PARAMETRIZATION OF CROP MODELS USING UAS CAPTURED DATA

by

Bilal Abughali

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science in Agricultural and Biological Engineering

School of Agricultural and Biological Engineering

West Lafayette, Indiana

December 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Keith A. Cherkauer, Chair
Department of Agricultural and Biological Engineering

Dr. Katy Martin Rainey
Department of Agronomy

Dr. Laura C Bowling
Department of Agronomy

Approved by:
Dr. Nathan Mosier

Dedicated to my loved ones

ACKNOWLEDGMENTS

The fulfillment of my master’s degree has been a very challenging yet rewarding experience. I have had the privilege to learn from and work with an amazing group of mentors and colleagues that made my journey easier and more enjoyable.

I would like to thank my advisor Dr. Cherkauer for his enormous amounts of patience. Dr. Cherkauer created a space where I could learn and develop my skills in a way that suited my learning style. His patience and willingness to help with even the most minute questions is a testament to his love for teaching, which I hope I can convey to my future students if I get the chance. I would also like to thank Dr. Laura Bowling and Dr. Katy Rainey, for pushing me to do my best work. Next, I would like to thank Stuart Smith, André Moraes and the rest of the Purdue Hydrologic Impacts Group (PHIG) for tolerating my nagging and unfathomable number of questions. Lastly, I would like to thank Vincent Seal, whom without I would have never finished the endless amounts of field work this project required.

The past 3 years have been though, especially with the COVID pandemic, but my mentors, colleagues and friends have been a great source of support and inspiration, in pushing me to strive.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Calibration of crop models is an expensive and time intensive procedure, which is essential to accurately predict the possible crop yields given changing climate conditions. One solution is the utilization of unmanned aircraft systems (UAS) deployed with Red Green Blue Composite (RGB), and multispectral sensors, which has the potential to measure and collect in field biomass and yield in a cost and time effective manner. The objective of this project was to develop a relationship between remotely sensed data and crop indices, similar to biomass, to improve the ability to parametrize crop models for local conditions, which in turn could potentially improve the quantification of the effect of hydrological extremes on predicted yield. An experiment consisting of 750 plots (350 varieties) was planted in 2018, and a subset of 18 plots (9 varieties) were planted in 2019. The in-situ above ground biomass along with multispectral and RGB imagery was collected for both experiments throughout the growing season. The imagery was processed through a custom software pipeline to produce spectrally corrected imagery of individual plots. A model was fit between spectral data and sampled biomass resulting in an R-square of 0.68 and RMSE of 160 g when the model was used to estimate biomass for multiple flight dates flights. The VIC-CropSyst model, a coupled hydrological and agricultural system model, was used to simulate crop biomass and yield for multiple years at the experiment location. Soybean growth was parametrized for the location using CropSyst's Crop Calibrator tool. Biomass values generated from UAS imagery, along with the in-situ collected biomass values were used separately to parametrize soybean simulations in CropSyst resulting in very similar parameter sets that were distinct from the default parameter values. The parametrized crop files along with the default files were used separately to run the VIC-CropSyst model and results were evaluated by comparing simulated and observed values of yield and biomass values. Both parametrized crop files (using in-situ samples and UAS imagery) produced approximately identical results with a max difference of 0.03 T/Ha for any one year, compared to a base value of 3.6 T/Ha, over a 12-year period in which the simulation was ran. The parametrized runs produced yield estimates that were closer to in-situ measured yield, as compared to unparametrized runs, for both bulk varieties and the run experiments, with the exception of 2011, which was a flooding year. The parametrized simulations consistently produced simulated yield results that were higher than the measured bulk variety yields, whereas the default parameters produced consistently lower yields. Biomass was only

assessed for 2019, and the results indicate that the biomass after parametrization is lower than the default, which is attributed to the radiation use efficiency parameter being lower in the parametrized files, 2.5 g/MJ versus 2.25 g/MJ. The improved accuracy of predicting yield is evidence that the UAS based methodology is a suitable substitute for the more labor intensive in-situ sampling of biomass for soybean studies under similar environmental conditions.

1. INTRODUCTION

1.1 Background

According to a World Bank census, global population numbers have risen from 6.11 billion in 2000 to 7.53 billion in 2018 (World Bank, 2017). This population growth requires increased crop production to sustain (Cohen, 2003). However, crop production increase is limited due multiple factors similar to fertility and area of cultivation, as well as water and climactic stresses (Cohen, 2003; Sinha et al., 1989). The Midwest is representative of this scenario, as climate models along with observational data, predict increased frequency of extreme weather in the future, which in turn could reduce crop yields (Cherkauer et al., 2010; Fan, 2014; Bowling et al., 2020). Two crops heavily affected by recent climate changes are corn and soybean, and the effects can be easily identified for the 2019 crop cycle. The combination of a long-wet spring and a dry summer resulted in decreases in both corn and soybean production by 5.3% and 19.8% respectively in the US (USDA, 2019). The hydrological climate extremes heavily affected corn in particular which experienced a drop in production even with an 8% rise in area planted for corn (USDA, 2020). Hence, there is a need to further our understanding to the effects of droughts and floods on crop growth and yield under different climate scenarios.

Research in agronomy, breeding and crop modeling further our knowledge into the dynamics governing crop growth as it is related to water stress and expected yield. This is achieved by relying on current knowledge along with the immense amount of agronomic and meteorological data collected annually. Currently crop breeders are utilizing different phenotypes on a field scale to evaluate the effects of water stress on growth (Araus and Cairns, 2014; Cabrera-Bosquet et al., 2012). This procedure involves planting hundreds or thousands of different varieties of the similar crop within a field and tracking the development of different physical traits throughout the plant's growth for each of the varieties (Moreira et al., 2019). The physical traits or phenotypes, which include biomass and leaf area index, serve as indicators to the performance of the planted varieties under the specific conditions witnessed throughout the planting season, which helps identify high yielding varieties (de Paiva Rolla et al., 2014). Biomass is the weight of the above ground dry matter in the canopy, and it directly correlates to the yield of a crop through a harvest index, such that as biomass increases, yield increases. Leaf area index is the measure of the amount of foliage in canopies, which is the surface area of leaves over a meter square of soil. Leaf area index determines the light interception capacity of a crop, in general higher leaf area index indicates higher crop productivity (Weraduwage et al., 2015). The relationship between plant phenotypes and plant growth is variable for the same crop varieties, as such multiple varieties must be screened simultaneously, under similar conditions so that varieties with desired characteristics can be easily identified. Increasing the speed and accuracy with which such measurements can be made will aid in rapidly identifying varieties that are more resistant to drought or flooding, while also identifying high yield potential.

Substantial advancements have been achieved by conducting phenotypic studies at a field and plot scale. The field scale in such studies spans a few hectares while the plot scale is just a few square meters in area (Borra-Serrano et al., 2020). Experiments are typically designed to study hundreds to thousands of varieties within a field. The field is divided into small spatial plots, and each plot represents a single crop variety as such each plot usually requires separate observations and analysis. A prominent and recurring problem faced in such setups is the high variation within the field itself, especially in larger fields where soil and hydrologic conditions might differ. In such cases, multiple observations of the same variety crops are needed to account for in field variations (the environment). The large number of precise observations needed for such studies traditionally required a substantial amount of manual labor (Araus & Cairns, 2014). The high cost of manual labor along with the requirement for precise data results in a need to find a solution that cuts labor cost while maintaining a high level of precision of the data collected.

To overcome the cost of labor and maintain a high level of data precision, researchers have started to exploit new methods for data collection. The use of Unmanned Aircraft Systems (UAS) equipped with imaging sensors are a possible solution (Borra-Serrano et al., 2020). UAS require minimal labor and can collect imagery for the entire field in a single flight. The spatial resolution, of approximately 1 cm/pixel to 3 cm/pixel, of the sensors allows researchers to assess individual plots within the field. Moreover, the UAS can be flown multiple times over the same fields, resulting in the introduction of a high temporal aspect to the data collected, for a lower cost than conventional remote sensing methods. The combination of both high spatial and temporal resolutions allows researchers to accurately assess plant growth throughout the growing season (Tsouros et al., 2019).

The initial difficulty in dealing with remote sensing data, retrieved via UAS, is the fact that the data are fundamentally different to datasets traditionally used for evaluating crop models, and relationships must be developed between what is measured from above the canopy by UAS imagery and the in-situ measurements collected manually within the field. Initial advancements in the field have linked canopy color and canopy cover size to crop characteristics similar to above ground biomass, LAI, and yield (Duchemin et al., 2008). Canopy coverage is defined as the proportion of the forest floor covered by the vertical projection of the tree crowns (USDA Forest Service, 2010). Canopy color is typically quantified using mathematical relationships between spectral bands to calculate UAS crop indices, which relate to chlorophyll content or photosynthetic activity. One example would be the normalized difference vegetation index (NDVI) which measures the state of plant health based on how the plant reflects light at certain frequencies (Rouse et al., 1974), given that the living green plants absorb solar radiation in the photosynthetically active radiation due to the chlorophyll content, and reflect at the near infra-red.

Canopy coverage, which is a measure of how much the plant canopy covers the ground, has been crucial for high-throughput phenotyping at a field scale due to the relative ease of accurately calculating canopy cover from UAS images as well as its strong relation to light interception and yield (Holben Compton et al., 1980). The use of canopy cover is prevalent, and models rely on it as a measure to estimate LAI for plant growth estimates (Stockle et al., 1994). Multiple crop models utilize canopy cover in their calibration and parameterization processes. One such model is AquaCrop, which was developed by the United Nation's Food and Agriculture Organization (FAO) to address food security and assess the effect of the environment and management on crop production. The model heavily depends on canopy cover estimates to approximate LAI and ultimately infer crop growth. The model utilizes canopy cover fractions as they are easier to collect than LAI measurements. Another older but more data intensive physically based simulation model is the Cropping Systems (CropSyst) simulation model. CropSyst is a multi-year, multi-crop, crop growth simulation model (Stockle, 1996). Originally it had been developed to serve as an analytic tool to study the effect of cropping systems management on productivity and the environment. The model requires multiple inputs to describe crop properties, soil characteristics and metrological conditions, which is similar to AquaCrop; however, CropSyst is more data intensive and requires larger and more precise datasets. The complex nature of the model allows for its use in studying the effects of changing environments and crop types to potential yields. Both models are used to assess the environments effect on crop growth; however, AqauCrop is mostly utilized in conditions where water is limited, while CropSyst is a more holistic model, which more accurately models crop growth under different management practices, and environmental conditions. Crop modelling has continued to develop rapidly, and more complex models now exist, an example being the Agricultural Production Systems sIMulator (APSIM), which contains interconnected models to simulate systems comprising soil, crop, tree, pasture, and livestock biophysical processes (Holzworth et al., 2014). Initially, APSIM was a cropping systems model that later evolved into an agro-ecosystem model, which means that APSIM not only models crop growth, but also contains soil and animal models. (Holzworth et al., 2014). APSIM relies on multiple input parameters that are hard to measure including radiation use efficiency and transpiration efficiency. CropSyst can simulate the effects of water stress on crops to better predict yield accordingly (Stockle et al., 2003). It is important to highlight that many of the crop models were built on the same fundamental equations, however the applications of these models through their integration with other models, and their customizability is what sets them apart.

One example of the coupling of models impactful to this study is the coupling of the macroscale Variable Infiltration Capacity (VIC) hydrologic model and the CropSyst model to predict yield values, the coupled model is known as the VIC-CropSyst model. The VIC-CropSyst model allows for more accurate water use simulation, as the VIC model can accurately identify the available water for crop growth as dictated by a full water and energy balance at the land surface, and feed that data into the CropSyst model to determine water use (Malek et al., 2017).

Crop parameters applied in these models are usually for general conditions, but parameter calibration can be conducted to better represent local conditions and crop varieties (Stockle et al., 2003). In addition, advancements in crop breeding have resulted in new crop varieties that may not be as well represented by the default parameterizations. The parameters required to populate these models are usually gathered in the field using both destructive and non-destructive methods. Field sampling is both costly and time consuming, which limits the number of samples taken both in space and time that what would be optimal to accurately model crop behavior. To increase the availability of data that can be used for parameterization of models, researchers have sought to supplant ground-based sampling with imagery-derived indices related to crop physical and physiological properties. The use of crop indices such as NDVI, the soil adjusted vegetation index (SAVI) and others have been successfully employed in estimating plant biomass development through the growing season and potential yield at the end of the season. In season biomass, in particular, is useful for quantifying water stress effects in near-real time. For example, the relationship between biomass and NDVI has been used to help identify crops that might be experiencing moisture stress due to droughts and floods (Jones, 2004; Smith, 2021). The acquisition of data for crop modeling is a daunting task, as the cost is high, and the time needed for collecting samples on a field scale is relatively lengthy and difficult. As such, utilizing Unmanned Aerial Systems (UAS) equipped with RGB, Multispectral and Thermal sensors is proposed as a way to cut down these costs. Use of inexpensive sensors collecting high-resolution imagery as needed through the growing season allows for the rapid calculation of multiple crop indices that could potentially be used to predict biomass, LAI and final yield (Sankaran et al., 2015).

High-throughput data, collected using UAVs, is usually more time efficient than in-situ sampling as data from the entire field can be collected within a few hours. This process generates a substantial amount of data when collecting high resolution imagery, as an experimental field can include upwards of a thousand plots that can be all collected and processed in a single flight operation. Handling all of the data, and quickly extracting useful agronomic information is necessary for remote sensing from UAS platforms to become an important tool. As such, this study aims to develop and evaluate a method for extracting phenotypic information from UAS imagery that has a direct application in the parametrization of crop models. This research will focus on soybean varieties currently being tested at Purdue's Agronomy Center for Research and Education (ACRE), which will support the simulation of multiple years of soybean growth and can potentially be used to quantify the impacts of hydrological extremes on crop yields. Development of these methods allows for production of a systematic procedure for the parametrization of different crop varieties under different hydrologic conditions.

1.2 Goals and Objectives

The overall objective of this study is to accurately quantify soybean growth, relative to local environmental conditions and to quantify the effect of model parametrization using those observations on final crop yields under normal and extreme hydrologic conditions. Climate change is altering weather patterns in some areas, and the ability to quantify the effects of these changes on crop production is crucial for management of farms and maintaining high yield to sustain the growing population. This research will be performed in order to test the following hypotheses:

1. Development of relationships between remotely sensed data and crop indices is expected to improve the ability to parametrize crops for local conditions.
2. Parametrization of crop models for local conditions is expected to improve their ability to quantify the impact of hydrological extremes on predicted crop yields.

These hypotheses will be addressed using the following procedures:

1. Field measurements of biomass and LAI were collected for multiple varieties of soybean. In addition, field measurements for reflectance of spectral panels were collected to atmospherically correct UAS collected data.
2. UAS data were collected regularly over soybean fields. The data collected was then atmospherically corrected using the empirical line method.
3. Band algorithms were generated from atmospherically corrected plot images, and regression lines were developed that link field measurements with UAS collected data.
4. The regression lines were used with data from previous seasons that demonstrate lower water stress for the same crop varieties, in order to estimate multiple values of biomass.
5. The estimated biomass and measured LAI was used within CropSyst's crop calibrator, to calibrate key parameters essential for crop growth.
6. The parametrized crop files were then used within the VIC-CropSyst model to simulate multiple growing seasons and estimate final yields.

1.3 Thesis Organization

This thesis is organized into five chapters. Chapter 1 provides an introduction to the developments of crop modelling as it pertains to the use of UAS imagery. It also provides the objective of this study and the hypotheses that will be assessed. Chapter 2 describes the process of high-throughput data collection and processing. The processes of plot extraction, atmospheric correction and development of regression lines using band indices and field measurements. Chapter 3 summarizes the VIC-CropSyst model along with CropSyst's built in crop calibrator tool. In chapter 4, simulation results of biomass and predicted yield are assessed to quantify the effect of using different data sources to parametrize the crop model. Finally, chapter 5 recaps and discusses the results of this study. In addition, it offers insight into future work and development that may be accomplished in this topic area.

2. HIGH-THROUGHPUT DATA COLLECTION AND PROCESSING

2.1 High-Throughput Data Overview

High-throughput data are information generated quickly and automatically resulting in massive datasets that are being used to address large and complex problems. When applied in the field of agronomy, it often refers to the rapid collection of large amounts of phenotypic data, mostly through the application of remote sensing tools. This offers a non-destructive approach to plant screening (White et al., 2012). Improvements in remote sensing technologies along with advances in data processing have increased the application and improved the accuracy of high-throughput phenotyping (Leinonen and Jones, 2004), which has the potential to allow crop breeders to study specific traits with the objective of breeding a more productive and resilient crop. Many high-throughput phenotyping platforms have been examined (e.g., Yang et al., 2013; Araus and Cairns, 2014), and most are in controlled and fully automated environments such as greenhouses and growth chambers. The problem with such environments is that they simulate conditions that are removed from the reality that is occurring on the field scale. The initial application of remote sensing technologies, for crop studies on a field scale, came in the form of satellite imaging technologies (Sankaran et al., 2015). However, currently available satellite sensors have major limitations due to the high cost, low spatial resolution for the identification of desirable traits, the influence of atmospheric effects and lengthy periods between revisits (Issei et al., 2010). Given the limitations of satellite imagery, researchers are focusing on unmanned aircraft systems (UAS), which provide the potential for large-scale crop monitoring with a high spatial, spectral, and temporal resolutions. Initially, the use of such systems has been limited to research activities due to the high cost and complexity of the platforms (Chapman et al., 2014). Nevertheless, in today's standards, these UAS are considered to be an affordable and powerful tool for crop phenotyping as compared to their satellite counterparts (Berni et al., 2009) as they offer a low-cost approach to meet the requirements of spatial, spectral, and temporal resolutions for a given site and study.

Phenotype, in terms of crops, is the expression of the genotype (genetic constitution), environmental effects and the management practices that influence growth and development. Some traits such as Leaf Area Index (LAI), plant height, lodging and Canopy Cover (CC) are considered geometric traits (related to the shape of the plant), while traits such as biomass and photosynthesis are considered physiological traits. In terms of management practices, those can relate to use of herbicides, quantity of supplied irrigation water, irrigation mode, pesticide application, tillage and other field management operations. In regards to this specific study, management practices are consistent for all fields and experiments, as such the effects of management practices are not considered.

Distinct methods have been developed and proposed to evaluate phenotypic traits in the field, whether it be dependent on spectral signature, canopy temperature or reflectance (Araus and Cairns, 2014). In general, geometric traits such as plant height or LAI are usually estimated by building a digital surface model or digital surface elevation of the canopy, along with image classification analysis (Hunt et al. 2005, 2010). Physiological traits are generally dependent on different plant indices built by using the reflectance and absorption from the canopies (Hunt et al. 2005, 2010). Estimating and predicting crop yield and growth will require the assessment of both geometric and physiological traits, as such both must be assessed, which would support their use in calibrating the crop model.

2.2 Experiment Setup and Field Sampling

To accurately predict crop growth and development, both field and UAS collected data are required. Data that was collected in the field includes biomass, LAI, dates at which crops reached a certain developmental stage and spectrometer reflectance from spectral calibration panels. UAS collected data included both multispectral and RGB images of the fields taken at different intervals during the growing season. Multiple flights for each sensor were performed per week, weather allowing, but not all flights resulted in data that are optimal in terms of both image quality and timing. The sensor data allows for the calculation of multiple crop indices including NDVI and SAVI, and the estimation of plant geometric traits such as canopy cover. Combining field and UAS data aids in the development of models that correlate band indices to crop growth phenotypes such as biomass accumulation. It is important to highlight that the fields are divided into multiple experiments, each focusing on a certain aspects of crop growth, and that this study considers only a subset of the overall set of experiments.

All high-throughput data along with field measurements were collected for the summer field seasons of 2018 and 2019, the locations of which are shown in Figure 1. These data sets were collected at the Agronomy Center for Research and Education (ACRE), which was established in 1949 for research groups interested in field crops. Multiple varieties of soybeans were used in this study. Upwards of 384 unique varieties were tested in the field season of 2018, while nine focused groups were assessed in the summer of 2019.

The 2018 experiment was made up of three sub experiments named RUE I, RUE II and RUE calibration. The three experiments were layout next to each other in the field as shown in Panel A in Figure 1. The RUE experiments are not complete replicates as the varieties used are not completely the same between experiments. RUE I and RUE II are the same size with approximately 350 varieties each, while RUE calibration is the smaller subset with 60 varieties in total.

The choice of the nine focus classes was dependent on the lodging rate, which is the dislocation of stems or roots from their upright and proper placement, and yield values from the 2018 varieties. All nine varieties selected for the 2019 experiment had low lodging in 2018. In terms of yield, the varieties were chosen to represent the range of 2018 yield values, there were three low yielding varieties, three high yielding, two medium yielding varieties and one control line. The setup as well as the varieties chosen can be seen in Table 1. In total, 18 plots were assessed in 2019 (nine varieties by two replications), where plots 1 to 9 are the first repetition, and plots 10 to 18 are the second repetition. Each plot is made of eight rows (eight rows per plot). The use of eight rows is crucial as rows 1, 5 and 8 are used as border rows, rows 2, 3 and 4 are used for destructive biomass sampling and rows 6 and 7 are used for non-destructive LAI measurements and for extracting information from remote sensing images. A similar plot layout was utilized for the 2018 RUE experiments.

Table 1: Genetic varieties used as well as their layout in the field and growth highlights for the 2019 experiment. Range is the location of the field north to south based on the southern edge (Range 1) of the experiment, so Range 18 is the northernmost plot.

Range	**Cultivar**	**Yield**	**Growth Highlights**
18	DS11-06174	High	Hand planted, low development
17	DS11-31160	Medium	Hand planted, low development
16	DS11-34110	Low	NA
15	DS11-06182	High	NA
14	IA-3048	Control	NA
13	DS11-42112	Low	NA
12	DS11-03007	High	NA
11	DS11-09043	Medium	NA
10	DS11-40064	Low	NA
9	DS11-06174	High	NA
8	DS11-31160	Medium	NA
7	DS11-34110	Low	NA
6	DS11-06182	High	NA
5	IA-3048	Control	NA
4	DS11-42112	Low	NA
3	DS11-03007	High	NA
2	DS11-09043	Medium	Planted late, late development
1	DS11-40064	Low	Planted late, late development

Plots in the field are named depending on their location in the experiment. Row describes the plot's location in the x-axis (East-West) while Range describes the plot's location in y-axis (North-South). The southwestern corner is considered the first plot and as such is given the identification of row 1, range 1. Table 1 shows the experimental setup of the 2019 experiment. This is a 1 row, 18 range experiment meaning that there are 18 plots in the y-axis all aligned along 1 row in the x-axis, so in total 18 plots constitute this experiment. Each plot within the experiment is planted with eight rows of the same genetic material. Given the setup of the experiment, plot names are simplified to only include the plot number, as such, plot 1 refers to row 1, range 1 and the same applies for all 18 plots for this experiment.

Biomass sampling occurred in the field approximately in two-week increments. Each sampling took place within one day of flight operations, so that the biomass sampled can be correlated with the imagery collected during the flights. The sampling procedure required the cutting of the plants from a specific location in the experiment and placing them in labelled mesh bags that were later placed in ovens for drying. After the drying process, the samples are weighed to obtain the weight of dry matter. Biomass collection is done using rows 2, 3 and 4 of each plot as mentioned earlier. Collection starts from row 2. The first 20 cm of the row were skipped, and then 50 cm were cut from the row for biomass collection. For the next data collection, the next 50 cm were skipped and the 50 cm after those were sampled. This process was repeated for each row until the end is reached, and no more locations from which 50 cm of plant biomass can be harvested; for this experiment, a row could be sampled three times. After row 2 was sampled, row 4 was sampled in a similar manner. When row 4 was fully sampled, sampling shifts to row 3. Sampling of row 3, unlike rows 2 and 4, starts 20 cm plus 50 cm from the end of the row so that sampling occurs in areas where rows 2 and 4 are undisturbed. Figure 2 shows the order in which each sample was taken. The first collection date was 30 days after planting, and a sample was taken every two weeks after that. In total, there were seven sample dates, four before maturity and three after. Biomass samples were dried for seventy-two hours in the oven set to 80ºC, which resulted in a sample with constant weight and no moisture. The samples were dried as soon as possible are harvesting since continued respiration could result in additional carbon loss, which would result in lower weights.

LAI measurements were collected in the field whenever biomass sampling occurred. Initially, LAI measurements were collected from rows 6 and 7 of each plot, which were also used for the UAS sensor measurements as those rows are undisturbed and no destructive sampling occurs in them. The LAI-2200C Plant Canopy Analyzer (Danner et al., 2015) was used for the measurements. The canopy analyzer estimates LAI by measuring the sunlight that is received by the device while the rod is placed under the canopy. The LAI readings were taken diagonally from row 6 to row 7. This was repeated twice in the plot, once for each of the two diagonals shown in Figure 1. Given issues with the canopy analyzer later in the season, around August 8th, 2019, leaf area (LA) was measured instead using the LI-3100C Area Meter. The Area Meter is basically a scanner in which individual leaves are scanned, and the total area of the leaves measured can be determined. Unlike the canopy analyzer, this is a destructive sampling method, so it could not be used on rows 6 and 7. Instead, area meter measurements were made using the leaves obtained from the biomass sample, so the leaves being measured are from rows 2, 3 and 4 rather than rows 6 and 7. During biomass sampling, leaves are separated from the plant stem for each sample, scanned, and then returned to the bag to be placed in the oven. To correlate between the LAI and leaf area values, leaf area was first converted to LAI by dividing the leaf area by the area of the plot that was sampled for biomass (Forest Ecosystems (Third edition), 2007).

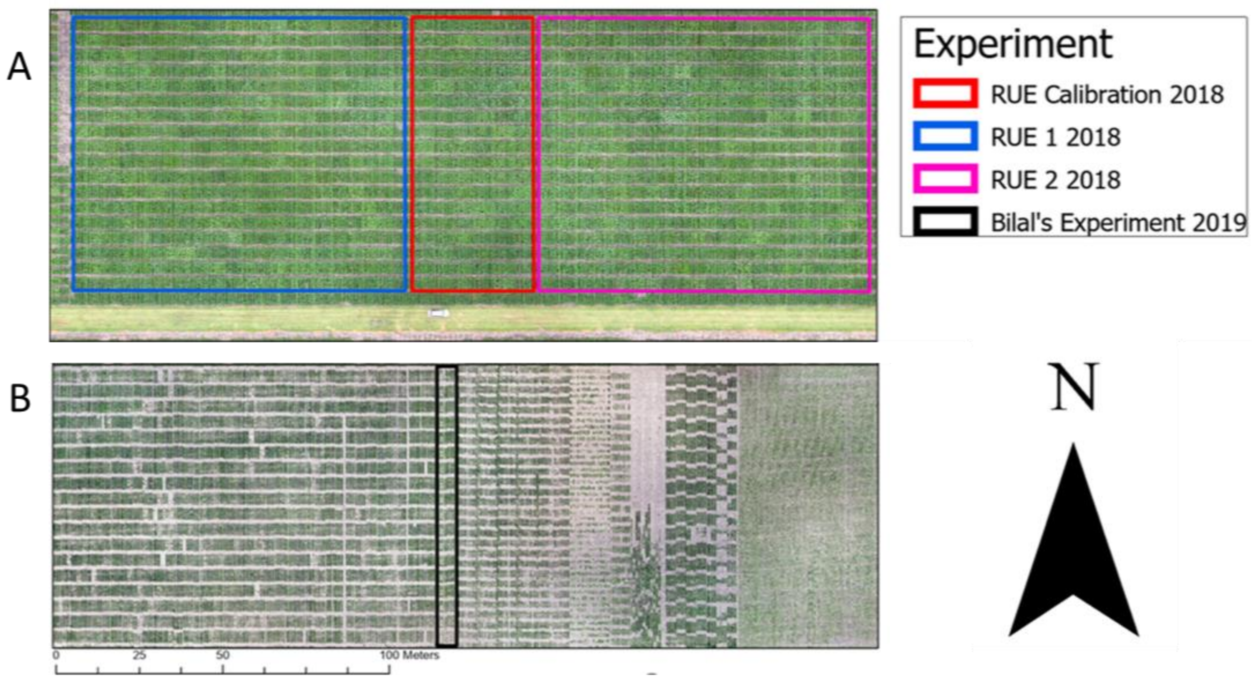

Figure 1: Experiment locations where panel A shows the 2018 experiments, and panel B shows the 2019 experiment. The three 2018 experiments are in the north field and the 2019 experiment is in the south field

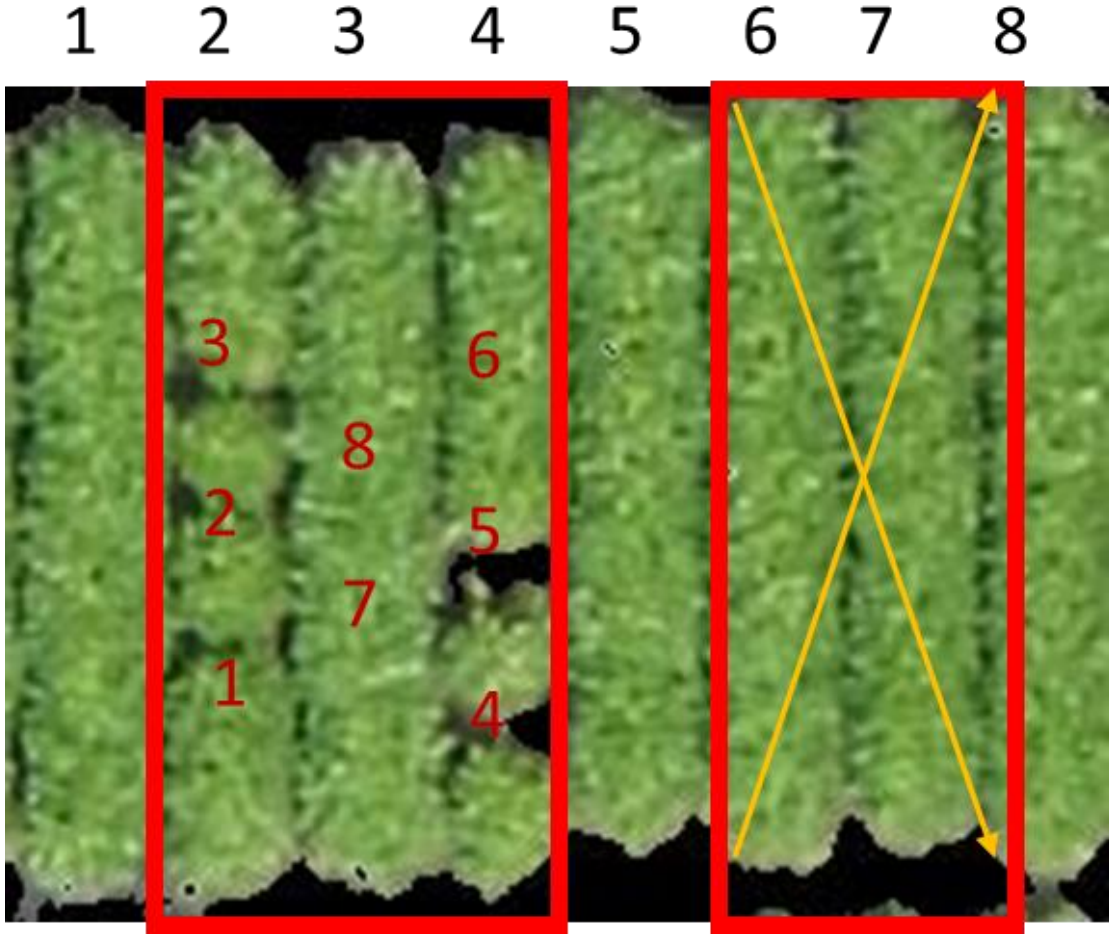

Figure 1: Biomass and LAI sampling procedure for the 2019 experiment. Image illustrates a single experimental plot within the field experiment. The top black numbers indicate the row number within the plot (each plot has 8 rows of the same genetic material), the numbers within the canopy show the order in which biomass sampling is conducted, and the diagonal lines illustrate how LAI measurements were taken using the LAI-2200C Plant Canopy Analyzer. This image was taken on August 1, 2019, after 5 biomass samples were collected, so gaps can be seen in the crop canopy for sampling locations 1-5. Neighboring plants expand into the canopy holes over time, thus the largest gap in the canopy is visible at location 5 where sampling just occurred.

Several criteria were considered prior to placing the 2019 experiment in its specific location in the field. The first consideration was that the experiment is not on the edge of the field as to avoid any edge effect on the plots (Langton, 1990). The second consideration was to avoid areas in the field that experienced ponding during the 2018 season, as prolonged saturation can severely lower productivity of soybean plants (Scott et al., 1989). This resulted in the experiment being placed in the center of the field in 2019, and its location within the field can be seen in Figure 2.

2.3 UAS Data Collection

The field in which the UAS data are collected was divided into multiple experiments as shown in Figure 1. In order to distinguish the experiment locations in the UAS images, ground control points (GCP) were installed on the edges of the experiments. The ground control points are 1-meter by 1-meter square panels and consist of a white center and a black border. Both the black and white material are made of cloth. The colors of the GCPs and their shape makes them distinguishable from their surroundings. In addition, the coordinates of the GCPs were captured using a TOPCON RTK (Real-time kinematic positioning) (Topcon, Tokyo Japan). The GCPs remain in their location for the entire season and are removed once the field season is over.

UAS data were collected throughout the planting season. Flights were conducted before planting to assess the flight plans for each field. In addition, these flights aided in checking if the installed ground control points (GCP) for the experiments were in the proper location in the images. Flights were generally performed close to solar noon, with the UAS system flying at an altitude of 120 meters. Overlap between images was kept consistent during each field season. In 2018 the forward and side overlap were set to 85% and 70%, respectively, this was changed in 2019 to improve plot extraction and was set to 90% and 90%, respectively.

The imagery collected was from two sensors, a Red Green Blue color scheme sensor (RGB) and a Multispectral sensor (MSP). The RGB data were collected using a S.O.D.A. camera (SenseFly Parrot Group, Switzerland), while the MSP data were collected with a 1.2 MP Parrot Sequoia camera (MicaSense Inc., Seattle, USA). The MSP has four discrete spectral bands: green (central wavelength = 550 nm, bandwidth = ± 20 nm), red (660 nm, ± 20 nm), red-edge (735 nm, ± 5 nm), and near-infrared (790 nm, ± 20 nm) (MicaSense Inc., Seattle, USA). Both sensors were attached to a SenseFLy eBee unmanned aerial vehicle (UAV). It is important to highlight that each sensor was flown on a separate flight with the RGB sensor flown over the field first, followed by the MSP sensor.

2.4 UAS Data Processing

The data collected by the UAS were individual images of parts of the field experiment as shown in panel A of Figure 3. These images needed to be flattened, rotated and stitched together into a single orthomosaic image with correct earth coordinates, a process that was completed using Pix4D mapper (Pix4D SA, 2018). Within Pix4D mapper, the images were flattened and stitched to form an orthomosaic of the field from the individual images. The flattening is crucial for the MSP images as the sensor has a Fisheye lens that significantly distorts the images (Bellas et al., 2009). The output of this procedure can be seen in panel B in Figure 3 where the images are stitched to form a single orthomosaic.

The orthomosaics generated for each field by Pix4D mapper included multiple experiments as shown in Figure 2. To extract data on a plot scale for each experiment, two pipelines were built within MATLAB: Crop Image Extraction version 2 (CIE 2.0) and Vegetation Indices Derivation version 1 (VID 1.0) (Lyu et al., 2019).

CIE extracts plot images from designed field experiments using RGB and multispectral (MSP) imagery captured by a UAS. The user configures the designed field experiment into CIE by providing metadata such as experiment location, number of rows and ranges, and size of plots. Once the experiment design is configured for the tool, CIE segments imagery to separate canopy from soil. The segmented images are used to accurately and precisely identify plot midpoints which enable the automated and rapid extraction of plot images similar to the ones shown in panel C in Figure 3. CIE works entirely in MATLAB and can run batch processes on Linux computer clusters. In addition, CIE allows for the extraction of multiple repetitions for the same plot. This is achieved by providing the original images captured by the sensors during the flights. Given that the Parrot Sequoia camera has a Fisheye lens, the flattened images from Pix4D mapper are used instead of the raw imagery. Both the flattened raw images and orthomosaics are used as inputs for CIE.

VID is the next step in the data processing pipeline. VID uses image attributes (e.g., row, range, date, image band) with customized functions such as band algorithms to quantify phenotypic traits from the extracted plot images. An automated and efficient trie structure is implemented in VID and allows for rapid processing of multiple images for each experimental plot (Lyu et al., 2019). For instance, an experiment of approximately 400 plots takes less than 30 seconds to process for each vegetation index being evaluated. VID can also calibrate spectral images resulting in a conversion from a digital number to reflectance values using several methods. The tool also can be used to extract certain rows within the plot (e.g., the middle 4 rows of an 8-row plot) and calculate vegetation indices only for the specified subset of the plot thus removing edge or sampled rows from the analysis. Any combination of spectral bands and supplementary information can be built into a VID equation, and results can be output from VID as text files or individual images such as the NDVI images illustrated in panel D of Figure 3.

Initially, this setup was developed for RGB imagery. When multispectral imagery was integrated into the system, it was evident that radiometric image calibration was needed to output accurate and meaningful results from VID, whether it be reflectance values or vegetation indices. This was resolved by developing a tool that would allow for the calibration of images before reflectance and vegetation index calculations.

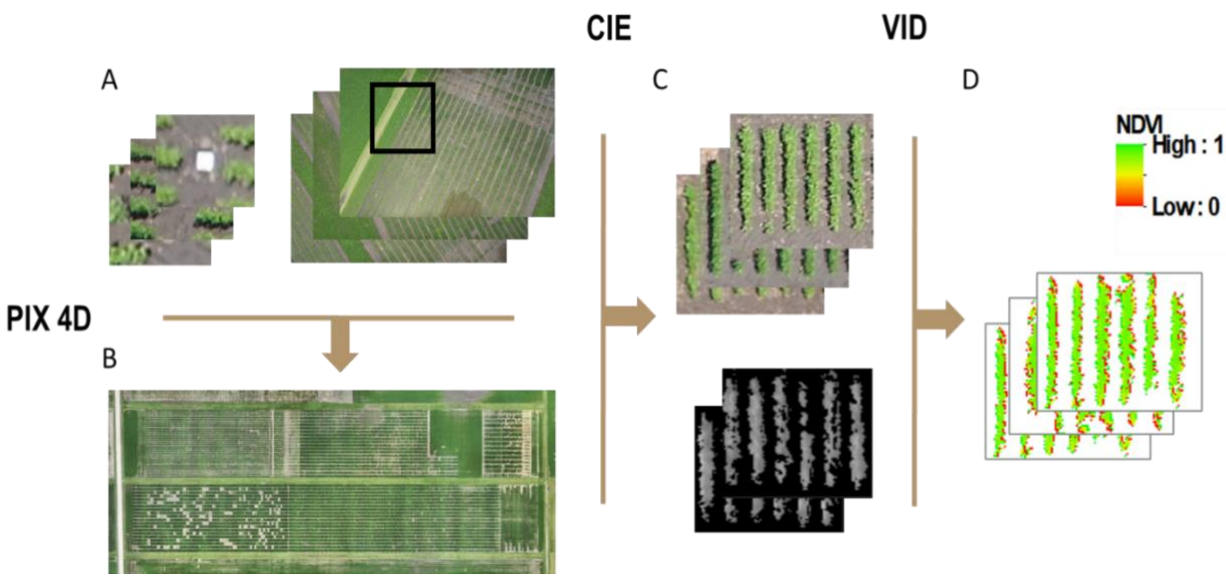

Figure 1: UAS data processing pipeline where panel A is a representative sample of raw images captured during flight operations. The raw images are used as inputs to the Pix4D Mapper software to generate the orthomosaic shown in panel B. Both the flattened raw images and orthomosaic are provided as inputs to CIE and used to generate the segmented row images and then extract each plot from the field experiment as shown in panel C. Outputs from CIE are used as inputs to VID and used to compute multiple vegetation indices, as illustrated with the sample images of NDVI in panel D.

2.4.1 UAS Image Calibration

Radiometric calibration is required to minimize the effects of atmospheric absorption and scattering on the reflectance values captured by the UAS cameras. Differences in lighting and atmospheric conditions can make it difficult to compare uncalibrated images between flight dates or even between different times of the day. Calibrating the images can remove atmospheric effects and potentially correct for any sensor sensitivity issues (Iqbal et al., 2018). Atmospheric calibration was not part of the original image processing pipeline, so this section describes the introduction of a supplemental step, the image calibration (IC) tool, where the images can be corrected for the effects of the atmosphere.

The radiometric calibration was carried out by utilizing a simplified empirical line method (Smith & Milton, 1999). The two factors considered for the empirical line method were the digital number (DN) of the raw images along with the reflectance for the plots from the UAS imagery (Iqbal et al. 2018). The relationship can be represented by the following simple linear equation.

$$Reflectance = slope \ \times DN \ \pm intercept \qquad (1)$$

The calibration process works on the individual plot images extracted by CIE and utilizes five unique spectral panels reflecting a specific and consistent percentage of light throughout the light spectrum (7% 12%, 22%, 36%, and 48% reflectance). These panels were chosen as the expected reflectance profile of the canopy should be between 7% and 12% for the red and green bands, between 22% and 36% for the red edge band and between 36% and 48% for the Near Infrared band (Bai et al., 2016). The 7% panel was used in the 2019 flights only as it was purchased just before the 2019 field season began, and as such was not available in the 2018 field season. The panels are laid out on the side of field before each flight and need to be visible in the final orthomosaic in order to be used for the calibration process. The digital numbers (DNs) for each panel for each camera spectral band are extracted from the orthomosaic. Mean DN can be different for the same spectral calibration panel depending on atmospheric conditions, camera orientation and the change of sun angle with the time of flight. A handheld spectrometer ASD FieldSpec ® 4 (ASD, Boulder, CO, USA) was used to measure the true reflectance of the panels while the multispectral images were collected. However, early in the 2019 season, the ASD broke, and a handheld GER 1500 (GER 1500 Spectra Vista Corporation, Poughkeepsie, NY, USA) was used instead to record the reflectance over the panels. It is important to highlight that these spectrometers have different spectral ranges and resolutions. The GER 1500 does not measure the near infrared and has a coarser spectral resolution (15 nm) compared to the ASD (1 nm). The spectrometer measurements were considered crucial for this experiment as there was concern that the panels might deviate from the intended reflectance under field conditions.

When the MSP sensor was flown, the following sampling procedure was followed for both spectrometers, when they were available. The sampling starts with three measurements over a white Spectralon panel along with one observation with the cap over the sensor of the spectrometer. The white panel was designed for full (100%; white) reflectance over the full range of both spectrometers, whereas the cap closed observation was used to simulate no reflectance (0%; black). After the initial calibration recordings, three measurements were collected over one of the spectral panels followed by three measurements over the white panel to recalibrate. After recalibration, the spectrometer was moved to the next spectral panel and another three recordings were taken over that panel followed by three over the white panel. This process was repeated until all spectral panels were sampled. In addition, the full procedure was repeated twice, once at the beginning of the flight and once at the end of the flight which totals six measurements for each panel for a single flight.

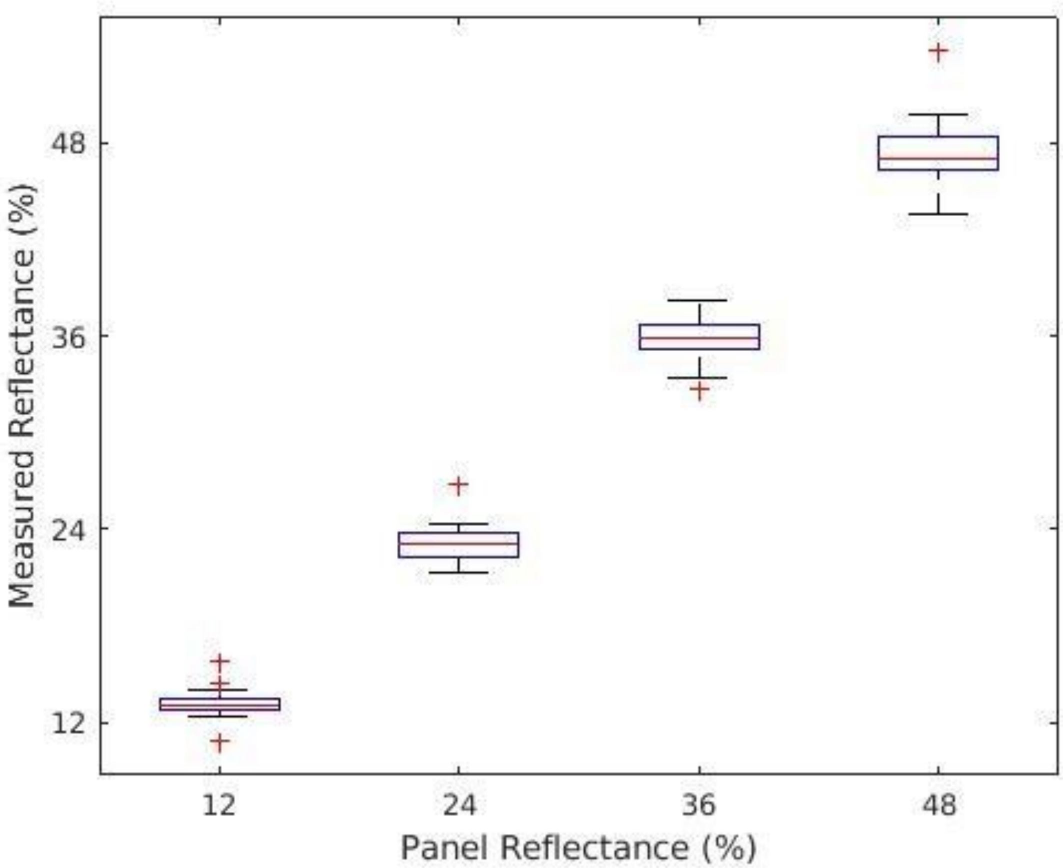

Figure 2: Boxplot of measured reflectance as a percentage for each spectral panel, for all 2018 flights, where the middle red line is the median, the box boundaries are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol.

Figure 4 illustrates the spread of reflectance values as measured by the ASD for all 2018 flights. From Figure 4 it can be seen that variability does in fact exist for different measurement dates for the same panels, but that variability is substantially smaller than the difference in reflectance between the different calibration panels. The low number of outliers and the fact that the values of the panels do not intersect is crucial as that clearly shows the robustness of the reflectance characteristics of the panels. In addition, Figure 4 illustrates the reliability of the panels across multiple flights. This is shown as the number of outliers is minimal with the varying atmospheric conditions between flights.

2.4.2 Image Calibration Tool Verification

In developing the image calibration tool, multiple different approaches for the model were assessed to determine the optimum procedure that would produce results closest to the true field reflectance. These approaches were compared to the reflectance values extracted from the orthomosaic produced by Pix4D mapper, the expected value as measured by the spectrometer over the crops and expected soybean reflectance as observed in other literary work (Bai et al., 2016). Three different approaches to building the models were considered: (1) a simple empirical line method (Simple ELM) where no additional data were added to come up with the regression lines relating DN to reflectance, (2) an altered empirical line method where the line is forced to pass through the point of zero reflectance and zero DN, (0,0), and (3) an altered empirical line method which added the point of zero reflectance and zero DN but did not force the line through it. The addition of the zero point was to simulate the reality of how the multispectral sensor and spectrometer work even though 0% reflectance is typically not found in the field. At zero reflectance the multispectral sensor regardless of band measured should output a DN of zero. Three dates were utilized to test these equations; the results are summarized in Figure 5.

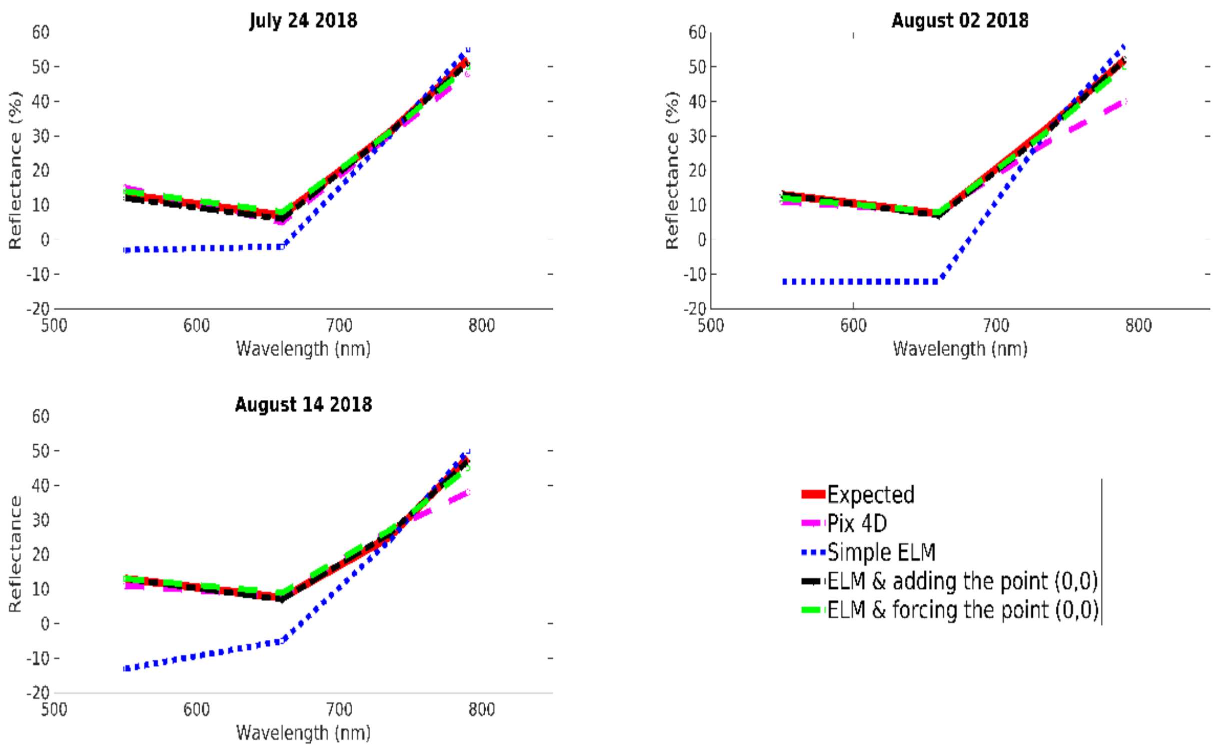

Figure 3: Image Calibration tool development, in which 3 different approaches are tested for 3 separate dates, where the x-axis is the wavelength in nanometers, and the y-axis is reflectance in %, the red line showing the expected (measured) reflectance values, the pink line showing the reflectance values as extracted from Pix4D's calibrated mosaics, the blue line showing the regression results if the simple ELM was used, the black line showing the regression results if the altered empirical line method which added the point of zero reflectance and zero DN was used, and the green line showing the regression results if the altered empirical line method which forced the line to pass through the point of zero reflectance and zero DN was used.

The Simple ELM produced regression lines that resulted in a negative value for both the Green and Red bands. This is due to the sensitivity of the sensor for the Green and Red bands being high, as compared to the RedEdge and NIR bands for the Parrot Sequoia (MicaSense Inc., Seattle, USA). The Parrot Sequoia sensor produces 16-bit images that are optimized for agricultural imaging, as such the sensitivity to green wavelengths is augmented relative to the other parts of the spectrum. By observing the canopy pixels within the green band, it is evident that the values are higher than what is to be expected. The canopy should reflect between 8% and 12% (Bai et al., 2016), which is equivalent to a DN between 5400 and 7800 for the Green band, however the canopy in the raw images from the Green band had a DN close to 33000 on average, which corresponds to 50%

reflectance. An object that would reflect at 36% reflectance in the field would show as saturated (65536 DN) in the Green Band. If the point (0,0) is not added, the regression lines would produce negative reflectance values for the Green and Red bands as the top value would be approximately 25000 DN rather than the full DN value 65536 based on the full bit depth. The sensor sensitivity depending on band can be observed in Figure 6 which was retrieved from Parrot Sequoia's official documentation (Micasense, 2016).

Figure 5 illustrates the results of using the three different methods in developing the regression lines for the reflectance values. The closer the estimated value is to the expected value, the more viable the method. In terms of the RedEdge band (730 nm to 740 nm), all methods produced results fairly similar to the expected values; however, by observing the results in Figure 5, it is clear that the Simple ELM is not an ideal option as it produces negative values in both the Green and Red bands for the three flights. The reflectance values were also extracted from Pix4D composite mosaics to compare those to the reflection values produces by the calibration approaches. The Pix4D extracted reflectance values are also not the best option as the NIR band reflectance is significantly lower than that of the measured value for two of the three flights, which is due to a combination of sensor sensitivity and calibration method utilized by Pix4D. The inclusion of the point of zero reflectance and zero DN eliminated the unwanted negative values for the Green and Red bands. Figure 5 illustrates the importance of including the point (0,0) in the empirical lines developed using this method. For all three dates tested, adding the point (0,0), while not forcing the line to pass through it performed better overall which can be linked to the sensitivity of the sensor. As such the IC tool uses only the simplified empirical line method after adding the point (0,0).

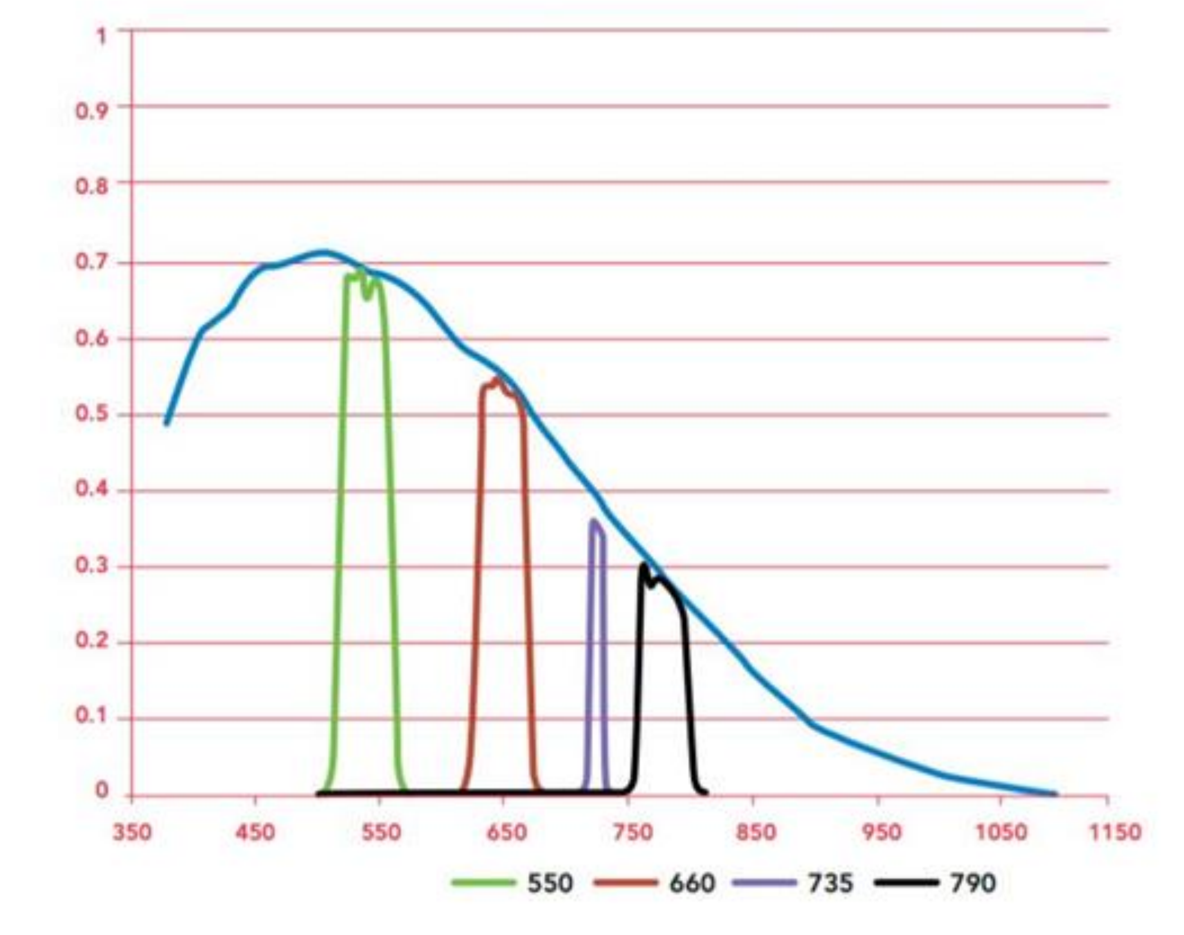

Figure 4: Parrot Sequoia's sensor sensitivity to sunlight for each band, where x-axis is wavelength and y-axis is reflectance (Micasense, 2016).

The IC tool is currently used to generate the regression lines used for calibration for all flights. The process is fully automated and compatible with the spectrometers used in this experiment. The code is flexible enough to allow the introduction of new spectrometers with minimal effort. The correction of images to calibrated reflectance occurs within VID when output from the IC is passed to it. While the IC uses only one method to generate a linear regression calibration line from ground spectral targets, it can generate two versions of the calibration line: one utilizing the spectrometer readings, and one utilizing the given spectral panel reflectance values. The former is considered more accurate, but the latter allows the user to perform image calibration if there is no spectrometer available. A comparison between calibrated reflectance from both methods was performed to evaluate the impact of not using spectrometer measurements in the process. These comparisons were performed during flights where spectrometer data were collected, and the IC was run generating the two distinct regression lines. The results are shown in Figure 7 as kernel density function (KDF) plots showing the difference between the reflectance values when using the spectrometer and when using the spectral panel values.

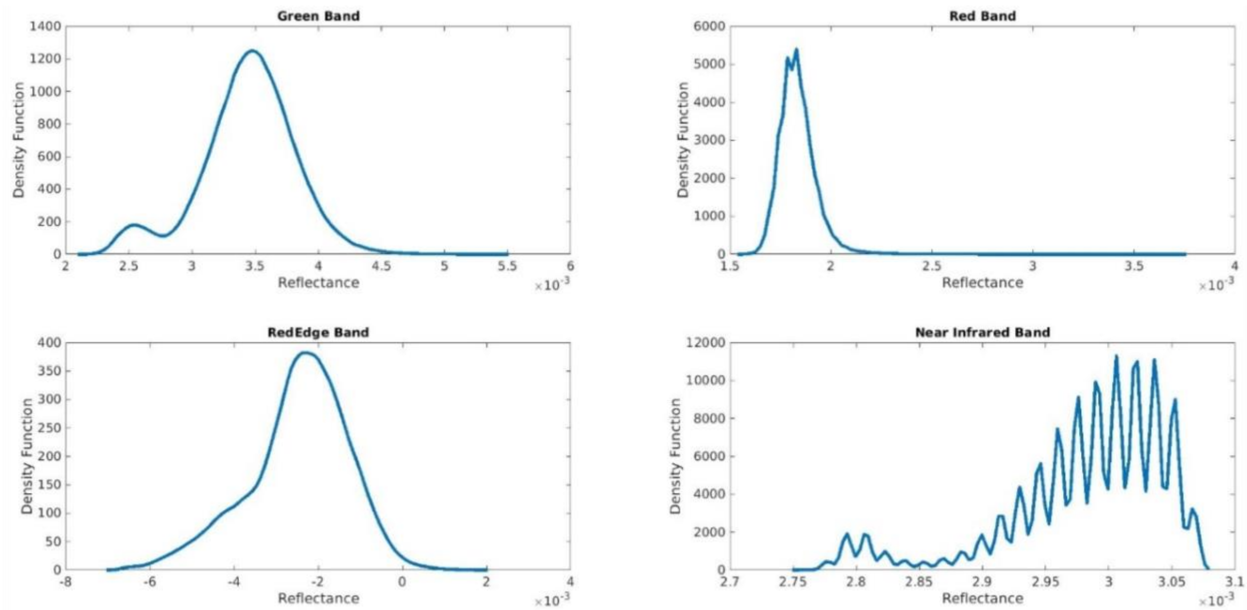

Figure 5: Kernel density function (KDF) produced by subtracting the pixel values produced when utilizing the spectrometer by the pixel values produced without the use of the spectrometer values, for all canopy pixels for all RUE calibration 2018 flights, for the Green, Red, RedEdge and NIR bands, while utilizing the 12%, 24%, 36% and 48% spectral panels.

By observing Figure 7, we can assume that using the given spectral panel values might be adequate given the low difference produced between both methods. The biggest difference for a single pixel (0.68%) occurred in the RedEdge band. The mean differences in reflectance for the Green, Red, Red-edge, and NIR bands were 0.38%, 0.20%, 0.22% and 0.29% respectively, which is close to negligible. An important observation is that the regression lines created while using the spectrometer data produced higher reflectance values for all bands except the RedEdge band, on average 0.01 reflectance for the Green, Red and NIR bands. The reason the RedEdge band did not show this trend has to do with the 22% reflectance panel. The 22% panel has been affected by dirt and dust and as such was not performing as expected, a problem discovered midway through the season. Given the fact that the panels are affected by dirt and sunlight conditions, it might be appropriate to utilize the spectrometer to accurately calibrate the images, rather than rely solely on the panel reflectance values, even though the improvements are minute enough, 0.1% reflectance for all bands, to where the given spectral panel values might be used if a spectrometer is not available. Sample output regression line for the method utilizing the spectral bands, for the NIR band, is shown in Figure 8. Only one method is shown as the regression lines are similar.

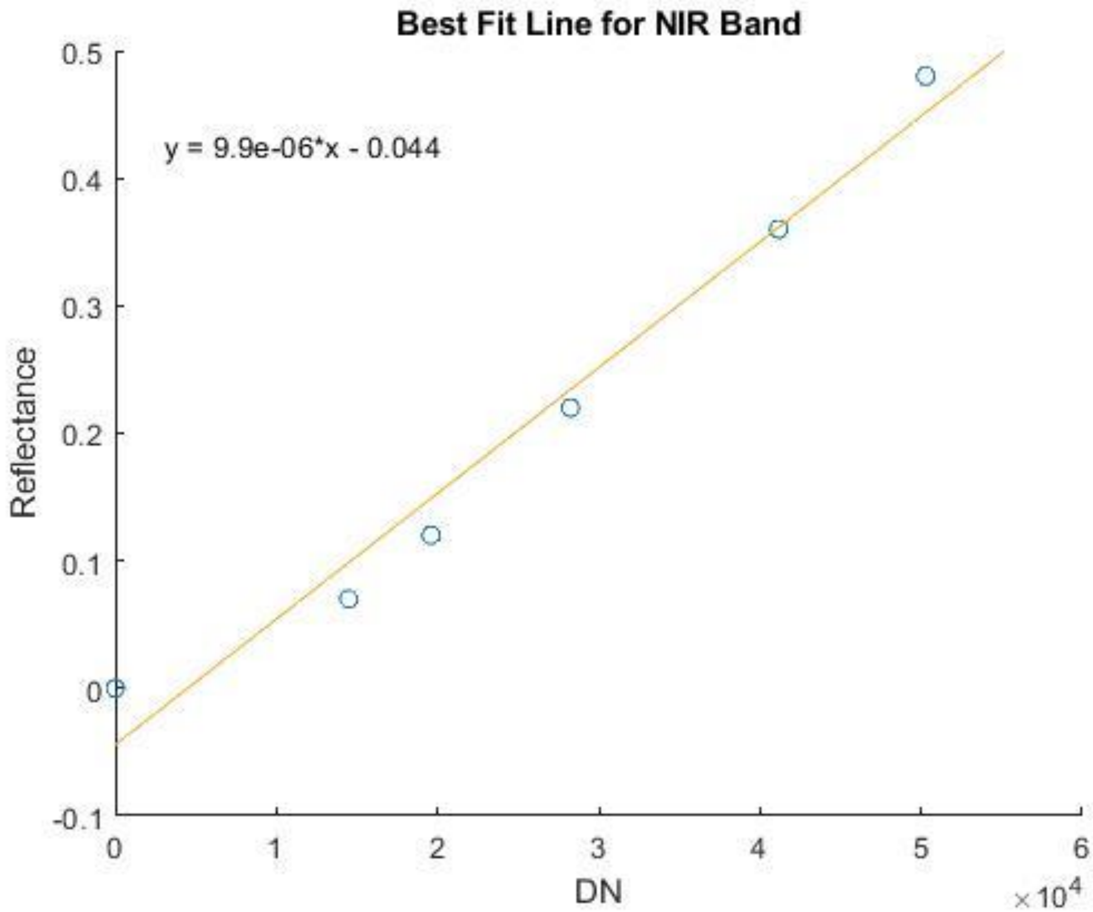

Figure 6: Sample output of the best fit line for the image calibration tool for the near infrared band for calibration utilizing the spectral bands as output by IC tool.

2.5 Flights Used for Image Analysis and Simulations

Flight operations were conducted throughout the growing season. Unless otherwise noted, RGB and MSP flights were collected for the same experiment on the same day. Exceptions were typically caused by problems with a specific sensor or the platform. The flight data used for this analysis spans from the V1 soybean growth stage, or first Trifoliate in which one set of unfolded trifoliate leaves form until R7 growth stage, or the Beginning Maturity in which some pods have reached their mature pod color. Before V1, the sensors are not able to capture greenness as the size of the crops is smaller than the resolution of the sensors. Beyond R7 is senescence which causes two problems. The first is that CIE cannot accurately extract the plots from the orthomosaic or images given that the coloration of the crops trends towards that of the soil (crops turn brown and loses leaves) which eventually makes them indistinguishable from the soil. The second problem is that after senescence biomass estimation models can no longer rely on NDVI and NIR values as the crop is no longer green and no longer transpiring. Instead, biomass is contained in the stems, not the leaves, which cannot be sensed using the same method developed for earlier in the growing season.

Initially 2019 flights were to be used for this study, however the 2019 field season had multiple complications. The first was the late planting date (mid-June rather than mid-May) which was caused by an especially wet spring resulting in substantial flooding in the field. The second complication was the visual stress on the crops caused by a drier than normal summer. Given that the crop parameter calibrator, which will be further expanded on in Chapter 3, requires datasets that are not stressed, that ruled out the use of the 2019 observations, so data from 2018 were used instead. In terms of the 2018 flights, seven dates on which flight operations occurred were used. The dates spanned June 18 to August 13. The flight dates are summarized in Table 2. All three RUE experiments highlighted in Figure 2 were used, as such only one flight per sensor is required to collect the image data. In terms of the spectrometer readings, only one run is required for each flight date during the MSP flight. In terms of spectrometer data used for IC, all flights had spectrometer data collected outside of July 17 where the spectrometer data were not available. For that date specifically, the factory spectral panel values were used.

Table 2: Dates of flights used in the study as well as the spectrometer data collection days for the 2018 Data

Date of collection	RGB flight	MSP flight	Spectrometer Readings
June 18	Collected	Collected	Collected
June 28	Collected	Collected	Collected
July 04	Collected	Collected	Collected
July 09	Collected	Collected	Collected
July 17	Collected	Collected	Not Collected
August 01	Collected	Collected	Collected
August 13	Collected	Collected	Collected

3. HYDROLOGIC AND CROP SYSTEM MODELING

3.1 VIC Model Overview

The Variable Infiltration Capacity (VIC) Model is a land surface macroscale hydrologic model (Hamman et al., 2018; Liang et al., 1994). The VIC model simulates multiple components of the water and energy balance including base flow generation, cold-season processes, evaporation, runoff generation, transpiration components and water movement in soil (Markert, 2017). The resolution at which the VIC model runs is variable, a grid cell resolution can range from 1/16th to 2 degrees latitude and longitude. The temporal resolution at which the VIC model runs is also variable and can range from hourly to daily. The VIC model can be run to process a small area which constitutes one grid cell (or point) or can be run to simulate multiple grid cells representing a large basin. To run large basins, the user needs to divide the area of interest to multiple grid cells of a consistent resolution. As of version 5 of VIC, there are two modes to run the model as it pertains to the space time continuum (Hamman et al., 2018). The first mode runs space before time. In this mode, for each time step, the model runs one individual grid cell for a single time step, and then moves on to the next cell until all cells are done for that time step. The model continues until all time steps have concluded. The second mode is time before space. In this mode, the model runs all the time steps for an individual grid cell, and then moves on to the next grid cell until all grid cells are simulated. In this study, version 4 of the VIC model is used and as such time before space is utilized. For the model to run, a minimum of three input types are required: meteorological data, soils data and land-use data. Each input is specified for the grid cell being run. Figure 9 shows a schematic overview of the process occurring within the model. Precipitation (P) is a driving factor as it is the only form of water that enters a grid cell. Precipitation that reaches the grid cell can have multiple outcomes. One possible outcome is the water that is captured by the canopy and then released as evaporation (Ec). Water that reaches the surface can infiltrate the soil (i), leave the grid cell as overland flow (R), or leave as evaporation from the soil, lakes or wetland (E). Water that infiltrates the soil surface might also leave the grid cell through transpiration by vegetation (Et), or as baseflow (B). Land cover is crucial as it determines the division between runoff and infiltration which is represented by the VIC model's namesake, the variable infiltration capacity curve. The curve represents the variability in infiltration rate across a large area such as a VIC model grid cell relative to the soil properties of that grid cell. Figure 9 illustrates the most common application of the VIC model with three soil layers. Layers 0 and 1 affect infiltration and surface runoff, baseflow comes from Layer 2 and vertical interflow is simulated between each of the soil layers (Liang et al., 1996). In theory the model can support an unlimited number of soil layers (Cherkauer et al., 2003) as well as vegetative classes for each grid cell.

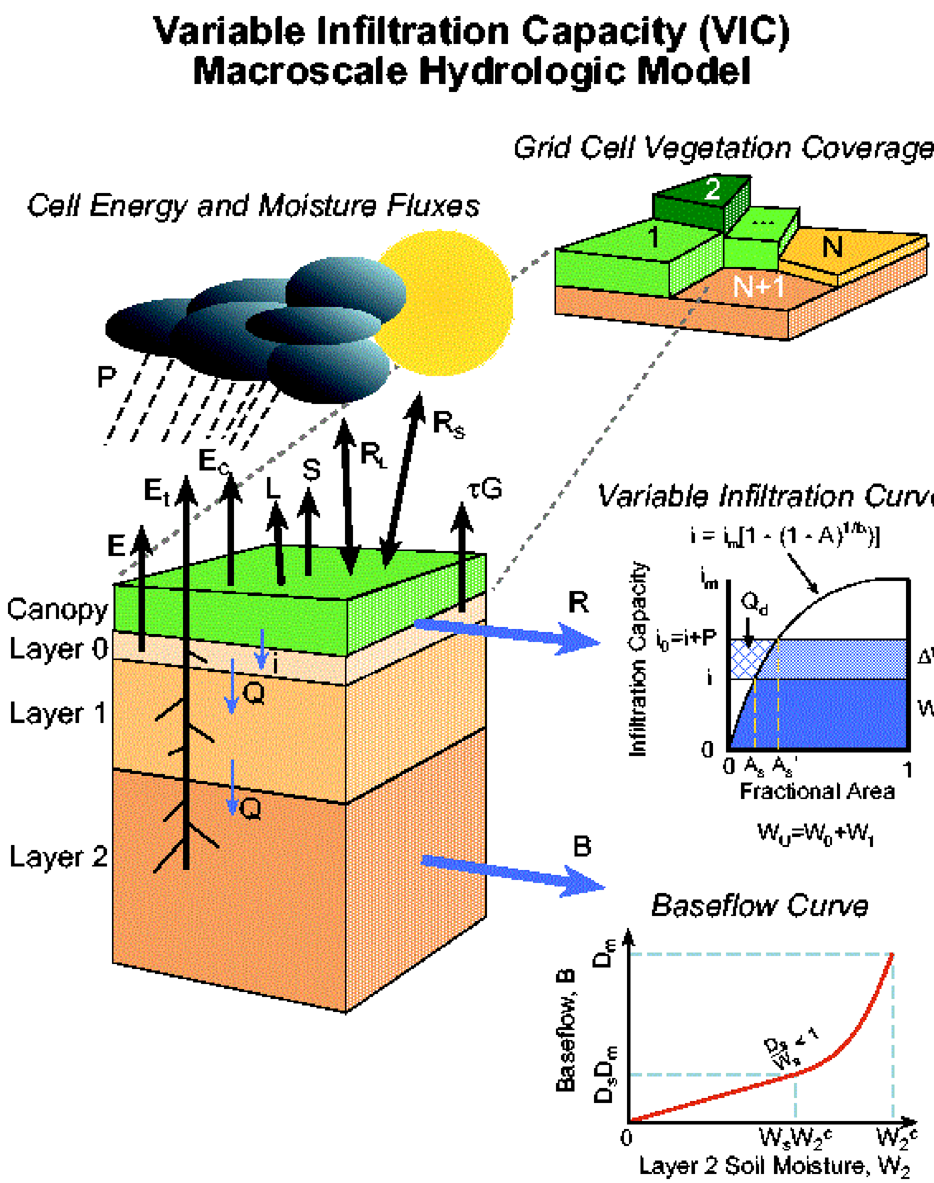

Figure 7: Schematic representation of the VIC model (Cherkauer et al., 2003)

3.2 CropSyst Model Overview

CropSyst is a cropping system simulation model which was initially developed to study the impact of climate, management and soils on crop productivity and the surrounding environment (Stockle et al. 1994; Stöckle et al. 2003). CropSyst does this by simulating crop growth, crop development, the nitrogen budget, residue production and soil water as well as other natural mechanisms on a daily time step. CropSyst allows for the simulation of multiple years and multiple crops under different management and rotation decisions. The model's primary purpose, however, is to simulate crop growth. Within CropSyst, crop development and growth depend on thermal time necessary to reach different growth stages (Stockle et al. 1994; Stöckle et al. 2003). The crop continues growing until it reaches maturity. Crop growth is expressed as biomass accumulation and is dependent on three factors: available radiation, available soil water, and available nitrogen. Potential growth for each day is estimated using both the intercepted plant available radiation (PAR)-dependent biomass growth (Donatelli et al., 1997), and potential crop transpiration-dependent biomass growth (Donatelli et al., 1997). Actual biomass growth is limited by transpiration-limited biomass growth (based on availability of soil water), and nitrogen-limited biomass growth (Pala et al., 1996). Each of these factors contributes to crop growth (biomass accumulation) as illustrated in Figure 10. This study assumes no Nitrogen limitation, so the components of CropSyst related to nutrient uptake and use are not discussed further. For this study, when calculating biomass accumulation, the two determining factors are crop potential transpiration-dependent biomass production and intercepted PAR-dependent biomass production. Initially, the model calculates the potential transpiration dependent biomass growth and PAR dependent biomass growth. The crop potential transpiration-dependent biomass is growth of the canopy dependent on transpiration, while the intercepted PAR-dependent biomass is the canopy development dependent on sunlight and radiation. The lower of the two values is then selected as the potential biomass growth as it is limiting the growth rate. Biomass is accumulated on a daily basis until the end of the season when biomass stops increasing either due to the maturation of the crop or harvest of that crop. Finally, yield is calculated by utilizing the biomass output and a harvest index.

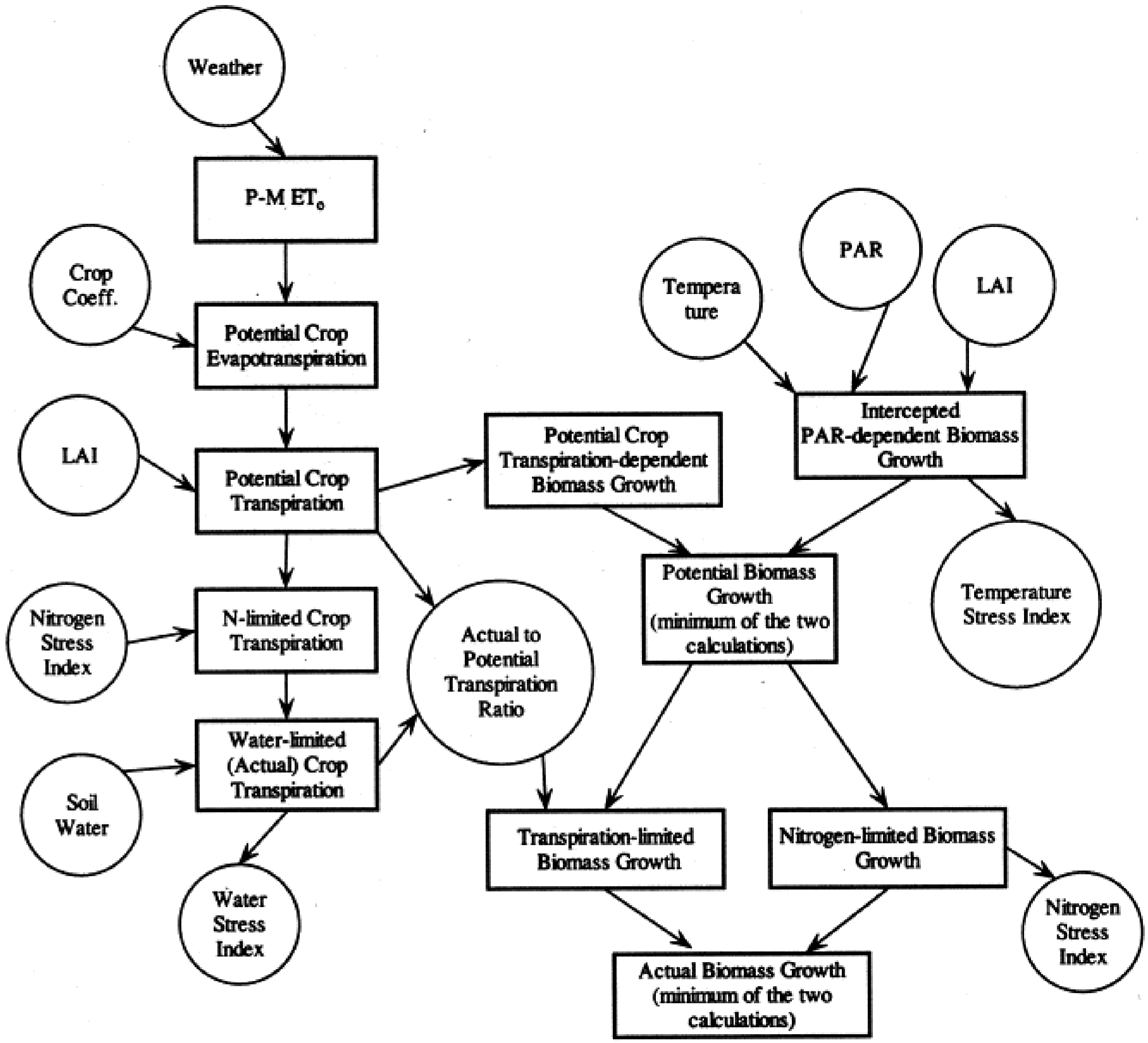

Figure 8: Schematic Representation of Biomass accumulation in CropSyst (Stöckle et al. 2003)

Intercepted PAR-dependent biomass production, GR; (kg/(m² day)) is the growth of the crop depending on the light intercepted and temperature. It is calculated as follows:

$$G_R = LtBC \times PAR \times FCCgreen \times T_{lim} \tag{2}$$

where *LtBC* (kg/MJ) is a coefficient that represents the conversion of *PAR* to aboveground biomass, *PAR* (MJ/(m² day)) is the photosynthetically active radiation assumed to be half of the total solar irradiance (CropSyst's Web Manual, Above-ground biomass accumulation). Solar irradiance is obtained from local weather data or estimated by the model if not provided by utilizing a utility tool within CropSyst Suite (Marcello Donatelli et al., 2003). *FCCgreen* is the fraction of incident *PAR* intercepted by the green canopy. *FCCgreen* is dependent on LAI which is determined internally as the crop develops. T_{lim} is the temperature limiting factor. This factor is a correction for radiation dependent growth based on radiation use efficiency. It is important to highlight that this factor does not impose heat stress, and is related to amount of light intercepted. The value of this factor depends on the actual air temperature versus the optimum air temperature in which the crop is intended to grow as well as the base temperature below which the crop stops growing. The value for T_{lim} is defined as:

$$\mathrm{T}_{lim} = \begin{cases} 1 \text{ for } T_{avg} > T_{opt} \\ \frac{T_{avg} - T_{base}}{T_{opt} - T_{base}} \text{ for } T_{opt} \geq T_{avg} \geq T_{base} \\ 0 \text{ for } T_{avg} < T_{base} \end{cases}$$

Where T_{avg} is the mean air temperature, T_{opt} is the optimum air temperature for growth and T_{base} is the base air temperature.

Crop potential transpiration-dependent biomass production, G_T (kg/(m² day)), is the growth of the crop depending on the actual available water. It is calculated as follows:

$$G_T = T_r \times \frac{BTR}{VPD} \tag{3}$$

Where Tr (m) is the actual transpiration or as defined within CropSyst, the crop water uptake. CropSyst assumes that there is no crop water storage within the leaves and stems; BTR (kPA/m) is the above ground biomass transpiration coefficient a value that is determined through parameterization of local crops; and VPD (kPa) is the daily mean vapor pressure deficit.

Biomass production on any day is controlled by the lower of the values G_T and G_R, except when VPD is greater than the maximum allowed VPD, which is set as a crop parameter, in which case G_R is chosen regardless of value, since transpiration is essentially shutdown. Equations 3 and 4 reflect the impact of radiation and water limitation on daily biomass production. Crop growth is represented by biomass accumulation over all days from planting until maturity at which time the daily accumulation of biomass stops. Maturity is controlled by crop specific parameters.

The biomass production is crucial as it feeds into the final crop yield. Currently there are two ways in which crop yield is calculated. The first is yield based on stress from flowering and grain filling periods and the second is yield based on translocation, which is the movement of materials from leaves to other tissues throughout the plant. Both are calculated by CropSyst and the greater yield value is considered the final yield achieved.
Yield based on stress from flowering and grain filling periods, is the first yield calculated within CropSyst as follows:

$$Yield = Biomass_{hrv} \times HI \times \left(1 - avgstress_{f}^{sf}\right) \times (1 - avgstress_{gf}^{sgf}) \quad (4)$$

Where *Biomass*$_{hrv}$ (kg/m²) is the total cumulative biomass at harvest; *HI* is the harvest index; *avgStress*$_{f}$ is the mean water stress index during the flowering period; *avgStress*$_{gf}$ is the mean water stress index during the grain filling period; *sf* is the harvest index adjustment parameter for water stress sensitivity during the flowering period; and *sgf* is the harvest index adjustment parameter for water stress sensitivity during the grain filling period.
Yield based on translocation, is the second yield calculated in CropSyst and is as follows:

$$Yield = Biomass_{flower} \times translocation \times avgstress_{f} \quad (5)$$

Where Biomass$_{flower}$ (kg/m²) is the total biomass achieved at the flowering stage; and translocation is a crop specific parameter.

Crop specific parameters are defined in separate crop parameterization files, one file for each type of crop being simulated. This allows CropSyst to represent many different crop types during a single simulation and allows the user to customize parameters to best reflect the varieties and management practices representing their study location.

3.3 VIC-CropSyst Model Overview

The coupled VIC-CropSyst model is utilized to run the simulations as part of this study. It is a coupling of the VIC model and CropSyst, in a spatially explicit manner (Malek et al., 2017). VIC-CropSyst-v2 which is utilized in this study, is a coupling of VIC version 4.1.2-e and CropSyst-v4.1.5 (Malek et al., 2017). All hydrologic processes excluding transpiration from defined crops are handled by the VIC model in this coupling. CropSyst handles crop growth and management practices as well as transpiration from crops.

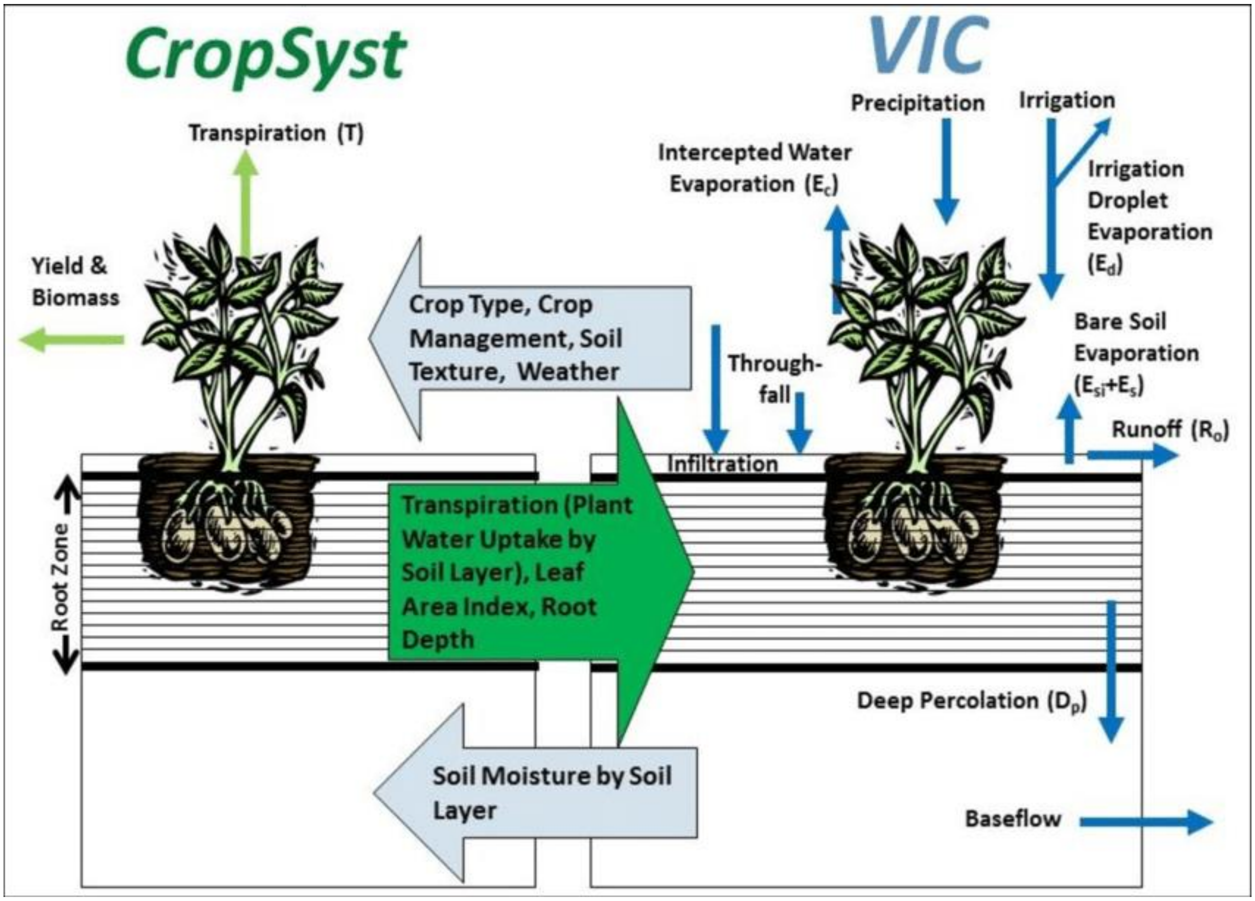

Figure 9: Schematic representation of the interaction between the VIC model and CropSyst (Malek et al., 2017)

Figure 11 illustrates how different processes are handled within the VIC-CropSyst coupling. The VIC model is first utilized to simulate the land surface energy balance and partition available energy into different energy flux and storage components which are explained in greater detail by Cherkauer et al. (2003). Any remaining energy is then determined to be available for potential evapotranspiration. Bare soil evaporation (E_s) and evaporation for water intercepted in the canopy

(E_c) are handled within the VIC model. Most agriculture in Indiana is rainfed and does not rely on irrigation, so we neglect irrigation sources and sinks, such as evaporation of irrigated water from bare soil (E_{si}) and from sprinkler droplets (E_d).

Once the VIC model has finished its calculations for the simulation time step, it passes potential transpiration and soil moisture content within each of its soil layers to CropSyst. CropSyst then calculates actual transpiration for each crop being simulated which in turn contributes to the estimation of crop biomass development and soil water extraction. After this, the amount of water extracted from the soil is passed back to the VIC model as well as updated LAI and simulated actual transpiration. This information is then used by the VIC model to continue its simulation and close the water balance for the grid cell. The simulation time step is limited to daily for VIC-CropSyst-v2, though most recent VIC applications are run at sub-daily time steps. It is important to highlight that soil properties are retrieved from the VIC model and not CropSyst; however, many soil processes within the VIC model have been modified to increase the coupling of the models. The most significant change is the number of soil layers represented within the VIC model. The VIC model can have three or more soil layers, but most applications utilize three soil layers. The first thin soil layer helps with the energy balance (Liang 1996), while the first two layers influence infiltration and baseflow is extracted from the bottom layer. Three layers are too course for CropSyst especially given that crop growth simulations are highly sensitive to soil moisture. VIC-CropSyst-v2 defaults to using 17 soil layers (Malek et al., 2017) in which the middle 15 layers are simulating the root zone where water uptake occurs. In the version of VIC-CropSyst utilized in this study, the top two layers are used layers for infiltration and runoff calculations, while all other layers are used for baseflow calculation, which aids in the simulation of the water balance. As CropSyst is focused on crop simulation, its internal simulation of water movement is simplified. By utilizing VIC-CropSyst, more accurate hydrologic modeling can be done which in turn will improve the results of biomass and yield estimates since there is a better simulation of the water than is available for the crops to be used as well as other hydrologic procedures (Malek et al., 2017).

3.4 CropSyst Crop Parameter Calibrator Overview

The Crop Parameter Calibrator is a program built into CropSyst's CropSuite, a set of tools that is distributed with the CropSyst model, and allows for the calibration of crop parameters to improve crop growth simulations for specific locations, management practices and varieties. The crop calibrator is crucial for this project as it will be used to set crop parameters based on soybean biomass measured destructively in the field and estimated from UAS imagery. The crop parameter files developed using the CropSyst calibrator will then be used within the VIC-CropSyst model to simulate soybean growth over multiple growing seasons, which will allow for the assessment of how different parameterization methods affect simulation of crops under local weather conditions. The Crop Parameter Calibrator is a standalone tool run separately from the VIC-CropSyst model. It requires inputs of planting dates, and biomass, LAI, and weather conditions with time during the growing season. Outputs from the calibrator include the above ground biomass-transpiration coefficient, the light to above ground biomass conversion factor, harvest index, leaf duration, and initial green leaf area index.

To parametrize the crop model using the Crop Parameter Calibrator, the user is required to provide biomass and final yield values. The biomass values were generated using the biomass estimation equation, discussed in Chapter 4, which utilized the flights from Table 2, as well in-situ destructive biomass samples. Yield values were measured in-situ after the growing seasons. Initially, LAI was collected in order to compare simulated and observed leaf development but given that the 2018 data were used to populate the model rather than 2019, and the Crop Parameter Calibrator requires the same number of LAI measurements as biomass measurements, that was deemed problematic as LAI was only collected on three dates in 2018. LAI was collected along with the biomass in 2019 multiple times, yet given the hydrologic conditions discussed earlier, the values in 2019 could not be used. LAI values were not further pursued, as VIC-CropSyst-v2 does not model LAI growth using the specific leaf area and stem/leaf partition coefficients, which are the outputs of the Crop Parameter Calibrator if LAI was included. VIC-CropSyst-v2 utilizes canopy cover development indices to model LAI development.

The crop calibrator currently consists of four steps or modules (CropSyst's Web Manual, Crop Calibrator) each that calibrates a specific component of crop growth for a location. It is recommended that the user run each step in the following order:

1) Step 1 - collect location specific information including weather. This step requires that the user provide weather files for the area in which the crop was grown. Soil is not needed as it does not have a direct effect on crop growth, but rather only governs the amount of water extracted from the soil layers, as such it is not needed here.
2) Step 2 – quantify the phenology of the crop. In this step the user provides information on the timing of phonologically important stages in the development of the crop being calibrated. Calibration can make use of approximate/typical seasonal values or specific dates collected from sites. If the latter, empirical data, are chosen, the user inputs the dates after planting at which major phenological events were observed for the crop, such as emergence date or the beginning of grain filling.
3) Step 3 - quantify the development of biomass with LAI. This step requires the input of LAI and biomass data collected from the crop in the field. As the crop parameters define optimal growth, the biomass and LAI data fed into the crop calibrator should be collected from unstressed crops. CropSyst will account for the effects of stress in simulations by reducing growth from the optimal path. This step is used to calibrate for specific leaf area and stem/leaf partition coefficients that control early growth.
4) Step 4 - calibrate biomass growth and yield. For this step the user inputs the final biomass and yield from a harvested crop. Preferably, this should be final (harvest) biomass and yield for multiple seasons to capture a range of conditions, which will in turn increase the reliability of the parameterization. For this study, the 2017, 2018 and 2019 data were used. The outputs of Step 4 are: the above ground biomass coefficient (BTR) that is used to calculate the transpiration dependent daily biomass growth (G_T), radiation use efficiency (LtBC) which is used to calculate the radiation dependent daily biomass growth (G_R), the unstressed harvest index ($HI_{unstressed}$) used for yield estimation, and specific leaf area and leaf duration which are used for LAI simulation.

3.5 Datasets Required For VIC-CropSyst

VIC-CropSyst is data intensive and requires data from multiple resources. VIC and CropSyst are coupled in this model and as such will share the same working space. The following datasets are required to run VIC-CropSyst:

1) Meteorological Forcing Files: The meteorological forcing file can be summarized as a weather file. One text file is required for each cell being processed. The weather files for this study contain precipitation, maximum and minimum temperature, and mean wind speed on a daily basis from 2000 to 2019. The weather data from 2000 to 2018 is a composite of weather station information from ACRE and Purdue's Throckmorton Purdue Agricultural Center (TPAC). A composite weather file was used as a significant amount of data from ACRE's weather station was missing and needed to be filled. The 2019 data were not corrected or filled and was retrieved from the REACCH weather data which is hosted on Northwest Knowledge Network. The dataset is a blend of spatial attributes of gridded climate data from PRISM, which is a dataset developed and maintained by PRISM climate group based in Oregon State University (Berteaux et al., 2006), as well as desirable temporal attributes from regional reanalysis using NASA's LDAS-2. The resolution of the grids is at 1/24th degree. Solar radiation is also required to run the model correctly, however the data are unavailable, as such it is to be estimated using the algorithm developed by Thornton and Running (1999), which is incorporated into the VIC model.
2) Soil Parameter Files: The soil parameter files used in this study were modified from Cherkauer et al. (2021) where they were used to simulate soil conditions based on drainage area of each grid cell and were parameterized for the three-layer model. The files were adjusted to simulate a seventeen-layer soil file where the soil is poorly drained. The top two layers were preserved from the original dataset where both were defined as being 0.1 m thick. The bottom layer was divided into fifteen layers, each at 0.083 m, to yield the 17 soil layers required for CropSyst. The higher number of soil layers helps limit water access to roots as they grow. If the root depth does not reach a specific soil layer, then moisture in that soil layer cannot be accessed.
3) Vegetation Files: The vegetation library and parameter files use in this study were also obtained from Cherkauer et al. (2021). The vegetation parameter file and library include

information regarding the attributes of the vegetation in each grid cell. For the VIC-CropSyst model simulations that default cropland class is replaced by crops defined using CropSyst's crop parameter files. For the crops simulated by CropSyst, default values in the VIC model vegetation parameter file and vegetation library will be replaced by those provided by CropSyst as it simulates their growth. The VIC model vegetation files are therefore mostly used to control non-agricultural vegetation that exists in the grid cell being simulated. In this study, the only crop considered is soybean, and it fills the entirety of the grid cell, so in practice we are simulating one field of soybean.

4) Crop Parameter Files: The crop parameter files contain all crop specific parameters required to simulate crop growth at the specified location. The crop parameter files used by the VIC-CropSyst model are transferred directly from CropSyst. The parameter files shared with CropSyst were developed by Stöckle et al. (2003) and represent default settings for multiple crop types in the U.S. The default soybean parameter file was used as a reference to compare how parameterization affects biomass and yield results. Modified crop parameter files created using CropSyst's crop parameter calibrator and two different sources of local measurements of crop development were also used in this simulation.

4. RESULTS

The focus of this study is to demonstrate the potential for using UAS derived phenotypes for parameterizing the VIC-CropSyst model for a field site in Indiana. Crop model parameterization using UAS derived phenotypes are evaluated in this chapter by (1) evaluating the relationship between crop biomass estimated using UAS imagery and ground reference samples of biomass; (2) quantifying changes in the parameters controlling the simulation of crop growth for different parameterization data sources; and (3) evaluating the results of a simulation experiment where the VIC-CropSyst model is used to predict biomass accumulation and yield at ACRE with the default and localized parameter sets.

4.1 Biomass Estimation Model

To accurately predict crop yield and biomass, the files used by the VIC-CropSyst coupled model to control crop growth must be parametrized for local environmental conditions and phenomics. The process of parametrization requires the collection of plant specific attributes (Biomass, LAI, yield, and timing of growth stages) that are input to CropSyst's Crop Parameter Calibrator (Chapter 3). These plant specific attributes are best collected in a representative field for the parametrization process to accurately represent crop physiology for the specific location of the simulation. Currently the majority of biomass measurements are done in-situ using destructive sampling methods that affect the canopy and the growth of plants around the sampling location. These destructive methods require a substantial number of man hours especially for large experiments which require researchers to balance the increased costs of a larger team with the amount of data it is possible to collect. For this study, a team of three working together was on average able to collect 100 destructive biomass samples in approximately three hours. The samples were then weighed, dried for three days in ovens, and then the dry weight was measured. Measurement of weights took approximately two hours to complete. In total, the procedure would take five days to complete.

An alternative for conventional in-situ biomass collection is the utilization of UAS based imagery to predict the above ground biomass of the crops. This process saves money by requiring the employment of fewer workers. It also reduced the time required in the field and in processing the samples, as the usual process requires a full day of sampling and three days of drying versus the images which can produce results within 24-hours utilizing the CIE-VID workflow described in Chapter 2. UAS imagery has been utilized previously to estimate canopy closure fractions and rates (Lopez et al., 2021) and improve genomic selection in soybeans (Freitas et al., 2021).

The equation used to estimate biomass from UAS imagery is based on the equation employed by Smith et al. (2021) to assess the relative change in soybean biomass from excess water stress. The biomass predictive model is defined as follows:

$$DM = (a \times NDVI^2 + b \times NDVI) + \left(\frac{NIR_{ref}}{Green_{ref}} \times \frac{NDVI_{max} - NDVI_{min}}{1 - NDVI_{min}}\right) + c \quad (6)$$

Where DM (g) is the estimated biomass in the field. NDVI is the normalized difference vegetation index, which is a value between 0-1 that is correlated with vegetation health. NDVI is a simple division of NIR – Red / NIR + Red (Zhou et al., 2009). NIR_{ref} is the mean plot near-infrared band reflectance value and $Green_{ref}$ is the mean plot green reflectance value. $NDVI_{max}$ is the maximum NDVI for all plots under investigation, whereas $NDVI_{min}$ is the minimum NDVI value for the same plots. Finally, a, b and c are parameters used to fit the model using multiple regression. For this study, these parameters are obtained from MATLAB's function to fit nonlinear regression models using the non-linear least squares method (The MathWorks, 2020).

The predictive model was developed in Smith et al. (2021) to quantify the effect of excess water stress on growth and is based on an earlier algorithm developed for satellite remote sensing imagery. Smith et al. (2021) modified it to make use of higher resolution imagery from UAS. For this study, we evaluate the model's ability to predict biomass at the experimental plot level for multiple flight dates. Accurate and consistent estimates of biomass through the growing season are critical for the model to produce data useful for parameterizing crops in the CropSyst model. In order to parameterize the model for a given location and crop, the model requires the collection of remote sensing imagery and biomass samples for the same plots.

Remote sensing data for fitting of the biomass model was extracted for two rows within each plot from the 2018 RUE experiments. For an eight-row experiment, that meant that rows 6 and 7 were utilized. Edge rows were not used to reduce interference from neighboring plots, and as rows 2-4 of the 8 row plots were used for destructive biomass sampling row 5 also counts as an edge row to minimize the effect of destructive sampling. The multispectral images were calibrated and then used to calculate NDVI for all image pixels. Cropped two-row images (undisturbed rows 6 and 7) were then used to calculate mean NDVI, NIR_{ref} and $Green_{ref}$ for each plot. Mean values of NDVI for all plots of interest were used to determine the maximum and minimum mean NDVI values for the field experiment. Biomass samples were collected for the calibration plots by physically removing plants within one day of UAS flight operations (see Chapter 2 for sampling methods). The sampled biomass of each plot for each date is then used along with the respective remote sensing data for that plot to calibrate the biomass estimation model. The use of multiple observation dates is different from the method employed by Smith et al. (2021), as they used only a single date of observations for model calibration and quantified only relative differences in estimated biomass between plots on other dates to confirm the effect of excess water stress. Because the CropSyst calibrator requires accurate estimates of biomass throughout the growing season, this calibration process will utilize all flight dates to fit a single model. This was done to allow for the prediction of biomass for flights where no biomass sampling occurred by using the output regression lines. A boxplot showing the distribution of biomass samples and date of samples, as days after planting, is shown in Figure 12. The boxplot clearly illustrates the difference in population means between samples within each experiment, which should create distinct populations especially for RUE I and II. Furthermore, differences in the timing of the samples might produce problems with the fitting of growth curves as there is data missing between the samples that might be crucial when modelling crop biomass growth. For example, only the RUE Calibration plot was sampled prior to 40-days after planting, resulting in the lowest sampled biomass values. That sampling was completed before the plots reached the R1 growth stage. For all plots, the spread in biomass values increases as the crop develops for all plots. The spread of biomass values is likely dominated by the fact that different plots come from distinct varieties of soybean which are expressing different phenotypic variance stemming from their genetic difference.

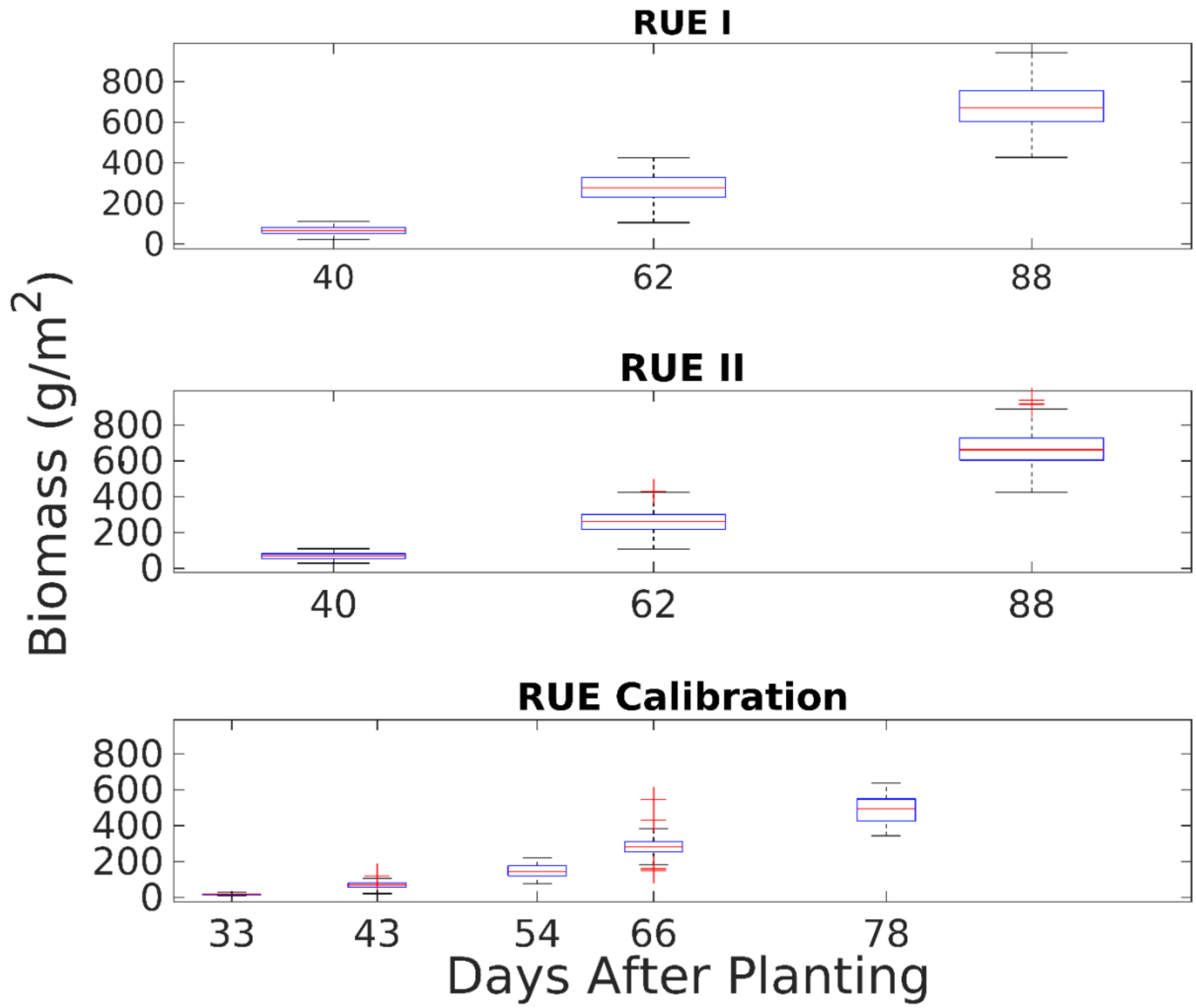

Figure 10: Boxplot of biomass samples for each experiment, where each boxplot is a population of sample done a certain number of days after planting, while the middle red line is the median, the box boundaries are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol.

4.2 Biomass Estimation Validation

In this section, two validation tests will be performed to assess the reproducibility of the biomass estimation model. If the biomass estimation is valid, it should be able to identify the differences between the distinct populations created by the sampling dates, and the distinct varieties used, where the values for one sampling date do not greatly intersect values from another sampling date. Cross-validation methods will be employed to evaluate the biomass estimation model performance. Cross-validation methods work by dividing the datasets into two parts, training data, and testing data. The training data are used to build the model, and the test data are used to validate the model.

The results of cross-validation are generally less biased or a less optimistic estimate of the model skill than other evaluation methods. The k-fold validation method (Anguita et al., 2012) will be utilized for this analysis. It divides the sample dataset into k unique equal classes. Each time the model is run, k minus one dataset is used to train the model and one dataset is reserved to test the model. The test dataset is unique for every run as it utilizes a different subset of the full data. A 10-fold validation was used for this study meaning that the test is run ten times, such that each of the 10 subsets was used for validation once while the model was built using the other 9 subsets (Refaeilzadeh et al., 2009).

The second test utilized for this analysis is the train/test split, otherwise known as holdout validation (Korjus et al., 2016). This method of validation randomly divides the dataset into 2 parts, a test, and a training dataset. The training dataset usually contains more data than the test set and is used to generate the model, which is then evaluated versus the testing data subset. This method is usually used for larger datasets where the method is run once. In this study a 70-30 holdout method will be rerun multiple times, where each run is using a new random training and test datasets and the spread of the values generated will be compared to the original model. The 70-30 method means that 70% of the dataset is used to establish the model, while evaluation makes use of the remaining 30%. The procedure was repeated 1000 times, where each application randomly divided the data into the two groups. The training dataset was approximately 735 points and the testing dataset consisted of 315points in each of the 1000 runs. It is important to highlight that the groups were randomly selected to create 1000 unique runs.

The model performance metrics used for the validations included the root mean square error (RMSE) (Korjus et al., 2016) as well as the 80% confidence interval (CI) of the coefficients produced by the biomass prediction model. The use of the 80% CI is due to the fact that intra-field yield variation is approximately 20% (Joernsgaard & Halmoe, 2003). If the results of the validation procedures produce coefficients within the CI and a RMSE value close to that of the original dataset, it is implied that the model is not overfitted. Results for both validation methods are presented in Tables 3 and 4, in addition the spread of the RMSE error can be visualized in the histograms in Figure 11. No visualization was produced for the k-fold validation as the experiments are only divided into ten groups. The validation was performed for RUE 1 and RUE

2 only. RUE calibration was excluded as the number of points was not sufficient to perform both validation tests reliably.

Table 3: Cross Validation results for RUE I experiment including Mean of the Model Coefficients in grams (Equation 6), as well as the 80% CI for the coefficients, for all subsets of the biomass sample dataset used for calibration and validation.

Test	Coefficient	Mean	CI
70 - 30 holdout (1000)	a	5028	NA
	b	-4904	
	c	1210	
k-fold (10)	a	5008	
	b	-4859	
	c	1189	
Full Dataset	a	5015	4614, 5433
	b	-4885	-(5425, 4369)
	c	1204	1043, 1373

Table 4: Cross Validation results for RUE II experiment including Mean of the Model Coefficients in grams (Equation 6), as well as the 80% CI for the coefficients, for all subsets of the biomass sample dataset used for calibration and validation.

Test	Coefficient	Mean	CI
70 - 30 holdout (1000)	a	4825	NA
	b	-4529	
	c	1080	
k-fold (10)	a	4833	
	b	-4539	
	c	1065	
Full Dataset	a	4888	4427, 5193
	b	-4569	-(4989, 4030)
	c	1071	929, 1219

The results of the validation tests indicate that the equation used to estimate biomass produces similar results using both validation methods, even though the number of days for which the data were collected is not uniform (Figure 12). Furthermore, the coefficients produced when using the validation tests for both experiments are approximately equal to the coefficients obtained when fitting the model using all outputs, which suggests that the equation produces unbiased results (Berrar, 2018). In addition, all coefficients fall within the 80% CI which further supports the use of the model as it does not over fit the data and is reproduceable.

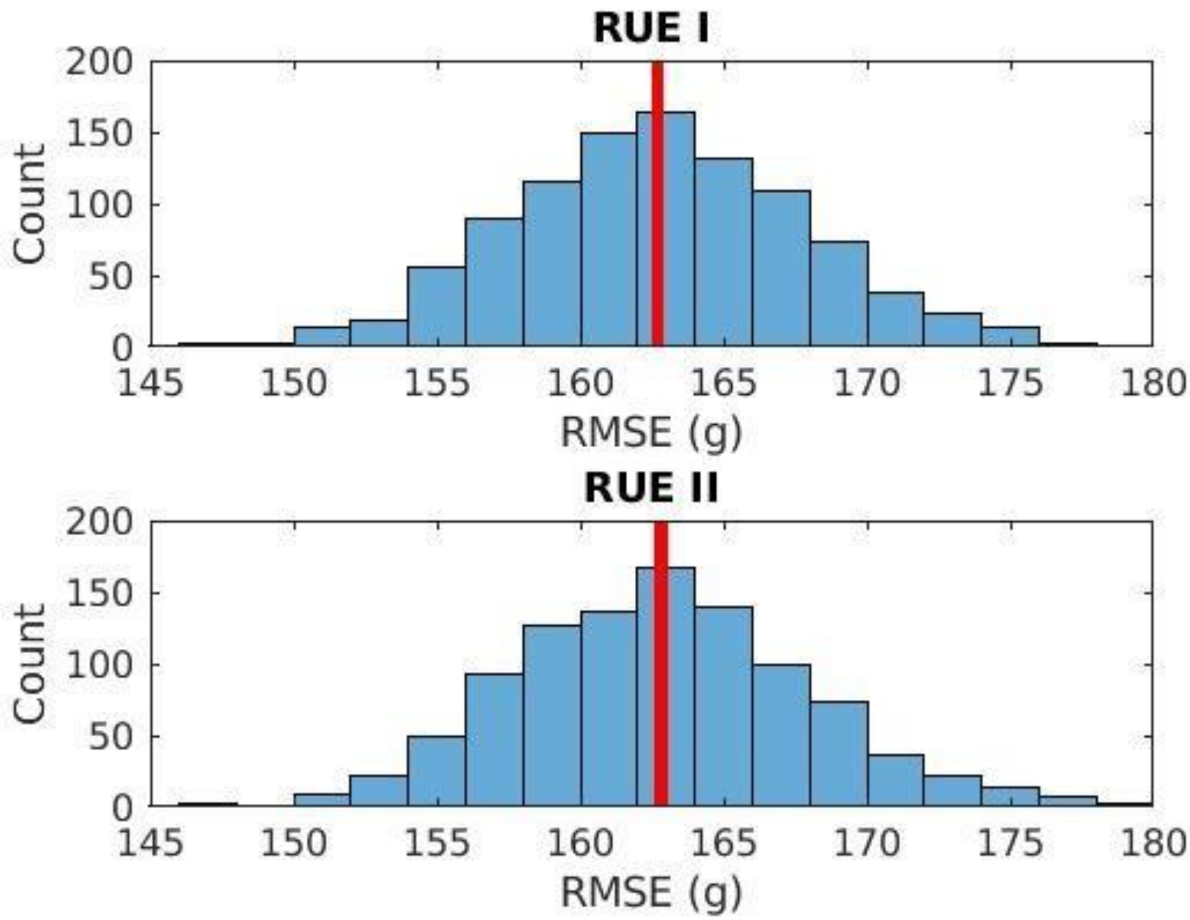

Figure 11: Histogram of Holdout (70-30) Validation for RUE I and RUE II RMSE, where the red line in the figure is the real RMSE when using the entirety of the data to create the model

The spread of the RMSE values within the histograms further supports the use of the biomass estimation equation as the RMSE values increase in count as they approach the mean value, which falls within the interval with the largest count. Furthermore, the spread of RMSE values follows a normal distribution for both validation methods, when looking at the histogram. This was further validated by running a Kolmogorov-Smirnov test on the datasets. The RMSE produced when using the entire dataset, 162 g and 163 g for RUE I and RUE II respectively, is similar to that of the mean RMSE for both validation experiments 163 g for both runs, which further supports the reliability of the model to produce the same results from different samples.

The results of the model using 70% of the data as a training set, was evaluated for each of the three RUE experiments in 2018, in addition a fourth evaluation utilizing all three experiments was conducted. The results are listed in Table 5 which includes the parameter estimates (a, b and c), the coefficient of determination (R-square), root mean square error (RMSE and the number of dates used for each. Furthermore, Figure 14 illustrates the results as a scatter plot comparing predicted versus observed values of biomass (g), where the 1:1 reference line is drawn to show overestimation or underestimation.

Table 5: Estimated Parameters for ACRE RUE experiments with coefficient of determination, root mean square error and number of dates used

Experiment	**a**	**b**	**c**	**R-square**	**RMSE (g)**	**Number of dates used**
RUE Calibration	4986	-4800	1075	0.75	87	5
RUE 1	5028	-4904	1210	0.63	162	3
RUE 2	4825	-4529	1080	0.65	163	3
All 2018	4958	-4705	1030	0.68	160	8

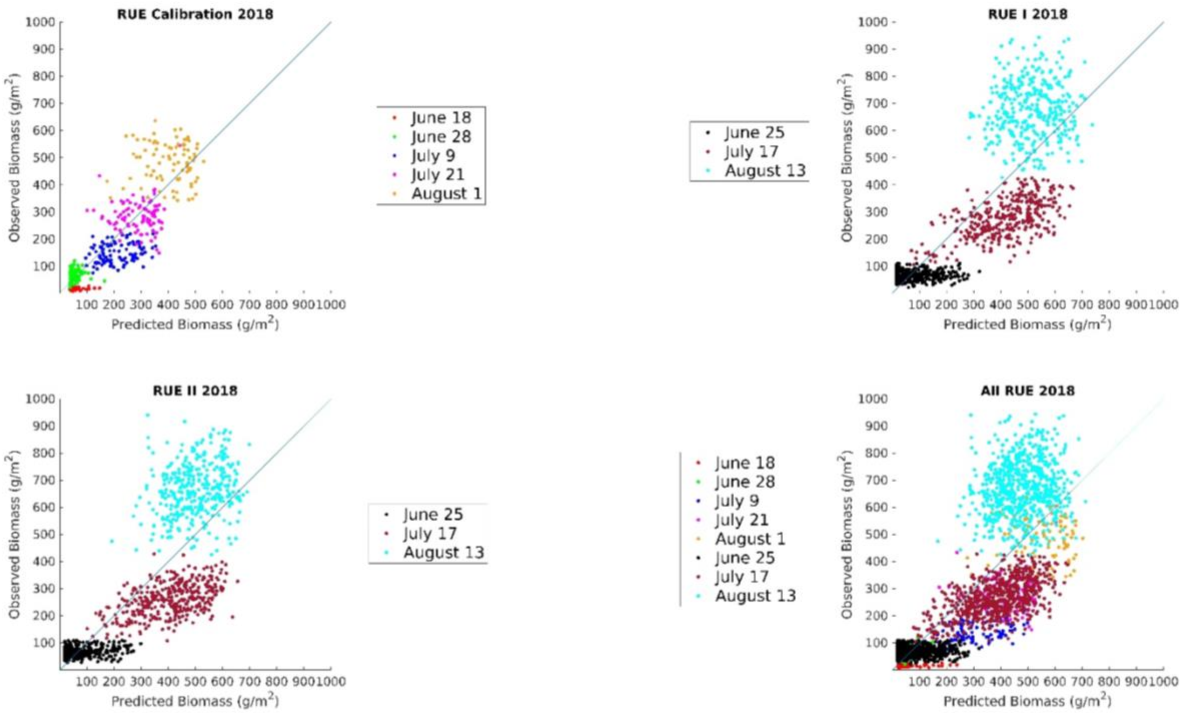

Figure 12: Biomass estimation results using only RUE Calibration, RUE I, and RUE II observations, and when using all observations. Dates indicate when data were sampled in the field. The diagonal is the 1:1 line.

Table 5 clearly shows that the regression line generated by using the data of RUE Calibration has a low RMSE as compared to the other regression varieties, which is due to multiple factors, First, RUE Calibration contains fewer genetic varieties (60 varieties) as compared to RUE I (350 varieties) and RUE II (350 varieties) which should lead to less genetic variation. The second reason is associated with the fact that more sampling dates (5 total) are available for RUE Calibration than for either RUE I or RUE 2 (3 dates each) which increases the number of points defining the curve. The third reason is associated with the actual dates at which the sampling took place, and their sparsity. The last sampling date for RUE Calibration was August 1, as compared to August 13 for RUE I and RUE II. The first signs of senescence appeared in mid-August for some of the plots, which could also explain why the observed biomass for these dates is higher than the predicted biomass. Senescence is important as the color of the canopy starts to turn yellow and can greatly skew our results.

When using all of the in-situ biomass measurements (Figure 14), some of the clusters overlap. This is an indication that the change in biomass between observation dates tends to be greater than that caused by genetic variability on a single sampling date. This is more evident when looking at only RUE I or RUE II as nearly no overlap exists between samples from different dates. Furthermore, the boxplots in Figure 12, for the separate sampling dates have close to no overlap. In viewing the RUE calibration data, it is clear that the early sampling date of June 18 does not fit the model, and it is assumed that this is because the crop canopy was still not sufficiently developed. The RUE Calibration plots were still at pre-R1 growth stage, which is when flowering occurs. As such, the June 18 flight was not utilized as part of the inputs to CropSyst's Crop Parameter Calibrator. The All 2018 figure which is the regression created by merging data from all 3 RUE experiments produces a slightly better fit than RUE I and RUE II, as the RMSE of All 2018 is lower while having a higher coefficient of determination (r^2). As such, All 2018 is the most suitable regression equation and will be used for the applications in this study. RUE Calibration was not considered as the dataset was relatively small and does not have the same amount of genetic diversity. For future applications, there might be a need for more data collection points, to produce more accurate regression lines as well as better planning of the sampling dates, where samples should be taken after the R1 growth stage, and before senescence, while maintaining a low number of days between samples, as to not create distinct populations, and have a continuous spectrum of values.

4.3 Crop Parametrization

In order the test the effectiveness of the UAS-derived biomass estimates, two sets of crop parameters were developed: (1) a set based on the UAS biomass model, and (2) a set based directly on the in-situ biomass samples. The UAS-derived biomass estimates used for the parametrization process were generated using the biomass estimation equation built on the 2018 UAS data (Table 2). The mean biomass of the plots for the entire field was generated for each of those dates that were used as input into CropSyst's Crop Parameter Calibrator for the calibration process described in Chapter 3 under the CropSyst Crop Parameter Calibrator Overview subsection. The output of this procedure was a new crop parameter file, that made use of the biomass estimates to tune simulated soybean growth to location conditions. The second parametrization utilizing the actual in-situ destructive biomass samples was completed using the same procedures as used for the UAS-derived biomass estimates, but the mean in-situ plot biomass across all varieties was used as the input for the parametrization process. The same LAI measurements, planting and harvesting dates were used for both parametrizations.

The results of the parametrization process are summarized in Table 6 and presented along with the uncalibrated (original CropSyst) parameters. The above ground biomass-transpiration coefficient (BTR) and Light to above ground biomass conversion (LtBC) are the primary parameters responsible for daily biomass accumulation. BTR is influential in the computation of the transpiration dependent biomass production, while LtBC is important in the calculation of the radiation dependent biomass production. If both values increased after parametrization that would indicate that biomass accumulation at the end of the season, under optimum conditions, would be higher when utilizing the parametrized crop parameter files. However, BTR increased, while LtBC decreased relative to the uncalibrated crop file for both parameterizations, which does not clearly indicate if there is a change in final biomass. As noted in Chapter 2, actual biomass accumulation in CropSyst is controlled by the lower of two values – radiation dependent biomass growth [LtBC] and transpiration dependent biomass growth [BTR]. A lower LtBC might indicate that radiation dependent growth is almost always the dominant control locally, thus the lower value results in radiation control on biomass production in most cases.

Three other parameters were also affected by the calibration process: the Harvest Index, the leaf duration and the initial green leaf area index. The Harvest Index also increased after both parameterizations from 0.30 to 035, which should result in higher overall yield values, as it is multiplied versus final biomass to obtain harvested yield. Leaf duration influences the length of time it takes for a crop to reach senescence. Within CropSyst, biomass stops accumulating when senescence is reached. As such, having a lower leaf duration (1100 versus 1200 degree-days) should lower the maximum potential biomass accumulation for both parameterization methods. The value of Initial green leaf area index will affect LAI development, which affects radiation dependent biomass production. Both parameterizations dropped from an uncalibrated value of 0.021 to 0.015. This parameter has an impact on the value of the FCC_{green} parameter, the fraction of incident PAR intercepted by the canopy, in each timestep. Decreasing the initial value means that it will take a little longer for leaves to capture as much radiation after parameterization.

When comparing between both parametrization methods, the primary difference is between parameters BTR and LtBC, which are concerned with biomass growth. All other factors yield the same values after calibration; thus, the source of biomass data does not directly affect LAI or yield. Both the calibrated BTR and LtBC values based on the UAS-derived biomass estimates are slightly higher than those of the files generated when using the in-situ sampled biomass values. BTR is 0.02 KPa/m larger, and LtBC is 0.01 g/MJ larger. The difference is low with respect to the overall base value; however, we should expect slightly, if not negligible, higher final biomass and yield values generated using the UAS estimated biomass. As such the use of UAS imagery to populate the crop calibrator tool is deemed appropriate as it provides approximately equal values to the case where the destructive biomass samples were used for the parametrization.

Another change for all scenarios (calibrated and uncalibrated) was to the planting and harvesting dates. By changing the planting and harvesting dates, the real field and experiment conditions can be simulated, as the model is localized for conditions at the site of the experiments. The planting and harvesting dates are not shown in Table 4.4 as they are kept the same for both runs, and they are not an output of the crop calibrator. The planting date was set to May 25, and the harvesting date was set to October 25.

Table 6: Parameter outputs of the parametrization process via CropSyst's Crop Parameter Calibrator as well the default files and description of the parameters

Variable	Uncalibrated	UAS-Biomass Calibration	In-Situ Biomass Calibration	Description
Above ground biomass-transpiration coefficient [BTR] (KPa/m)	3.50	3.71	3.69	This value represents the above ground biomass production per meter of transpiration under given conditions of atmospheric vapor density deficit.
Light to above ground biomass conversion [LtBC] (g/MJ)	2.50	2.25	2.24	This value represents the above ground biomass production per unit of light intercepted by the crop canopy. Radiation Use efficiency. In CropSyst the value is on a Photosynthetically Active Radiation (PAR) basis.
Harvest index	0.3	0.35	0.35	The ratio of yield to biomass for a crop without stress.
Leaf Duration (deg-days)	1200	1100	1100	This corresponds to the degree-days elapsed between the appearance and senescence of new green area index.
Initial green leaf area index	0.021	0.015	0.015	Initial green leaf area index.

4.4 VIC-CropSyst Simulation

This section introduces the results of the VIC-CropSyst model simulations using the new crop parameter files described in the previous section. As with the parameterization process, three scenarios will be evaluated: without parameterization and after parameterization using UAS-derived biomass estimates and using in-situ biomass, CropSyst parameters for each of these scenarios are summarized in Table 6. Evaluation of the simulations uses two datasets: (1) the in-situ measurements of biomass and yield measured during the 2017, 2018 and 2019 field seasons for all genetic varieties, and (2) yield values going back to 2007 for bulk soybean fields at ACRE. The first dataset allows for the assessment of biomass predictions in select years, while the second dataset will allow for the evaluation of yield from 2007 to 2019, including the drought year of 2012. A multi-year simulation allows for the evaluation of crop response to environmental variability. The results of this section are crucial as they are directly relevant to the hypotheses of this study.

The VIC-CropSyst simulation was run for each set of crop parameters shown in Table 6, while all other model parameters and inputs were kept constant. The outputs of each simulation are the daily biomass accumulation, LAI development across the season and yield at the end of the season. Evaluation of the simulated soybeans focused on multiple varieties for the years with additional data (2017-2019), and on final yield for all years of the simulation versus bulk soybean yields at ACRE.

When assessing the 2017, 2018 and 2019 experiments, one thing was clear, which was that yield after parametrization was always higher than that of before parametrization (Figure 13 and Table 7). This is expected given that both sets of calibration parameters had a harvest index greater than the uncalibrated parameter set. The in-situ measured yield is divided into two distinct values, one which is the mean of only the nine varieties picked for the 2019 experiment, and one which is the mean of the 30% training dataset for all the varieties, named all varieties, in the experiments. The distinction between the two in-situ yields is interesting as it sheds light on the variability, which was introduced by having different varieties. The difference between the yield of the nine- varieties chosen for 2019 and the yield of all varieties is small. This is to be expected as the criteria for the lines chosen for the 2019 experiment depended on observed yield from the previous season. The varieties in 2019 were chosen to represent different yield potentials, three low, three medium and three high yielding, which can be referred to in Table 2. The value that stands out was the 2018 yield estimation. Both crop files resulted in estimates that were significantly lower than the values measured in-situ. In addition, the 2018 results show the closest yield values before and after parametrizing. The reason this occured was linked to the T_{lim} value that was used in the radiation dependent biomass growth. One limitation of CropSyst is that it is simulated on a daily time step. Temperature on a daily basis was provided as a maximum and minimum, and within CropSyst, the mean temperature is derived from those two values. In 2018, there were many temperature fluctuations and colder periods that resulted in a drop in the radiation dependent biomass production. Furthermore, the parametrization process produced estimates that were closer to the in-situ measurements as compared to the results when the model was unparametrized. This can be seen from Table 7 as the yield after parametrization for all 3 years was closer to that of the in-situ measurements as compared to the yield produced by the model when not parametrized. This supports the first hypothesis that the development of relationships between crop indices and crop parameters is expected to improve the ability to simulate crop growth for local conditions. The 2019 varieties refer to the nine genetic varieties utilized for the 2019 calibration experiment as chosen from varieties used in 2018.

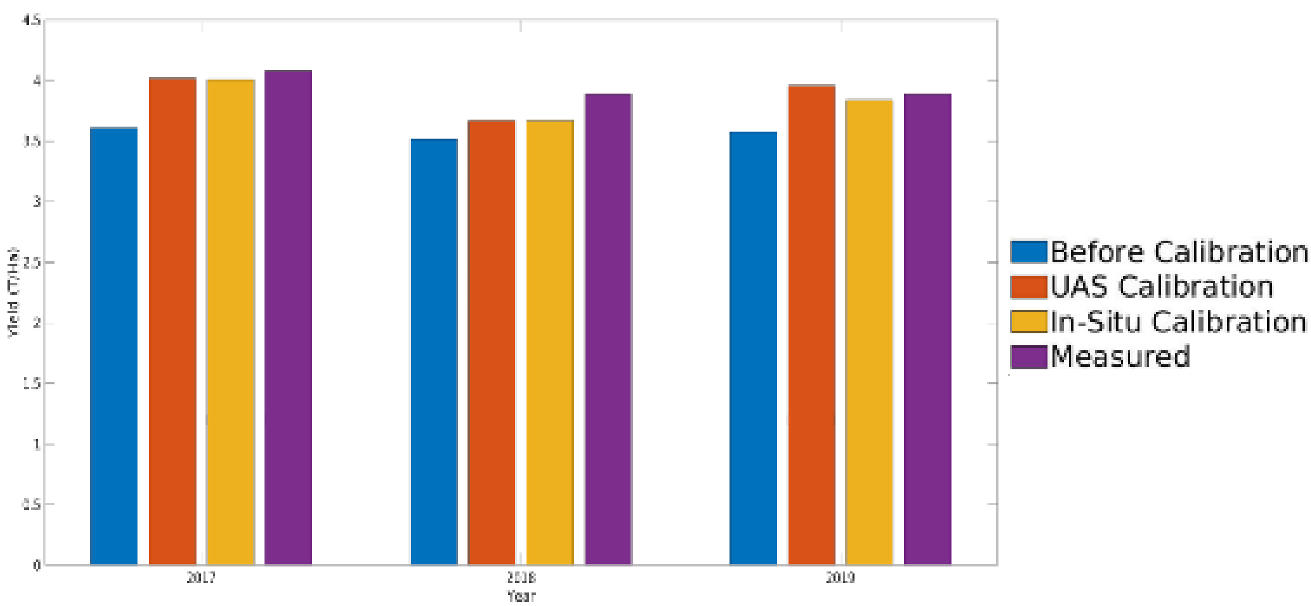

Figure 13: Measured and estimated yield values for 2017, 2018 and 2019

Table 7: Yield results for each simulation parameter set (uncalibrated, UAS calibrated, and in-situ calibrated) as well as the in-situ measured yield mean, median and standard deviation (SD) for the nine varieties selected for the 2019 calibration study (2019 varieties), as well as all varieties included in the experimental plot for each year's experiment (All varieties).

Year	Simulation mean yield (T/Ha)			2019 varieties yield (T/Ha)			All varieties (2018) yield (T/Ha)		
	Uncalibrated	UAS Calibration	In-Situ Calibration	Mean	Median	SD	Mean	Median	SD
2017	3.61	4.02	4.01	3.99	3.84	0.52	4.08	3.75	0.69
2018	3.52	3.67	3.67	3.96	3.92	0.42	3.89	3.9	0.59
2019	3.58	3.96	3.94	3.89	3.68	0.48	NA		

The mean, median and standard deviation of yields measured in the field from only the varieties included in the 2019 calibration experiment (2019 varieties) and for all genetic varieties in the full experiment (All varieties) demonstrate the yield variability that comes from the genetics (Table 7). Working with a more limited number of varieties resulted in less interannual variability in the mean and median yields values between 2017 and 2018, and a lower standard deviation. Varieties selected for the 2019 calibration experiment were not selected to be representative of the full extent of yield variance, so the lower standard deviation is expected. The variation in yield is also visible in Figure 15. Yields are relatively consistent in 2017 and 2018, while 2019 experienced a drop in yield. Yields for 2019 may be adversely affected due to plots 1, 2, 17 and 18 which had problems when planting that resulted in lower yields. Additionally, lower yields may also be due to the late planting season and unique hydrologic conditions that were witnessed in 2019 which had an extremely wet early season that delayed planting followed by a drought period. When the larger population of all varieties planted in 2018 is considered for the years 2017 to 2019 (Table 7), the abnormally low yields are less prevalent and instead the mean yield between years is found to be quite similar. The inclusion the 2019 experiment in the plot for all varieties is to further verify that the spread of yield values for the 2019 experiment is falls within the 2017 and 2018 yields.

Biomass development is harder to assess due to limited data collection. The only comparison of biomass development that can be properly evaluated is for the 2019 experiment. Within 2019, eight biomass measurements were taken for each of the eighteen plots in the experiment. The measured values were then averaged to obtain a single biomass value for the experiment. Plotting the mean biomass measurement with time yields the biomass development curve for the 2019 field season (Figure 14). Also shown is the simulated biomass accumulation for both the uncalibrated and UAS calibrated VIC-CropSyst simulations. As the simulation results from the in-situ calibration and the UAS calibration are virtually identical the remaining figures show only the results from the UAS calibrated simulations.

Only small differences in biomass were observed between the simulation runs (Figure 14). As the parametrization process increased BTR (ET controlled biomass production) while decreasing LtBC (light-based biomass production) from the uncalibrated to the calibrated parameter sets, these results do not clearly identify the limiting factor for biomass growth. Instead, these results confirm that for the year 2019, the simulated yield difference is almost entirely dependent on the harvest index.

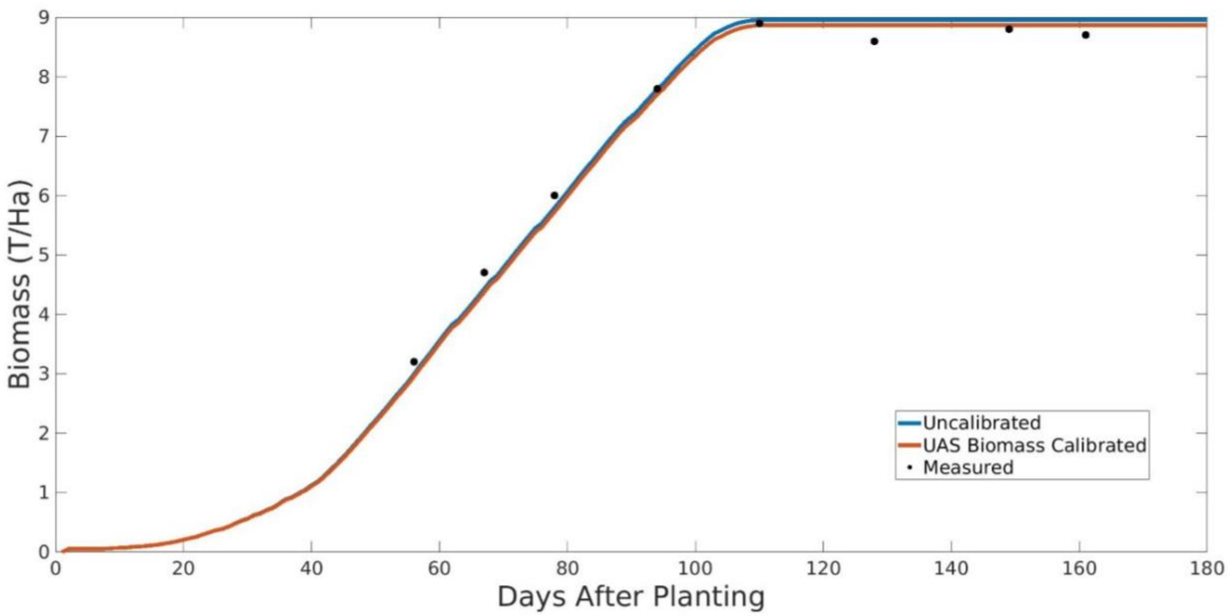

Figure 14: Biomass accumulation curve for the 2019 season for the Uncalibrated and UAS calibrated simulations, as well as in-situ measurements.

The crop calibration process focused only on data collected from experimental breeding plots. In A broader test of the calibrated crop parameter files is conducted by running the VIC-CropSyst model and comparing with bulk soybean yields. Bulk soybean is planted and managed by the farm staff and is not part of any experiment, instead the bulk soybean field use seed and management practices similar to those employed by farmers in the region. The biggest advantage of using the bulk soybean data are that it gives us access to multiple years of yield data, as illustrated in Figure 15. Also shown are simulated yield from two VIC-CropSyst simulations, the uncalibrated simulation and the simulation calibrated using UAS estimated biomass. The in-situ biomass calibrated simulation results are not shown as they are almost identical to those from the UAS calibrated simulation. Bulk yield values represent mean yields in all bulk soybean fields from the years 2007 to 2019, except 2011 when yield data were not measured.

On average, the uncalibrated model underestimated yield by 0.33 T/ha, while the calibrated model overestimated yield by 0.11 T/ha. The biggest contributor to this difference is the harvest index which is higher in the parametrized crop files resulting in more of the accumulated biomass being converted to yield. Additionally, the calibrated model yields track much closer to the observed yields, with a maximum deviation of 0.19 T/ha compared to the uncalibrated model with a maximum deviation of 0.42 T/ha. Though the calibrated model overestimates yield in all years with yield data except 2017 and 2018. Year 2010 and 2012 had lower than normal yields due to drought conditions. Drought formed in summer and continued into fall in 2010, while 2012 was a significant drought for much of the Midwestern U.S. for all of the growing season. The calibration process yielded the greatest improvement in simulated yield during these years impacted by drought. This is likely a combination of the BTR parameter being increased through calibration, which increases the relative importance of transpiration dependent biomass growth. Given that these are drought years, transpiration rather than radiation should be the more important limiting factor to biomass accumulation.

Low observed yields in 2011 were attributed to major flooding, which led directly to the bulk fields not being harvested and no yield being available for comparison. Yield for ACRE in 2011 was estimated from soybean yield in Indiana for the years 2011 to 2014, by obtaining the values from the USDA (“USDA National Agricultural Statistics Service,” 2020). By interpolating with the measured yield values from ACRE, for 2010 and 2012, and by considering the difference from the USDA report yield values, an estimated yield for 2011 of 3.2 T/ha is found. This value was substantially closer to the value before parametrization. The calibration process did not change simulated yield for that year as much as for the years affected by drought. A major shortcoming of CropSyst is that flooding stress is not represented in the default model, so there was no direct mechanism in CropSyst to represent the effect of this type of stress on yield.

These results partially support the studies second hypothesis that the parametrization of crop parameters is expected to improve our ability to quantify the impact of hydrological extremes on predicted crop yields. Under drought conditions the parametrization produced results that were closer to the actual values measured in the field as compared to the results when the model was not parametrized. The problem is the cases where flooding occurs. Given that CropSyst does not have a built-in way to account for flood stress, this cannot be further assessed.

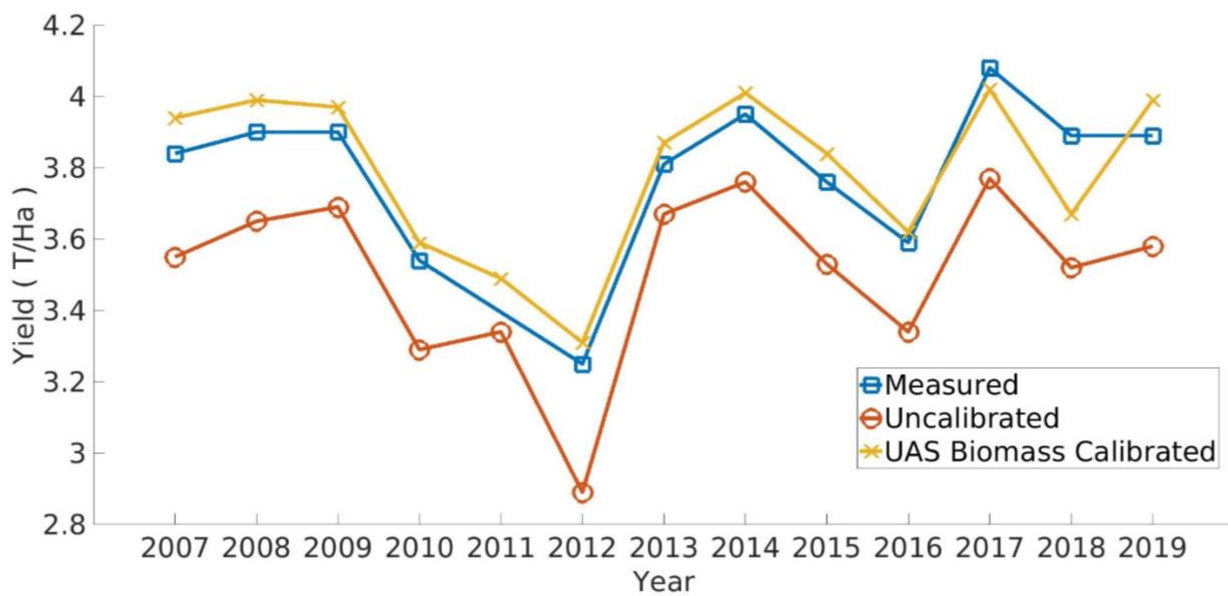

Figure 15: Yield development comparison for the uncalibrated run as well as the UAS calibrated run, along with the measured bulk yield for ACRE across multiple years.

5. DISCUSSION AND CONCLUSIONS

5.1 Discussion

The utilization of remotely sensed images from UAS was evaluated as a new technique to parametrize crops to improve to prediction of crop yields while requiring less labor and fewer man-hours than traditional in-situ sampling methods. In-situ and remotely sensed images were used to estimate biomass which was in turn used to parametrize crop growth via CropSyst’s crop parameter calibrator tool. The new soybean parameterizations were then used in VIC-CropSyst to simulate biomass and yield of soybeans at both the scale of a single field experimental plot and on average when compared to all bulk soybean plots at ACRE. The successful development of this procedure will potentially improve the ability to predict crop yields, while allowing a user to rapidly collect data at a relatively low cost and labor.

The development of a simplified procedure to calibrate remotely sensed MSP images was essential as it greatly improved the ability to consistently compute band algorithms across multiple dates that were vital in the biomass estimation equation. The procedure is relatively easy to use, as it only requires the use of spectral panels placed near the field being imaged to accurately calibrate the images. Furthermore, this procedure was incorporated into the CIE-VID processing pipeline which greatly improved its viability at producing accurate band indices and algorithms. All outputs and results can be credited to this calibration procedure.

The use of estimated biomass values within CropSyst’s crop parameter calibrator yielded crop growth parameters similar to those derived when using in-situ biomass samples. Both methods of calibration improved the representation of crop growth for the specific field site conditions. Five parameters were altered in the parametrization process. Two of these parameters, above ground biomass-transpiration coefficient and light to above ground biomass conversion were directly related to biomass production. Another two parameters, leaf duration and initial green leaf area index were directly related to leaf area development. The last parameter, harvest index, was directly influential in yield results as it controls the fraction of accumulated biomass that is harvested grain. Biomass production is controlled either by the availability of water or the plant's ability to make use of solar radiation, which is in turn related to LAI development and temperature. The above ground biomass-transpiration coefficient increased slightly during calibration, while the light to above ground biomass conversion decreased indicating that the calibration process increased the importance of water availability on the accumulation of biomass relative to incoming radiation. Neither the leaf duration or initial green leaf area index changed substantially with calibration, and those mostly impact LAI development and thus the radiation use efficiency of the plant. Harvest index was the parameter with the most substantial change. The change in harvest index was expected as the soybean parameters within the default crop parameter files were last calibrated in 2008 and major crop improvements have occurred since then. Furthermore, the phenotypes used to parametrize the models are indeterminate varieties which tend to have higher yields (Krashen, 1982), while the phenotypes used in the default files are determinate as they were developed based on southern soybean phenotypes.

The simulation run with a UAS parametrized crop parameter file produced yield and biomass values that were closer to the in-situ measured values, as opposed to the simulation run with the default parameters. In addition, the simulations run with the UAS parametrized crop parameter file and the in-situ parametrized parameter file produced similar yield values with a max difference of 0.02 T/Ha for both the research and bulk fields, which supports our first hypothesis that the development of relationships between remotely sensed data and crop indices is expected to improve the ability to parametrize crops for local conditions. This is to be expected as any site-specific calibration should improve model performance. The parametrization process utilized 2018 field data from multiple phenotypes. Initially the 2019 experiment was intended to be used for parametrization, but given the stress conditions witnessed in 2019, 2018 was used instead. This posed a problem as 2018 contained multiple distinct varieties, while 2019 contained only nine that were picked from 2018 to represent a wide variety of yield potentials. When comparing the measured in-situ yield of the 9 varieties and all 2018 varieties, small differences can be seen. This is because the varieties picked for 2019 were not within the extremes of yield potential. Furthermore, by comparing both values to the outputs of the parametrized simulation, the differences are minute. Picking the mean yield of the field was a necessary simplification to test the robustness of the model as well as test the influence of multiple varieties on the outputs.

The simulated yield after parametrization was much closer to that of both the nine varieties of 2019 and all varieties within the experiments as compared to the results of the default simulation, further supporting the first hypothesis. This validates the use of remote sensing images to parametrize the crop parameter files as the yield predictions greatly improved. When assessing for the yield of the bulk soybean within ACRE, more things become clear. First, the parametrization process produced yields that were generally overestimates of the measured yields, yet the mean differences to the true yield across twelve years was 0.11 T/ha which was a better estimate then the yield produced by the default parameter files which generally underestimated yield by an average of 0.33 T/Ha. This directly relates to the harvest index and suggests that the harvest index parametrized for is an overestimate to that of the bulk variety, as the biomass itself was well simulated, as the measured and simulated biomass were close in value.

In terms of years where extreme weather conditions were observed, three stand out, being 2010, 2011 and 2012. 2010 and 2012 were years where major droughts were experienced in Indiana. Drought stress is dependent on time in which the canopy received water as well as the quantity of water received. On average, we approximately expect 12.5 inches of rain in our study area. For 2010 and 2012, those were 9.5 and 7 inches respectively during the growing season. The yield results of the unparametrized simulation were farthest from the actual measured values in those years, while the parameterized model performed similar to other years. This validates the second hypothesis in which the parametrization of crop models for local conditions is expected to improve their ability to quantify the impact of hydrological extremes on predicted crop yields, however, this is only valid for drought years. 2011, which was a wet year, had major in-field ponding as the rain was more intense and sporadic with a total of 15 inches of rain during the growing season. 2011 was the only instance in which all simulations overestimated the 2011-estimated yield (Figure 17). In addition, the yield of the parametrized simulation was farther from the measured yield than the unparametrized run. This is due to CropSyst not having the ability to simulate excess water stress. Furthermore, given that above ground biomass conversion was higher in the parametrized soybean file, that meant that the transpiration dependent biomass growth would be higher, which in turn meant a higher biomass estimate thus a higher yield.

Biomass could only be assessed for 2019, as that is the only year with a complete biomass data collection. In terms of biomass prediction, both runs were able to produce values that were approximately equal to those measured in-situ, with a max difference of 0.5 T/Ha for any one date, and a mean difference of 0.1 T/Ha across all sampling dates. The difference between simulations was minute and was only distinguished at the end of the season. This was linked to the lower leaf duration within the parametrized simulation decreased biomass accumulation towards the end of the season. Given that the biomass estimates for both simulations for 2019 were fairly close, it can be easily concluded that harvest index is in fact the parameter that is generally affecting the yield most. An issue with biomass estimation within CropSyst, and other crop models in general, is the method in which biomass is accumulated. The model considers the lower of two values being radiation dependent biomass and transpiration dependent biomass. In reality, biomass growth will be impacted by both simultaneously as both will limit the crop growth not just one or the other. If more in-situ biomass data were available, a further thorough investigation could be performed.

The type of soybean used within CropSyst is modeled on southern soybean which is predominantly determinant. Soybean grown in much of the Midwestern U.S. and specifically at ACRE is indeterminate. Indeterminate varieties start flowering several weeks before they terminate vegetative growth, which could potentially lead to improved yield than simulated in the case where early season stress slows biomass accumulation, but improved conditions later in the growing season allow the indeterminant crop to continue flowering and producing pods. This is crucial as CropSyst models crop growth similar to how determinate crops would grow. A solution to this problem is to alter the days required for each growth stage to something more comparable to indeterminate crop growth.

Other research projects have used CropSyst, and parametrized crop files to simulate crop growth. Usually, parametrization occurs by measuring and estimating field values directly (Abi Saab et al., 2015; Confalonieri & Bocchi, 2005). This study is introducing a new method in which parametrization can make use of biomass estimates from remote sensing imagery. In terms of yield estimates, this study had a mean error of estimation of 0.11 T/Ha as compared to 0.15 T/Ha in Abi Saab et al. (2015) which fully utilized in-situ data collection, which is an indicator that utilizing UAS imagery produces results on par with in-situ measurements.

5.2 Conclusion

The overarching objective of this research project was to assess the ability of using remotely sensed data from UAS platforms to parametrize a crop model that could be used to simulate soybean biomass accumulation and yield. The initial hypothesis was that the development of relationships between remote sensing derived crop indices and parameters used by crop models would improve the simulation of crop growth for local conditions. This process is already established when using in-situ measurements of crop development, but the use of remote sensing derived measurements could significantly reduce the effort required to collect the necessary data. Increasing the amount of data available for calibration of crop models should improve their ability to quantify the impact on hydrologic extremes on predicted crop yields. The research analysis revealed the following:

- The implementation of image calibration for MSP images is crucial and a required to accurately analyze any images from the MSP sensor.
- Biomass estimation via remote sensing images currently requires a combination of MSP images and field measurements to accurately develop the model.
- The use of remote sensing derived estimated biomass to parametrize crop growth within CropSyst's crop parameter calibrator proved to be successful. The biggest difference post parametrization was the Harvest Index which is likely caused due to phenotypic traits and environmental conditions.
- Calibration of crop growth using both in-situ and remote sensing demonstrated that crop yield predictions can be improved substantially as compared to the unparametrized model. The yield from the parametrized simulations tended to slightly overestimate the measured yield especially for bulk varieties, which is to be expected as the parametrization was done for a different set of phenotypes and fields. Harvest Index proved to be the determining factor for yield predictions.
- Under drought stress, the parameterized simulations were able to produce a better yield prediction that that of unparametrized simulations. The difference of yield between both methods was greatest during drought stress years as the unparametrized simulation tended to output results that were significantly an underestimate of the actual measure yield, whereas the parametrization was just slightly overestimating the yield. In terms of excess water stress, CropSyst currently has no integrated way to accurately simulate such

conditions and as such both simulations produced yield estimates that were significantly off as compared to the estimated field yield.

- The work and methodology used in this study is crucial in the advancing of crop modeling. Not only does the parametrization using remote sensing images improve the predictions of crop models, but it also decreases the manpower and cost required to collect the different parameters needed to parametrize the cropping system models. The further use and development of remote sensing technology and band indices information will allow for a better understanding of the relation between remotely sensed data and crop parameters, and eventually the improvement of our ability to accurately predict crop growth which is essential to maintain future food securities.

5.3 Future Work

There are several ways in which the procedure and work introduced in this study can be expanded upon. The first would be the development of a biomass estimation equation that would be robust enough to handle multiple soybean phenotypes while requiring minimal field data to calibrate and to improve the prediction by decreasing RMSE and improving the fit of the model. Essentially this equation would be a combination of band indices and crop traits similar to plant height that would estimate soybean biomass. Some suggestions would be to follow a procedure similar to what is introduced in Maimaitijiang et al. (2019), where they introduce the use of stepwise multilinear regression of multiple crop indices from RGB cameras to estimate biomass. This can be altered to use MSP which would allow the use of the Red Edge and NIR bands which would improve the fit significantly. One issue with using multi linear regression models is the fact that often the multiple indices used are correlated, and as such there would be a need to decrease the number of inputs used to minimize correlation between the inputs. Another suggestion would be to incorporate different plant indices similar to plant height and LAI to improve the biomass estimation equation. In addition, the adaptation of a LAI estimation equation would allow the parametrization of LAI specific parameters and in turn the accurate modeling of LAI.

The second way to expand this work is to further our ability to calibrate crop growth by eliminating the need to use the crop calibrator. In this study CropSyst's crop parameter calibrator is used to output crop files with altered parameters. One possibility is to be able to estimate the specific crop indices directly from remotely sensed data. In the application of CropSyst this might be problematic as the parameters are hard to estimate even as field measurements (Abi Saab et al., 2015). This also introduces the possibility of expanding the work into different crop models similar to AquaCrop which uses crop parameters that are more easily estimated and readily available, yet might not be sufficient for hydrologic applications and climate change studies.

The third way to expand this work is by the utilization of satellite-based imagery as opposed to UAS platforms. This would require new algorithms to be developed as the scale of the imagery is very different between the two methods. Another venture would be to develop and test biomass estimation algorithms for crops other than Soybean.

One limitation of using CropSyst is the lack of ability to accurately simulate a plant's response to excessive soil moisture, which is shared by multiple cropping system models (e.g., Shaw et al., 2013; Li et al., 2019). As part of Pasley et al. (2020), the APSIM model was modified to include an excess water stress function that would be accounted for in the estimation of radiation use efficiency and root length. The implementation of such an alteration in VIC-CropSyst, would improve the yield predictions under conditions of excess soil moisture and waterlogging.

REFERENCES

Abi Saab, M. T., Todorovic, M., & Albrizio, R. (2015). Comparing aquaCrop and cropSyst models in simulating barley growth and yield under different water and nitrogen regimes: Does calibration year influence the performance of crop growth models? Agricultural Water Management, 147, 21–33. https://doi.org/10.1016/j.agwat.2014.08.001

Acevo-Herrera, R., Aguasca, A., Bosch-Lluis, X., Camps, A., Martínez-Fernández, J., Sánchez-Martín, N., & Pérez-Gutiérrez, C. (2010). Design and first results of an UAV-borne L-band radiometer for multiple monitoring purposes. Remote Sensing, 2(7), 1662–1679. https://doi.org/10.3390/rs2071662

Acevo-Herrera, R., Aguasca, A., Bosch-Lluis, X., Camps, A., Martínez-Fernández, J., Sánchez-Martín, N., & Pérez-Gutiérrez, C. (2010). Design and first results of an UAV-borne L-band radiometer for multiple monitoring purposes. Remote Sensing, 2(7), 1662–1679. https://doi.org/10.3390/rs2071662

Akenhead, S. A. (2017). Logic Model Overview. In University of Washington. http://vic.readthedocs.io/en/master/Overview/ModelOverview/

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., & Ridella, S. (2012). The “K” in K-fold Cross Validation. Undefined.

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. In Trends in Plant Science (Vol. 19, Issue 1, pp. 52–61). Trends Plant Sci. https://doi.org/10.1016/j.tplants.2013.09.008

Bai, G., Ge, Y., Hussain, W., Baenziger, P. S., & Graef, G. (2016, October 1). A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. Computers and Electronics in Agriculture; Elsevier B.V. https://doi.org/10.1016/j.compag.2016.08.021

Bellas, N., Chai, S. M., Dwyer, M., & Linzmeier, D. (2009). Real-time fisheye lens distortion correction using automatically generated streaming accelerators. Proceedings - IEEE Symposium on Field Programmable Custom Computing Machines, FCCM 2009, 149–156. https://doi.org/10.1109/FCCM.2009.16

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M. L., & Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation, 39, 79–87. https://doi.org/10.1016/j.jag.2015.02.012

Berni, J. A. J., Zarco-Tejada, P. J., Suárez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Transactions on Geoscience and Remote Sensing, 47(3), 722–738. https://doi.org/10.1109/TGRS.2008.2010457

Berrar, D. (2018). Cross-validation. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, 1–3, 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Borra-Serrano, I., Swaef, T. De, Quataert, P., Aper, J., Saleem, A., Saeys, W., Somers, B., Roldán-Ruiz, I., & Lootens, P. (2020). Closing the phenotyping gap: High resolution UAV time series for soybean growth analysis provides objective data from field trials. Remote Sensing, 12(10), 1644. https://doi.org/10.3390/rs12101644

Bowling, L. C., Cherkauer, K. A., Lee, C. I., Beckerman, J. L., Brouder, S., Buzan, J. R., Doering, O. C., Dukes, J. S., Ebner, P. D., Frankenberger, J. R., Gramig, B. M., Kladivko, E. J., & Volenec, J. J. (2020). Agricultural impacts of climate change in Indiana and potential adaptations. Climatic Change, 163(4), 2005–2027. https://doi.org/10.1007/s10584-020-02934-9

Byun, K., & Hamlet, A. F. (2018). Projected changes in future climate over the Midwest and Great Lakes region using downscaled CMIP5 ensembles. International Journal of Climatology, 38, e531–e553. https://doi.org/10.1002/joc.5388

C. L. Jones, P. R. Weckler, N. O. Maness, R. Jayasekara, M. L. Stone, & D. Chrz. (2007). Remote Sensing to Estimate Chlorophyll Concentration in Spinach Using Multi-Spectral Plant Reflectance. Transactions of the ASABE, 50(6), 2267–2273. https://doi.org/10.13031/2013.24079

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M. D., & Luis Araus, J. (2012). High-throughput Phenotyping and Genomic Selection: The Frontiers of Crop Breeding Converge. Journal of Integrative Plant Biology, 54(5), 312–320. https://doi.org/10.1111/j.1744-7909.2012.01116.x

Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L., & Suomalainen, J. (2015). Estimating plant traits of grasslands from UAV-acquired hyperspectral images: A comparison of statistical approaches. ISPRS International Journal of Geo-Information, 4(4), 2792–2820. https://doi.org/10.3390/ijgi4042792

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing of Environment, 62(3), 241–252. https://doi.org/10.1016/S0034-4257(97)00104-1

Cherkauer, K. A., Bowling, L. C., Byun, K., Chaubey, I., Chin, N., Ficklin, D. L., Hamlet, A. F., Kines, S. J., Lee, C. I., Neupane, R., Pignotti, G. W., Rahman, S., Singh, S., Femeena, P. V., & Williamson, T. N. (2021). Climate change impacts and strategies for adaptation for water resource management in Indiana. Climatic Change, 165(1–2). https://doi.org/10.1007/s10584-021-02979-4

Cherkauer, K. A., & Lettenmaier, D. P. (2003). Simulation of spatial variability in snow and frozen soil. Journal of Geophysical Research: Atmospheres, 108(22), 2003JD003575. https://doi.org/10.1029/2003jd003575

Cherkauer, K. A., & Sinha, T. (2010). Hydrologic impacts of projected future climate change in the Lake Michigan region. Journal of Great Lakes Research, 36(SUPPL. 2), 33–50. https://doi.org/10.1016/j.jglr.2009.11.012

Cobb, J. N., DeClerck, G., Greenberg, A., Clark, R., & McCouch, S. (2013). Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. In Theoretical and Applied Genetics (Vol. 126, Issue 4, pp. 867–887). Springer. https://doi.org/10.1007/s00122-013-2066-0

Cohen, J. E. (2003). Human Population: The Next Half Century. In Science (Vol. 302, Issue 5648, pp. 1172–1175). Science. https://doi.org/10.1126/science.1088665

Confalonieri, R., & Bocchi, S. (2005). Evaluation of CropSyst for simulating the yield of flooded rice in northern Italy. European Journal of Agronomy, 23(4), 315–326. https://doi.org/10.1016/j.eja.2004.12.002

CropSyst. (n.d.). CropSyst's Web Manual. Retrieved March 14, 2021, from http://modeling.bsyse.wsu.edu/CS_Suite_4/CropSyst/manual/index.htm

Danner, M., Locherer, M., Hank, T., & Richter, K. (2015). Measuring Leaf Area Index (LAI) with the LI-Cor LAI 2200C or LAI-2200. EnMAP Hyperspectral Imager, 2200. https://doi.org/10.2312/enmap.2015.009

de Paiva Rolla, A. A., de Fátima Corrêa Carvalho, J., Fuganti-Pagliarini, R., Engels, C., do Rio, A., Marin, S. R. R., de Oliveira, M. C. N., Beneventi, M. A., Marcelino-Guimarães, F. C., Farias, J. R. B., Neumaier, N., Nakashima, K., Yamaguchi-Shinozaki, K., & Nepomuceno, A. L. (2014). Phenotyping soybean plants transformed with rd29A:AtDREB1A for drought tolerance in the greenhouse and field. Transgenic Research, 23(1), 75–87. https://doi.org/10.1007/s11248-013-9723-6

Donatelli, M., Stockle, C., Ceotto, E., & Rinaldi, M. (1997). Evaluation of CropSyst for cropping systems at two locations of Northern and Southern Italy. European Journal of Agronomy, 6(1–2), 35–45. https://doi.org/10.1016/S1161-0301(96)02029-1

Donatelli, M., Bellocchi, G., & Fontana, F. (2003). RadEst3.00: Software to estimate daily radiation data from commonly available meteorological variables. European Journal of Agronomy, 18(3–4), 363–367. https://doi.org/10.1016/S1161-0301(02)00130-2

Duchemin, B., Maisongrande, P., Boulet, G., & Benhadj, I. (2008). A simple algorithm for yield estimates: Evaluation for semi-arid irrigated winter wheat monitored with green leaf area index. Environmental Modelling and Software, 23(7), 876–892. https://doi.org/10.1016/j.envsoft.2007.10.003

Fan, F., Bradley, R. S., & Rawlins, M. A. (2014). Climate change in the northeastern US: Regional climate model validation and climate change projections. Climate Dynamics, 43(1–2), 145–161. https://doi.org/10.1007/s00382-014-2198-1

Fehr, W. R., Caviness, C. E., Burmood, D. T., & Pennington, J. S. (1971). Stage of Development Descriptions for Soybeans, Glycine Max (L.) Merrill 1 . Crop Science, 11(6), 929–931. https://doi.org/10.2135/cropsci1971.0011183x001100060051x

Forest Ecosystems (Third edition). (2007). Leaf Area Index - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/earth-and-planetary-sciences/leaf-area-index

Freitas Moreira, F., Rojas de Oliveira, H., Lopez, M. A., Abughali, B. J., Gomes, G., Cherkauer, K. A., Brito, L. F., & Rainey, K. M. (2021). High-Throughput Phenotyping and Random Regression Models Reveal Temporal Genetic Control of Soybean Biomass Production. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.715983

Hamman, J. J., Nijssen, B., Bohn, T. J., Gergel, D. R., & Mao, Y. (2018). The variable infiltration capacity model version 5 (VIC-5): Infrastructure improvements for new applications and reproducibility. Geoscientific Model Development, 11(8), 3481–3496. https://doi.org/10.5194/gmd-11-3481-2018

Holben, B. N., Tucker, C. J., & Fan, C. J. (1980). Spectral assessment of soybean leaf area and lear biomass. In Photogrammetric Engineering & Remote Sensing (Vol. 46, Issue 5).

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore, A. D., Brown, H., Whish, J. P. M., Verrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L., Hochman, Z., … Keating, B. A. (2014). APSIM - Evolution towards a new generation of agricultural systems simulation. Environmental Modelling and Software, 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009

Hunt, E. R., Cavigelli, M., Daughtry, C. S. T., McMurtrey, J. E., & Walthall, C. L. (2005). Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. Precision Agriculture, 6(4), 359–378. https://doi.org/10.1007/s11119-005-2324-5

Hunt, E. R., Dean Hively, W., Fujikawa, S. J., Linden, D. S., Daughtry, C. S. T., & McCarty, G. W. (2010). Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. Remote Sensing, 2(1), 290–305. https://doi.org/10.3390/rs2010290

Iqbal, F., Lucieer, A., & Barry, K. (2018). Simplified radiometric calibration for UAS-mounted multispectral sensor. European Journal of Remote Sensing, 51(1), 301–313. https://doi.org/10.1080/22797254.2018.1432293

Joernsgaard, B., & Halmoe, S. (2003). Intra-field yield variation over crops and years. European Journal of Agronomy, 19(1), 23–33. https://doi.org/10.1016/S1161-0301(02)00016-3

Jones, H. G. (2004). Irrigation scheduling: Advantages and pitfalls of plant-based methods. Journal of Experimental Botany, 55(407), 2427–2436. https://doi.org/10.1093/jxb/erh213

Korjus, K., Hebart, M. N., & Vicente, R. (2016). An efficient data partitioning to improve classification performance while keeping parameters interpretable. PLoS ONE, 11(8). https://doi.org/10.1371/journal.pone.0161788

Krashen, S. D. (1982). A Comparison of Determinate and Indeterminate Soybean Lines for Double Cropping in Virginia. 46(2), 55. http://eprints.uanl.mx/5481/1/1020149995.PDF

Langton, S. (1990). Avoiding edge effects in agroforestry experiments; the use of neighbour-balanced designs and guard areas. Agroforestry Systems, 12(2), 173–185. https://doi.org/10.1007/BF00123472

Li, Y., Guan, K., Schnitkey, G. D., DeLucia, E., & Peng, B. (2019). Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. Global Change Biology, 25(7), 2325–2337. https://doi.org/10.1111/gcb.14628

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research, 99(D7), 14415. https://doi.org/10.1029/94jd00483

Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. Global and Planetary Change, 13(1–4), 195–206. https://doi.org/10.1016/0921-8181(95)00046-1

Liang, X., Wood, E. F., & Lettenmaier, D. P. (1999). Modeling ground heat flux in land surface parameterization schemes. Journal of Geophysical Research Atmospheres, 104(D8), 9581–9600. https://doi.org/10.1029/98JD02307

Lopez, M. A., Freitas Moreira, F., & Rainey, K. M. (2021). Genetic Relationships Among Physiological Processes, Phenology, and Grain Yield Offer an Insight Into the Development of New Cultivars in Soybean (Glycine max L. Merr). Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.651241

Lyu, B., Smith, S. D., Xue, Y., & Cherkauer, K. A. (2019). Deriving vegetation indices from high-throughput images by using unmanned aerial systems in soybean breeding. 2019 ASABE Annual International Meeting. https://doi.org/10.13031/aim.201900279

Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K. T., Maw, M. J. W., Shakoor, N., Mockler, T., & Fritschi, F. B. (2019). Vegetation Index Weighted Canopy Volume Model (CVM VI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 151, 27–41. https://doi.org/10.1016/j.isprsjprs.2019.03.003

Malek, K., Stöckle, C., Chinnayakanahalli, K., Nelson, R., Liu, M., Rajagopalan, K., Barik, M., & Adam, J. C. (2017). VIC-CropSyst-v2: A regional-scale modeling platform to simulate the nexus of climate, hydrology, cropping systems, and human decisions. Geoscientific Model Development, 10(8), 3059–3084. https://doi.org/10.5194/gmd-10-3059-2017

Micasense. (2016). Parrot Sequoia Datasheet. https://www.parrot.com/files/s3fs-public/sequoia_official_documentation_and_specifications_2016_0.pdf

Moreira, F. F., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2019). Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. Plant Methods, 15(1), 1–9. https://doi.org/10.1186/s13007-019-0519-4

Nakano, S., Purcell, L. C., Homma, K., & Shiraiwa, T. (2020). Modeling leaf area development in soybean (Glycine max L.) based on the branch growth and leaf elongation. Plant Production Science, 23(3), 247–259. https://doi.org/10.1080/1343943X.2019.1702468

Pala, M., Stockle, C. O., & Harris, H. C. (1996). Simulation of Durum wheat (Triticum turgidum ssp. Durum) growth under different water and nitrogen regimes in a Mediterranean environment using CropSyst. Agricultural Systems, 51(2), 147–163. https://doi.org/10.1016/0308-521X(95)00043-5

Pasley, H. R., Huber, I., Castellano, M. J., & Archontoulis, S. V. (2020). Modeling Flood-Induced Stress in Soybeans. Frontiers in Plant Science, 11, 62. https://doi.org/10.3389/fpls.2020.00062

Passricha, N., Saifi, S. K., Negi, H., Tuteja, R., & Tuteja, N. (2020). Transgenic approach in crop improvement. In Advancement in Crop Improvement Techniques (pp. 329–350). Elsevier. https://doi.org/10.1016/b978-0-12-818581-0.00020-6

PIX4D. (2021). Professional photogrammetry and drone mapping software | Pix4D. Pix4D. https://www.pix4d.com/

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. In Encyclopedia of Database Systems (pp. 532–538). Springer US. https://doi.org/10.1007/978-0-387-39940-9_565

Rouse. (1974). 1974_Rouse_Monitoring Vegetation Systems in the Great Plains with ERTS_NDVI_.pdf. NASA. Goddard Space Flight Center 3d ERTS-1 Symp., Vol. 1, Sect. A.

Sankaran, S., Khot, L. R., & Carter, A. H. (2015). Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. Computers and Electronics in Agriculture, 118, 372–379. https://doi.org/10.1016/j.compag.2015.09.001

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., & Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. In European Journal of Agronomy (Vol. 70, pp. 112–123). Elsevier. https://doi.org/10.1016/j.eja.2015.07.004

Scott, H. D., DeAngulo, J., Daniels, M. B., & Wood, L. S. (1989). Flood Duration Effects on Soybean Growth and Yield. Agronomy Journal, 81(4), 631–636. https://doi.org/10.2134/agronj1989.00021962008100040016x

Service, U. F. (n.d.). NRM Field Sampled Vegetation (FSVeg). Retrieved March 15, 2021, from https://www.fs.fed.us/nrm/fsveg/

Shafian, S., Rajan, N., Schnell, R., Bagavathiannan, M., Valasek, J., Shi, Y., & Olsenholler, J. (2018). Unmanned aerial systems-based remote sensing for monitoring sorghum growth and development. PLoS ONE, 13(5), e0196605. https://doi.org/10.1371/journal.pone.0196605

Shaw, R. E., Meyer, W. S., McNeill, A., & Tyerman, S. D. (2013). Waterlogging in Australian agricultural landscapes: A review of plant responses and crop models. In Crop and Pasture Science (Vol. 64, Issue 6, pp. 549–562). https://doi.org/10.1071/CP13080

Sinha, S. K., Rao, N. H., & Swaminathan, M. S. (1988). Food security in the changing global climate. Proceedings of the Conference on the Changing Atmosphere: Implications for Global Security. http://www.ciesin.org/docs/004-041/004-041.html

Spectravista. (n.d.). GER 1500 - Spectra Vista Corporation. Retrieved February 26, 2021, from https://www.spectravista.com/our-instruments/ger1500/

Stockle, C. O. (1996). GIS and simulation technologies for assessing cropping systems management in dry environments. American Journal of Alternative Agriculture, 11(2–3), 115–120. https://doi.org/10.1017/S0889189300006895

Stöckle, C. O., Donatelli, M., & Nelson, R. (2003). CropSyst, a cropping systems simulation model. European Journal of Agronomy, 18(3–4), 289–307. https://doi.org/10.1016/S1161-0301(02)00109-0

Stockle, C. O., Martin, S. A., & Campbell, G. S. (1994). CropSyst, a cropping systems simulation model: Water/nitrogen budgets and crop yield. Agricultural Systems, 46(3), 335–359. https://doi.org/10.1016/0308-521X(94)90006-2

The MathWorks. (2020). Nonlinear regression - MATLAB nlinfit. https://www.mathworks.com/help/stats/nlinfit.html

Thornton, P. E., & Running, S. W. (1999). An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agricultural and Forest Meteorology, 93(4), 211–228. https://doi.org/10.1016/S0168-1923(98)00126-9

Topcon. (n.d.). Topcon, Tokyo Japan. Retrieved March 15, 2021, from https://www.topconpositioning.com/na/accelerate-productivity?utm_source=InsightMediaLabs&utm_medium=AdWords&utm_campaign=434&utm_content=PositioningSolutions&gclid=Cj0KCQjwi7yCBhDJARIsAMWFScNeqhITeVZk0wgNfRti4t4VEzKL1rCm0I-rjRbSQv5t0kJlfvECOPsaAvagEALw_wc

Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A review on UAV-based applications for precision agriculture. Information (Switzerland), 10(11). https://doi.org/10.3390/info10110349

USDA. (2014). USDA National Agricultural Statistics Service. In Choice Reviews Online (Vol. 51, Issue 07). https://doi.org/10.5860/choice.51-3802

USDA. (2019). November Crop Production.

USDA. (3868 B.C.E.). Farm Bill | USDA. Usda. https://www.usda.gov/farmbill

USDA. (2020). Crop Production Orange Production Down 1 Percent from March Forecast.

Vargas, M., Combs, E., Alvarado, G., Atlin, G., Mathews, K., & Crossa, J. (2013). Meta: A suite of sas programs to analyze multienvironment breeding trials. Agronomy Journal, 105(1), 11–19. https://doi.org/10.2134/agronj2012.0016

Weraduwage, S. M., Chen, J., Anozie, F. C., Morales, A., Weise, S. E., & Sharkey, T. D. (2015). The relationship between leaf area growth and biomass accumulation in Arabidopsis thaliana. Frontiers in Plant Science, 6(APR), 167. https://doi.org/10.3389/fpls.2015.00167

Willmott, C. J., & Matsuura, K. (2005). CLIMATE RESEARCH Clim Res. CLIMATE RESEARCH Clim Res, 30, 79–82. www.int-res.com%0Awww.int-res.com%0Awww.usda.gov/%0Awww.int-res.com

Yang, G., Li, C., Yu, H., Xu, B., Feng, H., Gao, L., & Zhu, D. (2015). UAV based multi-load remote sensing technologies for wheat breeding information acquirement. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 31(21), 184–190. https://doi.org/10.11975/j.issn.1002-6819.2015.21.024

Zhou, X., Guan, H., Xie, H., & Wilson, J. L. (2009). Analysis and optimization of NDVI definitions and areal fraction models in remote sensing of vegetation. International Journal of Remote Sensing, 30(3), 721–751. https://doi.org/10.1080/01431160802392620