmans et al., 2009), despite efforts to establish the latter one as the norm. This leads to another source of uncertainty when comparing soil libraries and samples from various times and years. Wet oxidation methods fail to completely oxidize SOC, resulting in an underestimation of the SOC content. A recovery factor of 1.33 is usually found in the literature, used to estimate the total SOC, however this factor must be changed based on land use, soil type, climate, depth, and other variables (Tivet et al., 2012). Vaudour et al. (2022) advise adding this information to the metadata in order to compare varied sources and the same source across different periods in order to partially mitigate this issue. In addition, the intrinsic error associated with the laboratory (equipment and technicians) is another factor that should be included in the metadata, especially if data from different laboratories will be compared (Mountier et al., 1966).

The selection of an appropriate model of SOC estimation, has a significant impact on

the uncertainties associated with the final estimate. Various modeling techniques can be employed. Many published studies use multivariate methods, with partial least squares regression as the most prevalent option. In recent years more complex algorithms such as support vector machine, artificial neural networks, random forest, Bayesian analysis, hybrid and ensemble models have become increasingly common due to enhanced computer power and cloud computing (Tajik et al., 2020). These machine learning methods can effectively account for nonlinear data relationships and integrate supplementary variables in the learning process. Due to the fact that SOC dynamics and correlations with other variables are often non-proportional and even non-monotonic (Croft et al., 2012), nonlinear techniques become a core requirement for analyzing SOC dynamics. The proper model selection, hence, is determined by the available data, objectives, and level of understanding of underlying processes.

The remote sensing-based uncertainties (i.e., last group) are associated with (i) spatial resolution, (ii) image pre-processing, (iii) the spectral properties of the targeted soil (soil moisture, cover, soil type, etc.), and (iv) temporal analysis when assessing SOC. Insufficient spatial resolution can significantly limit the accuracy of SOC estimates via remote sensing. Remote sensing data may lack the granularity to differentiate between various SOC levels on the ground and this type of pixel mixing can result in erroneous estimations (Zhou et al., 2021). In addition, image pre-processing steps such as atmospheric correction, absorption effects, and topographic effects are frequently implied but not explicitly stated in the literature. Differences in pre-processing procedures and lack of universal ARD (analysis ready

data) standards can contribute to wrongful assessments, especially when looking at the same area at different times.

Soil moisture may alter the precision of remote sensing based SOC estimates (Jiang et al., 2016). The reflectance of the soil surface, depends on soil moisture that may indirectly control SOC retrievals via impacts on surface reflectivity (Nocita et al., 2013). Similarly, the presence of surface cover may further influence the precision of SOC estimates via its impact on surface reflectivity (Angelopoulou et al., 2019; Chabrillat et al., 2019). Variations in the spectral reflectance signature of the soil may also impact the precision of SOC given the influence of soil type, mineralogy, and SOC concentration (Dematte et al., 2009; J. Liu et al., 2018). These potentially confounding features and complexities become even more relevant when the assessments are extrapolated to larger areas, since heterogeneity is more likely to be part of the landscape.

In terms of temporal analysis (i.e., intra and inter-year) the literature is considerably scarce on the topic. For example from the perspective of the resolution and time of satellite data acquisition, Vaudour et al. (2019) found that the image date has an impact on the SOC assessment, while (Shi et al., 2022) highlighted that a multitemporal composite consistently outperformed single-date images.

To quantify changes in-field SOC and the impact of different management practices over C stocks, long term analysis is required along with diachronic assessments to improve the accuracy of SOC accumulation rates (Junior et al., 2013). The Food and Agriculture Organization of the United Nations (FAO), for example suggest 5 to 10 years before impacts

can be measured (Clara et al., 2017). In this sense the literature is still scarce even with decades old satellite missions, such as Landsat, being available. In addition, the conversion rate of biomass inputs to SOC can be highly variable based on management practices. For example, Fujisaki et al. (2018) assessing 48 tropical studies observed a mean conversion rate of C inputs to SOC of 8.2% with a range from - 7.3 to 35.6% with biomass-C inputs being the main predictor of SOC accumulation rate. The conversion rate of biomass to C could depend on factors such as no-tillage, use of cover crops, rotation, restitution of biomass inputs, and years of implementation. This highlights the need to have a clear description of agricultural practices along with soil attributes.

5.4 Future perspectives

Fujisaki et al. (2022) called for a harmonized thesaurus to give genericity to the terms describing land use and management practices across large soil database for the evaluation of soil carbon storage. The authors proposed a hierarchical tree based on three main management practices (i.e., annual and perennial croplands, grasslands, and forest and tree plantations) subdivided on land managers' point of view including plant, biomass, amendments management but also erosion, fire, and land clearing management. This being one example, the absence of a uniform and well-established ground-based protocol for SOC estimation is one of the major obstacles in the field. This is owing to a wide range of sampling strategies (such as sample depth, density, and bulk density) and laboratory tests (such as dry combustion and wet oxidation), which have led to inconsistent methodologies for SOC measurement

and evaluation. This prevents us from accounting for unbiased estimations and from establishing a foundation for more knowledgeable contributions of each biome, ecosystem, crop, and activity to the overall C system. A first step to help accelerate and upscale the use of remote sensing tools to address C dynamics, should lean towards a deeper understanding of the effect of sampling density to cover the variability and diversity of soil type at catena scale, sampling depth, soil thickness corrected by bulk density, sampling dates, historical information, management practices, and laboratory methods for different soil types around the world (Fig. 5.1).

The increasing presence and potential democratization of hyperspectral systems such as the Environmental Mapping and Analysis Program (EnMAP) launched on April 2022 and emerging HIS systems such as Copernicus Hyperspectral Imaging Mission for the Environment (CHIME, collaboration between European Space Agency and US National Aeronautics and Space Administration, ESA-NASA), Tanager (Planet), and TD-2 (Pixxel) embrace the opportunity for significant science innovations. These systems have the potential to deliver higher resolution and accuracy in SOC evaluations, playing an essential role in enhancing our comprehension of the effects that agricultural management practices have on the terrestrial storage of carbon. Still, uncertainties presented previously, associated with spatial resolution, image pre-processing, and temporal assessment hold valid for this technology as well.

SOC estimation is a complex and multidimensional issue, affected by the various factors mentioned above. To achieve more accurate and consistent SOC estimations, it is necessary

to establish a uniform and well-established protocol, that accounts for the variability and diversity of scenarios around the world. A concerted effort by the scientific community to tackle these challenges and develop a more comprehensive approach is crucial for achieving sustainable goals.

5.5 Conclusions

This study discussed the current challenges, limitations, and next steps related to advancing the use of remote sensing to monitor terrestrial C in croplands. The major limitations for expanding this discipline are linked to: (i) lack of consensus in sampling density and depth, (ii) lack of proper bulk density assessment, (iii) lack of standard laboratory procedures, (v) lack of details on imagery pre-processing steps, (vi) missing information on the spectral properties of the targeted soil, and (vii) lack of multi-temporal approaches. Our ability to accurately quantify SOC and its dynamics in space and time comes down to how well can address and resolve these obstacles.

In summary, this study highlights the need for standardized protocols in terms of ground-truth data collection and remote sensing approaches, along with an improved understanding of the interlinked impacts of differences in soil types, land use, and landscape heterogeneity on C assessment.

Figure 5.1: Current limitations and future perspectives for soil organic carbon (SOC) assessments from soil sample collection and analysis to remote sensing.

Chapter 6

Conclusions

The advancement of remote sensing technology has led to a revolution in the monitoring and management of agricultural systems in a remarkably brief span of 50 years. With thousands of satellites continuously orbiting the earth and collecting invaluable insights on multiple traits, the potential of this technology to facilitate agricultural practices, enhance food security, and promote sustainable agriculture knows no bounds. The present dissertation delved into the deployment of various satellite imagery on the optical range to undertake several classification tasks in the realm of plant-soil-management systems, with the dual aim of understanding current practices and identifying literature gaps that could contribute to advancing the state of the field.

6.1 Best combination (Chapter 2)

In an effort to efficiently report maize phenology, the second chapter investigates the use of Landsat 8 imagery and weather data over southwest Kansas. The chapter's aim is to identify the optimal combination of variables that would culminate in the most reliable classifier, availing ground-truth data from 2013 to 2018. This chapter focuses on two pivotal objectives: (i) model training and validation, which seeks to fathom how distinct variables influence model performance and to identify the best combination of spectral features, weather parameters, geolocation, and ground truth data, consequently yielding the most accurate model across years; and (ii) model testing for each phenology class, temporal transferability, and spatial transferability.

The analysis suggested by this study demonstrates that the combination of NDVI, EVI, GCVI, NDWI, GVMI, and vapor pressure deficit is the most efficacious ensemble, resulting in a high accuracy that surpasses 0.94, accompanied by comparable values for temporal transferability. Albeit the findings are encouraging, there exists potential for improvement. For example, the phenology classes used in the study do not correspond to the phenological stages typically employed in agronomy to describe crop development. Due to the limited availability of Landsat 8 images during the maize growing season in the region, entire stages of development were not captured during the physiological stage due to a lack of images. The limitations of this experiment culminated in the formulation of a hypothesis, which became the primary focus of chapter three, wherein temporal resolution took center stage.

6.2 Resolution matters (Chapter 3)

Chapter three aimed to test the hypothesis that emerged from chapter two, specifically whether satellite temporal and spatial resolution plays a significant role in classifying maize phenology. To carry out this experiment, daily 3-meter images from Planet Fusion were compared against Sentinel-2 for the year 2017 using more than 20 phenology stages, which are terms that are typically used in the field of agronomy to describe the development of crops.

The findings of this experiment unequivocally demonstrated the importance of high-cadence imagery for accurately capturing traits that change rapidly during the growing season, such as crop development. The model trained with Planet Fusion spectral data achieved an overall accuracy of 0.94 for all classes, while the Sentinel-2 counterpart was able to register an accuracy of merely 0.60 for the same task.

After identifying the significance of temporal resolution, the next pertinent question focuses on the potential significance of spatial resolution. To thoroughly scrutinize this matter, it becomes apparent that the Kansas fields, with an average size of 50 hectares do not present a sufficiently rigorous testing ground. Consequently, chapter four explored this feature in small-holder settings where field dimensions and landscape complexity exceed those examined in chapters 2 and 3.

6.3 Finding patterns (Chapter 4)

Chapter four endeavors to assess the spatial resolution of satellite imagery and its capacity to differentiate between two distinct agricultural practices in Cambodia: straw burning and bare soil after tillage.

The agricultural landscape of Cambodia's Battambang province is characterized by small fields, with an average size of 0.86 hectares, with two distinct seasons: the dry season from November to April and the wet season from May to October. This distinctive environment presented an optimal opportunity to evaluate the spatial and temporal resolution of Planet Fusion imagery, combined with deep learning techniques for image segmentation.

The results of this experiment evinced a similarity score of 0.72 for image segmentation and an accuracy score of 0.90 for the targeted practices during the selected time period. The availability of high spatial resolution images (3 meters) and ground labels for analysis bolstered the relevance of the achieved similarity score, attesting to the potential of these images to appraise land features in small-holder regions such as Cambodia where in-situ ground-truth data collection may pose challenges.

The impact of this chapter is aimed towards facilitating the implementation of sustainable practices that can aid farmers in preserving soil properties and fertility, thereby augmenting the sustainability and productivity of agriculture in the region.

The subsequent question concerns whether the scientific community's efforts can be seamlessly integrated and juxtaposed within the existing literature. In order to shed light on this conundrum, chapter five focuses on a topic that has gained popularity in recent years: the

use of remote sensing data for carbon assessment.

6.4 Why do we need protocols? (Chapter 5)

Using remote sensing tools, the fifth chapter, which is presented as a discussion paper, highlights the significance of standardized protocols in soil trait assessments, particularly soil organic carbon. The chapter draws upon a comprehensive review of the literature, revealing prominent discrepancies. The absence of a consensus regarding the sampling depth at which soil samples are collected for ground-truth data is one of the many causes for concern. Second, while the field of soil science has an established procedure for laboratory analysis, the literature on remote sensing employs a variety of methods for analyzing soil samples in order to compare them to remotely sensed features. In addition, the level of detail in image preprocessing and the methodologies used to extrapolate results from the field to the regional level are not adequately reported. Furthermore, large areas of the globe are underrepresented in soil trait assessments, posing a significant challenge for estimations. The absence of a protocol makes it impossible to compare the results of different studies. These findings underscore the pressing need for further research to improve our comprehension of the crucial role of remote sensing to study soil properties, as a proxy for food production and sustainable intensification, as well as the importance of implementing protocols that can be easily understood by scientists from various disciplines, promoting inter and multi-disciplinary efforts.

6.5 Looking ahead

The field of remote sensing is in a constant state of evolution, with commercial firms and governments around the world launching new missions each year. These endeavors have resulted in the development of novel sensors and innovative means of integrating pre-existing data, rendering this industry one of the most dynamic in the present era. However, the primary challenge remains in leveraging these advancements to amplify our food systems in a sustainable and cost-effective manner, thereby mitigating waste and safeguarding our valuable resources for future generations. Such an endeavor necessitates a concerted and inter and multi-disciplinary effort, requiring knowledge from diverse fields such as agricultural science, geography, statistics, among others. Still there is a long way ahead of us, but incentives are placed on the right track.

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131.

Abatzoglou, J. T., Rupp, D. E., & Mote, P. W. (2014). Seasonal climate variability and change in the pacific northwest of the united states. *Journal of Climate*, *27*(5), 2125–2142.

Adams, V. M., Iacona, G. D., & Possingham, H. P. (2019). Weighing the benefits of expanding protected areas versus managing existing ones. *Nature Sustainability*, *2*(5), 404–411.

Angelopoulou, T., Tziolas, N., Balafoutis, A., Zalidis, G., & Bochtis, D. (2019). Remote sensing techniques for soil organic carbon estimation: A review. *Remote Sensing*, *11*(6), 676.

Arun, P., Sadeh, R., Avneri, A., Tubul, Y., Camino, C., Buddhiraju, K. M., ... others (2022). Multimodal earth observation data fusion: Graph-based approach in shared latent space. *Information Fusion*, *78*, 20–39.

Asokan, A., & Anitha, J. (2019). Change detection techniques for remote sensing applications: A survey. *Earth Science Informatics*, *12*, 143–160.

Azzari, G., Jain, M., & Lobell, D. B. (2017). Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sensing of Environment*, *202*, 129–141.

Bai, J., Chen, X., Dobermann, A., Yang, H., Cassman, K. G., & Zhang, F. (2010). Evaluation of nasa satellite-and model-derived weather data for simulation of maize yield potential in china. *Agronomy Journal*, *102*(1), 9–16.

Bandaru, V., Yaramasu, R., Koutilya, P., He, J., Fernando, S., Sahajpal, R., ... Justice, C. (2020). Phenocrop: An integrated satellite-based framework to estimate physiological growth stages of corn and soybeans. *International Journal of Applied Earth Observation and Geoinformation*, *92*, 102188.

Basso, B., & Ritchie, J. T. (2018). Evapotranspiration in high-yielding maize and under increased vapor pressure deficit in the us midwest. *Agricultural & Environmental Letters*, *3*(1), 170039.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, *114*, 24–31.

Bhowan, U., Zhang, M., & Johnston, M. (2009). Genetic programming for image classification with unbalanced data. In *2009 24th international conference image and vision computing new zealand* (pp. 316–321). doi: 10.1007/978-3-642-12148-7_1

Bolton, D. K., & Friedl, M. A. (2013). Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and forest meteorology*, *173*, 74–84.

Branco, P., Torgo, L., & Ribeiro, R. P. (2016). A survey of predictive modeling on imbalanced domains. *ACM computing surveys (CSUR)*, *49*(2), 1–50.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Brinkhoff, J., Houborg, R., & Dunn, B. W. (2022). Rice ponding date detection in australia using sentinel-2 and planet fusion imagery. *Agricultural Water Management*, *273*, 107907.

Buscombe, D., & Goldstein, E. B. (2022). A reproducible and reusable pipeline for segmentation of geoscientific imagery. *Earth and Space Science*, *9*(9), e2022EA002332.

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., ... others (2019). Integrating satellite and climate data to predict wheat yield in australia using machine learning approaches. *Agricultural and forest meteorology*, *274*, 144–159.

Cai, Y., Guan, K., Nafziger, E., Chowdhary, G., Peng, B., Jin, Z., ... Wang, S. (2019). Detecting in-season crop nitrogen stress of corn for field trials using uav-and cubesat-based multispectral sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(12), 5153–5166.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., & Li, Z. (2018). A high-performance and in-season classification system of field-level crop types using time-series landsat data and a machine learning approach. *Remote sensing of environment*, *210*, 35–47.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*(6), 679–698.

Carletto, C., Gourlay, S., & Winters, P. (2015). From guesstimates to gpstimates: Land area measurement and implications for agricultural analysis. *Journal of African Economies*, *24*(5), 593–628.

Carvalho Júnior, O. A., Guimarães, R. F., Gillespie, A. R., Silva, N. C., & Gomes, R. A. (2011). A new approach to change vector analysis using distance and similarity measures. *Remote Sensing*, *3*(11), 2473–2493.

Casa, R., Castaldi, F., Pascucci, S., Basso, B., & Pignatti, S. (2013). Geophysical and hyperspectral data fusion techniques for in-field estimation of soil properties. *Vadose Zone Journal*, *12*(4).

Castaldi, F., Palombo, A., Santini, F., Pascucci, S., Pignatti, S., & Casa, R. (2016). Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. *Remote Sensing of Environment*, *179*, 54–65.

Ceccato, P., Gobron, N., Flasse, S., Pinty, B., & Tarantola, S. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach. *Remote sensing of environment*, *82*(2-3), 188–197.

Chabrillat, S., Ben-Dor, E., Cierniewski, J., Gomez, C., Schmid, T., & van Wesemael, B. (2019). Imaging spectroscopy for soil mapping and monitoring. *Surveys in Geophysics*, *40*, 361–399.

Chen, D., Chang, N., Xiao, J., Zhou, Q., & Wu, W. (2019). Mapping dynamics of soil organic matter in croplands with modis data and machine learning algorithms. *Science of the Total Environment*, *669*, 844–855.

Chen, L., An, G., Xin, M., & Lai, G. (2021). Crop classification based on image segmentation and phenological similarity using sar imagery. In *2021 ieee international geoscience and remote sensing symposium igarss* (pp. 5925–5928).

Ciampitti, I. A., Elmore, R. W., & Lauer, J. (2016). *Corn growth and development* (MF3305). Kansas State Research and Extension.

Clara, L., Fatma, R., Viridiana, A., & Liesl, W. (2017). Soil organic carbon: the hidden potential. *Food and Agriculture Organisation (FAO): Rome, Italy*.

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., ... Justice, C. (2018). The harmonized landsat and sentinel-2 surface reflectance data set. *Remote sensing of environment*, *219*, 145–161.

Clevers, J. (1997). A simplified approach for yield prediction of sugar beet based on optical remote sensing data. *Remote sensing of Environment*, *61*(2), 221–228.

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

Crocker, R. L. (1962). *Distribution of soil types in cambodia.* https://data.opendevelopmentcambodia.net/dataset/distribution-soil-type-in-cambodia?type=dataset. (Retrieved March 11, 2023)

Croft, H., Kuhn, N., & Anderson, K. (2012). On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. *Catena*, *94*, 64–74.

Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. In *2015 ieee symposium series on computational intelligence* (pp. 159–166).

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., ... Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *International Journal of Climatology: a Journal of the Royal Meteorological Society*, *28*(15), 2031–2064.

Davis, M. R., Alves, B. J., Karlen, D. L., Kline, K. L., Galdos, M., & Abulebdeh, D. (2017). Review of soil organic carbon measurement protocols: A us and brazil comparison and recommendation. *Sustainability*, *10*(1), 53.

Deines, J. M., Kendall, A. D., Crowley, M. A., Rapp, J., Cardille, J. A., & Hyndman, D. W. (2019). Mapping three decades of annual irrigation across the us high plains aquifer using landsat and google earth engine. *Remote Sensing of Environment*, *233*, 111400.

Dematte, J., R Fioriob, P., & Ben-Dorc, E. (2009). Estimation of soil properties by orbital and laboratory reflectance means and its relation with soil classification. *The Open Remote Sensing Journal*, *2*(1).

de Moraes Sá, J. C., Bürkner dos Santos, J., Lal, R., De Moraes, A., Tivet, F., Machado Sá, M. F., ... others (2013). Soil-specific inventories of landscape carbon and nitrogen stocks under no-till and native vegetation to estimate carbon offset in a subtropical ecosystem. *Soil Science Society of America Journal*, *77*(6), 2094–2110.

Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., ... Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, *10*(6), 291–297.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Multiple classifier systems: First international workshop, mcs 2000 cagliari, italy, june 21–23, 2000 proceedings* (Vol. 1, pp. 1–15). Springer Berlin Heidelberg.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... others (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46.

dos Santos Luciano, A. C., Picoli, M. C. A., Duft, D. G., Rocha, J. V., Leal, M. R. L. V., & Le Maire, G. (2021). Empirical model for forecasting sugarcane yield on a local scale in brazil using landsat imagery and random forest algorithm. *Computers and Electronics in Agriculture*, *184*, 106063.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... others (2012). Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote sensing of Environment*, *120*, 25–36.

Dubovik, O., Schuster, G. L., Xu, F., Hu, Y., Bösch, H., Landgraf, J., & Li, Z. (2021). Grand challenges in satellite remote sensing. *Frontiers in Remote Sensing*, *2*, 619818. doi: 10.3389/frsen.2021.619818

Dubovyk, O. (2017). The role of remote sensing in land degradation assessments: opportunities and challenges. *European Journal of Remote Sensing*, *50*(1), 601–613.

Duncan, J. (n.d.). A, dash, j., and atkinson, pm: The potential of satellite-observed crop phenology to enhance yield gap assessments in smallholder landscapes, front. *Environ. Sci*, *3*(10.3389).

European Space Agency. (n.d.). *Earth online: Third party missions.* ESA. Retrieved from https://earth.esa.int/eogateway/category/Missions/mission_programme=third+party+missions&sortby=RELEVANCE (Retrieved March 7, 2023)

FAOa. (n.d.). *What is conservation agriculture? — conservation agriculture — food and agriculture organization of the united nations.* Retrieved from https://www.fao.org/conservation-agriculture/overview/what-is-conservation-agriculture/en/ (Retrieved March 3, 2023)

FAOb. (n.d.). *Data sources: Country details - cambodia.* https://www.fao.org/family-farming/data-sources/dataportrait/country-details/en/?cnt=KHM. (Retrieved March 11, 2023)

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research*, *15*(1), 3133–3181.

Fletcher, A. L., Sinclair, T. R., & Allen Jr, L. H. (2007). Transpiration responses to vapor pressure deficit in well watered 'slow-wilting'and commercial soybean. *Environmental and Experimental Botany*, *61*(2), 145–151.

Frantz, D. (2019). Force—landsat+ sentinel-2 analysis ready data and beyond. *Remote Sensing*, *11*(9), 1124.

Fujisaki, K., Chevallier, T., Bispo, A., Laurent, J.-B., Thévenin, F., Chapuis-Lardy, L., ... others (2022). Semantics about soil organic carbon storage: Data4c+, a comprehensive thesaurus and classification of management practices in agriculture and forestry. *EGUsphere*, 1–41.

Fujisaki, K., Chevallier, T., Chapuis-Lardy, L., Albrecht, A., Razafimbelo, T., Masse, D., ... Chotte, J.-L. (2018). Soil carbon stock changes in tropical croplands are mainly driven by carbon inputs: A synthesis. *Agriculture, Ecosystems & Environment*, *259*, 147–158.

Gao, B.-C. (1995). Normalized difference water index for remote sensing of vegetation liquid water from space. In *Imaging spectrometry* (Vol. 2480, pp. 225–236).

Gao, F., Anderson, M., Daughtry, C., Karnieli, A., Hively, D., & Kustas, W. (2020). A within-season approach for detecting early growth stages in corn and soybean using high temporal and spatial resolution imagery. *Remote Sensing of Environment*, *242*, 111752.

Gao, F., Anderson, M. C., Zhang, X., Yang, Z., Alfieri, J. G., Kustas, W. P., ... Prueger, J. H. (2017). Toward mapping crop progress at field scales through fusion of landsat and modis imagery. *Remote Sensing of Environment*, *188*, 9–25.

Gao, F., & Zhang, X. (2021). Mapping crop phenology in near real-time using satellite remote sensing: Challenges and opportunities. *Journal of Remote Sensing*, *2021*.

Gerhards, M., Schlerf, M., Mallick, K., & Udelhoven, T. (2019). Challenges and future perspectives of multi-/hyperspectral thermal infrared remote sensing for crop water-stress detection: A review. *Remote Sensing*, *11*(10), 1240.

Ghamghami, M., Ghahreman, N., Irannejad, P., & Pezeshk, H. (2020). A parametric empirical bayes (peb) approach for estimating maize progress percentage at field scale. *Agricultural and Forest Meteorology*, *281*, 107829.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2004). Random forest classification of multisource remote sensing and geographic data. In *Igarss 2004. 2004 ieee international geoscience and remote sensing symposium* (Vol. 2, pp. 1049–1052).

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern recognition letters*, *27*(4), 294–300.

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

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

Goodin, D., Mitchell, J., Knapp, M., & Bivens, R. (1995). Climate and weather atlas of kansas: an introduction. educational ser. 12. *Kansas Geol. Survey. Lawrence*.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, 18–27.

Gourlay, S., Kilic, T., & Lobell, D. (2017). Could the debate be over? errors in farmer-reported production and their implications for the inverse scale-productivity relationship in uganda. *Errors in Farmer-Reported Production and Their Implications for the Inverse Scale-Productivity Relationship in Uganda (September 12, 2017). World Bank Policy Research Working Paper*(8192).

Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *Advances in information retrieval: 27th european conference on ir research, ecir 2005, santiago de compostela, spain, march 21-23, 2005. proceedings 27* (pp. 345–359).

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.-C., ... others (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors? *Remote sensing of environment*, *244*, 111806.

Hanks, E. M., Hooten, M. B., & Baker, F. A. (2011). Reconciling multiple data sources to improve accuracy of large-scale prediction of forest disease incidence. *Ecological applications*, *21*(4), 1173–1188.

Hao, P., Zhan, Y., Wang, L., Niu, Z., & Shakir, M. (2015). Feature selection of time series modis data for early crop classification using random forest: A case study in kansas, usa. *Remote Sensing*, *7*(5), 5347–5369.

Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *science*, *344*(6186), 862–867.

Hbirkou, C., Pätzold, S., Mahlein, A.-K., & Welp, G. (2012). Airborne hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale. *Geoderma*, *175*, 21–28.

Henebry, G. M., & de Beurs, K. M. (2013). Remote sensing of land surface phenology: A prospectus. *Phenology: An integrative environmental science*, 385–411.

Hobbs, P. R., Sayre, K., & Gupta, R. (2008). The role of conservation agriculture in sustainable agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 543–555.

Hoens, T. R., Qian, Q., Chawla, N. V., & Zhou, Z.-H. (2012). Building decision trees for the multi-class imbalance problem. In *Advances in knowledge discovery and data mining: 16th pacific-asia conference, pakdd 2012, kuala lumpur, malaysia, may 29-june 1, 2012, proceedings, part i 16* (pp. 122–134).

Homma, K., Horie, T., Shiraiwa, T., Supapoj, N., Matsumoto, N., & Kabaki, N. (2003). Toposequential variation in soil fertility and rice productivity of rainfed lowland paddy

fields in mini-watershed (nong) in northeast thailand. *Plant Production Science*, *6*(2), 147–153.

Hong, Y., Guo, L., Chen, S., Linderman, M., Mouazen, A. M., Yu, L., ... others (2020). Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon: Effects of fractional-order derivative and optimal band combination algorithm. *Geoderma*, *365*, 114228.

Hooten, M. B., Wikle, C. K., Dorazio, R. M., & Royle, J. A. (2007). Hierarchical spatiotemporal matrix models for characterizing invasions. *Biometrics*, *63*(2), 558–567.

Houborg, R., Fisher, J. B., & Skidmore, A. K. (2015). *Advances in remote sensing of vegetation function and traits* (Vol. 43). Elsevier.

Houborg, R., & McCabe, M. F. (2018a). A cubesat enabled spatio-temporal enhancement method (cestem) utilizing planet, landsat and modis data. *Remote Sensing of Environment*, *209*, 211–226.

Houborg, R., & McCabe, M. F. (2018b). Daily retrieval of ndvi and lai at 3 m resolution via the fusion of cubesat, landsat, and modis data. *Remote Sensing*, *10*(6), 890.

Hsiao, J., Swann, A. L., & Kim, S.-H. (2019). Maize yield under a changing climate: The hidden role of vapor pressure deficit. *Agricultural and Forest Meteorology*, *279*, 107692.

Huerta, R. E., Yépez, F. D., Lozano-García, D. F., Guerra Cobián, V. H., Ferriño Fierro, A. L., de León Gómez, H., ... Vargas-Martínez, A. (2021). Mapping urban green spaces

at the metropolitan level using very high resolution satellite imagery and deep learning techniques for semantic segmentation. *Remote Sensing*, *13*(11), 2031.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote sensing of environment*, *83*(1-2), 195–213.

Huete, A., Justice, C., & Liu, H. (1994). Develpmeclassification and soil indices for modis-eos. *Remote Sensing of Environment*, *49*, 224–234.

Huete, A. R. (1988). A soil-adjusted vegetation index (savi). *Remote sensing of environment*, *25*(3), 295–309.

Iakubovskii, P. (2019). *Segmentation models.* https://github.com/qubvel/segmentation_models.

Illarionova, S., Shadrin, D., Trekin, A., Ignatiev, V., & Oseledets, I. (2021). Generation of the nir spectral band for satellite images with convolutional neural networks. *Sensors*, *21*(16), 5646.

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent data analysis*, *6*(5), 429–449.

Ji, S., Zhang, C., Xu, A., Shi, Y., & Duan, Y. (2018). 3d convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sensing*, *10*(1), 75.

Jiang, Q., Chen, Y., Guo, L., Fei, T., & Qi, K. (2016). Estimating soil organic carbon of cropland soil at different levels of soil moisture using vis-nir spectroscopy. *Remote Sensing*, *8*(9), 755.

Jin, S., Su, Y., Gao, S., Hu, T., Liu, J., & Guo, Q. (2018). The transferability of random forest in canopy height estimation from multi-source remote sensing data. *Remote Sensing*, *10*(8), 1183.

Jin, Z., Azzari, G., Burke, M., Aston, S., & Lobell, D. B. (2017). Mapping smallholder yield heterogeneity at multiple scales in eastern africa. *Remote Sensing*, *9*(9), 931.

Jin, Z., Azzari, G., & Lobell, D. B. (2017). Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches. *Agricultural and forest meteorology*, *247*, 207–220.

Joshi, V. R., Kazula, M. J., Coulter, J. A., Naeve, S. L., & Garcia y Garcia, A. (2021). In-season weather data provide reliable yield estimates of maize and soybean in the us central corn belt. *International Journal of Biometeorology*, *65*(4), 489–502.

Junior, C. C., Corbeels, M., Bernoux, M., Piccolo, M. d. C., Neto, M. S., Feigl, B. J., ... Lal, R. (2013). Assessing soil carbon storage rates under no-tillage: comparing the synchronic and diachronic approaches. *Soil and Tillage Research*, *134*, 207–212.

Kennedy, R. E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W. B., & Healey, S. (2018). Implementation of the landtrendr algorithm on google earth engine. *Remote Sensing*, *10*(5), 691.

Kimm, H., Guan, K., Jiang, C., Peng, B., Gentry, L. F., Wilkin, S. C., ... others (2020). Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the us corn belt using planet labs cubesat and stair fusion data. *Remote Sensing of Environment*, *239*, 111615.

Kitchenham, B. (1998). A procedure for analyzing unbalanced datasets. *IEEE transactions on Software Engineering*, *24*(4), 278–301.

Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote sensing of environment*, *1*(3), 155–159.

Kumar, L., & Mutanga, O. (2018). Google earth engine applications since inception: Usage, trends, and potential. *Remote Sensing*, *10*(10), 1509.

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *science*, *304*(5677), 1623–1627.

Lal, R. (2006). Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land degradation & development*, *17*(2), 197–209.

Lal, R. (2015). Sequestering carbon and increasing productivity by conservation agriculture. *Journal of soil and water conservation*, *70*(3), 55A–62A.

Lanning-Rush, J. L. (2016). *Irrigation water use in kansas, 2013* (Tech. Rep.). US Geological Survey.

Lemenih, M., Karltun, E., & Olsson, M. (2005). Assessing soil chemical and physical property responses to deforestation and subsequent cultivation in smallholders farming system in ethiopia. *Agriculture, Ecosystems & Environment*, *105*(1-2), 373–386.

Leslie, C. R., Servina, L. O., & Miller, H. M. (2017). *Landsat and agriculture: Case studies on the uses and benefits of landsat imagery in agricultural monitoring and production.* US Department of the Interior, US Geological Survey Reston, VA, USA.

Li, J., & Roy, D. P. (2017). A global analysis of sentinel-2a, sentinel-2b and landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sensing*, *9*(9), 902.

Li, M., Zang, S., Zhang, B., Li, S., & Wu, C. (2014). A review of remote sensing image classification techniques: The role of spatio-contextual information. *European Journal of Remote Sensing*, *47*(1), 389–411.

Li, Y., Li, Z., Cui, S., & Zhang, Q. (2020). Trade-off between soil ph, bulk density and other soil physical properties under global no-tillage agriculture. *Geoderma*, *361*, 114099.

Liang, L., Schwartz, M. D., & Fei, S. (2011). Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. *Remote Sensing of Environment*, *115*(1), 143–157.

Liao, C., Wang, J., Dong, T., Shang, J., Liu, J., & Song, Y. (2019). Using spatio-temporal fusion of landsat-8 and modis data to derive phenology, biomass and yield estimates for corn and soybean. *Science of the total environment*, *650*, 1707–1721.

Lin, X., Harrington, J., Ciampitti, I., Gowda, P., Brown, D., & Kisekka, I. (2017). Kansas trends and changes in temperature, precipitation, drought, and frost-free days from the 1890s to 2015. *Journal of Contemporary Water Research & Education*, *162*(1), 18–30.

Liu, H. Q., & Huete, A. (1995). A feedback based modification of the ndvi to minimize canopy background and atmospheric noise. *IEEE transactions on geoscience and remote sensing*, *33*(2), 457–465.

Liu, J., Han, J., Zhang, Y., Wang, H., Kong, H., & Shi, L. (2018). Prediction of soil organic carbon with different parent materials development using visible-near infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *204*, 33–39.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, *164*, 324–333.

Long, J., Liu, Y., Xing, S., Qiu, L., Huang, Q., Zhou, B., ... Zhang, L. (2018). Effects of sampling density on interpolation accuracy for farmland soil organic matter concentration in a large region of complex topography. *Ecological Indicators*, *93*, 562–571.

LP DAAC. (n.d.). *Lp daac - hlsl30.* https://lpdaac.usgs.gov/products/hlsl30v015/. (Accessed on August 5, 2021)

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2cor for sentinel-2. In *Image and signal processing for remote sensing xxiii* (Vol. 10427, pp. 37–48).

Masialeti, I., Egbert, S., & Wardlow, B. D. (2010). A comparative analysis of phenological curves for major crops in kansas. *GIScience & Remote Sensing*, *47*(2), 241–259.

McFeeters, S. K. (1996). The use of the normalized difference water index (ndwi) in the delineation of open water features. *International journal of remote sensing*, *17*(7), 1425–1432.

Meersmans, J., Van Wesemael, B., & Van Molle, M. (2009). Determining soil organic carbon for agricultural soils: a comparison between the walkley & black and the dry combustion methods (north belgium). *Soil Use and Management*, *25*(4), 346–353.

Messina, C. D., Sinclair, T. R., Hammer, G. L., Curan, D., Thompson, J., Oler, Z., ... Cooper, M. (2015). Limited-transpiration trait may increase maize drought tolerance in the us corn belt. *Agronomy Journal*, *107*(6), 1978–1986.

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., ... others (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59–86.

Misra, G., Cawkwell, F., & Wingler, A. (2020). Status of phenological research using sentinel-2 data: A review. *Remote Sensing*, *12*(17), 2760.

Moeyersoms, J., & Martens, D. (2015). Including high-cardinality attributes in predictive models: A case study in churn prediction in the energy sector. *Decision support systems*, *72*, 72–81.

Montero, D., Aybar, C., Mahecha, M. D., & Wieneke, S. (2022). Spectral: Awesome spectral indices deployed via the google earth engine javascript api. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 301–306.

Morris, T. F., Murrell, T. S., Beegle, D. B., Camberato, J. J., Ferguson, R. B., Grove, J., ... others (2018). Strengths and limitations of nitrogen rate recommendations for corn and opportunities for improvement. *Agronomy Journal*, *110*(1), 1–37.

Mountier, N., Griggs, J., & Oomen, G. (1966). Sources of error in advisory soil tests: I. laboratory sources. *New Zealand journal of agricultural research*, *9*(2), 328–338.

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems engineering*, *114*(4), 358–371.

NASA. (n.d.). *Landsat Science.* Author. Retrieved from https://landsat.gsfc.nasa.gov/ (Retrieved March 7, 2023)

Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., & Huo, L.-Z. (2021). A modified bare soil index to identify bare land features during agricultural fallow-period in southeast asia using landsat 8. *Land*, *10*(3), 231.

Nguyen, G. H., Bouzerdoum, A., & Phung, S. L. (2009). Learning pattern classification tasks with imbalanced data sets. *Pattern recognition*, 193–208.

Nguy-Robertson, A., Gitelson, A., Peng, Y., Viña, A., Arkebauer, T., & Rundquist, D. (2012). Green leaf area index estimation in maize and soybean: Combining vegetation indices to achieve maximal sensitivity. *Agronomy journal*, *104*(5), 1336–1347.

Nieto, L., Houborg, R., Zajdband, A., Jumpasut, A., Prasad, P. V., Olson, B. J., & Ciampitti, I. A. (2022). Impact of high-cadence earth observation in maize crop phenology classification. *Remote Sensing*, *14*(3), 469.

Nieto, L., Schwalbert, R., Prasad, P. V., Olson, B. J., & Ciampitti, I. A. (2021). An integrated approach of field, weather, and satellite data for monitoring maize phenology. *Scientific Reports*, *11*(1), 15711.

Ninomiya, Y. (2002). Mapping quartz, carbonate minerals, and mafic-ultramafic rocks using remotely sensed multispectral thermal infrared aster data. In *Thermosense xxiv* (Vol. 4710, pp. 191–202).

Nocita, M., Stevens, A., Noon, C., & van Wesemael, B. (2013). Prediction of soil organic carbon for different levels of soil moisture using vis-nir spectroscopy. *Geoderma*, *199*, 37–42.

Olson, K. R., & Al-Kaisi, M. (2015). The importance of soil sampling depth for accurate account of soil organic carbon sequestration, storage, retention and loss. *Catena*, *125*, 33–37.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International journal of remote sensing*, *26*(1), 217–222.

Palm, C., Blanco-Canqui, H., DeClerck, F., Gatere, L., & Grace, P. (2014). Conservation agriculture and ecosystem services: An overview. *Agriculture, Ecosystems & Environment*, *187*, 87–105.

Parks, S. A., Dillon, G. K., & Miller, C. (2014). A new metric for quantifying burn severity: the relativized burn ratio. *Remote Sensing*, *6*(3), 1827–1844.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... others (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, *12*, 2825–2830.

Pelletier, C., Valero, S., Inglada, J., Champion, N., & Dedieu, G. (2016). Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, *187*, 156–168.

Peng, B., Guan, K., Pan, M., & Li, Y. (2018). Benefits of seasonal climate prediction and satellite data for forecasting us maize yield. *Geophysical Research Letters*, *45*(18), 9662–9671.

Peng, Y., Xiong, X., Adhikari, K., Knadel, M., Grunwald, S., & Greve, M. H. (2015). Modeling soil organic carbon at regional scale by combining multi-spectral images with laboratory spectra. *PloS one*, *10*(11), e0142295.

Perin, V., Roy, S., Kington, J., Harris, T., Tulbure, M. G., Stone, N., ... Yaeger, M. A. (2021). Monitoring small water bodies using high spatial and temporal resolution analysis ready datasets. *Remote Sensing*, *13*(24), 5176.

Peter, Z., & Pheap, A. (2015, September). *Unprotected areas—the cambodia daily.* https://english.cambodiadaily.com/unprotectedareas/.

Planet. (2020, September 14). *Introducing planetary variables: Soil water content, land surface temperature, and vegetation biomass.* Planet Pulse. Retrieved from https://tinyurl.com/5n8avfxr

Planet. (2022). *Fusion technical specification.* https://assets.planet.com/docs/Fusion-Tech-Spec_v1.0.0.pdf. (Retrieved March 7, 2023)

Planet Fusion Team. (2021). *Planet fusion monitoring technical specification, version 1.0.0-beta.3* (Tech. Rep.). San Francisco, CA. Retrieved from https://assets.planet.com/docs/Planet_fusion_specification_March_2021.pdfD

Rast, M., & Painter, T. H. (2019). Earth observation imaging spectroscopy for terrestrial systems: An overview of its history, techniques, and applications of its missions. *Surveys in Geophysics*, *40*(3), 303–331.

Rawls, W., Pachepsky, Y. A., Ritchie, J., Sobecki, T., & Bloodworth, H. (2003). Effect of soil organic carbon on soil water retention. *Geoderma*, *116*(1-2), 61–76.

Rezaei, E. E., Siebert, S., Hüging, H., & Ewert, F. (2018). Climate change effect on wheat phenology depends on cultivar change. *Scientific reports*, *8*(1), 4891.

Robinson, N. P., Allred, B. W., Smith, W. K., Jones, M. O., Moreno, A., Erickson, T. A., ... Running, S. W. (2018). Terrestrial primary production for the conterminous united states derived from landsat 30 m and modis 250 m. *Remote Sensing in Ecology and Conservation*, *4*(3), 264–280.

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, *71*, 804–818.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS journal of photogrammetry and remote sensing*, *67*, 93–104.

Rouse Jr, J. W., Haas, R. H., Deering, D., Schell, J., & Harlan, J. C. (1974). *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation* (Tech. Rep.).

Ruiz-Gutierrez, V., Hooten, M. B., & Campbell Grant, E. H. (2016). Uncertainty in biological monitoring: a framework for data collection and analysis to account for multiple sources of sampling bias. *Methods in Ecology and Evolution*, *7*(8), 900–909.

Ruml, M., & Vulić, T. (2005). Importance of phenological observations and predictions in agriculture. *Journal of Agricultural Sciences (Belgrade)*, *50*(2), 217–225.

Sakamoto, T., Gitelson, A. A., & Arkebauer, T. J. (2014). Near real-time prediction of us corn yields based on time-series modis data. *Remote Sensing of Environment*, *147*, 219–231.

Sakamoto, T., Wardlow, B. D., Gitelson, A. A., Verma, S. B., Suyker, A. E., & Arkebauer, T. J. (2010). A two-step filtering approach for detecting maize and soybean phenology with time-series modis data. *Remote Sensing of Environment*, *114*(10), 2146–2159.

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A crop phenology detection method using time-series modis data. *Remote sensing of environment*, *96*(3-4), 366–374.

Samkol, P., Sath, K., Patel, M., Windsor, P. A., & Holtenius, K. (2015). Survey of smallholder beef cattle production systems in different agro-ecological zones of cambodia. *Tropical animal health and production*, *47*, 1299–1306.

Sayão, V. M., & Demattê, J. A. (2018). Soil texture and organic carbon mapping using surface temperature and reflectance spectra in southeast brazil. *Geoderma Regional*, *14*, e00174.

Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., & Rundquist, D. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, *25*, 47–54.

Schramowski, P., Stammer, W., Teso, S., Brugger, A., Herbert, F., Shao, X., ... Kersting, K. (2020). Making deep neural networks right for the right scientific reasons by interacting with their explanations. *Nature Machine Intelligence*, *2*(8), 476–486.

Schwalbert, R. A., Amado, T. J., Nieto, L., Varela, S., Corassa, G. M., Horbe, T. A., ... Ciampitti, I. A. (2018). Forecasting maize yield at field scale based on high-resolution satellite imagery. *Biosystems engineering*, *171*, 179–192.

Seeley, M., & Asner, G. P. (2021). Imaging spectroscopy for conservation applications. *Remote Sensing*, *13*(2), 292.

Segarra, J., Buchaillot, M. L., Araus, J. L., & Kefauver, S. C. (2020). Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy*, *10*(5), 641.

Shan, Y., Paull, D., & McKay, R. (2006). Machine learning of poorly predictable ecological data. *Ecological Modelling*, *195*(1-2), 129–138.

Sharifi, H., Hijmans, R. J., Hill, J. E., & Linquist, B. A. (2017). Using stage-dependent temperature parameters to improve phenological model prediction accuracy in rice models. *Crop Science*, *57*(1), 444–453.

Shi, P., Six, J., Sila, A., Vanlauwe, B., & Van Oost, K. (2022). Towards spatially continuous mapping of soil organic carbon in croplands using multitemporal sentinel-2 remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *193*, 187–199.

Shihua, L., Jingtao, X., Ping, N., Jing, Z., Hongshu, W., & Jingxian, W. (2014). Monitoring paddy rice phenology using time series modis data over jiangxi province, china. *International Journal of Agricultural and Biological Engineering*, *7*(6), 28–36.

Shirvani, Z., Abdi, O., & Goodman, R. C. (2023). High-resolution semantic segmentation of woodland fires using residual attention unet and time series of sentinel-2. *Remote Sensing*, *15*(5), 1342.

Sibley, A. M., Grassini, P., Thomas, N. E., Cassman, K. G., & Lobell, D. B. (2014). Testing remote sensing approaches for assessing yield variability among maize fields. *Agronomy Journal*, *106*(1), 24–32.

Sourn, T., Pok, S., Chou, P., Nut, N., Theng, D., & Prasad, P. V. (2022). Assessment of land use and land cover changes on soil erosion using remote sensing, gis and rusle model: A case study of battambang province, cambodia. *Sustainability*, *14*(7), 4066.

Sourn, T., Pok, S., Chou, P., Nut, N., Theng, D., Rath, P., ... Prasad, P. V. (2021). Evaluation of land use and land cover change and its drivers in battambang province, cambodia from 1998 to 2018. *Sustainability*, *13*(20), 11170.

Sripada, R. P., Heiniger, R. W., White, J. G., & Meijer, A. D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*, *98*(4), 968–977.

Stockmann, U., Padarian, J., McBratney, A., Minasny, B., de Brogniez, D., Montanarella, L., ... Field, D. J. (2015). Global soil organic carbon assessment. *Global Food Security*, *6*, 9–16.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC bioinformatics*, *9*, 1–11.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC bioinformatics*, *8*(1), 1–21.

Tajik, S., Ayoubi, S., & Zeraatpisheh, M. (2020). Digital mapping of soil organic carbon using ensemble learning model in mollisols of hyrcanian forests, northern iran. *Geoderma Regional*, *20*, e00256.

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations—a review. *Remote Sensing*, *12*(7), 1135.

Tan, M., & Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105–6114).

Tatsumi, K., Yamashiki, Y., Torres, M. A. C., & Taipe, C. L. R. (2015). Crop classification of upland fields using random forest of time-series landsat 7 etm+ data. *Computers and Electronics in Agriculture*, *115*, 171–179.

Teillet, P. M. (2015). Overview of satellite image radiometry in the solar reflective optical domain. In *Remotely sensed data characterization, classification, and accuracies* (pp. 9–23). CRC Press.

Tiefenbacher, A., Sandén, T., Haslmayr, H., Miloczki, J., Wenzel, W., & Spiegel, H. (2021). *Optimizing carbon sequestration in croplands: A synthesis. agronomy 2021, 11, 882.* s Note: MDPI stays neutral with regard to jurisdic-tional claims in ....

Tivet, F., Carlos de Moraes Sá, J., Borszowskei, P. R., Letourmy, P., Briedis, C., Ferreira, A. O., & Burkner dos Santos Thiago Massao Inagaki, J. (2012). Soil carbon inventory by wet oxidation and dry combustion methods: effects of land use, soil texture gradients, and sampling depth on the linear model of c-equivalent correction factor. *Soil Science Society of America Journal*, *76*(3), 1048–1059.

Tucker, C. J. (1977). Asymptotic nature of grass canopy spectral reflectance. *Applied optics*, *16*(5), 1151–1156.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. doi: 10.1016/0034-4257(79)90013-0

Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M. (1999). Relationships between leaf area index and landsat tm spectral vegetation indices across three temperate zone sites. *Remote sensing of environment*, *70*(1), 52–68.

Union of Concerned Scientists. (n.d.). *Satellite database.* Retrieved from https://www.ucsusa.org/resources/satellite-database (Retrieved March 7, 2023)

United States Department of Agriculture (USDA). (2018a). *Crop progress and condition.* https://www.nass.usda.gov/Statistics_by_State/Kansas/Publications/Crop_Progress_and_Condition/historic.php. (Accessed on December 20, 2021)

United States Department of Agriculture (USDA). (2018b). *Usda/nass quick stats.* https://quickstats.nass.usda.gov/. (Accessed on July 15, 2021)

USDA-NASS. (2019, February). *Crop production 2018 summary.* Retrieved from https://www.nass.usda.gov/Publications/Todays_Reports/reports/cropan19.pdf

van Rossum, G. (2007). Python programming language. In *Usenix annual technical conference* (Vol. 41, p. 36).

Vaudour, E., Gholizadeh, A., Castaldi, F., Saberioon, M., Borůvka, L., Urbina-Salazar, D., ... et al. (2022). Satellite imagery to map topsoil organic carbon content over cultivated areas: an overview. *Remote Sensing*, *14*(12), 2917.

Vaudour, E., Gomez, C., Loiseau, T., Baghdadi, N., Loubet, B., Arrouays, D., ... Lagacherie, P. (2019). The impact of acquisition date on the prediction performance of topsoil organic carbon from sentinel-2 for croplands. *Remote Sensing*, *11*(18), 2143.

Viña, A., Gitelson, A. A., Nguy-Robertson, A. L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote sensing of environment*, *115*(12), 3468–3478.

Vina, A., Gitelson, A. A., Rundquist, D. C., Keydan, G., Leavitt, B., & Schepers, J. (2004). Monitoring maize (zea mays l.) phenology with remote sensing. *Agronomy Journal*, *96*(4), 1139–1147.

Viña, A., Peters, A. J., & Ji, L. (2003). Use of multispectral ikonos imagery for discriminating between conventional and conservation agricultural tillage practices. *Photogrammetric Engineering & Remote Sensing*, *69*(5), 537–544.

Vincini, M., & Frazzi, E. (2011). Comparing narrow and broad-band vegetation indices to estimate leaf chlorophyll content in planophile crop canopies. *Precision Agriculture*, *12*(3), 334–344.

Walter, K., Don, A., Tiemeyer, B., & Freibauer, A. (2016). Determining soil bulk density for carbon stock calculations: a systematic method comparison. *Soil Science Society of America Journal*, *80*(3), 579–591.

Wang, H., Ghosh, A., Linquist, B. A., & Hijmans, R. J. (2020). Satellite-based observations reveal effects of weather variation on rice phenology. *Remote Sensing*, *12*(9), 1522.

Wang, S., Azzari, G., & Lobell, D. B. (2019). Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote sensing of environment*, *222*, 303–317.

Wang, S., Di Tommaso, S., Deines, J. M., & Lobell, D. B. (2020). Mapping twenty years of corn and soybean across the us midwest using the landsat archive. *Scientific Data*, *7*(1), 307.

Wendt, J., & Hauser, S. (2013). An equivalent soil mass procedure for monitoring soil organic carbon in multiple soil layers. *European Journal of Soil Science*, *64*(1), 58–65.

White, E. V., & Roy, D. P. (2015). A contemporary decennial examination of changing agricultural field sizes using landsat time series data. *Geo: Geography and environment*, *2*(1), 33–54.

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global biogeochemical cycles*, *11*(2), 217–234.

Woolley, J. T. (1971). Reflectance and transmittance of light by leaves. *Plant physiology*, *47*(5), 656–662.

Wu, Z., Gao, Y., Li, L., Xue, J., & Li, Y. (2019). Semantic segmentation of high-resolution remote sensing images using fully convolutional network with adaptive threshold. *Connection Science*, *31*(2), 169–184.

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., ... others (2019). Current status of landsat program, science, and applications. *Remote sensing of environment*, *225*, 127–147.

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of landsat. *Remote Sensing of Environment*, *122*, 2–10.

Wulder, M. A., Roy, D. P., Radeloff, V. C., Loveland, T. R., Anderson, M. C., Johnson, D. M., ... others (2022). Fifty years of landsat science and impacts. *Remote Sensing of Environment*, *280*, 113195.

Xie, B., Zhang, H. K., & Xue, J. (2019). Deep convolutional neural network for mapping smallholder agriculture using high spatial resolution satellite image. *Sensors*, *19*(10), 2398.

Xu, L., He, N., & Yu, G. (2016). Methods of evaluating soil bulk density: Impact on estimating large scale soil organic carbon storage. *Catena*, *144*, 94–101.

Yang, L., Driscol, J., Sarigai, S., Wu, Q., Chen, H., & Lippitt, C. D. (2022). Google earth engine and artificial intelligence (ai): a comprehensive review. *Remote Sensing*, *14*(14), 3253.

Zhang, L., Liu, Z., Ren, T., Liu, D., Ma, Z., Tong, L., ... Li, S. (2020). Identification of seed maize fields with high spatial resolution and multiple spectral remote sensing using random forest classifier. *Remote Sensing*, *12*(3), 362.

Zhang, S., Tao, F., & Zhang, Z. (2017). Spatial and temporal changes in vapor pressure deficit and their impacts on crop yields in china during 1980–2008. *Journal of Meteorological Research*, *31*(4), 800–808.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., ... Huete, A. (2003). Monitoring vegetation phenology using modis. *Remote sensing of environment*, *84*(3), 471–475.

Zhong, L., Hu, L., Yu, L., Gong, P., & Biging, G. S. (2016). Automated mapping of soybean and corn using phenology. *ISPRS Journal of Photogrammetry and Remote Sensing*, *119*, 151–164.

Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote sensing of environment*, *221*, 430–443.

Zhong, L., Yu, L., Li, X., Hu, L., & Gong, P. (2016). Rapid corn and soybean mapping in us corn belt and neighboring areas. *Scientific reports*, *6*(1), 1–14.

Zhou, T., Geng, Y., Ji, C., Xu, X., Wang, H., Pan, J., ... Lausch, A. (2021). Prediction of soil organic carbon and the c: N ratio on a national scale using machine learning and satellite data: A comparison between sentinel-2, sentinel-3 and landsat-8 images. *Science of the Total Environment*, *755*, 142661.

Appendix A

Appendix A, Chapter 2

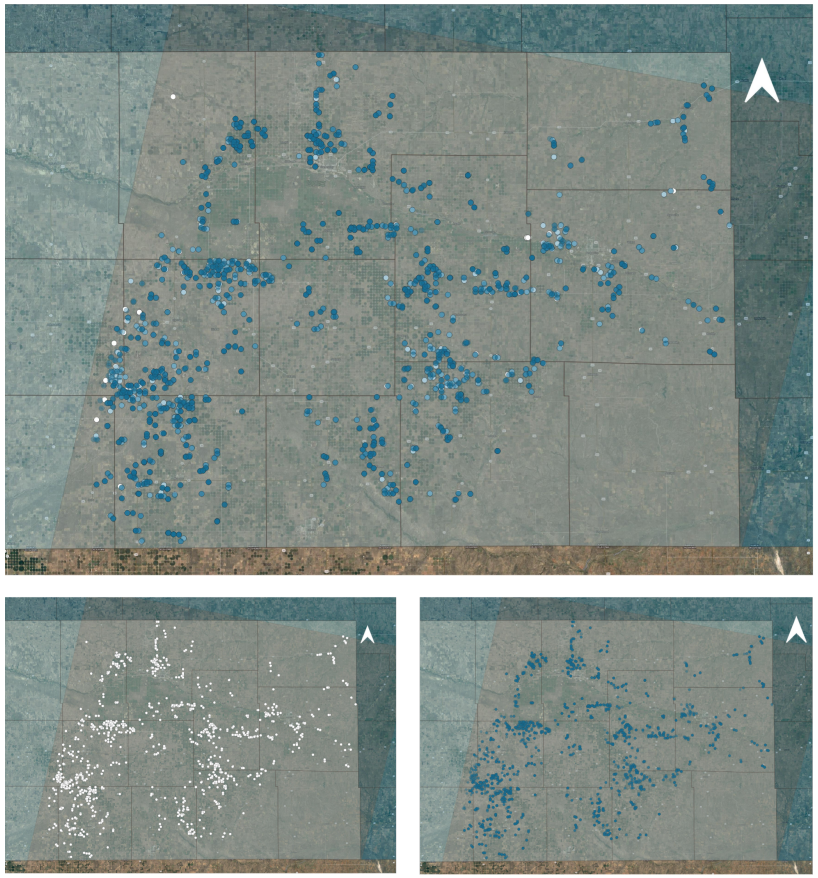

Figure A.1: Example of geographical distribution of the phenology samples for 2016. Panel (a) presents both the training and test datasets. Panel (b) white points present the distribution for the test dataset. Panel (c) blue points present the distribution for the training dataset.

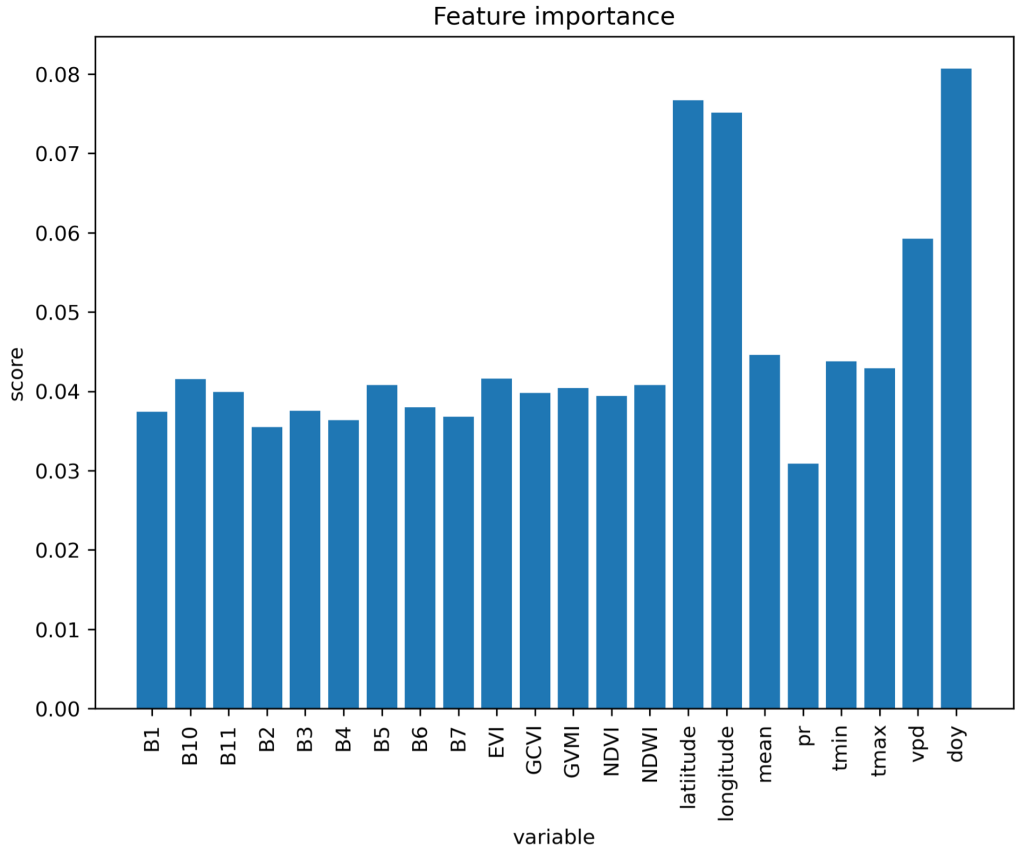

Figure A.2: Feature importance analysis for all the variables present in the dataset. The variables are described in Appendix A, Table 1.

Table A.1: Summary of variables used for random forest classification models. All variables were extracted from GEE archives with the exception of DOY, latitude, longitude, and growth classes, which were extracted from the ground truth dataset provided by Crop Quest Inc.

Group	**Variable name**	**Description**
Bands	B2	Reflectance value for band 2, blue (Landsat 8)
Bands	B3	Reflectance value for band 3, green (Landsat 8)
Bands	B4	Reflectance value for band 4, red (Landsat 8)
Bands	B5	Reflectance value for band 5, near infrared (Landsat 8)
Bands	B6	Reflectance value for band 6, short wave infrared 1 (Landsat 8)
Bands	B7	Reflectance value for band 7, short wave infrared 2 (Landsat 8)
Bands	B10	Reflectance value for band 10, thermal infrared 1 (Landsat 8)
Bands	B11	Reflectance value for band 11, thermal infrared 2 (Landsat 8)
Indices	NDVI	Normalized Difference Vegetation Index
Indices	EVI	Enhanced Vegetation Index
Indices	GCVI	Green Chlorophyll Vegetation Index
Indices	GVMI	Global vegetation Moisture Index
Indices	NDWI	Normalized Difference Water Index
Indices	DOY	Day of the year
Latitude	Latitude	Latitude value for the phenology data point
Longitude	Longitude	Longitude value for the phenology data point
Weather	GDU	Growing Degree Units
Weather	Precipitation	Cumulative Precipitation
Weather	TMax	Maximum Temperature
Weather	TMin	Minimum Temperature
Weather	VPD	Vapor Pressure Deficit

Table A.2: Summary of metrics presenting the results of overall accuracy, OOB, kappa value, precision, recall, F-score, and supports for each class (IV1, IV2, EV, LV, ER, LR1, LR2, LR3, H) from 2013 to 2018. Acronyms defined in Table 1 description

Year	**Class**	**Overall**	**OOB**	**Kappa**	**Precision**	**Recall**	**F1-score**	**Test**	**Training**
2013	IV1	0.95	0.95	0.95	1.00	0.96	0.98	195	455
	IV2				0.96	0.98	0.97	344	802
	EV				0.93	0.96	0.94	134	312
	LV				0.99	0.99	0.99	212	494
	ER				0.96	0.88	0.92	50	116
	LR1				0.94	0.95	0.95	389	907
	LR2				0.93	0.87	0.90	264	616
	LR3				0.96	0.99	0.97	546	1274
	H				0.96	0.85	0.90	60	140
2014	IV1	0.97	0.96	0.94	1.00	0.90	0.94	134	312
	IV2				0.98	0.99	0.99	252	588
	EV				0.96	0.97	0.96	111	259
	LV				0.95	0.99	0.97	225	525

Continuation of Table A.2

	ER				1.00	0.90	0.95	56	130
	LR1				0.93	0.98	0.95	477	1113
	LR2				0.97	0.98	0.93	146	340
	LR3				0.96	0.98	0.97	453	1057
	H				1.00	0.83	0.90	43	100
2015	IV1	0.92	0.92	0.88	0.97	0.84	0.90	189	441
	IV2				0.88	0.96	0.92	331	772
	EV				0.91	0.92	0.91	190	443
	LV				0.99	0.96	0.95	184	429
	ER				0.87	0.98	0.92	77	179
	LR1				0.93	0.90	0.91	296	690
	LR2				0.94	0.85	0.88	322	751
	LR3				0.92	0.98	0.95	501	1169
	H				0.92	1.00	0.96	34	79
2016	IV1	0.96	0.95	0.94	1.00	0.94	0.97	132	308
	IV2				0.97	0.97	0.97	427	996
	EV				0.94	0.95	0.95	264	616
	LV				0.98	0.97	0.98	177	413
	ER				0.98	0.94	0.96	100	223

Continuation of Table A.2

	LR1				0.93	0.95	0.94	417	973
	LR2				0.91	0.83	0.87	186	434
	LR3				0.96	0.99	0.97	444	1036
	H				0.94	1.008	0.97	17	39
2017	IV1	0.94	0.95	0.93	0.98	0.98	0.98	205	478
	IV2				0.97	0.97	0.97	613	1430
	EV				0.94	0.89	0.92	203	473
	LV				0.95	0.96	0.96	202	471
	ER				0.91	0.96	0.93	121	282
	LR1				0.94	0.98	0.96	477	1113
	LR2				0.96	0.85	0.90	271	632
	LR3				0.95	0.96	0.96	494	1152
	H				1.00	0.97	0.98	61	142
2018	IV1	0.95	0.94	0.91	1.00	0.82	0.90	91	212
	IV2				0.95	0.99	0.97	439	1024
	EV				0.96	0.88	0.92	161	375
	LV				0.86	1.00	0.93	66	154
	ER				1.00	0.86	0.92	42	98
	LR1				0.94	0.94	0.94	267	623

Continuation of Table A.2

LR2	0.90	0.88	0.90	159	371
LR3	0.96	0.98	0.97	410	956
H	1.00	1.00	1.00	6	14

Appendix B

Appendix B Chapter 4

Table B.1: Before and after dates, and the distance between both

before date	after date	distance
2021-03-18	2021-03-22	118.042
2020-02-13	2020-02-17	111.486
2020-02-18	2020-02-22	103.871
2020-04-26	2020-04-30	101.127
2021-02-13	2021-02-17	92.907
2020-02-23	2020-02-27	88.992
2020-01-28	2020-02-01	85.763
2021-03-28	2021-04-01	79.841

Continuation of Table B.1

2021-01-23	2021-01-27	79.495
2021-02-18	2021-02-22	78.895
2021-01-08	2021-01-12	78.096
2021-03-23	2021-03-27	77.168
2020-01-03	2020-01-07	76.286
2021-02-23	2021-02-27	75.964
2020-02-08	2020-02-12	75.505
2020-03-23	2020-03-27	73.172
2020-02-03	2020-02-07	72.820
2020-03-08	2020-03-12	72.102
2021-02-03	2021-02-07	66.642
2021-02-08	2021-02-12	64.498
2020-01-08	2020-01-12	63.243
2021-01-28	2021-02-01	61.874
2021-01-18	2021-01-22	57.514
2020-02-26	2020-03-01	56.530
2021-03-13	2021-03-17	56.204
2021-01-13	2021-01-17	55.467
2021-02-26	2021-03-02	53.712
2021-01-03	2021-01-07	53.675

Continuation of Table B.1

2020-03-13	2020-03-17	53.143
2020-04-18	2020-04-22	52.273
2020-01-18	2020-01-22	51.046
2020-03-03	2020-03-07	50.370
2020-03-18	2020-03-22	50.002
2021-03-08	2021-03-12	48.329
2020-01-13	2020-01-17	47.421
2020-04-23	2020-04-27	47.354
2020-03-28	2020-04-01	46.843
2020-01-23	2020-01-27	46.677
2020-04-03	2020-04-07	42.279
2021-03-03	2021-03-07	41.488
2020-04-13	2020-04-17	40.576
2020-04-08	2020-04-12	34.287