



UNIVERSITAT_{DE}
BARCELONA

Multi-scale and multi-sensor remote sensing in international agricultural development

Maria Luisa Buchailot



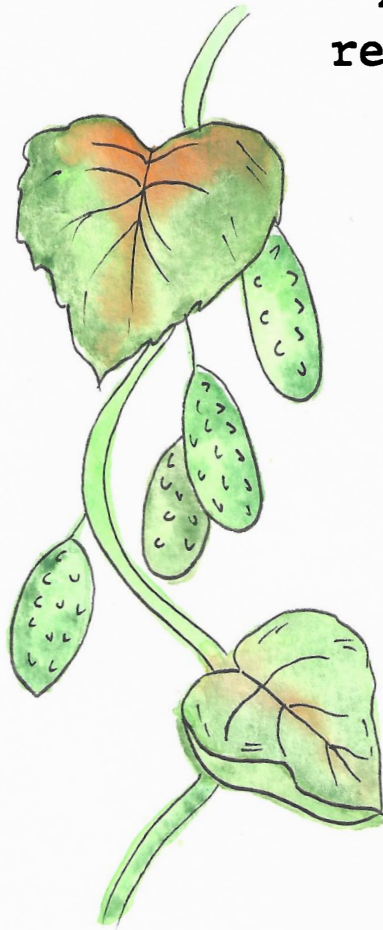
Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**

Multi-scale and multi-sensor remote sensing in international agricultural development

Maria Luisa Buchailot



Doctoral Thesis
BARCELONA, December 2022



UNIVERSITAT DE
BARCELONA

"Multi-scale and multi-sensor remote sensing in international
agricultural development"

Dissertation was presented by Maria Luisa Buchailot to obtain the degree of Doctor from the
University of Barcelona.

This work is part of the doctoral program in Ecology, Environmental Sciences, and Plant
Physiology of the Department of Evolutionary Biology, Ecology and Environmental Sciences
(BEECA) of the Faculty of Biology of the University of Barcelona.

This work has been carried out in the Integrative Crop Ecophysiology Group research group,
under the direction of Dr. Shawn C. Kefauver and Dr. Jose Luis Araus Ortega.

Doctorant

A handwritten signature in black ink, appearing to read 'Luisa Buchailot'.

Maria Luisa Buchailot

Director y tutor

A handwritten signature in black ink, appearing to read 'Shawn C. Kefauver'.

Dr. Shawn C. Kefauver

Director

A handwritten signature in black ink, appearing to read 'Jose Luis Araus Ortega'.

Dr. Jose Luis Araus Ortega

BARCELONA, DECEMBER 2022

*A Dios, es el guía que me acompaña
siempre, mi fiel consejero, padre y amigo.*

*A mis padres, quienes me dieron la vida y
apoyaron incondicionalmente
en este camino.*

*A mis hermanos y sus familias, a quienes
amo tanto, quienes siempre estuvieron
y estarán.*

*A mi familia de Barcelona, que me acogieron
y me amaron sin medida.*

*"No siempre podemos hacer grandes cosas, pero sí
podemos hacer cosas pequeñas con gran amor"*

Madre Teresa de Calcuta-

ABSTRACT

Personal food security means that they have physical and economic access to sufficient, safe, and quality food. On the other hand, there are three main causes of food insecurity: 1. the high vulnerability of agriculture to climate change; 2. local, national, or international conflicts; and 3. economic inequality. The United Nations (UN) Sustainable Development Goals (SDGs), which are objectives that should be attained by 2030, are targeted at both developing and developed nations. The second SDG, "Zero Hunger," aims to double small-scale food producers' productivity and earnings while promoting resilient agricultural methods and ensuring sustainable food production. Agriculture is heavily reliant on factors related to climate change conditions such as abiotic stress, which includes soil nutrient deficiencies, accelerating temperature rises, drought, and rising CO₂ concentrations; and biotic stress, which includes invasive pests, disease outbreaks, and decreased crop output. Remote sensing (RS) technologies can provide several non-destructive methods for identifying and quantifying various types of stress. For the application of RS, it is relevant to consider the different types of resolution: spectral, spatial, temporal, and radiometric. Also, the different observation scales are ground-based, aerial, space-based, or using orbital satellites. In this thesis, we evaluated the practical implementation of non-destructive methods using RS technologies across the four chapters. Moreover, we compared the maturity levels between different types of technologies using Technological Readiness Level (TRL) assessments. In the first chapter, our objective was to estimate the grain yield of the maize under low nitrogen using Vegetation Indices (VIs) from RGB (Red, Green, Blue composite color images) sensors at the ground and aerial levels in Sub-Saharan Africa (SSA). We developed an RS system in the second chapter to monitor an early warning fall armyworm (FAW) across SSA. In the third one, using leaf spectral reflectance and advanced regression models, we estimated the $V_{c,max}$ and J_{max} of soybean and peanut. And in the last chapter, we developed a user-friendly mobile app for the Middle East and North Africa (MENA) countries for plant disorders detection on tomatoes, cucumbers, peppers, and quinoa, covering everything from data collection to deep learning model creation, to web and mobile app launch. Different RS technologies were used in different countries at different scales and with different types of sensors. Nevertheless, it was very relevant to consider the objectives of each study because they determined the type of spectral, spatial, and temporal resolution and the scale of observation. Regarding the TRLs across the four chapters, they suggest that the level of technology readiness depends on the goals, the time to develop the project, the amount of data collection required, and the robustness of the validation.

ACKNOWLEDGEMENTS

Si tuviera que dar gracias a todas las personas que fueron parte de este viaje a lo largo de estos cuatro años, habría más cantidad de nombre personas que de palabras escritas en toda la tesis...

Voy a simplificar un poco, porque esta tesis es fruto de muchas cosas, de trabajo en equipo, de culturas nuevas en diferentes países, de resiliencia, de amor, de ganas de aprender, de descubrir que no todos tenemos el mismo DON, y de que cada uno en este mundo está para sembrar su granito de arena y hacer un mundo mejor; y sin nada de esto no hubiera sido posible. Por esto, GRACIAS a cada persona que paso por mi vida, enseñándome, corrigiéndome, ayudándome, sonriéndome, guiándome, ayudándome a enfrentar mis miedos, a ser mejor persona cada día, y a descubrir cosas nuevas; acogióndome en cada lugar nuevo por descubrir y amarme.

Gracias a **Shawn** que fue el mejor tutor/ director que me ha podido tocar. Desde que empecé a trabajar con él en la tesis de mi máster, siempre estuvo ahí para enseñarme a confiar en mí, enseñarme las habilidades desde cómo utilizar una cámara para sacar una foto, análisis de datos, volar un dron, escribir científicamente, mejorar mi inglés... hasta mandarme a Roma sola para presentar un proyecto, disfrutar los viajes de trabajo, a tener paciencia, a sacar siempre la parte positiva de cada situación de mierda. En fin, sin él no hubiera sido posible nada de todo este viaje.

Gracias a **Jose Luis** mi otro director que, sin su ayuda, guía, correcciones, enseñanzas no hubiera sido posible este viaje.

Gracias a todos mis colegas parte del grupo de **Integrative Ecophysiology Group** que fueron pilares en este viaje. De cada uno e aprendido cosas. **Adrián** fue el primero que conocí y me presento como era el mundo de la investigación y me enseñó habilidades básicas para poder seguir el cual se convirtió en un gran amigo y compañero. De **Fati y Melissa**, dos colegas que se convirtieron en grandes amigas, muy importante en los últimos años para poder seguir en la marcha y no rendirme, a escucharme cuando más lo necesitaba, y que la paciencia lo es todo. A **Jose Armando** colega latino el cual en los primeros años en los cuales me entendía mucho por similitudes de cultura y el mejor enseñándome la estadística o cada duda que no entendía, gran profesor. A **Joel** colega/ amigo en el cual el primer año compartimos investigaciones juntos, pero lo mas importante que el me hizo parte de su familia hermosa mostrándome muchas costumbres y cosas lindas de Catalunya que tanto extrañaba de mi familia. A **Omar** que compartimos poquito tiempo juntos en la universidad, pero gran amigo de cervecita después de los días estresantes. A **Cristina, Rut, Jordi, Thomas**, que lo que mas hicieron fue escucharme cuando lo más necesites en momentos que no podía más. Y a todos los estudiantes que pasaron

por el grupo, que cada uno me enseñó algo nuevo. Para destacar con respecto a este grupo lo mejor que compartí fueron los viajes de campaña a lo largo de España, el trabajo en equipo, cada anécdota que hay, los lugares donde íbamos, la comida espectacular que comíamos, grandes aventuras que las guardo en mi corazón.

Gracias a **Álvaro** que fue mi profesor en los 3 meses de estancia en Auburn University, donde me dio la oportunidad de aprender mucho sobre el uso de diferentes tecnologías que fueron muy útil para esta tesis, y a sus estudiantes que me ayudaron en los experimentos en el campo **Mery, Joan, Seth, Trammel**. Gracias a **David** que fue mi compañero de trailer/piso esos tres meses, y sin el no hubiera sobrevivido a la profunda Alabama. Gracias a **Gastón** cordobés argentino, que me trajo de esa familiaridad cuando lo necesitaba.

Quiero agradecer a todos los seres humanos que fueron parte de cada viaje que hice en Zimbabue, Tanzania, Kenia, Estados Unidos, Emiratos Árabes, Egipto, Tunes, para mi este fue el aprendizaje mas lindo, me sumergí en diferentes culturas, costumbre, comidas, convivencias. Todo esto me ayudo a tener otra viste del mudo y que somos seres diversos, que tenemos que aceptar a todos. Grandes científicos fueron partes de estos viajes donde he aprendido muchas cosas de ellos, gracias a **Jill, Esnath, Ken, David, John, Peter, Annalyse, Sumitha, Henda, Ghazi**.

Gracias gigantes a **SENSECO**, el cual Shawn me lo introdujo a ellos, grupo de científicos increíbles que me ayudaron a ser mejor científica, sintiéndome participe de una comunidad donde mi voz tiene peso. Gracias por los espectaculares viajes científicos juntos a Portugal, Letonia, Bulgaria, Rumania, donde aprendí mucho a nivel científico, pero también a nivel personal y cultural. Quiero destacar a **Katja, Verónica y Miriam**, tres mujeres las cuales me inspiraron en este viaje.

Gracias a mis amigos/colegas que me fui encontrando en todos los países recorridos.

Gracias a **Yaiza y Dani**, dos doctorandos de ecología que la vida me los presento, y se transformaron en los gurús amigos de como da una presentación oral, la estética de los ppt y la buena comida.

Gracias a **P. Antonio, Anna y Silvia**, que ellos me acompañaron desde lo espiritual, psicológico y físico para poder terminar esta tesis sin volverme loca.

Gracias a **FAMILIA de BARCELONA**, acá hay muchas personas involucradas que fui conociendo a lo largo de los cuatro años. Toda la comunidad de Santa Maria del Mar y a la comunidad Adsis por acogerme, escucharme y cuidarme teniendo a mi familia lejos. **Giacomo, Fer, Alberto, Sergio, Victoria, Dani, Isabel, Alina**. A mi familia que fui encontrando a Barcelona, a **Xavi, Azahara, Agus, Luli, Agus, Jona, Franci, Paula, Meli**.

Gracias a **Domi**, mi amor que apareció en el último año y que me apoyó desde el primer momento, y lo sigue haciendo hasta hoy.

Gracias a mis AMIG@S que son un soporte increíble. **Camila** mi amiga apreciada que con su hermoso arte me hace ver la vida de otra perspectiva, su amor incondicional, apoyándome siempre, escuchándome (Se puede apreciar su arte en las portadas de esta tesis). **Stefi** (con bebe), **Ceci**, **Florcita**, **Ludi** amigas geniales que están el día a día ayudándome a ser mejor, escuchándome en mis peores momentos a lo largo de estos años para poder terminar este viaje siempre viendo las cosas positivas y todo lo que he logrado. **Joaquín** amigo fiel, que también compartimos vida de doctorando, entonces aprendimos a acompañarnos en este viaje. **Juanmi** amigos que vive en otra parte del mundo, pero cada vez que sentía que no podía lo llamaba y él siempre estaba ahí para ayudarme a levantarme.

Gracias a **Marina**, amiga/confidente, que cuando deje mi país, fue la persona que me abrió su corazón desde el momento primero del máster, introduciéndome en todo su mundo sin cambio a nada. Gracias a Dios, elegimos las dos el mismo viaje del doctorado. Ella siempre estuvo y esta. Compartimos tantos momentos, decepciones, lloreras, pero aprendimos tanto juntas, la resiliencia es lo que la destaca a ella. Y claramente sin ella no hubiera sido posible este hermoso viaje.

Gracias a mi **FAMILIA**, mis papas, hermanos, hermanas, cuñadas, cuñados, mis hermosos sobrinos, por amarme, siempre apoyarme y creer en mí de que yo podía lograr esta tesis desde el otro lado del charco. Los extraño, los amo y siempre están presente en mi corazón, aunque estemos a miles de km. **MAMI** y **PAPI** sin ustedes dos nada hubiera sido posible, gracias por darme la vida y guiarme que siempre poniendo amor todo se puede.

TABLE OF CONTENTS

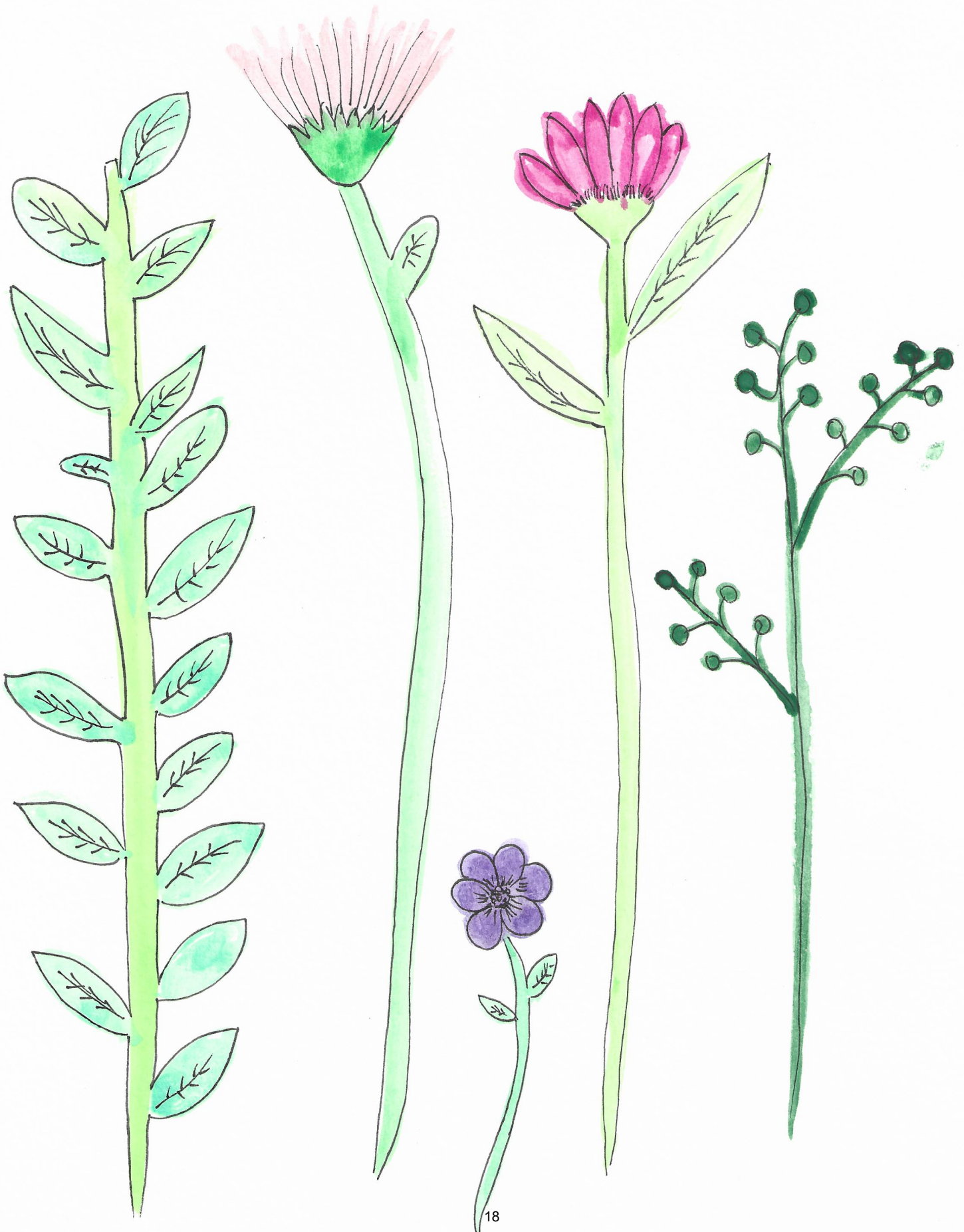
TABLE OF CONTENTS

ABBREVIATIONS	16
INTRODUCTION	18
OBJECTIVES	32
REPORT OF THE THESIS DIRECTORS	37
CHAPTER 1. Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques.	43
CHAPTER 2. Regional monitoring Fall Armyworm (FAW) using early warning systems.	75
CHAPTER 3. Estimating peanut and soybean photosynthetic traits using leaf spectral reflectance and advance regression models.	99
CHAPTER 4. Deep learning for detection of plant disorders on crop leaves: from data collection to framework tools.	123
DISCUSSION	151
CONCLUSIONS	162
RESUMEN GENERAL DE LA TESIS	167
REFERENCES	169

ABBREVIATIONS

ARDR , Automatic Relevance Determination Regression model	NDVI , Normalized Difference Vegetation Index
BR , Bayesian Ridge model	NIR , Near Infrared
DL , Deep Learning	PA , Precision Agriculture
DSS , Decision Support System	PLSR , Partial Least Squares Regression model
ESA , European Space Agency	R&D , Research and Development
FMIS , Farm Management Information System	RGB , Red-Green-Blue
FWHM , Full-width Half Maximum	RS , Remote Sensing
GA , Green Area	SCOPE , Soil-Canopy Observation of Photosynthesis and Energy
GGA , Greener Green Area	SDG , Sustainable Development Goals
GY , Grain Yield	SFTs , Smart Farming Technologies
HTPP , High Throughput Phenotyping Platforms	SSA , Sub-Sahara Africa
ITC , Information and Communication Technologies	SWIR , Short wave infrared
J_{max} , Maximum electron transport rate supporting RuBP regeneration	TIR , Thermal Infrared
LAI , Leaf Area Index	TRLs , Technological Readiness Level
LASSO , Least Absolute Shrinkage and Selection Operator model	UAE , United Arab Emirates
MENA , Middle East and North Africa	UAV , Unmanned Aerial Vehicles
ML , Machine Learning	UN , United Nations
NASA , National Aeronautics and Space Administration	V_{c,max} , Maximum rate of rubisco catalyzed carboxylation
	Vis , Vegetation Indices
	VIS , Visible

INTRODUCTION



INTRODUCTION

Food Security and the Second Sustainable Development Goal (SDG 2) “Zero Hunger”

Food security exists when all individuals have physical and economic availability to sufficient, safe, and quality food; and so that they are able to meet their own dietary needs for food and nutritional preferences to sustain an active and healthy life (Zolfaghari and Farzaneh 2021). New evidence shows that the number of hungry people in the world is growing, reaching more than 820 million in 2018 (von Grebmer et al.,2018). There are three main causes of food insecurity: (i) High exposure and vulnerability to climate extremes: the climate variability and extremes are a rising trend not only in terms of frequency but also particularly impacting less developed countries where agriculture is key but at the same time is more affected by the climate extremes. (ii) Conflicts are a key driver of severe food crises and recently re-emerged as famines, while hunger and undernutrition are significantly worse where conflicts are prolonged and institutional capacities weak. (iii) Economic slowdown: Income inequality increases the likelihood of severe food insecurity, and this effect is 20% higher for low-income countries compared with middle-income countries. (Boliko 2019)

The definition, principles, and dimensions of the United Nations’ sustainable development goals (SDG) were focused on how integrated social, economic, and environmental issues can be approached together in order to answer the challenges of sustainable development at the global level and create a better life and lifestyle for the people in this world regardless of their origin, race, or religion (Tanumihardjo et al. 2020). Regarding the SDGs, there are 17 that were established in 2015 by the United Nations (UN), which were set to be achieved by 2030, and were focused on developing and developed countries, since concerted action among all countries is the only way to achieve prosperity without threatening planetary boundaries (Steffen et al. 2015; Christian Kroll 2017)

With respect to SDG number 2, also called "Zero Hunger", it centers on bringing an end to hunger, food security, attaining better nutrition, and encouraging agricultural sustainability. This SDG is composed of eight targets, e.g., specific, measurable, and time-bound outcomes that directly contribute to the achievements of a goal. The 5 principal targets outlined by 2030 are: 2.1 End hunger and ensure access to safe nutritious and sufficient food; 2.2 End all the forms of malnutrition; 2.3 Double the productivity and incomes of small-scale food producers; 2.4 Ensure sustainable food production systems and implement resilient agricultural practices; and 2.5 Maintain the genetic diversity of seeds, plants, and animals. On the other hand, the last 3 targets are focused implementing mechanisms to 2.a) increase investment through enhanced

international cooperation; 2.b) correct and prevent trade restrictions and distortions in world agricultural markets; 2.c) ensure the proper functioning of food commodity markets and their derivatives, and also facilitate timely access to market information (Otekunrin et al. 2019; Tanumihardjo et al. 2020).

Agriculture is highly dependent on climatic conditions. There are studies that have indicated that most crop yields negatively respond to an increase in temperature (Ureta et al. 2020). Also, long-term changes in climatic conditions are also likely to increase the occurrence of extreme weather events (Cobián Álvarez and Resosudarmo 2019), and can initiate and alter the timing of pest and disease outbreaks (Nelson et al. 2009), reduce water nutrient use efficiency, and increase yield variability (Asplund et al. 2014). All these factors affect the lower limits of agricultural productivity and can undermine global efforts to reduce the SDG 2 goal of eradicating hunger. Negative effects of global change are becoming more frequent and lead, without swift and effective management responses, to decreases in crop productivity (Atzberger 2013). Due to climate change impacts, adaptation plans have been initiated against it.

With respect to the agricultural sector, farmers have been employing diverse adaptation plans and strategies. The major agricultural adaptation exercised by small farmers includes adjustments in farm operation timing, on-farm diversification, and soil-water management by improving irrigation, managing disease, and treating different nutrition deficiencies (Below et al. 2012; Jawid and Khadjavi 2019).

For the detection and quantification of biotic and abiotic stresses in agricultural crops, destructive and non-destructive methods can be used. Some quantitative methods provide sensitive analyses of molecules in biological systems, also known as high-throughput “-omics” techniques, i.e., metabolomics, genomics, proteomics, etc. (Fiehn 2001). However, these methods are not only destructive but also time-consuming and costly, limiting their use in continuous monitoring and scalable research (Galieni et al. 2021). As for non-destructive quantitative methods, remote sensing (RS) technologies can be considered either established or under development, such as imaging spectroscopy, fluorescence spectroscopy, and thermal and microwave RS, which all provide insights into the symptoms of stress in plants. In contrast to “omics” techniques, RS can be applied at larger spatial scales, with high revisit frequency, hence enabling the cost-effective detection of crop stress status and spatiotemporal dynamics across cultivated landscapes. Furthermore, RS is suitable for global coverage, affordable approaches exist and can thus potentially contribute to enhanced food security in developing countries

(Rembold et al. 2000). RS also provides effective means to help mitigate farmers' risks against production loss, for example using index-based insurance solutions (de Leeuw et al. 2014).

Remote Sensing for Agriculture

Plants interact with sunlight differently depending on the wavelength observed, in this sense, incident solar radiation can follow three pathways; it can be transmitted, reflected, or absorbed (Fig. 1). The electromagnetic radiation that is reflected by plants contains information about their biophysical composition and physiological status and can be measured using different sensors at different scales. Regarding the visible (VIS) domain, the leaf pigments present in green leaves are strongly absorbed in this region of the spectrum (400-700nm), especially in the blue and red wavelengths, where energy is captured for photosynthesis. Meanwhile, in the near infrared (NIR), from approximately 700 to 1300 nm, leaves exhibit high reflectance values and transmission, related to cell structure and biomass, and absorb less radiation in this spectral region. Also, plant canopy structure and leaf area are also fundamental traits and key parameters for monitoring growth. With respect to the part of the shortwave infrared (SWIR), from approximately 1300 to 2500 nm, the absorption of radiation is largely dominated by water contents and leaf biochemicals (Gerhards et al. 2019; Buchaillet et al. 2022).

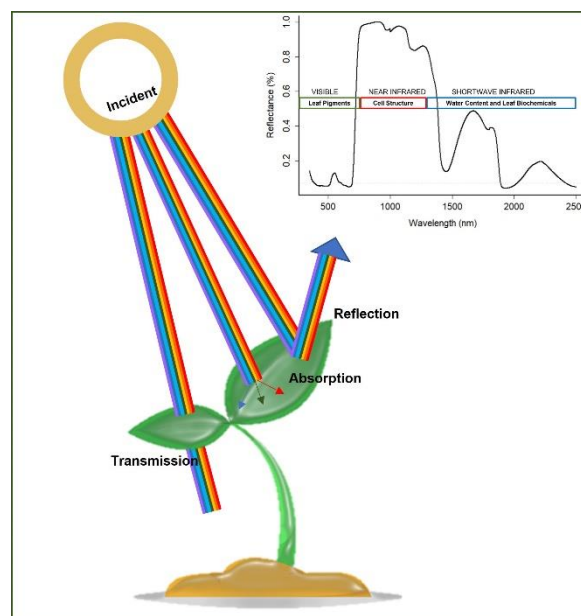


Fig. 1 The interaction between sunlight and plant; incident solar radiation can follow three pathways: it can be transmitted, absorbed, and reflected. The crop reflectance spectral features, the visible (VIS) leaf pigments, the near infrared (NIR) cell structure, and the shortwave infrared (SWIR) water content and leaf biochemicals.

Concerning agriculture, in situations where crops interact with any given aspect of their environment (seasonal climatic variations, meteorological extreme events, pests, soil properties, etc.) or as crops grow and pass through different phenological stages, the interactions between plants and light reflectance translate into changes in plant signal patterns that can be interpreted using different sensors. The use of RS instruments to monitor electromagnetic radiation reflectance changes in crops is well demonstrated in the scientific literature; for example, estimating contents of chlorophyll, xanthophyll, nitrogen, phosphorous, fiber, sucrose, and photosynthetic capacity (Gamon et al. 1997; Peñuelas and Filella 1998; Petisco et al. 2006; Koester et al. 2014; Serbin et al. 2014; Dechant et al. 2017; Yendrek et al. 2017a; Buchailot et al. 2018, 2022). In addition, the majority of proposed methods rely on parametric regressions, i.e. use of the spectral bands, vegetation indices (VIs) or spectral ratios and their relationships with functional traits linked to plant stress (Govender et al. 2009; Gerhards et al. 2016; Herrmann et al. 2020). However more advanced statistical methods, including machine learning may allow to fully explore the potential capabilities of the different RS approaches. Regarding VIs, the Normalized Difference Vegetation Index (NDVI), one of the most common remote sensing indexes, provides a simplified yet reliable estimate of green biomass (Tucker 1979). NDVI is derived from visible and near-infrared reflectance, which is closely related to vegetation presence or vigor and can also be measured at the ground level with portable sensors (Thenkabail et al. 2002). Similarly, for two of these well-known indexes, there are the low-cost alternatives in the Red-Green-Blue (RGB) indexes, taken from commercial RGB cameras. For example, the Green Area (GA) index represents the percentage of pixels in the image that ranges from yellow to bluish-green color as calculated by the open-source software BreedPix and FIJI (Casadesús and Villegas 2014; Kefauver et al. 2020).

With respect of RS, it is central to understand the different types of resolution and the different observation scales. Because they determine the monitoring capacities of the RS in the agricultural systems. Regarding resolution, four parameters are central to approaching a better understanding of the aims and scope of these instruments: the spectral, spatial, temporal, and radiometric.

Spectral resolution

The spectral resolution being the wavelength width and the number of the different frequency bands recorded by the detector which determine the spectral signatures used to assess the objects by remote sensing. The spectral data range could be from 400 to up to 2500 nm. Many

sensors only cover a limited part of this range, namely the VIS and NIR from about 400 to 1000 nm, on the other hand measuring information from 1100 to 2500 nm in reflectance is possible but requires different more expensive radiation detection sensors. There are different types of commercial sensors, RGB cameras with VIS spectral coverage (Buchailot et al. 2019; Gracia-Romero et al. 2019) and multispectral cameras for agriculture where they use VIS and some bands from the NIR (Gracia-Romero et al. 2017). There are also sensors that cover the whole spectro-radiometer range of 350-2500nm such as ADS Field Spec Hi-Res 4, e.g., Yendrek et al. 2017; Buchailot et al. 2018; Silva-Perez et al. 2018. In addition, there are hyperspectral cameras that have all continuous overlapping bands that provide full coverage bands from 400 to 900 part of VIS and NIR (Meacham-Hensold et al. 2019; Fu et al. 2020). In all these sensors with different bands, each band presents a specific width standardized measurement is at the point of the full-width half maximum (FWHM) of its full light sensitivity determines its spectral resolution. In addition, we may mention the thermal infrared (TIR) sensors which work in the wavelength range between 7000 and 14000 nm. Even when these sensors are traditionally very expensive their cost have decreased dramatically in recent years.

Spatial resolution

The spatial resolution, which is the measure of the smallest object that can be resolved by the sensor, depend on pixel size. In the Fig. 2 we can see different resolutions, from satellite images with 10 m of resolution per pixel to mobile phone pictures with 0.001 m of resolution per pixel. Besides, some satellites have 1 km or greater spatial resolution, and therefore those are not always useful for small farmers. The type of resolution depends on the sensor and the type of platform that you use to monitor your study. The spatial resolution can change on adjustable platforms, like unmanned aerial vehicles (UAVs) because they depend on the flying height and the type of the commercial sensor but are usually fixed on satellites.

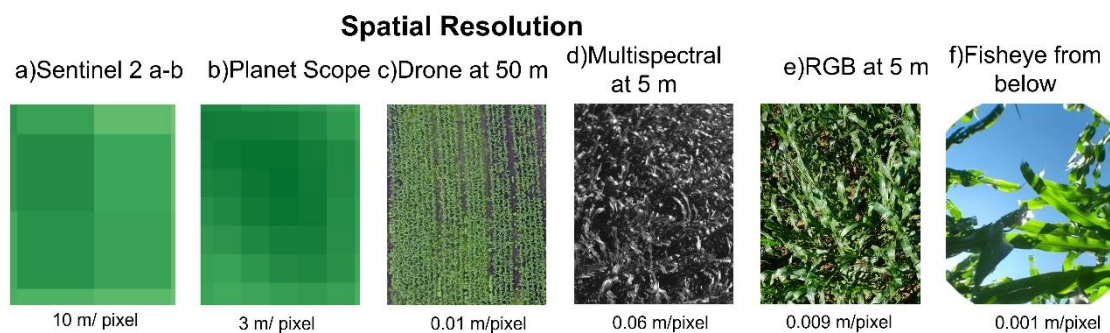


Fig. 2 Examples of different spatial resolutions. a) Calculated NDVI with Sentinel bands with a resolution of 10 m/pixel. b) Calculated NDVI with Planet Scope bands with a resolution of 3

m/pixel. c) Photo taken with a DJI Phantom 4 UAV at 50 m of altitude, and with a resolution of 0.01 m/pixel. d) Photo taken with a multispectral camera at 5 m of altitude, and with a resolution of 0.06 m/pixel. e) Photo taken an RGB camera at 5 m of altitude, and with a resolution of 0.009 m/pixel. f) Photo taken with a mobile phone with a fisheye-adapter from below, and with a resolution of 0.001 m/pixel.

Temporal resolution

Temporal resolution which is a measure of the frequency with which a sensor revisits the same area of the study. Temporal resolution is a relevant parameter to consider; an open-source satellite, such as Sentinel 2 a+b captures image data each 5 days and Landsat captures data each 16 days. A satellite's temporal resolution use to be fixed, regarding aerial or ground level platforms. These last platforms are more flexible since the temporal resolution can be adjusted as needed depending on the aim of the study.

Finally, the radiometric resolution is defined as the number of different intensities of radiation the sensor can distinguish. the ranges from 8 to 14 bits corresponding to 256 levels of greyscales and up to 16.384 intensities in each band (Thenkabail et al. 2002). Also, it is a combination of two components related to the design and operation of sensor systems (Verde et al. 2018). The first component is sensor system noise and the second is the number of quantizing levels present in the analog to digital converter of the sensor system (Rao et al. 2006).

Observation scales

Moreover, the RS technologies, with respective controllers and data loggers that complement the sensor systems, may be assembled into what is termed high throughput phenotyping platforms, HTPPs (Araus and Kefauver 2018; Araus et al. 2018; Jin et al. 2020). Regarding the matching of ecological, agricultural, and remote sensing scales, we currently find that, based on individual plants, plant communities, and agroecosystems, different RS technologies may provide coverage across canopy and landscape scales (Homolová et al. 2013). The concept of the phenotyping platform is widely used now and embraces a varied range of options in terms of placement: ground, aerial, and space level (Fig. 3, (Jin et al. 2020)). Within the category of ground levels, the range of options is very wide: from a simple hand-held sensor, including for example monopods, and tripods carrying any sensors from a simple yet effective RGB color camera, to a complex ground vehicle of diverse natures, which are generally named as "phenomobiles", and include tractor-mounted sensors and mobiles cranes. Moreover, there are the stationary platforms that are simultaneous and fully automated, fixed-site phenotyping equipment. They can carry sensors for the noninvasive estimation of crop growth, physiology,

morphology, and health (Jin et al. 2020). Also, the hand-held category platforms, smartphones are becoming an attractive alternative given that they may capture different images (e.g., RGB and thermal), data management activities, and geo-referencing functions (Araus and Kefauver 2018). In addition, in this category one may also include highly complex stationary facilities. Cable-based robotics system are also becoming an alternative for outdoor, i.e., field phenotyping (Virlet et al. 2017). Regarding aerial platforms, aerial vehicles of different nature are being widely used, particularly UAV, popularly known as drones, the sensors that they most frequently carry out include RGB cameras, multispectral, hyperspectral, and thermal sensors, or imagers (Aasen et al. 2018; Buchailot et al. 2019). Multi-rotor UAVs are often used in crop breeding programs because they offer convenient vertical take-off and landing capacities and adapt easily to carrying different payloads. Nevertheless, other unmanned options, e.g., fixed-winged UAVs, are often used where the crop area to monitor is larger than a few hectares and for more standardized and operative precision agriculture activities that don't require changing payloads. Other alternatives such as manned aircraft are less used by crop breeders given the lower costs of these unmanned alternative (Maes and Steppe 2019).

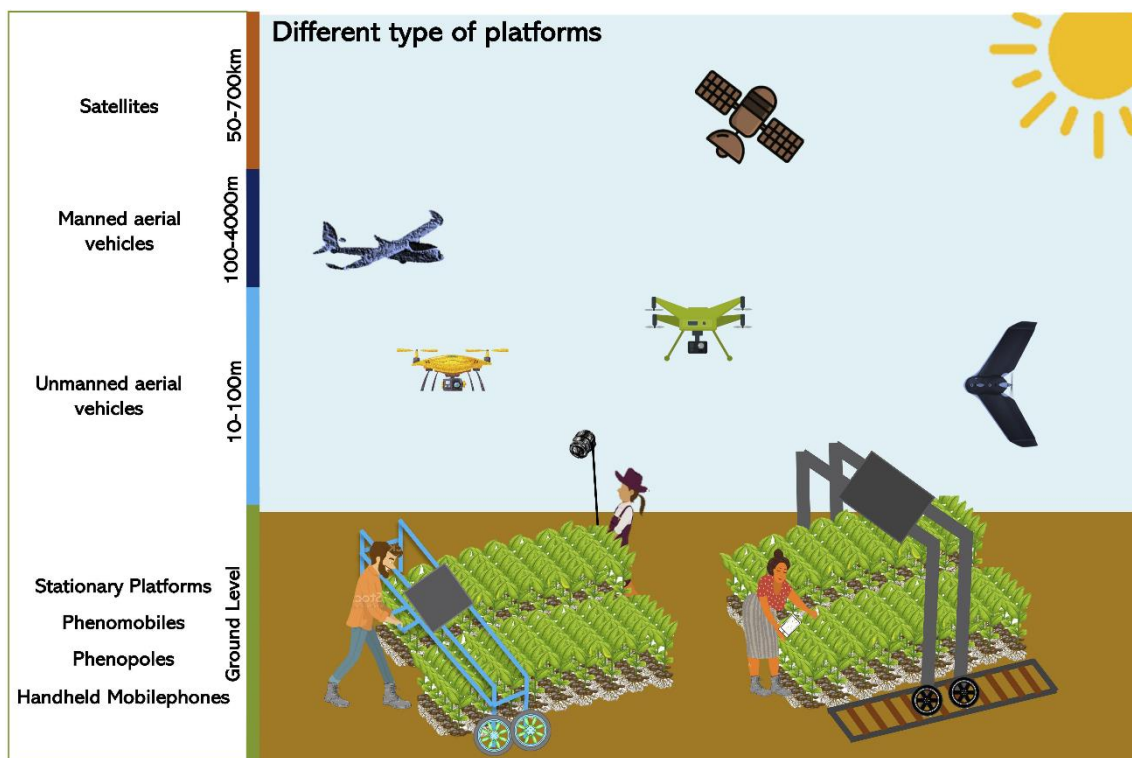


Fig. 3 Different Categories of Ground and Aerial Phenotyping Platforms. Ground level: these include from Handheld sensors (in this case just a person holding a mobile), to Phenopoles, Phenomobiles, and Stationary Platforms. From 10 to 100 m: Unmanned Aerial Vehicles, as drones of different sizes and compactness, fixed-wind drone. From 100 to 4000 m Manned

Aerial Vehicles as airplanes or helicopters. Different categories of satellites (Nanosatellite, Microsatellite and Satellites) from 50 to 700 km.

Space-based or orbital satellite sensors are not much used for crop phenotyping due to practical limitations related to the lack of free sub-meter resolution data (pixels are larger than the more standard phenotyping microplot sizes), and cloud cover; but they are of increasing interest in precision agriculture, crop classifications, and time-series phenology, etc. (Cheng et al. 2020; Segarra et al. 2020; Tsakmakis et al. 2021).

Furthermore, RS technologies provide coverage across a range of plant physiological processes from evapotranspiration and photosynthesis to phenology dynamics, productivity, and even quality traits such as those of a crop's harvestable component nutritional value. In conclusion, the spatial, temporal, and spectral resolutions of any sensor system are central to determining the scale(s) and precision(s) at which an instrument can monitor crops, and therefore the most adequate approach towards guiding agricultural monitoring.

Technological Readiness Levels

Technological Readiness Levels (TRL) are a set of management metrics that enable the assessment of the maturity of a particular technology (Fig.4). They provide a means for the consistent comparison of maturity levels between different types of technology all in the context of a specific system, application, and operational environment (Osinga et al. 2022). It is based on a scale from 1 to 9, with 9 being the most mature technology (EARTO 2014). The main objective of using TRLs is to help make decisions related to technology development. This concept was first developed at the National Aeronautics and Space Administration (NASA) during the 1970s for the space program and was subsequently formally adopted worldwide (Mankins 1995). In 2008, the scale was also used by the European Space Agency (ESA) (Kluk et al. 2008). And in 2013 the TRL scale was formalized through the ISO 16290 (EARTO 2014).

On the other hand, new opportunities for agriculture are emerging, as a result of the rapid development of communication networks and the availability of a wide range of remote and proximal sensors as mentioned above. (Wang et al. 2006; Aqeel-Ur-Rehman et al. 2014; Toth and Józków 2016). These new technologies can gather, process, and analyze data from the agro-environment (e.g., soil, crop, and climate), and, when combined with agro-climatic and economics models, technical interventions can be applied at the field level by either conventional means or automated solutions.

All these aspects are called "smart farming technologies" (SFTs) which represents the application of modern information and communication technologies (ITC) into agriculture (Walter et al. 2017; Bacco et al. 2019). Smart farming is based on precise and resource-efficient approaches and attempts to achieve higher efficiency in agricultural goods production with increased quality on a sustainable basis. Nevertheless, from the point of view of the farmers, it should also provide added value in the form of more accurate and timely decision-making and/or more efficient exploitation operations and management (Kerneckner et al. 2020). The SFTs can be divided into three main categories: farm management information systems (FMIS), precision agriculture (PA) systems, and agricultural automation and robotics.

Regarding PA, the general aim is at optimizing input use based on recording technologies to observe and measure inter- and intra-field spatial and temporal variability in crops, aiming to improve economic returns and reduce environmental impact (Balafoutis et al. 2020). PA is able to increase input efficiency for maintaining or even increasing production rates (Zhang et al. 2002; Schellberg et al. 2008; Balafoutis et al. 2017), using remote sensing technologies for data gathering with either satellite platforms for space imagery (Ge et al. 2011; Andrade-Sanchez et al. 2014) or aircrafts/UAVs for aerial applications (Zhang and Kovacs 2012; Mogili and Deepak 2018), combined use of sensors for ground data acquisition (Reyns et al. 2002), wireless networks for interconnecting them (Jawad et al. 2017), geospatial data analytics coming from different sources (Nash et al. 2009), decision support systems (DSSs) for optimized farming decision-making (Kuhlmann and Brodersen 2001; Lindblom et al. 2017), and others.

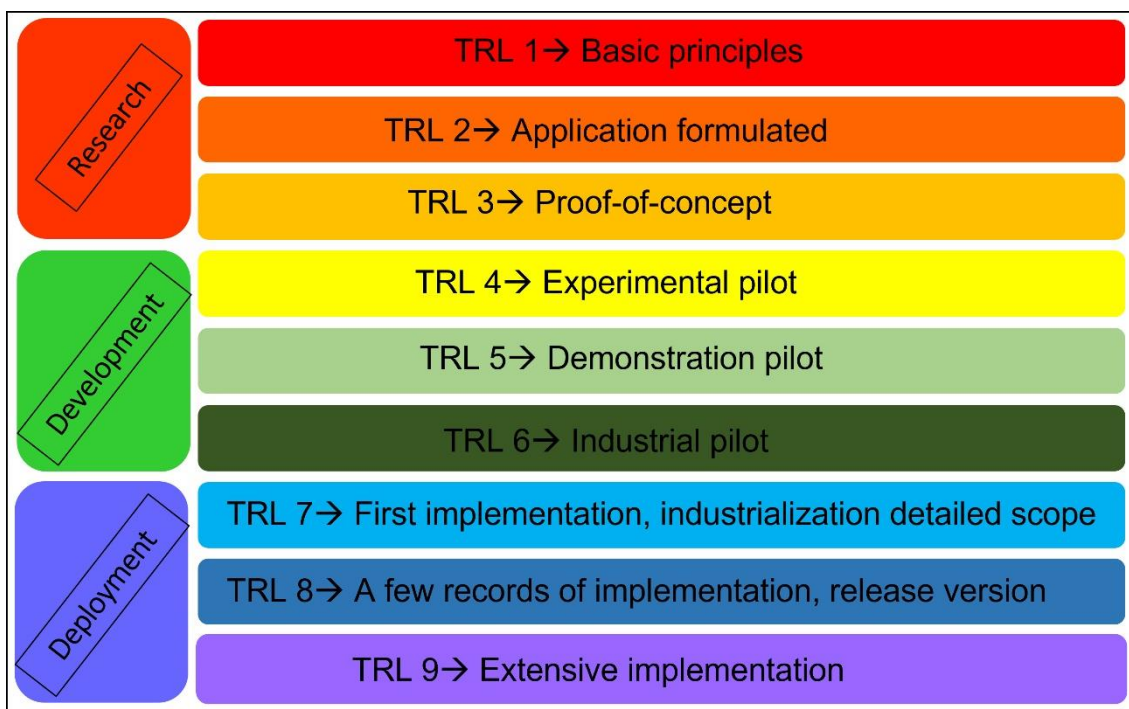


Fig. 4 The scale of Technological Readiness Levels (TRL) from 1 to 9 with the definitions.

At TRL 1, basic principles are observed and reported, and scientific research begins to be translated into applied research and development (R&D). In TRL 2, technology concepts and/or applications are formulated, practical applications can be invented, and research started. These applications are speculative and may be unproven. The TRL 3 tests analytical and experimental critical functions and/or characteristics as a proof-of-concept, with active R&D, initiated including analytical/laboratory studies to validate predictions regarding the applied technologies. These are the lowest categories of TRLs and relate to research, which applications that are experimental and have little more intention than to test a conceptual model or theory (1-3).

TRL 4 provides functional verification in the laboratory environment, where basic technological components are integrated to establish that they will work together, named the experimental pilot. TRL 5 is a demonstration pilot project, in which the components and validation are in the relevant environment. Where the basic technological components are integrated with reasonably realistic supporting elements so that they can be tested in a simulated environment. TRL 6 indicates a system/subsystem model or prototype demonstration in a relevant environment (ground or space). Where a representative model or a prototype system is tested in a relevant environment, industrial pilot. These categories consist of big data solutions that aim for development in the real world, but on a small scale, perhaps still in a controlled environment (levels 4-6).

TRL 7 is the first implementation with industrialization detailed scope. The prototype system is near/or at the planned operational system. Level 8 is when few records of implementation and the technology have been proven to work in its final form under expected conditions, release version. In the last TRL 9, extensive deployment, the system incorporating the new technology in its final form has been used under actual conditions. The highest TRL category is reserved only for applications that are operating in real-world environments and have all potential to be adopted on a large scale (levels 7-9) (Kluk et al. 2008; Maciejczak and Faltmann 2018; Pylianidis et al. 2021; Osinga et al. 2022).

As was mentioned before, the TRL indicates its maturity level and ranges from TRL1 (basic principles observed) to TRL 9 (actual system proven in operational environment). Based on this classification, the TRL was specified for the four chapters presented in the dissertation:

For the first chapter, can we choose the best performing cultivars and estimate maize grain yield under low nitrogen conditions utilizing VIs from an RGB camera at ground and UAV level in the first chapter? What TRL is it?

For the second chapter, can we develop an RS supplemental system to the FAO FAW monitoring and early warning systems (FAMEWS) mobile app for the second chapter? What TRL is it?

For the third chapter, using leaf spectral reflectance and advanced regression models, can we estimate the $V_{c,max}$ and J_{max} of soybeans and peanuts in the third chapter? What TRL is it?

For the fourth chapter, can we develop a system from data collection to recognition of diseases and nutrient deficiencies using RGB images and deep learning? Which TRL is it?



OBJETIVES

General aim

The evaluation of the remote sensing technology applied in the four distinct chapters and the comparison of technical maturity levels using the technological readiness levels (TRL) to advance the state of the art of remote sensing applications in international agricultural development.

Specific aims

Chapter 1. In this study we examined the potential of vegetation indices calculated from a RGB camera as an affordable HTPP tool to accurately phenotype commercial and pre-commercial maize genotypes under low- and optimal nitrogen fertilizer (N) conditions. Firstly, we evaluated the performance of a set of remote sensing RGB VIs from natural color images acquired at the ground level and from a UAV platform compared with the performance of the field based NDVI and SPAD sensors. Additionally, we evaluated how these different sets of plant phenotyping data contribute to improving multivariate model estimations of crop yield in combination with traditional agronomic field data, such as anthesis silking interval (ASI), anthesis data (AD), plant height (PH), and canopy senescence (SEN) in order to determine the level of improvements over traditional practices that they may provide. Field trials were carried out at the Southern Africa regional station of CIMMYT (International Maize and Wheat Improvement Center) located in Harare, Zimbabwe.

Chapter 2. The general aim of this study was to implement a cost-effective assessment for FAW Monitoring and Early Warning System (FAMEWS) on maize fields using different remote sensing technologies. We have divided this work into two parts: i) Development of satellite image-based monitoring algorithm through Google Earth Engine for Sentinel 2 a+b and FAO FAMEWS mobile application data. ii) Time series anomaly change detection and first derivative growth pattern analyses of NDVI using Sentinel 2 and Planet Scope image data during the maize vegetative growth stage, where FAW presence will result in a reduction of the LAI or total green biomass (NDVI) of the crop. Field trails were carried out on small farmers from three sub-Saharan African countries Zimbabwe, Kenya and Tanzania.

Chapter 3. The aims of this study were (i) to estimate photosynthetic capacity parameters, such as mid-day photosynthesis, leaf chlorophyll content (LCC), $V_{c,max}$, and J_{max} of two legume crops (soybean and peanut) using full-range leaf level reflectance spectra (VIS–NIR–SWIR, 400–2500 nm) with advance regression models and (ii) to simulate photosynthetic parameter model performance using four common types of sensors with more limited wavelength ranges: VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), and an advanced

multispectral sensor imitating the ESA Copernicus Sentinel 2 satellite with 12 spectral bands. Trails were carried in the University of Auburn and the USDA from Alabama, United States.

Chapter 4. The aim is to develop together four countries, Spain, the United Arab Emirates (UAE), Egypt, and Tunisia a user-friendly app for the automatic detection of plant diseases and disorders identification on tomato, cucumber, and pepper using project consortium local or cloud-based supercomputing. The study was divided into four parts: (I) dataset collection; (II) dataset curation; (III) development of the algorithm, and (IV) accessibility of the use of the CNN model.

REPORT DE THESIS DIRECTORS

REPORT OF THE THESIS DIRECTORS

Integrative Crop Ecophysiology Group

<https://integrativecropecophysiology.com>

Plant Physiology Section, Department of Evolutionary Biology, Ecology and Environmental Sciences, Faculty of Biology, University of Barcelona, Diagonal 643, 08028, Barcelona, Spain.
Tel. 934 021 46

Dr. Shawn C. Kefauver and Dr. José Luis Araus, as directors of the thesis entitled **“Multiscale and multi-sensor remote sensing in international agricultural development”** which was developed by the doctoral student Maria Luisa Buchailot, report about the impact factor and the participation of the doctoral student in the different chapters included in this doctoral thesis.

Chapter 1. This chapter is represented by the published scientific article. Buchailot, M.L., Gracia-Romero, A., Vergara-Diaz, O., Zaman-Allah, M.A., Tarekegne, A., Cairns, J.E., Prasanna, B.M., Araus, J.L., Kefauver, S.C., 2019. **“Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques.”** *Sensors* (Switzerland) 19. <https://doi.org/10.3390/s19081815>. *Sensors* MPDI in 2019 with an impact factor of 3.677, is a journal placed within the first decile of the Science Area: Agricultural and Biological Sciences: Plant Science. To date, this work has accumulated 56 citations (Google Scholar, revised in Nov. 22, 2022). In this study, we evaluated the performance of a set of remote sensing indices derived from Red-Green-Blue (RGB) images data as phenotypic traits and crop monitoring tools for assessment of maize performance under low nitrogen fertilization. This work reinforces the effectiveness of canopy remote sensing for plant phenotyping and crop management of maize under nitrogen nutrient conditions and suggests that the RGB indices are the best option. Maria Luisa Buchailot took a role participating in the field evaluation, further processing the information, and led in writing the draft of the manuscript.

Chapter 2. This chapter is represented by the published scientific article. Buchailot, M.L.; Cairns, J.; Hamadziripi, E.; Wilson, K.; Hughes, D.; Chelal, J.; McCloskey, P.; Kehs, A.; Clinton, N.; Araus, J.L.; Kefauver, S.C. **“Regional Monitoring of Fall Armyworm (FAW) Using Early Warning Systems.”** *Remote Sens.* 2022, 14, 5003. <https://doi.org/10.3390/rs14195003>. *Remote Sensing* MDPI, in October 2022, at which time the journal had an impact factor of 5.349 and a Cite score of 7.4 and is first quartile, Q1, in its fields of research. Maria Luisa Buchailot took a role participating in the field evaluation campaign in Tanzania, Kenya, and Zimbabwe. She wholly and

independently led the data curation, data processing, further statistical and graphical processing of all the relevant information for the study.

Chapter 3. This chapter is represented by the published scientific article. Buchailot, M.L.; Soba, D.; Shu, T.; Liu, J.; Aranjuelo, I.; Luis, J.; Runion, G.B.; Prior, S.A.; Kefauver, S.C.; Saez, A.S. *“Estimating peanut and soybean photosynthetic traits using leaf spectral reflectance and advance regression models.”* *Planta* 2022, 1–19. <https://doi.org/10.1007/s00425-022-03867-6>. *Planta* has an IF of 4.54 and which places it as Q1. Maria Luisa Buchailot led the field evaluation and independently managed all the data processing and statistical analyses of this study. This field data collection work was carried out during an exchange visit at the University of Auburn, Alabama in collaboration and supported in part by Professor Alvaro Sanz. Luisa was also fully responsible for the manuscript writing and preparation and the subsequent graphical and text revisions necessary for the submission and eventual publication of the manuscript in *Planta* on 24 March 2022.

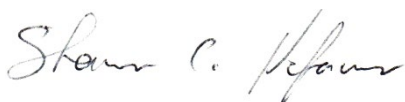
Chapter 4. The article *“Deep leaning for detection of plant disorders on crop leaves: from data collection to framework tools”*. Maria Luisa Buchailot took the lead role in the data gathering curation and further data processing and will be the first Autor of the article. This project was an important international collaboration with the International Center for Biosaline Agriculture (ICBA) of the United Arab Emirates in Dubai, together with other collaborative teams in Tunisia, Egypt, and Colombia as well. The project framework required a complex combined effort for data collection, which required extensive visits to each location for training exercises and conferences as well as coordination of research activities between parties in the large research consortium. Luisa organized all the input data from Northern Africa and the Middle East and curated it for statistical processing using supercomputing facilities through Google Colab. She has been fully responsible for the preparation of the text of the scientific manuscript resulting from this three-year collaborative research project. The manuscript has been prepared for submission to *Computers and Electronics in Agriculture*, a high IF 6.75 and Cite Score 11.8, is Q1 journal.

Other articles where the doctoral student participated as a co-author:

- Fernandez-Gallego, J.A., Buchailot, M.L., et al. “Cereal Crop Ear Counting in Field Conditions Using Zenithal RGB Images.” Article in *JoVE* 2019, URL: <https://www.jove.com/video/58695> <https://doi.org/10.3791/58695>.

- Fernandez-Gallego, J.A., Buchailot, M.L., et al. “Automatic wheat ear counting using thermal imagery” Article in Remote Sensing 2019, <https://doi.org/10.3390/rs11070751>.
- Segarra, J, Buchailot, M.L et al. “Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications” Review in Agronomy 2020, <https://doi.org/10.3390/agronomy10050641>.
- Araus J.L., Kefauver S.C., Vergara-Díaz O., Gracia-Romero A., Rezzouk F.Z., Segarra J., Buchailot M.L., Chang-Espino M., Vatter T., Sanchez-Bragado R., Fernandez-Gallego J.A., Serret M.D., Bort J., 2021. Crop phenotyping in a context of Global Change: what to measure and how to do it. J. Integr. Plant Biol. 2022 <https://doi.org/10.1111/jipb.13191>.
- Kamphorst, S.H., Amaral Júnior, A.T. do, Vergara-diaz, O., Gracia-romero, A., Fernandez-gallego, J.A., Chang-espino, M.C., Buchailot, M.L., Rezzouk, F.Z., de Lima, V.J., Serret, M.D., Araus, J.L., 2022. Heterosis and reciprocal effects for physiological and morphological traits of popcorn plants under different water conditions. Agric. Water Manag. 261, 1–14. <https://doi.org/10.1016/j.agwat.2021.107371>.
- Abrougui, K.; El, N.; Boughattas, H.; Belhaj, M.; Buchailot, M.L.; Segarra, J.; Amami, R.; Chehaibi, S.; Tarchoun, N.; Kefauver, S.C. 2022. Assessing Phytosanitary Application Efficiency of a Boom Sprayer Machine Using RGB Sensor in Grassy Fields. *Sustainability*, 14(6), 3666. <https://doi.org/10.3390/su14063666>.
- Hamdane, Y.; Gracia-Romero, A.; Buchailot, M.L.; Sanchez-Bragado, R.; Fullana, A.M.; Sorribas, F.J.; Araus, J.L.; Kefauver, S.C. Comparison of Proximal Remote Sensing Devices of Vegetable Crops to Determine the Role of Grafting in Plant Resistance to *Meloidogyne incognita*. *Agronomy* 2022, <https://doi.org/10.3390/agronomy12051098>.
- Berger, K., Machwitz, M., Kycko, M., Kefauver, S. C., Van Wittenberghe, S., Gerhards, M., Buchailot, M.L, .. & Schlerf, M. (2022). Multi-sensor spectral synergies for crop stress detection and monitoring in the optical domain: A review. *Remote sensing of environment*, 280, 113198.

To certify this for corresponding purposes,



Dr. Shawn C. Kefauver



Prof. Jose Luis Araus Ortega

CHAPTER 1

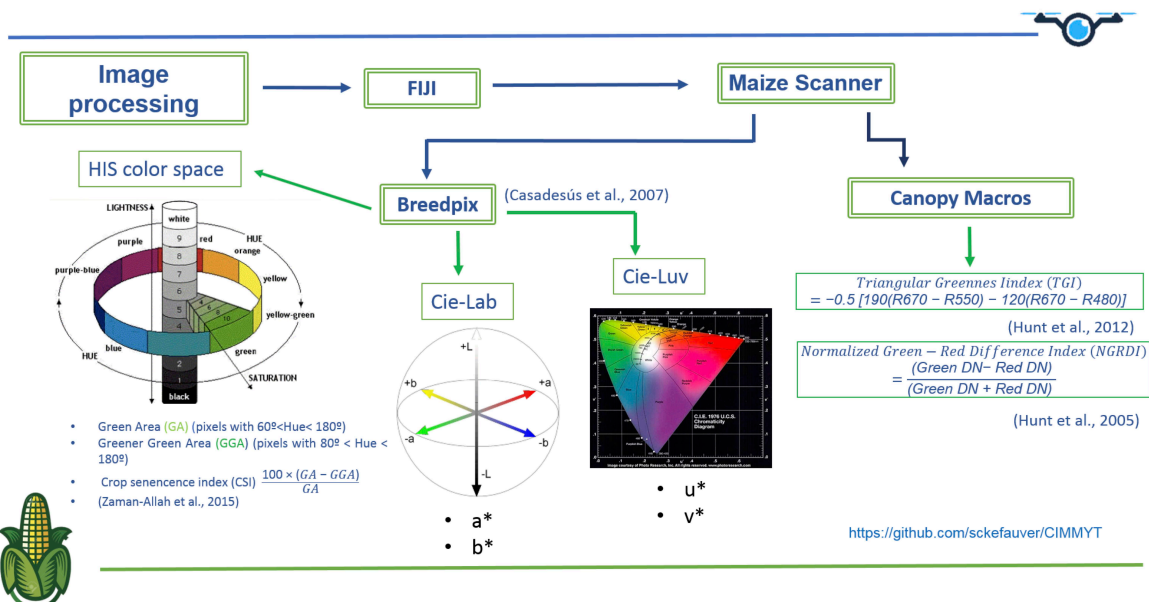


CHAPTER 1

Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques







Ma. Luisa Buchailot, Adrian Gracia-Romero, Omar Vergara-Diaz, Mainassara A. Zaman-Allah, Amsal Tarekegne, Jill E. Cairns, Boddupalli M. Prasanna, Jose Luis Araus and Shawn C. Kefauver

Published in:
Sensors (2019)



Article

Evaluating Maize Genotype Performance under Low Nitrogen Conditions Using RGB UAV Phenotyping Techniques

Ma. Luisa Buchailot ^{1,2}, Adrian Gracia-Romero ^{1,2}, Omar Vergara-Diaz ^{1,2},
Mainassara A. Zaman-Allah ³, Amsal Tarekegne ³, Jill E. Cairns ³, Boddupalli M. Prasanna ⁴,
Jose Luis Araus ^{1,2} and Shawn C. Kefauver ^{1,2,*}

- ¹ Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain; luisa.buchailot@gmail.com (M.L.B.); adriangraciaromero@hotmail.com (A.G.-R.); omarvergaradiaz@gmail.com (O.V.-D.); jaraus@ub.edu (J.L.A.)
- ² AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Spain
- ³ International Maize and Wheat Improvement Center, CIMMYT Southern Africa Regional Office, P.O. Box MP163 Harare, Zimbabwe; Z.mainassaraAbdou@cgiar.org (M.A.Z.-A.); a.tarekegne@cgiar.org (A.T.); J.Cairns@cgiar.org (J.E.C.)
- ⁴ International Maize and Wheat Improvement Center (CIMMYT), P.O. Box 1041 Nairobi, Kenya; b.m.prasanna@cgiar.org
- * Correspondence: sckefauver@ub.edu

Received: 8 January 2019; Accepted: 8 April 2019; Published: 16 April 2019



Abstract: Maize is the most cultivated cereal in Africa in terms of land area and production, but low soil nitrogen availability often constrains yields. Developing new maize varieties with high and reliable yields using traditional crop breeding techniques in field conditions can be slow and costly. Remote sensing has become an important tool in the modernization of field-based high-throughput plant phenotyping (HTPP), providing faster gains towards the improvement of yield potential and adaptation to abiotic and biotic limiting conditions. We evaluated the performance of a set of remote sensing indices derived from red–green–blue (RGB) images along with field-based multispectral normalized difference vegetation index (NDVI) and leaf chlorophyll content (SPAD values) as phenotypic traits for assessing maize performance under managed low-nitrogen conditions. HTPP measurements were conducted from the ground and from an unmanned aerial vehicle (UAV). For the ground-level RGB indices, the strongest correlations to yield were observed with hue, greener green area (GGA), and a newly developed RGB HTPP index, NDLab (normalized difference Commission Internationale de l'Éclairage (CIE)Lab index), while GGA and crop senescence index (CSI) correlated better with grain yield from the UAV. Regarding ground sensors, SPAD exhibited the closest correlation with grain yield, notably increasing in its correlation when measured in the vegetative stage. Additionally, we evaluated how different HTPP indices contributed to the explanation of yield in combination with agronomic data, such as anthesis silking interval (ASI), anthesis date (AD), and plant height (PH). Multivariate regression models, including RGB indices ($R^2 > 0.60$), outperformed other models using only agronomic parameters or field sensors ($R^2 > 0.50$), reinforcing RGB HTPP's potential to improve yield assessments. Finally, we compared the low-N results to the same panel of 64 maize genotypes grown under optimal conditions, noting that only 11% of the total genotypes appeared in the highest yield producing quartile for both trials. Furthermore, we calculated the grain yield loss index (GYLI) for each genotype, which showed a large range of variability, suggesting that low-N performance is not necessarily exclusive of high productivity in optimal conditions.

Keywords: maize; nitrogen; phenotyping; remote sensing; Africa; RGB; UAV; CIELab

1. Introduction

Maize is the most commonly cultivated cereal in Africa in terms of land area and production [1]. Low yields in this region are largely associated with drought stress, low soil fertility, weeds, pests, diseases, low input availability, low input use, and inappropriate seeds [2]. After water, nitrogen (N) is the single most important input for maize production, and the lack of N is considered to be the principal constraint to cereal yields in areas with more than 400 mm of average annual rainfall in Sub-Saharan Africa (SSA) [3], but fertilizer application in SSA is negligible, accounting for less than 1% of the global N fertilizer application [4]. As such, efforts to increase maize production capacities in low fertilizer conditions may contribute substantially to improving food security and well-being in the region [5]. One of the strategies considered for increasing maize yield with regards to N inputs in SSA is breeding to improve yield under nutrient deficiency or towards specific adaptation to increase performance under low-nitrogen conditions. Furthermore, the adaptation of maize to lower fertilizer conditions may improve agricultural economics at equal or even better levels of production with lower required inputs, less runoff, and resource extraction that may additionally result in reducing environmental degradation and the loss of ecosystem services [6–8].

Plant scientists, especially breeders and agronomists, face the challenge of solving these limitations while considering the additional implications of climate change on food security [2,9]. In that sense, affordable technologies capable of monitoring crop performance, improving yield prediction, or assessing phenotypic variability for breeding purposes are aimed at surpassing the bottlenecks in the way of full exploitation of this technology [10,11]. One of the first non-destructive and analytical tools was the chlorophyll meter, based on radiation absorbance by leaves in the red and near-infrared regions (usually at 650 and 940 nm). These leaf level relative chlorophyll content readings have an indirect and close relationship with leaf N and total chlorophyll concentrations [12,13]. Portable meters have been used for some time on crops as a fairly quick and reliable method for N management [14,15], but this technique is relatively slow compared with newer imaging techniques and does not include the whole plot, thereby capturing less variability than full canopy remote sensing techniques [16].

Remote sensing has become an important tool in the modernization of field-based high throughput plant phenotyping (HTPP), including improvements in yield potential, adaptation to abiotic stressors (drought, extreme temperatures, salinity), biotic limiting conditions (susceptibility to pests and diseases), and even quality traits [5,10,17]. Traditionally, the primary platforms used to obtain remote images of the Earth's surface were satellites and piloted aircrafts, but these instruments generally do not deliver data at adequate spatial and temporal resolutions necessary for more detailed agricultural applications, such as plant phenotyping [18]. Currently, these limitations can be overcome using more flexible unmanned platforms, such as unmanned aerial vehicles (UAVs), also called remotely piloted aircraft systems (RPASs) or unmanned aircraft systems (UASs) [19,20]. UAVs allow for many quick, precise, and quantitative observations at improved spatial and temporal resolutions and at lower costs with respect to airborne platforms or satellites.

The classical approach of remote sensing platforms, including UAVs, has involved the use of multispectral sensors and the calculation of different vegetation indices associated with plant physiological parameters, such as plant pigments, vigor, and above-ground biomass. In this sense, visible and near-infrared (VNIR) imaging spectroscopy has demonstrated a fairly reliable capacity in biophysical crop assessments in agriculture [21–27]. For example, the normalized difference vegetation index (NDVI) [28] is a well-known, broadband vegetation index derived from visible and near-infrared reflectance that is closely related to vegetation presence or vigor [29,30]. It can also be measured at the ground level with active portable sensors (e.g., GreenSeeker). Other examples use narrow-band reflectance values for more precise measurements [31] and are often found to be correlated with grain yield and provided reliable information for yield forecasting [32] or specific biophysical properties, such as canopy water content [33] or photosynthetically active pigments, [23,34,35] but require more advanced sensor technologies for adequate quantification. Similarly, thermal infrared (TIR) imaging

enables rapid remote observations of plant water status via their cooling capacity and stomatal conductance [36,37].

As a low-cost alternative, and at an order of magnitude less expensive than scientific multispectral VNIR or TIR sensors, various red–green–blue vegetation indices (RGB VIs), calculated from commercial RGB cameras, have demonstrated their ability to predict grain yield, quantify nutrient deficiencies, and measure disease impacts [38,39]. With respect to these commercially available RGB cameras, color calibration quality should be assessed prior to scientific use for checking and/or correcting variations in RGB color values, as illumination conditions may influence the accuracy of color reproduction [40,41]. On the other hand, when these types of camera are used for producing multi-image mosaics, within image vignetting should also be assessed, as brightness attenuates away from the image center and appears as artifacts in the image mosaics [42]. Still, RGB VIs can accurately quantify different properties of color and have often demonstrated performance levels similar to or better than NDVI [39]. RGB images can be processed using comparisons between red, green, and blue light broadband reflectance values or through the use of alternate color spaces, as with the Breedpix code suite [43]. The treatment of R, G, and B as separate spectral bands allows for the calculation of the triangular greenness index (TGI), which estimates chlorophyll concentration in leaves and canopies [1], and the normalized green–red difference index (NGRDI), which compares the differences between the green and red bands in a calculation similar to NDVI but with less marked differences and less signal saturation. In the hue–saturation–intensity (HSI) color space, where the hue (H) component describes color chroma traversing the visible spectrum in the form of an angle between 0° and 360° . Thus, the index green area (GA) is the percentage of pixels in the image in the hue range from 60° to 180° , ranging from yellow to bluish green, while the greener green area (GGA) includes a more restrictive range of hue from 80° to 180° , excluding yellowish-green tones that might be partially stressed or senescent. Hence comes the crop senescence index (CSI), which combines GA and GAA to provide a strong discrimination between tolerant and susceptible genotypes in various treatments [44,45].

To the end of better quantifying leaf pigment loss, and therefore color changes, due to nitrogen deficiency [46,47], further investigation of the capacities and techniques for accurate color quantification using digital images indicates that there are newer more advanced color models currently in use by photography professionals. In the Commission Internationale de l'Éclairage (CIE), CIELab color space model, dimension L^* represents lightness; the a^* component expresses green to red, with a more positive value representing red, and a more negative value indicating green; and the b^* component expresses blue to yellow, in which positive values are towards yellow, and negative values are closer to blue. Correspondingly, in the CIEluv color space model, dimensions u^* and v^* are perceptually uniform coordinates, where L is again lightness and u^* and v^* represent axes similar to a^* and b^* in separating the color spectrum, respectively. For more specific details on the development of these alternate color space RGB indices and their respective transformations, please see [48]. Both CIELab and CIEluv include color calibration corrections through the separation of the color hue from the illumination components of the input RGB signal; for that reason, we have developed two new vegetation indices using these color spaces in a way similar to the conceptual basis for NDVI, using the normalized difference between a^* and b^* (NDLab) and the normalized difference between u^* and v^* (NDLuv). Thus, the CIElab and CIEluv color spaces offer the ability to simultaneously contrast green vegetation quantity with both the reddish/brown soil background (fractional vegetation cover or plant growth) and yellowing caused by chlorosis (loss of foliar chlorophyll)—both common symptoms of nitrogen deficiency. Previously, RGB VIs have been employed at both the canopy and at the leaf levels for precise crop management or as effective HTPP techniques in breeding programs aimed to improve crop performance under a wide range of conditions [10].

In the research presented here, the RGB VIs described above, namely hue, a^* , b^* , GA, GGA, NGRDI, and the new NDLab and NDLuv, are examined for their potential as affordable HTPP tools to accurately phenotype commercial and pre-commercial maize genotypes under low- and optimal-N conditions. Firstly, we provide some maize genotype performance comparisons between the low-N

and optimal growing conditions in order to provide some initial insights on the potential of selecting for low-N-adapted maize genotypes. Then, we evaluate the performance of a set of remote sensing RGB VIs from natural color images acquired at the ground level and from a UAV platform compared with the performance of the field-based NDVI and SPAD sensors. Additionally, we evaluated how these different sets of plant phenotyping data contribute to improving multivariate model estimations of crop yield in combination with traditional agronomic field data, such as anthesis silking interval (ASI), anthesis data (AD), plant height (PH), and canopy senescence (SEN) in order to determine the level of improvements over traditional practices that they may provide.

2. Materials and Methods

2.1. Plant Material and Growing Conditions

Field trials for managed low-nitrogen and optimal fertilizer conditions were conducted at the International Center for Maize and Wheat Improvement (CIMMYT) regional station located in Harare, Zimbabwe ($-17,800$ S, $31,050$ E, 1498 m.a.s.l.) (Figure 1). The soil of the station is characterized by a pH slightly below 6, with low managed nitrogen (LOW) treatment for all plots at 25–35% less N compared with the optimal standard fertilization application of 200 kg/ha, here defined as the optimum nitrogen (OP) according to established standard CIMMYT protocols [17]. A set of 49 new maize genotypes that were developed at CIMMYT and 15 commercial maize genotypes in Zimbabwe were selected for the study (Table A1). Seeds were sown during the wet season, on 16 December 2015, in two rows per plot; the rows were 4 m long and 75 cm apart (5.25 m²/plot), with 14 planting points per row and 25 cm between the plants within a row. The experiment was carried out in 192 plots with 3 replicates per variety. Both trials were rainfed only, being grown in the Zimbabwe rainy season, with local weather station data recording growing season mean temperature, humidity, and total rainfall of 26°C , 68% , and 700 mm, respectively, effectively eliminating any chances of water stress even without irrigation.

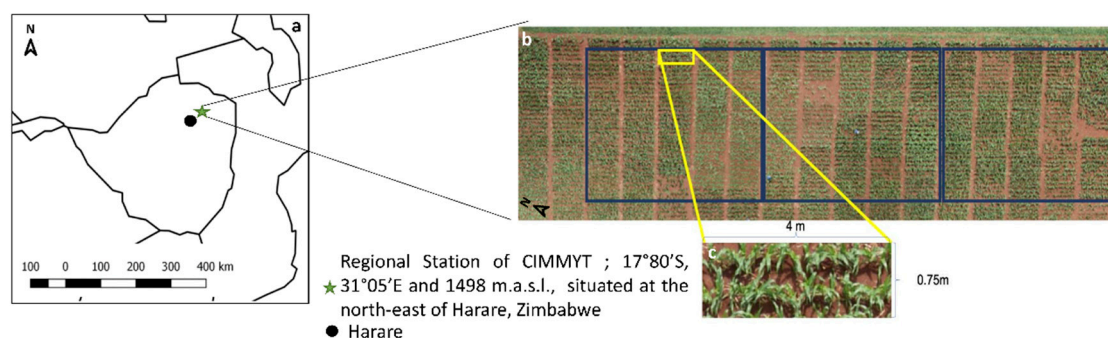


Figure 1. (a) Position of the regional station of International Center for Maize and Wheat Improvement (CIMMYT) in Harare, Zimbabwe. (b) Red–green–blue (RGB) aerial orthomosaic of the 192 plots, with 64 genotypes and 3 replicates (three blue box) per each one, under low managed nitrogen (LOW). (c) A plot with specific details of length and width.

The trials were harvested in mid-May of 2016, discarding 2 plants at each row end and harvesting the central 3.5 m of each row in order to reduce edge effects. Thus, the total harvested weight corresponded to an area of 5.25 m² (0.75 m apart \times 2 rows \times 3.5 m long), consisting of the same number of plants per plot (excepting locations of mortality). The cobs processed, and grains dried to approximately 12.5% moisture, such that grain yield (GY, t·ha⁻¹) was calculated as follows, where X is the grain weight per plot:

$$GY = (X \text{ (kg/plot)} * 10) / (5.25 \text{ m}^2) \quad (1)$$

The grain yield loss index (GYLI) as the stress index was calculated as:

$$GYLI = (GY \text{ at OP} - GY \text{ at LOW}) / (GY \text{ at OP}) \times 100 \quad (2)$$

where GY at OP represents the potential grain yield in optimum-nitrogen conditions and GY at LOW corresponds to grain yield in low managed nitrogen conditions [38].

2.2. Agronomic Parameters

PH was measured on 19 February 2016 as the length from the soil surface to the base of the tassel (excluding tassel length) using a ruler [49,50] on two representative plants per plot before all plants were hand harvested and grain yield was assessed. ASI was determined by the number of days from sowing until 50% of plants extruded anther AD and the number of days from sowing until 50% of the plants show silks (silking date, SD), such that $SD - AD = ASI$. SEN was measured visually on a plot basis as the proportion of green leaves 2–5 weeks after anthesis on a 0–100 scale, where 0 = 0% canopy senescence and 100 = 100% canopy senescence (Equation (3)). That technique is based on the different color classes, and given that any part of a leaf with yellow or brown (dry) color was classified as undergoing or having succumbed to senescence, a senescence index was proposed as the ratio between senesced canopy and the total canopy cover:

$$SEN = (YC + DC)/(YC + DC + GC) \quad (3)$$

where GC is green canopy cover, YC is yellow canopy cover, and DC is dry canopy cover [17,44,51,52]. This was measured 4 times during the experimental trial, but only the last SEN measurements from 5 April 2016 were used in this study.

2.3. Proximal and Aerial Data Collection

RGB remote sensing evaluations were performed on young maize plants (less than 5 leaves) on 28 January 2016, during the last week of January. For ground RGB VIs, vegetation indices were derived from one picture taken at the ground level for each plot (covering 40–50% of each plot), and UAV RGB VIs were derived from whole plot coverage from the UAV RGB aerial image mosaic of the whole study area as shown in Figure 2. At the ground level, one digital photograph was taken per plot with an Olympus OM-D E-M10 Mark III (Olympus, Tokyo, Japan), holding the camera at about 80 cm above the plant canopy in a zenithal angle and focused near the center of each plot. The images were acquired with a resolution of 16 megapixels with a Micro Four Thirds (M4/3) Live MOS sensor with a focal length of 14 mm, at a speed of 1/125 s with the aperture programmed in automatic mode at a resolution of 4608×3072 for a Ground Sample Distance (GSD) of 0.03 cm/pixel. RGB aerial images were acquired using an UAV (Mikrokopter OktoXL, Moormerland, Germany) flying under manual remote control at 50 m a.g.l. (altitude above ground level). The digital camera used for aerial imaging was a Lumix GX7 (Panasonic, Osaka, Japan), mounted on a two-axis gimbal with vibration reducers for stable image capture while in flight. Images were taken at a 16-megapixel resolution of 4592×2448 pixels using a 4/3" sensor and a 20 mm focal length lens for an estimated GSD 0.9419 cm/pixel. These images were taken with a 1/160 second shutter speed and auto-programmed mode for maximum aperture at a rate of every 2 s for the duration of the flight and stored locally on microSD cards for subsequent processing.

The measurements of the color calibration check and the vignetting calibration were taking the same day of the data collection. We used the ColorChecker Passport Photo (X-Rite, Inc. <https://www.xrite.com/es/categories/calibration-profiling/colorchecker-passport-photo/>), which has a panel of 24 industry standard color reference chips with published values in RGB, as well as the CIE Lab color space. The photos of this passport were taken with the cameras Olympus OM-D and Lumix GX7 in natural light conditions in a zenithal plane. With the software FIJI (Fiji is Just ImageJ, <https://fiji.sc/>, <https://imagej.nih.gov/ij/>), the calibration photos were imported and divided into the separate color channels of red, green, and blue and in the CIE Lab color space as lightness, a^* and b^* and then compared with the 24 published reference values of each standard chip with the photos of the passport taken with the different cameras.

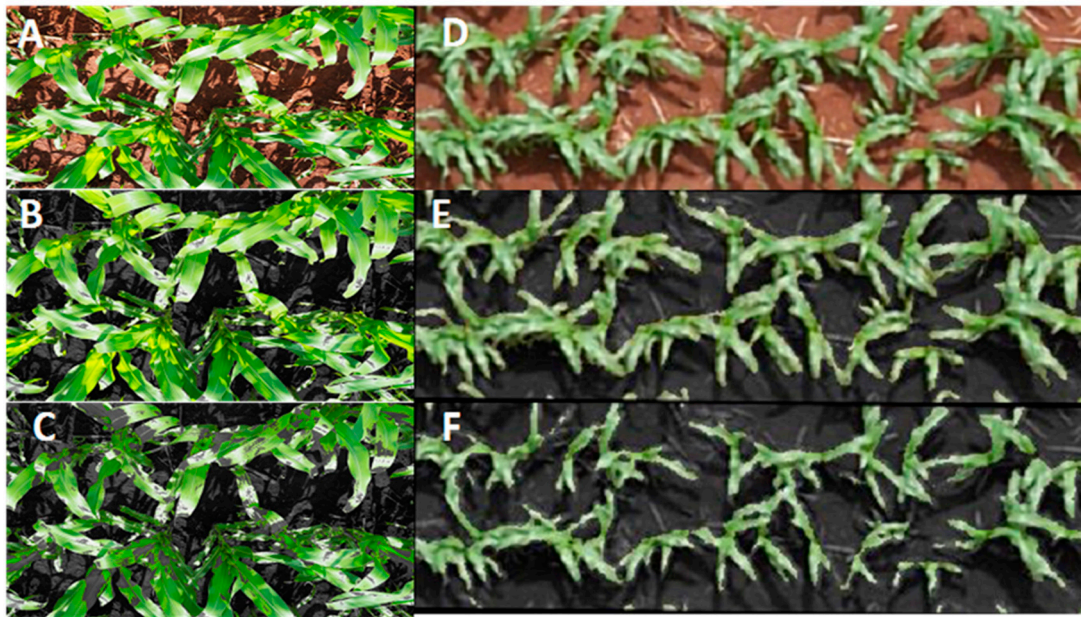


Figure 2. Examples of the differences in resolution between the images of maize taken at ground level (Ground Sample Distance 0.03 cm/pixel) and aerial level (Ground Sample Distance 0.9419 cm/pixel) in LOW. (A) Maize ground image from 80 cm or canopy level image averages. (B) Maize ground image from 80 cm showing green area (GA). (C) Maize ground image from 80 cm with greener green area (GGA). (D) Maize aerial image from 50 m or canopy level whole plot averages. (E) Maize aerial image from 50 m showing GA. (F) Maize aerial image from 50 m showing GGA.

With respect to vignetting calibration, one photo that was taken with the Lumix GX7 at 50 m was divided into the separate RGB and CIE Lab color space channels. On the R, G, B, hue, a^* , and b^* single band images, a line was drawn through the center for the X and Y axes in order to extract the cross-image transect. Then, a filter was created using the hue band from the HSI color space in order to select only sunlight soil pixels and applied for R, G, B, a^* , and b^* to eliminate vegetation and shadowed pixels, and the digital numbers (DN) were extracted from each remaining point along the line in order to observe changes in albedo across the image axes.

NDVI was measured on 28 January 2016 (at the same time as the RGB data) with the GreenSeeker active field sensor (GreenSeeker handheld crop sensor, Trimble, Ukiah, CA, USA), which uses a wavelength range of 650–670 nm and 765–795 nm for red and near-infrared, respectively. Additionally, SPAD chlorophyll meter (Minolta SPAD-502, Spectrum Technologies Inc., Plainfield, IL, USA) measurements were recorded on two different dates (at 3 and 5 weeks after the RGB and NDVI data), once on 18 February 2016 (SPAD vegetative stage, SPAD^V) and then again on 1 March 2016 (SPAD reproductive stage, SPAD^R). A total of 4 leaves were measured for each row for a total of 8 measurements per plot to provide a representative average value for each plot. Delayed SPAD sensor timing was due to availability and has been included for sensor technique as well as data capture timing comparisons. Different measurement timing details for the complete study are presented in Figure 3 for added clarity.

2.4. Image Processing

For the RGB images captured from the UAV platform, Agisoft PhotoScan Professional software (Agisoft LLC, St. Petersburg, Russia) was employed using a total of 63 overlapping images to produce an accurate image mosaic with at least 80% overlap, and this presented a resolution of 11772×4932 , as seen in Figure 1. As the aerial images were acquired in clear sky conditions at the same time as the ground RGB images, no cross-calibration radiometric corrections were deemed necessary. The open source image analysis platform FIJI [53] (Fiji is Just ImageJ; <http://fiji.sc/Fiji>) was used to segment

regions of interest for each row for the plots to be cropped in order to produce a single micro image per plot. RGB pictures were subsequently analyzed using a version of Breedpix 0.2 software adapted to JAVA8 and integrated as part of the MaizeScanner, an open-source and open access FIJI plugin that also provides for the implementation of TGI and NGRDI, as well as some specific analyses for maize research-related maize lethal necrosis impact quantification (<https://github.com/sckefauver/CIMMYT>).

Within FIJI, images were processed to convert RGB values into indices based on RGB broadband reflectance and also for color quantification from the HSI, CIELab and CIELuv color spaces. The TGI is calculated as the area of a triangle from the matrix determinants after factoring the terms:

$$A = \pm 0.5 [(\lambda_1 - \lambda_3) \times (R_1 - R_2) - (\lambda_1 - \lambda_2) \times (R_1 - R_3)] \quad (4)$$

where A is the triangular area; λ_1 , λ_2 , and λ_3 are the center wavelengths for the three image bands; and R1, R2, and R3 are reflectance values for the three image bands, respectively. The order of bands is not important, but the order will affect whether the result is positive or negative. Starting with R1 as R670 (red), R2 as R550 (green), and R3 as R480 (blue) for convenience:

$$TGI = -0.5 [190 \times (R_{670} - R_{550}) - 120 \times (R_{670} - R_{480})] \quad (5)$$

where TGI has units of wavelength \times reflectance, so using nm wavelength units or percent reflectance does not affect the value of TGI after the units are converted. We used digital camera bands of red, green, and blue broadband reflectance centered approximately at 670, 550, and 480 nm, respectively, so that λ_1 , λ_2 , and λ_3 were the centers of the wavebands, and R1–R3 were the waveband reflectance values [54].

We used the NGRDI to analyze the images from the digital camera:

$$NGRDI = (R_{550} - R_{670}) / (R_{550} + R_{670}) \quad (6)$$

where R550 and R670 are the reflectance values of the green and red bands of the RGB camera, respectively. The difference between green and red light reflectance differentiates well between plants and soil due to the absorption of chlorophyll at R670, and the sum normalizes for variations in light intensity resulting in a possible range from -1.0 to 1.0 , with NGRDI values mostly between -0.2 and 0.5 , ranging from soil to healthy vegetation [55].

As described previously, the HSI color space index GA is calculated as the percentage of pixels in the hue range from 60° to 180° , including from yellow to bluish green, while the GGA includes a more restrictive H range from 80° to 180° , excluding yellowish-green tones that might be partially stressed or senescent. Subsequently, the CSI was calculated in agreement with [38,44] as follows:

$$CSI = 100 \times (GA - GGA) / GA. \quad (7)$$

In addition, we developed two new different vegetation indices, modeled after NDVI, such that values of soil fall closer to 0 and vegetation closer to 1. In order to do so, because the a^* and the u^* image values for green are both negative, those values were placed first but using the complement of a^* so that greener vegetation gives a higher value, as would the near-infrared of NDVI. As b^* and v^* both have more yellowish values with higher values, no inversion was necessary [38,56–58]. The normalized difference between a^* and b^* (NDLab) through the color space CIELab is as follows:

$$NDLab = (((1 - a^*) - b^*) / ((1 - a^*) + b^*) + 1). \quad (8)$$

The normalized difference between u^* and v^* (NDLuv) through the color space CIELuv is as follows:

$$NDLuv = (((1 - u^*) - v^*) / ((1 - u^*) + v^*) + 1). \quad (9)$$

By inverting a^* and u^* , more green vegetation becomes a positive contribution to the index, while more red/brown soil background reduces the index value. Then, dividing by b^* and v^* , an increase in yellow chlorotic vegetation will reduce the index. The addition of 1 provides for a more balanced equation for positive values for crops from NDLab and ND Luv using CIELab and CIELuv; normalization then limits the index to values between -1 and 1 , with most crops between 0 and 1 .

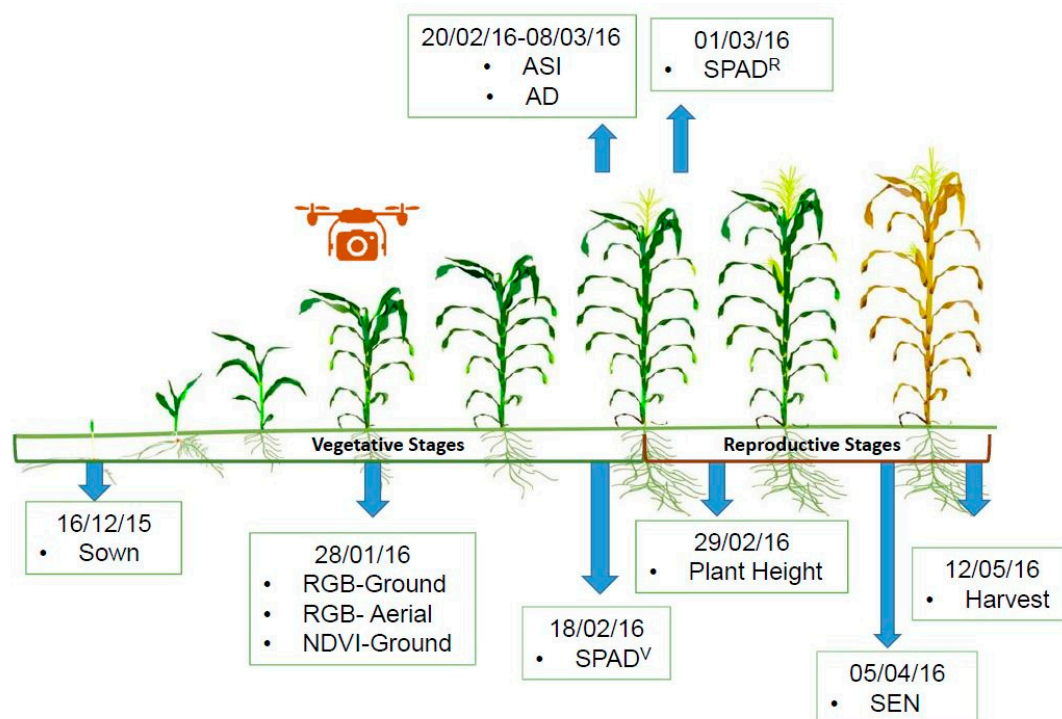


Figure 3. Field phenotyping, field imaging, and unmanned aerial vehicle (UAV) aerial image data capture chronogram for the controlled low-N field trial, showing dates for the measurement of all parameters at vegetative and reproductive stages, including red–green–blue (RGB), (high throughput plant phenotyping (HTPP) imaging, plant height (PH), canopy senescence (SEN), anthesis data (AD), and anthesis silking interval (ASI).

2.5. Statistical Analysis

Statistical analyses were conducted using the R project for statistical computing [59] in combination with R studio [60]. The maize crop physiological traits were analyzed using ANOVA and Fisher's Least Significant Difference (LSD) tests ($\alpha = 0.05$) in order to test the effects of growing conditions on the different traits. The results of the canopy level image averages per picture taken at the ground level were compared with the canopy level whole plot averages of the UAV images (Figure 2) with Pearson correlation coefficients and ANOVA analyses. Pearson correlation coefficients of the different remote sensing indices were additionally compared against grain yield. Multiple regressions were calculated with GY as the dependent variable and the different indices as independent variables using forward stepwise methods with the stepAIC () function of the MASS R package. The figures were also drawn using the R studio software.

3. Results

3.1. The Effect of Optimal Condition and Low Managed Nitrogen on grain yield

The range of yield in the LOW treatment was between 1.53 tn/ha and 4.43 tn/ha, while for OP it ranged between 6.68 tn/ha and 12.30 tn/ha; the GYLI range was from 46.88% to 85.22% (Table 1). On the other hand, significant differences in GY between genotypes were observed in this study for

the two different conditions (Table A1), but in order to standardize for comparisons between the two treatments, we divided the genotypes into quartiles by yield. Therefore, in Figure 4 the results show the 64 genotypes divided in quartiles as high yield (HY), medium high yield (MHY), medium low yield (MLY), and low yield (LY). The ANOVA for the OP and LOW treatment demonstrated that there were significant differences in GY between all of the quartiles of genotypes.

Table 1. Minimum, maximum, and average of grain yield (GY) and percentage of Nitrogen (N) of the two different treatments: LOW and optimum nitrogen (OP). Minimum, maximum, and average of grain yield loss index (GYLI).

	Minimum	Maximum	Average	N (%)
GY (Mg/ha) at LOW	1.53	4.43	2.93 ± 0.58	25–35
GY (Mg/ha) at OP	6.68	12.30	9.62 ± 1.24	100
GYLI (%)	46.88	85.22	69.01 ± 7.48	

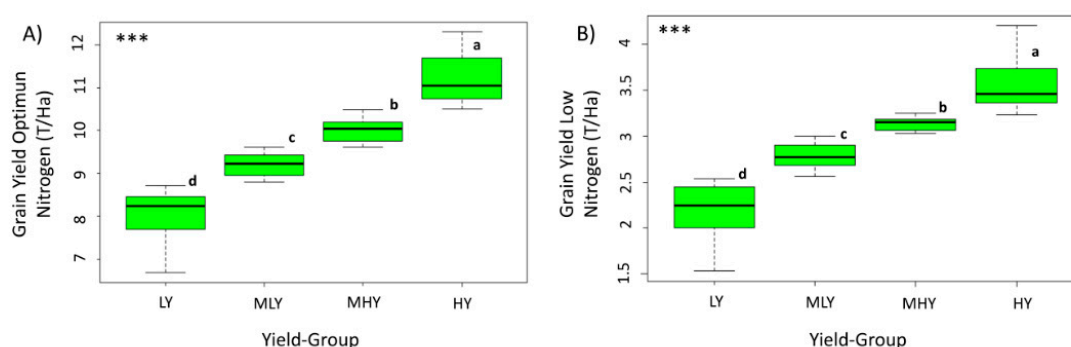


Figure 4. Box plot of grain yield for the set of 64 genotypes divided into four quartiles: low yield (LY), medium low yield (MLY), medium high yield (MHY), and high yield (HY) under OP (A) and LOW (B) conditions. The bottom and top of the box are lower and upper quartiles, respectively. The band near the middle is the median value across each group, and the bars are the standard deviation. Letters are significantly different according to Fisher Least Significant Difference (LSD) multiple range test ($P < 0.01$).

In Table 2, the results additionally demonstrate that 44% of the genotypes that belonged in the top HY group grown under OP conditions remained in the HY group in the LOW condition. Similarly, 19% of the genotypes in the OP condition LY group were also in the LOW condition LY group. This suggests that while high yield under both low-N and optimal-N conditions is not completely exclusive, previous breeding efforts have perhaps been more focused on yield in optimal conditions without considering the robust performance of a genotype in other potential growing conditions (i.e., low-N).

Table 2. Maize genotypes that were in both, the HY and LY groups with different applications of nitrogen: LOW and OP, with their GY.

Genotype	LOW		OP	
	GY(Mg/ha)	Yield Group	GY(Mg/ha)	Yield Group
CZH128	3.35	HY	10.93	HY
CZH15024	3.50	HY	11.13	HY
CZH15028	3.88	HY	10.60	HY
CZH15045	3.38	HY	10.13	HY
CZH15057	3.82	HY	11.40	HY
11C4393	3.37	HY	10.97	HY
LOCAL CHECK2	2.44	HY	10.59	HY
CZH15027	2.32	LY	8.44	LY
PHB30G19	2.28	LY	8.46	LY
MRI 634	2.29	LY	7.97	LY

3.2. The Performance of Remote Sensing Indices and Field Sensors

3.2.1. Color and Vignetting Calibration

The results shown in Figure 5 demonstrate that all the color calibration correlation R^2 were higher than 0.80, with most falling close to 0.90. With respect to the Lumix GX7, we can see the highest determination coefficient was the green channel, followed by the b^* ; for the Olympus Camera the highest was b^* , followed by the blue channel. While the L values from the CIELab color space were among the furthest from the 1:1 comparison line with high y-axis intercepts, the associated a^* and b^* linear correlations were among the closest to a 1:1 ratio with intercepts close to 0.

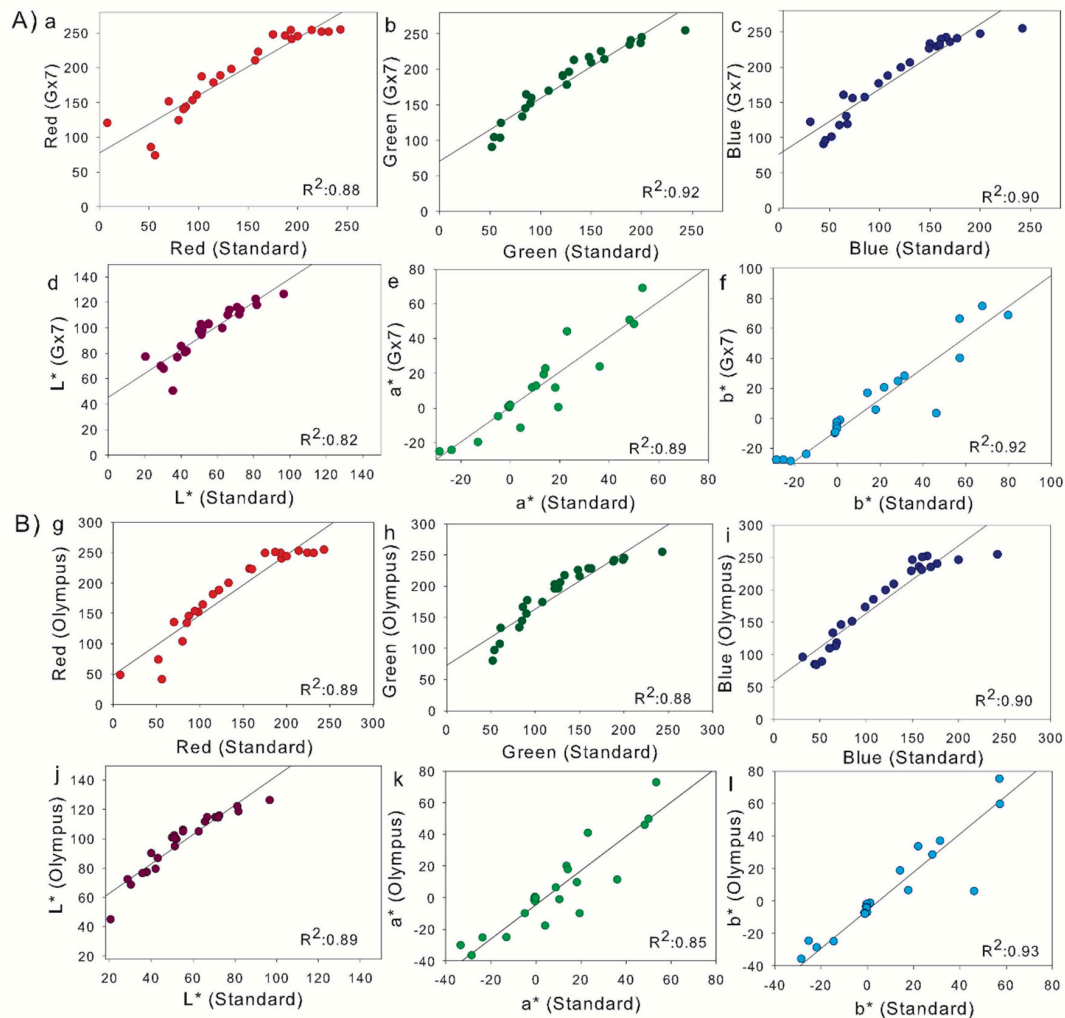


Figure 5. Color calibration through ColorChecker Passport. (A) The determination coefficients between the photo taken with Lumix GX7 camera for each channel red, green, and blue (a–c) and Commission Internationale de l’Éclairage (CIE) Lab color space (d–f) ($n = 24$). (B) The determination coefficients between the photo taken with the Olympus OM-D camera for each channel red, green, and blue (g–i) and CIE Lab color space (j–l) ($n = 24$).

Figure 6 shows that the vignetting effects observed for the both color spaces with respect to the y- and x-axes were minimally present for the RGB color spaces and reduced for a^* and b^* in CIE Lab.

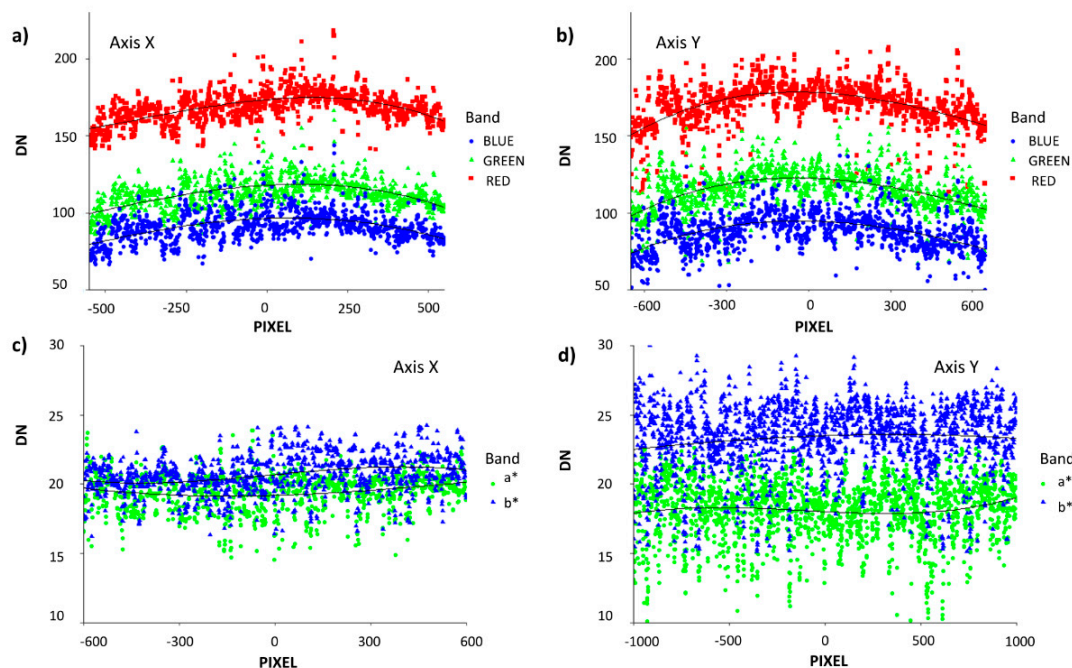


Figure 6. Vignetting effects observed in the RGB and CIELab color spaces from one example photo taken with Lumix GX7 from the UAV at 50 m, comparing x-axis and y-axis transect pixel digital number (DN) values. Hue-based color filters were used to identify only bare soil pixels along each transect. RGB channels (a,b). CIELab color space without lightness (c,d).

While the calibration check for the RGB images taken from both cameras demonstrated high correlations (Figure 5), the result of applying the calibration coefficients to the data resulted in both cases in lower correlations between the vegetation indices at different scales, as well as lower correlations between the RGB indices with GY (data not shown). Moreover, the results demonstrated low presence of vignetting effects, with reduced vignetting in the luminescence-controlled CIELab and CIELuv color spaces of particular interest for this study (Figure 6).

3.2.2. The Performance of Remote Sensing Indices and Field Sensors Assessing Grain Yield

No significant differences were found between quartile groups for any of the RGB indices from the aerial or ground level, as seen for GGA at ground level. The correlations were calculated for GY with both levels of RGB indices at LOW (Table 3). In the case of UAV RGB VIs, GGA was best correlated with GY followed by CSI, followed by saturation and GA. For ground RGB VIs the closest correlations were observed with hue, GGA (values between 0.248 and 0.685), and NDLab (values between 0.3953 and 0.8290). The rest of the RGB VIs were somewhat weaker with respect to the GY, but many were still significant. Additionally, the field sensors presented some close correlations with GY (Table 3), as well as significant differences between genotypes when grouped by quartile (data not shown), the strongest was SPAD^V (taken in the vegetative stage closer to the RGB VIs), which only indicated differences between HY and LY, followed by SPAD^V and finally NDVI, both of which were recorded in the vegetative stage (see Figure 3). SPAD measurements exhibited the highest correlations (Table 3) with respect to all indices with SPAD^V being the highest followed by SPAD^R (taken in the reproductive stage much later than the other measurements for temporal comparisons only).

Table 3. Grain yield correlations in LOW with all proximal remote sensing variables from the RGB images taken from the UAV aerial platform, RGB images from the ground, and leaf chlorophyll content (SPAD) and normalized difference vegetation index (NDVI) field sensors. These indices are defined in Sections 1 and 2. Levels of significance: *, $P < 0.05$; ***, $P < 0.001$. GGA; GA; NDLab, normalized difference between a* and b*; NDLUv, the normalized difference between u* and v*; CSI, crop senescence index; TGI, triangular greenness index; and NGRDI, normalized green–red difference index.

GY								
UAV RGB VIs	r	P	Ground RGB VIs	r	P	Additional Field Sensors	r	P
GGA	0.445	***	GGA	0.483	***	SPAD ^V (18/02/16)	0.542	***
GY	0.407	***	GA	0.466	***	SPAD ^R (01/03/16)	0.506	***
Hue	0.381	***	Hue	0.485	***	NDVI	0.375	***
Intensity	−0.305	***	Intensity	0.095				
Saturation	−0.427	***	Saturation	−0.227	*			
Lightness	−0.291	***	Lightness	0.144	*			
a*	−0.36	***	a*	−0.383	***			
b*	−0.397	***	b*	−0.089				
u*	−0.383	***	u*	−0.449	***			
v*	−0.297	***	v*	0.014				
NDLab	0.359	***	NDLab	0.468	***			
NDLUv	−0.378	***	NDLUv	0.442	***			
CSI	−0.428	***	CSI	−0.321	***			
TGI	0.229	*	TGI	−0.043				
NGRDI	0.406	***	NGRDI	−0.027				

The correlation coefficients between the hue, u*, GA, and GGA remote sensing indices evaluated at ground level versus the same indices measured from the UAV were quite strong (Table 4). In addition, most of these indices showed slopes close to 1:1 and correlations reaching $r = 0.766$. In contrast, the relationships reported for the remaining RGB indices, such as intensity, lightness, TGI, and NGRDI were lower. With regards to ANOVA, the results showed that there were statistically significant differences between all the RGB indices at ground level with the aerial observation level, except for saturation from the HSI color space.

Table 4. Correlation coefficients between the remote sensing UAV RGB VIs and ground RGB VIs. These indices are defined in Section 1 and further detailed in Section 2. Levels of significance: *, $P < 0.05$; **, $P < 0.01$, ***, $P < 0.001$; ns, not significant. GGA, GA, NDLab, NDLUv, CSI, TGI, and NGRDI.

	R	P	ANOVA
GGA	0.758	***	***
GA	0.766	***	***
Hue	0.731	***	***
Intensity	−0.062	ns	***
Saturation	0.509	***	ns
Lightness	−0.039	ns	***
a*	0.617	***	***
b*	0.424	***	***
u*	0.723	***	***
v*	0.33	***	***
NDLab	0.781	***	***
NDLUv	−0.676	***	***
CSI	0.457	***	***
TGI	−0.163	*	*
NGRDI	−0.223	*	**

3.3. Agronomic Parameters and Their Effect on Yield

Finally, for agronomic data parameters, the correlations with GY (Table 5), the results showed that all the agronomic data indicators performed differently between OP and LOW conditions. For LOW conditions, the indices that were better correlated with GY were ASI and AD. The other two indices (PH and SEN) showed very low correlations with GY. In OP conditions, the agronomic parameters showed very low correlations with GY.

Table 5. Grain yield correlations with different indices of agronomic data, such as PH, SEN, AD, and ASI. Correlations were studied across plots in LOW and OP conditions. Levels of significance: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant.

Agronomic Data	GY			
	LOW		OP	
	r		r	
PH	0.191	**	0.131	ns
SEN	−0.213	**	NA	ns
AD	−0.46	***	0.272	**
ASI	−0.53	***	0.161	*

3.4. Multivariate Models

Figure 7 shows the correlations of the most relevant agronomic parameters and indices with GY; it was considered that these could be complimentary in multivariate models because ASI and AD show negative correlations with respect to GY, whereas the other indices present positive correlations with GY. As such, in Table 6 we present the stepwise multivariate linear models for explaining grain yield using different selections of non-destructive UAV RGB VIs and ground RGB VIs at additional field sensor and agronomic data as indicated using both forward and backward stepwise selection techniques with a standard Akaike information criterion (AIC) selection criterion. We also present the determination coefficients (R^2) and the residual standard error (RSE). All three models presented were found to be significant at the $P < 0.001$ level. Where noted, (*) indicates simplified formulas, meaning that in using stepwise selection these formulas were considered to have an excess of nonsignificant parameters and were reduced accordingly (only significant parameters with the strongest individual correlation to yield were selected in the case of auto-correlation detected between two multivariate parameters).

Table 6. Multilinear regressions (stepwise) of GY in LOW as the dependent variable comparing the different categories of remote sensing traits: UAV and ground RGB VIs (these indices are defined in Section 1), agronomic data such as ASI, AD, SEN, and PH, and NDVI and SPAD. R^2 , determination coefficient; RSE, Residual Standard Error. Level of significance: ***, $P < 0.001$. (*) simplified formulas.

Parameters	Stepwise Equations	R^2	RSE	P
Agronomic Data + Field sensors	$GY = -AD*0.28 + SPAD^V*0.03 + SPAD^R*0.02 - ASI*0.78 + 5.97$	0.61	0.539	***
Agronomic Data + Ground RGB VIs (*)	$GY = -ASI*0.189 - AD*0.128 - SEN*0.237 + PH*0.01 + b*0.11 - v*0.064 + NDLab*15.20 - ND Luv*6.99 + 3.36$	0.588	0.556	***
Agronomic Data + UAV RGB VIs (*)	$GY = -ASI*0.20 - SEN*0.26 - AD*0.13 + PH*0.01 - Saturation*84.97 - u*1.37 + v*1.61 + TGI*0.02 + ND Luv*3.95 + 31.8$	0.604	0.546	***

In the combination of agronomic data with additional field sensors, such as $SPAD^V$, NDVI, and ASI, 61% of the yield could be explained. The multivariate stepwise models for explaining yield

variations regarding agronomic data plus ground RGB VIs level was 63%, but in Table 5, we show the simplified formula explaining 58% of variance. With respect to combining agronomic data with UAV RGB VIs, the result was similar at 62%, but we show the simplified formula with 60% of the yield that may be explained by mostly agronomic data and combined with other indices, such as NDLab, NDLuV, and Saturation. Combining the agronomic data, field sensors and RGB VIs provided little improvement in the multivariate model explaining the yield, whereas, in comparison, more parsimonious models combining only AD, ASI, and only either the NDLab or NDLuV RGB indices still explained over 50% of the variation in yield (data not shown).

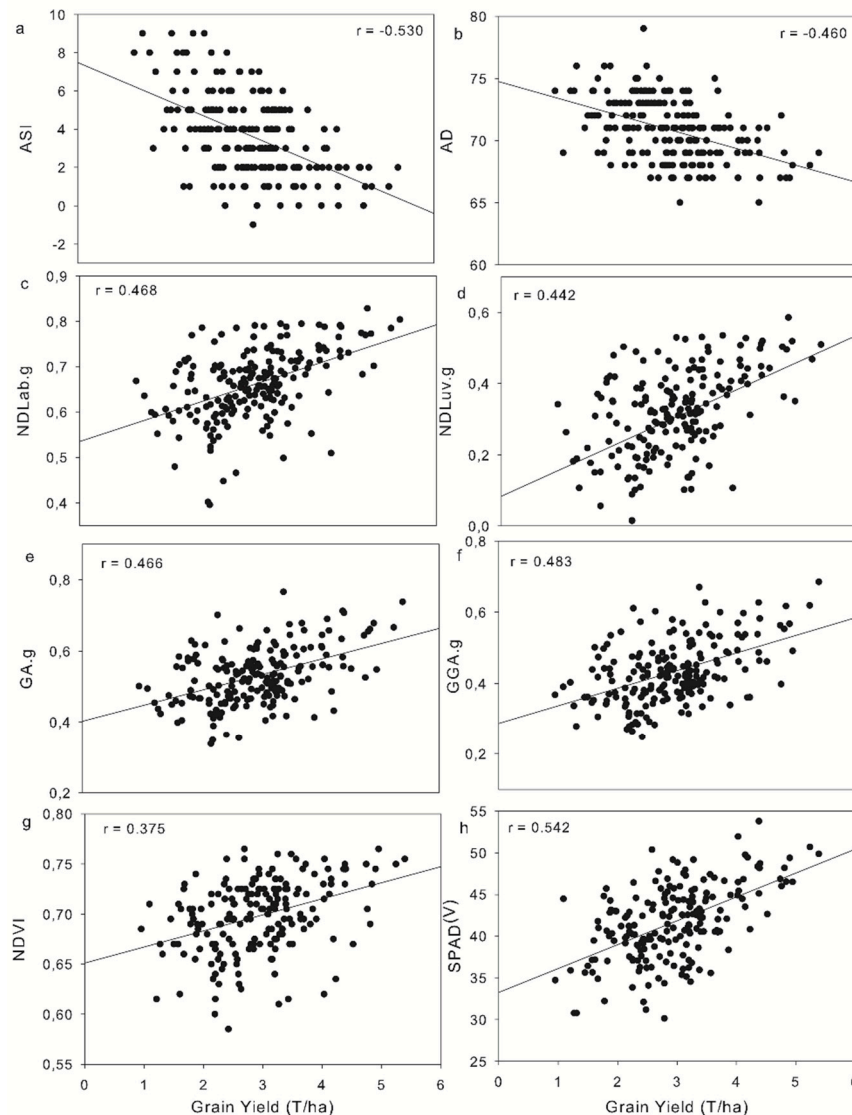


Figure 7. Correlation between GY in LOW and (a,b) traditional field plant physiology measurements ASI and AD; ground RGB VIs (c) NDLab; (d) NDLuV; (e) GA.g (green area ground); (f) GGA.g (greener green area ground); (g) plant vigor using GreenSeeker NDVI; and (h) relative leaf chlorophyll content in the vegetative stage (SPAD^V). Level of significance: ***, $P < 0.001$. (n = 192).

4. Discussion

4.1. The Effect of Managed Low Nitrogen on Grain Yield

Nitrogen (N), after water, is the single most important input for maize production. It plays a major role in establishing optimal photosynthetic capacity during key growth stages for crops to achieve high

yields [18,61]. N deficiency reduces leaf chlorophyll content, soluble protein content, photosynthetic rate, and related enzyme activities of the maize plant during grain filling [62–64]. For that reason, the GY of all the 64 genotypes was assumed to have been strongly affected by the lack of nitrogen in LOW conditions. There were significant differences noted in GY by genotype, but in Table A1 we can see that OP treatment presented many more different groups, and because of that, the genotypes were divided into quartiles for the sake of comparisons across growing conditions. (Figure 4). These four groups showed differences between each other with respect to grain yield. Nitrogen is especially plentiful in leaves, mainly in photosynthetic enzymes, where it may account for up to 4% of the dry weight. Because N uptake, biomass production, and grain yield are strongly correlated, the N requirement of a maize crop has even been directly related to grain yield; it has been estimated that 187 kg/ha N is required to produce 9.5 t/ha yield, 98 kg/ha is required for 5.0 t/ha, and 40 kg/ha is required for 2.0 t/ha [17]. Following these guidelines, the concentration for the optimal condition was around 200 kg/ha and that for the low managed nitrogen condition was around 40 kg/ha, thus indicating that many of the genotypes tested here have already been somewhat adapted to the managed low nitrogen conditions; however, those that showed the best and most consistent adaptation were the genotypes of the HY quartile, some of which appeared in HY in both LOW and OP (Table 2). In some cases, it has been reported that the genotypes selected under LOW fertilization input are not truly adapted to N-rich soils [65]; however, it has been suggested that when the plant material performs relatively well under low-N input conditions, it should be selected under N deficiency conditions for which yield reduction does not exceed 35–40% [66]. Here, in comparison with the same panel of maize genotypes grown under optimal conditions, 44% of the genotypes that were in the highest yield producing quartile under OP conditions remained in the highest quartile when grown under LOW conditions, further suggesting that low N productivity is not necessarily exclusive of high productivity in OP conditions. The GYLI results also show that there was a large amount of genotypic variability present (Table A1), again suggesting that the genotypes selected for this study behaved quite differently physiologically under the two nitrogen conditions.

4.2. Effect of Managed Low Nitrogen on Agronomic Parameters

A higher reduction in maize yield under stress environments can often be partly explained by the wider date range of ASI under stress, as ASI typically has a high negative correlation with GY under stress conditions [67,68]. In Figure 7, we demonstrated that AD and ASI exhibit negative significant relation in low-N conditions correlations when the plants were under stress, as these parameters decreased with increasing yield. Genotypes with a short ASI have been suggested to possess greater efficiency in biomass partitioning to ear and tassels at flowering than those with a long ASI [69]. On the other hand, the correlation results in OP conditions do not indicate any relationship between GY and ASI or AD (Table 5). Similar studies showed that high GY under a range of stress intensities is associated with a short ASI and earlier flowering dates, increased plant and ear height, increased number of ears per plant, and delayed leaf senescence [69,70].

Besides ASI, traits related to high photosynthetic capacity (e.g., chlorophyll content) and plant water content (e.g., stomatal conductance) have often been reported to contribute to higher GY under drought stress [71]. When maize flowers under drought, there is a delay in silking, and the period between male and female flowering increases giving rise to ASI. In this study, however, drought stress was minimal due to the adequate rainfall recorded during the study field season (700 mm), and also further supported by the lack of correlation between ASI and AD with GY in the OP conditions (Table 5). In generally optimal agronomic conditions, these phenological characteristics are not always good estimators of yield. In contrast, the rest of the agronomic parameters here showed similar relation with GY variability. These results suggest that this technology could be applied in an adapted way to water stress studies, even though it was not the specific aim of this research.

4.3. Remote Sensing Indices and Field Sensors

4.3.1. Color and Vignetting Calibration

Figure 5a–c,g–i show the correlations between the R, G, and B values from of each digital camera with respect to the standard values for R, G, and B from the X-Rite ColorChecker Passport. Some lamination effects can be observed as values from the camera appearing higher than the standard, though still with good correlation. However, in Figure 5e,f,k,l of the a^* and b^* images, the light effects are less pronounced, with a ratio closer to 1:1 due to the CIELab color space separation of lightness from color hue values [72]. Regarding the color calibration, the results show that the determination coefficients between the values of the ColorChecker passport and the values for each camera were high (most near $R^2 \approx 0.90$), suggesting that the photos were not in need of a separate color calibration [73]. Indeed, application of the color calibration coefficients did not result in any improvements in RGB vegetation index performance with regards to grain yield or yield loss estimation (data not shown).

With respect to vignetting effects, in Figure 6a,b, the graphics of three RGB channels present a small sad smile along the axis, showing some brightness attenuation away from the image center [74,75]. Nevertheless, this did not represent any significant differences between the DN values of extremes of each side respect to the DN of center (data not shown). On the other hand, Figure 6c,d present minimal variation along the axis. This was expected as a^* and b^* color spaces are independent of the image lightness and thus absent of vignetting effects. Thus, while the RGB VIs are passive sensors and dependent on ambient light conditions, the use of the alternate color spaces, such as CIELab, provides for inherent lightness correction and enables their use in variable conditions similar to active sensors, such as the NDVI GreenSeeker.

4.3.2. Performance of RGB VIs and Additional Field Sensors

The ground RGB VIs hue, GGA, and NDLab demonstrated the best correlations with GY, outperforming other ground RGB VIs (Table 3). GA and GGA quantify the portion of green pixels to the total pixels of the image and is a reliable estimation of vegetation cover [76]. The values of GA from both observation levels (field and aerial) were consistently below 60%. The ground and aerial measurements were taken at the same time on the same day, variation in environmental variables, such as light intensity and brightness can be assumed to be negligible. Thus, the main differences must be due to the resolution of the images (Figure 2); nevertheless, advances in digital photography allow for sufficiently high resolution for low-altitude aerial imaging to be a viable and economical monitoring tool for agriculture [77]. In this sense, correlations with GY by indices derived from aerial images were generally only slightly weaker than indices measured at ground level, most likely demonstrating a trade-off between them. Some of the RGB indices, such as NDLab, GGA, and GA produced coefficients of correlation higher than $r = 0.75$ when comparing the same indices measured from the ground level and from the aerial platform (Table 4). This is despite different acquisition/imaging techniques (full plot for UAV RGB VIs vs. one image per plot at higher resolution covering only a portion of the plot for the ground RGB VIs). On the other hand, none of the UAV RGB VIs and ground RGB VIs showed significant differences between adjacent quartiles. This may be best explained considering that the data for our study were collected at an early phenological (vegetative) stage, when the plants were not yet at full canopy cover, and they did not yet show the full range of symptoms of N deficiency, as may be observed in the reproductive stage (Figure 3). N deficiency can reduce plant growth rates, but also other later factors that affect GY, including leaf chlorophyll content, soluble protein content, photosynthetic rate, and related enzyme activities of the maize plant during grain filling in the reproductive stage [78–80], which may be a more optimal timing of remote sensing observations when phenotyping low N.

NDVI has been used with satisfactory results in many predictive models of yield in multiple crops, including wheat, barley, and maize at the field level [54,81], even at regional or state levels using field, airborne, and satellite imagery [82–84]. Regarding NDVI, the values clearly highlight (Figure 7)

that the variability was low, with more than 90% of values being in the range of 0.55–0.80. These results support the previously reported saturation of the index, such that increasing leaf area does not involve a parallel increase in NDVI values [83,85,86]. Furthermore, other authors have noted that the optimal stages for measuring NDVI vary depending on the germplasm and environment [87,88]. Better performance of NDVI usually occurs at earlier or later growth stages, depending on the crop and symptoms, because at maximum vegetation cover, NDVI values often saturate, and thus, correlation with GY decreases [38]. Furthermore, other studies comparing the performance of RGB VIs with frequent data acquisitions throughout the crop growing cycle and NDVI indicate that RGB VIs respond with higher correlations with GY earlier than NDVI [57].

SPAD is used to measure relative chlorophyll content in plant leaves, and it has been effectively used to diagnose N status and predict GY potential in maize [89]. Leaf level chlorophyll meters provide a convenient and reliable way to estimate leaf N content during vegetative growth [90] and over a large time range even after anthesis [91]. With specific mention regarding the two different measurement dates, SPAD^V and SPAD^R (Table A2), there are some notable differences related, in this case, to the date of measurement, having been one of the few sensors available to the field crews on-site to do multiple measurements. In the first measurement, SPAD^V, the results showed significant differences between the HY, MHY, and MLY groups in comparison with the LY group. These are interpreted as symptoms of lack of nitrogen. With regards to the second measurement, SPAD^R, it was possible to identify the differentiation of HY groups in comparison with MHY and MLY and, additionally, these three from the LY group. This is interpreted as a decline in relative chlorophyll content of the leaves measured between the two SPAD measurements. This is because when crops were younger at SPAD^V, when still developing roots and leaves, they may behave as sink organs for the assimilation of N and the synthesis of amino acids originating from N uptake before flowering [92]. After flowering, at SPAD^R, the N accumulates in the vegetative parts of the plant and is remobilized and translocated to the grain [93]. In most crop species, a substantial amount of N is absorbed after flowering to contribute to grain protein deposition [89] in V8–R1. Similarly, Teal et al. [94] also reported a strong association between grain yield and NDVI between the V6 and V8–R1 growth stages of maize, again between the timing of SPAD^V and SPAD^R, but after the remote sensing observations were recorded. Finally, the increased performance of SPAD^V compared with SPAD^R, closer to the RGB and NDVI data acquisition date, was promising in that earlier stage of image data capture for field phenotyping, which may reduce crop breeding costs with earlier variety selection and increased crop cycles per year.

4.4. Multivariate Model Assessment

The vegetation indices derived from conventional digital RGB images have been proposed as means of estimating green biomass and GY in maize and other cereals under stress conditions [95], and in other studies in wheat grown under different stress conditions [38,43,96]. The multivariate regression models revealed the most appropriate parameters for field phenotyping towards improving GY in managed low-nitrogen conditions. Using all the UAV RGB VIs and ground UAV VIS, the multivariate models explained GY at around $R^2 = 0.30$ (data not shown). That could be a result of the fact that the data capture of RGB VIs were taken earlier than the SPAD^V and SPAD^R, and at later growth stages, the plants may have presented more symptoms related to a lack of nitrogen. However, all of the regression models with a R^2 higher than 0.50 included some of the agronomic data as independent variables. Additionally, GY estimation was similar in the cases of agronomic data combined with field sensors with respect to agronomic data combined with UAV RGB VIs and ground RGB VIs (all having approximate R^2 values of 0.60). With respect to the RSE in these three different multivariable models, they adapt to a 50% coefficient of determination. As such, our study suggests that these phenological traits still provide useful information related to grain yield in abiotic stress conditions, but that they may be potentially supplemented by UAV RGB VIs and ground RGB VIs phenotyping platforms. Still, RGB image analyses were able to improve over agronomic data alone, increasing the R^2 values to explain more than half of the variance in the yield, suggesting that they are complimentary in the

information that they are able to provide. Furthermore, the UAV RGB VIs in this study were acquired quite early in the growing season, which may help to provide for faster selection of varieties, thus reducing costs and increasing the number of crop cycles per year.

Additionally, RGB VIs may provide considerable saving with regards to field equipment and human time, considering that RGB data capture and processing of 200 plots took approximately 10 min in the field (counting flight and preparations), 20 min to mosaic (unattended on the computer), and half an hour to extract and process the data (semi-automatic), totaling 60 min, excluding drone preparation prior to flying. In the case of RGB image data capture and processing, the field portion would be approximately double, while the computer processing would be about half, totaling about the same amount of time in the case of 200 plots. With respect to time costs while implementing the use of SPAD or field-based GreenSeeker NDVI, the estimated time for measurement would be over five times greater (e.g., 2 min per plot \times 200 plots = 400 minutes). Moreover, NDVI evaluation from the ground may not be easy to implement when plants are reaching the reproductive stage. Furthermore, with larger phenotyping trials, the time savings of the UAV RGB VIs would represent even greater time savings while retaining the same data quality; up to 1000 plots may approximately double the amount of time needed to process the data, while the field sensors would increase linearly and take five times as long, representing over a 10-fold time difference at larger study scales. Thus, the implementation of higher throughput UAV RGB VIs may make the most sense in combination with some of the quicker traditional agronomic measurements and can also result in substantial time cost savings when applied in large platform breeding programs.

In a recent study by Gracia-Romero et al. in 2017 [97], the effectiveness of UAVs for canopy level remote sensing for plant phenotyping of maize was similarly demonstrated under different phosphorus nutrient conditions and the results presented therein suggested that the RGB indices were the best option at early growth stages. In the case of low P, however, an equation using GA and u^* were the best indicators of GY ($R^2 = 0.82$). Even though applied to a different crop in that case, reflectance in the near-infrared (NIR) and blue regions was found to predict early season P stress between growth stages V6 and V8, much earlier than suggested for N deficiencies. With respect to plant N concentrations, the best correlations have been found using reflectance in the red and green regions of the spectrum, while grain yield was best estimated using reflectance in the NIR region, with the wavelengths of importance changing with growth stage (V14-R1) [98,99]. Furthermore, Ma et al. [100] showed that canopy light reflectance is strongly correlated with field “greenness” (similar to the GA and GGA used in this study) at almost all growth stages, with field greenness in that case being a product of plant leaf area and leaf greenness measured with a chlorophyll meter.

5. Conclusions

Modern phenotyping technologies may help in improving much-needed maize GY in low-N conditions, and the current range of variability in performance as indicated by the observed GYLI values suggests that low-N and optimal-N performance need not be considered mutually exclusive. For HTPP, RGB sensors can be considered to be functional technology with an advanced technology readiness level (TRL) from the ground or a UAV platform, but, similar to the current standard field sensors SPAD and NDVI, the data capture for RGB VIs must be planned accordingly in order to optimize their benefits in support of plant breeding. Several different RGB image-based vegetation indices, including the NDLab and NDluv indices new to this study, demonstrated similar correlations with GY and contributions to multivariate model GY estimates when compared with standard NDVI and SPAD field phenotyping sensors. This study presents possible uses of RGB color image analyses from the ground or from UAVs, with potential benefits compared with currently used field sensors, especially regarding time costs when applied to larger breeding platforms, here demonstrated in application to low-N phenotyping in maize.

Author Contributions: M.A.Z.-A., A.T., B.M.P., and J.C. managed and directed the maize trials at the CIMMYT Southern African Regional Office in Harare, Zimbabwe. S.C.K. carried out the UAV flights and image pre-processing for the obtainment of aerial measurements. A.G.-R., O.V.-D., and J.L.A. conducted the field measurements and the collection of samples. M.L.B. processed the images, analyzed the samples, and wrote the paper under the supervision of J.L.A. and S.C.K. with specific contributions from all of the other authors.

Funding: This research and APC was funded by Bill & Melinda Gates Foundation and USAID Stress Tolerant Maize for Africa program, grant number [OPP1134248], and the MAIZE CGIAR research program. The CGIAR Research Program MAIZE receives W1&W2 support from the Governments of Australia, Belgium, Canada, China, France, India, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Sweden, Switzerland, U.K., U.S., and the World Bank.

Acknowledgments: We thank the personnel from the CIMMYT Southern Africa Regional Office at Harare, Zimbabwe, for their support during the field measurements and sampling. Finally, we thank Jaume Casadesús for providing the BreedPix software and Esnath Hamadziripi for being instrumental in the trial implementation.

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

Abbreviations

The following abbreviations are used in this manuscript.

N	Nitrogen
SPAD	Relative leaf chlorophyll content
HTTP	High throughput plant phenotyping
UAV	Unmanned aerial vehicle
VNIR	Visible and near-infrared
NDVI	Normalized difference vegetation index
SR	Simple ratio
NDWI	Normalized difference water index
RGB	Red–green–blue
RGB VIs	Red–green–blue vegetation indices
UAV RGB VIs	Unmanned aerial vehicle red–green–blue vegetation indices
Ground RGB VIs	Ground level red–green–blue vegetation indices
HIS	Hue–intensity–saturation
H	Hue
GA	Green area
GGA	Greener green area
CSI	Crop senescence index
CIMMYT	International center for maize and wheat improvement
OP	Optimum nitrogen
LOW	Low managed nitrogen
L*	Lightness
NDLab	Normalized difference between a* and b*
NDLuv	Normalized difference between u* and v*
TGI	Triangular greenness index
NGRDI	Normalized green–red difference index
ASI	Anthesis silking interval
AD	Anthesis data
PH	Plant height
SEN	Canopy senescence
MOI	Moisture
GY	Grain yield
GYLI	Grain yield loss index
HY	High yield
MHY	Medium high yield
MLY	Medium low yield
LY	Low yield
ANOVA	Analyses of variance

Appendix A

Table A1. The LOW and OP treatment with ANOVA analysis, following with a post hoc.

LOW			OP		
Genotype	GY (Mg/ha)	Group	Genotype	GY (Mh/ha)	Group
CZH15062	4.43	a	PGS65	12.30	a
CZH15047	4.19	ab	CZH15026	12.08	ab
CZH15028	3.88	abc	CZH15054	12.04	abc
CZH15032	3.87	abc	CZH15022	11.98	abc
CZH15057	3.82	abc	CZH15057	11.40	abcd
CZH15058	3.65	abc	CZH15053	11.36	abcd
10C3271	3.54	abc	PAN53	11.27	abcde
CZH15055	3.51	abc	CZH15024	11.13	abcdef
CZH15024	3.5	abc	11C4393	10.97	abcdefg
CZH15052	3.44	abc	CZH128	10.93	abcdefgh
CZH142087	3.41	abc	MRI 614	10.79	abcdefghi
CZH15045	3.38	abc	CZH141029	10.75	abcdefghi
11C4393	3.37	abc	X40F424W	10.73	abcdefghi
CZH128	3.35	abc	LOCAL CHECK2	10.59	abcdefghij
CZH15060	3.3	abc	CZH132043	10.53	bcdefghij
CZH15050	3.28	abc	CZH132047	10.51	bcdefghijk
CZH15031	3.27	abc	CZH15029	10.49	bcdefghijk
CZH15046	3.25	abc	CZH15037	10.35	cdefghijkl
CZH132043	3.22	abc	10C3271	10.24	defghijklm
CZH15033	3.19	abc	CZH15033	10.23	defghijklm
LOCAL CHECK1	3.18	abc	MH1547	10.15	defghijklmn
CZH15051	3.17	abc	CZH15045	10.13	defghijklmn
CZH15054	3.16	abc	CZH15043	10.08	defghijklmn
CZH15029	3.16	abc	CZH15028	10.06	defghijklmno
CZH141022	3.15	abc	CZH15030	10.04	defghijklmno
CZH142010	3.15	abc	CZH15044	9.90	defghijklmno
CZH15042	3.06	abc	CZH15056	9.87	defghijklmno
CZH15039	3.05	abc	CZH15035	9.80	defghijklmnop
CZH15030	3.05	abc	CZH15036	9.71	defghijklmnop
CZH15035	3.03	abc	CZH15058	9.69	defghijklmnopq
CZH15038	3	abc	CZH15060	9.68	defghijklmnopq
CZH15059	3	abc	CZH15047	9.61	efghijklmnopq
CZH15044	2.97	abc	CZH15052	9.61	efghijklmnopq
PAN53	2.95	abc	CZH15025	9.56	efghijklmnopq
CZH15041	2.93	abc	CZH15048	9.48	fghijklmnopqr
CZH15040	2.87	abc	CZH15032	9.47	fghijklmnopqr
P2859W	2.85	abc	CZH15034	9.39	ghijklmnopqrs
MH1547	2.82	abc	CZH15061	9.39	ghijklmnopqrs
X40F423W	2.81	abc	CZH15031	9.37	ghijklmnopqrs
CZH15026	2.79	abc	CZH15038	9.24	hijklmnopqrs
CZH15053	2.75	abc	CZH15041	9.20	hijklmnopqrst
SC513	2.72	abc	CZH15023	9.17	ijklmnopqrst
CZH15061	2.71	abc	CZH15050	9.08	ijklmnopqrstu
CZH1227	2.68	abc	CZH141022	8.96	jklmnopqrstu
CZH15056	2.68	abc	X40F423W	8.93	jklmnopqrstu
PGS65	2.67	abc	CZH15046	8.88	jklmnopqrstu
X40F424W	2.67	abc	CZH15049	8.87	jklmnopqrstu
CZH15023	2.56	abc	MRI 624	8.79	klmnopqrstu
CZH15025	2.56	abc	CZH15039	8.72	lmnopqrstu
CZH141029	2.52	abc	CZH15042	8.59	mnopqrstu
CZH15048	2.45	abc	CZH142010	8.56	mnopqrstu
CZH132047	2.45	abc	PHB30G19	8.46	nopqrstu
LOCAL CHECK2	2.44	abc	CZH1227	8.45	nopqrstu
CZH15027	2.32	abc	CZH15027	8.44	nopqrstu

Table A1. Cont.

LOW			OP		
Genotype	GY (Mg/ha)	Group	Genotype	GY (Mh/ha)	Group
MRI 634	2.29	abc	CZH142087	8.42	nopqrstu
PHB30G19	2.27	abc	CZH15062	8.34	opqrstuv
MRI 614	2.24	abc	CZH15040	8.12	pqrstuv
CZH15049	2.21	abc	MRI 634	7.97	qrstuv
CZH15022	2.03	bc	P2859W	7.77	rstuv
CZH15043	2.02	bc	CZH15055	7.71	stuv
CZH15036	1.98	bc	CZH15059	7.68	stuv
CZH15034	1.92	bc	CZH15051	7.48	tuv
MRI 624	1.69	c	LOCAL CHECK1	7.39	uv
CZH15037	1.53	c	SC513	6.68	v
ANOVA: ***			ANOVA: ***		

Table A2. Descriptive data of GY, LOW, GYLI, ASI, PH, AD, NDVI, SPAD^V, SPAD^R, GA, and GGA according to the quartiles HY, MHY, MLY, and LY.

Entry name	GYLI (%)	GY (12/5/16)	ASI	PH (29/2/16)	AD	NDVI (28/1/16)	SPAD ^V (18/2/16)	SPAD ^R (1/3/16)	GA.a (28/1/16)	GA.g (28/1/16)	GGA.a (28/1/16)	GGA.g (28/1/16)
CZH128	69.32	3.35 ± 0.39	1.00	193.00	71.67	0.730	41.09	37.97	0.483	0.522	0.327	0.452
CZH142087	59.54	3.41 ± 0.75	4.00	192.38	67.33	0.705	42.86	37.56	0.457	0.573	0.307	0.479
CZH15024	68.59	3.50 ± 0.65	3.00	183.96	73.00	0.672	40.78	34.07	0.417	0.454	0.269	0.360
CZH15028	61.44	3.88 ± 0.43	0.67	200.63	69.33	0.702	44.41	43.00	0.358	0.496	0.225	0.415
CZH15031	65.11	3.27 ± 0.06	5.33	201.88	69.00	0.698	41.18	39.03	0.445	0.533	0.266	0.413
CZH15045	66.60	3.38 ± 0.21	2.00	208.04	70.00	0.707	41.80	36.65	0.460	0.540	0.324	0.434
CZH15047	56.32	4.20 ± 0.39	2.00	214.21	68.33	0.730	45.36	39.01	0.537	0.622	0.383	0.515
CZH15050	63.89	3.28 ± 0.35	3.00	172.88	68.67	0.725	43.14	39.29	0.410	0.471	0.256	0.391
CZH15052	64.19	3.44 ± 0.36	3.00	198.79	69.33	0.730	43.18	38.93	0.531	0.574	0.377	0.464
CZH15055	54.54	3.51 ± 0.69	2.00	188.75	70.67	0.675	43.25	40.68	0.397	0.497	0.232	0.397
CZH15057	66.49	3.23 ± 0.43	2.67	187.58	72.00	0.707	41.76	37.37	0.461	0.536	0.299	0.468
CZH15058	62.28	3.83 ± 0.44	0.33	185.63	71.67	0.712	46.04	38.54	0.421	0.558	0.278	0.456
CZH15061	71.13	3.48 ± 0.95	4.00	182.00	71.33	0.720	45.71	35.81	0.478	0.566	0.323	0.495
CZH15062	46.88	3.90 ± 0.49	1.67	185.42	68.33	0.697	42.51	37.50	0.484	0.557	0.326	0.477
11C4393	69.25	3.44 ± 0.22	3.00	204.29	67.67	0.720	41.01	33.84	0.529	0.622	0.367	0.534
Local check 2	76.92	3.64 ± 0.94	3.33	207.75	70.00	0.695	43.63	37.16	0.413	0.521	0.273	0.450
CZH132043	69.44	3.22 ± 0.56	3.67	202.33	72.00	0.692	43.48	37.98	0.431	0.548	0.285	0.442
CZH142010	63.25	3.15 ± 0.53	4.33	189.00	67.67	0.697	41.75	36.91	0.444	0.524	0.295	0.436
CZH141022	64.84	3.15 ± 0.15	4.33	169.75	68.67	0.697	42.57	36.05	0.487	0.490	0.331	0.410
CZH15029	69.91	3.16 ± 0.54	4.00	205.25	70.33	0.712	43.49	40.53	0.473	0.511	0.318	0.412
CZH15030	69.63	3.05 ± 0.60	3.33	197.46	70.67	0.685	44.07	37.67	0.387	0.472	0.246	0.378
CZH15032	59.12	3.87 ± 0.56	2.67	181.54	69.00	0.728	41.54	36.33	0.428	0.508	0.270	0.404
CZH15033	68.86	3.19 ± 0.61	4.33	195.00	73.00	0.690	41.56	36.17	0.495	0.534	0.321	0.414
CZH15035	69.07	3.03 ± 0.45	4.67	209.83	73.33	0.710	38.39	37.47	0.472	0.521	0.299	0.417
CZH15039	65.00	3.05 ± 0.46	4.00	196.21	67.67	0.687	42.23	41.02	0.379	0.502	0.243	0.418
CZH15042	64.35	3.06 ± 0.50	1.67	199.67	72.00	0.727	37.11	32.71	0.432	0.568	0.273	0.446

Table A2. Cont.

Entry name	GYLI (%)	GY (12/5/16)	ASI	PH (29/2/16)	AD	NDVI (28/1/16)	SPAD ^V (18/2/16)	SPAD ^R (1/3/16)	GA.a (28/1/16)	GA.g (28/1/16)	GGA.a (28/1/16)	GGA.g (28/1/16)
CZH15046	63.40	3.25 ± 0.39	3.00	173.17	67.33	0.687	46.32	38.29	0.425	0.555	0.285	0.468
CZH15051	57.62	3.17 ± 0.17	1.67	180.33	69.33	0.717	44.84	38.88	0.465	0.554	0.277	0.432
CZH15054	73.79	3.16 ± 0.31	3.00	207.33	71.00	0.710	38.12	32.19	0.478	0.605	0.312	0.451
CZH15059	60.88	3.22 ± 0.79	5.00	186.96	69.67	0.710	43.65	36.60	0.450	0.598	0.289	0.485
X40F424W	75.08	3.07 ± 0.66	5.33	213.71	72.00	0.727	43.71	33.73	0.551	0.627	0.400	0.539
10C3271	65.39	3.14 ± 0.15	1.67	198.83	66.67	0.723	40.27	39.00	0.492	0.612	0.350	0.510
PAN53	73.79	2.95 ± 0.23	3.67	204.25	70.67	0.702	43.55	34.76	0.451	0.576	0.318	0.504
P2859W	63.28	2.85 ± 0.33	2.67	187.67	70.00	0.643	41.21	35.78	0.367	0.453	0.239	0.388
SC513	59.23	2.72 ± 0.25	4.33	197.58	69.67	0.673	38.00	32.98	0.379	0.439	0.209	0.319
CZH1227	68.28	2.68 ± 0.42	2.67	173.21	69.00	0.702	40.32	37.79	0.508	0.585	0.340	0.481
CZH15023	72.05	2.56 ± 0.66	3.33	179.04	70.33	0.703	39.43	31.36	0.449	0.545	0.272	0.461
CZH15025	73.23	2.56 ± 0.71	1.67	196.38	72.67	0.725	35.41	31.33	0.471	0.529	0.306	0.447
CZH15026	76.90	2.79 ± 0.25	5.00	208.17	72.67	0.723	40.93	32.87	0.473	0.530	0.306	0.434
CZH15038	67.50	3.0 ± 0.52	5.33	199.00	69.00	0.717	44.98	37.76	0.431	0.574	0.287	0.492
CZH15040	64.71	2.87 ± 0.22	6.00	180.25	67.33	0.658	43.01	39.73	0.373	0.441	0.221	0.331
CZH15041	68.13	2.93 ± 0.37	4.67	184.21	69.00	0.673	43.31	37.73	0.351	0.499	0.212	0.412
CZH15044	70.02	2.97 ± 0.13	3.67	190.25	70.33	0.683	45.24	40.79	0.424	0.524	0.272	0.435
CZH15053	75.79	2.75 ± 0.37	2.00	202.00	71.67	0.715	38.50	33.27	0.533	0.552	0.367	0.449
CZH15056	72.87	2.68 ± 0.24	1.67	174.29	73.00	0.702	41.79	32.66	0.383	0.515	0.235	0.415
CZH15060	65.96	2.67 ± 0.20	2.00	181.13	69.00	0.685	43.64	40.30	0.458	0.543	0.306	0.461
X40F423W	68.52	2.81 ± 0.34	6.67	204.17	72.00	0.687	41.63	33.69	0.416	0.474	0.278	0.370
Local check 1	56.92	2.75 ± 0.47	3.00	201.96	72.67	0.678	44.61	35.68	0.407	0.467	0.246	0.357
PHB30G19	73.10	2.28 ± 0.60	6.67	203.75	69.33	0.712	41.80	35.77	0.478	0.561	0.335	0.469
CZH132047	76.68	2.45 ± 0.20	4.00	209.83	72.67	0.673	40.87	38.25	0.442	0.527	0.275	0.431
CZH141029	76.53	2.52 ± 0.04	1.33	189.67	74.33	0.672	38.99	35.54	0.361	0.467	0.207	0.341
CZH15022	83.03	2.03 ± 0.21	1.00	166.33	77.00	0.690	36.19	31.04	0.405	0.525	0.240	0.413
CZH15027	72.52	2.32 ± 0.27	3.67	191.38	74.00	0.665	39.63	31.78	0.399	0.465	0.254	0.352
CZH15034	79.59	1.92 ± 0.17	6.33	199.92	72.33	0.697	38.48	35.76	0.397	0.491	0.231	0.363
CZH15036	79.64	1.98 ± 0.23	3.67	205.13	72.67	0.695	38.90	31.69	0.436	0.543	0.276	0.408
CZH15037	85.22	1.53 ± 0.29	7.00	189.75	71.33	0.710	40.76	34.23	0.444	0.553	0.297	0.435
CZH15043	79.99	2.02 ± 0.32	7.33	196.38	72.67	0.710	39.74	33.60	0.493	0.616	0.332	0.517
CZH15048	74.16	2.45 ± 0.52	4.67	169.46	70.33	0.668	43.14	37.98	0.460	0.516	0.319	0.462
CZH15049	75.04	2.21 ± 0.59	5.00	174.67	70.67	0.700	39.81	36.25	0.415	0.462	0.268	0.381
PGS65	78.26	2.52 ± 0.45	4.00	197.46	75.00	0.703	37.86	28.82	0.426	0.484	0.258	0.381
MH1547	72.17	2.54 ± 0.01	3.00	194.58	72.33	0.675	36.77	33.97	0.386	0.468	0.235	0.379
MRI 624	80.82	1.84 ± 0.58	7.00	180.38	74.67	0.660	38.49	32.22	0.390	0.469	0.241	0.353
MRI 634	71.21	2.20 ± 0.33	5.33	196.17	72.33	0.668	43.86	42.19	0.435	0.540	0.299	0.467
MRI 614	79.22	2.37 ± 0.37	6.67	213.38	72.33	0.687	42.72	37.68	0.489	0.563	0.332	0.484

References

1. FAO. Food and Agriculture Organization of the United Nations; Statistic Division. 2017. Available online: <http://faostat.fao.org/> (accessed on 28 August 2017).
2. Cairns, J.E.; Hellin, J.; Sonder, K.; Araus, J.L.; MacRobert, J.F.; Thierfelder, C.; Prasanna, B.M. Adapting maize production to climate change in sub-Saharan Africa. *Food Secur.* **2013**, *5*, 345–360. [[CrossRef](#)]
3. Buerkert, A.; Bationo, A.; Piepho, H.P. Efficient phosphorus application strategies for increased crop production in sub-Saharan West Africa. *Field Crop. Res.* **2001**, *72*, 1–15. [[CrossRef](#)]
4. Leff, B.; Ramankutty, N.; Foley, J.A. Geographical distribution of major crops across the world. *Glob. Biogeochem. Cycles* **2004**, *18*, GB1009. [[CrossRef](#)]
5. Masuka, B.; Araus, J.L.; Das, B.; Sonder, K.; Cairns, J.E. Phenotyping for Abiotic Stress Tolerance in Maize. *J. Integr. Plant Biol.* **2012**, *54*, 238–249. [[CrossRef](#)]
6. Hergert, G.; Ferguson, R.; Gotway, C.; Peterson, T. The Impact of VRT-N application on N use efficiency of furrow irrigated corn. In Proceedings of the 3rd International Conference on Precision Agriculture, Minneapolis, MN, USA, 23–26 June 1996; pp. 389–397.
7. Delgado, J.A.; Follett, R.F.; Buchleiter, G.; Stuebe, A.; Sparks, R.T.; Dillon, M.A.; Thompson, A.; Thompson, K. Use of geospatial information for N management and conservation of underground water quality. In Proceedings of the 3rd International Conference on Geospatial Information in Agriculture and Forestry, Denver, CO, USA, 5–7 November 2001; pp. 5–7.
8. Wang, R.L.; Stec, A.; Hey, J.; Lukens, L.; Doebley, J. The limits of selection during maize domestication. *Nature* **1999**, *398*, 236–239. [[CrossRef](#)]
9. Cairns, J.E.; Sanchez, C.; Vargas, M.; Ordoñez, R.; Araus, J.L. Dissecting Maize Productivity: Ideotypes Associated with Grain Yield under Drought Stress and Well-Watered Conditions. *J. Integr. Plant Biol.* **2012**, *54*, 1007–1020. [[CrossRef](#)]
10. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [[CrossRef](#)]
11. Reynolds, M.; Foulkes, J.; Furbank, R.; Griffiths, S.; King, J.; Murchie, E.; Parry, M.; Slafer, G. Achieving yield gains in wheat. *Plant Cell Environ.* **2012**, *35*, 1799–1823. [[CrossRef](#)] [[PubMed](#)]
12. Debaeke, P.; Route, P.; Justes, E. Relationship between the normalize SPAD index and the nitrogen nutrition index: Application to durum wheat. *J. Plant Nutr.* **2006**, *29*, 75–92. [[CrossRef](#)]
13. Rorie, R.L.; Purcell, L.C.; Morteza, M.; Karcher, D.E.; King, C.A.; Marsh, M.C.; Longer, D.E. Association of “greenness” in corn with yield and leaf nitrogen concentration. *Agron. J.* **2011**, *103*, 529–535. [[CrossRef](#)]
14. Saleh, B. Effect of salt stress on growth and chlorophyll content of some cultivated cotton varieties grown in Syria. *Commun. Soil Sci. Plant Anal.* **2012**, *43*, 1976–1983. [[CrossRef](#)]
15. Vollmann, J.; Walter, H.; Sato, T.; Schweiger, P. Digital image analysis and chlorophyll metering for phenotyping the nodulation in soybean. *Comput. Electron. Agric.* **2011**, *75*, 190–195. [[CrossRef](#)]
16. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, 1–4. [[CrossRef](#)]
17. Bänziger, M.; Edmeades, G.O.; Beck, D.; Bellon, M. *Breeding for Drought and Nitrogen Stress Tolerance in Maize: From Theory to Practice*; CIMMYT: El Batán, Mexico, 2000; Available online: <https://repository.cimmyt.org/bitstream/handle/10883/765/68579.pdf?sequence=1&isAllowed=y> (accessed on 10 April 2019).
18. Nebiker, S.; Annen, A.; Scherrer, M.; Oesch, D. A light-weight multispectral sensor for micro UAV—Opportunities for very high resolution airborne remote sensing. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; International Society for Photogrammetry and Remote Sensing (ISPRS): Beijing, China, 2008; Volume 37(B1).
19. Nex, F.; Remondino, F. UAV for 3D mapping application: A review. *Appl. Geomat.* **2014**, *6*, 1–15. [[CrossRef](#)]
20. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [[CrossRef](#)]
21. Carter, G.A. Reflectance wavebands and indices for remote estimation of photosynthesis and stomatal conductance in pine canopies. *Remote Sens. Environ.* **1998**, *63*, 61–72. [[CrossRef](#)]
22. Huete, A.R. A soil adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
23. Gamon, J.A.; Peñuelas, J.; Field, C.B. 27 A narrow waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [[CrossRef](#)]

24. Huang, W.; Lamb, D.W.; Zheng, N.; Zhang, Y.; Liu, L.; Wang, J. Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precis. Agric.* **2007**, *8*, 187–197. [[CrossRef](#)]
25. Peñuelas, J.; Gamon, J.A.; Griffin, K.L.; Field, C.B. Assessing type, biomass, pigment composition and photosynthetic efficiency of aquatic vegetation from spectral reflectance. *Remote Sens. Environ.* **1993**, *46*, 110–118. [[CrossRef](#)]
26. Peñuelas, J.; Filella, I.; Gamon, J.A. Assessment of photosynthetic radiation—Use efficiency with spectral reflectance. *New Phytol.* **1995**, *131*, 291–296. [[CrossRef](#)]
27. Chappelle, E.W.; Kim, M.S.; McMurtrey III, J.E. Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. *Remote Sens. Environ.* **1992**, *39*, 239–247. [[CrossRef](#)]
28. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
29. Thenkabail, P.S.; Smith, R.B.; De Pauw, E. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sens. Environ.* **2000**, *71*, 158–182. [[CrossRef](#)]
30. Thenkabail, P.S.; Smith, R.B.; De Pauw, E. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 607–622.
31. Peñuelas, J.; Filella, I. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci.* **1998**, *3*, 151–156. [[CrossRef](#)]
32. Serrano, L.; Filella, I.; Peñuelas, J. Remote Sensing of Biomass and Yield of Winter Wheat under Different Nitrogen Supplies. *Crop Sci.* **2000**, *40*, 723. [[CrossRef](#)]
33. Gao, B.-C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
34. Gitelson, A.A.; Merzlyak, M.N. Signature Analysis of Leaf Reflectance Spectra: Algorithm Development for Remote Sensing of Chlorophyll. *J. Plant Physiol.* **1996**, *148*, 494–500. [[CrossRef](#)]
35. Murchie, E.H.; Kefauver, S.; Araus, J.L.; Muller, O.; Rascher, U.; Flood, P.J.; Lawson, T. Measuring the dynamic photosynthetic. *Ann. Bot.* **2018**. [[CrossRef](#)]
36. Araus, J.L.; Slafer, G.A.; Royo, C.; Serret, M.D. Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* **2008**, *27*, 377–412. [[CrossRef](#)]
37. Weber, V.; Araus, J.; Cairns, J.; Sanchez, C.; Melchinger, A.; Orsini, E. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. *Field Crops Res.* **2012**, *128*, 82–90. [[CrossRef](#)]
38. Vergara-Diaz, O.; Kefauver, S.C.; Elazab, A.; Nieto-Taladriz, M.T.; Araus, J.L. Grain yield to SSES in yellow-rusted durum wheat estimated using digital and conventional parameters under field conditions. *Crop J.* **2015**, *3*, 200–210. [[CrossRef](#)]
39. Kefauver, S.C.; El-Haddad, G.; Vergara-Diaz, O.; Araus, J.L. RGB picture vegetation indexes for High-Throughput Phenotyping Platforms (HTPPs). In Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology XVII, Toulouse, France, 14 October 2015. [[CrossRef](#)]
40. Chang, Y.-C.; Reid, J.F. RGB calibration for color image analysis in machine vision. *IEEE Trans. Image Process.* **1996**, *5*, 1414–1422. [[CrossRef](#)]
41. Fernández, P.D.M.; Peña, F.A.G.; Ren, T.I.; Leandro, J.J.G. Fast and Robust Multiple ColorChecker Detection using Deep Convolutional Neural Networks. *Image Vis. Comput.* **2018**. [[CrossRef](#)]
42. Lebourgeois, V.; Bégué, A.; Labbé, S.; Mallavan, B. Can Commercial Digital Cameras Be Used as Multispectral. *Sensors* **2008**, *8*, 7300–7322. [[CrossRef](#)]
43. Casadesús, J.; Kaya, Y.; Bort, J.; Nachit, M.M.; Araus, J.L.; Amor, S.; Ferrazzano, G.; Maalouf, F.; Maccaferri, M.; Martos, V.; et al. Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. *Ann. Appl. Biol.* **2007**, *150*, 227–236. [[CrossRef](#)]
44. Zaman-Allah, M.; Vergara, O.; Araus, J.L.; Tarekegne, A.; Magorokosho, C.; Zarco-Tejada, P.J.; Cairns, J. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* **2015**, *11*, 35. [[CrossRef](#)]

45. Kefauver, S.C.; Vicente, R.; Vergara-Díaz, O.; Fernandez-Gallego, J.A.; Kerfal, S.; Lopez, A.; Melichar, J.P.E.; Serret Molins, M.D.; Araus, J.L. Comparative UAV and Field Phenotyping to Assess Yield and Nitrogen Use Efficiency in Hybrid and Conventional Barley. *Front. Plant Sci.* **2017**, *8*. [[CrossRef](#)]
46. Pinter, P.J., Jr.; Hatfield, J.L.; Schepers, J.S.; Barnes, E.M.; Moran, M.S.; Daughtry, C.S.T.; Upchurch, D.R. Remote Sensing for Crop Management. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 647–664. [[CrossRef](#)]
47. El-Shikha, D.M.; Barnes, E.M.; Clarke, T.R.; Hunsaker, D.J.; Pinter, P.J.; Waller, P.M.; Thompson, T.L. Remote sensing of cotton nitrogen status using the canopy chlorophyll content index (CCCI). *Trans. ASABE* **2008**, *51*, 73–82. [[CrossRef](#)]
48. Trussell, H.J.; Vrhel, M.J.; Saber, E. Color image processing. *IEEE Signal Process. Mag.* **2005**, *22*, 14–22. [[CrossRef](#)]
49. Masuka, B.; Atlin, G.N.; Olsen, M.; Magorokosho, C.; Labuschagne, M.; Crossa, J.; Bänziger, M.; Pixley, K.V.; Vivek, B.S.; Von Biljon, A.; et al. Gains in maize genetic improvement in eastern and southern Africa: I. CIMMYT hybrid breeding pipeline. *Crop Sci.* **2017**, *57*, 168–179. [[CrossRef](#)]
50. Zaidi, P.H.; Zaman-Allah, M.; Trachsel, S.; Seetharam, K.; Cairns, J.E.; Vinayan, M.T. *Phenotyping for Abiotic Stress Tolerance in Maize—Heat Stress*; A Field Manual; CIMMYT: Hyderabad, India, 2016.
51. Makanza, R.; Zaman-Allah, M.; Cairns, J.; Magorokosho, C.; Tarekegne, A.; Olsen, M.; Prasanna, B. High-Throughput Phenotyping of Canopy Cover and Senescence in Maize Field Trials Using Aerial Digital Canopy Imaging. *Remote Sens.* **2018**, *10*, 330. [[CrossRef](#)]
52. Bänziger, M.; Lafitte, H. Efficiency of Secondary Traits for Improving Maize for Low-Nitrogen Target Environments. *Crop Sci.* **1997**, *37*, 1110. [[CrossRef](#)]
53. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Meth.* **2012**, *9*, 671–675. [[CrossRef](#)]
54. Hunt, E.R.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *21*, 103–112. [[CrossRef](#)]
55. Hunt, E.R.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey, J.E.; Walthall, C.L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
56. Bai, X.D.; Cao, Z.G.; Wang, Y.; Yu, Z.H.; Zhang, X.F.; Li, C.N. Crop segmentation from images by morphology modeling in the CIE L^a a^b b^b color space. *Comput. Electron. Agric.* **2013**, *99*, 21–34. [[CrossRef](#)]
57. Fernandez-Gallego, J.A.; Kefauver, S.C.; Vatter, T.; Gutiérrez, N.A.; Nieto-Taladriz, M.T.; Araus, J.L. Low-cost assessment of grain yield in durum wheat using RGB images. *Eur. J. Agron.* **2019**, *105*, 146–156. [[CrossRef](#)]
58. Gracia-Romero, A.; Vergara-Díaz, O.; Thierfelder, C.; Cairns, J.E.; Kefauver, S.C.; Araus, J.L. Phenotyping Conservation Agriculture Management Effects on Ground and Aerial Remote Sensing Assessments of Maize Hybrids Performance in Zimbabwe. *Remote Sens.* **2018**, *10*, 349. [[CrossRef](#)]
59. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2016.
60. Racine, J. RStudio: A Platform-Independent IDE for R and Sweave. *J. Appl. Econom.* **2011**, *27*, 167–172. [[CrossRef](#)]
61. Below, F.E. Nitrogen metabolism and crop productivity. In *Handbook of Plant and Crop Physiology*; Pessaraki, M., Ed.; Marcel Dekker, Inc.: New York, NY, USA; Basel, Switzerland, 1995; pp. 275–301.
62. Uribe-larrea, M.; Crafts-Brandner, S.J.; Below, F.E. Physiological N response of field-grown maize hybrids (*Zea mays* L.) with divergent yield potential and grain protein concentration. *Plant Soil* **2009**, *316*, 151–160. [[CrossRef](#)]
63. Ding, L.; Wang, K.J.; Jiang, G.M.; Biswas, D.K.; Xu, H.; Li, L.F.; Li, Y.H. Effects of nitrogen deficiency on photosynthetic traits of maize hybrids released in different years. *Ann. Bot.* **2005**, *96*, 925–930. [[CrossRef](#)]
64. Hirel, B.; Martin, A.; Terce-Laforgue, T.; Gonzalez-Morob, M.; Estavillo, J. Physiology of maize I: A comprehensive and integrated view of nitrogen metabolism in a C4 plant. *Physiol. Plant.* **2005**, *124*, 167–177. [[CrossRef](#)]
65. Muruli, B.I.; Paulsen, G.M. Improvement of nitrogen use efficiency and its relationship to other traits in maize. *Maydica* **1981**, *26*, 63–73.
66. Gallais, A.; Coque, M. Genetic variation and selection for nitrogen use efficiency in maize: A synthesis. *Maydica* **2005**, *50*, 531–537.

67. Westgate, M.E. Physiology of flowering in maize: Identifying avenues to improve kernel set during drought. In *Developing Drought and Low-N Tolerant Maize*; Edmeades, G.O., Bänziger, M., Mickelson, H.R., Pena-Valdivia, C.B., Eds.; CIMMYT: El Batán, Mexico, 1997; pp. 136–141.
68. Beyene, Y.; Mugo, S.; Semagn, K.; Asea, G.; Trevisan, W.; Tarekegne, A.; Tefera, T.; Gethi, J.; Kiula, B.; Gakunga, J.; et al. Genetic distance among doubled haploid maize lines and their testcross performance under drought stress and non-stress conditions. *Euphytica* **2013**, *192*, 379–392. [[CrossRef](#)]
69. Edmeades, G.O.; Bolaños, J.; Chapman, S.C.; Lafitte, H.R.; Bänziger, M. Selection improves tolerance to mid/late season drought in tropical maize populations. Gains in biomass, grain yield and harvest index. *Crop Sci.* **1999**, *39*, 1306–1315. [[CrossRef](#)]
70. Bänziger, M.; Edmeades, G.O.; Lafitte, H.R. Selection for drought tolerance increases maize yields over a range of N levels. *Crop Sci.* **1999**, *39*, 1035–1040. [[CrossRef](#)]
71. Araus, J.L.; Slafer, G.A.; Reynolds, M.P.; Royo, C. Plant breeding and drought in C3 cereals: What should we breed for. *Ann. Bot.* **2002**, *89*, 925–940. [[CrossRef](#)]
72. Pattee, H.E.; Giesbrecht, F.G.; Young, C.T. Comparison of Peanut Butter Color Determination by CIELAB L * a * b * and Hunter Color-Difference Methods and the Relationship of Roasted Peanut Color to Roasted Peanut Flavor Response. *J. Agric. Food Chem.* **1991**, *39*, 519–523. [[CrossRef](#)]
73. Penczek, J.; Boynton, P.A.; Splett, D. Color Error in Digital Camera Image Capture Process. *J. Digit. Imaging* **2014**, *27*, 182–191. [[CrossRef](#)]
74. Berra, E.F.; Gaulton, R.; Barr, S. Commercial Off-the-Shelf Digital Cameras on Unmanned Aerial Vehicles for Multitemporal Monitoring of Vegetation Reflectance and NDVI. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4878–4886. [[CrossRef](#)]
75. Yu, W. Practical Anti-Vignetting Methods for Digital Cameras. *IEEE Trans. Consum. Electron.* **2004**, *50*, 975–983. [[CrossRef](#)]
76. Lukina, E.V.; Stone, M.L.; Raun, W.R. Estimating vegetation coverage in wheat using digital images. *J. Plant Nutr.* **1999**, *22*, 341–350. [[CrossRef](#)]
77. Sankaran, S.; Khot, L.R.; Espinoza, C.Z.; Jarolmasjed, S.; Sathuvalli, V.R.; Vandemark, G.J.; Miklas, P.N.; Carter, A.H.; Pumphrey, M.O.; Knowles, R.R.N.; et al. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *Eur. J. Agron.* **2015**, *70*, 112–123. [[CrossRef](#)]
78. Aparicio, N.; Villegas, D.; Casadesus, J.; Araus, J.L.; Royo, C. Spectral vegetation indices as nondestructive tools for determining durum wheat yield. *Agron. J.* **2000**, *92*, 83–91. [[CrossRef](#)]
79. Maresma, Á.; Ariza, M.; Martínez, E.; Lloveras, J.; Martínez-Casasnovas, J. Analysis of Vegetation Indices to Determine Nitrogen Application and Yield Prediction in Maize (*Zea mays* L.) from a Standard UAV Service. *Remote Sens.* **2016**, *8*, 973. [[CrossRef](#)]
80. Vergara-Díaz, O.; Zaman-Allah, M.; Masuka, B.; Hornero, A.; Zarco-Tejada, P.; Prasanna, B.; Cairns, J.; Araus, J. A Novel Remote Sensing Approach for Prediction of Maize Yield Under Different Conditions of Nitrogen Fertilization. *Front. Plant Sci.* **2016**, *7*, 666. [[CrossRef](#)]
81. Araus, J.L.; Kefauver, S.C.; Zaman-Allah, M.; Olsen, M.S.; Cairns, J.E. Translating high throughput phenotyping into genetic gain. *Trend. Plant Sci.* **2018**, *23*, 451–466. [[CrossRef](#)]
82. Moriondo, M.; Maselli, F.; Bindi, M. A simple model of regional wheat yield based on NDVI data. *Eur. J. Agron.* **2007**, *26*, 266–274. [[CrossRef](#)]
83. Gitelson, A.; Viña, A.; Arkebauer, T.; Rundquist, D.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*. [[CrossRef](#)]
84. Daughtry, C.; Hunt, E.; McMurtrey, J. Assessing crop residue cover using shortwave infrared reflectance. *Remote Sens. Environ.* **2004**, *90*, 126–134. [[CrossRef](#)]
85. Hobbs, T.J. The use of NOAA-AVHRR NDVI data to assess herbage production in the arid rangelands of Central Australia. *Int. J. Remote Sens.* **1995**, *16*, 1289–1302. [[CrossRef](#)]
86. Elazab, A.; Bort, J.; Zhou, B.; Serret, M.D.; Nieto-Taladriz, M.T.; Araus, J.L. The combined use of vegetation indices and stable isotopes to predict durum wheat grain yield under contrasting water conditions. *Agric. Water Manag.* **2015**, *158*, 196–208. [[CrossRef](#)]
87. Marti, J.; Bort, J.; Slafer, G.A.; Araus, J.L. Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index. *Ann. Appl. Bot.* **2007**, *150*, 253–257. [[CrossRef](#)]
88. Li, Y.; Chen, D.; Walker, C.N.; Angus, J.F. Estimating the nitrogen status of crops using a digital camera. *Field Crops Res.* **2010**, *118*, 221–227. [[CrossRef](#)]

89. Vetsch, J.A.; Randall, G.W. Corn production as affected by nitrogen application timing and tillage. *Agron. J.* **2004**, *96*, 502–509. [[CrossRef](#)]
90. Chapman, S.; Baretto, H. Using a chlorophyll meter to estimate specific leaf nitrogen of tropical maize during vegetative growth. *Agron. J.* **1997**, *89*, 557–562. [[CrossRef](#)]
91. Dwyer, L.M.; Anderson, A.M.; Ma, B.L.; Stewart, D.W.; Tollenaar, M.; Gregorich, E. Quantifying the nonlinearity in chlorophyll meter response to corn leaf nitrogen concentration. *Can. J. Plant Sci.* **1995**, *75*, 179–182. [[CrossRef](#)]
92. Hirel, B.; Lea, P.J. Ammonia assimilation. In *Plant Nitrogen*; INRA Springer: Berlin/Heidelberg, Germany, 2001; pp. 79–99.
93. Masclaux, C.; Quillere, I.; Gallais, A.; Hirel, B. The challenge of remobilization in plant nitrogen economy. A survey of physio- agronomic and molecular approaches. *Ann. Appl. Biol.* **2001**, *138*, 69–81. [[CrossRef](#)]
94. Teal, R.K.; Girma, B.K.; Freeman, K.W.; Arnall, D.B. In-season prediction of corn grain yield potential using normalized difference vegetation index. *Agron. J.* **2006**, *98*, 1488–1494. [[CrossRef](#)]
95. Ahmad, I.S.; Reid, J.F. Evaluation of Colour Representations for Maize Images. *J. Agric. Eng. Res.* **1996**, *63*, 185–195. [[CrossRef](#)]
96. Morgounov, A.; Gummadov, N.; Belen, S.; Kaya, Y.; Keser, M.; Mursalova, J. Association of digital photo parameters and NDVI with winter wheat grain yield in variable environments. *Turk. J. Agric. For.* **2014**, *38*, 624–632. [[CrossRef](#)]
97. Gracia-Romero, A.; Kefauver, S.; Vergara-Díaz, O.; Zaman-Allah, M.; Prasanna, B.; Cairns, J.; Araus, J. Comparative Performance of Ground vs. Aerially Assessed RGB and Multispectral Indices for Early-Growth Evaluation of Maize Performance under Phosphorus Fertilization. *Front. Plant Sci.* **2017**, *8*, 2004. [[CrossRef](#)]
98. Osborne, S.; Schepers, J.; Francis, D.; Schlemmer, M. Detection of Phosphorus and Nitrogen Deficiencies in Corn Using Spectral Radiance Measurements. *Agron. J.* **2002**, *94*, 1215–1221. [[CrossRef](#)]
99. Blackmer, T.M.; Schepers, J.S.; Varvel, G.E.; Walter-Shea, E.A. Nitrogen deficiency detection using reflected shortwave radiation from irrigated corn canopies. *Agron. J.* **1996**, *88*, 1–5. [[CrossRef](#)]
100. Ma, B.L.; Morrison, M.J.; Dwyer, M. Canopy light reflectance and field greenness to assess nitrogen fertilization and yield of maize. *Agron. J.* **1996**, *88*, 915–920. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

CHAPTER 2

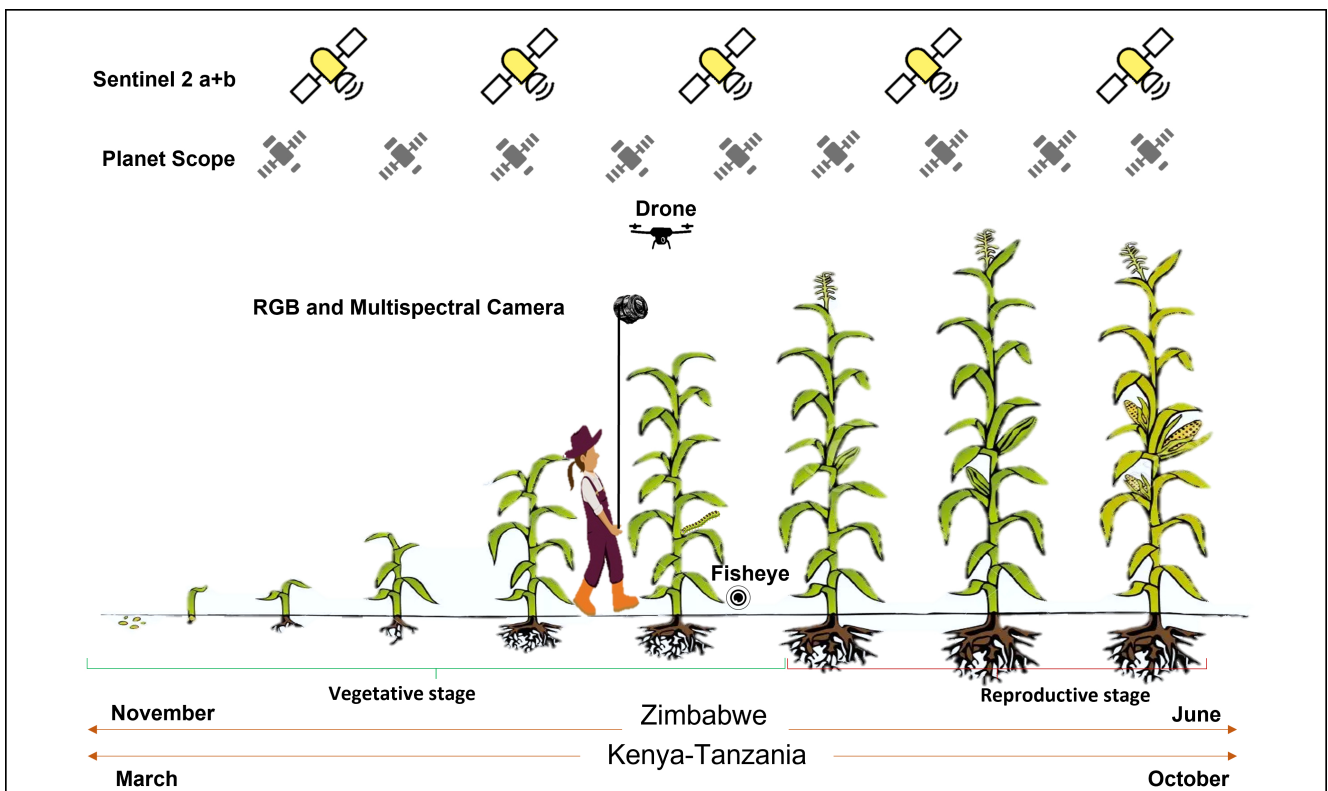


CHAPTER 2

Regional Monitoring of Fall Armyworm (FAW) Using Early Warning Systems

Ma. Luisa Buchailot, Jill Cairns,
Esnath Hamadziripi, Kenneth Wilson,
David Hughes, John Chelal, Peter McCloskey,
Annalyse Kehe, Nicholas Clinton,
José Luis Araus and Shawn C. Kefauver

Published in:
Remote Sensing (2022)





Article

Regional Monitoring of Fall Armyworm (FAW) Using Early Warning Systems

Ma. Luisa Buchailot ^{1,2}, Jill Cairns ³ , Esnath Hamadziripi ³ , Kenneth Wilson ⁴, David Hughes ⁵, John Chelal ⁶, Peter McCloskey ⁵, Annalyse Kehs ⁵, Nicholas Clinton ⁷, José Luis Araus ^{1,2} and Shawn C. Kefauver ^{1,2,*}

¹ Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain

² AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Spain

³ CIMMYT Western Africa Regional Office, Harare P.O. Box MP163, Zimbabwe

⁴ Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK

⁵ Department of Entomology & Biology, Penn State University, University Park, PA 16801, USA

⁶ Department of Agriculture and Biotechnology, Moi University, Eldoret 3900-30100, Kenya

⁷ Google, 1600 Amphitheatre Pkwy, Mountain View, CA 94043, USA

* Correspondence: sckefauver@ub.edu

Abstract: The second United Nations Sustainable Development Goal (SDG2), zero hunger, aims to improve the productivity, food security, nutrition, and sustainability of small-scale farmers. The fall armyworm (FAW, *Spodoptera frugiperda*) has been devastating to smallholder farmer food security since it spread to sub-Saharan Africa in 2016, who have suffered massive crop losses, particularly maize, an important staple for basic sustenance. Since the FAW mainly devours green leaf biomass during the maize vegetative growth stage, the implementation of remote sensing technologies offers opportunities for monitoring the FAW. Here, we developed and tested a Sentinel 2 a+b satellite-based monitoring algorithm based on optimized first-derivative NDVI time series analysis using Google Earth Engine. For validation, we first employed the FAO Fall Armyworm Monitoring and Early Warning System (FAMEWS) mobile app data from Kenya, and then subsequently conducted field validation campaigns in Zimbabwe, Kenya, and Tanzania. Additionally, we directly observed loss of green biomass during maize vegetative growth stages caused by the FAW, confirming the observed signals of loss of the leaf area index (LAI) and the total green biomass (via the NDVI). Preliminary analyses suggested that satellite monitoring of small-scale farmer fields at the regional level may be possible with an NDVI first-derivative time series anomaly analysis using ESA Sentinel 2 a+b ($R^2 = 0.81$). Commercial nanosatellite constellations, such as PlanetScope, were also explored, which may offer benefits from greater spatial resolution and return interval frequency. Due to other confounding factors, such as clouds, intercropping, weeds, abiotic stresses, or even other biotic pests (e.g., locusts), validation results were mixed. Still, maize biomass anomaly detection for monitoring the FAW using satellite data could help confirm the presence of the FAW with the help of expanded field-based monitoring through the FAO FAMEWS app.

Keywords: maize; fall armyworm; *Spodoptera frugiperda*; remote sensing; Google Earth Engine; Sentinel 2; planet; sustainable development goals; Africa



Citation: Buchailot, M.L.; Cairns, J.; Hamadziripi, E.; Wilson, K.; Hughes, D.; Chelal, J.; McCloskey, P.; Kehs, A.; Clinton, N.; Araus, J.L.; et al. Regional Monitoring of Fall Armyworm (FAW) Using Early Warning Systems. *Remote Sens.* **2022**, *14*, 5003. <https://doi.org/10.3390/rs14195003>

Academic Editor: Jianxi Huang

Received: 28 July 2022

Accepted: 30 September 2022

Published: 8 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The sustainable development goals (SDGs) developed by the UN are seventeen prioritized challenges regarding poverty, inequality, climate, environmental degradation, prosperity, peace, justice, and their interactions. SDG2 is “Zero Hunger”, which aims to end hunger globally, achieve food security, improve nutrition, and promote sustainable agriculture. SDG2 specifically prioritizes the support of subsistence and sustainable small-scale farmers with a focus on staple agro-food production [1]. Maize (*Zea mays* L.), after wheat and rice, is the third most important staple crop in the world, providing half of

the daily food energy to Africa and America [2]. In Africa alone, more than 300 million people depend on maize as their main food crop and feed for livestock [3]. In Sub-Saharan Africa (SSA), there are 37 million ha of annually cultivated maize, with 95% of farms being less than 2 ha in size and belong to smallholder farmers [4]. Smallholder farmers grow their plots of maize using almost exclusively family labor, using mostly harvested seeds from previous seasons with few purchased inputs and their production is mostly for household consumption [5].

The fall armyworm (FAW), *Spodoptera frugiperda*, is native to the tropical and subtropical Americas, where it has been known for many decades as an economically detrimental pest of many crop species, especially maize. The FAW is a polyphagous lepidopteran pest widely considered to be one of the most damaging pests; it feeds on over 350 crops species to diverse families, including maize, rice, cabbage, soybean, tomato, cotton, etc. [6]. This pest has a strong preference for maize in the early vegetative stages when the plant presents tender green leaves [7]. Nevertheless, at the reproductive stage, the FAW tends to move to the cob, where it devastates yield quantity and quality [8]. This pest arrived in 2016 to Western Africa and spread quickly across Africa and has continued eastward through the Middle East by 2017, Asia by 2018–2019, and Australia by 2020 (Figure 1), and is expected to threaten Europe eventually [9,10]. The rapid spread of this species is related to its sporadic and long-distance migratory behavior, with the adult moths capable of flying over 100 km on a single night [11]. Some studies have indicated that the effect of FAW damage can cause USD 2.5–6.2 billion losses annually in yield maize across Africa [10,12]. In sub-Saharan Africa (SSA), this has been estimated to have increased by up to USD 13 billion per year after 2018 [13], thus threatening the livelihoods of millions of small farmers.

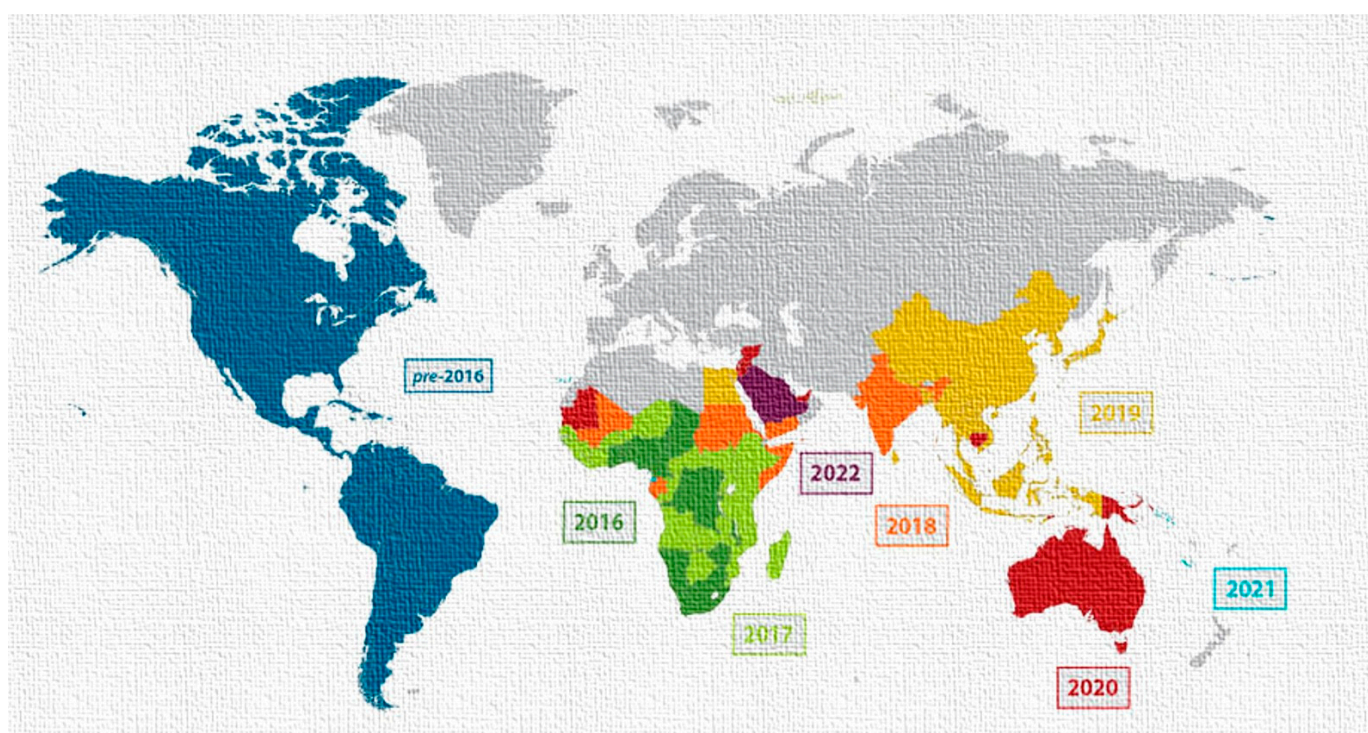


Figure 1. Geographic invasion/expansion of the fall armyworm (FAW), *Spodoptera frugiperda*, across the world (source from [9] adapted).

In America, FAW pest management using chemical pesticides and transgenic maize has succeeded in keeping FAW impacts under control [14,15]. However, for smallholder farmers in SSA affected by the FAW, efforts made in applying available insecticides have been deemed largely ineffective or not economically feasible [11,13]. For that reason,

environmentally friendly and cost-effective strategies, such as integrated pest management (IPM) for FAW management, have been proposed to maintain productivity efficiently and sustainably according to the UN SDG2 guidelines [11]. There are different proposals on how to best manage biotic pests in accordance with different local cultural practices, such as high-quality seeds; optimizing the time of crop planting; rotating the target crop; or intercropping the main staples crop, e.g., maize or wheat mixed with another crop, such as beans, that repels the FAW (called push–pull), which moreover can increase the population of natural enemies such as predators and parasitoids. Besides these different IPM approaches, the application of sand/ash in the maize leaf funnels has also been shown to be a natural deterrent to the FAW when applied at specific larval growth stages [5,10,16].

As mentioned before, the FAW eats the tender leaves of maize during the vegetative stage. The newly hatched FAW larvae that feed on the young maize leaves do not directly cause yield losses because the plant is able to compensate for at least some loss of green biomass and a subsequent reduction in the total leaf area in terms of the leaf area index (LAI) of the crop [16,17]. However, if the older larvae attack the apical meristem or cob, the result is a dead heart, which means wilting and death of the unfurled leaves [10]. For that reason, repeated observations of the LAI and aboveground biomass during the early stages of maize may offer one possibility for the early warning and monitoring of FAW infestation. Here, remote sensing assessments of the LAI may serve to capture this effect at a large scale.

The LAI is a biophysical variable that has long been considered a key crop observation goal by remote sensing researchers working at different scales of observation [6], and there are a number of different technologies available for its measurement [18–20]. For instance, the European Space Agency (ESA) Sentinel-2 a+b satellites were launched relatively recently, coinciding with the invasion of the FAW in Africa, which provides unprecedented remote sensing global coverage at 10 m spatial resolution and 5 day repeat intervals [21]. The Sentinel-2 Toolbox Sentinel Application Platform (SNAP, <https://step.esa.int/main/toolboxes/snap/>, last accessed 25 June 2022) includes radiative transfer model based advanced algorithm calculations of the LAI, with control for calculations over specific geographic areas and time series ranges. In a similar fashion, the normalized difference vegetation index (NDVI), one of the most common remote sensing indexes, provides a simplified yet reliable estimate of green biomass [22–24]. The NDVI is derived from visible and near-infrared reflectance, which is closely related to vegetation presence or vigor, and can also be measured at the ground level with portable sensors [25]. Sentinel-2 a+b, as well as other satellites, provides spectral bands for the calculation of the NDVI, including PlanetScope microsatellite data, which offer a better 3 m spatial resolution and daily coverage of SSA [26]. Similarly, there are low-cost alternatives to the LAI and the NDVI from red–green–blue (RGB) visible spectrum indexes that can be calculated from commercial cameras in the field or on UAVs. In this case, the green area (GA) index represents the percentage of pixels ranging from yellow to bluish-green color, as calculated by the open-source software BreedPix and FIJI [27].

Several studies have been conducted using NDVI time series anomalies, including an assessment of the ecological response to global warming [28], phenological change [29], crop status ([30,31], land cover change [32], and early drought detection [23]. An NDVI time series may be affected by acquisition conditions, such as cloudy (or cloud shadow) vs. clear days. An NDVI time series curve for a crop season cycle should have an increasing slope and maximum NDVI values at peak crop growth and then decrease towards the end of the crop season, representing senescence. NDVI progression across a crop season is sensitive to increasing biomass in the first half of the season and then dominated by decreasing chlorophyll in the second half of the season [28]. For that reason, an NDVI time series relevant to FAW monitoring information should focus on crop development stages (both at the start and in the middle) and consider potential anomalies apart from the FAW. Two NDVI time series anomaly analyses are suggested: one to compare two different years and observe the differences and another to compare an average over a longer time period

to one single year. In both of these cases, observed anomalies in growth and senescence or reductions in grain yield should also be considered against potential variations in climatic conditions season over season [33–37]. One other potential time series anomaly detection approach is to focus on rapid unusual changes, meaning to normalize the change rate of the NDVI curve through plotting the NDVI first derivative, which should help to remove the effects of climatic conditions, and other potential outside effects such as a lack of fertilizer or intercropping, which may vary between seasons [38–42].

The general aim of this study was to implement a cost-effective assessment for FAW monitoring and early warning system (FAMEWS) on sub-Saharan maize fields using different remote sensing technologies. We divided this work into two parts: (i) the development of a preliminary satellite image-based monitoring algorithm combining Google Earth Engine for Sentinel 2 a+b and validated by FAO FAMEWS mobile application data and (ii) a multi-scale field validation campaign of different time series anomaly change detection approaches, focusing on the first-derivative growth pattern analyses of the NDVI using Sentinel 2 a+b and PlanetScope image data during the maize vegetative growth stage, where FAW presence will result in a reduction in the LAI or total green biomass (NDVI) of the crop. During the field campaign and also in manual satellite image analyses, the data will also be further inspected in order to better decipher between potential confounding anomalies apart from the expected FAW impacts.

2. Materials and Methods

2.1. Study Sites

The measurements to monitor the FAW on maize fields were taken in three different countries: Zimbabwe, Tanzania, and Kenya (Figure 2), with efforts focused locally on Mashonaland and Masvingo (Zimbabwe), Arusha (Tanzania), and Western provinces (Kenya).

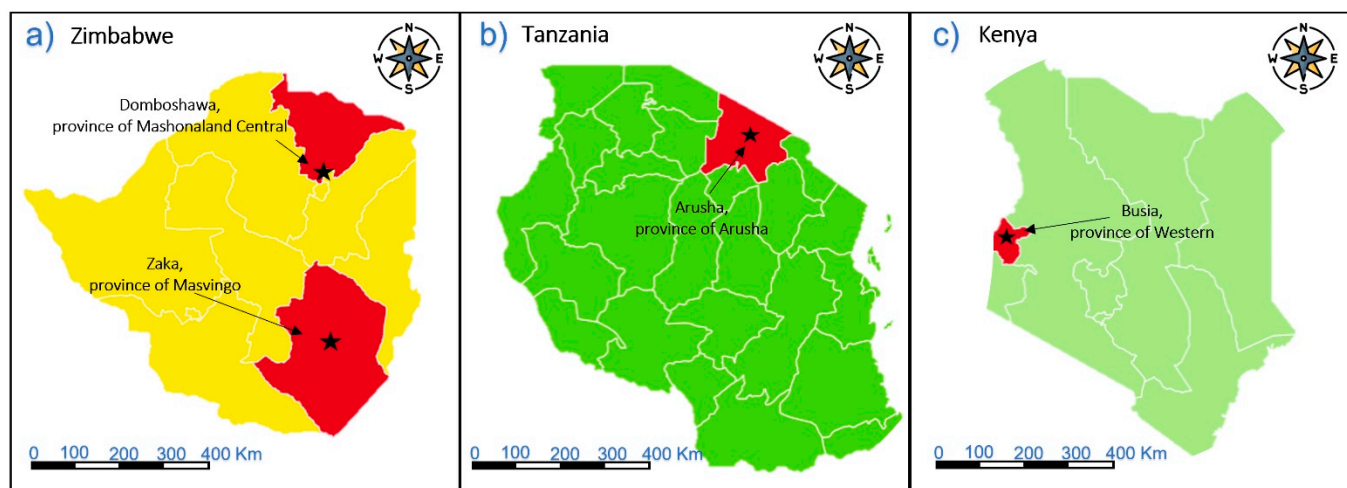


Figure 2. (a) Map of Zimbabwe, in red provinces of Mashonaland Central and Masvingo; (b) map of Tanzania, in the red province of Arusha; (c) map of Kenya, in the red Western province. Asterisks mark the locations where we took the data.

Firstly, we went to Zimbabwe (Figure 2a) from 30 March to 5 April of 2019, where we sampled a total of eight maize fields of different sizes in the two provinces. The maize growing season lasted from November to the end of June (Figure 3). All the fields were rainfed, and the average rainfall was 360 mm across the season.

Secondly, we went to Tanzania (Figure 2b) from 23 to 24 of May of 2019 and sampled twelve fields in the targeted province. The growing season was from March to October (Figure 3). All the fields were rainfed, cultivated during the unimodal season, and the average rainfall was 665 mm over the whole season.

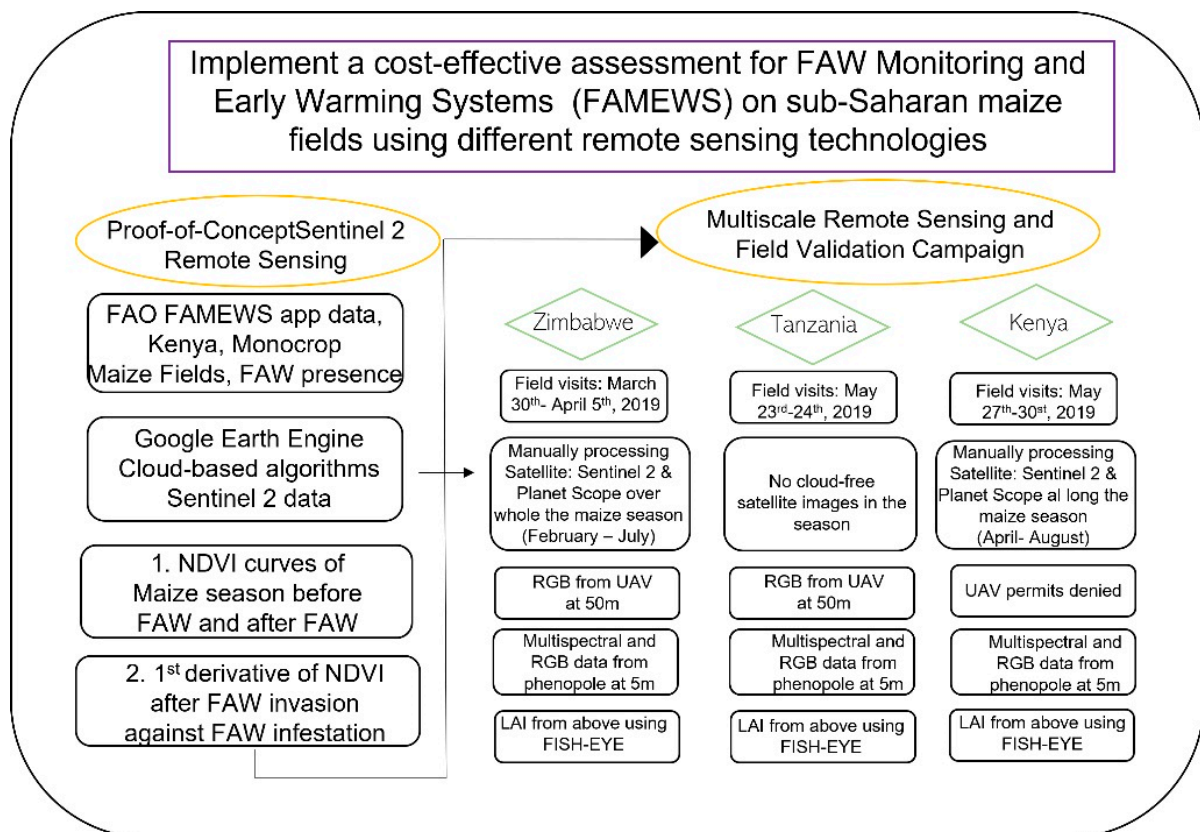


Figure 3. Workflow summary of the principal objectives and the two separate specific phases of the project with all the processes for each location and study phase.

Finally, we went to Kenya (Figure 2c) from 27 to 30 of May of 2019, where we sampled nineteen farms of different sizes in one province. The growing season lasted from March to October (Figure 3). All the fields were rainfed with 690 mm as the average rainfall across the season.

On the other hand, we classified the fields sampled according to the FAO FAMEWS protocols so that the data would be comparable. The data that we requested from the smallholder farmers or extension officers were the date of sowing; the stage of the crop at the same time samples were taken: vegetative, vegetative–reproductive (transition), or reproductive; farming systems: seasonal, rotation, or intercropping; types of the other crops if intercropped; and finally, any pest controls that the farmer used to control FAW: pesticide, IPM, or no treatments. Regarding levels of FAW infestation, we classified them into LOW, MEDIUM, and HIGH. This classification was based on choosing 50 plants from the middle of the field and counting how many of them were infested. We also classified the fields as monocropping and intercropping attending the type of crop. All the fields of maize that we selected were in the vegetative or vegetative–reproductive growth stages as the FAW impacts in the reproductive growth stage were less detectable using remote sensing. In addition, we collected data at each site using the FAMEWS app, which contributed to larger database efforts (<https://www.fao.org/fall-armyworm/monitoring-tools/famews-mobile-app>, last accessed 25 June 2022).

An important highlight could be that the most frequently high levels of FAW infestation occurred with monocropping and the most frequent low level of FAW infestation occurred when intercropping with common beans, which was also identified as part of the IPM strategy for the FAW [13]. In Zimbabwe, farmers had the option of using pesticides and their fields were mostly monocropping compared to the other two countries, where they controlled FAW using cultural controls with ash or handpicking, if any preventative measures were attempted.

2.2. Cloud-based Process Satellite Data: Sentinel 2 a+b

We worked with Google Earth Engine's (GEE's) application programming interface (API), which allows for the direct integration of Google-based server processing with satellite remote sensing data, such as NASA's Landsat series and the ESA Copernicus EO program, but also includes extensive geographic, census, and climate data. Before the validation, through data captured during field campaigns conducted in 2019, we developed our theoretical framework into a fully functional code (<https://code.earthengine.google.com/281c50b864af3836ca26345e09c4c248>, last accessed 25 June 2022) in the GEE platform for extracting the NDVI and a time-lagged NDVI first derivative. This code was used to power image analyses from Sentinel 2 and study the impact and distribution of the FAW across Africa on a country-by-country basis. In the context of agriculture in Africa, cloud cover is an issue, but unique GEE algorithms allowed us to produce cloud-free image mosaics by combining multiple satellite image scenes using cloud masks applied at a pixel level. In collaboration with the United Nations Food and Agriculture Organization (FAO), we used the data from the FAMEWS mobile app (<https://www.fao.org/fall-armyworm/monitoring-tools/famews-mobile-app>, last accessed 25 June 2022), which we also helped to develop and refine, and was developed specifically for gathering data on the invasive pest FAW and assessing impacts on maize fields, and other crops, across the African continent (more specific details provided below in the farmer field sampling and questionnaire protocols) (Figure 3).

The left of Figure 4a is a graphical representation of the theoretical signal of the FAW in terms of vegetation change (Figure 4a); this is not the NDVI, but rather a derivative of the NDVI, (Δ NDVI), where we anticipate opportunities to separate FAW impacts from a normal crop growth phenology. The rationale of the shape of this curve is that the FAW will decrease the crop leaf area during the vegetative growth stage, when it primarily eats the young leaves; however, the FAW will preferentially move to the cob once the crop enters the reproductive stage. In this moment, crop yield will be reduced but the crop leaf area may yet recover/regrow partially. The FAW has also been observed to infest a crop area and then quickly move on to another area, leaving the young crop, even still in a vegetative stage, to recover afterwards. Likewise, a similar recovery signal is observed in the case of a FAW infestation that is then treated afterwards. In Figure 4b, this part was based on the preliminary FAW data from known outbreak sites that were collected by the FAO FAMEWS app prior to our research initiation (<https://www.fao.org/fall-armyworm/monitoring-tools/famews-mobile-app/>, last accessed 25 June 2022).

At first, only a small amount of data was available for testing the capacities of the Google Earth Engine platform, resulting in the loss of >50% of the data available for looking at more detailed changes in crop vegetation growth dynamics. The combined Sentinel 2 a+b satellites, with a 5 day repeat interval once both satellites allowed for improved cloud-free mosaics every 10–15 days. Still, the data points provided by the early FAMEWS app FAW data extracted over one growing season was limited once filtered for cloud masking. Still, this was deemed to be minimally sufficient for the testing of the proof-of-concept for the two selected anomaly detection calculations. This derivative analysis approach may additionally prove useful as an approach for the detection of any crop phenology anomalies related to extreme weather events as suggested above, whereas the fusion of Sentinel 2 satellite data with weather anomalies may help to further improve the data analyses presented here.

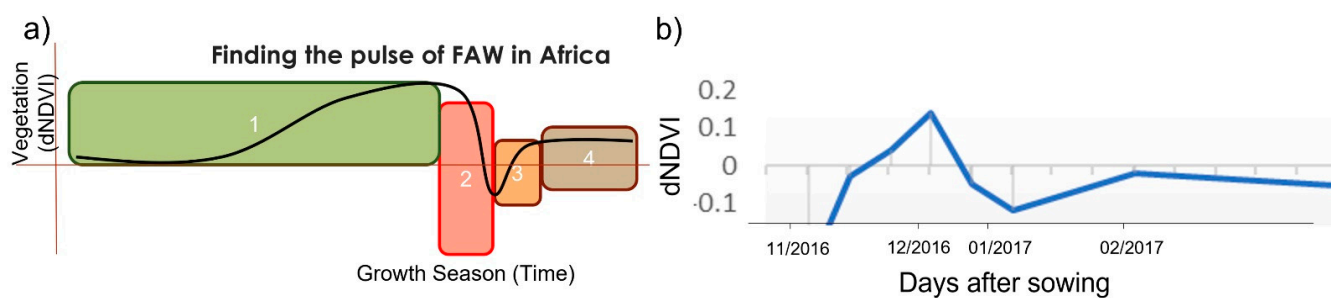


Figure 4. (a) Conceptual framework of the time series analyses that was expected through the four stages of maize response to the FAW. (1) The crop growth as normal, emergence, increased rate of growth until reaching a plateau at near full growth. (2) FAW infestation results in the loss of vegetative cover, marked as “negative growth.” (3) The negative values of the first derivative end with treatment or cob development (as the FAW will move to the cob during the maize reproductive phase). (4) The crop may recover and the FAW may attack again up until the maize reproductive phase-cob development, when subsequent FAW impacts will result in symptoms not observable by remote sensing techniques, being a somewhat hidden yield loss. (b) Exemplary actual Sentinel 2 NDVI time change first-derivative time series at a location of documented high FAW damage, according to FAMEWS. The satellite image analysis results using the GEE API from the preliminary data point of FAW infestation from the 2016–2017 maize field season where the most data points remained after cloud mask filtering at the pixel level.

2.3. Manually Processed Satellite Data: Sentinel 2 a+b and PlanetScope

Satellite image analysis was conducted manually, with data extractions and analysis corresponding to each maize field in vegetative stages from the different campaigns in Zimbabwe and Kenya. No cloud-free satellite images were available for the field campaign dates in Tanzania. For each country, the data were gathered during the respective maize field season. Sentinel-2 satellite data have 12 spectral bands: 443, 494, 560, 665, 704, 740, 781, 834, 944, 1375, 1612, and 2194 nm. The data were captured every 5 days at 10 m of resolution and the images were free to download. We used the Earth Explorer portal of the United States Geological Survey (USGS, <https://earthexplorer.usgs.gov>, last accessed 25 June 2022) to download them, and we applied a 50% cloud filter. For Zimbabwe, we downloaded seven images for each field between 1 February 2019 and 17 May 2019. Regarding Kenya, we downloaded five images for each field between 12 April 2019 and 6 July 2019. Table S1 shows the day of each image of Sentinel 2 a+b. We also used the commercial microsatellite PlanetScope with four bands (485, 545, 630, and 820nm) which offer daily images at 3 m resolution. We used the Planet Explorer (<https://www.planet.com/explorer>, last accessed 25 June 2022, requires access account) to download the images. For Zimbabwe, we downloaded 7 images for each field between 31 January 2019 and 17 May 2019. Regarding Kenya, we downloaded 8 images for each field between 11 April 2019 and 14 July 2019. Table S1 shows the day of each image of PlanetScope. We used the QGIS program (<https://qgis.org>, last accessed 25 June 2022) and we calculated the NDVI from the images of Sentinel-2 and PlanetScope.

For the NDVI [21], we calculated the following equation from Sentinel-2a+b bands:

$$\text{NDVI} = \frac{(\text{B8NIR} : 834\text{nm} - \text{B8RED} : 665\text{nm})}{(\text{B8NIR} : 834\text{nm} + \text{B8RED} : 665\text{nm})} \quad (1)$$

For the NDVI, we calculated the following equation from PlanetScope bands [43]:

$$\text{NDVI} = \frac{(\text{B4NIR} : 820\text{nm} - \text{B3RED} : 630\text{nm})}{(\text{B4NIR} : 820\text{nm} + \text{B3RED} : 630\text{nm})} \quad (2)$$

We used the program QGIS [44] to calculate the NDVI from the images of Sentinel-2 and PlanetScope. Once we had the NDVI for each day, we processed the time series curves, and we calculated the first derivative of the difference between each NDVI along the time series curve.

2.4. Unmanned Airborne Vehicle (UAV) Data Collection and Analysis

We managed to arrange for hiring locally licensed UAV pilots in two different countries as we were not approved with permits to fly UAVs in any of the three project countries. One was in Zimbabwe from 1 to 5 of April, and the second was in Tanzania from 23 to 24 of May. In both Zimbabwe and Tanzania, UAV data were collected 50 m above ground level (a.g.l.) using a DJI Phantom 4, with an RGB camera with a sensor of 1/2.3" CMOS (0.01 m/pixel). We carried out 6 flights in Zimbabwe and 12 flights in Tanzania. With the photographs taken with the UAV, we created orthomosaics for each field with the structure-from-motion Agisoft Metashape (<https://www.agisoft.com>, last accessed 25 June 2022). Once we had the orthomosaics, we used FIJI (<https://imagej.net/software/fiji/>, last accessed 25 June 2022) and the Mosaic tool (<https://gitlab.com/sckefauver/MosaicTool>, last accessed 25 June 2022) to cut the orthomosaics, where we obtained the area of interest. Then, we extracted the green area (GA) index values by applying the Breedpix algorithms incorporated into our free open-source software developed in previous research in collaboration with CIMMYT (<https://github.com/sckefauver/CIMMYT>, last accessed 25 June 2022), which can assess the total green biomass similar to the NDVI, but based on RGB images. In the hue–saturation–intensity (HSI) color space, the hue (H) component describes color chroma traversing the visible spectrum in the form of an angle between 0° and 360°. Thus, the green area (GA) index is the percentage of pixels in the image in the hue range from 60° to 180°, ranging from yellow to bluish green [45,46]. In some countries of Africa, all UAV flights were prohibited, such as in Kenya; thus, we used a QX1 RGB camera on a 5-meter-high telescoping pole and then used the MaizeScanner FIJI plugin, as described above, to calculate the GA index from these images for comparison purposes. The difference between the spatial resolution can be observed in Figure 5.

Spatial Resolution

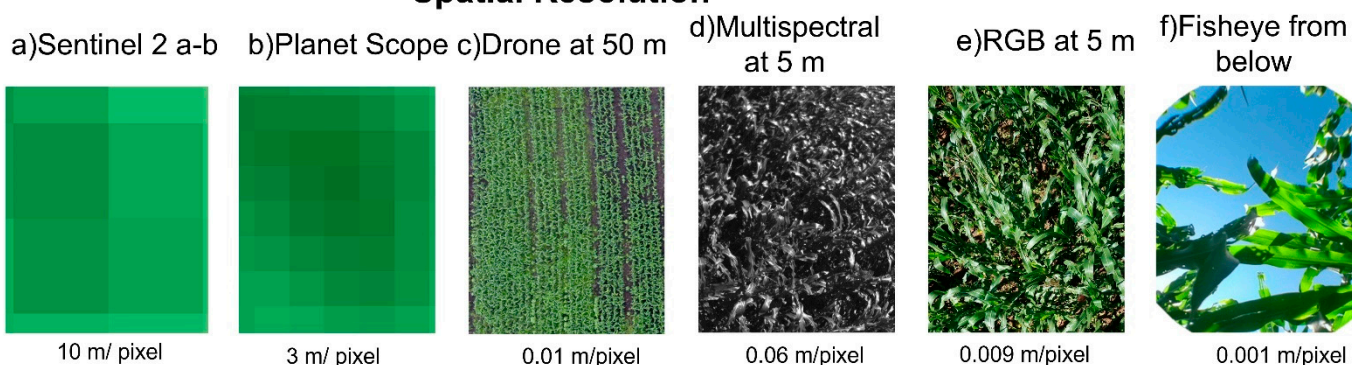


Figure 5. The different spatial resolutions that were used a long the article. (a) Calculated NDVI with Sentinel bands with a resolution of 10 m/pixel. (b) Calculated NDVI with PlanetScope bands with a resolution of 3 m/pixel. (c) Photo taken with a DJI Phantom 4 UAV at an altitude of 50 m and with a resolution of 0.01 m/pixel. (d) Photo taken with a multispectral camera at 5 m of altitude and with a resolution of 0.06 m/pixel. (e) Photo taken with a RGB camera at 5 m of altitude, and with a resolution of 0.009 m/pixel. (f) Photo taken with a mobile phone with a fisheye adapter from below, and with a resolution of 0.001 m/pixel.

2.5. Multispectral Data Collection and Analysis

The multispectral field photos were taken in three different campaigns. First, in Zimbabwe, we collected multispectral data from 1 to 5 of April. Second, we collected multispectral data in Tanzania from 23 to 24 of May. Finally, we collected multispectral

data in Kenya from 27 to 30 of May. For the multispectral data, we used an AIRINOV MultiSpec4-C multispectral camera with four bands (550, 660, 735, and 790 nm) on a 5-m pole. The images were acquired in continuous capture mode every 2 s, while crossing the fields in simulated UAV flight lines, so as to provide as close as full coverage of each field as possible in order to compare field averages for the NDVI across scales. In total, we took images in a total of 34 maize fields: 9 fields in Zimbabwe with an average of 150 images per field, 12 fields in Tanzania, with an average of 200 images per field, and 13 fields in Kenya with an average of 100 images per field. The number of images acquired changed depending on the size of the maize field. After the acquisition of the data, we customized an image processing code in FIJI to align the separate sensor images (550, 660, 735, and 790 nm) and to calculate the NDVI [47] following the equation below:

$$\text{NDVI} = \frac{(\text{B4NIR} : 790\text{nm} - \text{B3RED} : 735\text{nm})}{(\text{B4NIR} : 790\text{nm} + \text{B3RED} : 735\text{nm})} \quad (3)$$

2.6. Hemispherical (Fisheye) Lens Image Processing in CAN-EYE

For this technique, we took a digital image from under crops looking skywards using a hemispherical (fisheye) lens adapter that enables the collection of calibration/validation data if performed correctly using an adequately calibrated lens and complex calculations, as provided by the free software CAN-EYE (<https://www6.paca.inrae.fr/can-eye>, last accessed 25 June 2022) [18,48]. The CAN-EYE software provides an effective LAI, in which pixels are classified interactively, as well as easy processing and camera calibration protocols. We acquired these data across 39 fields, including maize farms from three different field campaigns in Zimbabwe, Tanzania, and Kenya. Hemispherical photographs were taken in jpeg format at the highest possible resolution (4632 × 3474 pixels) with a mobile camera Moto G (S5). We captured the three RGB photos in the middle of each field from below in the middle of four plants.

3. Results

3.1. Sentinel 2 a+b Time Series Analyses

Seasonal averages for known maize fields compared before and after the invasion of the FAW and the first-derivative negative incidences of the LAI/NDVI time series over the whole of the local maize season at each study location in Africa.

Figure 6 shows a comparison of the time series means of the NDVI values of December 2015, before the FAW, and the time series means of the NDVI values of December 2018 against the level of infestation with an $R^2 = 0.401$. The values of the NDVI decrease as the level of infestation increases, following a logical negative correlation.

The resulting signal is a pulse of vegetation change that may be quantified in terms of depth and timing. In the graphic conceptualization of the FAW pulse signal, Sentinel 2 NDVI time change derivative series presented on the right side of Figure 4a shows the satellite image analysis results using the GEE API from the preliminary data point of FAW infestation during the 2016–2017 maize field season (Figure 4b) where the most data points remained after cloud mask filtering at the pixel level.

Figure 7 shows the results for the fall season of 2018 in Kenya using the FAMEWS app v1 data. In this case, the analysis of the first derivative of the NDVI time series of Sentinel-2 satellite image data shows an R^2 of 0.81, with the infestation level of the FAW in each field exhibiting a markedly negative relationship. This was conducted with filtered data to include only vegetative state maize data for Kenya in these months. This suggests that using derivative NDVI time series analysis appears to be less sensitive to year over year factors, and thus better correlates with FAW severity, besides also potentially providing FAW infestation timing, which would be useful for developing an early warning system. No correlation was observed between the level of FAW infestation and NDVI observations.

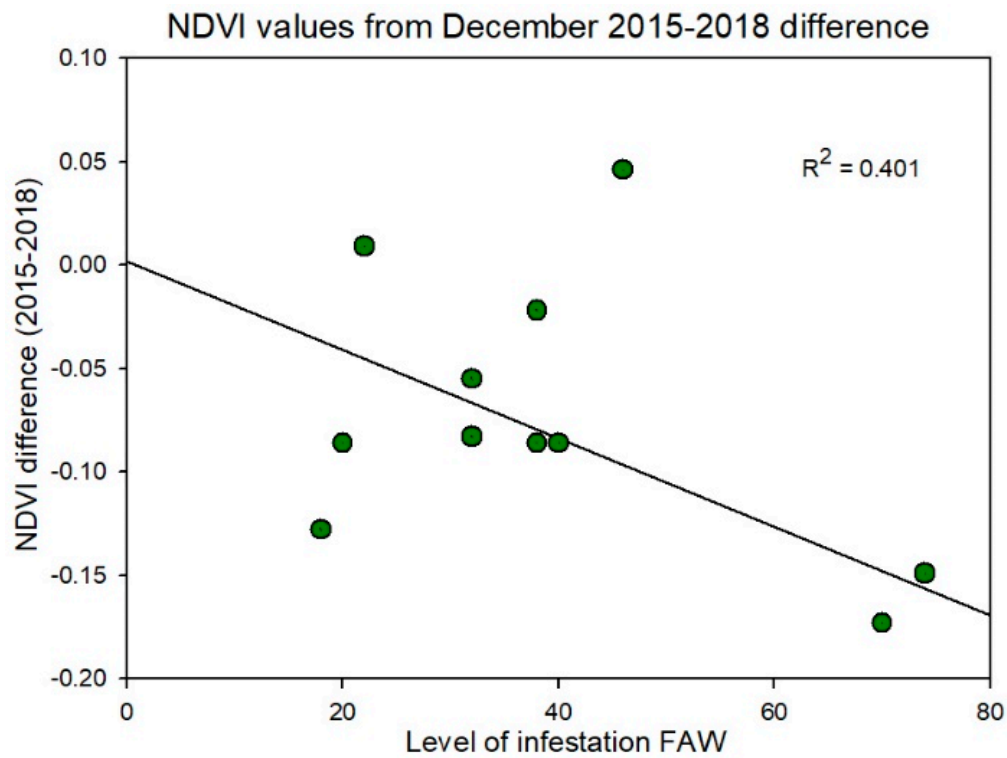


Figure 6. Difference between the NDVI average from December 2015 (farms without the FAW) and the NDVI average from December 2018 (farms with the FAW) against the level infestation of the FAW (0 to 100) in the maize field collected by the FAMEWS app.

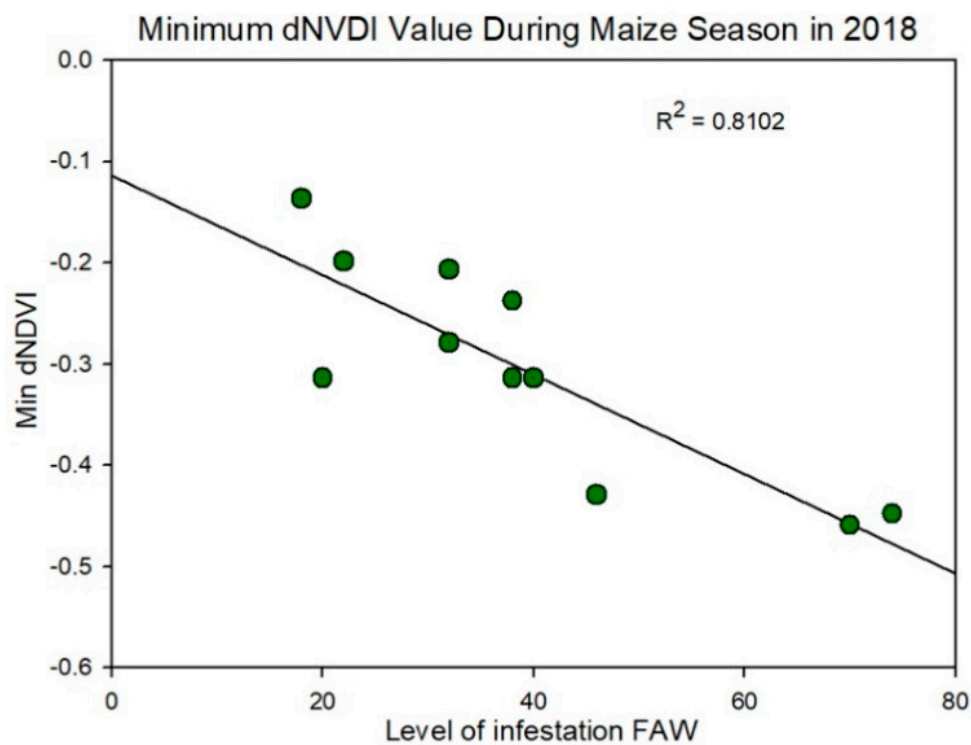


Figure 7. Relation between the minimum value from the entire growth season of the first-derivative NDVI time series against the level of infestation of the FAW (0-100) in the maize field collected by the FAMEWS app.

Since the launch of Sentinel 2b, the repeat interval has been reduced by half, but a very high loss of data was still observed in relation to cloud cover over the three project study sites where satellite information was used, indicating the somewhat limited utility for satellite-based FAW monitoring at a large scale, except in largely cloud-free areas. Vegetation time series and anomaly change detection alone have not proved to be sufficiently reliable for uniquely identifying FAW; however, promising results were achieved from vegetation time series first-derivative analysis based on time step interval-based cloud-free mosaics. Subsequently, these analyses were expanded to include the whole of Africa using the same FAMEWS data for the same time, with poor reproducibility of results potentially related to environmental factors (data not shown).

3.2. Description of the Maize Fields That We Visited in Zimbabwe, Tanzania, and Kenya

For the in-situ validation of the satellite image anomaly detection results presented in Section 3.1, we visited 39 maize fields in the three different countries in total. Table S2 shows details of each field, including country, province, approximate sowing data, and crop stage as vegetative or vegetative–reproductive, and the approximate BBCH crop stage as well. Moreover, we described the farming system of each individual field, i.e., if there was rotation, seasonal crop, or intercropping. However, if the farming system was intercropping, the maize crop shared the field with other crops, including pumpkins, common beans, groundnuts, and sunflowers. We also included information on the level of the FAW infestation in each field, as estimated using the FAMEWS protocol for the field-level FAW infestation and how the smallholder farmers treated FAW, if any treatment was applied.

In Zimbabwe, most of the farmers used pesticides for control, and only one used ash as an alternative treatment. With respect to the field for Tanzania, only one farmer attempted to control the FAW using hand picking, but more detailed information was not available. In Kenya, the majority used ash to try to reduce the FAW infestation, and others did not use anything. Regarding the date of sowing, in Zimbabwe where the sowing months of maize were usually November–December, we can see that the sowing was late in January (Table S1), which could be because the year 2018 was relatively dry [49], which encouraged late sowing practices. On the other hand, in Tanzania and Kenya, almost all the maize fields were planted between March and April, with the exceptions of three fields at beginning of May. Given the present high level of FAW infestation, avoiding late planting is one of the IPM strategies, as some farmers attempted to avoid FAW infestation by adjusting their crop phenological cycles [11].

3.3. Comparison of the NDVI, GA Index, and LAI at Different Spatial Resolution

We conducted observations (NDVI, LAI, and GA index) at three different scales (PlanetScope, phenopole, and fisheye hemispherical lens taken from below the plant and looking upwards) from each country to field scales, all taken on the same day. Therefore, these data were collected in Kenya in the maize fields from 28 to 29 of May. The difference between the spatial resolution can be observed in Figure 5.

In Figure 8, the results show a comparison between the different indices taken with different sensors at different scales. In Figure 8a, we show the determination coefficient (R^2) of the LAI taken from the fisheye in a mobile phone against the NDVI extracted from images from PlanetScope (nanosatellite) with an $R^2 = 0.737$ (RSE: 0.05), the LAI against the NDVI extracted from images taken with a multispectral camera (at an altitude of 5 m) averaged over each field with an $R^2 = 0.617$ (RSE: 0.05), and the LAI against the GA index extracted from the photo taken from an RGB camera (at an altitude of 5 m) with an $R^2 = 0.684$ (RSE: 0.15). In Figure 8b, we can see the relationship between the GA (as we mentioned before) against the NDVI extracted from PlanetScope presents an $R^2 = 0.936$ (RSE: 0.02) and the GA against the NDVI extracted from the multispectral camera presents an $R^2 = 0.708$ (RSE: 0.05). Finally, the results presented in Figure 8c show the direct relation between the NDVI from the PlanetScope nanosatellite and NDVI values from the multispectral

camera with an $R^2 = 0.713$ (RSE: 0.05). R^2 was higher than 60% between each category, despite differences in the observation level and image capture or averaging technique. Regarding the NDVI extracted from Sentinel 2 a+b, we could not compare these data easily to any other sources, as the day that we captured these field data was 28 May 2019, and there were no data from Sentinel 2 a+b on this specific day, whereas the closest possible (i.e., unclouded) day was 7 May 2019 and afterwards in July. Moreover, the GA that was taken from the UAV was not comparable either, as the UAV flights in Kenya were prohibited at the time of this field campaign.

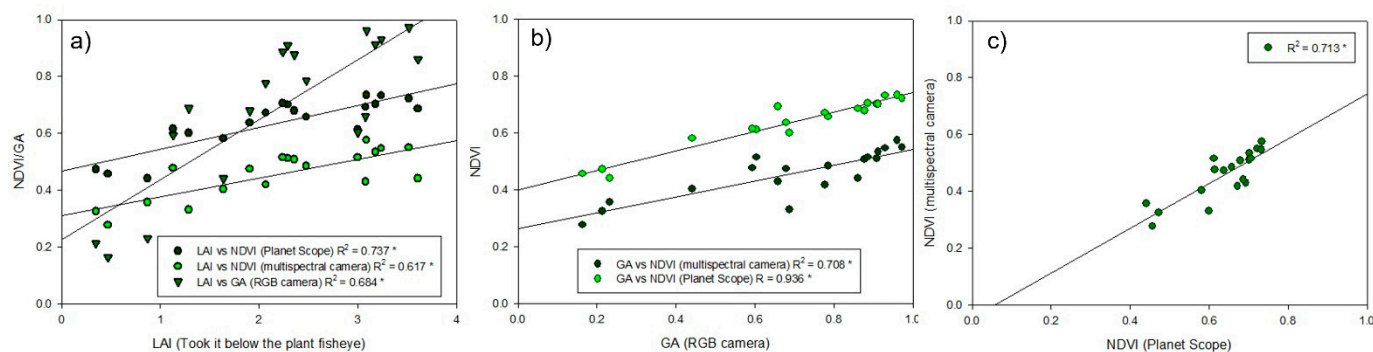


Figure 8. Coefficient of determination (R^2) of the different indices of aboveground biomass against each other, all taken on the same day (28 May 2019). (a) In the x axis, the LAI is taken with the fisheye (from below the plant) and the y axis shows the NDVI calculated from PlanetScope (nanosatellite), the NDVI calculated from the multispectral sensor (at 5 m), and the GA index extracted from the RGB camera (at 5 m). (b) The GA index from the camera compared against the two different NDVI sources: the multispectral sensor and the nanosatellite. (c) The comparison between the NDVI from the nanosatellite and the NDVI from the field multispectral camera. (* = Pearson correlation significant with $p < 0.05$).

3.4. Vegetation Growth Curves Based on Manually Processed Sentinel 2a+b and PlanetScope Image Data

Figure 9 shows the results of four different fields visited on 28 May 2019 in Kenya (first season of maize), showing the NDVI vegetation growth curve (continuous green line) and the first derivative of the NDVI curve first derivative (dashed light green line) along the season. With respect to the curves from Sentinel 2 shown in Figure 9, we can see that the curves do not present anomalies on the values of NDVI in the middle of the curve. These increase as the plants grow until the maize starts to dry (the color is not green anymore). The same happens for the first derivative of the NDVI curve.

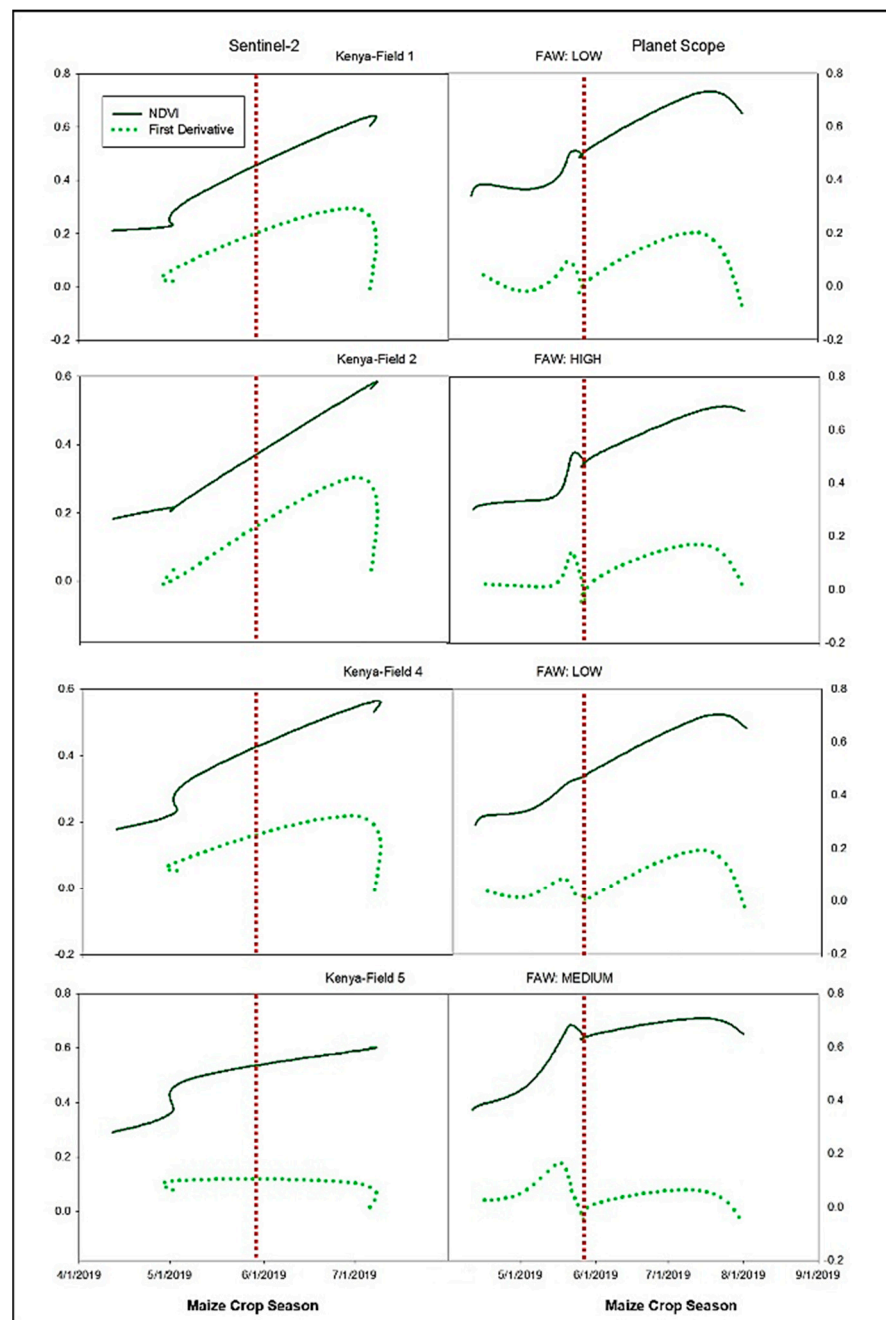


Figure 9. NDVI curves of vegetation from Kenya (continue green line) and the first derivative of NDVI curves (dashed light green line) were taken from Sentinel 2 a+b (**left**) and PlanetScope (**right**) along the first season of maize from the end of April to the end of July. The field presented three levels of infestations: LOW, MEDIUM, and HIGH. The vertical red dotted line is the day that we were in the field taking the data (28 May 2019). The x-axis (**left**) values between 0 and 0.8 belong to the NDVI index and x-axis (**right**) values between -0.2 and 0.8 belong to the NDVI first derivative. The Y axis belongs to the maize crop season from the sowing day to the harvest day.

4. Discussion

4.1. Sentinel 2 a+b Time Series Analyses

Improved data filtering can exclude other possible causes of sudden increases in NDVI/LAI values of maize crops. One study assessed the spatial distribution of the differences in the time lag between a dry year with negative precipitation and vegetation NDVI anomalies and a wet year with positive precipitation and vegetation NDVI anomalies

in a drought-prone region in China [23]. In addition, in one study in a maize field in northeastern Indiana, by comparing NDVI values for individual image dates with the derived normal curve, the response of crop growth to environmental factors was quantified as NDVI residuals, and regression analysis revealed a significant relationship between the yield and NDVI residual during the pre-silking period, indicating that NDVI residuals reflect crop stress in the early growing period that impacts yield [50]. Furthermore, in a study in Zimbabwe, the development of the regional NDVI time series was composited for cultivated areas and adjusted temporally, according to the timing of the rainy season and this adjustment standardized the NDVI response vis-à-vis the expected phenological response of maize. As a consequence of that, a national time series index was developed and this national time series provided an effective summary of vegetation response in agricultural areas, allowing for the identification of NDVI green-up during grain filling [33].

All Sentinel 2 images for this period were used in pairs to create cloud cover reduced image composites. These were then compared with the previous composites in a time step of 11 days to calculate the NDVI change derivative. At the time of the first analyses, there was only one Sentinel 2 satellite sensor launched, and the code was prepared such that it would benefit from the launch of the Sentinel 2b sensor and be prepared for the analysis of new FAMEWS data starting in the summer of 2018. Meanwhile, the powerful computing capability of the GEE platform using greatly improved efficiency of dense image series analysis was in line with previous studies on crop and forest monitoring using GEE and satellite data [51–53]. A comparison between scalable techniques (Figure 4) (e.g., with nanosatellites, aircraft, and UAVs), all of which could provide higher spatial resolution coverage and higher frequency, which indicates better temporal coverage, is proposed for the supplemental or replacement acquisition of FAW monitoring data, as these techniques were not dependent on cloud cover [54–56]. These analyses which compare sensors at different scales could allow for temporal gap filling with real data and could improve the potential use of NDVI derivative anomaly detection, possibly related to the FAW. Furthermore, these techniques could help to confirm the qualitative observation of better performance in lowland areas where geographical variability is minimized.

4.2. Comparison of the NDVI, GA Index, and LAI at Different Spatial Resolutions

Once we had the preliminary data from GEE and some curves of the NDVI showed promise for FAW monitoring, we visited three different countries in Africa: Zimbabwe, Tanzania, and Kenya. In this part, we visited the small farmers' fields and recorded the geographic coordinates of each one, because afterwards we needed to extract the related satellite information. At the same time, we used different sensors to measure the biomass in the maize fields. A close but not perfect relationship between LAI values and the spectral index NDVI is generally expected when compared for the same crop and across similar growing conditions. The slightly lower values of correlation with the field multispectral sensor were most apparent due to a lack of full-field coverage when deployed with a telescoping pole (FOV approx. 5m), but with limited coverage compared to data from a UAV, which provided more precise full coverage. We carried out this comparison to demonstrate the relationships between different technologies and sensors for validation because satellite data were often not available for our monitoring purposes due to cloud cover issues, especially when manually processing the satellite image data. The initial proposal was to use UAVs combined with satellites for data gap filling; however, in the countries in Africa that were included in this project, UAVs were either prohibited or highly limited, as we mentioned before. Moreover, gap-filling time series interpolation techniques would not work for using a first-derivative NDVI approach, but the scalability of the proof of concept for LAI assessments also needed some further investigation, even if the relationship under normal conditions is fairly well understood.

The parameters LAI, NDVI, and GA can be considered as similar total green biomass assessments, when measured in the vegetative growth stage for the same crop. These indices should be comparable enough to enable gap filling with real data in the case

of cloud cover during the FAW critical monitoring periods. There are many studies that compared the LAI using the LI-COR LAI-2000 meter (LI-COR, Inc., Lincoln, NE, USA), an approach similar to our fisheye technique, against NDVIs calculated from different satellites. Kovacs [57] showed that the NDVI calculated from a high-resolution IKONOS satellite (1x1 m the pixel) against the LAI taken from below the plant presents an $R^2 = 0.70$. Moreover, a multiscale comparison between the LAI taken with a LI-COR LAI-2000 against the NDVI (Sentinel-2) showed an $R^2 = 0.77$, and the same LAI against the NDVI from PlanetScope presented a correlation of 0.80. In addition, LAI compared to NDVI taken from the multispectral camera taken with a UAV showed an $R^2 = 0.67$ [58]. Moreover, there are many studies in precision viticulture which compare NDVI at different scales evaluated by Matese [56], and they presented a correlation between the nanosatellite and the UAV ($R^2 = 0.70$).

More advanced LAI assessment techniques with better calibration between the different approaches may provide improvements on the results presented above, even though we implemented a calibrated and standardized classic scientific approach using hemispherical lens photography. The LAI can be measured even by a smartphone using a variety of different apps and techniques, at specific angles (LAI above vs LAI below), or with hemispherical lens photography, as indicated here [20,48]. In fact, all of our LAI data were collected using the Android mobile app that we developed using ODK Collect for metadata recording, image capture, and geolocation, as well as data backup, all of which could allow for very low-cost citizen science-based data filling of LAI datasets for improving FAW monitoring in theory, incorporated as an optional module within the existing FAMEWS app, and potentially validated using the ESA Sentinel 2 SNAP Toolbox. This proof-of-concept scaling extension went beyond the scope of this project but may be pursued in future collaborative work.

4.3. Vegetation Growth Curves Based on Manually Processed Sentinel 2a+b and PlanetScope Image Data

This section focuses on the fields that we visited in the different target countries; we also estimated the FAW damage for the field as a categorical variable. We calculated the NDVI time series curves and first-derivative times-series curves using the data from Sentinel 2 and PlanetScope from the day that the maize was sowed until the reproductive stage was finished. In the case of negative values in the first derivative of the NDVI, something caused a marked and sudden reduction in the total green biomass (NDVI), whether by the FAW or some other biotic or abiotic factor.

The PlanetScope curves in Figure 9 presented possible FAW-related anomalies in the values of NDVI, where the first derivative of NDVI values is negative. After these values start to increase again, indicating that the crop recovered its growth again, and that something caused a marked and sudden reduction in the total green biomass (NDVI), probably the timing and impact of a FAW pest invasion, notably because field measurements and surveys were conducted during those same days. In fact, field data were taken on 28 May 2019, indicated by the vertical red line in Figure 9.

Supplemental Figures S1–S4 show the NDVI vegetation curve and NDVI first-derivative curve along the season of maize. Figure S1a–c shows values from the fields measured in Zimbabwe during the maize season. Figures S2–S4 show the fields measured in Kenya during the maize season.

They presented negative values on the NDVI first derivative calculated with PlanetScope data but could not confirm this pattern caused by the FAW, abiotic stress, or biotic stress. One of the preventative methods used by smallholder farmers was intercropping, using ground nuts or beans because they repelled and helped prevent FAW infestation [11,13]. On the other hand, repeating the same analysis for the 2020 crop season to test these concepts further for their potential for FAW monitoring was not possible due to the Locust (*Schistocerca gregaria* Forskål) invasion that year, which similarly attacked the plant through

herbivory, often massively reducing the crop aboveground biomass and thus reducing both NDVI and LAI values [59–61].

5. Conclusions and Future

The first derivative of the Sentinel-2 a+b and PlanetScope NDVI time series showed the most promising results but only in select locations and is yet to be fully implemented with adequate spatial and false-positive anomaly filters. Again, the analyses presented here are best limited to vegetative growth stages with mostly maize as the main crop, though possible feasible with intercropping. Intercropping may be considered a source of variability, which will always be present for any type of image analysis. Cloud cover also reduced some of the useable data and is a major limitation using passive remote sensing approaches. One of the possible suggestions for future continuations of this work could be integration at different scales of remote sensing. For example, companies or governments could have regional control of the small farmer's fields through monitoring with Sentinel 2 a+b and/or commercial nanosatellites with more detailed resolutions and shorter return intervals (depends on the type of satellite) to calculate the NDVI vegetation and derivative curves. Then, if the analysis showed that the NDVI was down in the middle of the season, they could travel to the field and double check using a mobile app, such as Nuru from PlantVillage [62,63] (see <https://plantvillage.psu.edu/projects> for more details, last accessed 25 June 2022) to recognize if the cause is the FAW or not, and verify that the dip in the NDVI growth curves was not caused by intercropping, weeds, or abiotic and other biotic factors (Figure 10).

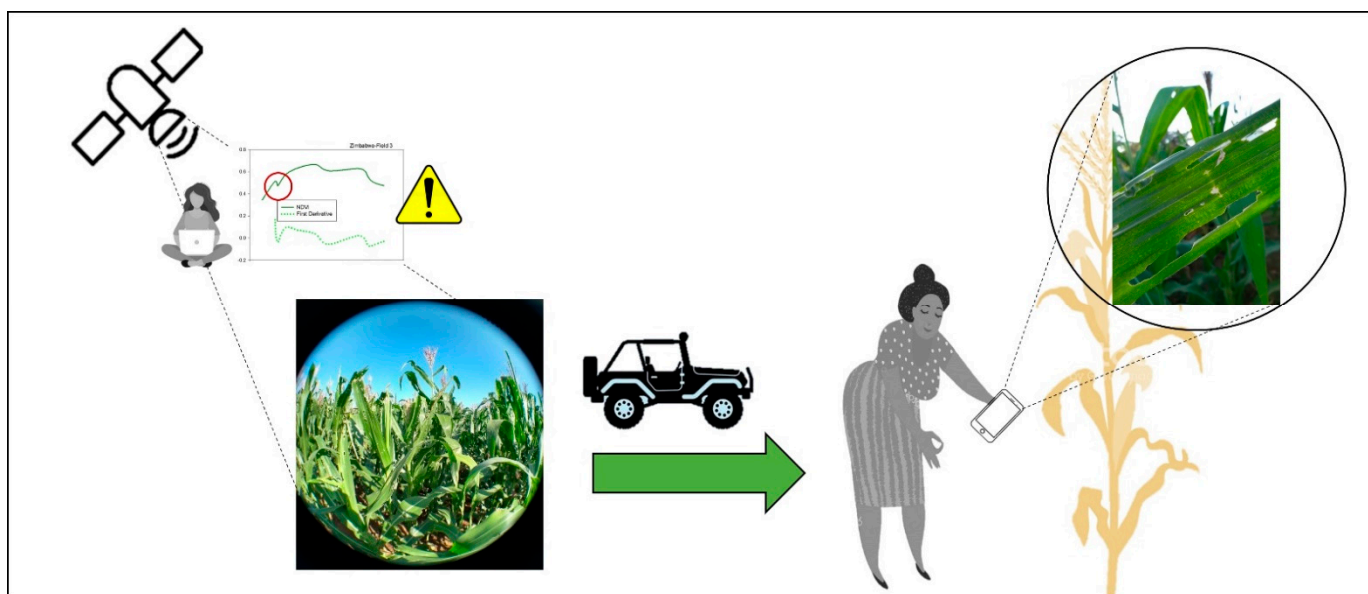


Figure 10. One tentative “solution” could be integrating at different scales of remote sensing.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs14195003/s1>. Table S1. Days of the images from Sentinel 2 a+b and PlanetScope downloaded from USGS. Table S2. Relevant data taken from each field that we visited, using the FAMEWS app. Figure S1. (a–c) NDVI curves of vegetation from Zimbabwe (continue green line) and the first derivative of the NDVI curves (dashed light green line) were taken from Sentinel 2 a+b (left) and PlanetScope (right) along the first season of maize from January to the end of March and April. Figure S2. (a,b) NDVI curves of vegetation from Kenya (continue green line) and the first derivative of NDVI curves (dashed light green line) were taken from Sentinel 2 a+b (left) and PlanetScope (right) along the first season of maize from the end of April to the end of July. Figure S3. NDVI curves of vegetation from Kenya (continue green line) and the first derivative of NDVI curves (dashed light green line) were taken from Sentinel 2 a+b (left) and PlanetScope (right)

along the first season of maize from the end of April to the end of July. Figure S4. NDVI curves of vegetation from Kenya (continue green line) and the first derivative of NDVI curves (dashed light green line) were taken from Sentinel 2 a+b (left) and PlanetScope (right) along the first season of maize from the end of April to the end of July.

Author Contributions: Conceptualization, M.L.B. and S.C.K.; methodology, M.L.B., S.C.K., J.C. (Jill Cairns), E.H., K.W., A.K., D.H., P.M. and J.C. (John Chelal); software N.C., S.C.K. and M.L.B.; validation, M.L.B. and S.C.K.; formal analysis, M.L.B. and S.C.K.; investigation, M.L.B. and S.C.K.; resources, M.L.B.; data curation, M.L.B. and S.C.K.; writing—original draft preparation, M.L.B. and S.C.K.; writing—review and editing, M.L.B., S.C.K. and J.L.A.; visualization, M.L.B. and S.C.K.; supervision, M.L.B. and S.C.K.; project administration, M.L.B. and S.C.K.; funding acquisition, S.C.K. All authors have read and agreed to the published version of the manuscript.

Funding: Funded by the Food and Agriculture of United Nations (FAO). S.C.K. is supported by the Ramon y Cajal RYC-2019-027818-I research fellowship from the Ministerio de Ciencia e Innovación, Spain. This research was also supported by the COST Action CA17134 SENSECO (Optical synergies for spatiotemporal sensing of scalable ecophysiological traits) funded by COST (European Cooperation in Science and Technology, www.cost.eu accessed on 29 April 2022).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Thanks to the CIMMYT from Zimbabwe, University of Moi from Kenya; the Nelson Mandela Institute from Arusha, Tanzania; Pennsylvania University from the U.S; and Lancaster University from the U.K.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

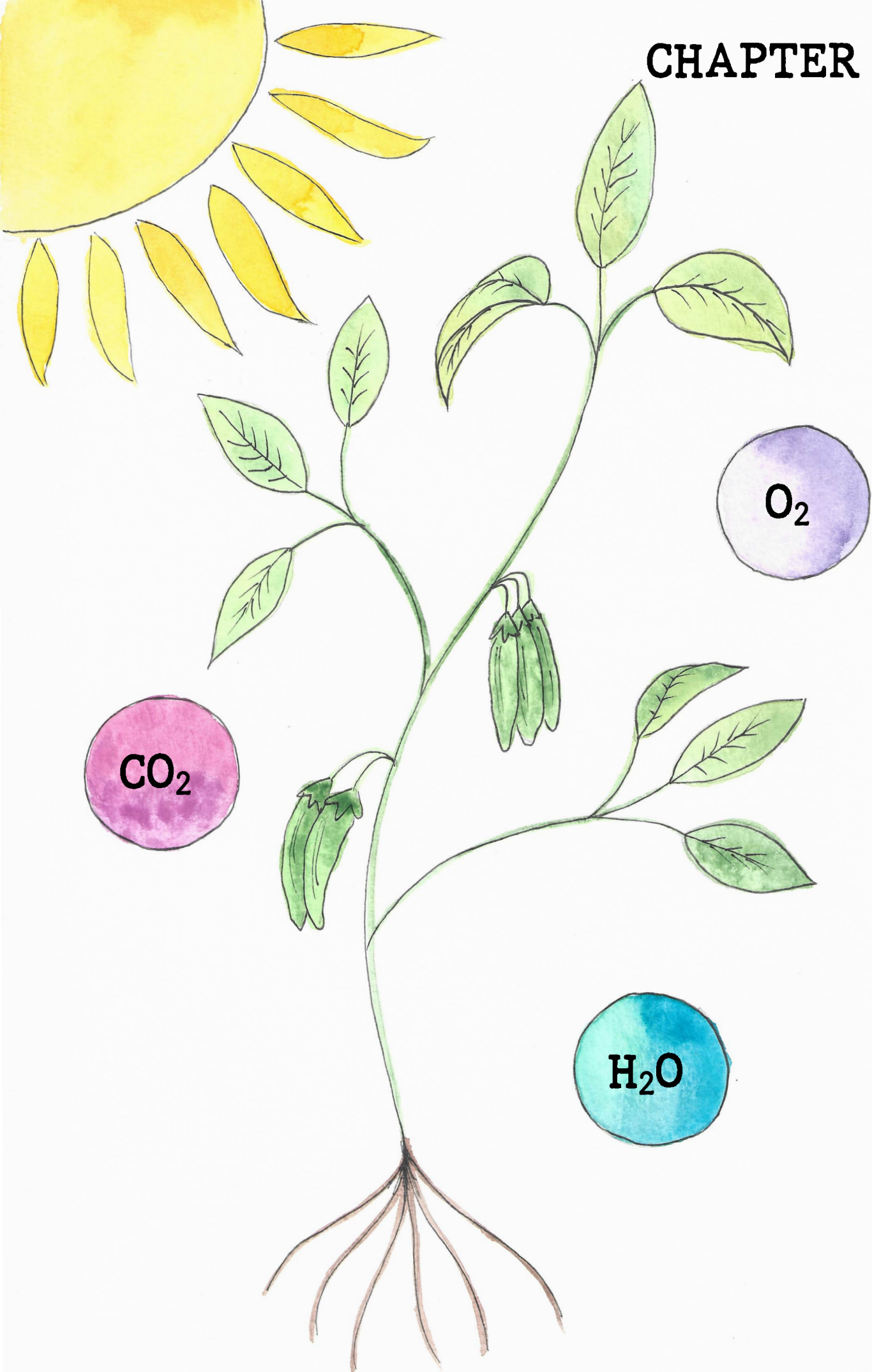
References

1. Brown, M.E. Metrics to accelerate private sector investment in sustainable development goal 2—Zero hunger. *Sustainability* **2021**, *13*, 5967. [[CrossRef](#)]
2. Tanumihardjo, S.A.; McCulley, L.; Roh, R.; Lopez-Ridaura, S.; Palacios-Rojas, N.; Gunaratna, N.S. Maize agro-food systems to ensure food and nutrition security in reference to the Sustainable Development Goals. *Glob. Food Secur.* **2020**, *25*, 100327. [[CrossRef](#)]
3. Prasanna, B.; Huesing, J.E.; Eddy, R.; Peschke, V.M. *Fall Armyworm in Africa: A Guide for Integrated Pest Management*, 1st ed.; CIMMYT: Mexico City, Mexico, 2018; pp. 45–62.
4. FAO. *The State of Food and Agriculture 2016 (SOFA): Climate Change, Agriculture and Food Security*; FAO: Rome, Italy, 2016; ISBN 9789251062159.
5. Hruska, A.J. Fall armyworm (*Spodoptera frugiperda*) management by smallholders. *CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour.* **2019**, *14*, 1–11. [[CrossRef](#)]
6. FAO. *The Global Action for Fall Armyworm Control: Action Framework 2020–2022*; FAO: Rome, Italy, 2020; ISBN 9789251326961.
7. Balla, A.; Bhaskar, M.; Bagade, P.; Rawal, N. Yield losses in maize (*Zea mays*) due to fall armyworm infestation and potential IoT-based interventions for its control. *J. Entomol. Zool. Stud.* **2019**, *7*, 920–927.
8. Williams, W.P.; Davis, F.M.; Buckley, P.M.; Hedin, P.A.; Baker, G.T.; Luthe, D.S. Factors Associated with Resistance to Fall Armyworm (Lepidoptera: Noctuidae) and Southwestern Corn Borer (Lepidoptera: Crambidae) in Corn at Different Vegetative Stages. *J. Econ. Entomol.* **1998**, *91*, 1471–1480. [[CrossRef](#)]
9. Sisay, B.; Simiyu, J.; Mendesil, E.; Likhayo, P.; Ayalew, G.; Mohamed, S.; Subramanian, S.; Tefera, T. Fall armyworm, (*spodoptera frugiperda*) infestations in East Africa: Assessment of damage and parasitism. *Insects* **2019**, *10*, 195. [[CrossRef](#)] [[PubMed](#)]
10. Day, R.; Abrahams, P.; Bateman, M.; Beale, T.; Clotley, V.; Cock, M.; Colmenarez, Y.; Corniani, N.; Early, R.; Godwin, J.; et al. Fall armyworm: Impacts and implications for Africa. *Outlooks Pest Manag.* **2017**, *28*, 196–201. [[CrossRef](#)]
11. Gebrezihher, H.G. Review on management methods of fall armyworm (*Spodoptera frugiperda* J.E. Smith) in Sub-Saharan Africa. *Int. J. Entomol. Res.* **2020**, *5*, 09–14.
12. Hailu, G.; Niassy, S.; Zeyaur, K.R.; Ochatum, N.; Subramanian, S. Maize–legume intercropping and push–pull for management of fall armyworm, stemborers, and striga in Uganda. *Agron. J.* **2018**, *110*, 2513–2522. [[CrossRef](#)]
13. Harrison, R.D.; Thierfelder, C.; Baudron, F.; Chinwada, P.; Midega, C.; Schaffner, U.; van den Berg, J. Agro-ecological options for fall armyworm (*Spodoptera frugiperda* J.E. Smith) management: Providing low-cost, smallholder friendly solutions to an invasive pest. *J. Environ. Manag.* **2019**, *243*, 318–330. [[CrossRef](#)]

14. Abate, T.; van Huis, A.; Ampofo, K.O. Pest Management Strategies in Traditional Agriculture: An African Perspective. *Annu. Rev. Entomol.* **2000**, *45*, 631–659. [[CrossRef](#)] [[PubMed](#)]
15. Wyckhuys, K.A.G.; O’Neil, R.J. Social and ecological facets of pest management in Honduran subsistence agriculture: Implications for IPM extension and natural resource management. *Environ. Dev. Sustain.* **2010**, *12*, 297–311. [[CrossRef](#)]
16. Yigezu, G.; Wakgari, M. Local and indigenous knowledge of farmers management practice against fall armyworm (*Spodoptera frugiperda*) (J.E. Smith) (Lepidoptera: Noctuidae): A review. *J. Entomol. Zool. Stud.* **2020**, *8*, 765–770.
17. Ayra-Pardo, C.; Huang, S.; Kan, Y.; Wright, D.J. Impact of invasive fall armyworm on plant and arthropod communities and implications for crop protection. *Int. J. Pest Manag.* **2021**, 1–12. [[CrossRef](#)]
18. Gross, H.R.; Young, J.R.; Wiseman, B.R. Relative Susceptibility of a Summer-Planted Dent and Tropical Flint Corn Variety to Whorl Stage Damage by the Fall Armyworm (Lepidoptera: Noctuidae). *J. Econ. Entomol.* **1982**, *75*, 1153–1156. [[CrossRef](#)]
19. Jonckheere, I.; Fleck, S.; Nackaerts, K.; Muys, B.; Coppin, P.; Weiss, M.; Baret, F. Review of methods for in situ leaf area index determination Part I. Theories, sensors and hemispherical photography. *Agric. For. Meteorol.* **2004**, *121*, 19–35. [[CrossRef](#)]
20. Van Gardingen, P.R.; Jackson, G.E.; Hernandez-Daumas, S.; Russell, G.; Sharp, L. Leaf area index estimates obtained for clumped canopies using hemispherical photography. *Agric. For. Meteorol.* **1999**, *94*, 243–257. [[CrossRef](#)]
21. Garrigues, S.; Shabanov, N.V.; Swanson, K.; Morisette, J.T.; Baret, F.; Myneni, R.B. Intercomparison and sensitivity analysis of Leaf Area Index retrievals from LAI-2000, AccuPAR, and digital hemispherical photography over croplands. *Agric. For. Meteorol.* **2008**, *148*, 1193–1209. [[CrossRef](#)]
22. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
23. Van Hoek, M.; Jia, L.; Zhou, J.; Zheng, C.; Menenti, M. Early drought detection by spectral analysis of satellite time series of precipitation and Normalized Difference Vegetation Index (NDVI). *Remote Sens.* **2016**, *8*, 422. [[CrossRef](#)]
24. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS. In Proceedings of the Third ERTS Symposium, Washington, DC, USA, 10–14 December 1973; pp. 309–317.
25. Thenkabail, P.S.; Smith, R.B.; De Pauw, E. Evaluation of Narrowband and Broadband Vegetation Indices for Determining Optimal Hyperspectral Wavebands for Agricultural Crop Characterization. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 607–621.
26. Casadesús, J.; Villegas, D. Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. *J. Integr. Plant Biol.* **2014**, *56*, 7–14. [[CrossRef](#)] [[PubMed](#)]
27. Segarra, J.; Buchailot, M.L.; Araus, J.L.; Kefauver, S.C. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* **2020**, *10*, 641. [[CrossRef](#)]
28. Cheng, Y.; Vrieling, A.; Fava, F.; Meroni, M.; Marshall, M.; Gachoki, S. Phenology of short vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sens. Environ.* **2020**, *248*, 112004. [[CrossRef](#)]
29. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510. [[CrossRef](#)]
30. White, M.A.; De Beurs, K.M.; Didan, K.; Inouye, D.W.; Richardson, A.D.; Jensen, O.P.; O’keefe, J.O.; Zhang, G.; Nemani, R.R.; Leeuwen, W.J.D.V.; et al. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Glob. Chang. Biol.* **2009**, *15*, 2335–2359. [[CrossRef](#)]
31. Tottrup, C.; Rasmussen, M.S. Mapping long-term changes in savannah crop productivity in Senegal through trend analysis of time series of remote sensing data. *Agric. Ecosyst. Environ.* **2004**, *103*, 545–560. [[CrossRef](#)]
32. Qader, S.H.; Dash, J.; Alegana, V.A.; Khwarahm, N.R.; Tatem, A.J.; Atkinson, P.M. The role of earth observation in achieving sustainable agricultural production in arid and semi-arid regions of the world. *Remote Sens.* **2021**, *13*, 3382. [[CrossRef](#)]
33. Hüttich, C.; Herold, M.; Schmullius, C.; Egorov, V.; Bartalev, S.A. Indicators of Northern Eurasia’s land-cover change trends from SPOT-VEGETATION time-series analysis 1998–2005. *Int. J. Remote Sens.* **2007**, *28*, 4199–4206. [[CrossRef](#)]
34. Funk, C.; Budde, M.E. Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sens. Environ.* **2009**, *113*, 115–125. [[CrossRef](#)]
35. Eerens, H.; Haesen, D.; Rembold, F.; Urbano, F.; Tote, C.; Bydekerke, L. Image time series processing for agriculture monitoring. *Environ. Model. Softw.* **2014**, *53*, 154–162. [[CrossRef](#)]
36. Reinermann, S.; Gessner, U.; Asam, S.; Kuenzer, C.; Dech, S. The effect of droughts on vegetation condition in Germany: An analysis based on two decades of satellite earth observation time series and crop yield statistics. *Remote Sens.* **2019**, *11*, 1783. [[CrossRef](#)]
37. Rembold, F.; Atzberger, C.; Savin, I.; Rojas, O. Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote Sens.* **2013**, *5*, 1704–1733. [[CrossRef](#)]
38. Serra, P.; Pons, X. Monitoring farmers’ decisions on Mediterranean irrigated crops using satellite image time series. *Int. J. Remote Sens.* **2008**, *29*, 2293–2316. [[CrossRef](#)]
39. Meroni, M.; Atzberger, C.; Vancutsem, C.; Gobron, N.; Baret, F.; Lacaze, R.; Eerens, H.; Leo, O. Evaluation of agreement between space remote sensing SPOT-VEGETATION fAPAR Time Series. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1951–1962. [[CrossRef](#)]
40. Winkler, K.; Gessner, U.; Hochschild, V. Identifying droughts affecting agriculture in Africa based on remote sensing time series between 2000–2016: Rainfall anomalies and vegetation condition in the context of ENSO. *Remote Sens.* **2017**, *9*, 831. [[CrossRef](#)]

41. Rembold, F.; Meroni, M.; Urbano, F.; Royer, A.; Atzberger, C.; Lemoine, G.; Eerens, H.; Haesen, D. Remote sensing time series analysis for crop monitoring with the SPIRITS software: New functionalities and use examples. *Front. Environ. Sci.* **2015**, *3*, 1–11. [[CrossRef](#)]
42. Ramos, A.P.M.; Gomes, F.D.G.; Pinheiro, M.M.F.; Furuya, D.E.G.; Gonçalves, W.N.; Marcato, J., Jr.; Michereff, M.F.F.; Blassioli-Moraes, M.C.; Borges, M.; Alaumann, R.A.; et al. Detecting the attack of the fall armyworm (*Spodoptera frugiperda*) in cotton plants with machine learning and spectral measurements. *Precis. Agric.* **2021**, *23*, 470–491. [[CrossRef](#)]
43. Brown, M.E.; Mugo, S.; Petersen, S.; Klausner, D. Designing a Pest and Disease Outbreak Warning System for Farmers, Agronomists and Agricultural Input Distributors in East Africa. *Insects* **2022**, *13*, 232. [[CrossRef](#)]
44. Pisman, T.I.; Erunova, M.G.; Botvich, I.Y.; Shevyrnogov, A.P. Spatial Distribution of NDVI Seeds of Cereal Crops with Different Levels of Weediness According to PlanetScope Satellite Data. *J. Sib. Fed. Univ. Eng. Technol.* **2020**, *13*, 578–585. [[CrossRef](#)]
45. Moyroud, N.; Portet, F. Introduction to QGIS. In *QGIS Generic Tools*; Wiley: Hoboken, NJ, USA, 2018; pp. 1–17.
46. Casadesus, J.; Kaya, Y.; Bort, J.; Nachit, M.M.; Araus, J.L.; Amor, S.; Ferrazzano, G.; Maalouf, F.; Maccaferri, M.; Martos, V.; et al. Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. *Ann. Appl. Biol.* **2007**, *150*, 227–236. [[CrossRef](#)]
47. Buchailot, M.L.; Gracia-Romero, A.; Vergara-Diaz, O.; Zaman-Allah, M.A.; Tarekgegne, A.; Cairns, J.E.; Prasanna, B.M.; Araus, J.L.; Kefauver, S.C. Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. *Sensors* **2019**, *19*, 1815. [[CrossRef](#)] [[PubMed](#)]
48. Cao, S.; Danielson, B.; Clare, S.; Koenig, S.; Campos-Vargas, C.; Sanchez-Azofeifa, A. Radiometric calibration assessments for UAS-borne multispectral cameras: Laboratory and field protocols. *Int. Soc. J. Photogramm. Remote Sens.* **2019**, *149*, 132–145. [[CrossRef](#)]
49. Demarez, V.; Duthoit, S.; Baret, F.; Weiss, M. Dedieu Estimation of leaf area and clumping indexes of crops with hemispherical photographs. *Agric. For. Meteorol.* **2008**, *148*, 644–655. [[CrossRef](#)]
50. Baudron, F.; Zaman-Allah, M.A.; Chaipa, I.; Chari, N.; Chinwada, P. Understanding the factors influencing fall armyworm (*Spodoptera frugiperda* J.E. Smith) damage in African smallholder maize fields and quantifying its impact on yield. A case study in Eastern Zimbabwe. *Crop Prot.* **2019**, *120*, 141–150. [[CrossRef](#)]
51. Wang, R.; Cherkauer, K.; Bowling, L. Corn response to climate stress detected with satellite-based NDVI time series. *Remote Sens.* **2016**, *8*, 269. [[CrossRef](#)]
52. Li, H.; Jia, M.; Zhang, R.; Ren, Y.; Wen, X. Incorporating the plant phenological trajectory into mangrove species mapping with dense time series Sentinel-2 imagery and the Google Earth Engine platform. *Remote Sens.* **2019**, *11*, 2479. [[CrossRef](#)]
53. Mahdianpari, M.; Jafarzadeh, H.; Granger, J.E.; Mohammadimanesh, F.; Brisco, B.; Salehi, B.; Homayouni, S.; Weng, Q. A large-scale change monitoring of wetlands using time series Landsat imagery on Google Earth Engine: A case study in Newfoundland. *GISci. Remote Sens.* **2020**, *57*, 1102–1124. [[CrossRef](#)]
54. Decuyper, M.; Chávez, R.O.; Lohbeck, M.; Lastra, J.A.; Tsendbazar, N.; Hackländer, J.; Herold, M.; Vågen, T.G. Continuous monitoring of forest change dynamics with satellite time series. *Remote Sens. Environ.* **2022**, *269*, 112829. [[CrossRef](#)]
55. Nonni, F.; Malacarne, D.; Pappalardo, S.E. Sentinel-2 Data Analysis and Comparison with UAV Multispectral Images for Precision Viticulture Study areas. *GI Forum* **2018**, *1*, 105–116.
56. Berra, E.F.; Gaulton, R.; Barr, S. Assessing spring phenology of a temperate woodland: A multiscale comparison of ground, unmanned aerial vehicle and Landsat satellite observations. *Remote Sens. Environ.* **2019**, *223*, 229–242. [[CrossRef](#)]
57. Matese, A.; Toscano, P.; Di Gennaro, F.S.; Genesio, L.; Vaccari, F.P.; Primicerio, J.; Belli, C.; Zaldei, A.; Bianconi, R.; Gioli, B. Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. *Remote Sens.* **2015**, *7*, 2971–2990. [[CrossRef](#)]
58. Kovacs, J.M.; Flores-Verdugo, F.; Wang, J.; Aspden, L.P. Estimating leaf area index of a degraded mangrove forest using high spatial resolution satellite data. *Aquat. Bot.* **2004**, *80*, 13–22. [[CrossRef](#)]
59. Ihuoma, S.O.; Madramootoo, C.A.; Kalacska, M. Integration of satellite imagery and in situ soil moisture data for estimating irrigation water requirements. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102396. [[CrossRef](#)]
60. Salih, A.A.M.; Baraibar, M.; Mwangi, K.K.; Artan, G. Climate change and locust outbreak in East Africa. *Nat. Clim. Chang.* **2020**, *10*, 584–585. [[CrossRef](#)]
61. Meynard, C.N.; Lecoq, M.; Chapuis, M.P.; Piou, C. On the relative role of climate change and management in the current desert locust outbreak in East Africa. *Glob. Chang. Biol.* **2020**, *26*, 3753–3755. [[CrossRef](#)] [[PubMed](#)]
62. Odhiambo, K.; Lewis, J.; Tefera, N.; Thomas, A.; Meroni, M.; Rembold, M. *Impacts of COVID-19 and Desert Locusts on Smallholder Farmers Food Systems and Value Chains in Kenya*; Publications Office of the EU: Luxembourg, 2021; ISBN 978-92-76-32073-9.
63. Kehs, A.; McCloskey, P.; Chelal, J.; Morr, D.; Amakove, S.; Mayieka, J.; Ntango, G.; Nyongesa, K.; Pamba, L.; Mugo, J.; et al. From village to globe: A dynamic real-time map of African fields through PlantVillage. *Front. Sustain. Food Syst.* **2019**, *5*, 514785. [[CrossRef](#)]

CHAPTER 3

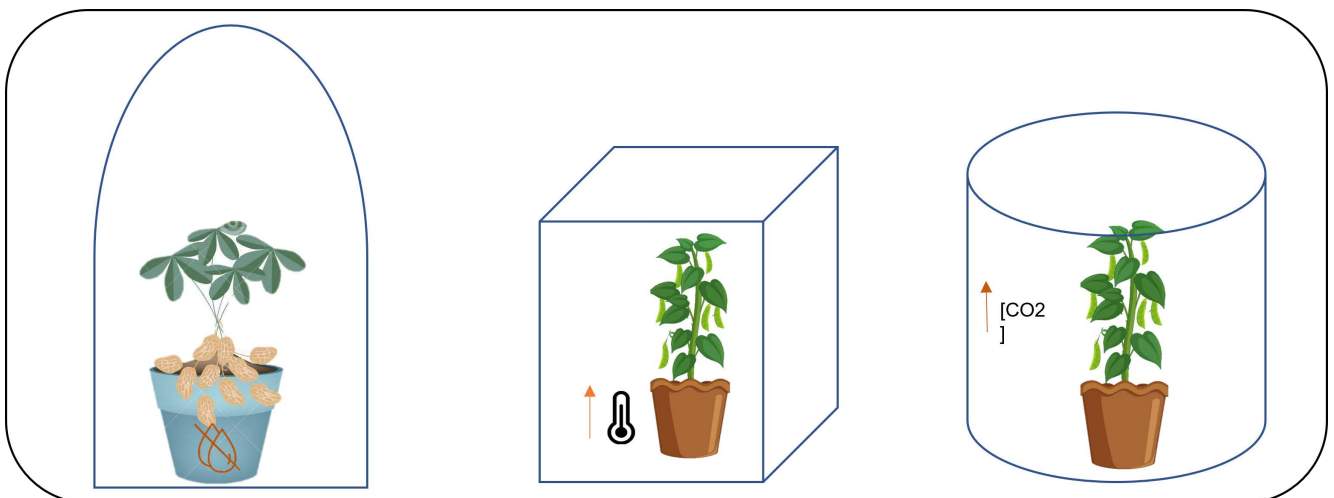


CHAPTER 3

Estimating peanut and soybean photosynthetic traits using leaf spectral reflectance and advance regression models

Ma. Luisa Buchailot, David Soba, Tianchu Shu, Juan Liu, Iker Aranjuelo, José Luis Araus, G. Brett Runion, Stephen A. Prior, Shawn C. Kefauver, Alvaro Sanz-Saez

Published in:
Planta (2022)





Estimating peanut and soybean photosynthetic traits using leaf spectral reflectance and advance regression models

Ma. Luisa Buchailot^{1,2} · David Soba³ · Tianchu Shu⁴ · Juan Liu⁵ · Iker Aranjuelo³ · José Luis Araus^{1,2} · G. Brett Runion⁶ · Stephen A. Prior⁶ · Shawn C. Kefauver^{1,2} · Alvaro Sanz-Saez⁴

Received: 16 November 2021 / Accepted: 3 March 2022
© The Author(s) 2022

Abstract

Main conclusion By combining hyperspectral signatures of peanut and soybean, we predicted $V_{c,max}$ and J_{max} with 70 and 50% accuracy. The PLS was the model that better predicted these photosynthetic parameters.

Abstract One proposed key strategy for increasing potential crop stability and yield centers on exploitation of genotypic variability in photosynthetic capacity through precise high-throughput phenotyping techniques. Photosynthetic parameters, such as the maximum rate of Rubisco catalyzed carboxylation ($V_{c,max}$) and maximum electron transport rate supporting RuBP regeneration (J_{max}), have been identified as key targets for improvement. The primary techniques for measuring these physiological parameters are very time-consuming. However, these parameters could be estimated using rapid and non-destructive leaf spectroscopy techniques. This study compared four different advanced regression models (PLS, BR, ARDR, and LASSO) to estimate $V_{c,max}$ and J_{max} based on leaf reflectance spectra measured with an ASD FieldSpec4. Two leguminous species were tested under different controlled environmental conditions: (1) peanut under different water regimes at normal atmospheric conditions and (2) soybean under high [CO₂] and high night temperature. Model sensitivities were assessed for each crop and treatment separately and in combination to identify strengths and weaknesses of each modeling approach. Regardless of regression model, robust predictions were achieved for $V_{c,max}$ ($R^2=0.70$) and J_{max} ($R^2=0.50$). Field spectroscopy shows promising results for estimating spatial and temporal variations in photosynthetic capacity based on leaf and canopy spectral properties.

Keywords Advanced regression models · ARDR · Bayesian ridge model · High-throughput phenotyping · J_{max} · Lasso · Leaf reflectance · Peanut · Photosynthesis · PLS · Soybean · $V_{c,max}$

Abbreviations

ARDR	Automatic relevance determination regression model
BR	Bayesian ridge model

Communicated by Dorothea Bartels.

✉ Shawn C. Kefauver
sckefauver@ub.edu

✉ Alvaro Sanz-Saez
azs0223@auburn.edu

¹ Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain

² AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Spain

³ Instituto de Agrobiotecnología (IdAB), Consejo Superior de Investigaciones Científicas (CSIC)-Gobierno de Navarra, Av. Pamplona 123, 31192 Mutilva, Spain

⁴ Department of Crop, Soil, and Environmental Sciences, Auburn University, Alabama, USA

⁵ Industrial Crops Research Institute, Henan Academy of Agricultural Sciences, Henan, China

⁶ U.S. Department of Agriculture–Agricultural Research Service, National Soil Dynamics Laboratory, Auburn, AL 36832, USA

J_{\max}	Maximum electron transport rate supporting RuBP regeneration
LASSO	Least absolute shrinkage and selection operator model
NIR	Near-infrared spectral reflectance
PLSR	Partial least squares regression model
RuBP	Ribulose 1,5-bisphosphate
SWIR	Shortwave infrared spectral reflectance
$V_{c,\max}$	Maximum rate of rubisco-catalyzed carboxylation
VIS	Visible spectral reflectance

Introduction

One of the great challenges for the future is the production of sufficient food for a growing population. From 1961 to 2012, the human population more than doubled from approximately 3 billion to 7 billion people and a further increase to 9.3 billion is projected for the year 2050 (FAOSTAT 2016). This means that crop production must double by 2050 to meet the predicted production demands of the global population. However, achieving this goal will be a significant challenge for agriculture since crop yields would have to increase at a rate of 2.4% per year, yet the average rate of increase is only 1.3%, with yields stagnating in up to 40% of land under cereal production (Araus and Cairns 2014). Further, climate change will exacerbate this challenge by intensifying field crop exposure to abiotic stress conditions, including rising temperature, drought, and increased CO₂ concentration [CO₂] (Christensen et al. 2007). This is a major issue because climatic factors since the end of the 1980s have counterbalanced the wheat genetic progress of recent decades in Europe (Oury et al. 2012). Indeed, as observed by Oury et al. (2012) and Gray and Brady (2016), the beneficial effects expected from the increase in atmospheric [CO₂] in the World's crop production during recent decades have been constrained by the effects of temperature increases and extended drought.

Grain legumes are the main source of proteins, minerals, and fibers for animals and humans (Meena et al. 2018). To achieve significant improvements in crop yield, breeding strategies aiming to increase biomass gains and crop productivity need to focus on radiation uptake, photosynthetic efficiency, and harvest index (HI) (Reynolds et al. 2012; Koester et al. 2014). However, to date, breeding for higher photosynthetic efficiency or for tolerance to different environmental stresses has only played a minor role in increasing crop productivity over past decades (Zhu et al. 2010). In a rational sense, plant physiology research should focus on improving photosynthesis due to its central part in plant productivity (Long et al. 2004). Recently, different studies

have advanced how to optimize photosynthetic processes in different crops (Ort et al. 2015; Simkin et al. 2019).

One way to improve crop photosynthesis is to increase our knowledge of genomic control of photosynthesis under different environmental conditions. To achieve this, diverse crop populations representing hundreds of cultivars need to be screened (phenotyped) under different environments to associate traits of interest (i.e., photosynthetic parameters) with specific genomic regions. With the rise of genomic and bioinformatics technologies, phenotyping entire populations for traits of interest is the bottleneck that delays scientific advancement in genomics (Adachi et al. 2011; Yan et al. 2015; de Oliveira Silva et al. 2018; Oakley et al. 2018). Therefore, genomic approaches and breeding solutions need to implement new high-throughput phenotyping techniques that allow rapid measurement of photosynthetic traits for screening cultivars in the shortest amount of time (Araus and Cairns 2014; Araus et al. 2018). By improving techniques for measuring photosynthetic traits, more efficient cultivar selection will likely improve both yield potential and resilience to abiotic stresses.

Photosynthetic performance is frequently measured with an infrared gas analyzer that assesses plant CO₂ assimilation rate. Photosynthetic parameters, such as leaf mid-day photosynthesis and leaf diurnal photosynthesis, can be used to assess in situ plant performance under different abiotic stresses (Sanz-Sáez et al. 2012, 2017). More detailed photosynthetic parameters, such as maximum rate of rubisco-catalyzed carboxylation ($V_{c,\max}$) and maximum electron transport rate supporting RuBP regeneration (J_{\max}), have been identified as selection parameters for tolerance to abiotic stress, such as drought (Aranjuelo et al. 2009, 2013), elevated tropospheric ozone (Yendrek et al. 2017), or for improved performance under elevated atmospheric CO₂ (Ainsworth et al. 2004; Soba et al. 2020). Depending on the parameter to be measured, sampling can take a few minutes each (e.g., mid-day photosynthesis) or 20–60 min per sample for photosynthetic parameters, such as $V_{c,\max}$ and J_{\max} , which are calculated using photosynthesis to intercellular CO₂ curves or A–Ci curves (Farquhar et al. 1980; Long and Bernacchi 2003). In addition, $V_{c,\max}$ and J_{\max} are essential input parameters for the FvCB model (Farquhar et al. 1980) that relates photosynthetic biochemistry responses to known environmental conditions (Von Caemmerer 2013). This model has also been used in earth systems models for predicting ecosystem responses to environmental changes (Rogers 2014).

Reflectance spectra at leaf and canopy levels can facilitate assessment of plant's structure, nutritional status, and certain stress parameters. This includes estimating contents of chlorophyll, xanthophylls, nitrogen, phosphorus, fiber, sucrose (Gamon et al. 1997; Peñuelas and Filella 1998; Petisco et al. 2006; Asner and Martin 2008; Colombo et al.

2008; Ainsworth et al. 2014; Serbin et al. 2014; Dechant et al. 2017; Yendrek et al. 2017), and plant secondary metabolites (Couture et al. 2016; Vergara-Diaz et al. 2020). In addition, leaf level spectral reflectance has been used to predict photosynthetic parameters, such as $V_{c,max}$ and J_{max} in soybean (Ainsworth et al. 2014), wheat (Silva-Perez et al. 2018), maize (Heckmann et al. 2017; Yendrek et al. 2017), and trees (Serbin et al. 2012) as well as dark respiration in wheat (Coast et al. 2019).

Although translating data acquired with a field spectrometer using a leaf clip to scalable imaging approaches using multispectral or hyperspectral cameras in drones or other aerial platforms (frequently limited to the 350–1000 nm spectral range) may be further complicated by the heterogeneous nature of canopies, such techniques could greatly expand the scope of applicability of these measurements. In the above-mentioned research, relationships between photosynthetic parameters and complex data arrays captured by leaf level spectrometers need to be analyzed using complex multivariate statistical models. Partial least squares regression (PLSR) is the most commonly used model (Serbin et al. 2012; Ainsworth et al. 2014; Heckmann et al. 2017; Silva-Pérez et al. 2017; Yendrek et al. 2017). However, Fu et al. (2020) recently reported that other machine learning algorithms such as Least Absolute Shrinkage and Selection Operator (LASSO) can estimate photosynthetic parameters as accurately or better than PLSR, since LASSO is more robust when comparing different environments or plant species (Tibshirani 1996). Therefore, to bypass PLSR performance problems, we propose to explore other powerful machine learning algorithms with appropriate feature extraction capacities, which include LASSO (Vergara-Diaz et al. 2020), Bayesian Ridge (BR; Neal 1996), and Automatic Relevance Determination Regression (ARDR; Tipping 2001).

For these multivariate models, utilized data must represent enough phenotypic variability to support proper model functioning. To achieve sufficient phenotypic variability, several researchers have applied a range of growth conditions, including different levels of abiotic stresses, such as drought (Silva-Perez et al. 2017), elevated tropospheric ozone (Ainsworth et al. 2014; Yendrek et al. 2017), or high temperature (Serbin et al. 2012). Another means for increasing phenotypic variation is by including several related species in the same model. For example, Doughty et al. (2011) used 149 tropical tree species to create a PLSR model to estimate mid-day photosynthesis using canopy hyperspectral imaging; and Serbin et al. (2012) combined hyperspectral data of two tree species to estimate $V_{c,max}$. However, to the best of our knowledge, no published study has combined multiple leguminous row crops species. In our research, we focused on soybean (*Glycine max*) and peanut (*Arachis hypogea*), which are leguminous crops often grown under high abiotic stress levels (drought and elevated temperature)

in the southeastern United States. These legume crops are also important in rotation with corn and cotton.

The aims of this study were (i) to estimate photosynthetic capacity parameters, such as mid-day photosynthesis, leaf chlorophyll content (LCC), $V_{c,max}$, and J_{max} of two legume crops (soybean and peanut) using full-range leaf level reflectance spectra (VIS–NIR–SWIR, 400–2500 nm) with PLSR, BR, ARDR and LASSO models and (ii) to simulate photosynthetic parameter model performance using four common types of sensors with more limited wavelength ranges: VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), and an advanced multispectral sensor imitating the ESA Copernicus Sentinel 2 satellite with 12 spectral bands.

Materials and methods

Trial setup and design

Experiments were conducted in field trials and controlled conditions located at Auburn University (Alabama, USA). The study was carried out with two leguminous crops (soybean and peanut) that were exposed to different growth conditions. The first experiment involved two soybean (*Glycine max*. L.) cultivars grown under ambient and elevated [CO_2] at an Open Top Chamber Facility. The second experiment involved four soybean cultivars grown under high night temperature in growth chambers. The third experiment was performed with 6 peanut (*Arachis hypogea* L.) cultivars grown under well-watered and water-stress conditions in a greenhouse.

Experiment 1: soybean cultivar response to elevated [CO_2]

Two soybean cultivars representing high (PI398223) and low (PI567201A) water use efficiencies (WUE) were chosen for the study based on previous screening by Dhanapal et al. (2015). The two cultivars were planted on 16 May 2019 in 20 L pots filled with commercial growth media (Pro-Mix, Premier Tech, Quebec, Canada) at the Open Top Chamber Facility located at the USDA-ARS National Soil Dynamics Laboratory, Auburn, AL, USA. Open top chambers (OTC) (Rogers et al. 1983), encompassing 7.3 m² of ground surface area, were used to deliver target [CO_2] of ~410 ppm (ambient) or ambient plus 200 ppm (elevated) [CO_2] during light hours using a delivery and monitoring system described elsewhere (Mitchell et al. 1995). There were four replicate chambers of each CO_2 level for a total of eight experimental plots. Each OTC held two pots of each cultivar to have two sub-replicates for each plot. The experiment was conducted as a split plot design with CO_2 level being

the main plot factor and cultivar being the split plot factor. Mid-day photosynthesis and A–Ci curves were performed when plants were at the beginning of pod development (R3, Fehr et al. 1971, 15 July) and at the beginning of seed filling (R5, 26 July) according to growth stages defined by Fehr et al. (1971). Relative chlorophyll content and leaf hyperspectral reflectance measurements were performed concurrently with photosynthetic parameter measurements. More detailed information on experimental design was previously reported by Soba et al. (2020).

Experiment 2: soybean cultivar response to high night temperatures

Four soybean cultivars (PI360846, DS25-1, PI458098, and Agx9) were planted in 3.8 L pots containing a peat-moss: perlite potting mixture (2:1) on 1 May 2019. Plants were grown at the Auburn University Plant Science Research Center greenhouse complex. Temperatures were maintained at 28/20 °C (day/night) until plants reached the first flowering stage (R1). To impose night temperature treatments, plants were then moved to two Conviron CMP 6010 growth chambers (Conviron, Manitoba, Winnipeg, Canada) maintained on a 12 h photoperiod (1200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ PAR) with 50/70% RH (day/night). Control plants were grown at 30/20 °C (day/night) and high night temperature plants were grown at 30/30 °C (day/night). Three replicates per cultivar and chamber were used and the whole experiment was repeated twice. Fourteen days after temperature treatments were imposed, mid-day photosynthesis, A–Ci curves, LCC, and leaf hyperspectral reflectance were performed as explained below.

Experiment 3: peanut cultivar response to drought

Six peanut cultivars (AUG16-28, AU17, 18H19-3738, G06-G, AU8-19, and AU18-21) were planted at the Auburn University Plant Science Research Center greenhouse complex on 21 April 2019. Plants were grown in 20 L pots containing a mixture of sand and sandy-loam field soil (1:1, w/w) collected from EV-Smith Research Center, Shorter, AL, USA. Plants were maintained under well-watered conditions (80% relative soil water content, RSWC) until 60 days old; at this time, the drought experiment was initiated. Weighing pots every 2–3 days initially and every day towards the end of the experiment allowed RSWC to be gravimetrically maintained. Well-watered plants were maintained at 80% RSWC while drought plants were maintained at a 30% RSWC. Four replicates per cultivar and stress treatment were used in this experiment. At 20 and 40 days after drought initiation (i.e., 80- and 100-day-old plants), mid-day photosynthesis, A–Ci curves, LCC, and leaf hyperspectral reflectance measurements were performed as explained below.

Physiological parameter assessments

In this study, mid-day photosynthesis, A–Ci curves, and SPAD measurements were taken from 3 different experiments and coupled with full-range (350–2500 nm), high-resolution (3–8 nm) spectral reflectance measurements taken with a Field Spec Hi-Res four field spectrometer (Analytical Spectral Devices, Boulder, CO, USA) to predict physiological parameters that characterize photosynthetic traits.

Mid-day photosynthesis measurements

Depending on experiment size, mid-day photosynthesis measurements were taken one day before A–Ci curves using two or three LI-6400 (Li-Cor Biosciences, Lincoln, NE, USA) systems. Measurements were performed on fully expanded young leaves corresponding with the third/fourth leaf from the top in soybean, and second/third leaf from the top of the main stem in peanut. Prior to measurements, systems were set to match environmental growth conditions (light intensity and temperature) and maintained at a relative humidity of 60–70%. While photosynthesis measurements were in progress, relative chlorophyll content and spectral reflectance measurements were also performed on the same leaves using a SPAD meter (Minolta SPAD-502, Spectrum Technologies Inc., Plainfield, IL, USA) and the Field Spec Hi-Res 4 field spectrometer, respectively.

A–Ci curves

To calculate maximum rate of rubisco-catalyzed carboxylation ($V_{c,\text{max}}$) and maximum electron transport rate supporting RuBP regeneration (J_{max}), A–Ci curves were performed at different developmental stages in each experiment. In general, the A–Ci curves were the same for peanut and soybean except for different light saturation points: 1750 $\mu\text{mol m}^{-2} \text{s}^{-1}$ PAR for soybean (Ainsworth et al. 2004) and 2000 $\mu\text{mol m}^{-2} \text{s}^{-1}$ PAR for peanuts (Ferreira et al. 2000). Photosynthesis was initially induced at the growth [CO_2] (410 ppm for ambient and 610 ppm for elevated CO_2 treatments), and then [CO_2] was reduced stepwise to the lowest concentration of 50 ppm. Afterwards, [CO_2] was increased stepwise to the highest CO_2 concentration of 1500 ppm. A total of 11 measurements per curve were recorded (Sanz-Sáez et al. 2017). During measurements, block temperature was set at 28 °C (i.e., mean mid-day temperature at Auburn, AL). The equations and spreadsheet developed by Sharkey et al. (2007) were used to calculate $V_{c,\text{max}}$ and J_{max} normalized at 25 °C as it has been demonstrated by (Khan et al. 2021) that different temperatures and the effect on reflectance does not affect prediction of these

normalized parameters. While A–Ci curves were taken, concurrent spectral reflectance measurements were performed on the same leaves.

Relative chlorophyll content

Relative chlorophyll content was taken on the same mid-day photosynthesis leaves using a SPAD-502 chlorophyll meter (Konica Minolta, Tokyo, Japan). Five subsample measurements per leaf were collected and averaged.

Leaf spectral reflectance measurements

Leaf spectral reflectance was measured with a FieldSpec Hi-Res 4 concurrently on the same leaves used for photosynthetic measurements. This device has three sensors with a full spectro-radiometer range of 350–2500 nm, with a resolution of 3 nm in visible (VIS; 350–700) and near-infrared (NIR; 700–1000 nm) and 8 nm in shortwave-infrared (SWIR; 1000–2500 nm). Measurements were taken via a leaf clip coupled to a fiber-optic cable. The FieldSpec has a radiometrically calibrated internal light source, which was standardized for relative reflectance using white reference measurements every 15 min. For each leaf, 6 reflectance measurements were recorded on different regions of a single leaf per pot. We used the FieldSpectra package in R to average the six samples and align the VIS, NIR, SWIR sensors with a spectral splice correction (Serbin et al. 2014; Yendrek et al. 2017).

To accomplish the second research aim, we simulated if a more limited spectral range (corresponding to other remote sensing devices) would be able to estimate photosynthetic parameters with the same accuracy as the full-range spectra achieved with the Field Spec HiRes4. Simple spectral resampling of four different sensors was performed to simulate commercial spectrophotometer sensors, such as the UniSpec-DC VIS/NIR (310–1100 nm; PP Systems, Amesbury, MA, USA), the USB 2000 VIS/NIR (340–1014 nm; Ocean Optics, Dunedin, FL, USA), and the Liga SWIR spectrophotometer (850–1888 nm; STEAG Micro Parts, Dortmund, Germany). We also included a resampling simulation for the bands and bandwidths of the ESA Copernicus Sentinel-2 satellite, with 12 spectral bands (443, 494, 560, 665, 704, 740, 781, 834, 944, 1375, 1612, and 2194 nm) representing VIS, NIR, and SWIR (see more in Drusch et al. 2012; Segarra et al. 2020).

Statistical analysis of measured and estimate values

Statistical analyses were conducted using R Studio (RStudio Team 2020) and Python 3.7 (Python Software Foundation, <https://www.python.org>) via a Jupiter notebook (Wofford et al. 2019). Effects of abiotic stress treatments

and differences between cultivars on studied variables were assessed using analysis of variance (ANOVA) in R Studio. We also analyzed correlations between photosynthetic parameters against each spectrum band by Pearson's correlation using R Studio.

With respect to the different advance regression models, we used the SciPy module (Jones et al. 2001; Varoquaux et al. 2015) in Python 3.7 and the Scikit-Learn library for the estimation of different parameters to estimate determination (R^2) and the root means squared error (RMSE). For cross-validation, we used the “train test split method” where, we split our data into training (60% of the data used to build the model) and testing (40% of the data used to test the model). This method quantifies the prediction error, the RMSE, which measures the average prediction error made by the model in predicting the outcome for an observation. That is, the average difference between the observed known outcome values and the values predicted by the model. Associations between photosynthetic parameters (response variables) and the leaf reflectance spectrum (explanatory) variables were analyzed using four advances models: (i) Partial Least Squares Regression (PLSR) is based on the dimension reduction method (Wold et al. 2001). For this model, we used between 5 and 11 components, choosing the number of components that gave the highest R^2 and the lower RSME; (ii) Least Absolute Shrinkage and Selection Operator (LASSO) is a shrinkage method (Tibshirani 1996); (iii) Bayesian ridge (BR) and (iv) Automatic relevance determination regression (ARDR) are both high-dimensional methods (Neal 1996; Tipping 2001). Figures were prepared using the matplotlib (Hunt 2019) and Seaborn Python (Waskom et al. 2017) modules in Python 3.7.

Results

Effect of abiotic stress and cultivar on photosynthetic parameters

Analyzing the effect of abiotic stress and cultivars can yield valuable insights into phenotypic range of variation within each experiment. In Experiment 1, the two soybean cultivars showed significant effects of $[\text{CO}_2]$ on mid-day photosynthesis and LCC, but not on $V_{c,\text{max}}$ and J_{max} (Table 1 and Fig. S2). We observed treatment effects for mid-day photosynthesis and LCC (Table 1a). In summary, phenotypic variation was noticeable with a range of 17.01–36.22 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for mid-day photosynthesis, 34.55–51.35 for LCC, 182.9–348.4 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for $V_{c,\text{max}}$, 174.7–263.7 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for J_{max} , and 29.4–30.37 °C for leaf temperature. In Experiment 2, four soybean cultivars were grown under high night temperature (30/30 °C day/night) for comparison to controls (30/20 °C day/night).

Table 1 Mean values of mid-day photosynthesis ($\mu\text{mol m}^{-2} \text{s}^{-1}$), leaf chlorophyll content (LCC, arbitrary units), maximum rate of rubisco-catalyzed carboxylation ($V_{c,\text{max}}$, $\mu\text{mol m}^{-2} \text{s}^{-1}$), maximum electron transport rate supporting RuBP regeneration (J_{max} , $\mu\text{mol m}^{-2} \text{s}^{-1}$), and leaf temperature ($^{\circ}\text{C}$) per each treatment. (a) Experiment 1: two varieties of soybean grown at 410 ppm and 610 ppm of $[\text{CO}_2]$; $n=32$. (b) Experiment 2: four soybean varieties grown at low (20°C) and high (30°C) night temperature; $n=48$

(A)						
Genotype	Treatment	Photosynthesis ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	LCC (arbitrary unit)	Vcmax ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Jmax ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Leaf Temperature ($^{\circ}\text{C}$)
Pi398223	410	23.3 \pm 3.6 b	45.5 \pm 3.6 a	249.5 \pm 44.5 ab	212.2 \pm 23.3 a	30.31 \pm 0.388 a
Pi567201	410	25.1 \pm 3.9 b	41.6 \pm 3.3 b	269.8 \pm 50.5 a	227.8 \pm 31.9 a	30.06 \pm 0.708 a
Pi398223	610	30.4 \pm 2.1 a	46.3 \pm 3.5 a	227.2 \pm 22.7 b	211.4 \pm 19.7 a	30.37 \pm 0.706 a
Pi567201	610	31.2 \pm 4.8 a	46.7 \pm 2.2 a	257.4 \pm 40.9 ab	219.2 \pm 16.2 a	29.42 \pm 1.377 a
ANOVA	[CO₂]	0.001***	0.15*	0.242	0.569	0.354
ANOVA	Varieties	0.344	0.13	0.092	0.173	0.06
ANOVA	[CO₂]*Varieties	0.678	0.077	0.733	0.646	0.265
(B)						
Genotype	Treatment	Photosynthesis ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	LCC (arbitrary unit)	Vcmax ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Jmax ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Leaf Temperature ($^{\circ}\text{C}$)
PI360846	Low T	17.7 \pm 5.2 b	45.1 \pm 0.6 ab	72.7 \pm 28.6 a	102.6 \pm 33.7 bc	30.06 \pm 0.059 ab
PI458098	Low T	21.4 \pm 4.8 ab	49.8 \pm 3.7 a	91.7 \pm 37.2 a	128.6 \pm 43.1 ab	29.98 \pm 0.418 ab
DS25-1	Low T	22.2 \pm 1.3 ab	38.4 \pm 3.9 c	82.7 \pm 9.5 a	111.0 \pm 10.2 abc	29.86 \pm 0.450 b
AG48 \times 9	Low T	27.2 \pm 3.9 a	45.9 \pm 1.5 a	107.6 \pm 13.1 a	149.0 \pm 16.6 a	29.98 \pm 0.216 ab
PI360846	High T	16.9 \pm 5.3 b	47.3 \pm 4.7 a	100.6 \pm 29.1 a	136.3 \pm 34.3 ab	30.33 \pm 0.170 a
PI458098	High T	16.9 \pm 0.5 b	50.1 \pm 4.1 a	84.7 \pm 29.7 a	108.0 \pm 14.7 abc	29.91 \pm 0.188 ab
DS25-1	High T	15.6 \pm 3.6 b	39.7 \pm 3.2 bc	68.7 \pm 7.4 a	78.67 \pm 16.3 c	30.08 \pm 0.202 ab
AG48 \times 9	High T	27.5 \pm 5.1 a	45.7 \pm 3.2 a	109.3 \pm 26.0 a	146.6 \pm 25.2 ab	30.01 \pm 0.146 ab
ANOVA	Temperature	44.81	0.522	0.833	0.624	0.312
ANOVA	Varieties	0.010**	0.002**	0.303	0.042*	0.561
ANOVA	Temp*Varieties	0.999	0.999	0.999	0.999	0.999

Levels of significance: x, $P < 0.1$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Cultivar effects with treatment showed a significant effect on mid-day photosynthesis, LCC, and J_{max} (Table 1b). Overall, phenotypic variation was noticeable with a range of 11.52–32.68 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for mid-day photosynthesis, 34.01–53.95 for LCC, 48.01–135.2 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for $V_{c,\text{max}}$, 61.01–165.1 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for J_{max} , and 29.9–30.33 $^{\circ}\text{C}$ for leaf temperature. In Experiment 3, the effect of drought was significant for all measured peanut parameters except for $V_{c,\text{max}}$ and J_{max} (Table 2). Cultivars only showed significant effects for LCC and J_{max} . The interaction effects of drought and cultivars was only slightly significant for $V_{c,\text{max}}$ ($P = 0.094$). Phenotypic variation was perceptible since mid-day photosynthesis ranged from 5.051 to 26.41 $\mu\text{mol m}^{-2} \text{s}^{-1}$, LCC varied from 42.30 to 52.45, $V_{c,\text{max}}$ varied from 64.38 to 171.3 $\mu\text{mol m}^{-2} \text{s}^{-1}$, J_{max} ranged from 79.3 to 206.1 $\mu\text{mol m}^{-2} \text{s}^{-1}$, and 28.6 to 30.5 $^{\circ}\text{C}$ for leaf temperature. When phenotypic variation of all three experiments was considered together, the range for mid-day photosynthesis was 5.051–36.22 $\mu\text{mol m}^{-2} \text{s}^{-1}$, 34.55–53.95 for LCC, 48.01–348.4 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for $V_{c,\text{max}}$, 61.01–263.7 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for J_{max} , and 26.33–31.55 $^{\circ}\text{C}$

for leaf temperature (Fig. S2, shows the Box plot for each experiment).

Relationships between spectral signatures and photosynthetic parameters

Figure 1 presents the sensitivity of leaf reflectance spectrum for different species and abiotic stresses. Under high night temperature, soybean reflectance spectrum shows higher variability than the control with a larger peak at ~ 550 nm and wider reflectance band between ~ 750 –1400, 1550–1800 and 2000–2300 nm (Fig. 1a, b). Elevated CO_2 in soybean tended to reduce variability of the reflectance spectrum between ~ 500 –600 and 750–1400 while maintaining the variability in the reflectance spectrum between 1550–1800 and 2000–2300 nm (Fig. 1c, d). In peanut, drought increased variability at all wavelengths with the exception of the 500–600 nm range (Fig. 1e, f). When comparing reflectance of the two legume species, we noted that peanut added a lot of spectral variation in the range from 750 to 2300 nm,

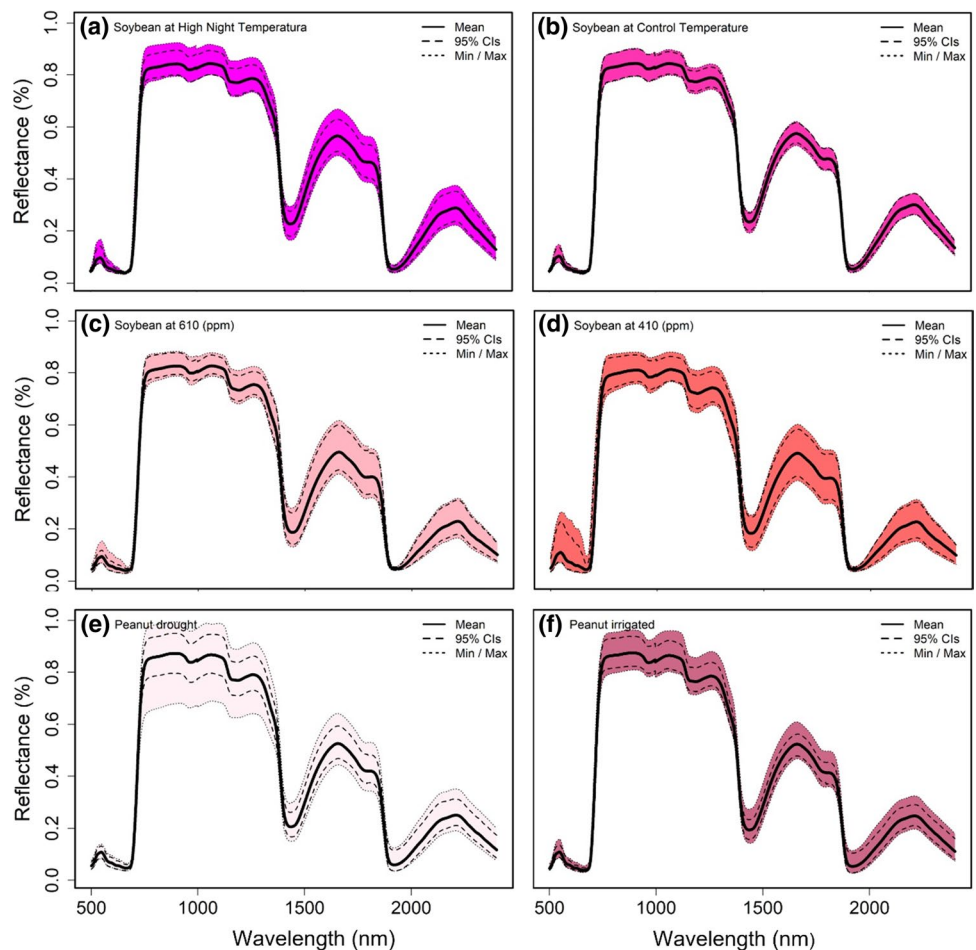
Table 2 Mean values of midday photosynthesis ($\mu\text{mol m}^{-2} \text{s}^{-1}$), leaf chlorophyll content (LCC, arbitrary units), maximum rate of rubisco-catalyzed carboxylation ($V_{c,\text{max}}$, $\mu\text{mol m}^{-2} \text{s}^{-1}$), maximum electron transport rate supporting RuBP regeneration (J_{max} , $\mu\text{mol m}^{-2} \text{s}^{-1}$),

and leaf temperature ($^{\circ}\text{C}$) in six varieties of peanut grown under well-watered (WW, 80% SWC) and water-stress (WS, 30% SWC) conditions

Genotype	Treatment	Photosynthesis ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	LCC (arbitrary unit)	$V_{c,\text{max}}$ ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	J_{max} ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Leaf Temperature ($^{\circ}\text{C}$)
18H19-3738	WW	22.2 ± 2.1 a	50.7 ± 1.4 bcd	126.2 ± 17.1 a	169.2 ± 14.9 ab	28.93 ± 0.906 ab
AU17	WW	21.1 ± 4.8 a	49.6 ± 1.9 cde	135.9 ± 19.9 a	179.9 ± 29.8 a	28.67 ± 0.727 b
AU18-21	WW	21.1 ± 1.9 ab	45.7 ± 0.8 e	129.6 ± 15.1 a	159.5 ± 31.4 abc	29.96 ± 0.662 ab
AU16-28	WW	20.2 ± 4.8 ab	46.6 ± 3.6 de	136.1 ± 20.4 a	183.4 ± 18.9 a	29.01 ± 0.974 ab
AU18-19	WW	17.9 ± 2.5 abc	46.9 ± 2.7 de	108.9 ± 19.2 abc	150.2 ± 22.6 abc	29.58 ± 1.377 ab
G-06-G	WW	17.9 ± 5.0 abc	45.6 ± 2.2 e	123.1 ± 35.2 ab	139.8 ± 45.1 bcd	29.37 ± 1.489 ab
18H19-3738	WS	15.4 ± 3.5 bcd	56.6 ± 1.2 a	125.7 ± 40.4 ab	153.94 ± 31.5 abc	30.40 ± 0.640 ab
AU17	WS	14.1 ± 2.1 cd	53.1 ± 4.6 abc	131.1 ± 22.6 a	152.3 ± 5.8 abc	29.24 ± 0.688 ab
AU18-21	WS	10.7 ± 5.4 d	52.1 ± 4.1 abc	85.4 ± 32.8 c	109.1 ± 22.8 d	29.24 ± 0.688 ab
AU16-28	WS	12.5 ± 2.8 d	54.9 ± 2.7 ab	113.7 ± 14.6 abc	121.3 ± 26.7 cd	30.56 ± 0.707 a
AU18-19	WS	11.9 ± 3.2 d	54.8 ± 4.1 ab	135.4 ± 15.9 a	133.6 ± 37.3 bcd	30.58 ± 0.224 a
G-06-G	WS	10.3 ± 1.8 d	49.3 ± 2.8 cde	89.1 ± 25.8 bc	126.7 ± 15.9 cd	29.47 ± 1.351 ab
ANOVA	Drought	0.001***	0.001***	0.46	0.275	0.02*
ANOVA	Varieties	0.154	0.001***	0.196	0.092	0.837
ANOVA	Drought*Varieties	0.884	0.353	0.094	0.352	0.461

Levels of significance: x, $P < 0.1$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; $n = 48$

Fig. 1 **a** Mean, ± standard deviation ($n = 24$), and minimum and maximum leaf reflectance for soybean at high night temperature grown in growth chambers. **b** Mean, ± standard deviation ($n = 24$), and minimum and maximum leaf reflectance for soybean at control temperature grown in growth chambers. **c** Mean, ± standard deviation ($n = 18$), and minimum and maximum leaf reflectance for soybean at 610 ppm grown at an Open Top Chamber Facility. **d** Mean, ± standard deviation ($n = 18$), and minimum and maximum leaf reflectance for soybean at 410 ppm grown at an Open Top Chamber Facility. **e** Mean, ± standard deviation ($n = 24$), and minimum and maximum leaf reflectance for peanut drought grown under greenhouse conditions. **f** Mean, ± standard deviation ($n = 24$), and minimum and maximum leaf reflectance for peanut irrigated grown under greenhouse conditions



probably due to the drought treatment; meanwhile soybean added more variability in the 500–600 nm range (Fig. S1).

Pearson’s correlations were performed to highlight which zones of spectral signatures presented negative or positive correlations with each measured parameter. Pearson’s correlations between the parameter and each wavelength were

presented separately for soybean (Fig. 2a), peanut (Fig. 2b), and both species combined (Fig. 2c). Regarding soybean $V_{c,max}$ and J_{max} values, correlation against each band showed significant ($P < 0.05$) negative values (Pearson coefficient around -0.6) in the VIS (400 nm) and in almost all SWIR (1400–2500 nm) bands (Fig. 2a). On the other hand, mid-day

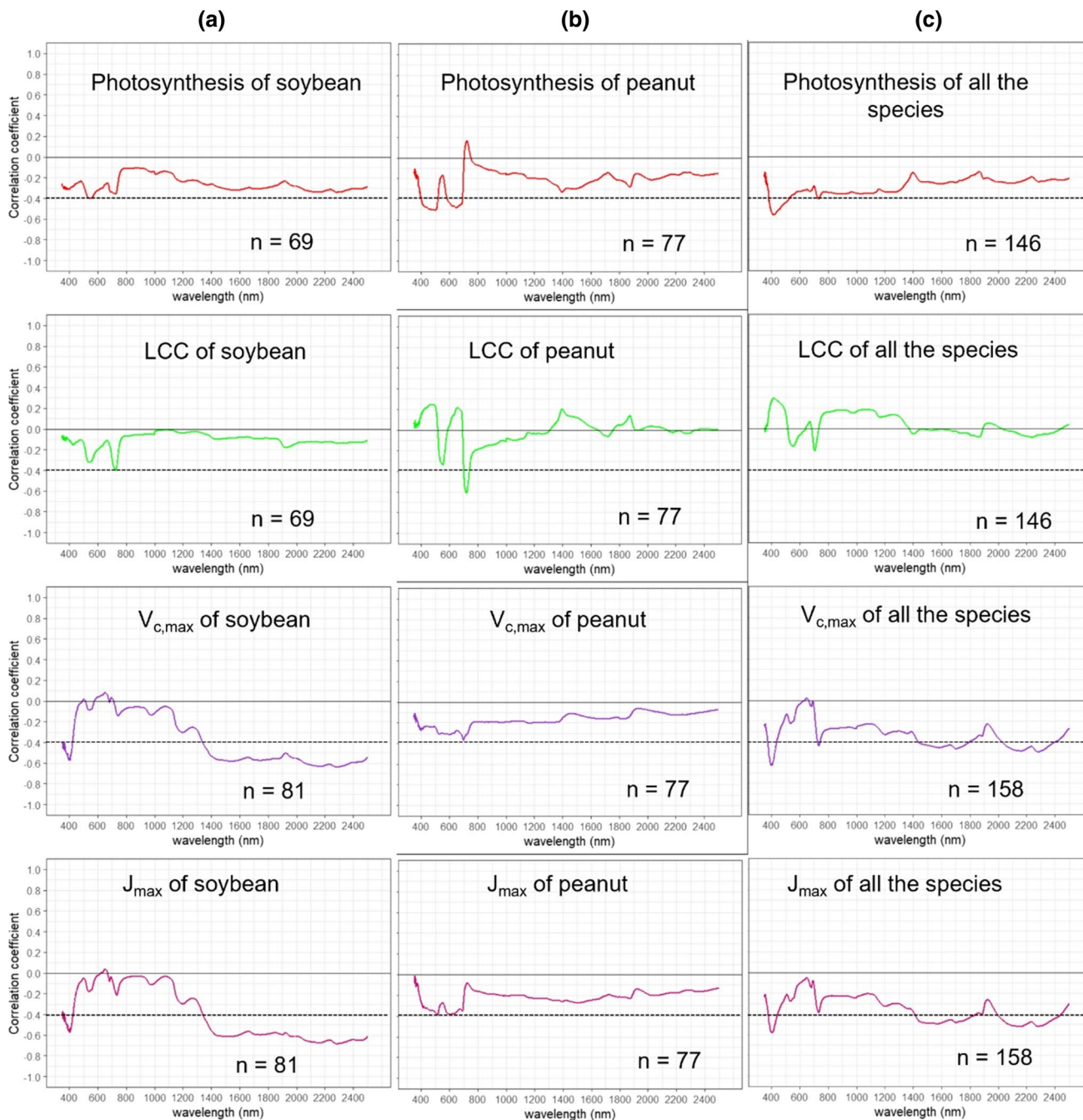


Fig. 2 Pearson’s correlation coefficients (r) between photosynthetic parameters and each wavelength from the leaf reflectance spectrum for each species and both species combined. **a** Soybean varieties under two treatments, one at high $[CO_2]$ and the other at high temperature. **b** Peanut varieties at water stress. **c** Soybean and peanut data

pooled together. Each graphic presents in the x-axis the wavelength spectrum between 350 and 2500 nm and in the y-axis the Pearson’s correlation coefficient from -1 to 1 . The discontinuous line in each graphic means the significance level $P < 0.05$ below the x-axis

photosynthesis and LCC presented lower and no significant correlation coefficients against each band from the reflectance spectrum. In the case of peanut (Fig. 2b), photosynthesis values against each wavelength band showed significant correlation ($r = -0.6$, $P < 0.05$) in VIS–NIR (400–1000 nm) bands. LCC and each wavelength showed strong correlation ($r = -0.7$, $P < 0.05$) in the NIR (700 nm). For $V_{c,max}$ and J_{max} , the correlation against each wavelength was very low or non-significant (Fig. 2b). With increased variability from combining all experiments, we could observe that mid-day photosynthesis against each wavelength showed a significant correlation ($r = -0.5$, $P < 0.05$) in the VIS (400 nm). Regarding the coefficient of correlation between $V_{c,max}$ and J_{max} , significance ($r = 0.6$, $P < 0.05$) in the VIS (400 nm) and most of the SWIR (1400–2500 nm) bands indicated an improvement relative to species analyzed separately. For this reason, we ran all advance models using combined phenotypic and spectral data from each species and environmental condition.

Estimating photosynthetic parameters using field spectroscopy and advance regression models

To test how accurately a given model estimated different photosynthetic parameters, we presented the coefficient of determination (R^2) and RMSE for each model and mean parameter, i.e., interpreted as the proportion of information in data that is explained by each model (Fig. 3). Since estimation of the $V_{c,max}$ and J_{max} parameters did not work well in the peanut experiment but worked well for the soybean (Table S1), and since the LCC estimation does not work with soybean, we decided to combine these three experiments and focus on the combination of the two crop species in this manuscript (Fig. 3). Mid-day photosynthesis showed a higher R^2 (0.62) and low RMSE (4.79) using the PLSR model using 10 components, followed by BR ($R^2 = 0.41$ and RMSE = 5.92) with the worst model being the ARDR ($R^2 = 0.28$ and RMSE = 6.55) (Fig. 3a). LCC was better assessed by PLSR ($R^2 = 0.56$ and RMSE 3.83) using 10 components, followed by ARDR ($R^2 = 0.34$ and RMSE = 4.71) with the BR model showing the worst performance ($R^2 = 0.08$ and RMSE = 5.55; Fig. 2b). The best $V_{c,max}$ model was obtained by PLSR ($R^2 = 0.70$ and RMSE = 42.80) using nine components followed by the other three models with similar values ($R^2 = 0.56$ – 0.59 ; RMSE = 50.11–52.03). Regarding J_{max} , the best model was PLSR ($R^2 = 0.50$ and RMSE = 35.83) using nine components closely followed by Lasso ($R^2 = 0.46$ and RMSE = 37.1) and BR ($R^2 = 0.45$ and RMSE = 37.41), with ARDR ($R^2 = 0.40$ and RMSE = 39.29) being the worst model.

For each of the four models, we calculated the coefficient of weight for each band and model (Fig. 4). These coefficients showed waveband contributions along the

VIS–NIR–SWIR spectrum for photosynthetic parameter estimations using leaf reflectance spectrum of pooled species, cultivars, and growing conditions. The coefficient of weight for estimating mid-day photosynthesis using PLSR showed maximum values around 400, 750, and 1750 nm, while ARDR and LASSO showed high coefficient weights at 400 nm. On the other hand, BR did not show any remarkable coefficient weights for mid-day photosynthesis (Fig. 4a). With respect to LCC, PLSR showed maximum coefficients at 400, 750, and 1750 nm, while ARDR showed a peak around 400 nm (Fig. 4b). LASSO and BR showed very low coefficients at all wavelengths (Fig. 4b). In Fig. 4c, we can observe the different coefficients of each band for $V_{c,max}$, where the maximum peaks were at 400, 700 and around 2000 nm for PLSR, BR and LASSO, while for ARDR it was only at 400 and 750 nm. For estimates of J_{max} , the highest coefficient weights for PLSR were located in SWIR (2200–2300), followed by NIR (900–1100). For the LASSO model, the strongest areas were at 400, 750, and 1750 nm (Fig. 4d), while the highest coefficients were found in the SWIR (1400–2500 nm) for BR and ARDR.

Scaling up estimations of photosynthetic parameters for potential hyperspectral aerial or satellite applications

To assess their ability to estimate photosynthetic parameters compared to full spectra captured by the Field Spec Hi-Res4 (VIS–NIR–SWIR, 350–2500 nm), we simulated other sensors with limited wavelength ranges, specifically VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), and the 12 wavelength bands of Sentinel-2 satellites (Table S2). To test this, we used reflectance data acquired by the Field Spec Hi-Res4 and separated the reflectance data according to the wavelength range of each before mentioned sensor. We then performed photosynthetic estimations using the same 4 models (PLSR, BR, ARDR, and LASSO).

Table 3, Figs. 5, and 6 show estimations of photosynthetic parameters using pooled data from both species. For mid-day photosynthesis and LCC, simulations with different sensors with just the VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), and SWIR (1400–2500 nm) spectrum regions were best performed using PLSR compared to BR, ARDR, and LASSO models (Table 3; Figs. 5, 6). However, LCC was estimated best by BR, ARDR, and LASSO using the simulated ESA Copernicus Sentinel-2 satellite multi-spectral bands (Table 3; Figs. 5, 6). Concerning estimation of $V_{c,max}$ within the VIS–NIR range (350–1000 nm) and the ESA Copernicus Sentinel-2 satellite sensors, the best performing model was PLSR using 10 components ($R^2 = 0.63$ and 0.53, respectively). For simulations of the NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), BR was the best

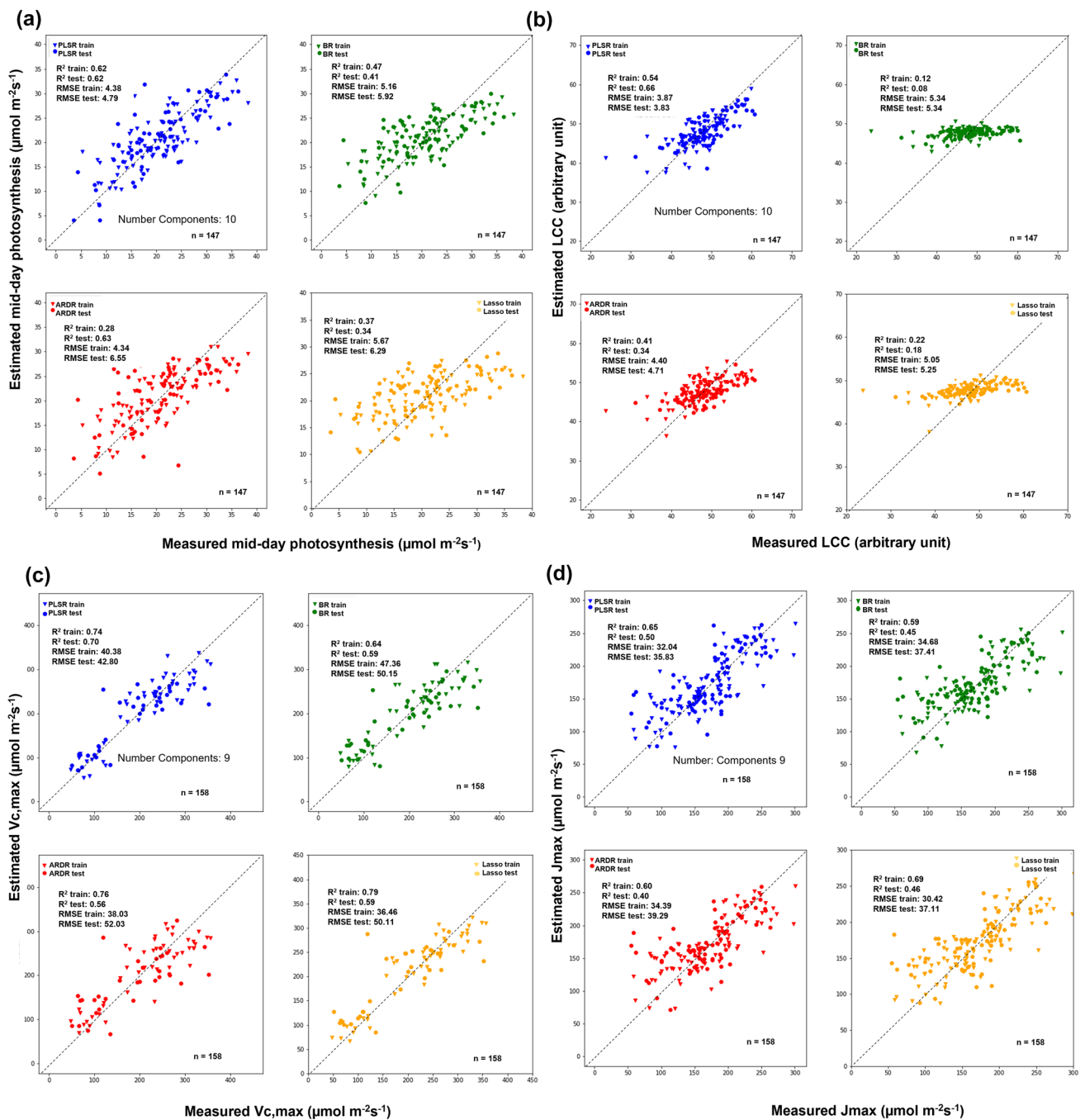


Fig. 3 Measured against estimated values correlation for different physiological parameters estimated with PLSR (blue), BR (green), ARDR (red), and LASSO (yellow) predictive models. The estimated physiological parameters are: mid-day photosynthesis (**a**), leaf chlorophyll content (**b**), maximum rate of Rubisco catalyzed carboxylation ($V_{c,max}$, **c**) and maximum electron transport rate supporting RuBP

regeneration (J_{max} , **d**) for soybean and peanut cultivars all pooled together. All the models were built using train and test data splitting them into 60 and 40%, respectively. In each graph, the R^2 , the RMSE of the train and test of the model are shown along with the size of the train and test population and number of model components (comp) used in each PLSR model. The gray dashed line shows the 1:1 line

Fig. 4 Spectral-specific coefficients for each prediction model (PLSR, BR, ARDR and LASSO) used to predict the following photosynthetic parameters of the two species pooled together. **a** Mid-day photosynthesis. **b** Leaf chlorophyll content (LCC). **c** Maximum rate of Rubisco carboxylation ($V_{c,max}$). **d** Maximum electron transport rate supporting RuBP regeneration (J_{max}). Continuous vertical lines delineate different regions of the spectrum: VIS = 450–700 nm, NIR = 700–1400, and SWIR = 1400–2500 nm

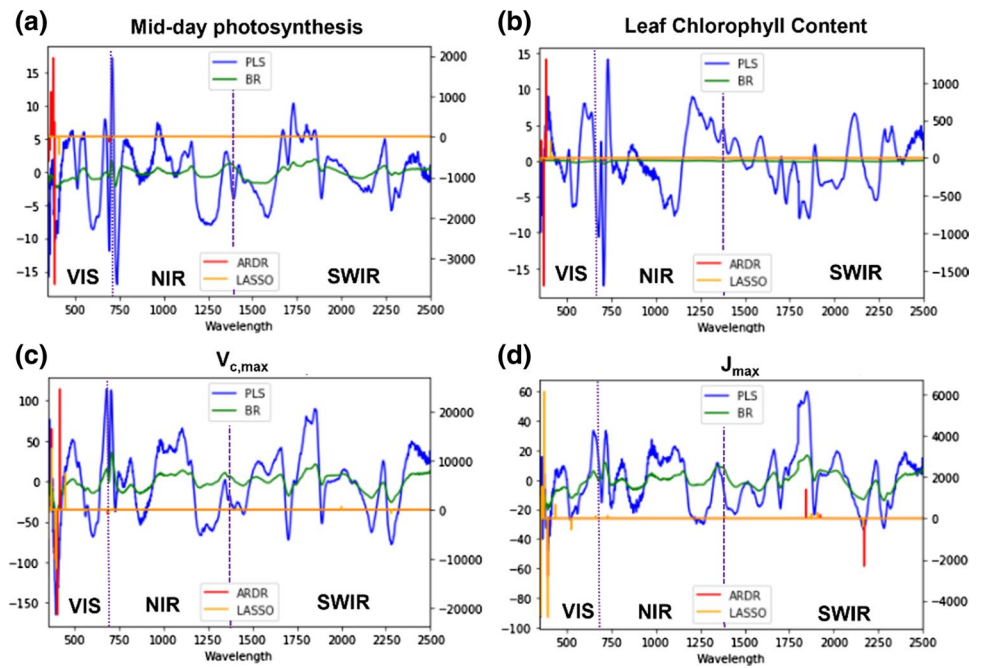


Table 3 Coefficient of determination (R^2) and root mean squared error (RMSE) of mid-day photosynthesis ($\mu\text{mol m}^{-2} \text{s}^{-1}$), leaf chlorophyll content (arbitrary units) of all species pooled together based on leaf reflectance spectra at different ranges [VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), and Sentinel-2

bands] through advance regression models: Partial Least Squares Regression (PLSR), Bayesian Ridge (BR), the Automatic Relevance Determination Regression (ARDR), and Least Absolute Shrinkage and Selection Operator (LASSO)

Estimation using the species, cultivars, and growing conditions together									
n = 146	From 350 to 1000 nm		From 1000 to 2500 nm		From 1400 to 2500 nm		Simulation of Sentinel-2		
	Model	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
Mid-day photosynthesis									
PLSR	0.47	5.61	0.42	6.1	0.52	5.53	–	–	
BR	0.28	6.54	0.47	5.84	0.50	5.69	–	–	
ARDR	0.27	6.59	0.40	6.20	0.49	5.70	–	–	
LASSO	0.34	6.29	–	–	–	–	–	–	
Leaf chlorophyll content									
PLSR	0.35	4.98	0.33	4.09	0.22	4.41	–	–	
BR	0.26	4.98	–	–	–	–	0.23	7.04	
ARDR	–	–	–	–	0.33	40.90	0.26	6.90	
LASSO	0.15	5.33	–	–	–	–	0.28	6.79	

“–” indicates that the estimation model did not yield results

model for assessing $V_{c,max}$ ($R^2=0.62$ and 0.60 , respectively). For estimating J_{max} with the VIS–NIR (350–1000 nm) sensor, the best model was LASSO ($R^2=0.42$). For the range NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm) ARDR estimated J_{max} similarly ($R^2=0.51$). PLSR, BR, and LASSO presented the same coefficient of determination

($R^2=0.41$) when using ESA Copernicus Sentinel-2 satellite simulated wavebands to assess J_{max} .

Regarding comparison of different sensors (VIS–NIR, NIR–SWIR, and SWIR) against original FieldSpec data (VIS–NIR–SWIR), we observed that estimation of mid-day photosynthesis by the different models was similar to

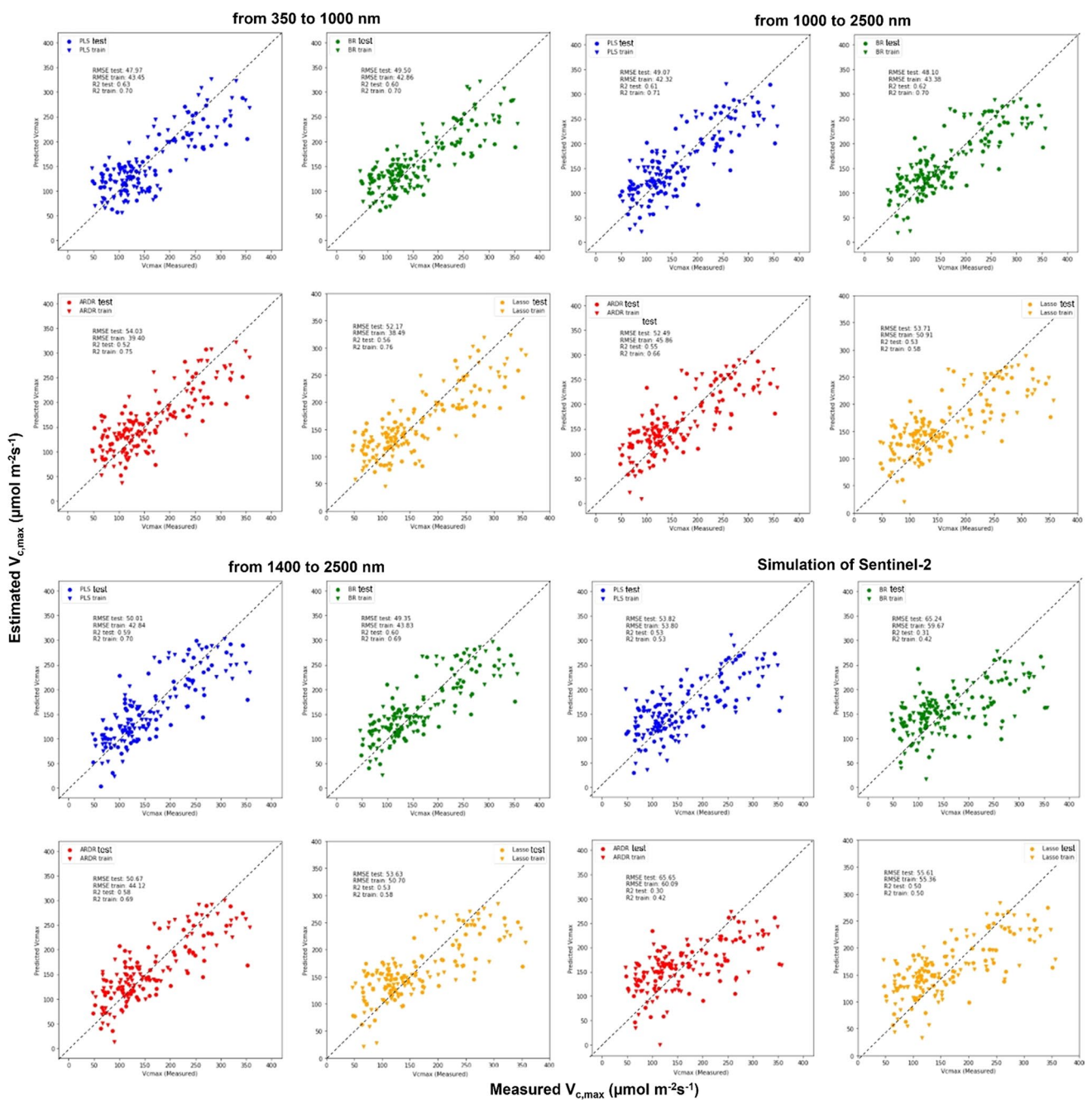


Fig. 5 Measured (X-axis) against estimated (Y-axis) correlation of maximum rate of rubisco-catalyzed carboxylation ($V_{c,max}$) estimated with PLSR (blue), BR (green), ARDR (red), and LASSO (yellow) predictive models. These models were based on leaf reflectance spectra at different ranges [VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), and Sentinel-2 bands] for

soybean and peanut cultivars all pooled together. All the models were built using the training and test split method (60 and 40%, respectively). Each graph shows the train and test R^2 and the RMSE values of for each model. For PLSR models, we used 10 components. Size of population is $n = 158$. The gray dashed line shows the 1:1 line

that of simulated sensors (Figs. 3, 5, 6 and Table 3). With ESA Copernicus Sentinel-2 satellite, estimation of mid-day photosynthesis did not work. Estimation of LCC using ESA Copernicus Sentinel-2 satellite was lower than when the whole spectrum was used. Regarding the estimation of the

$V_{c,max}$ simulating ESA Copernicus Sentinel-2 satellite, the PLSR and LASSO presented an R^2 (0.50) that was a little lower than the FieldSpec ($R^2 = 0.70$). With respect to J_{max} estimation, we observed that coefficients for the simulated NIR–SWIR and SWIR sensor ranges were very similar (but

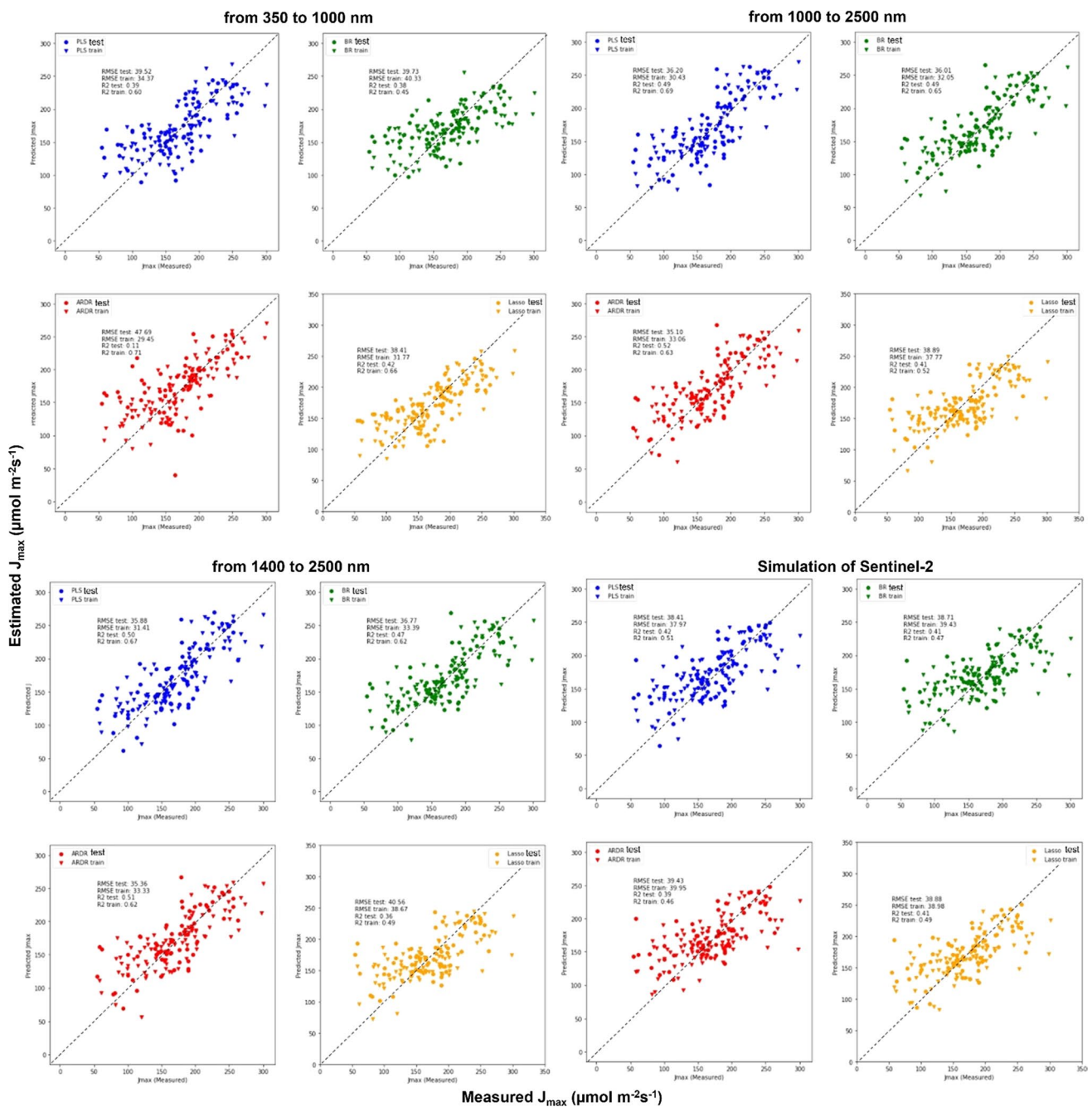


Fig. 6 Measured (axis X) against estimated (axis Y) correlation of maximum electron transport rate supporting RuBP regeneration (J_{max}) estimated with PLSR (blue), BR (green), ARDR (red), and LASSO (yellow) predictive models. These models were based on leaf reflectance spectra at different ranges [VIS–NIR (350–1000 nm), NIR–SWIR (1000–2500 nm), SWIR (1400–2500 nm), and Sentinel-2

bands] for soybean and peanut cultivars all pooled together. All the models were built using the training and test split method (60 and 40%, respectively). Each graph shows the train and test R^2 and the RMSE values of for each model. For PLSR models, we used 10 components. Size of population is $n=158$. The gray dashed line shows the 1:1 line

slightly lower) to the full-range FieldSpec (Figs. 3, 5, 6 and Table 3). The VIS–NIR and ESA Copernicus Sentinel-2 satellite simulations presented values that were lower than using the whole spectrum (Figs. 3, 5, 6 and Table 3).

Discussion

Estimating photosynthetic parameters using field spectroscopy and advance regression models

The main objective of this research was to assess which advanced statistical model (PLSR, BR, ARDR, and LASSO) was the most successful in estimating different photosynthetic parameters using leaf reflectance spectra (VIS–NIR–SWIR, 350–2500 nm) from two legume species. The use of advance regression models to predict different physiological parameters needs ample phenotypic variation to be accurate (Kuhn and Johnson 2013). Since statistical effects of different treatments over some variables were not significant (Table 1), we combined findings from three experiments (two different species) to increase phenotypic range for better parameter estimation with all models rather than examining each species separately (Table S1). Similar approaches have been used recently to increase phenotypic variation and obtain a better prediction model by including different species and/or cultivars (Doughty et al. 2011; Serbin et al. 2012; Choquette et al. 2019), different abiotic stresses such as drought (Silva-Perez et al. 2018), or elevated atmospheric ozone concentrations (Ainsworth et al. 2014; Yendrek et al. 2017).

In our study, when data from both legumes were combined, almost all of the advanced models were able to estimate $V_{c,max}$ and J_{max} at greater than $R^2 > 0.50$ (Fig. 3). Of the four models used to predict these two parameters, PLSR was the overall best model for $V_{c,max}$ ($R^2 = 0.70$ and RMSE 42.80) and J_{max} ($R^2 = 0.50$ and 35.83), followed by LASSO and BR for $V_{c,max}$ ($R^2 = 0.59$ with RMSE 50.11; 0.59 with a RMSE 50.15, respectively), and BR and LASSO for J_{max} ($R^2 = 0.45$ with a RMSE 37.11; 0.46 with a RMSE 37.41, respectively) (Fig. 3). This may be because the PLSR model does not estimate shrinkage when performing variable selection (spectral wavebands) as do BR, ARDR, and LASSO (Neal 1996; Tipping 2001; Wold et al. 2001). Others have also found that PLSR and LASSO had similar estimation capacities, showing that LASSO band block contribution was similar to the PLSR model (Fu et al. 2020). Specific reasons why PLSR was more efficient at estimating photosynthetic parameters assessed in this study are discussed in detail below.

Successful predictions of $V_{c,max}$ ($R^2 = 0.89$ with a RMSE 15.4) and J_{max} ($R^2 = 0.93$ with a RMSE 18.67) using PLSR have been previously obtained by combining two tree species (Serbin et al. 2012); this study showed statistically

significant phenotypic variation due to temperature treatments as well as species. In our study, the lower R^2 associated with $V_{c,max}$ and J_{max} estimates could be attributed to the lack of effect of some environmental treatments (temperature, elevated CO_2 , and drought) and cultivars over these parameters (Table 1). However, Ainsworth et al. (2014) showed a significant correlation between measured and estimated $V_{c,max}$ ($R^2 = 0.88$ with a RMSE 13.4) with the effect of treatments (elevated ozone) and cultivars not being significant. This demonstrated that good parameter estimation and significant treatment or cultivar effects are not mutually exclusive and that it is only necessary to have sufficient range in variation of phenotypic data. For example, Ainsworth et al. (2014) and Serbin et al. (2012) noted $V_{c,max}$ variation (60–280 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and 40–170 $\mu\text{mol m}^{-2} \text{s}^{-1}$, respectively) similar to the values obtained in this study when all three experiments were combined (48–348 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for the current experiment). Since the ranges in variation of $V_{c,max}$ and J_{max} data are similar but higher to those obtained in the above-mentioned research, why are R^2 values in the current study for $V_{c,max}$ ($R^2 = 0.70$) and J_{max} ($R^2 = 0.50$) lower and RSME (42.80 and 35.83, respectively) higher than in those studies? Tibshirani (1996) has noted that PLSR models lose accuracy when estimating parameters across different environments. Research by Serbin et al. (2012) and Ainsworth et al. (2014) were each performed in one environment (greenhouse and field, respectively) for one growing season, while our study combined information from three experiments representing distinct environments (greenhouse, growth chambers, and open top chambers) with plants grown at very different environmental conditions. In an experiment with several corn breeding lines grown under ambient and elevated ozone repeated over three growing seasons, Yendrek et al. (2017) obtained $V_{c,max}$ estimations ($R^2 = 0.55$ with RMSE 6.61, and 0.65 with a RMSE 6.60) similar to those reported in our study but with a RMSE lower than ours. This was probably due to the effects of changing environments on PLSR performance (Serbin et al. 2012; Ainsworth et al. 2014). Regarding the lower RMSE obtained in the above-mentioned publications (Serbin et al. 2012; Ainsworth et al. 2014; Yendrek et al. 2017) in comparison with those obtained in our research, this could be due to the different cross-validation used in our approach. In our cross-validation, the test error rate can be highly variable, depending on which observations are included in the training set and which observations are included in the validation set. This may be the reason for the higher RMSE values observed in $V_{c,max}$ and J_{max} . Also the high RMSE values can be due to a higher phenotypic range as a result of including two crop species grown in three very different environments. This highlights the importance of performing calibration experiments under multiple environments. Other issue that can arise is the use of these models with completely new set of

cultivars and experimental conditions as was tested in Yendrek et al. (2017). In such a case, it would be recommendable to test model precision by measuring spectral reflectance under new conditions and corroborating model estimates of extreme values for $V_{c,max}$ with ground truth measurements of the photosynthetic parameter. Although this extra step will take more time, this procedure could serve to test model accuracy and help improve the model with new training data.

To solve this multiple environment/location problem, new approaches need to be developed and implemented. For example, Fu et al. (2020) increased prediction model accuracy by stacking different machine learning algorithms (i.e., R^2 increases of 0.1–0.2 over single prediction models). Another alternative would be creation of a consortium of scientists interested in using hyperspectral reflectance technology to predict physiological traits. Their combined expertise would create strong standardized calibrations that could be used across multiple environments as has been done for assessing forage quality traits using NIRS technology (i.e., NIRS Consortium; <https://www.nirsconsortium.org/>).

Estimation of mid-day photosynthesis using PLSR, BR, ARDR, and LASSO presented lower R^2 values (≈ 0.29 – 0.62) than for $V_{c,max}$ and J_{max} (Fig. 3) since in situ photosynthetic measurements are likely more influenced by environment (Sanz-S  ez et al. 2017; Soba et al. 2020) than by leaf structure and biochemistry (Serbin et al. 2012; Ainsworth et al. 2014). Thus, a looser estimation was expected. Due to environmental variability, few reports have estimated mid-day photosynthesis. However, our PLSR estimation was better than the observations of Vitrack-Tamam et al. (2020) for cotton stomatal conductance ($R^2 = 0.23$); this was likely due to the lower range spectral reflectance device used in their experiment (633–1659 nm). Similar estimations of net photosynthesis were accomplished using the scaled photochemical reflectance index and a FieldSpec Hi-Res Device (Kumari et al. 2012).

Regarding spectral wavelength specific coefficients for each estimation model for $V_{c,max}$ and J_{max} , the most frequent selection for the four models was the VIS waveband (Fig. 4) where chlorophyll and other pigments have strong absorption features (Pe  uelas and Filella 1998). However, these models also used wavebands in the NIR and SWIR, similar to other studies (Hansen and Schjoerring 2003; Doughty et al. 2011; Serbin et al. 2012; Ainsworth et al. 2014; Yendrek et al. 2017). In addition, Rubisco has several relatively broad spectral absorption features in the NIR and SWIR (Elvidge 1990). These selections of spectral region combinations indicate that $V_{c,max}$ and J_{max} spectral signatures are not simply a function of chlorophyll content, which suggests that more information is needed beyond the VIS–NIR wavebands to estimate such complex processes. The inclusion of a broader range of wavebands, due in part to less penalizations, is likely why the PLSR model outperformed

BR, ARDR, and LASSO by more effectively capturing the broader spectral absorption features of Rubisco. For example, the $V_{c,max}$ LASSO model only selected specific coefficients at 540, 680, 720, 2000, and 2250 nm (Fig. 4c), while the PLSR model had significant coefficient ranges between 400–450, 700–800, and 1750–1900 (Fig. 4c). Photosynthesis and LCC also presented the highest selection of spectral peaks in the VIS, followed by NIR; this has been extensively documented through both vegetation indices that estimate chlorophyll pigment content and also by the Photochemical Reflectance Index (PRI) that predicts photosynthetic efficiency through a zeaxanthin absorption feature (Gamon et al. 1997; Gitelson et al. 2005; Schlemmera et al. 2013).

We also present a more in-depth comparison of the four models. As shown in Fig. 3, the R^2 of models do not present significant differences between each other, although we can see that the models used different numbers of coefficients to estimate each parameter (Fig. 4). This was reflected in the algorithm differences in each model approach to parsimony, the simple explanation of an occurrence involving the fewest entities, assumptions, or changes. This means that a fewer number of weight coefficients were used to estimate the different parameters (Vandekerckhove and Matzke 2015). In our study, all PLSR models (blue line in Fig. 4) used VIS, NIR, and SWIR wavelengths, but potentially over-fitted by an over-inclusion of predictor variables (Geladi et al. 1986; Wold et al. 2001). This contrasts to the BR (in green), ARDR (in red), and LASSO (in yellow) models (Fig. 4), which used more specific and limited spectra than restricted models that penalize the lesser coefficients (Neal 1996; Tibshirani 1996; Tipping 2001).

Scaling up estimations of photosynthetic parameters for potential hyperspectral aerial or satellite applications

The second aim of this study was to simulate different sensors with more limited spectral coverage (VIS–NIR, NIR–SWIR, and SWIR), including the ESA Copernicus Sentinel-2 satellite 13 bands. We found that estimation of $V_{c,max}$ using three different sensor ranges (VIS–NIR–SWIR) with the four models performed ($R^2 = 0.50$) surprisingly similar to the whole spectrum (Figs. 3 and 5). For J_{max} , the highest estimation ($R^2 = 0.51$) used NIR–SWIR and SWIR data in ARDR. This was quite similar to Meacham-Hensold et al. (2020) who used PLSR models and canopy-level spectra with three different spectral ranges (500–900, 500–1700, and 500–2400 nm) to achieve $V_{c,max}$ estimations near $R^2 = 0.60$ and J_{max} estimations around $R^2 = 0.40$.

We also resampled FieldSpec data to cover the 12 spectral bands of the ESA Copernicus Sentinel-2 satellite; these were quite similar to spectral ranges selected by the coefficients used by the different models to estimate photosynthetic

parameters. Concerning the different photosynthetic parameters, only $V_{c,max}$ was estimated at more than $R^2=0.50$. This could be related to the carboxylation process ($V_{c,max}$) having several relatively broad spectral absorption features in NIR and SWIR centered at 1.5, 1.68, 1.74, 1.94, 2.05, 2.29 μm , etc. (Elvidge 1990), which are in close proximity to several Sentinel-2 wavelength bands (Table S2). Supplementary data (Table S3) and Serbin et al. (2012) showed that wavelengths (490, 610, 690, 710, 1680, 1940, 2200, 2400 nm) used to estimate $V_{c,max}$ have some bands similar to Sentinel-2. Figure 5d also shows that the spectral regions used in PLSR models were similar to Sentinel bands (Yendrek et al. 2017). The limited success of single-leaf-level estimations of photosynthetic capacities using point-based spectral analysis (Serbin et al. 2015) found considerable promise in airborne and potential promise in space-borne imaging spectroscopy such as the NASA HypSIRI mission (Mariotto et al. 2013). In this regard, hyperspectral imagery through inversion of the Soil-Canopy Observation of Photosynthesis and Energy (SCOPE) model to estimate $V_{c,max}$ also uses sensor resolutions available in airborne or even precision agriculture technologies (Camino et al. 2019). Recently, one plot-level study using sunlit vegetative reflectance pixels from a single visible near infra-red (VNIR; 400–900 nm) hyperspectral camera reported determination coefficients of $R^2=0.79$ for $V_{c,max}$ and $R^2=0.59$ for J_{max} (Meacham-Hensold et al. 2020). Thus, our simulation analyses and other recent literature suggest that the wide range of variability in VIS, NIR, and SWIR sensors and the Sentinel-2 multispectral sensor (to a more limited extent) could be employed to estimate photosynthetic parameters (including $V_{c,max}$ and J_{max}) with advanced regression models. However, more research needs to be done in this area as one of the limitations of this work was that we measured leaf reflectance with a leaf clip, while UAV and satellites measure canopy reflectance that can be different from single leaf reflectance. For the future, we suggest to test if canopy reflectance measurements at different precision levels can predict leaf level photosynthetic measurements or even canopy-level photosynthesis as has been done with models such as PROSAIL (Berger et al. 2018).

Conclusion and future directions

In this study, we estimated $V_{c,max}$ and J_{max} using leaf spectral reflectance data and different advanced regression models with determination coefficients higher than $R^2=0.50$ – 0.70 . The combination of different species and environmental conditions (elevated $[\text{CO}_2]$, high temperature, and drought) increased phenotypic variation and improved model estimations where treatment effects were not significant. To achieve higher coefficients of determination and model performance, this research demonstrated

that it is more important to have a wider range of phenotypic variation than a significant effect of a treatment or cultivar. We suggest that estimating photosynthetic capacity from reflectance spectra may be considered sufficiently robust to be useful for several different plant physiological applications, such as abiotic stress detection, improved characterization of photosynthesis process-based crop models, and a prescreening tool in breeding programs. We demonstrated that PLSR was the best model for predicting photosynthetic parameters in comparison to other advanced regression models (BR, ARDR and LASSO). However, new advance regression approaches that combine different regression models may be employed to increase phenotype estimation using this technology. Based on simulation of four limited spectral range sensors (VIS–NIR, NIR–SWIR and SWIR) using a leaf level spectrophotometer, we demonstrated that it is possible to estimate $V_{c,max}$ with similar precision compared to using the whole VIS–NIR–SWIR spectrum. This research should encourage future studies using different imaging sensors (hyperspectral and multispectral) at different scales for estimating $V_{c,max}$ and J_{max} .

Author contribution statement MLB Experimentation, curation of the data, formal analysis, writing original draft. DS Experimentation, review and editing. TS Experimentation, review and editing. JL Experimentation, review and editing. IA Resource managing, review and editing. JLA Resource managing, review and editing. GBR Experimentation, review and editing. SAP Experimentation, review and editing, resource managing. SCK Conceptualization, data curation, resource managing, formal analysis, supervision, writing original draft. ASS Conceptualization, experimentation, data curation, resource managing, formal analysis, supervision, project administration, writing original draft.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00425-022-03867-6>.

Acknowledgements The authors would like to thank the technical help during the experiment of Mr. Robert Icenogle, Barry Dorman (USDA-ARS), Seth Johnston, and Mary Durstock (Crop Physiology Laboratory, Auburn University). The authors also would like to thank to Dr. Jose A. Jimenez Berni for statistical support to analyze the data. This research was supported by the Action CA17134 SENSECO (Optical Synergies for Spatiotemporal Sensing of Scalable Ecophysiological Traits) funded by COST (European Cooperation in Science and Technology, www.cost.eu). This research was also supported by Auburn University and Alabama Agricultural Experimental Station Seed Grant.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data availability statement The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

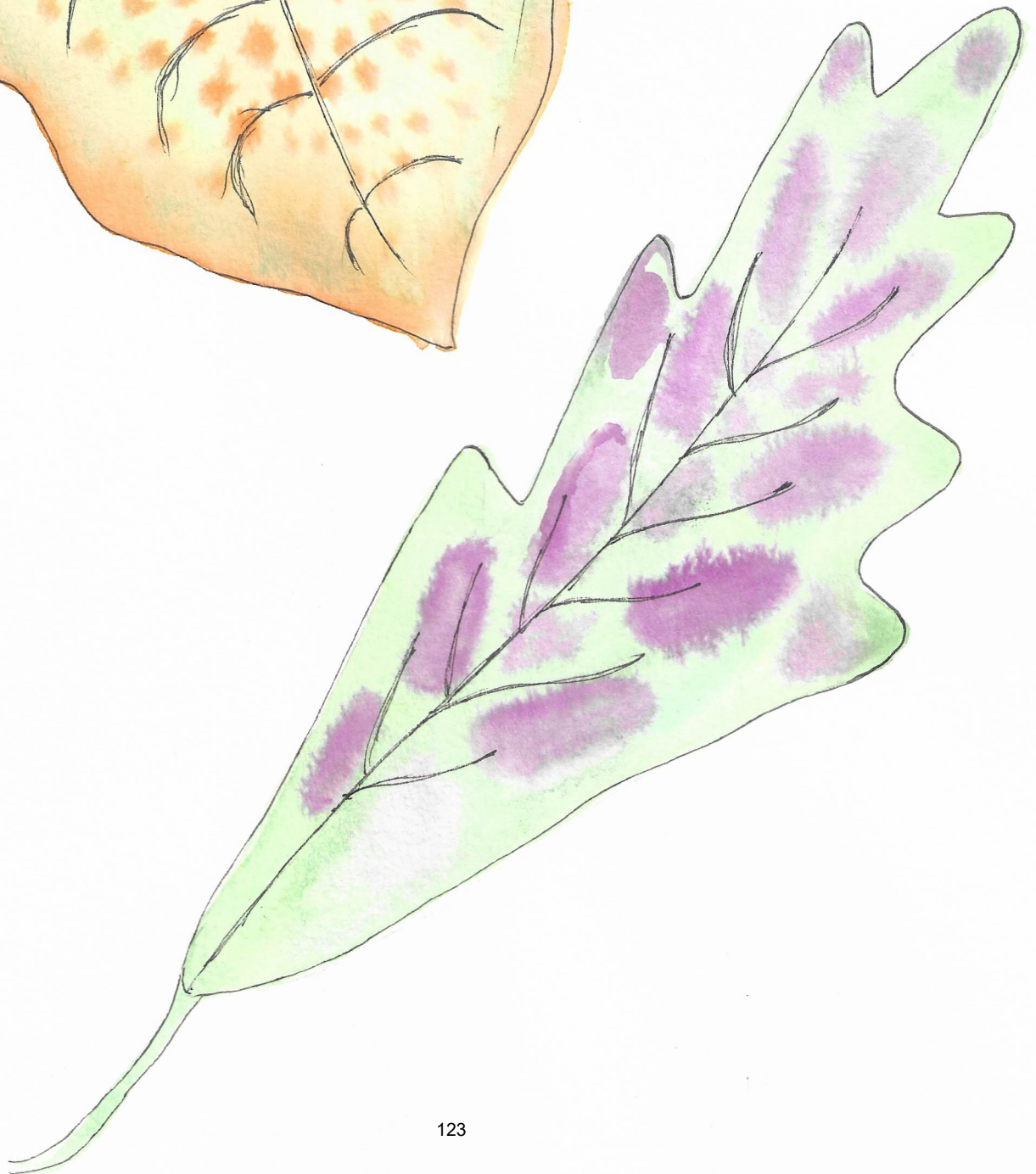
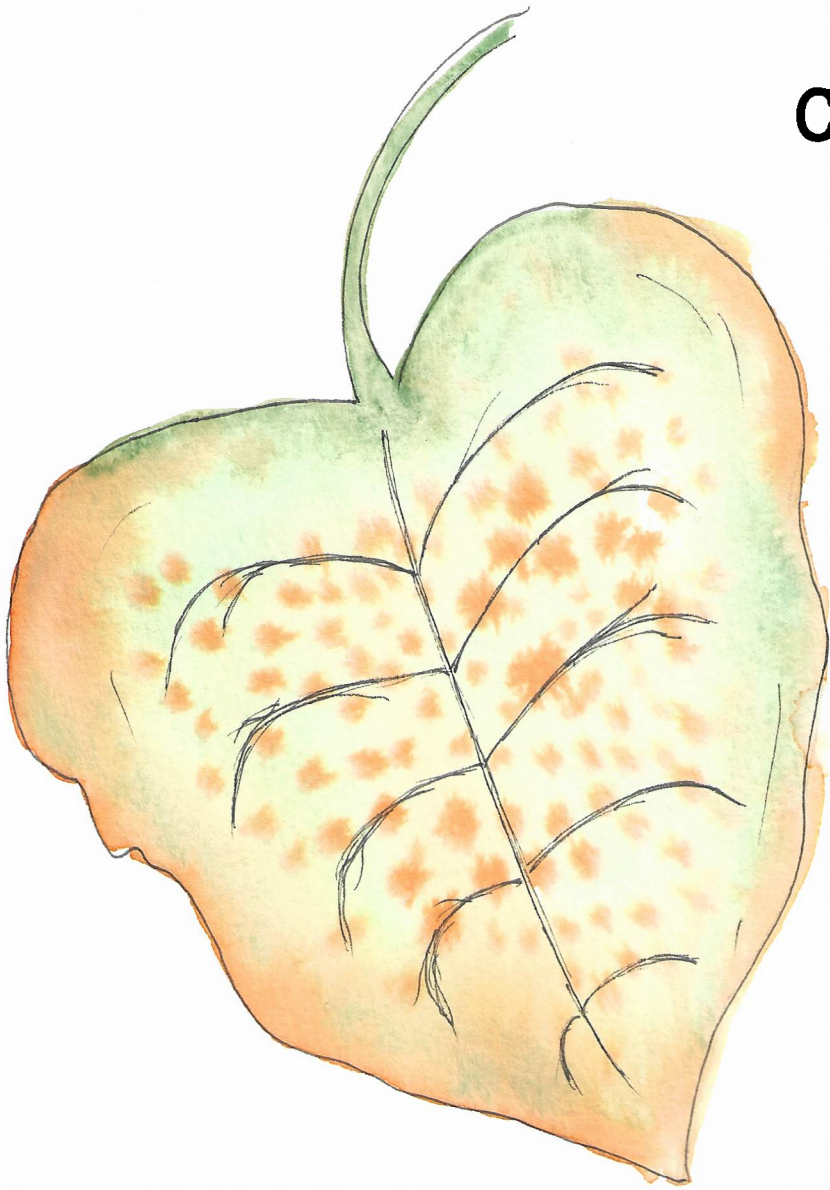
- Adachi S, Tsuru Y, Nito N et al (2011) Identification and characterization of genomic regions on chromosomes 4 and 8 that control the rate of photosynthesis in rice leaves. *J Exp Bot* 62:1927–1938. <https://doi.org/10.1093/jxb/erq387>
- Ainsworth EA, Rogers A, Nelson R, Long SP (2004) Testing the “source-sink” hypothesis of down-regulation of photosynthesis in elevated [CO₂] in the field with single gene substitutions in *Glycine max*. *Agric for Meteorol* 122:85–94. <https://doi.org/10.1016/j.agrformet.2003.09.002>
- Ainsworth EA, Serbin SP, Skoneczka JA, Townsend PA (2014) Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynth Res* 119:65–76. <https://doi.org/10.1007/s11120-013-9837-y>
- Aranjuelo I, Pardo A, Biel C et al (2009) Leaf carbon management in slow-growing plants exposed to elevated CO₂. *Glob Chang Biol* 15:97–109. <https://doi.org/10.1111/j.1365-2486.2008.01829.x>
- Aranjuelo I, Cabrerizo PM, Arrese-Igor C, Aparicio-Tejo PM (2013) Pea plant responsiveness under elevated [CO₂] is conditioned by the N source (N₂ fixation versus NO₃- fertilization). *Environ Exp Bot* 95:34–40. <https://doi.org/10.1016/j.envexpbot.2013.06.002>
- Araus JL, Cairns JE (2014) Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci* 19:52–61. <https://doi.org/10.1016/j.tplants.2013.09.008>
- Araus JL, Kefauver SC, Zaman-Allah M et al (2018) Translating high-throughput phenotyping into genetic gain. *Trends Plant Sci* 23:451–466. <https://doi.org/10.1016/j.tplants.2018.02.001>
- Asner GP, Martin RE (2008) Spectral and chemical analysis of tropical forests: scaling from leaf to canopy levels. *Remote Sens Environ* 112:3958–3970. <https://doi.org/10.1016/j.rse.2008.07.003>
- Berger K, Atzberger C, Danner M et al (2018) Evaluation of the PRO-SAIL model capabilities for future hyperspectral model environments: a review study. *Remote Sens* 10:85. <https://doi.org/10.3390/rs10010085>
- Camino C, Gonzalez-Dugo V, Hernandez P, Zarco-Tejada PJ (2019) Radiative transfer V_{max} estimation from hyperspectral imagery and SIF retrievals to assess photosynthetic performance in rain-fed and irrigated plant phenotyping trials. *Remote Sens Environ* 231:111186. <https://doi.org/10.1016/j.rse.2019.05.005>
- Choquette NE, Ogut F, Wertin TM et al (2019) Uncovering hidden genetic variation in photosynthesis of field-grown maize under ozone pollution. *Global Chang Biol* 25:4327–4338. <https://doi.org/10.1111/gcb.14794>
- Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held R, Jones R, et al (2007) Regional climate projections. In: *Climate Change 2007: The physical science basis*. Contribution of Working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, University Press, Cambridge, Chapter 11, ISBN: 978-0-521-88009-1
- Coast O, Shah S, Ivakov A et al (2019) Predicting dark respiration rates of wheat leaves from hyperspectral reflectance. *Plant Cell Environ* 42:2133–2150. <https://doi.org/10.1111/pce.13544>
- Colombo R, Meroni M, Marchesi A et al (2008) Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. *Remote Sens Environ* 112:1820–1834. <https://doi.org/10.1016/j.rse.2007.09.005>
- Couture JJ, Singh A, Rubert-Nason KF et al (2016) Spectroscopic determination of ecologically relevant plant secondary metabolites. *Methods Ecol Evol* 7:1402–1412. <https://doi.org/10.1111/2041-210X.12596>
- de Oliveira Silva FM, Lichtenstein G, Alseikh S et al (2018) The genetic architecture of photosynthesis and plant growth-related traits in tomato. *Plant Cell Environ* 41:327–341. <https://doi.org/10.1111/pce.13084>
- Dechant B, Cuntz M, Vohland M et al (2017) Estimation of photosynthesis traits from leaf reflectance spectra: correlation to nitrogen content as the dominant mechanism. *Remote Sens Environ* 196:279–292. <https://doi.org/10.1016/j.rse.2017.05.019>
- Dhanapal AP, Ray JD, Singh SK et al (2015) Genome-wide association study (GWAS) of carbon isotope ratio ($\delta_{13}\text{C}$) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. *Theor Appl Genet* 128:73–91. <https://doi.org/10.1007/s00122-014-2413-9>
- Doughty CE, Asner GP, Martin RE (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia* 165:289–299. <https://doi.org/10.1007/s00442-010-1800-4>
- Drusch M, Del Bello U, Carlier S et al (2012) Remote sensing of environment Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens Environ* 120:25–36. <https://doi.org/10.1016/j.rse.2011.11.026>
- Elvidge CD (1990) Visible and near infrared reflectance characteristics of dry plant materials. *Int J Remote Sens* 11:1775–1795. <https://doi.org/10.1080/01431169008955129>
- FAOSTAT (2016) The State of Food and Agriculture 2016 (SOFA): Climate change, agriculture and food security. Rome. <https://www.fao.org/3/i6030e/i6030e.pdf>
- Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* 149:78–90. <https://doi.org/10.1007/BF00386231>
- Fehr WR, Caviness CE, Burmood DT, Pennington JS (1971) Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci* 11:929–931. <https://doi.org/10.2135/cropsci1971.0011183X001100060051x>
- Ferreira RA, Pachepsky LB, Collino D, Acock B (2000) Modeling peanut leaf gas exchange for the calibration of crop models for different cultivars. *Ecol Model* 131:285–298. [https://doi.org/10.1016/S0304-3800\(00\)00252-0](https://doi.org/10.1016/S0304-3800(00)00252-0)
- Fu P, Meacham-Hensold K, Guan K et al (2020) Estimating photosynthetic traits from reflectance spectra: a synthesis of spectral indices, numerical inversion, and partial least square regression. *Plant Cell Environ*. <https://doi.org/10.1111/pce.13718>
- Gamon JA, Serrano L, Surfus JS (1997) The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* 112:492–501. <https://doi.org/10.1007/s004420050337>
- Geladi P, Kowalski B (1986) Partial least-squares regression: a tutorial. *Anal Chim Acta* 185:1–17
- Gitelson AA, Viña A, Ciganda V et al (2005) Remote estimation of canopy chlorophyll content in crops. *Geophys Res Lett* 32:1–4. <https://doi.org/10.1029/2005GL022688>
- Gray SB, Brady SM (2016) Plant developmental responses to climate change. *Dev Biol* 419:64–77. <https://doi.org/10.1016/j.ydbio.2016.07.023>

- Hansen PM, Schjoerring JK (2003) Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens Environ* 86:542–553. [https://doi.org/10.1016/S0034-4257\(03\)00131-7](https://doi.org/10.1016/S0034-4257(03)00131-7)
- Heckmann D, Schlüter U, Weber APM (2017) Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Mol Plant* 10:878–890. <https://doi.org/10.1016/j.molp.2017.04.009>
- Hunt J (2019) Introduction to Matplotlib. In: *Advanced guide to Python 3 programming*. Springer, Cham, pp 35–42. <https://doi.org/10.1007/978-3-030-25943-3>
- Jones E, Oliphant T, Peterson P (2001) SciPy: Open source scientific tools for Python, online
- Khan HA, Nakamura Y, Furbank RT, Evans JR (2021) Effect of leaf temperature on the estimation of photosynthetic and other traits of wheat leaves from hyperspectral reflectance. *J Exp Bot* 72:1271–1281. <https://doi.org/10.1093/jxb/eraa514>
- Koester RP, Skoneczka JA, Cary TR et al (2014) Historical gains in soybean (*Glycine max* Merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *J Exp Bot* 65:3311–3321. <https://doi.org/10.1093/jxb/eru187>
- Kuhn M, Johnson K (2013) *Applied predictive modeling*. Springer, New York. <https://doi.org/10.1007/978-1-4614-6849-3>
- Kumari M, Patel NR, Raj R et al (2012) Parametric estimation of net photosynthesis in rice from in-situ spectral reflectance measurements. *Curr Sci* 103:55–61
- Long SP, Bernacchi CJ (2003) Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *J Exp Bot* 54:2393–2401. <https://doi.org/10.1093/jxb/erg262>
- Long SP, Ainsworth EA, Rogers A, Ort DR (2004) Rising atmospheric carbon dioxide: plants FACE the future. *Annu Rev Plant Biol* 55:591–628. <https://doi.org/10.1146/annurev.arplant.55.031903.141610>
- Mariotto I, Thenkabail PS, Huete A et al (2013) Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission. *Remote Sens Environ* 139:291–305. <https://doi.org/10.1016/j.rse.2013.08.002>
- Meacham-Hensold K, Fu P, Wu J et al (2020) Plot-level rapid screening for photosynthetic parameters using proximal hyperspectral imaging. *J Exp Bot* 71:2312–2328. <https://doi.org/10.1093/jxb/eraa068>
- Meena RS, Das A, Yadav GS, Lal R (2018) *Legumes for soil health and sustainable management*. Springer, Singapore
- Mitchell RJ, Runion GB, Prior SA et al (1995) Effects of nitrogen on *Pinus palustris* foliar respiratory responses to elevated atmospheric CO₂ concentration. *J Exp Bot* 46:1561–1567. <https://doi.org/10.1093/jxb/46.10.1561>
- Neal RM (1996) *Bayesian learning for neural networks*. Springer Verlag, New York
- Oakley CG, Savage L, Lotz S et al (2018) Genetic basis of photosynthetic responses to cold in two locally adapted populations of *Arabidopsis thaliana*. *J Exp Bot* 69:699–709. <https://doi.org/10.1093/jxb/erx437>
- Ort DR, Merchant SS, Alric J et al (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc Natl Acad Sci USA* 112:8529–8536. <https://doi.org/10.1073/pnas.1424031112>
- Oury FX, Godin C, Mailliard A et al (2012) A study of genetic progress due to selection reveals a negative effect of climate change on bread wheat yield in France. *Eur J Agron* 40:28–38. <https://doi.org/10.1016/j.eja.2012.02.007>
- Peñuelas J, Filella L (1998) Technical focus: visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci* 3:151–156. [https://doi.org/10.1016/S1360-1385\(98\)01213-8](https://doi.org/10.1016/S1360-1385(98)01213-8)
- Petisco C, García-Criado B, Mediavilla S et al (2006) Near-infrared reflectance spectroscopy as a fast and non-destructive tool to predict foliar organic constituents of several woody species. *Anal Bioanal Chem* 386:1823–1833. <https://doi.org/10.1007/s00216-006-0816-4>
- Reynolds M, Foulkes J, Furbank R et al (2012) Achieving yield gains in wheat. *Plant Cell Environ* 35:1799–1823. <https://doi.org/10.1111/j.1365-3040.2012.02588.x>
- Rogers A (2014) The use and misuse of $V_{c,max}$ in earth system models. *Photosynth Res* 119:15–29. <https://doi.org/10.1007/s11120-013-9818-1>
- Rogers HH, Bingham GE, Cure JD et al (1983) Responses of selected plant species to elevated carbon dioxide in the field. *J Environ Qual* 12:569–574. <https://doi.org/10.2134/jeq1983.00472425001200040028x>
- RStudio Team (2020) RStudio: integrated development environment for R. RStudio. PBC, Boston
- Sanz-Sáez Á, Erice G, Aguirreolea J et al (2012) Alfalfa yield under elevated CO₂ and temperature depends on the *Sinorhizobium* strain and growth season. *Environ Exp Bot* 77:267–273. <https://doi.org/10.1016/j.envexpbot.2011.11.017>
- Sanz-Sáez Á, Koester RP, Rosenthal DM et al (2017) Leaf and canopy scale drivers of genotypic variation in soybean response to elevated carbon dioxide concentration. *Glob Chang Biol* 23:3908–3920. <https://doi.org/10.1111/gcb.13678>
- Schlemmer M, Gitelson A, Schepers J et al (2013) Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *Int J Appl Earth Obs Geoinf* 25:47–54. <https://doi.org/10.1016/j.jag.2013.04.003>
- Segarra J, Buchailot ML, Araus JL, Kefauver SC (2020) Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* 10:1–18. <https://doi.org/10.3390/agronomy10050641>
- Serbin SP (2012) Spectroscopic determination of leaf nutritional, morphological, and metabolic traits. <https://doi.org/10.6084/M9.FIGSHARE.745311.V1>
- Serbin SP, Dillaway DN, Kruger EL, Townsend PA (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *J Exp Bot* 63:489–502. <https://doi.org/10.1093/jxb/err294>
- Serbin SP, Singh A, McNeil BE et al (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecol Appl* 24:1651–1669. <https://doi.org/10.1890/13-2110.1>
- Serbin SP, Singh A, Desai AR et al (2015) Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. *Remote Sens Environ* 167:78–87. <https://doi.org/10.1016/j.rse.2015.05.024>
- Sharkey TD, Bernacchi CJ, Farquhar GD, Singsaas EL (2007) Fitting photosynthetic carbon dioxide response curves for C₃ leaves. *Plant Cell Environ* 30:1035–1040. <https://doi.org/10.1111/j.1365-3040.2007.01710.x>
- Silva-Pérez V, Furbank RT, Condon AG, Evans JR (2017) Biochemical model of C₃ photosynthesis applied to wheat at different temperatures. *Plant Cell Environ* 40:1552–1564. <https://doi.org/10.1111/pce.12953>
- Silva-Perez V, Molero G, Serbin SP et al (2018) Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *J Exp Bot* 69:483–496. <https://doi.org/10.1093/jxb/erx421>
- Simkin AJ, López-Calcagno PE, Raines CA (2019) Feeding the world: Improving photosynthetic efficiency for sustainable crop production. *J Exp Bot* 70:1119–1140. <https://doi.org/10.1093/jxb/ery445>

- Soba D, Shu T, Runion GB et al (2020) Effects of elevated [CO₂] on photosynthesis and seed yield parameters in two soybean genotypes with contrasting water use efficiency. *Environ Exp Bot* 178:104154. <https://doi.org/10.1016/j.envexpbot.2020.104154>
- Tibshirani R (1996) Regression shrinkage and selection via the Lasso. *J R Stat Soc Ser B* 58:267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. *J Mach Learn Res* 1:211–244
- Vandekerckhove J, Matzke D, Wagenmakers E-J (2015) Model comparison and the principle of parsimony. In: Busemeyer JR et al (eds) *Oxford handbook of computational and mathematical psychology*. Oxford University Press, Oxford
- Varoquaux G, Buitinck L, Louppe G et al (2015) Scikit-learn. *Get-Mobile Mob Comput Commun* 19:29–33. <https://doi.org/10.1145/2786984.2786995>
- Vergara-Diaz O, Vatter T, Vicente R et al (2020) Metabolome profiling supports the key role of the spike in wheat yield performance. *Cells* 9:1025. <https://doi.org/10.3390/cells9041025>
- Vergara-diaz O, Araus L (2020) Assessing durum wheat ear and leaf metabolomes in the field through hyperspectral data. *Plant J* 102:615–630. <https://doi.org/10.1111/tpj.14636>
- Vitrack-Tamam S, Holtzman L, Dagan R et al (2020) Random forest algorithm improves detection of physiological activity embedded within reflectance spectra using stomatal conductance as a test case. *Remote Sens* 12:2213. <https://doi.org/10.3390/rs12142213>
- Von Caemmerer S (2013) Steady-state models of photosynthesis. *Plant Cell Environ* 36:1617–1630. <https://doi.org/10.1111/pce.12098>
- Waskom M, Botvinnik O, O’Kane D (2017) mwaskom/seaborn: seaborn v0. 8.1, Zenodo
- Wofford MF, Boscoe BM, Borgman CL et al (2019) Jupyter notebooks as discovery mechanisms for open science: citation practices in the astronomy community. *Comput Sci Eng* 22:5–15
- Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 58:109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)
- Yan D, Wang Y, Murakami T et al (2015) *Auxenochlorella protothecoides* and *Prototheca wickerhamii* plastid genome sequences give insight into the origins of non-photosynthetic algae. *Sci Rep* 5:1–9. <https://doi.org/10.1038/srep14465>
- Yendrek CR, Tomaz T, Montes CM et al (2017) High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. *Plant Physiol* 173:614–626. <https://doi.org/10.1104/pp.16.01447>
- Zhu X-G, Long SP, Ort DR (2010) Improving photosynthetic efficiency for greater yield. *Annu Rev Plant Biol* 61:235–261. <https://doi.org/10.1146/annurev-arplant-042809-112206>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

CHAPTER 4

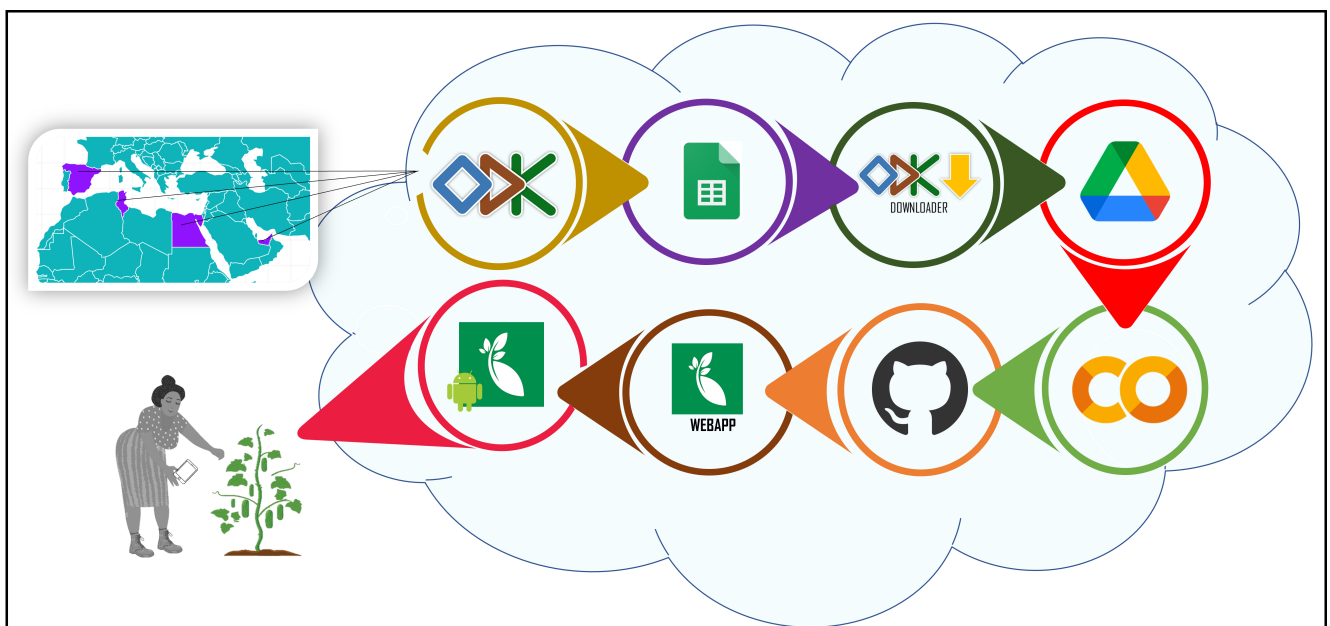


CHAPTER 4

Deep learning for detection of plant disorders on crop leaves: from data collection to framework tools

Ma. Luisa Buchailot, Jose A. Fernandez-Gallego, Henda Mahmoudi, Sumitha Thushar, Ghazi Al Jabri, Amna Abdulnoor Aljanaahi, M. Isabel Trillas, Alexi Akl, Zied Hammami, Jose Luis Araus, Shawn C. Kefauver

Submitted to:
Computers and Electronics
in Agriculture



Deep leaning for detection of plant disorders on crop leaves: from data collection to framework tools

Ma. Luisa Buchailot^{1&2}, Jose A. Fernandez-Gallego³, Henda Mahmoudi⁴, Sumitha Thushar⁴, Amna Abdulnoor Aljanaahi⁴, Ghazi Al Jabri⁴, Alexi Akl⁵, Zied Hammami⁴, M. Isabel Trillas¹, Jose Luis Araus^{1&2}, Shawn C. Kefauver^{1&2}

¹Section of Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain.

²AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198, Lleida, Spain, Lleida, Spain.

³Department of Electronics, University of Ibagué, Colombia.

⁴International Center for Biosaline Agriculture, Dubai, UAE.

⁵Full-stack software engineers, Lodz, Poland.

*Correspondence: Shawn C. Kefauver, sckefauver@ub.edu

Abstract

The second of the United Nations Sustainable Development Goals (SDG 2) highlights the importance of food security with the target of doubling the productivity and incomes of small shareholder farmers and food producers. In the Middle East and North Africa (MENA), horticultural crops are grown extensively year-round for local consumption, both outdoors and in protective shade houses, but due to climate change pests and diseases appear increasingly, and local farmers do not possess adequate knowledge to treat them, resulting in significant reductions in crop productivity. In accordance with SDG 2, there is a real demonstrable need for developing a user-friendly crop diagnosis mobile application in the MENA region that can also provide accurate decision support for more rapid and effective disease and pest control in order to reduce crop harvest losses. Meanwhile, the use of machine learning and deep learning approaches to process simple mobile phone RGB images has shown promise for automatic plant disease and pest diagnosis. In this study, we have developed a system for the big data collection that is required for accurately training deep learning models using a combination of open-source tools and proprietary software developments. Then, we developed deep learning models based on open-source deep learning architectures which can classify 21 different leaf disorders (diseases, pests, and nutritional deficiencies) in cucumber, tomato, pepper, and quinoa with an overall accuracy of 94%. Finally, we incorporated our deep learning model into a web app and an Android based mobile app, called Doctor X UB and Doctor X Nabat, respectively. Both applications include an integrated database with a decision support system that provides confirmatory symptomology and ecological, traditional/integrated pest management, and conventional treatment guidance in the most common three languages of the MENA region.

Keywords: SDG 2, MENA, Horticultural Crops, Data Collection, Deep Learning, Mobile App

1. Introduction

The Sustainable Development Goals (SDGs) were adopted by the member countries of the United Nations in 2015 as a universal call to "transform the world" (Brown, 2021). There are 17 goals, which cover the challenges that we are facing with regards to poverty, inequality, climate, environmental degradation, prosperity, peace, justice and their interactions (Lee et al., 2016). SDG number 2 was defined to address the importance of food security and nutrition; one of the principal targets by 2030 is to double the productivity and incomes of small-scale food producers (Ayoola et al., 2016). In addition, SDG number 17 promotes the development, transfer, dissemination and diffusion of environmentally sound technologies to developing countries on favorable terms (Humphreys et al., 2019). Sustainable agricultural development that aims to tackle SDG 2 in developing countries must incorporate the ideals set forth in SDG 17 in order to produce lasting benefits and results towards feeding the future.

In the African continent, the severe pre-and post-harvest losses due to the high incidence of pests, and diseases in crops (Otekunrin et al., 2019) is one of the major causes of hunger. Crop losses due to pests and diseases for major food and market crops were estimated between 20% and 40% at country and regional levels in the continent (Oerke, 2006). In the Middle East and North Africa (MENA) region, agriculture has many challenges; land and water are in short supply, and rainfed and irrigated lands are continually degrading due to unsustainable farming methods. In addition, as climate change makes the area hotter and drier, agriculture is more vulnerable to climate change. For that reason, crop production has been reduced by up to 30 to 35% of the potential productivity. On one hand, MENA region yields for horticultural crops, like tomatoes and cucumbers, are comparable to those of the rest of the world (FAO, 2018). It happened because many of these nations, including the United Arab Emirates, increased the practice of vegetables crop production in greenhouses. On the other hand, climate change has been observed to have a close relationship to the increase in prevalence of diseases in horticultural and crops; often, diseases are occurring for the first time in areas where they had not been observed before (Johannes et al., 2017). For example, local farm extension support offices who are not familiar with these new pests and diseases are unable to offer any support to small farmers. For that reason, crop losses are a major threat to the well-being of small farmers (rural families), local and national economy and stability, and to food security (Avelino et al., 2015; Savary and Willocquet, 2014).

In addition, these small farmers spend considerable amounts of money on chemical fertilizers and pest and disease management, sometimes without adequate technical support, resulting in poor control, pollution, and harmful effects. Because of that, the detection and diagnosis of plant stressors are urgently needed for the rapid and robust application of crop management in precision agriculture (Tanumihardjo et al., 2020). However, in the best scenario plant diagnosis is generally conducted through visual examination by experts, sometimes followed by a subsequent (metabolic or genetic) test applied as necessary; therefore this approach may result in expensive and time-consuming (Tani et al., 2018). Also, there is a risk for error due to a subjective perception even if the visual identification is pursued by experts (Dutot et al., 2013). Regarding this, different remote sensing (spectroscopic and imaging) techniques have been proposed to replace visual examination for detecting plant diseases. Nevertheless, they often require precise instruments and bulky sensors, which are expensive (Mahlein et al., 2013; Yuan et al., 2014).

On the other hand, due to the quick advancement of communication networks and the accessibility of several new remote, proximal, and touch sensors, new opportunities are opening up in the farming industry (Aqeel-Ur-Rehman et al., 2014; Wang et al., 2006). These technologies facilitate the low-cost geo-localized real-time transmission of information in the agricultural

context (Kumar and Ilango, 2018; Mahan and Yeater, 2008; Polo et al., 2015). A farm management information system, which primarily consists of a software system for gathering, processing, and storing data in the format needed to carry out farms' operations and activities, is one of the smart farming technologies (Fountas et al., 2015; Kaloxylou et al., 2012; Kitchen, 2008; Lewis, 1998). Another class of smart farming technology is the precision agriculture system, which is a farming management concept based on recording technologies to watch and measure crop field spatial and temporal variability with the purpose of increasing financial returns and minimizing environmental effects (Finger et al., 2019). With a focused and resource-conscious strategy, smart farming aims to produce agricultural products more effectively, qualitatively, and sustainably (Balafoutis et al., 2017). Smart farming should, however, generate value from the perspective of the low-income farmers by enabling more precise and faster decision-making as well as/or more effective exploitation operations and management at an affordable cost (Kerneck et al., 2020). Advances in decision-making and management are taking advantage not only of the increased capacities and decrease in cost of the information technologies and remote sensing capacities but also in the fast advance of artificial intelligence (AI), particularly of the avenues that machine learning is creating for a fast and efficient use of the huge amount of information generated and transmitted through the remote sensing and information techniques.

With respect to data and information processing, machine learning (ML) is the study of strategies and techniques for computational applications that can adjust or modify their activities to be more accurate. The process of machine learning consists of the following steps: data preparation and acquisition, selection of characteristics of interest, algorithm selection, parameters selection, training, and rating. (Marsland, 2011). ML algorithms can be divided into the following groups: reinforcement learning, evolutionary learning, unsupervised learning, and supervised learning (Patrício and Rieder, 2018). The unsupervised learning algorithms compare inputs to one another to find similarities and categorize them. Deep learning (DL) is one of these; it consists of numerous connected nodes, each with a weight that controls the signal and strength of the connection. Additionally, deep learning incorporates automatic parameter extraction during training and requires powerful processing capacity to develop an application model directly from image data, one of the major differences between machine learning and deep learning approaches (Russell and Norvig, 2010). Unsupervised Pretrained Networks, Convolutional Neural Networks (CNN), Recurrent Neural Networks, and Recursive Neural Networks are the four basic architectures employed, with CNN being one the most used. The main goal of a CNN is to learn data features via convolutional operations. Moreover, CNN architectures with many processing layers and neurons can efficiently perform high-complexity tasks such as image recognition by processing large-size data (Atila et al., 2021). As a result, this kind of network is better suited for identifying patterns in images. LeNet-5 was one of the initial CNN architectures that were suggested (LeCun et al., 1998). The nodes with the probabilities of the original image falling into one of the ten-digit classes are found in the final layer (Kirk and Wen-Mei, 2016). AlexNet (2012) (Krizhevsky et al., 2012), ZF Net (2013) (Zeiler and Fergus, 2014), VGG Net (2014) (Simonyan and Zisserman, 2015), GoogLeNet (2015) (Szegedy et al., 2015), and Microsoft ResNet (2015) (He et al., 2016) are a few examples of other network architectures that are described in the literature.

Regarding the popularization of the RGB digital cameras, there are many studies (Arivazhagan et al., 2013; Chuanlei et al., 2017; Islam et al., 2017; Qin et al., 2016; Rajan et al., 2017) about automatic plant disease diagnosis via ML and DL that have been widely applied as an alternative to the expensive sensors. The combination of RGB and ML or DL in disease detection and recognition has shown immense potential to address the problem of early and accurate detection of pest and diseases (Arnal Barbedo, 2019; Barbedo, 2018; Bierman et al., 2019; Brahimi et al., 2017; Cruz et al., 2017; DeChant et al., 2017; Dhakal and Shakya, 2018; Farjon et

al., 2020; Ferentinos, 2018; Johannes et al., 2017; Liu et al., 2018; Ozguven and Adem, 2019; Qin et al., 2016). In addition, there already exist a few plant disease detection mobile apps, some of which include multi-language support and offer a tutorial for users; for example Plantix, PlantVillage Nuru, MyPestGuide Disease, and Crop doctor (Siddiqua et al., 2022); however none of the above are focused specially on vegetable crops in the MENA region. Plant leaves are considered the first loci of infection for the germination of bacterial, and fungal capsules due to the suitable macro environments. Therefore, plant leaves are the most common part used to visually detect and identify diseases (Saikawa et al., 2019). These symptoms are often indicative of specific plant pathogens (e.g., fungi, bacteria, and viruses); insect feeding (e.g., sucking insect pests); or lack plant nutrition (e.g., lack of macro and micronutrients) (El-Helly et al., 2003). As vegetable crops show visible symptoms of diseases or pests, they suffer many different leaf-level patterns of symptoms, such as for example discolorations that often differ in color, shape, and size according to the causal agent, making them discernable using ML or DL image analyses.

As was mentioned before, SDG 17 is about strengthening the means of implementation and revitalizing global partnerships for sustainable development, and the target is to use multidisciplinary partnerships to improve agro-food systems. (Lee et al., 2016). In this study, the aim is to develop together between four countries, Spain, United Arab Emirates (UAB), Egypt, and Tunisia, a user-friendly application for smallholder farmers for the detection of plant diseases and nutritional disorders in tomatoes (*Solanum lycopersicum*), cucumber (*Cucumis sativus*), pepper (*Capsicum annuum*) and quinoa (*Chenopodium quinoa*). We present here a collaborative multi-actor approach that leverages mobile phone, local, and cloud-based supercomputing across multiple countries. As no database existed in order to develop the AI DL models for the detection of plant diseases and nutritional disorders in our target crops, we first designed a mobile phone to cloud data collection pipeline to collect image datasets and curate them. Once a sufficient image database was developed, preliminary AI models could be developed and tested for implementation in the final mobile phone and web application for automatic identification of our targeted plant diseases, pests, and nutritional disorders. A multi-language database was then integrated for actionable decision support following conventional, organic, and integrated pest management approaches.

2. Materials and Methods

2.1. Sequence of events

1. We first produced a data collection and storage pipeline using mobile phone RGB cameras and then held several training programs on disorder (diseases, pests and nutritional deficiencies) identification and data capture and upload. Then, we did multiple iterations of optimization of the data collection process - this was the creation of the first app of the project - the app for data collection improved through feedback to build the image database.
2. During the database curation process, the images were reviewed by a team of plant pathologists which evaluated each image batch more thoroughly thanks to the ODK Downloader.
3. We developed the DL model using the curated and labeled centralized image database and then we added new curated and labeled images of diseases and disorders to update the model as the project progressed. The whole data collection, curation, and labeling pipeline allowed for improved feedback and update capacities once in place.
4. We later launched both outreach and training activities for the improvement of the decision support system (DSS) plant disease, pest, and nutritional disorder database of the final mobile

application: particular focus was devoted to new ecological and IPM approaches in top of the conventional chemical treatment options, and new language support was added to include English, French and Arabic.

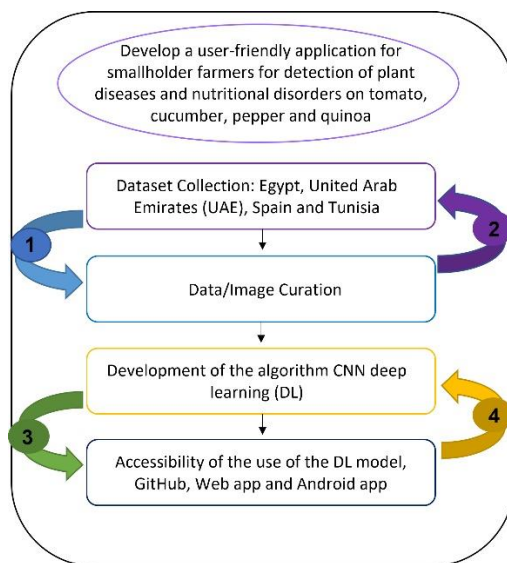


Figure 1. Experimental workflow of the process from the data collection to the Doctor X Nabat application. Feedback was critical in the refinement of the (1&2) data collection/image curation ODK app and also in the (3&4) development of the final user-friendly plant disease/pest/disorder detection web-based and Android mobile phone application.

2.2. Dataset collection

2.2.1 Site of study

The first part of this project was to collect the image dataset; in this case, images of the leaves of tomato, cucumber, and pepper showing symptoms of the targeted diseases were gathered in order to train the deep learning algorithm (Table 1). Deep learning requires big data for model training and development, and this is key to the overall success of the project. The data were collected from trained agriculture extension offices and technicians from four different countries including Spain, Tunisia, Egypt, and United Arab Emirates (UAE).

2.2.2 Dataset

To arrange all the separately gathered datasets from each country, we used “Open Data Kit” (ODK), an open-source application (<https://getodk.org/>, last accessed November 2022), which builds a data collection form or survey (built using XLSForm, for example) to collect the data and images on a mobile device and upload to an ODK Aggregate (<https://docs.getodk.org/aggregate-intro/>, last accessed November 2022) open-source server running on the Google Cloud Platform (<https://universitatbarcelonadoctorxlai.appspot.com/>). The aggregate server receives and organizes the collected survey and image data and can publish the organized data to other useful formats, such as Google Sheets, for further image data image processing. Regarding the forms, we built the survey forms to include relevant ancillary information such as location, species, type of stress, the potential cause of the stress, one photo of the whole plant, and most importantly five images of five different leaves of the plant with visible symptoms of the specific stress taken with a standardized background and zenithal (Figure 2). The information contained in the forms was used at the time to organize the images. Also, to capture the images in standardized conditions, we designed a background with a white balance panel and QR codes for scaling (for

instance; Figure 5). The use of ODK and cloud data storage servers for data collection across continents and cultures enabled the continuation of data collection efforts between all 4 collaborating countries despite limitations in mobility or funding. ODK provided the tools for a successful decentralized data collection with a centralized and controlled data curation process, all the while working entirely remotely.

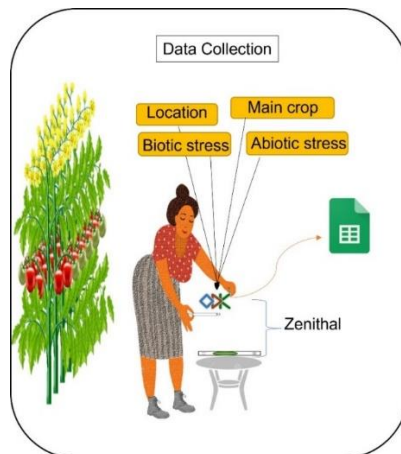


Figure 2. Visual explanation of important information that people added in the ODK form surveys, how they captured the images. After finishing the form and image data capture, they uploaded in the data to the ODK aggregate server.

2.2.3 Technical validation

We confirmed the identity of the diseases and other disorders by having an expert plant pathologist in each country identify and determine the diseases in the field. The experts worked in the greenhouses with postdocs and pre-doctoral students providing diagnoses determined based on standard approaches used by plant pathologists. Also, for some of the diseases, plant pathologists inoculated the crop directly using standard experimental approaches. Lastly, additional trained plant pathologists and technicians reviewed the images and forms during the data curation process.

2.3. Data/image curation

The second part of this project was to sort the image and survey dataset uploaded using the ODK application to the aggregate server. As we mentioned above, the ODK mobile app and form build a database collection (i.e., a survey); once the ODK survey form and pictures were uploaded, they were organized by the UB ODK Aggregate server and published to a shared Google Spreadsheet document. For downloading and sorting the spreadsheet database, we developed an open-source standalone JAVA software application for PC called the ODK Downloader (<https://gitlab.com/sckefauverl/odk-downloader>) to optimally organize our collective data (Figure 3). The ODK Downloader allows for fast and efficient management of our "big data" repository as well as quality control through improved organization and visualization. Besides, it allows the user to export, organize and download the data from the Google Spreadsheet database with adjustments for a multi-tiered file folder system (to organize images by crop type and then pest species, for example), selection of the form submissions to download (rows), the images to download (columns) and options for image file naming and labeling. After this part, we uploaded the dataset with the different species with diseases in each folder in a drive.

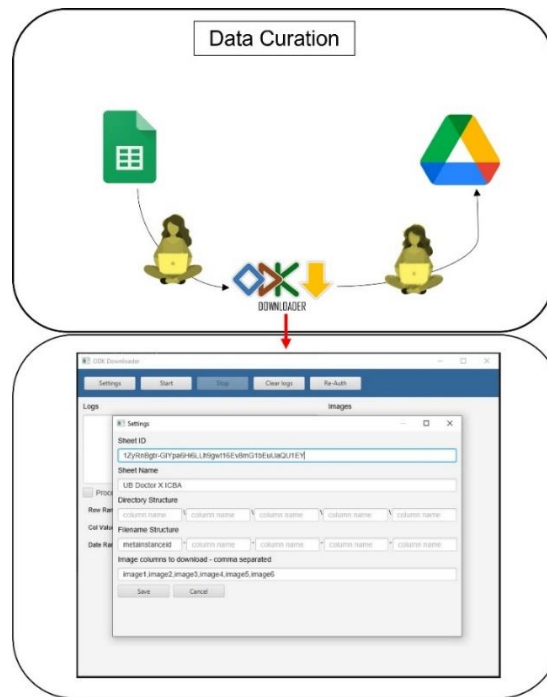


Figure 3. Data curation from the google sheet, down from the could to PC and sort with the ODK Downloader and after upload in the drive for the next step.

2.4. Development of the algorithm (Jose A) needs to describe all this part.

The third part of the project was the development of the DL model using a Convolutional Neural Network (CNN) architecture for object detection. We were first focused on the design a pipeline system (Fig. 4) from the database to classification in order ensure quick and efficient updates as more forms and image data future expansions.

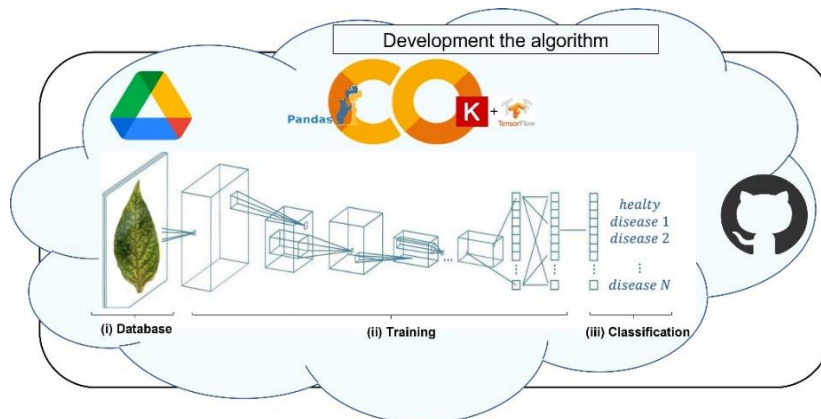


Figure 4. Pipeline system: from database (ODK) to disease and pest recognition (labels).

2.4.1 Database

Load and process images from Google Drive (ODK images); for this step, we used Google Colab as an open-source programming environment, and additionally, as a link between ODK images and data processing (Fig. 4).

2.4.2 Training

Features extraction and model parameters (weights) calculation; in this step, the convolutional layer, also called the feature map, is calculated and the neural network weights are updated. The algorithm codes are being written under Python programming language using open-source libraries, such as TensorFlow, Keras, Pandas, Numpy, and OpenCV in order to ensure the continued development and use of these algorithms in the future. For the training, we used specific stresses for a given plant species (Fig. 5), where image acquisition is more advanced, on the order of 200s of images per plant species and disease combination. The study investigated a deep learning architecture based on a recent convolutional neural network called ResNET50 to classify leaf healthy/disease/deficiency images. The model used 25 M of features, ResNet50 architecture (He et al., 2016), which won the ILSVRC-2015 competition in 2015, is an architecture proposed to solve the problem of multiple non-linear layers not learning identity maps and degradation problem. ResNet50 is a network in network architecture based on many stacked residual units. Residual units are used as building blocks to build the network. These units consist of convolution and pooling layers. This architecture uses 3×3 filters as VGG16 and takes input images of 224×224 pixels.

2.4.3 Classification

Prediction and object detection; this final step is used for discriminating between labels and evaluating the network performance. The system is being developed with the goal to detect as many diseases or pests as the leaf-image is affected.

2.4.4. Model Accuracy Assessments

The performance of our model in predicting the affirmative class is described by the precision ratio. By dividing the total number of true positives by the sum of true positives and false positives, it is determined:

$$\text{Precision} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalsePositives}} \quad (1)$$

The recall ratio is determined by dividing the quantity of true positives by the total of true positives and false negatives, as shown below:

$$\text{Recall} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}} \quad (2)$$

A weighted average of recall and precision is used to determine the F1-score ratio as follows:

$$\text{F1 - score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

2.5. Accessibility of the use of the DL model

2.5.1 GitHub

Once we had the algorithm training and with a high score of accuracy of detection of diseases in the different crops. We added in the GitHub: <https://github.com/sckefauver/ml-icba>.

2.5.2 Web-based Application

Once the algorithms were considered full functional for a minimal selection of crop species and pests, we developed a portal web app pages to support the different database languages <https://doctorxub.com/icbafr>, <https://doctorxub.com/icbaar>, <https://doctorxub.com/icba>,

accessible with any PC, tablet, Windows, Ubuntu, and Apple software functionality with connection to internet service, where images may be uploaded and crop diseases, pests or deficiencies identified. This part was where we integrated the extensive knowledge base of the ICBA research team concerning best practices for the treatment of these crop diseases, pests, or deficiencies across MENA.

2.5.3 Android Mobile Application

Lastly, the fully functional and integrated crop diseases, pests or deficiency deep learning algorithms and MENA best practices knowledge base of the web app testing environment was ported to a native Android (mobile phone) application format so that the whole system can be tested and launched for use offline with any Android phone in the field (<https://play.google.com/store/apps/details?id=com.doctorxnabat&hl=en&gl=US&pli=1>).

3. Results and Discussion

3.1 Dataset collection & curation

Using the ODK app, we gathered a total of 20,718 pictures of the 4 distinct plant species from Spain, the UAE, Egypt, and Tunisia for this study. Table 1 shows the visual symptoms in leaves caused by the different 14 disorders we have included to date in the App. These include nutritional disorders, in this case, potassium and phosphorous deficiencies, and salinity stress. Moreover, different fungi, bacteria, viruses, and pests responsible for other visible symptoms on leaves were included. Figure 5 shows example leaves of the species that have diseases or nutritional issues.

Table 1. The visual symptoms of the leaf in response to biotic and abiotic stressors

Stress Type	Agent	Nutritional Deficiencies/Disease	Plant Affected	Visual Symptoms
Abiotic	-	Potassium deficiency	Cucumber	Leaves have sunken main veins and are bronzed and discolored, turning yellowish green at the margins. Additionally, the margins dry out while the veins continue to be green. Later, interveinal chlorosis worsens and spreads into the heart of the leaf and is then followed by necrosis.
Abiotic	-	Phosphorous deficiency	Tomato	The most typical sign is leaf purpling, especially in the leaf veins. Clearly visible purpling on the petiole, stem, and undersides of the leaves. The undersides of the leaves on older plants turn reddish-purple, while the leaves on young plants turn purplish. leaves that are older and have some necrotic patches. The tendency for leaves to turn blue gray under severely deficient conditions. Plants grow slowly and are stunted or dwarfed.
Abiotic	-	Salinity	Tomato	Plants affected by high salt concentrations often appear dark green in the early stages, but rapidly develop marginal yellowing and necrosis of older leaves.
Biotic	Fungus	Downy mildew (<i>Pseudoperonosoa cubensis</i>)	Cucumber and Pepper	The spots start out yellow and angular on the upper side, then they turn brown and sometimes into sizable brown patches that damage the leaves. The dots on the underside of the leaf blade are brown and covered in sporangia, giving the leaf blade a purplish-brown hue. Old leaves are the first targets of attacks, then fresh leaves.

Biotic	Fungus	Sooty mold (<i>Capnodium citri</i>)	Pepper	A blackish powdery material, resembling black soot, coats the leaves. The first develops on leaves, which persist for the entire leaf's lifespan. The second variety, which persists on the stems and twigs of woody plants, is replenished using leftovers from the fungus' production from the previous season. Although it's not hazardous to the plant, if it's present in sufficient amounts, it can hinder photosynthesis in the leaves, suffocating plants until they wither and die.
Biotic	Fungus	Powdery mildew (<i>Oidiopsis taurica</i>)	Tomato	Leaf with circular to angular, pale green to yellow dots on the upper side that are coated in a white mycelium and conidial down. The impacted areas eventually turn brown, develop a necrotic center, deteriorate, and tear readily. Attacked leaves become brittle and may even fold up their blades.
Biotic	Fungus	Early blight (<i>Alternaria solani</i>)	Tomato	On leaves, 4 to 7 mm-diameter black necrotic patches with a yellow halo surround them as they grow in concentric circles. We may observe elliptical brown, gray, and concentric dots on the stems.
Biotic	Fungus	Leaf mold (<i>Passalora fulva</i>)	Tomato	First to be infected are the oldest leaves. On the upper sides of leaves, little, pale greenish-yellow dots with no discernible edges, usually less than 1/4 inch across, develop. Below the leaf spots, an olive-green to brown velvety mold develops on the bottom leaf surface. Leaf splotches coalesce and darken as a group. Although they wither and die, leaves frequently stay affixed to the plant. Infected blooms lose their color and turn black.
Biotic	Bacteria	Wilt (<i>Erwinia tracheiphila</i> or <i>Curtobacterium flaccumfaciens</i>)	Cucumber	Initially dull green in color leaves wilt during the day before recovering at night. Additionally, the edges turn brown and yellow and entirely wither and die. Every crop wilt at a different rate. Cucumbers quickly wilt and perish.
Biotic	Virus	Tomato Yellow Leaf Curl Virus (<i>Begomovirus spp.</i>)	Pepper, Tomato and Quinoa	Small leaves exhibit interveinal and marginal yellowing as well as significant crumpling. The margins develop yellow, leave upward-curling curls, and become stunted. severe necrosis and blighting of the leaves. In seedlings, the shoots shorten and then become stunted, and the plant grows upright.
Biotic	Pest	Spider mite (<i>Tetranychus urticae</i>)	Cucumber and Pepper	Leaves elongate, thicken, and darken. On the underside of the leaf, regions of corky brown color can be seen between the primary veins. Young vegetation is distorted. When the spider mite's population density is high, it is possible to see slowed growth and shoot dieback.
Biotic	Pest	Leaf hopper (<i>Coccinella septempunctata</i>)	Pepper	An excess of leafhoppers can cause leaves and shoot tips to turn yellow, then brown, curl, and eventually die. leaves to get light flecks on them. Leafhoppers also exude honeydew, which supports the growth of a blackish-sooty mold. This can contaminate the fruit, foliage, and surfaces that are under infected plants. Stunted and deformed plant growth.

Biotic	Pest	Aphid (<i>Aphis fabae</i>)	Pepper	Aphids secrete honeydew, a sugary liquid waste that, in high enough concentrations, can inhibit photosynthesis in the leaves and cause stunted or delayed plant growth. Honeydew supports the formation of a dark-colored fungus that appears on honeydew secretions (sooty mold). Additionally, they weaken plants in high numbers by suckling up the sap they consume and causing leaf deformation, twisting, and curling, as well as yellowing and withering.
Biotic	Pest	Leaf miner (<i>Tuta absoluta</i>)	Tomato	Completely eaten leaves. Additionally, reduced photosynthesis and production.

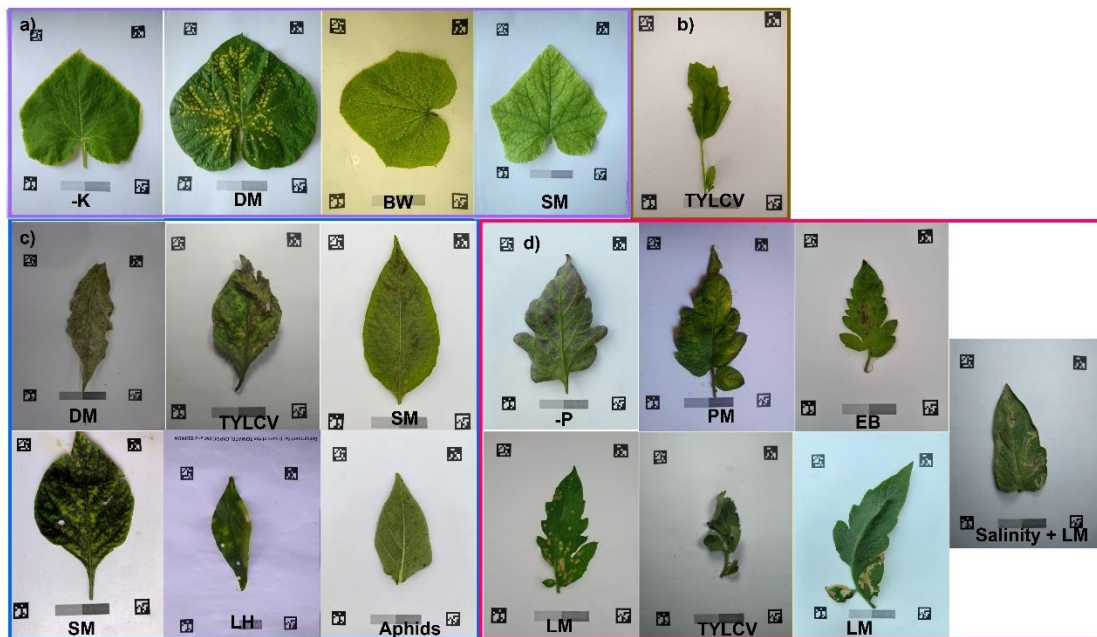


Figure 5. Examples of images of the leaves from different crops with diseases, pests, or nutritional disorders. **a)** Cucumber leaves with Potassium deficiency (-K), Downy Mildew (DM), Bacteria Wilt (BW) and Spider Mite (SM). **b)** Quinoa leaves with Tomato Yellow Leaf Curl Virus (TYLCV). **c)** Pepper leaves with Downy Mildew (DM), Tomato Yellow Leaf Curl Virus (TYLCV), Sooty Mold (SM), Spider Mite (SM), Leaf Hopper (LH) and aphids. **d)** Tomato leaves with Phosphorus deficiency (-P), Powdery Mildew (PM), Early blight (EB), Leaf Mold (LM), Tomato Yellow Leaf Curl Virus (TYLCV), Leaf Miner (LM) and Salinity and Leaf Miner (LM).

We collected healthy leaves of the three horticultural species: cucumber, pepper, and tomato leaves, two classes of leaves with nutrition deficiencies, and sixteen classes of leaves with diseases caused by fungi, bacteria, viruses, and pests. The images in the dataset had different quality and sizes because they were taken from different types of mobile phone cameras by different people. Additional photos of a few complicated combinations of disorders were also gathered and kept for a potential future upgrade to the models and the app. In comparison, there are only two main datasets with trusted labels online: Plant Village (PV) (Hughes and Salathe, 2015) and Digipathos (Barbedo et al., 2018). Many studies of deep learning are based on the PV data set (Atila et al., 2021; Ferentinos, 2018; Kehs et al., 2019; Mohanty et al., 2016), which is currently the largest dataset containing 54305 images. Nevertheless, the dataset presented in this study was collected in a constrained environment with a homogeneous

background and may offer some advantages. While the Digipathos dataset addresses the above problem, also the number of images is too small with only 2326 images.

3.2 Development of the algorithm

The dataset was divided into three parts: training, validation, and testing. Table 2 shows the number of images used in the three phases across the 21 classes in 4 different crop species. The number of the images in each phase is determined based on the fine-tuned hyperparameters and structure of the CNN model. In Table 2 we can see the total number of the images; the lower number is sooty mold in pepper with 197 images and the higher is the powdery mildew on tomato with 2940 images.

Table 2. Database details: the species, disorder, type of disease or deficiency, number of Images used in the training, validation, and testing phases across the all the healthy, diseases and deficiencies classes that we collect a cross the two years.

Species	Disorder	Type	Train Set Count	Validation Set Count	Test Set Count	Total Number of Images
Cucumber	Healthy	-	140	46	47	233
	Nutrient Deficiency	Potassium	597	199	199	995
	Fungus	Downy Mildew	183	61	61	305
	Bacteria	Wilt	171	57	57	285
	Pest	Spider Mite	600	200	200	1000
Pepper	Healthy	-	189	62	63	314
	Fungus	Downy Mildew	699	233	233	1165
		Sooty mold	119	39	38	197
	Virus	TYLCV	135	45	45	225
	Pest	Spider Mite	608	202	202	1012
		Leaf Hopper	168	55	56	279
Aphid		675	224	225	1124	
Tomato	Healthy	-	610	203	183	1016
	Nutrient Deficiency	Phosphorous	1229	409	410	2048
	Fungus	Powdery Mildew	1764	588	494	2940
		Early Blight	339	113	113	565
		Leaf Mold	669	223	223	1115
	Virus	TYLCV	1475	491	491	2457
	Pest	Leaf Miner	1425	475	475	2375
	Abiotic + Biotic	Salinity + Leaf Miner	171	56	57	284
Quinoa	Virus	TYLCV	471	156	157	784
Total			12.437	4137	4144	20.718

The confusion matrix for the CNN model in Figure 6 gives a detailed analysis of how the model performance changes for different disease classes. The matrix rows the actual (true) classes, and the columns correspond to the predicted classes. The diagonal cells show the proportion of the

correct predictions of our CNN model, whereas the of diagonal cells illustrate the error rate of our model.

	Cucumber Healthy	Cucumber Potassium Deficiency	Cucumber Downie Mildew	Cucumber Wilt	Cucumber Spider Mites	Pepper Healthy	Pepper Downie Mildew	Pepper Yellow Leaf Curl Virus	Pepper Sooty Mold	Pepper Spider Mites	Pepper Leaf Hopper	Pepper Aphids	Tomato Healthy	Tomato Phosphorous Deficiency	Tomato Powdery Mildew	Tomato Early Blight	Tomato Leaf Mold	Tomato Yellow Leaf Curl Virus	Tomato Leaf Miner	Tomato Salinity + Leaf Miner	Quinoa Yellow Leaf Curl Virus
Cucumber Healthy	42	0	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cucumber Potassium Deficiency	0	192	1	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cucumber Downie Mildew	0	0	61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cucumber Wilt	0	0	0	57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cucumber Spider Mites	0	0	0	1	199	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pepper Healthy	0	0	0	0	0	60	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0
Pepper Downie Mildew	0	0	0	0	0	0	210	0	1	0	0	6	0	0	2	0	10	2	0	0	2
Pepper Yellow Leaf Curl Virus	0	0	0	0	0	0	0	41	0	0	2	0	0	0	0	0	0	0	0	2	0
Pepper Sooty Mold	0	0	0	0	0	0	0	0	37	0	0	1	0	0	0	0	0	0	0	0	0
Pepper Spider Mites	0	0	0	0	0	0	0	0	0	202	0	0	0	0	0	0	0	0	0	0	0
Pepper Leaf Hopper	0	0	0	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0	0	0	0
Pepper Aphids	0	0	0	0	0	0	0	0	0	0	0	220	0	0	0	0	0	0	0	0	0
Tomato Healthy	0	0	0	0	0	0	0	0	0	0	0	3	178	0	0	0	2	0	0	0	0
Tomato Phosphorous Deficiency	0	0	0	0	0	0	0	0	1	0	0	2	0	393	0	0	0	10	4	0	0
Tomato Powdery Mildew	0	0	0	3	0	0	0	0	0	0	1	1	0	2	447	2	3	6	25	0	5
Tomato Early Blight	0	0	0	0	0	0	0	0	0	0	5	7	0	0	88	12	0	1	0	0	
Tomato Leaf Mold	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	211	0	0	0	0	
Tomato Yellow Leaf Curl Virus	0	0	0	7	0	0	0	1	0	0	1	14	0	27	2	0	0	415	15	6	3
Tomato Leaf Miner	0	0	0	0	0	0	0	0	0	0	7	0	0	1	0	6	0	457	0	4	
Tomato Salinity + Leaf Miner	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	57	0	
Quinoa Yellow Leaf Curl Virus	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	1	0	153

Figure 6. The Confusion matrix for the CNN Model for 21 classes.

In this study, the confusion matrix demonstrates in most cases, that our model can differentiate between the healthy/disease/deficiency classes and achieve high levels of prediction accuracy (Fig. 6). Nevertheless, the model in some cases algorithms confuses the diseases of each crop such as the algorithm confuses the cucumber healthy with twice times wilt and 3 times spider mites.

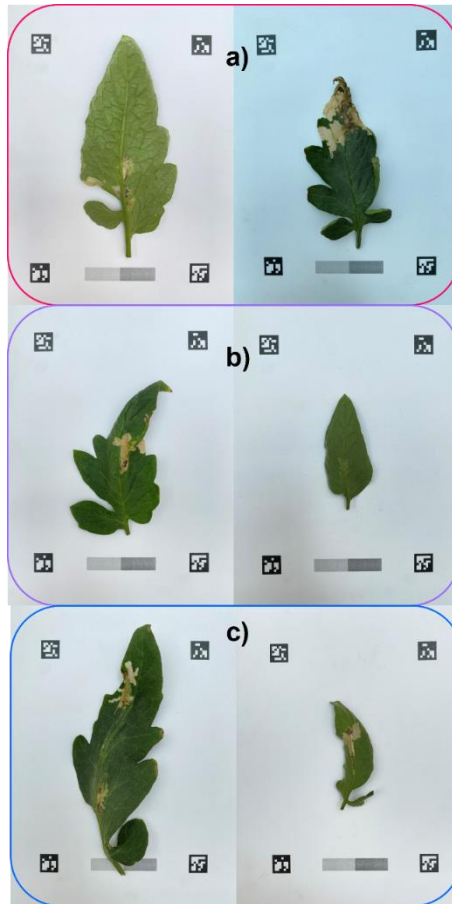


Figure 7. Examples of tomato leaves with Leaf Miner. a) In left leaf with brighter lighting and in the right with slightly darker lighting. b) In the left, older leaf with higher impact the pest and in the right younger leaf with a relatively low impact of pest. c) In the left an older and bigger leaf and on the right a smaller leaf with a similar level of pest impact.

Figure 7 illustrates three distinct scenarios that the model application may encounter during classification. Figure 7a) shows the variations in light that can affect an image because of the collecting of data from many individuals in various locations, light conditions and cameras around the world. Additionally, Fig. 7b) demonstrates that the disease may be in an early or late stage. Finally, we can see how the size of the leaves differs in Fig. 7c). This range of variation was included in the model training on purpose as part of the data collection planning process in order to capture the full range of conditions that might need to be identified for each specific pest, disease or deficiency.

Table 3. The Precision vs. Recall vs. f1-score values of the CNN Model for all the healthy, disease, pest, and deficiencies classes.

Species	Type	Precision	Recall	f1-score	Support
Cucumber	Healthy	1.00	0.89	0.94	47
	Potassium	1.00	0.96	0.98	199
	Downy Mildew	0.98	1.00	0.99	61
	Wilt	0.80	1.00	0.89	57
	Spider Mites	0.96	0.99	0.98	200
Pepper	Healthy	0.98	0.95	0.97	63
	Downy Mildew	1.00	0.90	0.95	233

	TYLCV	0.98	0.91	0.94	45
	Sooty mold	0.84	0.97	0.90	38
	Spider Mites	1.00	1.00	1.00	202
	Leaf Hopper	0.90	0.98	0.94	56
	Aphids	0.84	0.98	0.91	225
Tomato	Healthy	0.90	0.97	0.94	183
	Phosphorous	0.93	0.96	0.94	410
	Powdery Mildew	0.99	0.90	0.95	494
	Early Blight	0.98	0.78	0.87	113
	Leaf Mold	0.86	0.95	0.90	223
	TYLCV	0.96	0.85	0.90	491
	Leaf Miner	0.91	0.96	0.93	475
	Salinity + Leaf Miner	0.88	1.00	0.93	57
Quinoa	TYLCV	0.92	0.97	0.94	157
Accuracy				0.94	4029
Macro avg		0.93	0.95	0.94	4029
Weighted avg		0.94	0.94	0.94	4029

Table 3 shows the accuracy, recall, and F1-score ratios for our CNN model using various probability thresholds and reflect the trade-off between the true-positive rate and the positive predictive values. Precision shows the positive predictive value of our model, the recall measures (i.e. how many true positives are accurately detected), while the F1-score considers the number of false positives and false negatives.

We confirmed that our model showed good diagnostic performance not only for single disease and deficiency cases but also for multiple infection cases. The precision for each class ranged from 80 to 100%, the recall varies between 78 to 100% and the f1-score was from 87 to 100% (Table 3). These data could be compared with the confusion matrix, in which the algorithm confuses the classes, and with the cases in Fig. 6. The model development, RedNet50 using 25 M features, can classify 21 classes with an average accuracy of 94%. This means our model can recognize 15 single diseases, caused by fungi, viruses, bacteria, and pests, also 1 class with multiple disorders, caused by a combination of salinity stress and leaf miner, besides 2 nutritional deficiencies and 3 healthy leaves of tomato, pepper, and cucumber.

There are many studies where that used different models to detect/identify/recognize diseases/deficiencies in plants. Joshi et al., (2021) proposed the CNN layers model for the detection de viral infection with a maximum accuracy of 98.27%, maximum precision of 98.42%, maximum recall of 96.54%, and maximum f1-score of 97.37%. Another study developed an identification system for mildew disease in pearl millet using a model VGG16 CNN where the maximum accuracy was 95%, the maximum precision was 90.50%, the maximum recall was 94.50%, and the maximum f1-score of 91.75% (Coulibaly et al., 2019). For recognition of plants diseases using the dataset create for Plant Village (Hughes and Salathe, 2015), Sladojevic et al., (2016) used the model Caffe deep learning and achieved an accuracy of 96.3%. Other study used a deep CNN and they presented an accuracy of 93.4% (Ma et al., 2018)

3.3. Accessibility of the use of the DL model

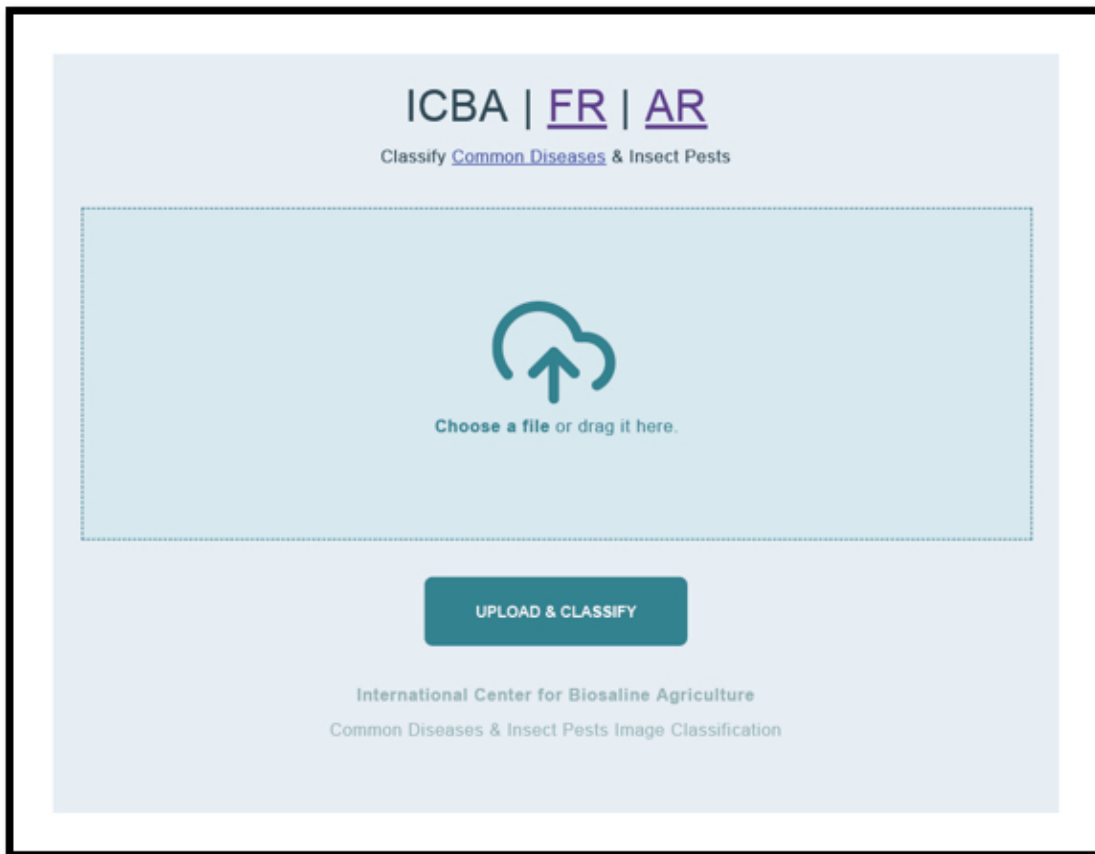


Figure 8. Image of the Doctor X UB web app, that is in three different languages: English, French and Arabic.

Regarding Figure 8, we show the front end of the web app that we released before the Android app. As part of the project multi-actor feedback approach, we first did a trial web app with 8 common diseases and nutritional disorders. Then, we conducted training and feedback meetings in Egypt, Tunisia, and UAE with researchers, government extension officers, and smallholder farmers, and they give us the suggestions to add in more languages to the app, in this case, we added French and Arabic to English. In addition, we collected more images subsequently and soon thereafter expanded the model to include 21 different classes in the web app with a fully expanded database for the decision support systems (DSS) in all three languages.

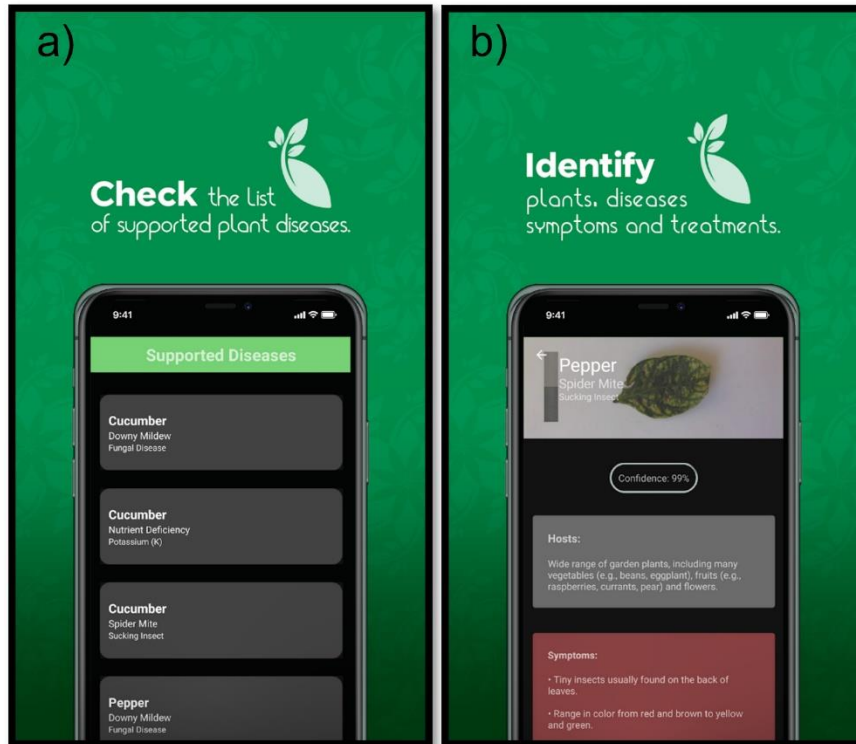


Figure 9. Images of the Doctor X Nabat Android app, that is in three different languages English, French and Arabic. a) The list of the supported plant diseases/deficiencies. b) Detection the pepper leaf with spider mite in this example was with 99% of confidence.

Once we developed the web app, we built the Android application. Figure 9 shows two images of the Doctor X Nabat on Google Play. The app is free and publicly developed and supported application that can be used as a digital assistant to help MENA farmers diagnose major horticultural diseases and other disorders through the acquisition of images with smartphones. The Android application currently uses the same cloud-based image processing server for applying the model, but it may work eventually without an internet connection. When online it gives feedback in real-time. Further, the presented high accuracy for the target disorder is more than 94% for the entire set of plants disorders encompassing the app. Based on recommendations from ICBA, the app has been specifically designed to meet the demands of the MENA region. It supports Arabic, French, and English, the three most important languages.

In contrast to other previously released apps, we used a bottom-up, multi-actor stakeholder-driven development method to make sure the app satisfies the unique requirements of our end customers in MENA. The end users have directly contributed to its establishment, fostering trust in its usefulness and functionality. On the other hand, Plantix is currently the most complete app, in addition to being free. However, it is worth to mention that it has been developed for the past 5-7 years, involving a larger multidisciplinary team fully devoted to build up and maintain the app, which means the budgetary support behind has been much larger. Moreover, apps such as Plant Village Nuru or MyPestGuide Diseases focus on a more limited number of diseases and other disorders for other regions of the world. Some apps are not fully free to use or include paid services (Plant Disease Identifier, Agrio). Other apps may have a different focus, and, instead of identifying disorders, the app performs quantitative assessments for plant diseases that have been a priority identified (Leaf Doctor) or have focused on three species mostly for the USA environment (Purdue Tree Doctor). Some others also depend on access to

multispectral or hyperspectral images captured by a drone (Plant Scope), or provide only general information upon request, rather than identifying disorders directly with a built in DSS (Crop Doctor; MyPestGuide Reporter).

4. Conclusions and future work

The ODK data collection demonstrated that multiple countries' data may be gathered using a reliable open-source system. More than 20.000 images of quinoa, cucumber, pepper, and tomato leaves in good health, with a wide range of disease or pest foliar symptoms, or in need of improved nourishment were gathered. As an added bonus, we developed the ODK Downloader, a PC based tool that facilitated the arrangement of the complete ODK dataset online by scrapping details out of a Google Spreadsheet shared document and may be adapted to support other similar ODK databases. The CNN DL algorithm developed for classifying the classes was successful in achieving its 94% accuracy goal while covering 21 different plant-symptom combinations. Additionally, the creation of a user-friendly, local language-supporting web and Android apps makes cutting-edge technologies more available to MENA countries and both of these were launched with an integrated DSS. For future work, we may consider expanding this app to cover more diseases, pests, and nutritional disorders, and also the combinations of different diseases, as this has been indicated as another common complication in real world field conditions. This paper illustrates the experience of building up an affordable albeit efficient mobile phone- hosted app for automatic identification of plant disorders. The app represents an example of successful participatory endeavor, embracing the collaborative work of end users (i.e., farmers) together with international institution with a global mandate to improve MENA agriculture and the academia.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

5. References

- Aqeel-Ur-Rehman, Abbasi, A.Z., Islam, N., Shaikh, Z.A., 2014. A review of wireless sensors and networks' applications in agriculture. *Comput. Stand. Interfaces* 36, 263–270. <https://doi.org/10.1016/j.csi.2011.03.004>
- Arivazhagan, S., Shebiah, R.N., Ananthi, S., Vishnu Varthini, S., 2013. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agric. Eng. Int. CIGR J.* 15, 211–217.
- Arnal Barbedo, J.G., 2019. Plant disease identification from individual lesions and spots using deep learning. *Biosyst. Eng.* 180, 96–107. <https://doi.org/10.1016/j.biosystemseng.2019.02.002>
- Atila, Ü., Uçar, M., Akyol, K., Uçar, E., 2021. Plant leaf disease classification using EfficientNet deep learning model. *Ecol. Inform.* 61, 101182. <https://doi.org/10.1016/j.ecoinf.2020.101182>
- Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A.J., Morales, C., 2015. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Secur.* 7, 303–321. <https://doi.org/10.1007/s12571-015-0446-9>

- Ayoola, J., Ayoola;G., Okike;, I., Dashiell;, K., Ogbodo, J., 2016. A Policy Situation Analysis for Achieving the SDG2 (Zero Hunger) Targets In Selected States of Nigeria. 30th Int. Conf. Agric. Econ. 1, 1–21.
- Balafoutis, A., Beck, B., Fountas, S., Vangeyte, J., Van Der Wal, T., Soto, I., Gómez-Barbero, M., Barnes, A., Eory, V., 2017. Precision agriculture technologies positively contributing to ghg emissions mitigation, farm productivity and economics. *Sustain.* 9, 1–28. <https://doi.org/10.3390/su9081339>
- Barbedo, J.G.A., 2018. Factors influencing the use of deep learning for plant disease recognition. *Biosyst. Eng.* 172, 84–91. <https://doi.org/10.1016/j.biosystemseng.2018.05.013>
- Barbedo, J.G.A., Koenigkan, L.V., Halfeld-Vieira, B.A., Costa, R.V., Nechet, K.L., Godoy, C.V., Junior, M.L., Patricio, F.R.A., Talamini, V., Chitarra, L.G., Oliveira, S.A.S., Ishida, A.K.N., Fernandes, J.M.C., Santos, T.T., Cavalcanti, F.R., Terao, D., Angelotti, F., 2018. Annotated plant pathology databases for image-based detection and recognition of diseases. *IEEE Lat. Am. Trans.* 16, 1749–1757. <https://doi.org/10.1109/TLA.2018.8444395>
- Bierman, A., LaPlumm, T., Cadle-Davidson, L., Gadoury, D., Martinez, D., Sapkota, S., Rea, M., 2019. A high-throughput phenotyping system using machine vision to quantify severity of grapevine powdery mildew. *Plant Phenomics* 2019. <https://doi.org/10.34133/2019/9209727>
- Brahimi, M., Boukhalifa, K., Moussaoui, A., 2017. Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. *Appl. Artif. Intell.* 31, 299–315. <https://doi.org/10.1080/08839514.2017.1315516>
- Brown, M.E., 2021. Metrics to accelerate private sector investment in sustainable development goal 2—zero hunger. *Sustain.* 13, 4–9. <https://doi.org/10.3390/su13115967>
- Chuanlei, Z., Shanwen, Z., Jucheng, Y., Yancui, S., Jia, C., 2017. Apple leaf disease identification using genetic algorithm and correlation based feature selection method. *Int. J. Agric. Biol. Eng.* 10, 74–83. <https://doi.org/10.3965/j.ijabe.20171002.2166>
- Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., Traore, D., 2019. Deep neural networks with transfer learning in millet crop images. *Comput. Ind.* 108, 115–120. <https://doi.org/10.1016/j.compind.2019.02.003>
- Cruz, A.C., Luvisi, A., De Bellis, L., Ampatzidis, Y., 2017. X-FIDO: An effective application for detecting olive quick decline syndrome with deep learning and data fusion. *Front. Plant Sci.* 8, 1–12. <https://doi.org/10.3389/fpls.2017.01741>
- DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E.L., Yosinski, J., Gore, M.A., Nelson, R.J., Lipson, H., 2017. Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology* 107, 1426–1432. <https://doi.org/10.1094/PHYTO-11-16-0417-R>
- Dhakal, A., Shakya, S., 2018. Image-Based Plant Disease Detection with Deep Learning. *Int. J. Comput. Trends Technol.* 61, 26–29. <https://doi.org/10.14445/22312803/ijctt-v61p105>
- Dutot, M., Nelson, L.M., Tyson, R.C., 2013. Predicting the spread of postharvest disease in stored fruit, with application to apples. *Postharvest Biol. Technol.* 85, 45–56. <https://doi.org/10.1016/j.postharvbio.2013.04.003>
- El-Helly, M., Rafea, A.A., El-Gammal, S., 2003. An Integrated Image Processing System for Leaf Disease Detection and Diagnosis., in: *IICAI*. pp. 1182–1195.

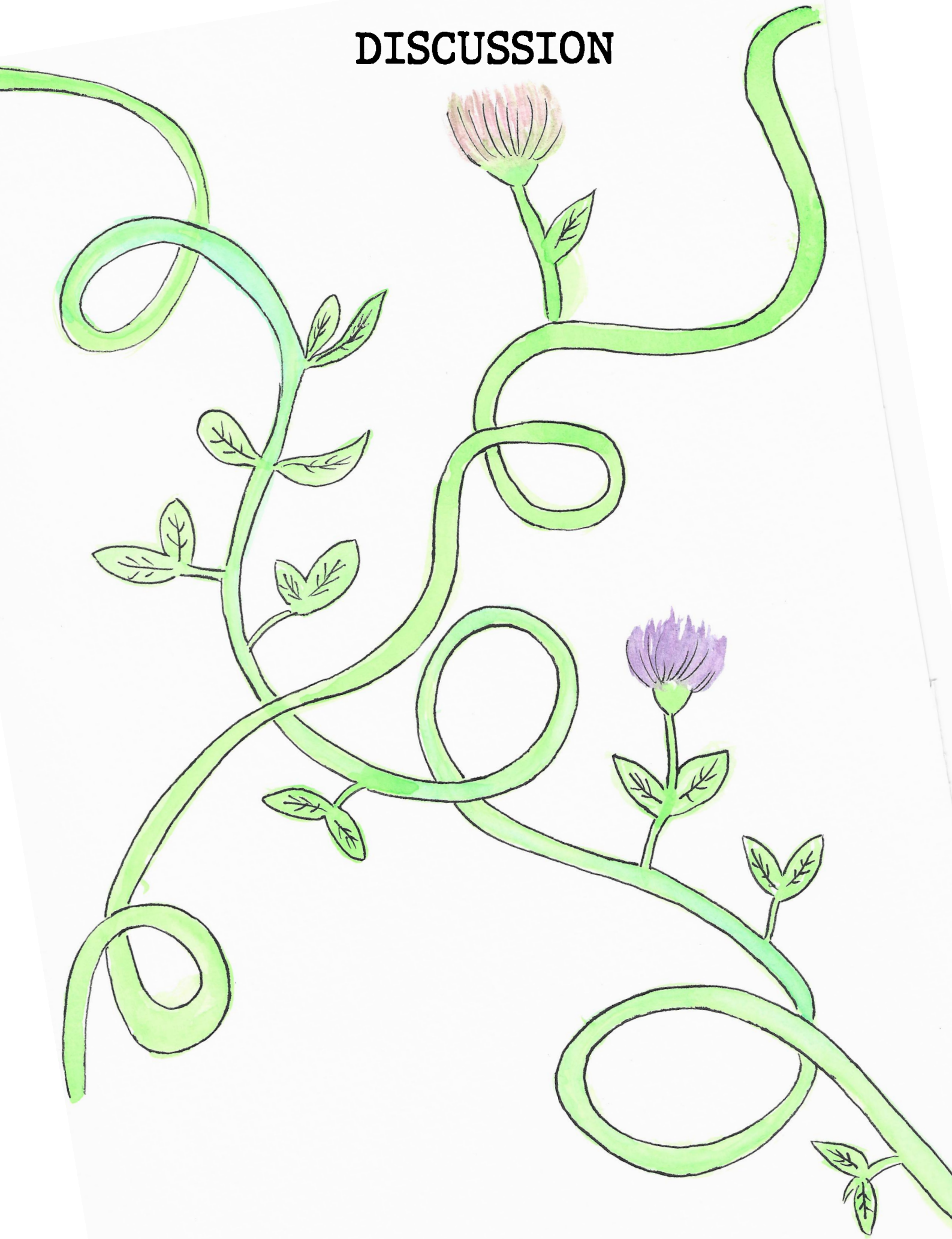
- FAO, 2018. The Middle East and North Africa: Prospects and challenges, in: OECD-FAO AGRICULTURAL OUTLOOK 2018-2027. pp. 67–107. https://doi.org/10.1787/agr_outlook-2018-5-en
- Farjon, G., Krikeb, O., Hillel, A.B., Alchanatis, V., 2020. Detection and counting of flowers on apple trees for better chemical thinning decisions. *Precis. Agric.* 21, 503–521. <https://doi.org/10.1007/s11119-019-09679-1>
- Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* 145, 311–318. <https://doi.org/10.1016/j.compag.2018.01.009>
- Finger, R., Swinton, S.M., El Benni, N., Walter, A., 2019. Precision Farming at the Nexus of Agricultural Production and the Environment. *Annu. Rev. Resour. Econ.* 11, 313–335. <https://doi.org/10.1146/annurev-resource-100518-093929>
- Fountas, S., Carli, G., Sørensen, C.G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebenson, J., Tisserye, B., 2015. Farm management information systems: Current situation and future perspectives. *Comput. Electron. Agric.* 115, 40–50. <https://doi.org/10.1016/j.compag.2015.05.011>
- He, K., Zhang, X., Shaoqing, R., Sun, J., 2016. Deep Residual Learning for Image Recognition. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* 770–778.
- Hughes, D.P., Salathe, M., 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics.
- Humphreys, D., Singer, B., Lawlor, K., Smith, R., Budds, J., Gabay, M., Bhagwat, S., de Jong, W., Newing, H., Cross, C., Satyal, P., 2019. SDG 17: Partnerships for the goals-Focus on forest finance and partnerships, Sustainable Development Goals: Their Impacts on Forests and People. <https://doi.org/10.1017/9781108765015.019>
- Islam, M., Dinh, A., Wahid, K., Bhowmik, P., 2017. Detection of potato diseases using image segmentation and multiclass support vector machine. *Can. Conf. Electr. Comput. Eng.* 8–11. <https://doi.org/10.1109/CCECE.2017.7946594>
- Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A.D., Ortiz-Barredo, A., 2017. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Comput. Electron. Agric.* 138, 200–209. <https://doi.org/10.1016/j.compag.2017.04.013>
- Joshi, R.C., Kaushik, M., Dutta, M.K., Srivastava, A., Choudhary, N., 2021. VirLeafNet: Automatic analysis and viral disease diagnosis using deep-learning in Vigna mungo plant. *Ecol. Inform.* 61, 101197. <https://doi.org/10.1016/j.ecoinf.2020.101197>
- Kaloxylou, A., Eigenmann, R., Teye, F., Politopoulou, Z., Wolfert, S., Shrank, C., Dillinger, M., Lampropoulou, I., Antoniou, E., Pesonen, L., Nicole, H., Thomas, F., Alonistioti, N., Kormentzas, G., 2012. Farm management systems and the Future Internet era. *Comput. Electron. Agric.* 89, 130–144. <https://doi.org/10.1016/j.compag.2012.09.002>
- Kehe, A., McCloskey, P., Chelal, J., Morr, D., Amakove, S., Mayieka, J., Ntango, G., Nyongesa, K., Pamba, L., Mugo, J., Tsuma, M., Onyango, W., Hughes, D., 2019. From village to globe : A dynamic real-time map of African fields through PlantVillage.
- Kernecker, M., Knierim, A., Wurbs, A., Kraus, T., Borges, F., 2020. Experience versus expectation: farmers' perceptions of smart farming technologies for cropping systems across Europe. *Precis. Agric.* 21, 34–50. <https://doi.org/10.1007/s11119-019-09651-z>

- Kirk, D.B., Wen-Mei, W.H., 2016. Programming massively parallel processors: a hands-on approach. Morgan kaufmann.
- Kitchen, N.R., 2008. Emerging technologies for real-time and integrated agriculture decisions. *Comput. Electron. Agric.* 61, 1–3. <https://doi.org/10.1016/j.compag.2007.06.007>
- Krizhevsky, B.A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* 60, 84–90.
- Kumar, S.A., Ilango, P., 2018. The Impact of Wireless Sensor Network in the Field of Precision Agriculture: A Review. *Wirel. Pers. Commun.* 98, 685–698. <https://doi.org/10.1007/s11277-017-4890-z>
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2323. <https://doi.org/10.1109/5.726791>
- Lee, B.X., Kjaerulf, F., Turner, S., Cohen, L., Donnelly, P.D., Muggah, R., Davis, R., Realini, A., Kieselbach, B., MacGregor, L.S., Waller, I., Gordon, R., Moloney-Kitts, M., Lee, G., Gilligan, J., 2016. Transforming Our World: Implementing the 2030 Agenda Through Sustainable Development Goal Indicators. *J. Public Health Policy* 37, S13–S31. <https://doi.org/10.1057/s41271-016-0002-7>
- Lewis, T., 1998. Evolution of farm management information systems. *Comput. Electron. Agric.* 19, 233–248. [https://doi.org/10.1016/S0168-1699\(97\)00040-9](https://doi.org/10.1016/S0168-1699(97)00040-9)
- Liu, B., Zhang, Y., He, D.J., Li, Y., 2018. Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry (Basel)*. 10. <https://doi.org/10.3390/sym10010011>
- Ma, J., Du, K., Zheng, F., Zhang, L., Gong, Z., Sun, Z., 2018. A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Comput. Electron. Agric.* 154, 18–24. <https://doi.org/10.1016/j.compag.2018.08.048>
- Mahan, J.R., Yeater, K.M., 2008. Agricultural applications of a low-cost infrared thermometer. *Comput. Electron. Agric.* 64, 262–267. <https://doi.org/10.1016/j.compag.2008.05.017>
- Mahlein, A.K., Rumpf, T., Welke, P., Dehne, H.W., Plümer, L., Steiner, U., Oerke, E.C., 2013. Development of spectral indices for detecting and identifying plant diseases. *Remote Sens. Environ.* 128, 21–30. <https://doi.org/10.1016/j.rse.2012.09.019>
- Marsland, S., 2011. Machine learning: an algorithmic perspective. Chapman and Hall/CRC.
- Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7, 1–10. <https://doi.org/10.3389/fpls.2016.01419>
- Oerke, E.C., 2006. Crop losses to pests. *J. Agric. Sci.* 144, 31–43. <https://doi.org/10.1017/S0021859605005708>
- Otekunrin, Olutosin A., Otekunrin, Oluwaseun A., Momoh, S., Ayinde, I.A., 2019. How far has Africa gone in achieving the zero hunger target? Evidence from Nigeria. *Glob. Food Sec.* 22, 1–12. <https://doi.org/10.1016/j.gfs.2019.08.001>
- Ozguven, M.M., Adem, K., 2019. Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. *Phys. A Stat. Mech. its Appl.* 535, 122537. <https://doi.org/10.1016/j.physa.2019.122537>
- Patrício, D.I., Rieder, R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Comput. Electron. Agric.* 153, 69–81.

- <https://doi.org/10.1016/j.compag.2018.08.001>
- Polo, J., Hornero, G., Duijneveld, C., García, A., Casas, O., 2015. Design of a low-cost Wireless Sensor Network with UAV mobile node for agricultural applications. *Comput. Electron. Agric.* 119, 19–32. <https://doi.org/10.1016/j.compag.2015.09.024>
- Qin, F., Liu, D., Sun, B., Ruan, L., Ma, Z., Wang, H., 2016. Identification of alfalfa leaf diseases using image recognition technology. *PLoS One* 11, 1–26. <https://doi.org/10.1371/journal.pone.0168274>
- Rajan, P., Radhakrishnan, B., Padma Suresh, L., 2017. Detection and classification of pests from crop images using Support Vector Machine. *Proc. IEEE Int. Conf. Emerg. Technol. Trends Comput. Commun. Electr. Eng. ICETT 2016*. <https://doi.org/10.1109/ICETT.2016.7873750>
- Russell, S., Norvig, P., 2010. *Intelligence artificielle: Avec plus de 500 exercices*. Pearson Education France.
- Saikawa, T., Cap, Q.H., Kagiwada, S., Uga, H., Iyatomi, H., 2019. AOP: An Anti-overfitting Pretreatment for Practical Image-based Plant Diagnosis. *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019* 5177–5182. <https://doi.org/10.1109/BigData47090.2019.9006567>
- Savary, S., Willocquet, L., 2014. Simulation Modeling in Botanical Epidemiology and Crop Loss Analysis. *Plant Heal. Instr.* <https://doi.org/10.1094/phi-a-2014-0314-01>
- Siddiqua, A., Kabir, M.A., Ferdous, T., Ali, I.B., Weston, L.A., 2022. Evaluating Plant Disease Detection Mobile Applications: Quality and Limitations. *Agronomy* 12. <https://doi.org/10.3390/agronomy12081869>
- Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* 1–14.
- Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Comput. Intell. Neurosci.* 2016. <https://doi.org/10.1155/2016/3289801>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 07-12-June, 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
- Tani, H., Kotani, R., Kagiwada, S., Uga, H., Iyatomi, H., 2018. Diagnosis of Multiple Cucumber Infections with Convolutional Neural Networks. *Proc. - Appl. Imag. Pattern Recognit. Work. 2018-October*, 13–16. <https://doi.org/10.1109/AIPR.2018.8707385>
- Tanumihardjo, S.A., McCulley, L., Roh, R., Lopez-Ridaura, S., Palacios-Rojas, N., Gunaratna, N.S., 2020. Maize agro-food systems to ensure food and nutrition security in reference to the Sustainable Development Goals. *Glob. Food Sec.* 25, 100327. <https://doi.org/10.1016/j.gfs.2019.100327>
- Wang, N., Zhang, N., Wang, M., 2006. Wireless sensors in agriculture and food industry - Recent development and future perspective. *Comput. Electron. Agric.* 50, 1–14. <https://doi.org/10.1016/j.compag.2005.09.003>
- Yuan, L., Huang, Y., Loraamm, R.W., Nie, C., Wang, J., Zhang, J., 2014. Spectral analysis of winter wheat leaves for detection and differentiation of diseases and insects. *F. Crop. Res.* 156, 199–207. <https://doi.org/10.1016/j.fcr.2013.11.012>
- Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks. *Lect.*

Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)
8689 LNCS, 818–833. https://doi.org/10.1007/978-3-319-10590-1_53

DISCUSSION



DISCUSSION

The Sustainable Development Goals (SDGs) of the United Nations (UN)– targets aimed to be achieved by 2030 – are focused on both developing and developed countries. The “Zero Hunger” SDG is number two, focused on doubling the productivity and incomes of small-scale food producers, ensuring sustainable food production, and implementing resilient agricultural practices. Agriculture in the world is highly dependent on climate change conditions, such as abiotic stress which includes nutrient deficiency in soils, intensifying rising temperatures, drought, and increasing CO₂ concentrations. Also, biotic stress increases can also be linked to climate change and can have large impacts on agricultural sustainability, including invasive pests, disease outbreaks, and decreases in crop productivity. This doctoral Thesis represents the practical implementation of non-destructive methods using remote sensing (RS) technologies applied at different scales and using different types of sensors, which all provide insights into nutrient deficiency, pest invasions, estimation of photosynthesis parameters, and detection of crop diseases and other plant disorders. Besides, this dissertation furthermore provides a comparison of maturity levels between different types of technologies through technological readiness levels (TRL) assessments along the four chapters.

Food Security and the second Sustainable Development Goal (SDG 2) “Zero Hunger”

With respect to food security in the world, developing countries are more vulnerable to climate change, and depend more on local agricultural production. We developed the campaigns/science of 3 of the 4 results chapters of this dissertation in developing countries. In Fig. 5 we can see where are the countries that each chapter are involved in each project.

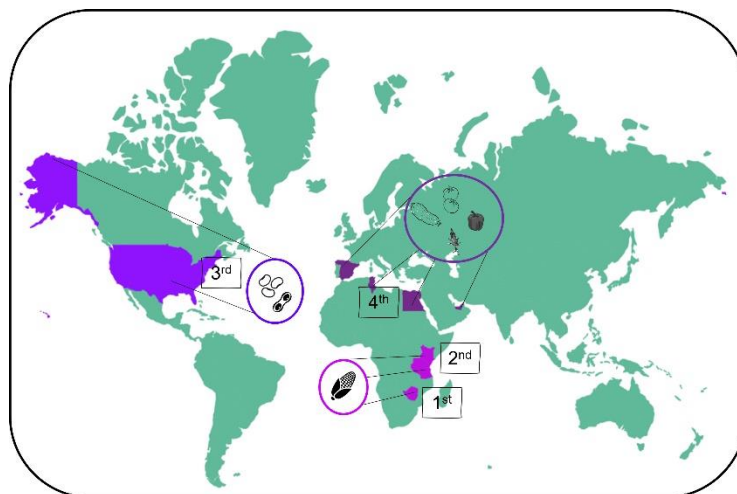


Fig. 5. World map indicates the different countries where research included in this thesis was development. The 1st chapter field campaign was conducted in Zimbabwe and focused on maize. The 2nd chapter was carried out in Zimbabwe, Tanzania, and Kenya using remote sensing and also focused on maize. The 3rd chapter was completed during an exchange in United States and focused on soybean and peanut. Finally, the 4th chapter was carried out in Spain, Tunisia, Egypt, and United Arab Emirates (UAE).

The first and second chapters were conducted primarily in Sub-Saharan Africa (SSA) and focused on the maize crop under abiotic (1st) and biotic (2nd) stress. In SSA there are 37 million ha of maize produced annually, of which 95% of the farms, belong to small farmers (less than 2 ha, FAOSTAT 2016). After water, the most important input for maize production is nitrogen, and in SSA fertilizer applications are often negligible, and this has become one of the most important abiotic stresses that directly affects the production capacity of small holder farmers (Masuka et al. 2012). On the other hand, in 2016 the Fall Armyworm (FAW) arrived and spread fast across Africa; one of the most damaging pests ever to invade the African continent, and one that feeds preferentially on maize and is still causing billions US\$ of losses annually in maize yield in SSA (Day et al. 2017; Hailu et al. 2018). These two important abiotic and biotic stresses continue to threaten the livelihoods of millions of small farmers in SSA.

The third chapter was developed in the United States, in collaboration with the Auburn University in Alabama. There, advanced technologies were made available for a more in-depth study of the potential remote sensing capacities under field conditions using advanced portable spectroscopy equipment. The study focused on soybean, which is the main source of proteins at the World level, minerals for humans and animals. Second species evaluated was peanut, a main crop in the southeast of the United States and a major source of proteins and lipids not only in the USA but also in many developing countries of africa (Onuegbu and Ibeabuchi 2021)

The fourth chapter was developed between four countries: Spain, Tunisia, Egypt, and the United Arab Emirates, and focused on the principal vegetable crops eaten in the Middle East and North Africa (MENA) region, plus an herbaceous crop introduced during the last decades in the region. We based the study on tomato, pepper, cucumber, and quinoa. In the region of MENA, the control of diseases and pests in common horticulture has been a challenge for small farmers (Zolfaghari and Farzaneh 2021). Rapid response in the treatment of crops is often critical in order to ensure sustainable productivity in adverse conditions. Thus, we focused here on developing tools for the quick identification of pests, diseases, and common nutrient deficiencies in these four crops.

The following summary of the four chapters is focused on each as an effort to contribute to the UN SDG 2 “Zero Hunger,” the goal of which, as mentioned before, is to double the productivity and incomes of small scales producers, and thus this dissertation has attempted to contribute to through international cooperation and applying remote sensing technologies.

Remote Sensing in agriculture

In this part, we compared the different types of resolution: spectral, spatial, and temporal and the observation scales to determine the monitoring capacities of the different chapters.

Spectral resolution

The spectral resolution refers to the wavelength width and the number of the different frequency bands. Concerning the VIS (visible light), it ranges from 400-700 nm, it refers to the images acquired by the conventional RGB (Red, Green, Blue) cameras. Vegetation Indices (VIs) derived from conventional digital RGB images have been proposed as means of estimating green biomass and grain yield (GY) of crops under stress conditions (Casadesus et al. 2007; Vergara-Diaz et al. 2015). In the first chapter, by taking photos of the maize canopy under low nitrogen, we used RGB VIs, extracted from the photos, for yield estimation. In addition, RGB digital cameras have been used for automatic disease diagnosis via machine learning (ML) and deep learning (DL). One study proved the potential of the proposed algorithms, where five different cucumber leaf diseases are considered and classified with an accuracy of 98.08% in just 10.52 seconds using DL (Khan et al. 2020). In the fourth chapter, we developed an algorithm that can identify 21 different plant disorders on leaves of cucumber, tomato, pepper, and quinoa with an overall classification accuracy of 94%, all with just RGB observations limited to the VIS spectral region.

On the other hand, the VNIR (VIS+NIR, near infrared) reflectance (400-1000nm) covers a larger spectral range and is also very common approach for phenotyping. Within the VNIR region, the Normalized Difference Vegetation Index (NDVI) is an classical standard in remote sensing and a very powerful vegetation signal. This index is one of the most common remote sensing indices, that provides a simplified yet reliable estimate of green biomass (Tucker 1979; Van Hoek et al. 2016). Similarly, this may be also compared to the Green Area (GA) index from the VIS range, which quantify the portion of green pixels to the total pixels of the image and can provide a reliable estimation of vegetation cover (Lukina et al. 1999). In chapter one we compared the RGB VIs mentioned above with NDVI, and the data showed that the estimation of the grain yield of the maize was higher through RGB VIs than with NDVI. One reason could be that the data of the NDVI was saturated, and they do not present differences significative between the values.

Regarding the second chapter, we used these two indices that mention before NDVI and GGA, and in addition the Leaf Area Index (LAI). That is the area of leaves over a corresponding area on the ground (Area of leaves/square meter on the ground (Van Gardingen et al. 1999). The LAI values were calculated from images captured with a mobile camera RGB (Demarez et al. 2008; Garrigues et al. 2008). We analyzed these three separate indices at various scales, and the results showed that despite variations in observation scale and picture quality, there was a coefficient of determination higher than $R^2 = 0.60$ between each category. That suggests that even using different of range of the spectrum, we can estimate the biomass of the plant.

Regarding the VIS-NIR-SWIR (SWIR, short wave infrared) full spectral range (400-2500 nm), one may acquire the full leaf optical reflectance spectra, with which advanced regression models like the Partial Least Squares Regression (PLSR) may be applied to estimate more complex traits as the maximum rate of rubisco-catalyzed carboxylation ($V_{c,max}$) and maximum electron transport rate supporting RuBP regeneration (J_{max}) in different species/or cultivars (Doughty et al. 2011; Serbin et al. 2012; Choquette et al. 2019) under different abiotic stresses such as drought (Silva-Perez et al. 2018) or elevated atmospheric ozone concentrations (Ainsworth et al. 2014; Yendrek et al. 2017b). In the third chapter, we estimated the $V_{c,max}$ and J_{max} of two legume crops (soybean and peanut) using full range leaf level reflectance spectra with four different advance regression models: PLSR, Automatic relevance determination regression (ARDR), Bayesian ridge (BR), and least absolute shrinkage and selection operator (Lasso).

Spatial resolution

The spatial resolution is the resolution of the sensor in terms of area on the ground, and this depends on the sensor and the remote sensing platform that you use to monitor your crop. In the first and second chapters, this concept of spatial resolution was considered relevant.

In the first chapter we compared the RGB VIs taken from ground level and unmanned aerial vehicle (UAV) platform level. The Pearson correlation between similar indices acquired at ground and aerial levels was higher than $r = 0.75$ between the relevant indices such as GA, and Green Greener Area (GGA). We did this comparison to observe if the resolution of the pixels from the cameras can affect the results and if we could replace the RGB photos from ground level for RGB photos at UAV levels due to the difference in the pixels. One similar study showed an estimation of grain yield of maize through RGB images from a UAV with $R^2 = 0.82$ (Gracia-Romero et al. 2017). Additionally, one study showed a good correlation of grassland yield using RGB VIs taken from UAV with $R^2 = 0.62$ (Lussem et al. 2018). Regarding the second chapter, we

conducted an observation at three different scales. In analyzing correlations of NDVI from microsatellites (3m/pixel resolution), NDVI from phenopoles at 5 m (0.06m/pixel), GA from phenopoles at 5 m (0.009 m/pixel), and LAI from the fisheye hemispherical lens were taken from below the plant and looking upwards (0.001m/pixel). We did these comparisons between the different technologies and sensors validation because often the satellite data was not available for our monitoring purposes due to the cloud cover issues. These parameters were considered similar to total green biomass assessments as we mention above. There are many studies that compare these indices at different scales, and most were with an R^2 higher than 0.70 (Kovacs et al. 2004; Matese et al. 2015; Ihuoma et al. 2021). This could suggest that the spatial resolution does not affect the results with data from different scales.

Concerning of third chapter, we did a simulation of the bands of the Sentinel 2 a+b using the data from the ASD FieldSpec VIS-NIR-SWIR data to see the limitation of the single-leaf-level estimations of photosynthetic capacities using point-based spectral analysis. For that reason, we resampled FieldSpec data to cover the 12 of the 13 bands of Sentinel 2 a+b to the estimation of the $V_{c,max}$ and J_{max} , and the results showed a $R^2 = 0.50$.

In the fourth chapter, all the data that we used was very high resolution and were captured at ground level using the RGB cameras of mobile phones for the estimation of leaf diseases/deficiencies through DL. The images also had a range of different resolutions because these images were taken by different persons and different types of mobile phones as part of the international collaboration.

Temporal resolution

The temporal resolution is considered a measure of the frequency at which the sensor revisits the same area in the study.

In the first chapter, this temporal resolution was an interesting factor because we only took the data with UAV at an early stage of the maize season and the effects of the under nitrogen mostly started to appear after. This study suggests that if a research plan has only one shot to capture data, it is better to study first what is the most suited crop growth stage and then estimate when the key growth stage would be to assess the abiotic or biotic stress.

In the second chapter, we applied a time-series anomaly change detection and first derivative growth pattern analyses of NDVI using Sentinel 2 a+b with a temporal resolution fixed by 5 days and Planet Scope with a temporal resolution of 1 day during the maize vegetative growth stage to observe a reduction of the NDVI caused by FAW. One of the biggest problems of this study

was cloud cover, because it reduced the data frequency and was a major limitation of using a passive remote sensing approach.

For the third chapter, as we mentioned in the spatial resolution considerations, we did a simulation using the bands of the Sentinel 2 a+b, which exhibited a determination coefficient higher than 0.5. In this case, using a spectrophotometer sensor was flexible and we could adjust the timing to capture data. We compared this to using an open-source satellite, having a defined temporal frequency, as they are in orbit. Satellites capture data frequently, but sometimes the data is not useful for agriculture, such as if there are clouds present. Also, measurements at a fixed temporal resolution may not fit the actual crop or stressor temporal resolution for a specific purpose. Regarding the fourth chapter, we also used a hand-held platform to capture the images, a mobile phone provided with a RGB camera, and the temporal resolution was not as relevant, though very flexible for a quick assessment of the crop disease and timely decision support.

Observation scales

Regarding remote sensing platforms in this thesis, we can find the three types: ground, aerial, and space level along of the four chapters

In the first chapter, both ground-based and airborne sensors were utilized. We examined RGB photos collected from 50 meters away and from one meter away from crop and discovered a strong association between them. When we considered the information discerned from the RGB data from the UAV and processing of 200 plots, it would suggest that deploying the UAV could result in significant savings in terms of time, requiring about 10 minutes in the field, 20 minutes to mosaic, and 30 minutes to extract and process the data. The field part would roughly quadruple time spent if the data were gathered from the ground-level for image capture and processing. Due to advancements in digital photography, including image processing, low-altitude aerial imagery is now feasible and an affordable monitoring tool for agriculture.

We employed three separate platform categories for the second chapter. Before moving to the field, we first used the proof-of-concept approach, which simply uses data from Sentinel 2, such the frequency of image capture and percentage of clouds. Then, we traveled to other fields and employed RGB cameras on a UAV, but this information was not displayed due to the poor quality of the data and the fact that UAV flight was prohibited in one country. We then used as alternative a 5-meter pole equipped with an RGB and multispectral sensor to replace the images cultured from sensors placed at UAV. Finally, we calculated the LAI and attempted to validate the satellite data using a cell phone with a fisheye lens.

In the third chapter, we observed at the single leaf level and performed resampling using this data to recreate the Sentinel 2 bands, which revealed a greater than 50% estimation of the photosynthetic parameters. This demonstrated the strong potential of imaging spectroscopy that is airborne or spaceborne, like the NASA AVIRIS and HypIRI programs (Mariotto et al. 2013; Chapman et al. 2019). Additionally, the Soil-Canopy Observation of Photosynthesis and Energy (SCOPE) models were used to invert hyperspectral data (Camino et al. 2019).

In the final chapter, we used a low-level platform—a mobile camera—to collect leaf images for the purpose of identifying illnesses and deficiencies. Many studies have employed unmanned aerial vehicles (UAVs) to acquire pictures, some of which do so at low altitudes. For example, one study uses a UAV at a height of 6 meters to capture photos of maize with Northern Leaf Blight (Stewart et al. 2019). Additionally, another study deployed a UAV at a height of 2 meters to automatically detect illnesses in soybean leaves (Tetila et al. 2020).

Technological Readiness Levels

As was mentioned before, the TRL indicates its maturity level and ranges from TRL 1 (basic principles observed) to TRL 9 (actual system proven in operational environment). Based on this classification, TRL was specified for the four chapters presented a long of the dissertation.

The first chapter, we have assessed as TRL 2-TRL 3 (application formulated-proof of concept), denotes that the application was carried out as planned, and the study was successful in demonstrating why the best-performing quartile of maize varieties should be chosen. However, the control treatment for the study was planted in a separate year. Additionally, at the time when we took the data from RGB cameras at observation levels the maize crop under lower nitrogen was in an early growth stage and much of the low nitrogen symptoms appeared after that date, as evidenced by other ground measurements. In order to advance the TRL further, we recommend capturing the imagery at a slightly later growth stage when the visual symptoms start to have appeared.

We have an example of a TRL 3 proof-of-concept in the second chapter, as it was considered inadequate and unspecific even though we saw a decline in the NDVI curve, which could have been brought on by intercropping, weeds, abiotic, or other biotic factors. Moreover, it could be a bit of TRL 4 because we have an experimental pilot, because we used the UAV, but the data was not very good due to the pilot's lack of knowledge on how do flights for remote sensing data capture. Further work would be needed could include the ESA Sentinel 1 SAR and possibly have more results and data regarding the crop canopy to advance a TRL 4-5

In the third chapter, estimated at TRL 4, we ran an experimental pilot and showed that we can estimate the $V_{c,max}$, and J_{max} of soybean and peanut with 70% accuracy by utilizing leaf spectral reflectance and advance regression models. Furthermore, we recommend a specific suggestion for increasing the TRL. In this case, scaling the leaf to a canopy level and then moving on to the demonstration pilot (TRL 5).

The fourth chapter demonstrated a good system design that makes it scalable, extremely accurate, and user-friendly. Through this work, we consider that the concept moved from TRL 4 to TRL 8. There are four processes in this system: data collecting, data/image curation, development of the algorithm CNN DL, and accessibility of the use of the DL model GitHub, web app and android app. This project will continue to conduct experimental, demonstration, and industrial pilots for the new additional diseases or inadequacies. On the other hand, the mobile app Doctor X Nabat, which will debut in December of 2022, has a 94% accuracy rating for the categorization and recognition of 21 healthy/disease/deficiency conditions, making it a TRL 8 (initial release version).

CONCLUSIONS



CONCLUSIONS

The evaluation of the practical applications of non-destructive techniques utilizing remote sensing (RS) technology has been used in different countries at various scales and using various types of sensors. However, it is crucial to consider the study goals when applying these approaches to agriculture because the specific goals of the study determine the kinds of spectral, spatial, and temporal resolutions considered optimal for each project.

Regarding the evaluation of the TRLs in the four chapters of this dissertation, it may be argued that the level of technology depends on the goals and techniques of the study, the time invested in developing the research, the volume of data collected, and the statistical and validation approaches employed.

Chapter 1.

As with the SPAD and NDVI sensors (ground level and VIS-NIR), the RGB (VIS) sensors from the ground and from the aerial level can be regarded as functional technology.

In order to maximize the benefits of utilizing RGB sensors at the ground and aerial levels in support of plant phenotyping, the data-collecting process must be structured and timed appropriately.

We evaluated the efficiency of these technologies to estimate grain yield under low nitrogen. The TRL was identified as between 2 and 3, which could be because the RGB (VIS) sensors data was not taken at the optimal moment when the symptoms of low nitrogen appeared in the leaves.

Chapter 2.

The most promising findings were shown by the Sentinel 2 a+b and Planet Scope NDVI-time-series first derivative analysis, but only in a few areas of study, and they have yet to be completely integrated with suitable spatial and false positive anomaly filters. Although intercropping may be possible, it is advisable to keep the analyses to the vegetative development stage and predominantly maize as the major crop. Actually, our study focused on this species, crop management and crop stage conditions. In fact, intercropping and weeds are considered sources of variability that may affect remote sensing studies of agriculture. Also, cloud cover is a significant barrier to adopting RS methods and can decrease the amount of data that is useful.

We evaluated the efficiency of several technologies to develop an RS system for monitoring and early warning systems of an invasive pest. The TRL was assessed between number 3 and 4, due to the NDVI-time-series first derivative analysis that could show the FAW was eating the maize. Nevertheless, the signal could be distorted by other environmental factors.

Chapter 3.

The photosynthetic parameters of soybean and peanut together were successfully estimated using data of leaf spectral reflectance and various advanced regression models with determination coefficients ranging between 0.50 and 0.70.

To achieve higher coefficients of determination and model performance, this research showed that it is more crucial to have a larger range of phenotypic variation than any significant effect of a treatment or cultivar in order to achieve greater coefficients of determination and model performance. In contrast to other advanced regression models, we showed that PLSR was the most effective model for estimating photosynthetic parameters (BR, ARDR, and LASSO).

We evaluated that through leaf spectral reflectance we can estimate the $V_{c,max}$ and J_{max} of soybeans and peanuts. The TRL was estimated at 4, because we demonstrate the technology that we estimated a 70% of the $V_{c,max}$ and J_{max} . To increase the level of the TRL, the objective could be to estimate the J_{max} or $V_{c,max}$ of all varieties of legumes in a non-controlled environment.

Chapter 4.

The ODK data collection showed that a robust open-source technology could be used to collect data from different countries. More than 20,000 photos of healthy, diseased, and nutritionally deficient leaves from quinoa, cucumber, pepper, and tomato were gathered.

Additionally, we created the ODK Downloader, a program that made it easier to organize the entire ODK dataset and did a great job of organizing and curating the data.

Furthermore, an artificial intelligence approach, through deploying deep learning, was created for the categorization of the different classes of disorders, achieving an overall 94% accuracy.

For MENA nations, the development of a user-friendly app and with local language support should make advanced technologies more accessible and thus increase their TRL.

For chapter four, we estimated the TRL as between 4 and 8, as this project was developed over almost 3 years and we improved our research process several times with feedback between the different steps such as data collection, data curation, and development of the algorithm until producing the final mobile app.

RESUMEN DE LA TESIS DOCTORAL

La seguridad alimentaria personal significa que tienen acceso físico y económico a alimentos suficientes, seguros y de calidad. Por otro lado, existen tres causas principales de la inseguridad alimentaria: 1. la alta vulnerabilidad de la agricultura al cambio climático; 2. conflictos locales, nacionales o internacionales; y 3. desigualdad económica. Los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas (ONU), que son objetivos que deben alcanzarse para 2030, están dirigidos tanto a las naciones en desarrollo como a las desarrolladas. El segundo ODS, "Hambre cero", tiene como objetivo duplicar la productividad y los ingresos de los productores de alimentos a pequeña escala, al tiempo que promueve métodos agrícolas resilientes y garantiza la producción sostenible de alimentos. La agricultura depende en gran medida de factores relacionados con las condiciones del cambio climático, como el estrés abiótico, que incluye deficiencias de nutrientes del suelo, aumentos acelerados de temperatura, sequía y aumento de las concentraciones de CO₂; y estrés biótico, que incluye plagas invasoras, brotes de enfermedades y disminución de la producción de cultivos. Las tecnologías de teledetección remota (RS) pueden proporcionar varios métodos no destructivos para identificar y cuantificar varios tipos de estrés. Para la aplicación de RS, es relevante considerar los diferentes tipos de resolución: espectral, espacial, temporal y radiométrica. Además, las diferentes escalas de observación son terrestres, aéreas, espaciales o utilizando satélites orbitales. En esta tesis, evaluamos la implementación práctica de métodos no destructivos utilizando tecnologías RS en los cuatro capítulos. Además, comparamos los niveles de madurez entre diferentes tipos de tecnologías utilizando evaluaciones de nivel de preparación tecnológica (TRL). En el primer capítulo, nuestro objetivo era estimar el rendimiento de grano del maíz bajo un nivel bajo de nitrógeno usando índices de vegetación (VI) de sensores RGB (imágenes compuestas de color rojo, verde y azul) a nivel del suelo y del aire en el África subsahariana (SSA). Desarrollamos un sistema RS en el segundo capítulo para monitorear un gusano cogollero de advertencia temprana (FAW) en SSA. En el tercero, usando reflectancia espectral de hoja y modelos de regresión avanzada, estimamos la $V_{c,max}$ y J_{max} de soja y maní. Y en el último capítulo, desarrollamos una aplicación móvil fácil de usar para los países de Medio Oriente y África del Norte (MENA) para la detección de trastornos de las plantas en tomates, pepinos, pimientos y quinua, que cubre todo, desde la recopilación de datos hasta la creación de modelos de aprendizaje profundo, al lanzamiento de aplicaciones web y móviles. Se utilizaron diferentes tecnologías RS en diferentes países a diferentes escalas y con diferentes tipos de sensores. Sin embargo, fue muy relevante considerar los objetivos de cada estudio porque determinaron el tipo de resolución espectral, espacial, temporal y la escala de observación. Con respecto a los

TRL en los cuatro capítulos, sugieren que el nivel de preparación tecnológica depende de los objetivos, el tiempo para desarrollar el proyecto, la cantidad de recopilación de datos requerida y la solidez de la validación.

REFERENCES

- Aasen H, Honkavaara E, Lucieer A, Zarco-Tejada PJ (2018) Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: A review of sensor technology, measurement procedures, and data correction workflows. *Remote Sens* 10:1–42. <https://doi.org/10.3390/rs10071091>
- Ainsworth EA, Serbin SP, Skoneczka JA, Townsend PA (2014) Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynth Res* 119:65–76. <https://doi.org/10.1007/s11120-013-9837-y>
- Andrade-Sanchez P, Gore MA, Heun JT, et al (2014) Development and evaluation of a field-based high-throughput phenotyping platform. *Funct Plant Biol* 41:68–79. <https://doi.org/10.1071/FP13126>
- Aqeel-Ur-Rehman, Abbasi AZ, Islam N, Shaikh ZA (2014) A review of wireless sensors and networks' applications in agriculture. *Comput Stand Interfaces* 36:263–270. <https://doi.org/10.1016/j.csi.2011.03.004>
- Araus JL, Kefauver SC (2018) Breeding to adapt agriculture to climate change: affordable phenotyping solutions. *Curr Opin Plant Biol*. <https://doi.org/10.1016/j.pbi.2018.05.003>
- Araus JL, Kefauver SC, Zaman-Allah M, et al (2018) Translating High-Throughput Phenotyping into Genetic Gain. *Trends Plant Sci*. <https://doi.org/10.1016/j.tplants.2018.02.001>
- Asplund L, Bergkvist G, Weih M (2014) Proof of concept: Nitrogen use efficiency of contrasting spring wheat varieties grown in greenhouse and field. *Plant Soil* 374:829–842. <https://doi.org/10.1007/s11104-013-1895-6>
- Atzberger C (2013) Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sens* 5:949–981. <https://doi.org/10.3390/rs5020949>
- Bacco M, Barsocchi P, Ferro E, et al (2019) The Digitisation of Agriculture: a Survey of Research Activities on Smart Farming. *Array* 3–4:100009. <https://doi.org/10.1016/j.array.2019.100009>
- Balafoutis A, Beck B, Fountas S, et al (2017) Precision agriculture technologies positively contributing to ghg emissions mitigation, farm productivity and economics. *Sustain* 9:1–28. <https://doi.org/10.3390/su9081339>
- Balafoutis AT, van Evert FK, Fountas S (2020) Smart farming technology trends: Economic and environmental effects, labor impact, and adoption readiness. *Agronomy* 10:1–26. <https://doi.org/10.3390/agronomy10050743>
- Below TB, Mutabazi KD, Kirschke D, et al (2012) Can farmers' adaptation to climate change be explained by socio-economic household-level variables? *Glob Environ Chang* 22:223–235. <https://doi.org/10.1016/j.gloenvcha.2011.11.012>
- Boliko MC (2019) FAO and the situation of food security and nutrition in the world. *J Nutr Sci Vitaminol (Tokyo)* 65:S4–S8. <https://doi.org/10.3177/jnsv.65.S4>
- Buchaillet ML, Gracia-Romero A, Vergara-Diaz O, et al (2019) Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. *Sensors (Switzerland)* 19:. <https://doi.org/10.3390/s19081815>
- Buchaillet ML, Gracia-Romero A, Zaman-Allah MA, et al (2018) Evaluating the Performance of Different Commercial and Pre-Commercial Maize Varieties under Low Nitrogen Conditions Using Affordable Phenotyping Tools. *Proceedings* 2:366. <https://doi.org/10.3390/ecrs-2-05180>
- Buchaillet ML, Soba D, Shu T, et al (2022) Estimating peanut and soybean photosynthetic traits using leaf spectral reflectance and advance regression models. *Planta* 1–19. <https://doi.org/10.1007/s00425-022-03867-6>
- Camino C, Gonzalez-Dugo V, Hernandez P, Zarco-Tejada PJ (2019) Radiative transfer V_{cmax} estimation from hyperspectral imagery and SIF retrievals to assess photosynthetic performance in rainfed and irrigated plant phenotyping trials. *Remote Sens Environ* 231:111186. <https://doi.org/10.1016/j.rse.2019.05.005>
- Casadesus J, Kaya Y, Bort J, et al (2007) Using vegetation indices derived from conventional

- digital cameras as selection criteria for wheat breeding in water-limited environments. *Ann Appl Biol* 150:227–236. <https://doi.org/10.1111/j.1744-7348.2007.00116.x>
- Casadesús J, Villegas D (2014) Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. *J Integr Plant Biol* 56:7–14. <https://doi.org/10.1111/jipb.12117>
- Chapman JW, Thompson DR, Helmlinger MC, et al (2019) Spectral and radiometric calibration of the Next Generation Airborne Visible Infrared Spectrometer (AVIRIS-NG). *Remote Sens* 11:. <https://doi.org/10.3390/rs11182129>
- Cheng Y, Vrieling A, Fava F, et al (2020) Phenology of short vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sens Environ* 248:112004. <https://doi.org/10.1016/j.rse.2020.112004>
- Choquette NE, Ogut F, Wertin TM, et al (2019) Uncovering hidden genetic variation in photosynthesis of field-grown maize under ozone pollution. *Glob Chang Biol* 25:4327–4338. <https://doi.org/10.1111/gcb.14794>
- Christian Kroll (2017) Sustainable Development Goals: Are the rich countries ready? *Sustain Gov Indic* 106
- Cobián Álvarez JA, Resosudarmo BP (2019) The cost of floods in developing countries' megacities: a hedonic price analysis of the Jakarta housing market, Indonesia. *Environ Econ Policy Stud* 21:555–577. <https://doi.org/10.1007/s10018-019-00242-w>
- Day R, Abrahams P, Bateman M, et al (2017) Fall armyworm: Impacts and implications for Africa. *Outlooks Pest Manag* 28:196–201. https://doi.org/10.1564/v28_oct_02
- de Leeuw J, Vrieling A, Shee A, et al (2014) The potential and uptake of remote sensing in insurance: A review. *Remote Sens* 6:10888–10912. <https://doi.org/10.3390/rs61110888>
- Dechant B, Cuntz M, Vohland M, et al (2017) Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism. *Remote Sens Environ* 196:279–292. <https://doi.org/10.1016/j.rse.2017.05.019>
- Demarez V, Duthoit S, Baret F, et al (2008) Estimation of leaf area and clumping indexes of crops with hemispherical photographs. *148:644–655*. <https://doi.org/10.1016/j.agrformet.2007.11.015>
- Doughty CE, Asner GP, Martin RE (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. *Oecologia* 165:289–299. <https://doi.org/10.1007/s00442-010-1800-4>
- EARTO (2014) European Association of Research and Technology Organisations. The TRL Scale as a Research & Innovation Policy Tool , EARTO Recommendations. *Earto Impact Deliv* 1–17
- FAOSTAT (2016) The State of Food and Agriculture 2016 (SOFA): Climate change, agriculture and food security
- Fiehn O (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics* 2:155–168. <https://doi.org/10.1002/cfg.82>
- Fu P, Meacham-Hensold K, Guan K, et al (2020) Estimating photosynthetic traits from reflectance spectra: A synthesis of spectral indices, numerical inversion, and partial least square regression. *Plant Cell Environ* 1–18. <https://doi.org/10.1111/pce.13718>
- Galieni A, D'Ascenzo N, Stagnari F, et al (2021) Past and Future of Plant Stress Detection: An Overview From Remote Sensing to Positron Emission Tomography. *Front Plant Sci* 11:1–22. <https://doi.org/10.3389/fpls.2020.609155>
- Gamon JA, Serrano L, Surfus JS (1997) The photochemical reflectance index: An optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* 112:492–501. <https://doi.org/10.1007/s004420050337>
- Garrigues S, Shabanov N V., Swanson K, et al (2008) Intercomparison and sensitivity analysis of Leaf Area Index retrievals from LAI-2000, AccuPAR, and digital hemispherical photography over croplands. *Agric For Meteorol* 148:1193–1209.

- <https://doi.org/10.1016/j.agrformet.2008.02.014>
- Ge Y, Thomasson JA, Sui R (2011) Remote sensing of soil properties in precision agriculture: A review. *Front Earth Sci* 5:229–238. <https://doi.org/10.1007/s11707-011-0175-0>
- Gerhards M, Rock G, Schlerf M, Udelhoven T (2016) Water stress detection in potato plants using leaf temperature, emissivity, and reflectance. *Int J Appl Earth Obs Geoinf* 53:27–39. <https://doi.org/10.1016/j.jag.2016.08.004>
- 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 Sens* 11:. <https://doi.org/10.3390/rs11101240>
- Govender M, Dye PJ, Weiersbye IM, et al (2009) Review of commonly used remote sensing and ground-based technologies to measure plant water stress. *Water SA* 35:741–752. <https://doi.org/10.4314/wsa.v35i5.49201>
- Gracia-Romero A, Kefauver SC, Fernandez-Gallego JA, et al (2019) UAV and ground image-based phenotyping: A proof of concept with durum wheat. *Remote Sens* 11:. <https://doi.org/10.3390/rs11101244>
- Gracia-Romero A, Kefauver SC, Vergara-Díaz O, et al (2017) Comparative Performance of Ground vs. Aerially Assessed RGB and Multispectral Indices for Early-Growth Evaluation of Maize Performance under Phosphorus Fertilization. *Front Plant Sci* 8:1–13. <https://doi.org/10.3389/fpls.2017.02004>
- Hailu G, Niassy S, Zeyaur KR, et al (2018) Maize–legume intercropping and push–pull for management of fall armyworm, stemborers, and striga in Uganda. *Agron J* 110:2513–2522. <https://doi.org/10.2134/agronj2018.02.0110>
- Herrmann I, Bdolach E, Montekyo Y, et al (2020) Assessment of maize yield and phenology by drone-mounted superspectral camera. *Precis Agric* 21:51–76. <https://doi.org/10.1007/s11119-019-09659-5>
- Homolová L, Malenovský Z, Clevers JGPW, et al (2013) Review of optical-based remote sensing for plant trait mapping. *Ecol Complex* 15:1–16. <https://doi.org/10.1016/j.ecocom.2013.06.003>
- Ihuoma SO, Madramootoo CA, Kalacska M (2021) Integration of satellite imagery and in situ soil moisture data for estimating irrigation water requirements. *Int J Appl Earth Obs Geoinf* 102:102396. <https://doi.org/10.1016/j.jag.2021.102396>
- Jawad HM, Nordin R, Gharghan SK, et al (2017) Energy-efficient wireless sensor networks for precision agriculture: A review. *Sensors (Switzerland)* 17:. <https://doi.org/10.3390/s17081781>
- Jawid A, Khadjavi M (2019) Adaptation to climate change in Afghanistan: Evidence on the impact of external interventions. *Econ Anal Policy* 64:64–82. <https://doi.org/10.1016/j.eap.2019.07.010>
- Jin X, Zarco-Tejada PJ, Schmidhalter U, et al (2020) High-Throughput Estimation of Crop Traits. *IEEE Geosci Remote Sens Mag* 1–33
- Kefauver SC, Gracia Romero A, Buchailot ML, et al (2020) OPEN-SOURCE SOFTWARE FOR CROP PHYSIOLOGICAL ASSESSMENTS. pp 4359–4362
- Kernecker M, Knierim A, Wurbs A, et al (2020) Experience versus expectation: farmers’ perceptions of smart farming technologies for cropping systems across Europe. *Precis Agric* 21:34–50. <https://doi.org/10.1007/s11119-019-09651-z>
- Khan MA, Akram T, Sharif M, et al (2020) An automated system for cucumber leaf diseased spot detection and classification using improved saliency method and deep features selection. *Multimed Tools Appl* 79:18627–18656. <https://doi.org/10.1007/s11042-020-08726-8>
- Kluk AF, Johnson HC, Mcginnis CP, et al (2008) *Technology Readiness Levels Handbook for space applications*
- Koester RP, Skoneczka JA, Cary TR, et al (2014) Historical gains in soybean (*Glycine max* Merr.) seed yield are driven by linear increases in light interception, energy conversion, and

- partitioning efficiencies. *J Exp Bot* 65:3311–3321. <https://doi.org/10.1093/jxb/eru187>
- Kovacs JM, Flores-Verdugo F, Wang J, Aspden LP (2004) Estimating leaf area index of a degraded mangrove forest using high spatial resolution satellite data. *Aquat Bot* 80:13–22. <https://doi.org/10.1016/j.aquabot.2004.06.001>
- Kuhlmann F, Brodersen C (2001) Information technology and farm management: Developments and perspectives. *Comput Electron Agric* 30:71–83. [https://doi.org/10.1016/S0168-1699\(00\)00157-5](https://doi.org/10.1016/S0168-1699(00)00157-5)
- Lindblom J, Lundström C, Ljung M, Jonsson A (2017) Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies. *Precis Agric* 18:309–331. <https://doi.org/10.1007/s11119-016-9491-4>
- Lukina E V, Stone ML, Raun WR (1999) Estimating vegetation coverage in wheat using digital images. *J Plant Nutr* 22:341–350. <https://doi.org/10.1080/01904169909365631>
- Lussem U, Bolten A, Gnyp ML, et al (2018) Evaluation of RGB-based vegetation indices from UAV imagery to estimate forage yield in Grassland. *Int Arch Photogramm Remote Sens Spat Inf Sci - ISPRS Arch* 42:1215–1219. <https://doi.org/10.5194/isprs-archives-XLII-3-1215-2018>
- Maciejczak M, Faltmann J (2018) Assessing Readiness Levels of Production Technologies for Sustainable Intensification of Agriculture. *Appl Stud Agribus Commer* 12:47–52. <https://doi.org/10.19041/apstract/2018/1-2/7>
- Maes WH, Steppe K (2019) Perspectives for Remote Sensing with Unmanned Aerial Vehicles in Precision Agriculture. *Trends Plant Sci* 24:152–164. <https://doi.org/10.1016/j.tplants.2018.11.007>
- Mankins J (1995) TECHNOLOGY READINESS LEVELS. *Adv Concepts Off Off Sp Access Technol NASA*. <https://doi.org/10.1080/08956308.2010.11657640>
- Mariotto I, Thenkabail PS, Huete A, et al (2013) Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypsIRI mission. *Remote Sens Environ* 139:291–305. <https://doi.org/10.1016/j.rse.2013.08.002>
- Masuka B, Araus JL, Das B, et al (2012) Phenotyping for Abiotic Stress Tolerance in Maize. *J Integr Plant Biol* 54:238–249. <https://doi.org/10.1111/j.1744-7909.2012.01118.x>
- Matese A, Toscano P, Di Gennaro FS, et al (2015) Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. 2971–2990. <https://doi.org/10.3390/rs70302971>
- Meacham-Hensold K, Montes CM, Wu J, et al (2019) High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. *Remote Sens Environ* 231:111176. <https://doi.org/10.1016/j.rse.2019.04.029>
- Mogili UR, Deepak BBVL (2018) Review on Application of Drone Systems in Precision Agriculture. *Procedia Comput Sci* 133:502–509. <https://doi.org/10.1016/j.procs.2018.07.063>
- Nash E, Korduan P, Bill R (2009) Applications of Open Geospatial web services in precision agriculture: A review. *Precis Agric* 10:546–560. <https://doi.org/10.1007/s11119-009-9134-0>
- Nelson GC, Rosegrant MW, Koo J, et al (2009) Food Policy Report: Climate Change- Impact on Agriculture and and Costs of Adaptation
- Onuegbu NC, Ibeabuchi JC (2021) The Role of Indigenous Food Species in Achieving Food Security in South-Eastern Nigeria
- Osinga SA, Paudel D, Mouzakitis SA, Athanasiadis IN (2022) Big data in agriculture: Between opportunity and solution. *Agric Syst* 195:103298. <https://doi.org/10.1016/j.agsy.2021.103298>
- Otekunrin OA, Otekunrin OA, Momoh S, Ayinde IA (2019) How far has Africa gone in achieving the zero hunger target? Evidence from Nigeria. *Glob Food Sec* 22:1–12. <https://doi.org/10.1016/j.gfs.2019.08.001>

- Peñuelas J, Filella L (1998) Technical focus: Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci* 3:151–156. [https://doi.org/10.1016/S1360-1385\(98\)01213-8](https://doi.org/10.1016/S1360-1385(98)01213-8)
- Petisco C, García-Criado B, Mediavilla S, et al (2006) Near-infrared reflectance spectroscopy as a fast and non-destructive tool to predict foliar organic constituents of several woody species. *Anal Bioanal Chem* 386:1823–1833. <https://doi.org/10.1007/s00216-006-0816-4>
- Pylaniadis C, Osinga S, Athanasiadis IN (2021) Introducing digital twins to agriculture. *Comput Electron Agric* 184:105942. <https://doi.org/10.1016/j.compag.2020.105942>
- Rao NR, Garg PK, Ghosh SK (2006) The effect of radiometric resolution on the retrieval of leaf area index from agricultural crops. *GIScience Remote Sens* 43:377–387. <https://doi.org/10.2747/1548-1603.43.4.377>
- Rembold F, Carnicelli S, Nori M, Ferrari GA (2000) Use of aerial photographs, Landsat TM imagery and multidisciplinary field survey for land-cover change analysis in the lakes region (Ethiopia). *ITC J* 2:181–189. [https://doi.org/10.1016/S0303-2434\(00\)85012-6](https://doi.org/10.1016/S0303-2434(00)85012-6)
- Reyns P, Missotten B, Ramon H, De Baerdemaeker J (2002) A review of combine sensors for precision farming. *Precis Agric* 3:169–182. <https://doi.org/10.1023/A:1013823603735>
- Schellberg J, Hill MJ, Gerhards R, et al (2008) Precision agriculture on grassland: Applications, perspectives and constraints. *Eur J Agron* 29:59–71. <https://doi.org/10.1016/j.eja.2008.05.005>
- Segarra J, Buchailot ML, Araus JL, Kefauver SC (2020) Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* 10:1–18. <https://doi.org/10.3390/agronomy10050641>
- Serbin SP, Dillaway DN, Kruger EL, Townsend PA (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *J Exp Bot* 63:489–502. <https://doi.org/10.1093/jxb/err294>
- Serbin SP, Singh A, McNeil BE, et al (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecol Appl* 24:1651–1669. <https://doi.org/10.1890/13-2110.1>
- Silva-Perez V, Molero G, Serbin SP, et al (2018) Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *J Exp Bot* 69:483–496. <https://doi.org/10.1093/jxb/erx421>
- Steffen W, Richardson K, Rockström J, et al (2015) Planetary boundaries: Guiding human development on a changing planet. *Science* (80-) 347:. <https://doi.org/10.1126/science.1259855>
- Stewart EL, Wiesner-Hanks T, Kaczmar N, et al (2019) Quantitative Phenotyping of Northern Leaf Blight in UAV Images Using Deep Learning. *Remote Sens* 11:1–10. <https://doi.org/10.3390/rs11192209>
- Tanumihardjo SA, McCulley L, Roh R, et al (2020) Maize agro-food systems to ensure food and nutrition security in reference to the Sustainable Development Goals. *Glob Food Sec* 25:100327. <https://doi.org/10.1016/j.gfs.2019.100327>
- Tetila EC, Machado BB, Menezes GK, et al (2020) Automatic Recognition of Soybean Leaf Diseases Using UAV Images and Deep Convolutional Neural Networks. *IEEE Geosci Remote Sens Lett* 17:903–907. <https://doi.org/10.1109/LGRS.2019.2932385>
- Thenkabail PS, Smith RB, Pauw E De (2002) Evaluation of Narrowband and Broadband Vegetation Indices for Determining Optimal Hyperspectral Wavebands for Agricultural Crop Characterization. 68:607–621
- Toth C, Józkó G (2016) Remote sensing platforms and sensors: A survey. *ISPRS J Photogramm Remote Sens* 115:22–36. <https://doi.org/10.1016/j.isprsjprs.2015.10.004>
- Tsakmaki ID, Gikas GD, Sylaios GK (2021) Integration of Sentinel-derived NDVI to reduce uncertainties in the operational field monitoring of maize. *Agric Water Manag* 255:106998. <https://doi.org/10.1016/j.agwat.2021.106998>
- Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation.

- Remote Sens Environ 8:127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Ureta C, González EJ, Espinosa A, et al (2020) Maize yield in Mexico under climate change. *Agric Syst* 177:102697. <https://doi.org/10.1016/j.agry.2019.102697>
- Van Gardingen PR, Jackson GE, Hernandez-Daumas S, et al (1999) Leaf area index estimates obtained for clumped canopies using hemispherical photography. *Agric For Meteorol* 94:243–257. [https://doi.org/10.1016/S0168-1923\(99\)00018-0](https://doi.org/10.1016/S0168-1923(99)00018-0)
- Van Hoek M, Jia L, Zhou J, et al (2016) Early drought detection by spectral analysis of satellite time series of precipitation and Normalized Difference Vegetation Index (NDVI). *Remote Sens* 8:. <https://doi.org/10.3390/rs8050422>
- Verde N, Mallinis G, Tsakiri-Strati M, et al (2018) Assessment of radiometric resolution impact on remote sensing data classification accuracy. *Remote Sens* 10:1–17. <https://doi.org/10.3390/rs10081267>
- Vergara-Diaz O, Kefauver SC, Elazab A, et al (2015) Grain yield losses in yellow-rusted durum wheat estimated using digital and conventional parameters under field conditions. *Crop J* 3:200–210. <https://doi.org/10.1016/j.cj.2015.03.003>
- Virlet N, Sabermanesh K, Sadeghi-Tehran P, Hawkesford MJ (2017) Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. *Funct Plant Biol* 44:143–153. <https://doi.org/10.1071/FP16163>
- von Grebmer; K., Bernstein; J, Hammond; L, et al (2018) 2018 Global Hunger Index: Forced Migration und Hunger, Global Hunger Index. Global Hunger Index PP - Germany, Germany
- Walter A, Finger R, Huber R, Buchmann N (2017) Smart farming is key to developing sustainable agriculture. *Proc Natl Acad Sci U S A* 114:6148–6150. <https://doi.org/10.1073/pnas.1707462114>
- Wang N, Zhang N, Wang M (2006) Wireless sensors in agriculture and food industry - Recent development and future perspective. *Comput Electron Agric* 50:1–14. <https://doi.org/10.1016/j.compag.2005.09.003>
- Yendrek CR, Tomaz T, Montes CM, et al (2017a) High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. *Plant Physiol* 173:614–626. <https://doi.org/10.1104/pp.16.01447>
- Yendrek CR, Tomaz T, Montes CM, et al (2017b) High-Throughput Phenotyping of Maize Leaf Physiological and Biochemical Traits Using Hyperspectral Reflectance 1 [OPEN]. 173:614–626. <https://doi.org/10.1104/pp.16.01447>
- Zhang C, Kovacs JM (2012) The application of small unmanned aerial systems for precision agriculture: A review. *Precis Agric* 13:693–712. <https://doi.org/10.1007/s11119-012-9274-5>
- Zhang N, Wang M, Wang N (2002) Precision agriculture - A worldwide overview. *Comput Electron Agric* 36:113–132. [https://doi.org/10.1016/S0168-1699\(02\)00096-0](https://doi.org/10.1016/S0168-1699(02)00096-0)
- Zolfaghari M, Farzaneh J (2021) Food Security in the Middle East and North Africa (MENA)

