4-2023

Estimating Crop Stomatal Conductance Through High-Throughput Plant Phenotyping

Junxiao Zhang
University of Nebraska-Lincoln, jzhang95@huskers.unl.edu

ESTIMATING CROP STOMATAL CONDUCTANCE THROUGH HIGH-THROUGHPUT PLANT PHENOTYPING

by

Junxiao Zhang

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Master of Science

Major: Agricultural and Biological Systems Engineering

Under the Supervision of Professor Yufeng Ge

Lincoln, Nebraska

April 2023

ESTIMATING CROP STOMATAL CONDUCTANCE THROUGH HIGH-THROUGHPUT PLANT PHENOTYPING

Junxiao Zhang, M.S.

University of Nebraska, 2023

Advisor: Yufeng Ge

During photosynthesis and transpiration, crops exchange carbon dioxide and water with the atmosphere through stomata. When a crop experiences water stress, stomata are closed to reducing water loss. However, the closing of stomata also negatively affects the photosynthetic efficiency of the crop and leads to lower yields. Stomatal conductance (g_s) quantifies the degree of stomatal opening and closing by using the rate of gas exchange between the crop and the atmosphere, which helps to understand the water status of the crop for better irrigation management. Unfortunately, g_s measurement typically requires contact measuring instruments and manual collection in the field, which is time-consuming and labor-intensive. Thus, this study estimates g_s in two ways. Firstly, plant phenotypic data and weather information were used to estimate g_s for various types of crops. The plant phenotypic data were extracted from images captured by a thermal infrared camera, a multispectral camera, and a visible and near-infrared spectrometer integrated on field phenotyping platform. Weather information was obtained from a field weather station. The random forest regression (RFR) model performed the best with R^2 of 0.69 and RMSE of 0.135 $mol*m^{-2}*s^{-1}$, while the model using weather parameters alone had R^2 of 0.58 and RMSE of 0.161, and the model using phenotypic data alone had R^2 values of 0.59 and RMSE of 0.158

mol*m^{-2}*s^{-1}. The results indicated that there was a complementary relationship between plant phenotypic data and weather information in estimating g_s. The second aspect of the study was to estimate maize and soybean g_s directly from near-infrared, thermal-infrared and RGB (Red Green Blue) images collected by the same platform. The results showed that the convolutional neural network (CNN) model outperformed the other models with an R^2 of 0.52. In addition, adding soil moisture as a variable to the model improved its accuracy, which decreased the RMSE from 0.147 to 0.137 mol*m^{-2}*s^{-1}. This study highlights the potential of estimating g_s from remote sensing and field phenotyping platforms to help growers obtain information about the water status of crops and plan irrigation more efficiently.

Keywords: remote sensing, plant phenotyping, stomatal conductance, vegetation indices, machine learning, deep learning

ACKNOWLEDGMENT

First of all, I would like to express my deep appreciation to my advisor, Dr. Yufeng Ge, for the guidance he has given me throughout my academic journey. His expertise and dedication have helped me accomplish my research goals. I am grateful for his help in getting me through the challenges and opportunities of a Master's degree.

I would also want to thank my committee members, Dr. Geng (Frank) Bai, Dr. Yeyin Shi, and Dr. Stephen Scott, for their assistance and comments throughout my graduate studies. Their expertise and insights have helped me to improve the quality of my research.

I would also like to thank my dear friends Bidhan, Gautam, Kantilata, Nipuna, Shiva, Shubham and Zhaocheng for their support and encouragement. Their discussions and feedback to help me overcome the challenges I faced were invaluable.

I would also like to express my gratitude to my family for their love and support. Their unwavering faith in me has been a source of motivation to complete my studies and I am forever grateful for their presence in my life.

Finally, I would like to thank all those who have contributed to my academic and personal growth. Your support has been an integral part of my life's journey.

FUNDING SOURCES

The grant for my research was supported by (1) USDA-NIFA grant (High Intensity Phenotyping Sites: Transitioning to a Nationwide Plant Phenotyping Network, Award No. 2020-68013-32371), and (2) the Nebraska Agricultural Experiment Station through the Hatch Multistate funding program (Accession Number 7000908) from USDA-NIFA.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 HIGH-THROUGHPUT MEASUREMENT OF PLANT LEAF STOMATAL CONDUCTANCE – A REVIEW

1.1 Introduction

As the global population increases, so does the demand for food, feed, fiber, and fuel produced from the agriculture sector. One study points out that the demand for cereal will increase by 102% between 2009 and 2050 (Fukase & Martin, 2020). At the same time, climate change has brought increasingly more frequent extreme weathers that threaten agricultural production. Among them, severe drought events with prolonged duration are particularly concerning, as the availability of water is the single most important factor in determining crop productivity (Kimm et al., 2020). Therefore, water related issues in crop production, ranging from the breeding of drought-tolerant varieties, to improved decision-making in irrigation scheduling, have been the focus of extensive research in the past few decades. With the development of technology, people's attention is gradually shifting to sustainable smart agriculture (Campbell et al., 2014; Lipper et al., 2014).

It is well known that different genotypes can result in varied physiological and morphological traits in plants. The comprehensive evaluation of plant traits related to their growth, development, and physiology is called plant phenotyping (Araus & Cairns, 2014; Bai et al., 2016; Ge et al., 2019; L. Li et al., 2014; Pandey et al., 2017). To better elucidate the interaction between plant genotype and environment, and

resultant phenotype, can potentially play a significant role in improving agricultural yield. The conventional way of phenotypic data collection requires a great deal of repeated efforts by individuals using specialized equipment, which can be expensive, time-consuming, and error-prone. In typical plant breeding programs, hundreds to thousands of genetic lines need to be screened in a short time window, making traditional low-throughput phenotyping infeasible. This throughput problem is especially pronounced when target traits are rapidly changing, or a number of different traits are to be measured to improve selection or decision-making.

In the recent years, scientists have developed numerous robotic or automated systems for high-throughput phenotyping of plants in controlled environments and in the fields, thanks to the advent and democratization of new sensing and imaging technologies, as well as software to effectively process and analyze these data (Atefi et al., 2021; W. Yang et al., 2020). In a controlled environment or a greenhouse, plants are moved by a conveyor belt to imaging stations where the sensor data are recorded. Another strategy is to move sensors around the plants to collect data. Greenhouses have the benefit of controlled temperature, light intensity and duration, humidity, and other environmental variables, as well as reducing the impacts of sunshine, wind, etc. on the sensors (Ge et al., 2016). The relationship between plant phenotypic data and genotypes in natural habitats must also be investigated, as controlled environments can never be identical to the actual field (Basu et al., 2015). Unlike indoor platforms, measurements cannot be made in a field environment by moving plants. The research can only be conducted by using a stationery or mobile platform. These platforms

consist of cable suspension systems, gantry systems, satellite systems, agricultural Internet of Things (IoT) systems, unmanned aerial vehicles (UAVs), and ground vehicles (Araus & Cairns, 2014; Bao et al., 2021; Shakoor et al., 2017). Cable suspension systems are similar to gantry systems in that they use cables or gantries to carry several sensors above the canopy for measurements, a facility that enables repeatable crop experiments (Bai et al., 2019; Virlet et al., 2016). Satellite systems utilize images captured by small satellites in low orbit and have the ability to standardize measurements on a global scale (C. Zhang et al., 2020). IoT systems often employ field-fixed nodes with several sensors, which offer the advantage of continually collecting and transferring data with high spatial and temporal precision (Chamara et al., 2022). On the other hand, UAVs and ground vehicles such as tractors carry sensors for data collecting via such mobile platforms. The advantage of these platforms is their portability, but the battery sustainability of UAVs and the crop and soil damage caused by tractors cannot be ignored (R. Xu et al., 2019; G. Yang et al., 2017).

Stomata, being a key structure utilized by plants for carbon and water exchange, play a crucial function in plant growth and development. Stomata are microscopic pores founded in the epidermis of leaves and stems of plants. They are surrounded by two guard cells, which control the pore size and the gas exchange rate of the plant by changing the shape of the guard cells (Hetherington & Woodward, 2003; Kollist et al., 2014). The carbon dioxide and water exchange rate of stomata can be used to monitor the water use of plants, and this rate is called stomatal conductance (g_s). In modern

smart agriculture, g_s monitoring can therefore be used to optimize irrigation scheduling during crop growth, or to assist researchers in studying and breeding drought and heat tolerance of crops.

This paper reviews a series of high-throughput phenotyping studies on crop stomatal conductance performed by researchers in recent years and the techniques used. It is hoped that this article will help the readers understand the research gaps and potential of phenotyping of crop g_s.

1.2 Factors that Impact Stomatal Conductance

As an important structure for physiological functions such as transpiration and photosynthesis, g_s is primarily determined by the size and density of stomata (Bertolino et al., 2019). In addition to the differences in stomata morphology among species and genotypes (Casson & Gray, 2008), stomata are also regulated by environmental signals such as temperature, light intensity, atmospheric CO_2 concentration, relative humidity, and soil moisture content (Chaerle et al., 2005; Chaves et al., 2016; Roelfsema & Hedrich, 2005). Researchers have tried to understand and model the relationship between g_s and these environmental factors since several decades ago. Take two widely used g_s model as example, the Jarvis model (a model of the effect of different environmental factors on g_s) (Jarvis, 1976) and the Ball-Berry model (a model of the response of g_s to photosynthesis) (Ball et al., 1987). Table 1 shows the equations of these two models and the variables used in them. The next part of this section discusses in detail the reasons for the impact of

these variables on g_s.

Table 1: The Jarvis and Ball-Berry models for empirical modeling of plant stomatal conductance

Model Name	Function
Jarvis	$g_s = f_1(Q_p) * f_2(T) * f_3(VPD) * f_4(\Psi_1) * f_5(C_a)$
Ball-Berry	$g_s = f_6 A \frac{h_s}{C_a}$

Where: Q_p is the photon flux density, T is the leaf temperature, VPD is the vapor pressure deficit, Ψ_1 is the leaf water potential, C_a is ambient CO_2 concentration, A is CO_2 assimilation rate, and h_s is the relative humidity at the leaf surface.

Stomata are used to maximize photosynthesis and minimize water loss (Buckley et al., 2017). Photosynthesis, an essential biological process in plants, uses light energy to transform water and inorganic elements such as CO_2 into adenosine triphosphate (ATP) for plant growth. Consequently, light tensity, temperature, humidity, and CO_2 concentration influence g_s at all times. When leaf detects light, it opens stomata to absorb more CO_2 (Huang et al., 2021; Q. Zhang et al., 2019) and transport the water collected through the root system to the leaf for photosynthesis by transpiration. However, transpiration and photosynthesis are often in conflict because stomata open for photosynthesis and also promote water loss to the atmosphere through transpiration. When the temperature is too high, plants must cool themselves by evaporation (Jagadish et al., 2021). Nevertheless, when soil water content is less than the amount of water required for plant transpiration, g_s will be decreased to

prevent desiccation and control water loss (Huang et al., 2021). Hence, stomatal closure must be balanced in both open and close stages, which make it difficult to simply assess the impacts of these environmental conditions on g_s, and there are still gaps in our understanding of the physiological mechanisms behind stomatal regulation (Damour et al., 2010). Moreover, as some variables in these empirical models cannot be measured readily on a wide scale, the use of easily measurable variables for g_s estimation in agriculture offers considerable promise.

1.3 Contact Sensing Techniques

Based on the ecosystem structure of plant stomata and the physiological principles, the g_s measurements that have been developed can be divided into three methods: (1) Gas exchange measurements, (2) Porometry measurements, and (3) Sap flow measurements (Gowdy et al., 2022; Toro et al., 2019). In this section, the principle of each measurement method is described in detail and its advantages and disadvantages are discussed.

1.3.1 Gas Exchange Measurement

Gas exchange measurement quantifies the rate of photosynthesis and transpiration in real time by measuring the exchange of gases between the atmosphere and the interior airspace of the leaf. Today's commercial gas exchange measurement (Figure 1) instruments are mainly infrared gas analyzers (IRGA) and termed portable/compact photosynthesis systems, such as CIRAS-4 (PP Systems International, Inc.

Massachusetts, USA), LI-6800 (LI-COR Inc., Nebraska, USA), and LCi T (ADC BioScientific Ltd., Hoddesdon, UK). They are based on the principle that CO_2 and water vapor are absorbed by specific infrared wavebands (Barillot et al., 2010; Cernusak, 2020; Chieppa et al., 2021; Fayezizadeh et al., 2021; Lavoie-Lamoureux et al., 2017; Long & Bernacchi, 2003). These instruments calculate the difference in gas concentration between the sample chamber containing the leaf and the atmosphere, then combine it with the air flow rate to calculate g_s (Busch, 2018; Gaastra, 1959).

Figure 1: Examples of the gas exchange instrument for plant leaf stomatal conductance measurement. Left is LI-6800 and right is CIRAS-4.

1.3.2 Porometry Measurement

Porometry measurement, on the other hand, calculates g_s by measuring the water vapor produced by plants during transpiration (Askari et al., 2021). Porometers can be divided into two ways according to its measuring principle, dynamic method and steady-state method (Bell & Squire, 1981). The common dynamic porometers on the market today are AP4 (Delta-T Devices Ltd., Burwell, UK) and LI-600 (LI-COR Inc.,

Nebraska, USA), which calculates g_s by measuring the diffusing rate of water vapor from a sensor set under the leaf. By adding a confined space with a humidity sensor under the leaf blade, the g_s can be determined by measuring how quickly the humidity increases in that space. The steady-state porometer (for example SC-1, METER Group, Inc., Washington, USA) is relatively simple (Figure 2). Similar to the dynamic porometer, the steady-state porometer also measures the water vapor below the leaf blade. The difference between the two approaches is that the steady-state porometer measures the flow rate of water vapor in the path through two humidity sensors connected in series to obtain the g_s. Therefore the dynamic porometer is more complex than the steady-state porometer (Bell & Squire, 1981; Toro et al., 2019). In an early report (Bell & Squire, 1981) the accuracy of these two methods was compared and the results showed that the difference between these measurements was in the range of 2-20% but with the same random error (15%).

Figure 2: Examples of leaf porometers. Left is AP4 and right is SC-1.

1.3.3 Sap Flow Measurements

Sap flow measurement is commonly used in g_s measurement of trees, where the sap flow is derived by heating the stem and measuring the temperature at several points along the stem wood (Čermák et al., 2004; Ghimire et al., 2018; Motzer et al., 2005; Vandegehuchte & Steppe, 2012). This is then combined with climate data to derive g_s using the Penman-Monteith-type equation in the reverse direction (Jones et al., 1988; Kučera et al., 2017; Su et al., 2019). The common thermodynamic methods include, but are not limited to, the Trunk segment heat balance and Head dissipation. But since the focus of this paper is on field crops such as corn and soybeans, we will not go into the details of the sap flow sensors.

1.3.4 Comparison between the Gas Exchange and Porometry Instruments

Compared to the gas exchange and porometry measurement discussed in the previous sections, IRGA using the gas exchange method requires longer measurement time and a stable environment due to its complex principle (Toro et al., 2019). The porometer, on the other hand, is popular among researchers because of its portability. It only requires the use of relative humidity sensors compared to IRGA that requires the use of infrared light source, a PAR sensor, and a drying column for drying the gas. However, unlike the porometer, which can only measure g_s, the IRGA can also measure sub-stomatal CO_2 concentration and CO_2 assimilation rate while obtaining g_s due to its measurement principle. Despite this, it is still not possible to measure g_s

accurately with today's technology. One of the reasons is that stomata are not evenly distributed on the leaf surface, and even if measurements are made on the same leaf the results can still be biased due to differences in position (Dumont et al., 2014). Moreover, the instrument used for measurement can only cover a small area of the leaf blade, which also makes the measurement difficult. Another reason is that it is difficult to measure uniform and accurate g_s values between the two measurement methods because of their different principles. The result of Lavoie-Lamoureux et al., (2017) showed that porometer measurements are significantly higher than those of IRGA, and there is evidence that the accuracy decreases when the ambient humidity is too high (McDermitt, 1990). The final reason is leaf selection, even the same plant under natural conditions can have independent evaporation requirements depending on irradiance, wind, and temperature (Richardson et al., 2017), which may make it difficult for researchers to design experiments.

1.4 Remote Sensing Techniques

Since the 1980s, remote sensing has been utilized as one of the non-destructive methods for monitoring crop conditions. The development of remote sensing technology has expanded agriculture possibilities (Mulla, 2013; Weiss et al., 2020). As an emerging technique, remote sensing directly or indirectly collects plant traits such as height, canopy temperature, and vegetation indices by measuring electromagnetic energy emitted or reflected by plants at a distance. With the development of information and communication technologies, this technology, which

can collect data rapidly and on a vast scale, has become widely used in all facets of precision agriculture (Khanal et al., 2020; Sishodia et al., 2020). The collection of g_s in the field is challenging, but due to its importance, modeling employing g_s-associated environmental parameters and plant phenotypic data has become popular. This section discusses the parameters used in some modeling as well as forecasting studies in recent years. Sensor platforms commonly used for g_s monitoring can be classified into three types: (1) Satellite platform, (2) Unmanned aerial vehicles (UAVs), and (3) stationary platform in the field. The main types of sensors used include, RGB (Red Green Blue), multispectral, hyperspectral, and thermal imaging. This section briefly describes the operating principles of these sensors and their related applications.

1.4.1 Spectral Imaging

Spectral imaging works by scanning the target in different bands of the electromagnetic spectrum and composing the images of different bands into a single image set. It can be divided into multispectral imaging and hyperspectral imaging according to the number of bands measured. Conventionally, imaging with less than 20 bands is called multispectral imaging while imaging with more than 20 bands is called hyperspectral imaging. In agriculture, plant spectral information gathered by satellites, UAVs and other aerial platforms using multispectral or hyperspectral information has been utilized extensively to infer plant structure and biochemical properties. Normalized Difference Vegetation Index (NDVI), for instance, is

computed from the spectral reflectance of a visible red band and a near infrared band. A higher NDVI represents a greener vegetation, and it is now proven that vegetation indices derived from spectral reflectance, such as the NDVI, are efficient in monitoring plant water stress levels (Ballester et al., 2019; Ihuoma & Madramootoo, 2019). Thus, spectral remote sensing images are also often used for stomatal conductance estimation and modeling (Espinoza et al., 2017; J. Li et al., 2022; H. Li et al., 2022; Matese et al., 2018; Panda et al., 2014; Sobejano-Paz et al., 2020; Y. Zhang et al., 2022; Zhao et al., 2021; Zhu et al., 2020).

1.4.2 Thermal Imaging

Thermal imaging technology measures the infrared radiation generated by an object in the range of 8-14 μm and converts it into a visual image using an infrared detector (Ishimwe et al., 2014). The principle is that any object with a temperature above absolute zero emits radiation to the outside world. Due to the fact that it measures radiation, thermal imaging technology has a quick response time and high accuracy. The technology has been widely used in many fields such as civil engineering, aerospace, and medical industry (Vadivambal & Jayas, 2011). Thermal imaging is often used in agriculture to estimate the canopy temperature and water status of plants (Deery et al., 2016; Zhou et al., 2021; Zia et al., 2013). Although thermal infrared remote sensing cannot measure g_s directly, it is often used for g_s estimation based on the close relationship between the canopy temperature and plant water potential (Berni et al., 2009; Ellsäßer et al., 2020; Espinoza et al., 2017; J. Li et

al., 2022; Matese et al., 2018; Sobejano-Paz et al., 2020; Struthers et al., 2015; Zhao et al., 2021; Zhu et al., 2020).

Table 2: Summary of remotely sensed spectral and thermal imaging for stomatal conductance estimation for various crops.

Reference	Crop	Spectral imaging	Thermal imaging	R^2
Berni, J. A.J (2009)	Olive	×	✓	0.59
Ellsäßer, Florian (2020)	Oil palm	×	✓	0.5
Li, Haojie (2022)	Maize, wheat, rice, soybean	✓	×	0.86
Sobejano-Paz, Verónica (2020)	Maize, soybean	✓	✓	0.66
Zhang, Yuan (2022)	Temperate species	✓	×	0.57
Zhao, Lin (2021)	Maize, soybean, sorghum, sunflower	✓	✓	0.81

1.5 Machine Learning Models

Due to the prevalence of high-performance computing and high-throughput plant phenotyping, various types of Machine Learning (ML) methods have been widely used in different research areas. ML has been shown to accurately predict barley yield (R^2=0.84) and leaf area index (LAI) of trees (R^2=0.75) (Omer et al., 2016; Sharifi, 2021). Unlike the various formulations mentioned earlier that focus differently on the relationships between variables, ML is more like a black box, which tends to focus more on predicting the outcome and ignore the process in between. Although ML models have low explanatory power, they tend to provide more accurate estimations

(Houshmandfar et al., 2021). ML models can be classified into two categories based on their learning methods: supervised learning and unsupervised learning. Supervised learning, as its name suggests, labels the data in the training set and then the algorithm summarizes the method of deducing conclusions from the labeled answers and applies the method to the test set. Therefore, this type of method is mainly used in various regression and classification problems. Unsupervised learning, on the other hand, is a training method with no defined objective, so it does not require labeling of the data, which makes it impossible to evaluate the model's performance. This type of learning method is frequently used for identifying outliers and classification. According to the principles of these two methods, supervised learning is the most prevalent technique employed in agriculture research, and this section describes several models that have been used.

1.5.1 Random Forest

The random forest regression algorithm was proposed by Breiman in 2001 (Breiman, 2001), which is a machine learning algorithm derived from regression trees. Similar to decision trees, regression trees provide output values based to input qualities. In contrast to decision trees, which provide categorical labels, regression trees utilize squared errors to produce regression values. A random forest is called a forest because it uses an integrated algorithm to add many trees and average the final output. Therefore, random forests would have higher accuracy than simple regression trees. With the increased availability of big data, random forests are widely used in the

field of spectral remote sensing, for example to predict soil heavy metal concentration (Tan et al., 2019), crop yield (Prasad et al., 2021), or leaf nitrogen content (Liang et al., 2018) using hyperspectral sensor data. Consequently, this useful machine learning model is also frequently employed by researchers with environmental variables and plant phenotypic data mentioned in the previous section to create estimations on g_s (Brewer et al., 2022; Ellsäßer et al., 2020; Houshmandfar et al., 2021c; Saunders et al., 2021; Zhao et al., 2021).

1.5.2 Support Vector Regression

Support vector regression (SVR) is a machine learning algorithm similar to support vector machine (SVM). Unlike SVM, which is often used in classification problems, SVR is employed in regression situations. In contrast to RF, it takes an entirely different method, which locates the optimal hyperplane in an n-dimensional space that meets the objective, and the fitted line that exists in this hyperplane is the model. It has the advantage of transforming nonlinear problem into linear ones (Geng et al., 2020). Like RF, SVR is often used for yield, plant nitrogen content and g_s estimation utilizing environmental variables as well as spectral data (Houshmandfar et al., 2021; Jarolmasjed et al., 2018, 2018b; Khosla et al., 2020; Minaei et al., 2022; Tian et al., 2018; Y. Zhang et al., 2022).

1.5.3 Partial Least Squares Regression

Partial Least Squares regression (PLSR) is a common machine learning technique

that is frequently employed for multiple-variable data analysis (Mehmood et al., 2012). As a dimensionality reduction approach, it can transform highly correlated independent variables into a smaller number of uncorrelated variables and improve the performance of the model. PLSR is often used in spectral data modeling, by using this dimensionality reduction method, a large number of co-linear spectral variables can be transformed into uncorrelated variables (Darvishzadeh et al., 2008; Fu et al., 2014). In fact, PLSR is the de facto standard approach to model hyperspectral data and to estimate a large array of soil and crop attributes (Wijewardane et al., 2016, 2018, 2023), including leaf water content or g_s (Rapaport et al., 2015; Y. Zhang et al., 2022).

1.5.4 Artificial Neural Network

Artificial neural network (ANN) is a machine learning algorithm inspired by biological neuronal systems. It consists of an input layer, an output layer, and hidden layers, with each layer containing a different number of nodes connected to all the nodes in the next layer. With a significant amount of training data, different weights are assigned to each node to categorize or predict the data. ANN, a modern technique for artificial intelligence, has been effectively applied to a variety of applications, including plant disease detection, plant phenotypic analysis, etc. (Pathan et al., 2020). ANN has also been tested with satisfactory results for g_s estimation (Ellsäßer et al., 2020; H. Li et al., 2022; Vitrack-Tamam et al., 2020; Y. Zhang et al., 2022).

1.5.5 Conclusion

The author reviewed a number of articles that used machine learning methods for g_s estimation, and RF had the highest relative accuracy among the four machine learning methods mentioned previously. Ellsäßer et al. (2020) used weather information and plant phenotypic data derived from thermal and RGB images to predict the g_s of different tree species using multiple linear regression (MLR), RF bagging, RF boosting, SVM, and ANN. RF bagging offered the highest accuracy (R^2 around 0.5) among these four approaches, and ANN was the second-best algorithm (Ellsäßer et al., 2020). Zhao et al. (2021) utilized data such as relative humidity, vapor pressure deficit, air temperature, and canopy temperature to predict g_s for a variety of field crops. Gradient boosting machine (GBM) exhibited the highest accuracy (R^2=0.80) among the three machine learning algorithms tested (MLR, RF, and GBM). In the meantime, researchers have demonstrated that machine learning models are more accurate than classic empirical models. Houshmandfar et al. (2021) and Saunders et al. (2021) compared Jarvis model and Ball-Berry model to machine leaning techniques. By using the same variables in both the machine learning and empirical models, the studies showed that RF increased the R^2 value of the Jarvis model from 0.76 to 0.97, and the Ball-Berry model from 0.41 to 0.75 (Houshmandfar et al., 2021; Saunders et al., 2021). Consequently, we conclude that the application of machine learning techniques can complement and further improve the efficacy of empirical models.

1.6 Challenge and Potential Improvements

The future of digital agriculture is contingent on the development of cutting-edge technologies such as artificial intelligence, IoT, remote sensing technology, etc. However, there are still considerable challenges in the process of development. The first issue to be considered of g_s estimation based on high-throughput phenotypic data is the ground truth data collection. As mentioned in the preceding section, the two contact-type technologies that are primarily employed for field crops, gas exchange measurement and porometry measurement, need researchers to utilize handheld equipment to collect data from individual plants. Both of these data collection technologies need approximately 30 seconds to record a single data point, but a substantial amount of training data is required to allow the modeling of high-throughput phenotypic data. Therefore, researchers are burdened with a great deal of labor due to the vast amount of data necessary. In addition, as pointed out in Section 1.3, the current accuracy of g_s measurements still needs to be improved, and the measurement bias of handheld instruments is one of the reasons for the impact on the accuracy of g_s estimation. Another challenge that needs to be considered is the choice of training variables. Through article search, the authors found that the training variables at this stage are focused on environmental variables as well as plant phenotypic data in the canopy. There is evidence that g_s responds differently based on the duration of radiation exposure (J. Xu et al., 2021). Hence, in the future, time series data or other data that reflect plant water potential, such as soil moisture, might be

used to estimate g_s. Overall, g_s estimation may be achieved efficiently by combining high-throughput plant phenotypic data with machine learning. As the need for food and crops produced for biofuel rises, good estimation of stomatal conductance can significantly boost crop yields and reduce the negative effects of climate change on plants. With the constant development of computers and sensors, technological developments can improve the accuracy of g_s estimation. In the research field, accurate and quick collection of g_s data can assist researchers in identifying drought-tolerant genotypes. In production, these tools can help farmers optimize irrigation and fertilization schedules to maximize crop yields while minimizing cost.

CHAPTER 2 ESTIMATING CROP STOMATAL CONDUCTANCE USING PHENOTYPIC TRAITS AND WEATHER VARIABLES THROUGH MACHINE LEARNING

2.1 Introduction

On the epidermis layer of higher plants, two small symmetric guard cells are present known as stomata, which play a crucial role in gas exchange between inner air space of the leaf and the outer atmosphere (Lawson, 2009). The opening and closing of stomata to regulate the flux of carbon dioxide into and water vapor out of the leaf is measured by stomata conductance (g_s). In addition to carbon dioxide and water vapor, stomata also control the exchange of various trace gases. Determining and estimating stomata conductance of plants is therefore imperative for us to monitor and understand plant growth, development, and productivity (Otu-Larbi et al., 2021). Finding the best model to determine g_s has been a key factor for phenotypic studies and optimization of crop water use in plants (Thorp et al., 2018). The crop's estimated g_s can be used to determine the drought sensitivity (Buckley & Mott, 2013), and also formulate decisions for irrigation scheduling (Jones, 1999). Under conditions of high temperature, the heat stress response of the plant cools the leaves by increasing their g_s (Jagadish et al., 2021). Thus, estimating g_s can be useful for conducting heat tolerance studies on plants and responses to extreme climate conditions.

Canopy temperature has long been used to make an empirical estimate of plant water

stress. Decreased water uptake in plants closes stomata, which cause reduction in transpiration rate in plants and increases leaf temperature (Blonquist et al., 2009). Therefore, stomatal conductance could be inferred from canopy temperature through photosynthesis, respiration and other biophysiological processes (Jones, 2004). Stomata creates a strong influence on plant characteristics associated with photosynthesis and transpiration, and controls temperature and water use efficiency (WUE) which is vital for the survival of plants. During short term water stress, plant increases their WUE by reducing stomatal aperture and transpiration rate (Y. Li et al., 2017). When trees are transplanted from their native environment to a drier one, plants would reduce stomatal conductance caused by a reduction of stomata density (Gindel, 1969). Similarly, environmental conditions play a significant role in g_s. Increase in atmospheric CO_2 concentration and temperature results in a decrease of stomatal density (Franks et al., 2015). Therefore, g_s is an important target for crop improvement, irrigation management and development of drought-tolerant cultivars (Doheny-Adams et al., 2012).

Due to climate change, precipitation patterns have altered drastically around various regions of the world, causing severe and frequent drought stress to many crop species (Wehner et al., 2011). WUE can be used to monitor plant growth, development and drought stress (Thorp et al., 2018), and knowledge of WUE provides an effective understanding of plant water content. Remote sensing and precision agriculture has developed precision irrigation technology, which can be used to estimate crop evapotranspiration (ET) and estimate drought stress in crop at

multiple scales (Maes & Steppe, 2012). Thermal infrared images captured by IR cameras mounted on unmanned aerial vehicles (UAVs) provide sub-meter spatial resolution, capable of retrieving pure canopy temperature and minimizing soil thermal effects. These high-resolution thermal IR images retrieve the energy fluxes from pure vegetation on open canopies where most remote sensing-based method are not able to perform well (Berni, Zarco-Tejada, Suarez, et al., 2009; Herwitz et al., 2004; Sugiura et al., 2005).

Contact-type g_s measurements in the field can be divided into two categories: (1) Gas exchange measurements: g_s is derived by measuring the gas exchange rate of water vapor and CO_2 between the leaves and the atmosphere using infrared spectroscopy. (2) Porometry measurements: Using several humidity sensors measure the diffusion rate of water vapor from the leaves (Toro et al., 2019). Canopy temperature is a useful indicator of stomata conductance which can be measured by using thermometers in field plots (Takai et al., 2010; Yu et al., 2016), but periodic measurement is labor intensive, time consuming and expensive. Improvements in remote sensing techniques in the past decades have enabled image acquisition technologies such as LANDSAT, SPOT and thermal imaging to estimate stomata conductance, surface temperature and various vegetation indices (VIs) (Sadler et al., 2002). Nowadays, sensors fitted on small, portable, light weight platforms (Bai et al., 2019) , and field sensors combined with Internet of Things (IoT) (Chamara et al., 2022) and data transmission networks were designed to acquire thermal infrared images. Additionally, the problem of coarse spatial resolution from satellite imagery is

alleviated by these IR sensors installed on platforms and UAVs (Colaizzi et al., 2017). Therefore, it will be possible to quantify g_s across a vast area using remote sensing images.

Three approaches have been used to model g_s: empirical model (data-dependent statistic model), mechanistic model (models that rely on the biology principle of plants), and economic model (describe the water exchange of the plants from cost and benefit under a controlled environment). The method presented by Jarvis (1976) was an empirical model to include five variables; light, temperature, carbon dioxide, vapor presser deficit, and leaf water status; to estimate g_s (Buckley & Mott, 2013). Environmental variables such as vapor pressure deficit, relative humidity, canopy temperature, dew point temperature, air temperature, solar radiation, and soil moisture influence g_s. In recent years, studies have been focused more on measuring canopy temperature using remote sensing, which provides a basis for estimating g_s in plants (van Dijk et al., 2021; Virnodkar et al., 2020; Zhao et al., 2021).

In recent years, efforts have been made to find associations between these variables from empirical modeling, by leveraging the cutting-edge machine learning approach. Some studies combined machine learning with empirical models like Jarvis and Ball-Berry. The results of the study showed that machine learning models, such as random forests, performed better than traditional statistical models. Machine learning can improve the R^2 values of two empirical models from 0.76 to 0.97 (Jarvis model) and from 0.41 to 0.75 (Ball-Berry model) when using the same variables. (Houshmandfar et al., 2021; Saunders et al., 2021). However, it is important to note

that most studies have used variables from current empirical models for estimation (which include data that are difficult to gather, such as CO_2 concentration), whereas just a small number of studies have linked remote sensing data to g_s. If g_s could be estimated in real time, it would be feasible to optimize plant irrigation schedules and determine plant drought resistance using rapid in situ measurements. (Jones, 1999). In this project, two sources of data were used, near real-time high spatial resolution plant phenotyping data and weather data collected from a field phenotyping platform and an on-site weather station. Three ML models were used including Random Forest Regression (RFR), Support Vector Regression (SVR), and Partial Least Squares Regression (PLSR). The objectives of this study are three folds: (1) investigate the relationship between g_s and a series of environmental and plant phenotypic variables acquired via automated sensors; (2) build models to estimate g_s from these environmental and phenotypic variables; and (3) explore the benefit of including plant phenotypic data on g_s estimation.

2.2. Materials and Methods

2.2.1 Site Description and Data Collection

The research was conducted in a research field (41°08’41.9”N 96°26’20.5”W) at the University of Nebraska-Lincoln’s Eastern Nebraska Research, Extension and Education Center in 2020 and 2021. This field is dedicated for high throughput plant phenotyping research and equipped with a fully autonomous sensor platform (Referred to as NU-Spidercam) to acquire plot-scale phenotyping data including

thermal-infrared images, multi-spectral images, and canopy reflectance spectra (Bai et al., 2019).

This study involved five crop types: maize, soybean, sunflower, sorghum, and winter wheat. All these crops were planted in plots of 2 rows, 3 rows, or 6 rows, with a row spacing of 0.76 m and a row length of 6.1 m. The exception was winter wheat which was planted as 5 rows in each plot, with a row spacing of 0.19 m and a row length of 4.6 m. Detailed information on the crop type, plot size, date of data collection, and number of genotypes in each year is given in Table 3 and Figure 3B. Handheld leaf porometers (SC-1, METER Group, Pullman, WA) were used to measure g_s from all experimental plots, The leaf porometer sets up a diffusion pathway for water vapor in a chamber with a fixed diameter and length and derives leaf g_s by two relative humidity sensors along the diffusion pathway. In June, July, August, and September 2020, data on soybean, sorghum, sunflower, and corn were collected for 18 days. In May, June, July, August, and September of 2021, data on winter wheat and maize were collected for a total of 24 days. In 2020, readings from three plant leaves per plot were made and the average was taken. In 2021, only one reading per plot was made, due to the relatively long time needed for each g_s measurement. The sensors needed to be kept in equilibrium with the external humidity after each measurement. Depending on the weather conditions of the day, each measurement took approximately 4 minutes.

Table 3: Crop type, plot setup, and measurement date of the study

Crop	Plot Setup	Measurement	Genotypes	Number of

Type		date		observations
Soybean	12 plots, 6-row plots	June 23, 29, July 6, 22, August 3, 14, 26, 28 in Year 2020	8	116
Sorghum	12 plots, 6-row plots	June 17, July 2, 6, 22, August 14, 26, 28 in Year 2020	1	99
Sunflower	12 plots, 3-row plots	July 22, August 3, 14 in Year 2020	2	64
Maize	28 plots (6 rows) and 18 subplots (2 rows)	June 25, July 2, 8, 16, 23, August 11 in year 2020, and August 2,12, 16 in Year 2021	26	299
Winter Wheat	80 Plots	May 7, 12, 13, 26, June 1, 3, 7, 9 in Year 2021	20	592

The variables used in this experiment can be categorized into two groups: plant phenotypic data and weather data. The plant phenotypic data were collected from the NU-Spidercam platform equipped with a thermal infrared camera, a multispectral camera, and a VNIR spectrometer (Figure 3A) to collect canopy temperature, soil temperature, plant coverage, and several canopy VIs at five meters above each plot. For each plot, the sensor system moved to the above area based on the predetermined map coordinates and captured RGB and near-infrared (NIR) images, an IR image, point clouds, and incoming and canopy reflectance VNIR spectra using the onboard sensors.

Canopy coverage rate was calculated as a ratio of the canopy pixel counts over the image size. A semi-automatic image segmentation algorithm, which utilized RGB and NIR images, was applied to segment the canopy pixel from the background (Bai et al., 2019). Co-registration of the multispectral images and the IR image was done

by an affine transformation, which was possible due to the fixed layout of the two cameras. The average canopy temperature was subsequently calculated after carrying out the segmentation on the transformed multispectral images. The average temperature of the background pixels was computed as the soil temperature (Figure. 3C).

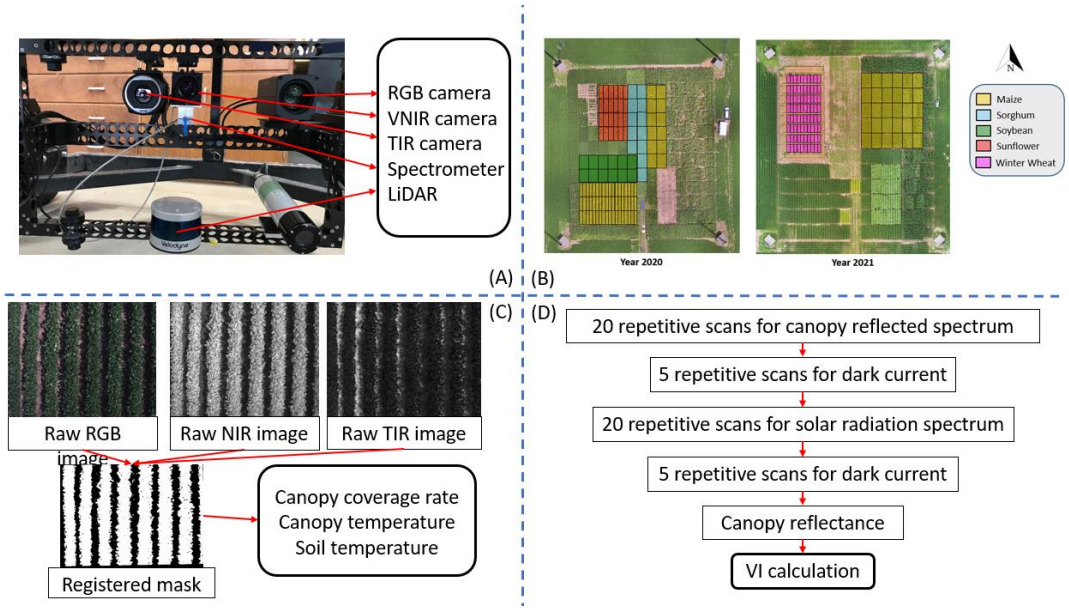

Figure 3: NU-Spidercam data collection and preprocessing. (A) Spidercam sensing platform; (B) Detailed map of the plant planted in 2020 and 2021; (C) Image processing from VNIR and TIR images; (D) The calculation protocol of Vegetation Indices.

A dual-head spectrometer-fiber system collected the incoming and reflected solar radiation at the visible-NIR region, the VIs are calculated as shown in Figure 3D. Wavelength and spectral calibration were done at a seasonal basis using NIST-certified light sources in the lab and reference targets with known reflectance in the field. Seven VIs (Table 4) were calculated from the canopy reflectance, including

Normalized Difference Vegetation Index (NDVI), Red-Edge NDVI (NDRE), Optimized Soil Adjusted VI (OSAVI), Reflectance-based Solar-Induced Fluorescence (SIF), Photochemical Reflectance Index (PRI), Near-infrared Reflectance of Vegetation (NIR_V), and Transformed Chlorophyll Absorption in Reflectance Index (TCARI) (Badgley et al., 2017; Haboudane et al., 2004; Kim et al., 2011; Rondeaux et al., 1996).

Table 4: Equations used to calculate the vegetation indices (VI) used in this study. The numbers after R represent the central wavelengths of the spectral bands involved in calculating the VI.

NDVI	$\frac{\mathbf{R750 - R705}}{\mathbf{R750 + R705}}$	(1)
NDRE	$\frac{R770 - R730}{R770 + R730}$	(2)
OSAVI	$(1 + 0.16) * \frac{R800 - R670}{R800 + R670}$	(3)
SIF	$\frac{R800 - R740}{R800 + R740}$	(4)
PRI	$\frac{R570 - R531}{R570 + R531}$	(5)
NIR_V	$R800 * \frac{R800 - R670}{R800 + R670}$	(6)
TCARI	$\frac{3 * (a - 0.2 * b * c)}{(1 + 0.16) * \frac{d}{e}}$	(7)
	*Where: $a = R700 - R670$ $b = R700 - R550$ $c = R700/R670$ $d = R800 - R670$ $e = R800 + R670 + 0.16$	

The weather variables used in the estimations were obtained from an on-site weather station, which is located approximately 30m west of the field and part of Nebraska Mesonet (https://mesonet.unl.edu). All The environmental variables were

recorded by the weather station at a 1-minute interval.

2.2.2 Data Preprocessing

Air temperature (Ta), dew point temperature (Tdew), relative humidity (RH), wind speed, and shortwave solar radiation were used as weather variables. Ta, Tdew, and RH were measured 2m above the ground; and wind speed and shortwave solar radiation were measured 3 m above the ground. Vapor pressure deficit (VPD) was calculated as the difference between saturation vapor pressure and actual vapor pressure (Allen et al., 1998).

$$VPD = SVP - AVP$$

$$SVP = 0.6108 * \exp\left(\frac{17.27 * T_a}{T_a + 237.3}\right)$$

$$AVP = 0.6108 * \exp\left(\frac{17.27 * T_{dew}}{T_{dew} + 237.3}\right)$$

Where:

SVP is saturation vapor pressure

AVP is actual vapor pressure

Since the NU-Spidercam is running continuously, it is necessary to match the g_s data as well as the variables in time, both plant phenotype data and weather data were used closest to the time of g_s collection. The weather station in the field records data every minute, so the weather data and the g_s can be matched perfectly. Spidercam sensors took less then 30s to collect data for one plot and had other experiment plots

to cover. Therefore, there were differences of the measuring time between NU-Spidercam and the manual g_s data collection. In this study, only data with time intervals less than 3 hours were used in the analysis.

2.2.3 Model Training and Optimization

Prior to model training, all variables were first standardized to speed up the training as well as convergence of ML models. Three ML models were attempted in this study, namely, PLSR, RFR, and SVR. PLSR models are generally considered linear whereas RFR and SVR models are nonlinear. Because one of our objectives was to explore the benefit of phenotyping data in g_s estimation, models were trained and compared with three scenarios: (1) only weather data (T_a, T_{dew}, RH, wind speed, solar radiation, and VPD) was used as input variables, (2) only the sensor-based phenotypic data (plant coverage, canopy temperature, soil temperature, and seven VIs) was used as input variables, and (3) combining weather data and sensor-based phenotypic data as input variables.

The RFR algorithm builds a large number of decision trees by randomly selecting variables in the training set, and finally averaging the results of each decision tree to obtain the final estimation value (Wang et al., 2016). RFR employs the bagging ensemble learning method by dividing the training set into distinct subsets for separate training. Multiple models are ultimately combined to enhance performance. SVR, on the other hand, fits the data by constructing a hyperplane in a multidimensional space (Tian et al., 2018). The final model of SVR is the best-fit line

that contains the most data points within a specified range on both sides of the hyperplane. PLSR is the algorithm that reduce the estimator variables into several uncorrelated inputs then apply the least square regression (Rasooli et al., 2022).

Optuna (Akiba et al., 2019) was used to perform hyperparameter tuning under Python environment for RFR and SVR model, which improves the accuracy of the model by sampling and pruning algorithms to continuously reduce the search space of hyperparameters. The number of estimators and the depth of the trees for RFR was the hyperparameters for turning. Additionally, kernel, C, and Gamma (kernel is the function used in the model, C is the regularization parameter, and Gamma is the kernel coefficient) were used for SVR. After dividing the dataset into a 75% and 25% training and testing set, a 5-fold-cross-validation was performed on the training set to determine the best hyperparameters with the lowest mean squared error. These hyperparameters were used in the training set to obtain the model, which was evaluated using root mean squared error (RMSE) and coefficient of determination (R^2) obtained from the test set.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}$$

$$RMSE = \sqrt{\sum_i^n \frac{(y_i - \hat{y}_i)^2}{n}}$$

Where:

y_i is the measured value and $\hat{y}_i$ is the estimated value.

$\bar{y}_i$ is the average value of n samples in the test set.

2.3 Results

The box plot (Figure 4) shows the distribution of measured g_s for different crop types. The soybean has the largest fluctuation range (0.063 mol m^{-2} s^{-1} to 1.855 mol m^{-2} s^{-1}). According to the image, maize has a large number of outliers, which may be due to the differences in crop characteristics between genotypes. In addition to maize, sunflower and winter wheat also exhibited a small number of outliers.

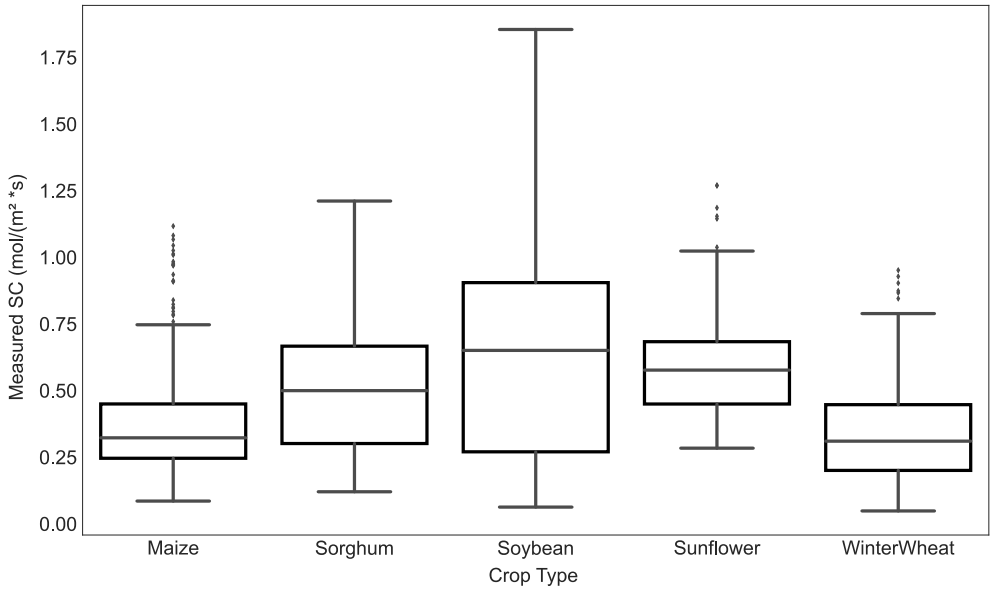

Figure 4: Boxplot for measured stomatal conductance between all crop types.

Table 5 shows the performance of three different input variables under five different models and six different crop types. Analysis of the model results showed that RFR performed best among all five ML models for g_s estimation for all crops. For the RFR model that makes estimation for all six crops together, combination inputs have an R^2 value of 0.69, with RMSE of 0.135 mol m^{-2} s^{-1}, Plant phenotyping data estimation has an R^2 value of 0.59, RMSE of 0.158 mol m^{-2} s^{-1}, and weather only estimation has an R^2 value of 0.58, with RMSE of 0.161 mol m^{-2} s^{-1}. The result shows

that although the Spidercam collection time has an interval ranging from 0-3 hours with the g_s collection time, the estimation result from plant phenotypic data is still better than the result from weather information. The model with the poorest performance for the entire crop was PLSR, with estimations of $R^2 = 0.41$, $R^2 = 0.26$, and $R^2 = 0.17$ for three sets of variables. In the terms of single species estimation, the RFR model for sorghum shows the best performance, with an R^2 value of 0.86, RMSE of 0.131 mol m^{-2} s^{-1}. This is consistent with the findings regarding the entire crop, but it is unknown if the sorghum result is affected by the validation bias of the small data set.

Table 5: The performance for estimating stomatal conductance with three different machine learning models and five crop types.

		Random Forest Regression			**Support Vector Regression**			**Partial Least Squares Regression**		
		All	Spider only	Weather only	All	Spider only	Weather only	All	Spider only	Weather only
Total	R^2	**0.69**	0.59	0.58	0.66	0.61	0.53	0.41	0.26	0.17
	RMSE	**0.14**	0.16	0.20	0.15	0.16	0.21	0.20	0.22	0.23
Maize	R^2	0.62	0.61	0.35	**0.67**	0.30	0.32	0.48	0.35	0.10
	RMSE	0.12	0.12	0.16	**0.11**	0.17	0.16	0.15	0.16	0.19
Sorghum	R^2	**0.86**	0.80	0.78	0.71	0.49	0.58	0.69	0.46	0.55
	RMSE	**0.13**	0.16	0.15	0.17	0.24	0.21	0.18	0.23	0.21
Soybean	R^2	**0.79**	0.49	0.67	0.62	0.35	0.58	0.53	0.36	0.53
	RMSE	**0.18**	0.17	0.26	0.24	0.30	0.19	0.28	0.31	0.23
Sunflower	R^2	0.83	0.60	**0.86**	0.74	0.62	0.77	0.77	0.53	0.81
	RMSE	0.10	0.15	**0.09**	0.12	0.14	0.11	0.11	0.15	0.10
Winte	R^2	**0.31**	0.22	0.20	0.21	0.13	0.21	0.20	0.13	0.21

r wheat	RMSE	**0.14**	0.14	0.14	0.15	0.15	0.14	0.15	0.15	0.14

Figures 5, 6, and 7 show the estimation results based on all data, Spidercam data, and weather information in RFR model, respectively. The best performing model is all data model of sorghum, while the worst performing model is the weather information model of winter wheat, with an R^2 value of only 0.20. Sunflower and soybean's weather information models outperformed their phenotype data models. The plant phenotype models for winter wheat, sorghum, maize, and the entire crop set performed better than the weather information models. The images indicate that the combined model is closer to the measured values than either the model of plant phenotype data or the model of the weather information alone. In general, the performance of the plant phenotype data model is slightly superior to that of the weather information model. In conclusion, the results indicate that by merging weather information with plant phenotyping data, the accuracy of results' estimation can be significantly enhanced.

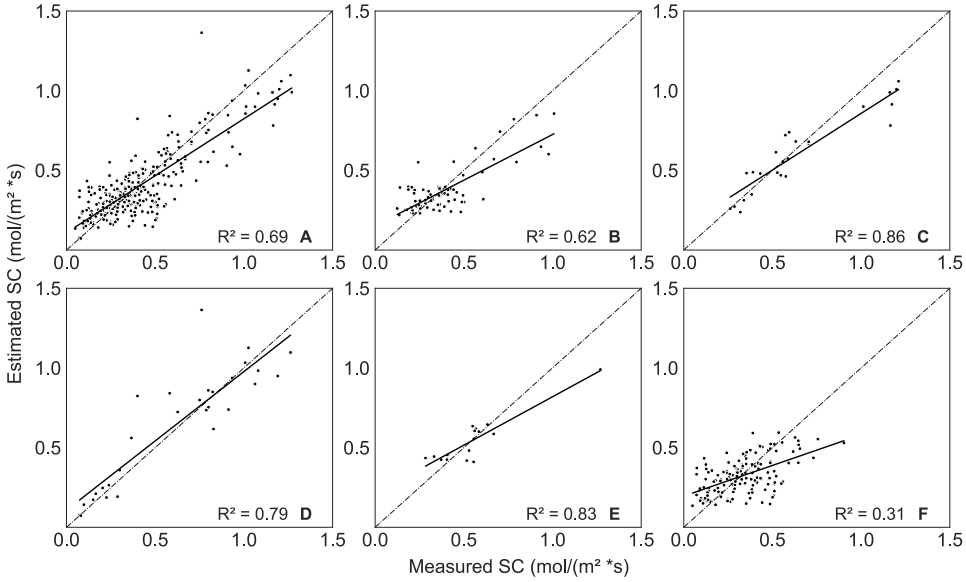

Figure 5: g_s estimated based on all data using the random forest model. (A) All crop

types. (B) Maize. (C) Sorghum. (D) Soybean. (E) Sunflower. (F) Winter wheat.

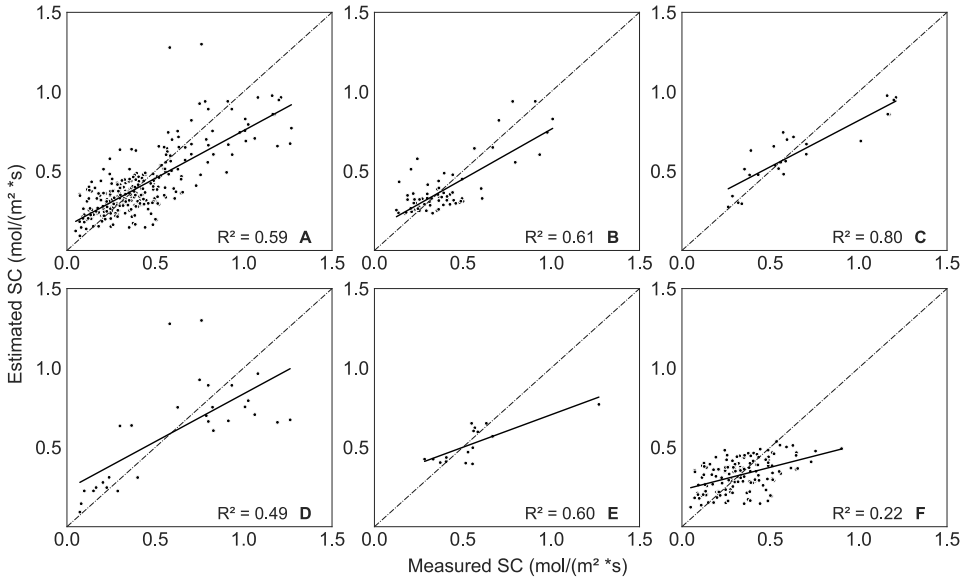

Figure 6: g_s estimated based on Spidercam data. (A) All crop types. (B) Maize. (C) Sorghum. (D) Soybean. (E) Sunflower. (F) Winter wheat.

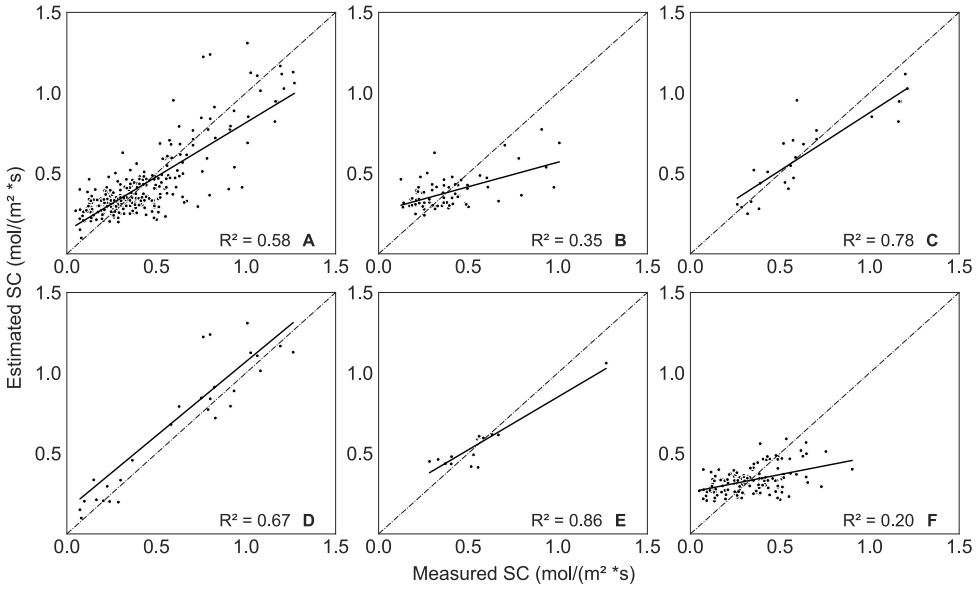

Figure 7: Stomatal conductance estimated based on weather information. (A) All crop types. (B) Maize. (C) Sorghum. (D) Soybean. (E) Sunflower. (F) Winter wheat.

2.4 Discussion

2.4.1 Demonstrated the Potential of Estimating Stomatal Conductance Based on Plant Phenotyping Information

Compared to estimate g_s using weather information alone, the estimation using plant phenotypic data and the combination of both data as estimator variables improved the R^2 values 17% (0.10) and 19% (0.11), respectively. The temperature of the canopy is affected by environmental factors such as air temperature, wind speed, etc. (Rebetzke et al., 2012), so canopy temperature can represent weather data to some extent. In our study, we also discovered that plant phenotypic data generally made more accurate estimations than weather data. Since stomata is mainly used for water and CO_2 exchange between plants and atmosphere, the water status of the plant and the atmosphere CO_2 concentration are two influential factors (Buckley, 2017). Plants subjected to water stress and elevated CO_2 levels will close their stomata to reduce the transpiration water loss. Therefore, numerous models incorporate these two crucial variables to improve their accuracy. Take Jarvis model as an example, the variables are light intensity, leaf temperature, VPD, atmospheric CO_2 concentration, and leaf water potential. Among them, leaf water potential as well as leaf temperature belong to leaf phenotypic information, the NDVI and OSACI/TCARI used in this experiment can be used for plant water status assessment by showing the change in chlorophyll content of the crop (Baluja et al., 2012). Other VIs included in this study have likewise been shown to quantify the state of crop under biotic as well as abiotic

stresses (Khan et al., 2018; Motohka et al., 2010; Rutkoski et al., 2016). However, dueto the fact that each species or genotype responds differently to environmental changes (Mathe-Gaspar et al., 2005), using weather information alone might not adequately capture the state of the crop compared to plant phenotypic information. This is likely the reason why the ML models using weather information had overall lower accuracy than those using plant phenotypic information only. Meanwhile, the ability of plants to regulate the opening and closing of stomata varies among different crops and genotypes (Banerjee et al., 2020; Lopes et al., 2011). This represents an additional source of variation in the results when plant traits are used in an attempt to make a unifying estimation of g_s for different genotypes of crops. There is also a study demonstrated that canopy temperature and g_s differed between different wheat varieties in sunny and cloudy conditions (Takai et al., 2010). This argument is also supported by the estimation results for a single genotype crop, sorghum, in this study. When using RFR for sorghum estimation, it has the highest accuracy among all crop types with an R^2 value of 0.86. Our study shows that the use of plant phenotype data can be used for g_s estimation and reduces the need for handheld instruments when collecting data in the field.

2.4.2 Performance of Machine Learning Models

In this study, three ML models were used to estimate crop g_s from sensor-based plant phenotypic data and weather data. Where RFR demonstrated the best estimation performance, the results show that the model can be used for g_s estimation for a wide

range of crops. This suggests that decision tree-like machine learning algorithms perform better on similar problems, a conclusion that is also supported by other paper (Houshmandfar et al., 2021; Saunders et al., 2021), which tend to have better performance of RFR models when estimating stomatal conductance This class of algorithms can also cope with complex interrelationships among variables. Since the effect of weather information on crops is nonlinear, the PLSR model does not perform as well as other models (Cai et al., 2019). In other words, the nonlinear model is more suitable than the linear model for g_s estimation. From the experimental results, it can be observed that RFR can estimate up to $R^2 = 0.69$ for all crops while PLSR is only $R^2 = 0.41$.

Figure 8 shows the importance of the variables in the RFR, SVR, and PLSR model, for the RFR model, it shows that T_{dew} occupies the most important role in the model, followed by SIF, RH and NDRE in order of importance. It is noteworthy that all weather variables except T_{dew} as well as RH have low importance. However, for SVR and PLSR, the most important variables were NDRE and RH, respectively. Although the feature importance varies for each machine learning model, it can be observed that the RH and several VIs plays an important role in estimation. The first six important variables in all three models included NDRE, relative humidity, and SIF, which covered several important factors affecting g_s, i.e., chlorophyll content, environmental water content, and photosynthetic efficiency.

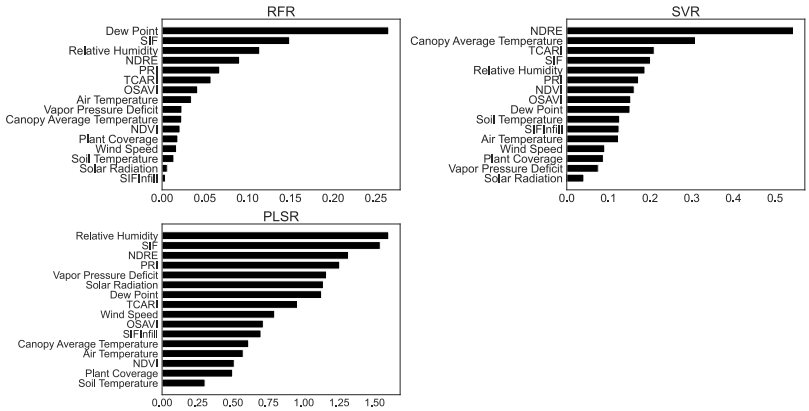

Figure 8: Feature importance of all variables used in RFR, SVR, and PLSR

In contrast to conventional empirical formulas like Jarvis model, ML models are primarily concerned with estimating outcomes rather than discovering the relationship between individual variables. Empirical formulae based on scientific theory tend to be more explanatory and have less risk and uncertainty, and with the addition of computer technology the accuracy of the model can be further improved (Houshmandfar et al., 2021). These ML models are not generalizable because they only apply to data generated under similar conditions. During the model training, it is also important to consider the amount of data for each crop in the dataset, and the quality of the dataset will directly impact the results (Saunders et al., 2021). Indicators of the dataset's quality include accuracy and the representative of ground truth data, among others (He & Garcia, 2009). The proportion of each data type in the dataset will have a direct impact on the performance of the model, including but not limited to crop species and their growth stages. Theoretically, the better the data set, the better the model performs (R^2 value for the all-species data set is 0.70 in our study). It is worth noting that the sunflower model with only 64 observations ($R^2 = 0.83$) and the sorghum model with 99 observations ($R^2 = 0.86$) still perform very well, this might

because of the fewer genotypes included in the data set. Similarly, expanding the scope of the data to include additional regions, data collection devices and species would increase the generalizability of the model. Consequently, the performance of the ML models could be further improved with a comprehensive data set comprising samples from different times, weather conditions, and field conditions.

2.4.3 Challenges and Future Improvements

In our study, we took three measurements per plot in 2020 and calculated the average value, whereas in 2021, only one measurement per plot was taken for all crops. Since stomata are induced by light, measurements performed in the same plot can also differ depending on the leaf orientation to sunlight (Bagley et al., 2015). So a single measurement does not accurately represent the average g_s of each plot (winter wheat has the lowest accuracy). In the future, data will be gathered by randomly measuring more readings per plot and averaging the result. In addition, noise reduction is one of the primary focuses of the data processing in the field (van Dijk et al., 2021). Data collection in the field will inevitably be affected by environmental factors such as wind, so how to make the collected data more stable is also a key point to consider. Although g_s can be measured quickly using a leaf porometer, the accuracy of the measurement is still affected by the dependence on environmental factors such as temperature and relative humidity due to the technical limitations (Fanourakis et al., 2016; Toro et al., 2019). Van Dijk et al. (2021) also mentioned the results from the end-to-end learning model, so we can also consider skipping the explicit feature

extraction from raw sensor data, such as extracting canopy temperature and VIs, and using multispectral and NIR images directly for g_s estimation. Another limitation is the quality of the data set, as stated in the preceding section, the diversity of the data has an effect on the model results. Therefore, determining how to filter the existing data and expand the size of the dataset should be a top priority for the future research. It has been demonstrated that weakly associated estimator variables reduce the accuracy of regression models (Boulesteix et al., 2015), so the selection of variables also needs to be considered according to their relevance to g_s. In the future, a comparison between end-to-end models and variable filtered models could be considered in order to evaluate their performance differences.

Numerous models based on the Jarvis model incorporate soil water status to determine the impact of water stress on g_s (Damour et al., 2010). Therefore, one missing component in this study was the soil data – mainly soil water content but could also include soil textures, organic matter, and others. Currently we treat all measurements in the model as independent samples. In other words, our models do not have the ability to account for the time dependence among these variables. On the other hand, photosynthetic acclimation is one of the ways in which plants adapt to environmental changes (Bagley et al., 2015). It has been reported that the interaction between CO_2 levels and water stress has a positive effect on g_s over time (Thruppoyil & Ksiksi, 2020). So we speculate that the dynamics of stomata opening and closing are not only dependent on the current state, but also the past state of the weather variables. Therefore, we will investigate more advanced approaches (such as recurrent

neural network) to accommodate the time-series weather data in the modeling.

2.5 Conclusion

The study employed three ML algorithms to estimate the leaf g_s of five crops using near real-time, sensor-based plant phenotypic data and weather data. Combining the plant phenotypic data with weather data significantly improved the predication accuracy of g_s, with the model testing R^2 value of 0.69 (increased from testing R^2 of 0.58 and 0.59 for each data set alone). This result indicated the usefulness of sensor-based plant phenotypic data in boosting the accuracy of plant g_s modeling and estimation. Because these sensor-based plant phenotypic data could be acquired rapidly and relatively inexpensively, thanks to the advancement in high-throughput plant phenotyping, this study provides a rapid method to quantify leaf g_s more accurate than relying on weather data alone. In terms of ML algorithms, nonlinear RFR and SVR models outperformed the linear PLSR models, suggesting intrinsic nonlinear relationships between leaf g_s and the estimator variables, as well as the complex control and feedback mechanisms among these variables. The plant phenotypic data in our study was collected from the NU-Spidercam platform. However, the approach described in this study could also be readily used for UAVs or other ground platforms with similar sensors mounted. Rapid and accurate estimation of plant leaf g_s, as demonstrated in this study, have implications on applications such as screening for drought-tolerant genotypes and precision irrigation management.

CHAPTER 3 ESTIMATING MAIZE STOMATAL CONDUCTANCE FROM RGB, NIR, AND THERMAL INFRARED IMAGE

3.1 Introduction

As the global population increases, so does the demand for food production. By 2050, the demand for food will double compared to what it was in 2009 (Fukase & Martin, 2020). Hence, increasing food production has been a subject of growing concern. Corn, for instance, is the third most-consumed cereal in the world and is utilized for food, feed, and biofuel production (Erenstein et al., 2022). Another widely grown field crop is soybean, which is used as one of the largest sources of vegetable oil and animal feed due to its high content of vegetable protein and oil (Pagano & Miransari, 2016). In order to increase the corn and soybean yield, the factors affecting yield have been widely explored. It has been demonstrated that the yield of maize and soybean is directly related to its drought index during the grain filling and reproductive growth periods (Mishra & Cherkauer, 2010). Hence, it would be advantageous to increase yield if the water status of crops could be accurately measured.

The stomata are one of the primary channels for water and gas exchange between the plant and the atmosphere (Lawson, 2009). The stomata consist of two guard cells in the epidermis of higher plants, which are regulated to control the opening and closing of the stomata. In photosynthesis, the plant uses light energy to transform

water and CO_2 absorbed through the stomata into energy for plant growth. However, the opening of stomata also leads to accelerated loss of water into the atmosphere due to transpiration. Therefore, when the crop is under water stress, stomata are closed to conserve water. Stomatal conductance (g_s), which indicates the rate of gas exchange between the crop and the atmosphere per unit area, is frequently used to determine the degree of stomatal closure. Hence, if stomatal conductance can be measured effectively and precisely, it can aid in understanding the crop's current water status and improve the irrigation planning.

However, accurate g_s measurement requires contact-type instruments to be clamped on the leaf blade, which is a time-consuming and labor-intensive measurement approach. Gas exchange measurement and porometry measuring are two frequent ways of measurement. Using an infrared spectrometer and relative humidity sensors, respectively, these two methods of measurement determine the gas exchange rate between the leaf and the atmosphere (Toro et al., 2019). In order to estimate g_s rapidly, researchers have attempted to model g_s using empirical models. Using environmental variables including light intensity, leaf temperature, vapor pressure deficit, leaf water content, and atmospheric CO_2 concentration, the Jarvis model (1976) has produced accurate estimations of g_s (Jarvis, 1976). As the model relies solely on environmental factors for assessment, it greatly simplifies the estimation of g_s.

Along with the rise in popularity of machine learning, it has been discovered that the estimation accuracy of models is greatly enhanced when machine learning

models are combined with conventional empirical models (Houshmandfar et al., 2021). With the advancement of remote sensing technology, platforms such as satellites and Unmanned Aerial Vehicles (UAVs) have increased the potential for rapid, large-scale data collection. Thermal infrared cameras fitted into these platforms enable large-scale crop canopy temperature observations. As a fundamental way to regulate transpiration, there is a significant relationship between g_s and canopy temperature (Dai et al., 2004). If g_s can be monitored by these remote sensing platforms, growers will have a better understanding of the current crop status and would be able to plan accordingly.

In this study, we evaluated the estimation accuracy of multiple linear regression (MLR), support vector regression (SVR), and convolutional neural network (CNN) for maize and soybean g_s employing near-infrared, thermal-infrared, and RGB images gathered from a remote sensing platform in the field. We also evaluated whether omitting soil moisture from the model has a substantial impact.

3.2 Material and methods

In 2022, the research was done at the University of Nebraska-Lincoln's Eastern Nebraska Research, Extension and Education Center (41°08'41.9"N 96°26'20.5"W). The total experimental field was 60 by 67 meters and was divided into 128 plots, each plot of 4.6 by 6.1 meters. The two crops in this experiment were maize and soybean, with 18 maize plots and 12 soybean plots (Figure. 9). Throughout the months of August and September, a total of five days were devoted to the collection of data.

Table 6 lists the crop types, plot situation and dates of collection.

Figure 9: Location and detailed map for plots

Table 6: Crop type, plot setup, and measurement date of the study

Crop Type	Plot situation	Measurement date	Number of observations
Soybean	12 plots (6 irrigated plot, 6 non-irrigated plots)	August 8, 10, 17, 22, September 4 in Year 2022	60
Maize	18 plots (9 irrigated plot, 9 non-irrigated plots)	August 8, 10, 17, 22, September 4 in Year 2022	86

Handheld leaf porometers were used to collect all ground truth data (SC-1, METER Group, Pullman, WA). The g_s was calculated using two humidity sensors to measure the diffusion rate of water vapor on the abaxial surface of the leaf. For each measurement, the uppermost fully developed leaf of three randomly selected plants was measured and averaged for each plot. To prevent the effect of shadows on photosynthetic efficiency, only leaves that were totally exposed to sunlight were selected.

The images as well as the data used for estimation came from a fully autonomous

sensor platform (NU-Spidercam) in the experimental field. The platform is equipped with a thermal infrared camera, an RGB camera, and a near-infrared camera (Figure. 10A). The images captured by these cameras are used in the CNN model, while the variables utilized by the MLR and SVR models must be computed. These variables include canopy coverage rate, canopy temperature, and soil temperature. By importing the RGB and NIR images into a semi-automatic image segmentation algorithm, the canopy pixels are segmented from the background to calculate the canopy coverage rate. The segmented canopy mask is then merged with the IR image to obtain canopy and soil temperatures (Figure. 10B) (Bai et al., 2019b). Soil moisture data for comparing model performance were obtained from soil moisture sensors (TEROS 10, METER Group, Pullman, WA) installed at the center of each plot (Figure. 10C). These sensors continuously collected soil moisture readings at 15-minute intervals, and we selected the subset of soil moisture measurements that corresponded most closely to the collection time of the ground truth data.

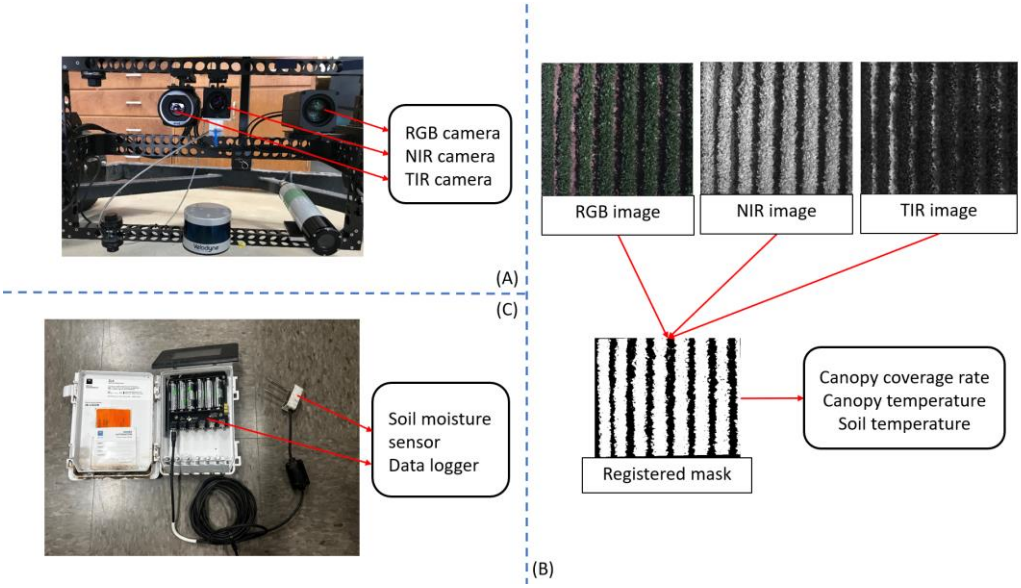

Figure 10: NU-Spidercam data collection and preprocessing. (A) Spidercam sensing

platform; (B) Example image and image processing procedure for calculating canopy coverage rate, canopy temperature and soil temperature; (C) Data logger and soil moisture sensor.

In this study, three machine learning models were evaluated: MLR, SVR, and CNN. MLR fits all the observations to a line by finding the relationship between the independent and dependent variables. SVR is similar to MLR, except it seeks the best-fit hyperplane in the n-dimensional space of all independent variables, making it a nonlinear model as opposed to MLR. CNN is an artificial neural network algorithm inspired by the visual neuron system that obtains the required output by downscaling the image while keeping the image's important features. Figure 11 shows the architecture of two CNN models for comparing the inclusion or exclusion of soil moisture data. Figure 11A, for example, illustrates the architecture without soil moisture. The TIR, NIR, and RGB images captured by the cameras are scaled to 256*256 pixels and input into the model, where they are subjected to a convolution layer and a max-pooling layer, before being connected by a concatenate layer. The connected data is then sent through a convolution layer and a max-pooling layer, and after feature extraction and dimensionality reduction, the final output is acquired through the dense layer. Figure 11B shows the architecture with soil moisture, which is similar to Figure 11A with the addition of soil moisture as an additional input variable in the final dense layer.

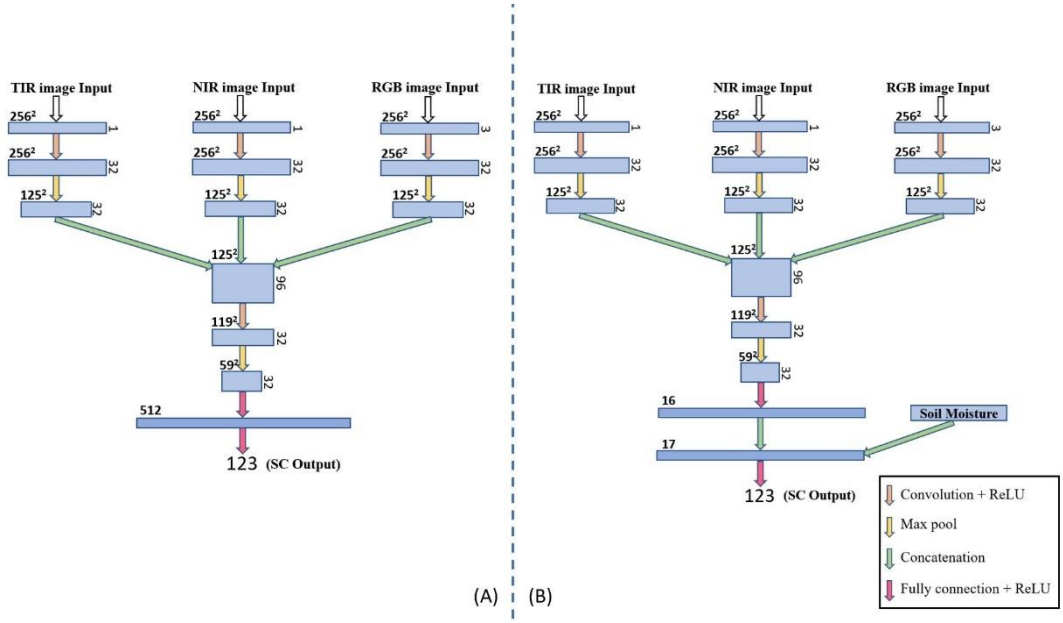

Figure 11: Architecture for CNN models. (A) Without soil moisture; (B) With soil moisture.

The whole data set is divided into training and test sets in the ratio of 80% and 20%. During the training process, the grid search method was used to perform hyperparameter tuning on the SVR, and the hyperparameters involved in the tuning were regularization parameter: C, and kernel coefficient: Gamma. In the study, the performance of these three models was evaluated for maize and soybean crops with and without soil moisture by using coefficient of determination (R^2) and root mean squared error (RMSE).

3.3 Results and Discussion

Figure 12 shows the distribution of maize and soybean at different measurement dates. As indicated in the figure, the mean g_s values of soybean are greater than those of maize on the same day of measurement. It is noteworthy that both maize and

soybean exhibit short-term seasonal variation in g_s, with soybean showing the most significant variation. It increased abruptly from a low of 0.4 mol/(m^2*s) on August 8th to 0.81 mol/(m^2*s) on August 10th and then decreased gradually to reach 0.27 mol/(m^2*s) on September 4th. The performance of Maize showed the same trend although it was not obvious. We also found that the g_s of both crops under irrigation was significantly higher than that of the non-irrigated crop during the same period, so we can conclude that the decrease in soil water potential leads to a decrease in g_s. In other words, under drought stress, stomata will close to reduce water loss by transpiration to increase the water use efficiency. This behavior could also cause irreversible wilting of leaves due to water deficiency (Blackman et al., 2009). With the closure of stomata, the developmental process will inevitably lead to a reduction in photosynthetic efficiency in the reproductive stage and failure to develop more biomass (Onyemaobi et al., 2021). The result from two other studies also showed that crops experienced the yield reductions due to the decrease in g_s caused by the decrease in water availability during water shortage conditions (Gleason et al., 2021; Liao et al., 2022).

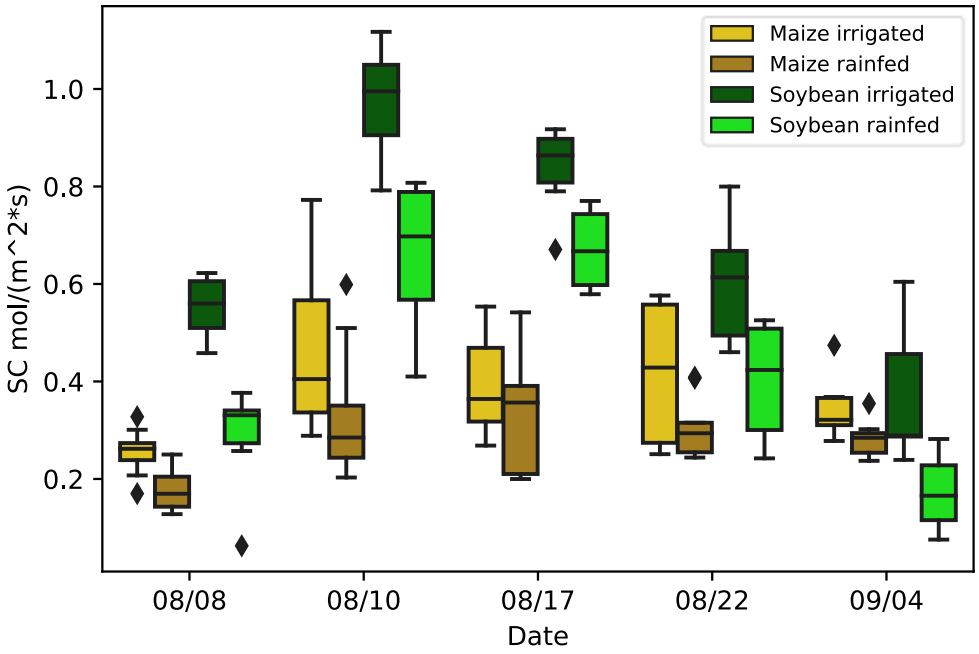

Figure 12: Boxplot for measured g_s

Table 7 gives the results of three models for both crops under conditions including and excluding soil moisture. All models for maize performed poorly, while SVR performed the best of all models with an R^2 value of 0.16 and an RMSE of 0.112 mol/(m^2*s). The accuracy of the SVR and CNN models for soybean was comparable, however the CNN model performed better when soil moisture was not considered, with a R^2 value of 0.40 and an RMSE of 0.137 mol/(m^2*s). The SVR model performed best with an R2 value of 0.62 and an RMSE of 0.116 mol/(m^2*s) when soil moisture was included. The CNN model is much superior to the other two models when both crops are used together for estimating purposes. The R^2 value was the same whether soil moisture was included or not, but the RMSE increased from 0.147 mol/(m^2*s) to 0.137 mol/(m^2*s) when soil moisture was included in the variables.

Table 7: The performance for estimating stomatal conductance with three different machine learning models.

	Mazie				**Soybean**				**Both Crops**			
	W/O moisture		**W/ moisture**		**W/O moisture**		**W/ moisture**		**W/O moisture**		**W/ moisture**	
	R^2	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
MLR	0.07	0.111	0.07	0.112	0.32	0.177	0.35	0.172	0.20	0.167	0.20	0.167
SVR	**0.13**	**0.114**	**0.16**	**0.112**	0.35	0.163	**0.62**	**0.116**	0.25	0.153	0.30	0.148
CNN	0.10	0.114	0.14	0.110	**0.40**	**0.177**	0.52	0.137	**0.52**	**0.147**	**0.52**	**0.137**

Figure 13, on the other hand, shows the performance results when estimating g_s for both crops at the same time. As seen in the graph, the MLR model performs the worst, the R^2 value is 0.20 and an RMSE of 0.167 mol/(m^2*s) for with or without the inclusion of soil moisture. Although the non-linear relationship between photosynthesis and g_s (Barnard & Bauerle, 2013; Lamour et al., 2022) makes the SVR model slightly better than MLR, it is still not comparable to the accuracy of the CNN model. This may be because the CNN model with image input can determine the species of the crop according to those images, whereas the MLR and SVR models cannot determine the species using the plant coverage rate, canopy temperature, and soil temperature. Nonetheless, because the water use efficiency and light response vary among species (Y. Zhang et al., 2022), it is not possible to reliably estimate g_s if the species is not included in the model. In the case of the C3 species soybean and C4 species maize used in this experiment, the transpiration rate and g_s

were higher in soybean than in maize due to their lower water use efficiency under the same weather conditions (Knapp, 1993; Ye et al., 2020). Therefore, it is essential to use an image input model similar to CNN or to add species information when using SVR models if a g_s estimation model with multiple crops is needed. Also, such as other additional factor that not included in the MLR and SVR models, Normalized difference vegetation index (NDVI), was positively correlated with g_s. Hence, CNN models utilizing images as input will always contain more predictive information than MLR and SVR models. This is a possible explanation for why the CNN model performs better than the other two models.

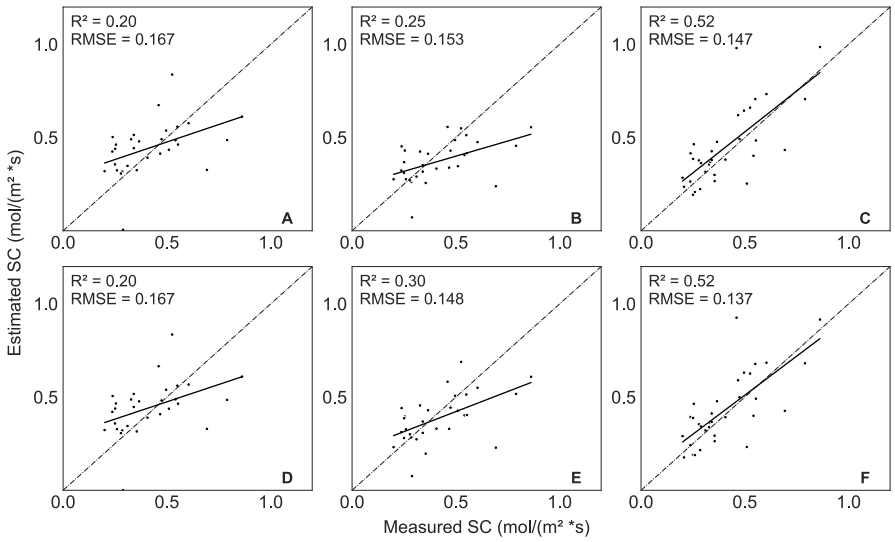

Figure 13: g_s estimated for both maize and soybean. (A) MLR without soil moisture. (B) SVR without soil moisture. (C) CNN without soil moisture. (D) MLR with soil moisture. (E) SVR with soil moisture. (F) CNN with soil moisture.

As mentioned previously for the differences in g_s caused by these two crop species under different irrigation conditions, are also apparent in the model estimation results. As shown in Table 1, estimation results that include soil moisture as a variable

are often more accurate than those that do not include soil moisture (e.g., for the CNN model with the two crop combinations, the RMSE is increased by 0.01 mol/(m^2*s)). Consequently, despite the fact that remote sensing images may currently provide relatively accurate estimates of g_s, other factors can be added to the model if further efficiency gains are required. In addition to the soil moisture already incorporated into the model, vapor pressure deficit (VPD) can also be considered for inclusion. In other words, even when the soil moisture is low, the crop may still undergo transpiration because of the low VPD (Kimm et al., 2020; Liao et al., 2022; J. Zhang et al., 2021). Therefore, we can further expand the size of the dataset in the future, while trying to see if adding VPD as a variable will also improve the efficiency of the model.

3.4 Conclusion

In this study, three machine learning approaches were employed to estimate g_s by merging data from RGB, NIR, and TIR cameras. The results indicate that adding soil moisture as a variable to the estimated model can enhance the model performance. The best performing model for both crops had an R^2 value of 0.52 and an RMSE of 0.137 mol/(m^2*s). Therefore, this study provides a possibility to quantify leaf-level g_s rapidly by using remote sensing images collected from platforms such as UAVs or satellites to help growers or researchers collect g_s information over large areas.

CHAPTER 4 OVERALL CONCLUSION AND FUTURE PROSPECT

With the growth of the global population and the decrease of non-renewable fossil fuels, the demand for the crops to support the food and energy needs increases accordingly. Thus, how to increase the crop yield has also become the focus of researchers and growers. One of the factors that affects crop yield is water availability. Therefore, how to ensure the healthy growth of crops has become a top priority in today's water scarce world. Two effective ways to overcome this challenge are (1) to develop drought-tolerant species, and (2) to improve the efficiency of water management in crop production. Both of these potential ways required the understanding of current water potential of the crops. Because of the principle of photosynthesis and transpiration, g_s represents the water and gas exchange between the crop and the atmosphere can be used to monitor the crop's current water status. The traditional contact-type sensing techniques such as IRGA or porometer requires long measurement time in the field for the data collection, which brings several difficulties in quantifying and understanding the water status of the crop.

Over the past few decades, rapid advancements have been archived in remote sensing and now it is widely used in agricultural research. Compared with the traditional data collection methods, the application of the remote sensing technology brings more possibilities for g_s collection. Along with the development of various sensor platforms such as satellites or UAVs, this method, which combines automation and data analysis, also offers a greater potential for high-throughput plant

phenotyping. This remote sensing-based approach for g_s estimation has several advantages: Large-scale and rapid data collection, and low labor involvement.

In both studies we conducted in this research, we tried to estimate g_s with machine learning models using different types of remote sensing data. For example, in the first study we used the weather information collected from an on-site weather station and plant phenotypic data gathered from the spectrometer, multispectral camera and thermal infrared camera. The result showed that g_s estimation was more accurate when combining the two types of data than when only using a single type of data. The random forest model had higher estimation accuracy than the other two machine learning models, for the five crop types investigated, maize, sorghum, soybean, sunflower, and winter wheat. The estimation result was able to reach about the R^2 of 0.69.

However, instruments such as spectrometers are difficult to install on UAVs and other mobile platforms because of their high price and non-portability. So, we decided to only use the sensors that are frequently used on those platforms. In the second study, we used only RGB, NIR, and thermal infrared images to estimate g_s. Also, in order to reduce the difficulties of transmitting real-time data like weather information, we used soil moisture to instead. The results showed that, after including the soil moisture, the estimation accuracy using these three kinds of images with the CNN model has a R^2 of 0.52.

From the results of these two studies we can infer the feasibility of using those remote sensing data to estimate g_s. Although there is still potential for the

improvement in estimation accuracy, the data from the high-throughput plant phenotyping platform demonstrated the success of this estimation approach. Compared with the empirical model proposed previously, we added various spectral indices in study one, including SIF, NDRE, and PRI, which have important relationships with g_s. We also included soil moisture as an additional variable in study two. Both of the studies showed that these crop and soil information were conducive to estimating g_s.

Admittedly, these two studies had some limitations that could be addressed in future studies For example, the amount of data used for model training is still relatively small. Another point is that although the second study is proposed based on the first one, there is no strong connection between them. The study two explored the estimation of g_s using images and soil moisture, but it ignored the influence of weather information. Therefore, the direction of the research in the future can be discussed in three points. First, expand the size of the dataset, which includes expanding the research to other crops and continuing to collect more data to train the model. Second, conduct variable selection in model training. Since each variable has different degrees of correlation with g_s, certain variables that have low correlation with g_s will be removed. Third, consider estimate g_s using timer series data that are continuously collected by the fixed sensor stations in the field, which can study the response of stomata under prolonged drought.

With the further development of various sensors and automation platforms, the estimation accuracy of g_s from the remote sensing data can be further improved.

Imagine if there is an integrated system that could collect weather, soil and crop information at the same time, it has a potential to effectively estimate the g_s. In this way, the water status of the crop is monitored and can be used to understand water use efficiency and the drought conditions of the crop. This information allows to accomplish the objectives mentioned previously, i.e., screening for the drought-tolerant species or improving the efficiency of irrigation in agriculture production.

REFERENCES

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2623–2631. https://doi.org/10.1145/3292500.3330701

Allen, R., Pereira, L., Raes, D., & Smith, M. (1998). FAO Irrigation and drainage paper No. 56. *Rome: Food and Agriculture Organization of the United Nations*, *56*, 26–40.

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*, *19*(1), 52–61. https://doi.org/10.1016/J.TPLANTS.2013.09.008

Askari, S. H., De-Ville, S., Hathway, E. A., & Stovin, V. (2021). Estimating evapotranspiration from commonly occurring urban plant species using porometry and canopy stomatal conductance. *Water*, *13*(16), 2262. https://doi.org/10.3390/W13162262

Atefi, A., Ge, Y., Pitla, S., & Schnable, J. (2021). Robotic technologies for high-throughput plant phenotyping: Contemporary reviews and future perspectives. *Frontiers in Plant Science, 12*, 1082. https://www.frontiersin.org/article/10.3389/fpls.2021.611940

Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances*, *3*(3), e1602244. https://doi.org/10.1126/sciadv.1602244

Bagley, J., Rosenthal, D. M., Ruiz-Vera, U. M., Siebers, M. H., Kumar, P., Ort, D. R., & Bernacchi, C. J. (2015). The influence of photosynthetic acclimation to rising CO2 and warmer temperatures on leaf and canopy photosynthesis models. *Global Biogeochemical Cycles*, *29*(2), 194–206. https://doi.org/10.1002/2014GB004848

Bai, G., Ge, Y., Hussain, W., Baenziger, P. S., & Graef, G. (2016). A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. *Computers and Electronics in Agriculture*, *128*, 181–192. https://doi.org/10.1016/j.compag.2016.08.021

Bai, G., Ge, Y., Scoby, D., Leavitt, B., Stoerger, V., Kirchgessner, N., Irmak, S., Graef, G., Schnable, J., & Awada, T. (2019). NU-Spidercam: A large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research. *Computers and Electronics in Agriculture*, *160*, 71–81. https://doi.org/10.1016/j.compag.2019.03.009

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. *Progress in Photosynthesis Research*, 221–224. https://doi.org/10.1007/978-94-017-0519-6_48

Ballester, C., Brinkhoff, J., Quayle, W. C., & Hornbuckle, J. (2019). Monitoring the effects of water stress in cotton using the green red vegetation index and red edge ratio. *Remote Sensing 2019*, *11*(7), 873. https://doi.org/10.3390/RS11070873

Baluja, J., Diago, M. P., Balda, P., Zorer, R., Meggio, F., Morales, F., & Tardaguila, J.

(2012). Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV). *Irrigation Science*, *30*(6), 511–522. https://doi.org/10.1007/s00271-012-0382-9

Banerjee, K., Krishnan, P., & Das, B. (2020). Thermal imaging and multivariate techniques for characterizing and screening wheat genotypes under water stress condition. *Ecological Indicators*, *119*, 106829. https://doi.org/10.1016/j.ecolind.2020.106829

Bao, Y., Gai, J., Xiang, L., Tang, L., Gai, J., Xiang, · L, Tang, · L, & Bao, Y. (2021). Field robotic systems for high-throughput plant phenotyping: A review and a case study. *High-Throughput Crop Phenotyping*. 13–38. https://doi.org/10.1007/978-3-030-73734-4_2

Barillot, R., Frak, E., Combes, D., Durand, J. L., & Escobar-Gutiérrez, A. J. (2010). What determines the complex kinetics of stomatal conductance under blueless PAR in Festuca arundinacea? Subsequent effects on leaf transpiration. *Journal of Experimental Botany*, *61*(10), 2795–2806. https://doi.org/10.1093/JXB/ERQ115

Barnard, D. M., & Bauerle, W. L. (2013). The implications of minimum stomatal conductance on modeling water flux in forest canopies. *Journal of Geophysical Research: Biogeosciences*, *118*(3), 1322–1333. https://doi.org/10.1002/JGRG.20112

Basu, P. S., Srivastava, M., Singh, P., Porwal, P., Singh, J., & Kant, R. (2015). High-precision phenotyping under controlled versus natural environments. *Phenomics in Crop Plants: Trends, Options and Limitations*, 27–40.

https://doi.org/10.1007/978-81-322-2226-2_3

Bell, C. J., & Squire, G. R. (1981). Comparative measurements with two water vapour diffusion porometers (dynamic and steady-State). *Journal of Experimental Botany*, *32*(6), 1143–1156. https://doi.org/10.1093/JXB/32.6.1143

Berni, J. A. J., Zarco-Tejada, P. J., Sepulcre-Cantó, G., Fereres, E., & Villalobos, F. (2009). Mapping canopy conductance and CWSI in olive orchards using high resolution thermal remote sensing imagery. *Remote Sensing of Environment*, *113*(11), 2380–2388. https://doi.org/10.1016/J.RSE.2009.06.018

Berni, J. A. J., Zarco-Tejada, P. J., Suarez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote wensing 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

Bertolino, L. T., Caine, R. S., & Gray, J. E. (2019). Impact of stomatal density and morphology on water-use efficiency in a changing world. *Frontiers in Plant Science*, *10*, 225. https://doi.org/10.3389/FPLS.2019.00225

Blackman, C. J., Brodribb, T. J., & Jordan, G. J. (2009). Leaf hydraulics and drought stress: response, recovery and survivorship in four woody temperate plant species. *Plant, Cell & Environment*, *32*(11), 1584–1595. https://doi.org/10.1111/J.1365-3040.2009.02023.X

Blonquist, J. M., Norman, J. M., & Bugbee, B. (2009). Automated measurement of canopy stomatal conductance based on infrared temperature. *Agricultural and Forest Meteorology*, *149*(11), 1931–1945.

https://doi.org/https://doi.org/10.1016/j.agrformet.2009.06.021

Boulesteix, A.-L., Janitza, S., Hapfelmeier, A., Van Steen, K., & Strobl, C. (2015). Letter to the Editor: On the term ‘interaction’ and related phrases in the literature on Random Forests. *Briefings in Bioinformatics*, *16*(2), 338–345. https://doi.org/10.1093/bib/bbu012

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brewer, K., Clulow, A., Sibanda, M., Gokool, S., Odindi, J., Mutanga, O., Naiken, V., Chimonyo, V. G. P., & Mabhaudhi, T. (2022). Estimation of maize foliar temperature and stomatal conductance as indicators of water stress based on optical and thermal imagery acquired using an unmanned aerial vehicle (UAV) Platform. *Drones 2022*, *6*(7), 169. https://doi.org/10.3390/DRONES6070169

Buckley, T. N. (2017). Modeling stomatal conductance. *Plant Physiology*, *174*(2), 572–582. https://doi.org/10.1104/pp.16.01772

Buckley, T. N., & Mott, K. A. (2013). Modelling stomatal conductance in response to environmental factors. *Plant, Cell & Environment*, *36*(9), 1691–1699. https://doi.org/https://doi.org/10.1111/pce.12140

Buckley, T. N., Sack, L., & Farquhar, G. D. (2017). Optimal plant water economy. *Plant, Cell & Environment*, *40*(6), 881–896. https://doi.org/10.1111/PCE.12823

Busch, F. A. (2018). Photosynthetic gas exchange in land plants at the leaf level. *Methods in Molecular Biology*, *1770*, 25–44. https://doi.org/10.1007/978-1-4939-7786-4_2

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., & Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, *274*, 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010

Campbell, B. M., Thornton, P., Zougmoré, R., van Asten, P., & Lipper, L. (2014). Sustainable intensification: What is its role in climate smart agriculture? *Current Opinion in Environmental Sustainability*, *8*, 39–43. https://doi.org/10.1016/J.COSUST.2014.07.002

Casson, S., & Gray, J. E. (2008). Influence of environmental factors on stomatal development. *New Phytologist*, *178*(1), 9–23. https://doi.org/10.1111/J.1469-8137.2007.02351.X

Čermák, J., Kučera, J., & Nadezhdina, N. (2004). Sap flow measurements with some thermodynamic methods, flow integration within trees and scaling up from sample trees to entire forest stands. *Trees - Structure and Function*, *18*(5), 529–546. https://doi.org/10.1007/S00468-004-0339-6

Cernusak, L. A. (2020). Gas exchange and water-use efficiency in plant canopies. *Plant Biology*, *22*(S1), 52–67. https://doi.org/10.1111/PLB.12939

Chaerle, L., Saibo, N., & van der Straeten, D. (2005). Tuning the pores: towards engineering plants for improved water use efficiency. *Trends in Biotechnology*, *23*(6), 308–315. https://doi.org/10.1016/J.TIBTECH.2005.04.005

Chamara, N., Islam, M. D., Bai, G. (Frank), Shi, Y., & Ge, Y. (2022). Ag-IoT for crop

and environment monitoring: Past, present, and future. *Agricultural Systems*, *203*, 103497. https://doi.org/10.1016/J.AGSY.2022.103497

Chaves, M. M., Costa, J. M., Zarrouk, O., Pinheiro, C., Lopes, C. M., & Pereira, J. S. (2016). Controlling stomatal aperture in semi-arid regions—The dilemma of saving water or being cool? *Plant Science*, *251*, 54–64. https://doi.org/10.1016/J.PLANTSCI.2016.06.015

Chieppa, J., Brown, T., Giresi, P., Juenger, T. E., Resco de Dios, V., Tissue, D. T., & Aspinwall, M. J. (2021). Climate and stomatal traits drive covariation in nighttime stomatal conductance and daytime gas exchange rates in a widespread C4 grass. *New Phytologist*, *229*(4), 2020–2034. https://doi.org/10.1111/NPH.16987

Colaizzi, P. D., O'Shaughnessy, S. A., Evett, S. R., & Mounce, R. B. (2017). Crop evapotranspiration calculation using infrared thermometers aboard center pivots. *Agricultural Water Management*, *187*, 173–189. https://doi.org/https://doi.org/10.1016/j.agwat.2017.03.016

Dai, Y., Dickinson, R. E., & Wang, Y.-P. (2004). A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, *17*(12), 2281–2299. https://doi.org/10.1175/1520-0442(2004)017<2281:ATMFCT>2.0.CO;2

Damour, G., Simonneau, T., Cochard, H., & Urban, L. (2010). An overview of models of stomatal conductance at the leaf level. *Plant, Cell & Environment*, *33*(9), 1419–1438. https://doi.org/10.1111/J.1365-3040.2010.02181.X

Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., & Cho, M. (2008). LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(4), 409–426. https://doi.org/10.1016/J.ISPRSJPRS.2008.01.001

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., James, R. A., Condon, A. G., Bovill, W. D., Hutchinson, P., Scarrow, J., Davy, R., & Furbank, R. T. (2016). Methodology for high-throughput field phenotyping of canopy temperature using airborne thermography. *Frontiers in Plant Science*, *7*(1808), https://doi.org/10.3389/FPLS.2016.01808

Doheny-Adams, T., Hunt, L., Franks, P. J., Beerling, D. J., & Gray, J. E. (2012). Genetic manipulation of stomatal density influences stomatal size, plant growth and tolerance to restricted water supply across a growth carbon dioxide gradient. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1588), 547–555. https://doi.org/10.1098/rstb.2011.0272

Dumont, J., Cohen, D., Gérard, J., Jolivet, Y., Dizengremel, P., & le Thiec, D. (2014). Distinct responses to ozone of abaxial and adaxial stomata in three Euramerican poplar genotypes. *Plant, Cell and Environment*, *37*(9), 2064–2076. https://doi.org/10.1111/PCE.12293

Ellsäßer, F., Röll, A., Ahongshangbam, J., Waite, P. A., Hendrayanto, Schuldt, B., & Hölscher, D. (2020). Predicting tree sap flux and stomatal conductance from drone-recorded surface temperatures in a mixed agroforestry system—A machine learning approach. *Remote Sensing 2020*, *12*(24), 4070.

https://doi.org/10.3390/RS12244070

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: trends and R&D implications. *Food Security 2022*, *14*(5), 1295–1319. https://doi.org/10.1007/S12571-022-01288-7

Espinoza, C. Z., Khot, L. R., Sankaran, S., & Jacoby, P. W. (2017). High resolution multispectral and thermal remote sensing-based water stress assessment in subsurface irrigated grapevines. *Remote Sensing 2017, 9*(9), 961. https://doi.org/10.3390/RS9090961

Fanourakis, D., Bouranis, D., Giday, H., Carvalho, D. R. A., Rezaei Nejad, A., & Ottosen, C. O. (2016). Improving stomatal functioning at elevated growth air humidity: A review. *Journal of Plant Physiology*, *207*, 51–60. https://doi.org/10.1016/J.JPLPH.2016.10.003

Fayezizadeh, M. R., Ansari, N. A. Z., Albaji, M., & Khaleghi, E. (2021). Effects of hydroponic systems on yield, water productivity and stomatal gas exchange of greenhouse tomato cultivars. *Agricultural Water Management*, *258*, 107171. https://doi.org/10.1016/J.AGWAT.2021.107171

Franks, P. J., W. Doheny-Adams, T., Britton-Harper, Z. J., & Gray, J. E. (2015). Increasing water-use efficiency directly through genetic manipulation of stomatal density. *New Phytologist*, *207*(1), 188–195. https://doi.org/10.1111/nph.13347

Fu, Y., Yang, G., Wang, J., Song, X., & Feng, H. (2014). Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements. *Computers and Electronics in*

Agriculture, *100*, 51–59. https://doi.org/10.1016/J.COMPAG.2013.10.010

Fukase, E., & Martin, W. (2020). Economic growth, convergence, and world food demand and supply. *World Development*, *132*, 104954. https://doi.org/10.1016/J.WORLDDEV.2020.104954

Gaastra, P. (1959). *Photosynthesis of crop plants as influenced by light, carbon dioxide, temperature, and stomatal diffusion resistance*. [internal PhD, WU, Wageningen University]. Veenman. https://edepot.wur.nl/183268

Ge, Y., Atefi, A., Zhang, H., Miao, C., Ramamurthy, R. K., Sigmon, B., Yang, J., & Schnable, J. C. (2019). High-throughput analysis of leaf physiological and chemical traits with VIS-NIR-SWIR spectroscopy: A case study with a maize diversity panel. *Plant Methods*, *15*(1), 1–12. https://doi.org/10.1186/S13007-019-0450-8

Ge, Y., Bai, G., Stoerger, V., & Schnable, J. C. (2016). Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Computers and Electronics in Agriculture*, *127*, 625–632. https://doi.org/10.1016/j.compag.2016.07.028

Geng, J., Gan, W., Xu, J., Yang, R., & Wang, S. (2020). Support vector machine regression (SVR)-based nonlinear modeling of radiometric transforming relation for the coarse-resolution data-referenced relative radiometric normalization (RRN). *Geo-spatial Information Science,* 23:3, 237-247 https://doi.org/10.1080/10095020.2020.1785958

Ghimire, C. P., Bruijnzeel, L. A., Lubczynski, M. W., Zwartendijk, B. W., Odongo, V.

O., Ravelona, M., & van Meerveld, H. J. I. (2018). Transpiration and stomatal conductance in a young secondary tropical montane forest: contrasts between native trees and invasive understorey shrubs. *Tree Physiology*, *38*(7), 1053–1070. https://doi.org/10.1093/TREEPHYS/TPY004

Gindel, I. (1969). Stomatal number and size as related to soil moisture in tree Xerophytes in Israel. *Ecology*, *50*(2), 263–267. https://doi.org/10.2307/1934854

Gleason, S. M., Nalezny, L., Hunter, C., Bensen, R., Chintamanani, S., & Comas, L. H. (2021). Growth and grain yield of eight maize hybrids are aligned with water transport, stomatal conductance, and photosynthesis in a semi-arid irrigated system. *Physiologia Plantarum*, *172*(4), 1941–1949. https://doi.org/10.1111/PPL.13400

Gowdy, M., Pieri, P., Suter, B., Marguerit, E., Destrac-Irvine, A., Gambetta, G., & van Leeuwen, C. (2022). Estimating bulk stomatal conductance in grapevine canopies. *Frontiers in Plant Science*, *13*.839378 https://doi.org/10.3389/FPLS.2022.839378

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3), 337–352. https://doi.org/10.1016/j.rse.2003.12.013

He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

https://doi.org/10.1109/TKDE.2008.239

Herwitz, S. R., Johnson, L. F., Dunagan, S. E., Higgins, R. G., Sullivan, D. V, Zheng, J., Lobitz, B. M., Leung, J. G., Gallmeyer, B. A., Aoyagi, M., Slye, R. E., & Brass, J. A. (2004). Imaging from an unmanned aerial vehicle: agricultural surveillance and decision support. *Computers and Electronics in Agriculture*, *44*(1), 49–61. https://doi.org/10.1016/j.compag.2004.02.006

Hetherington, A. M., & Woodward, F. I. (2003). The role of stomata in sensing and driving environmental change. *Nature 2003*, *424*(6951), 901–908. https://doi.org/10.1038/nature01843

Houshmandfar, A., O'Leary, G., Fitzgerald, G. J., Chen, Y., Tausz-Posch, S., Benke, K., Uddin, S., & Tausz, M. (2021). Machine learning produces higher prediction accuracy than the Jarvis-type model of climatic control on stomatal conductance in a dryland wheat agro-ecosystem. *Agricultural and Forest Meteorology*, *304–305*, 108423. https://doi.org/10.1016/j.agrformet.2021.108423

Huang, G., Yang, Y., Zhu, L., Peng, S., & Li, Y. (2021). Temperature responses of photosynthesis and stomatal conductance in rice and wheat plants. *Agricultural and Forest Meteorology*, *300*, 108322. https://doi.org/10.1016/J.AGRFORMET.2021.108322

Ihuoma, S. O., & Madramootoo, C. A. (2019). Sensitivity of spectral vegetation indices for monitoring water stress in tomato plants. *Computers and Electronics in Agriculture*, *163*, 104860. https://doi.org/10.1016/J.COMPAG.2019.104860

Ishimwe, R., Abutaleb, K., & Ahmed, F. (2014). Applications of thermal imaging in

agriculture—A review. *Advances in Remote Sensing*, *3*(3), 128–140. https://doi.org/10.4236/ARS.2014.33011

Sadler, E. J., Camp, C. R., Evans, D. E., & Millen, J. A. (2002). Corn canopy temperatures measured with a moving infrared thermometer array. *Transactions of the ASAE*, *45*(3), 581-591. https://doi.org/10.13031/2013.8855

Jagadish, S. V. K., Way, D. A., & Sharkey, T. D. (2021). Plant heat stress: Concepts directing future research. *Plant, Cell & Environment*, *44*(7), 1992–2005. https://doi.org/10.1111/pce.14050

Jarolmasjed, S., Sankaran, S., Kalcsits, L., & Khot, L. R. (2018). Proximal hyperspectral sensing of stomatal conductance to monitor the efficacy of exogenous abscisic acid applications in apple trees. *Crop Protection*, *109*, 42–50. https://doi.org/10.1016/J.CROPRO.2018.02.022

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *273*(927), 593–610. https://doi.org/10.1098/RSTB.1976.0035

Jones, H. G. (1999a). Use of infrared thermometry for estimation of stomatal conductance as a possible aid to irrigation scheduling. *Agricultural and Forest Meteorology*, *95*(3), 139–149. https://doi.org/10.1016/S0168-1923(99)00030-1

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

Jones, H. G., Hamer, P. J. C., & Higgs, K. H. (1988). Evaluation of various heat-pulse methods for estimation of sap flow in orchard trees: comparison with micrometeorological estimates of evaporation. *Trees*, *2*(4), 250–260. https://doi.org/10.1007/BF00202380

Khan, Z., Rahimi-Eichi, V., Haefele, S., Garnett, T., & Miklavcic, S. J. (2018). Estimation of vegetation indices for high-throughput phenotyping of wheat using aerial imaging. *Plant Methods*, *14*(1), 1–11. https://doi.org/10.1186/S13007-018-0287-6

Khanal, S., Kushal, K. C., Fulton, J. P., Shearer, S., & Ozkan, E. (2020). Remote sensing in agriculture—accomplishments, limitations, and opportunities. *Remote Sensing 2020, 12*(22), 3783. https://doi.org/10.3390/RS12223783

Khosla, E., Dharavath, R., & Priya, R. (2020). Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. *Environment, Development and Sustainability*, *22*(6), 5687–5708. https://doi.org/10.1007/S10668-019-00445-X

Kim, Y., Glenn, D. M., Park, J., Ngugi, H. K., & Lehman, B. L. (2011). Hyperspectral image analysis for water stress detection of apple trees. *Computers and Electronics in Agriculture*, *77*(2), 155–160. https://doi.org/10.1016/j.compag.2011.04.008

Kimm, H., Guan, K., Gentine, P., Wu, J., Bernacchi, C. J., Sulman, B. N., Griffis, T. J., & Lin, C. (2020). Redefining droughts for the U.S. corn belt: The dominant role of atmospheric vapor pressure deficit over soil moisture in regulating

stomatal behavior of Maize and Soybean. *Agricultural and Forest Meteorology*, *287*, 107930. https://doi.org/10.1016/J.AGRFORMET.2020.107930

Knapp, A. K. (1993). Gas exchange dynamics in C3 and C4 grasses: consequences of differences in stomatal conductance. *Ecology*, *74*(1), 113–123. https://doi.org/10.2307/1939506

Kollist, H., Nuhkat, M., & Roelfsema, M. R. G. (2014). Closing gaps: linking elements that control stomatal movement. *New Phytologist*, *203*(1), 44–62. https://doi.org/10.1111/NPH.12832

Kučera, J., Brito, P., Jiménez, M. S., & Urban, J. (2017). Direct Penman–Monteith parameterization for estimating stomatal conductance and modeling sap flow. *Trees - Structure and Function*, *31*(3), 873–885. https://doi.org/10.1007/S00468-016-1513-3

Lamour, J., Davidson, K. J., Ely, K. S., Le Moguédec, G., Leakey, A. D. B., Li, Q., Serbin, S. P., & Rogers, A. (2022). An improved representation of the relationship between photosynthesis and stomatal conductance leads to more stable estimation of conductance parameters and improves the goodness-of-fit across diverse data sets. *Global Change Biology*, *28*(11), 3537–3556. https://doi.org/10.1111/GCB.16103

Lavoie-Lamoureux, A., Sacco, D., Risse, P. A., & Lovisolo, C. (2017). Factors influencing stomatal conductance in response to water availability in grapevine: a meta-analysis. *Physiologia Plantarum*, *159*(4), 468–482. https://doi.org/10.1111/PPL.12530

Lawson, T. (2009). Guard cell photosynthesis and stomatal function. *New Phytologist*, *181*(1), 13–34. https://doi.org/10.1111/j.1469-8137.2008.02685.x

Li, H., Zhang, J., Zhang, S., & Bai, Y. (2022). Machine learning and remote sensing-based modeling of the optimal stomatal behavior of crops. *Computers and Electronics in Agriculture*, *200*, 107261. https://doi.org/10.1016/J.COMPAG.2022.107261

Li, J., Schachtman, D. P., Creech, C. F., Wang, L., Ge, Y., & Shi, Y. (2022). Evaluation of UAV-derived multimodal remote sensing data for biomass prediction and drought tolerance assessment in bioenergy sorghum. *The Crop Journal*, *10*(5), 1363-1375. https://doi.org/10.1016/j.cj.2022.04.005

Li, L., Zhang, Q., & Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors 2014, 14*(11), 20078–20111. https://doi.org/10.3390/S141120078

Li, Y., Li, H., Li, Y., & Zhang, S. (2017). Improving water-use efficiency by decreasing stomatal conductance and transpiration rate to maintain higher ear photosynthetic rate in drought-resistant wheat. *The Crop Journal*, *5*(3), 231–239. https://doi.org/10.1016/J.CJ.2017.01.001

Liang, L., Di, L., Huang, T., Wang, J., Lin, L., Wang, L., & Yang, M. (2018). Estimation of leaf nitrogen content in wheat using new hyperspectral indices and a random forest regression Algorithm. *Remote Sensing 2018, 10*(12), 1940. https://doi.org/10.3390/RS10121940

Liao, Q., Ding, R., Du, T., Kang, S., Tong, L., & Li, S. (2022). Stomatal conductance

drives variations of yield and water use of maize under water and nitrogen stress. *Agricultural Water Management*, *268*, 107651. https://doi.org/10.1016/J.AGWAT.2022.107651

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., … Torquebiau, E. F. (2014). Climate-smart agriculture for food security. *Nature Climate Change 2014 4:12*, *4*(12), 1068–1072. https://doi.org/10.1038/nclimate2437

Long, S. P., & Bernacchi, C. J. (2003). Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *Journal of Experimental Botany*, *54*(392), 2393–2401. https://doi.org/10.1093/JXB/ERG262

Lopes, M. S., Araus, J. L., van Heerden, P. D. R., & Foyer, C. H. (2011). Enhancing drought tolerance in C4 crops. *Journal of Experimental Botany*, *62*(9), 3135–3153. https://doi.org/10.1093/jxb/err105

Maes, W. H., & Steppe, K. (2012). Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. *Journal of Experimental Botany*, *63*(13), 4671–4712. https://doi.org/10.1093/jxb/ers165

Matese, A., Baraldi, R., Berton, A., Cesaraccio, C., Di Gennaro, S. F., Duce, P., Facini, O., Mameli, M. G., Piga, A., & Zaldei, A. (2018). Estimation of water stress in grapevines using proximal and remote sensing methods. *Remote Sensing*

2018, 10(1), 114. https://doi.org/10.3390/RS10010114

Mathe-Gaspar, G., Fodor, N., Pokovai, K., & Kovacs, G. J. (2005). Crop modelling as a tool to separate the influence of the soil and weather on crop yields. *Physics and Chemistry of the Earth, Parts A/B/C*, *30*(1–3), 165–169. https://doi.org/10.1016/J.PCE.2004.08.024

McDermitt, D. K. (1990). Sources of error in the estimation of stomatal conductance and transpiration from porometer data. *HortScience, 25*(12), 1538–1548. https://doi.org/10.21273/HORTSCI.25.12.1538

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62–69. https://doi.org/10.1016/J.CHEMOLAB.2012.07.010

Minaei, S., Soltanikazemi, M., Shafizadeh-Moghadam, H., & Mahdavian, A. (2022). Field-scale estimation of sugarcane leaf nitrogen content using vegetation indices and spectral bands of Sentinel-2: Application of random forest and support vector regression. *Computers and Electronics in Agriculture*, *200*, 107130. https://doi.org/10.1016/J.COMPAG.2022.107130

Mishra, V., & Cherkauer, K. A. (2010). Retrospective droughts in the crop growing season: Implications to corn and soybean yield in the Midwestern United States. *Agricultural and Forest Meteorology*, *150*(7–8), 1030–1045. https://doi.org/10.1016/J.AGRFORMET.2010.04.002

Motohka, T., Nasahara, K. N., Oguma, H., & Tsuchida, S. (2010). Applicability of

green-red vegetation index for remote sensing of vegetation phenology. *Remote Sensing 2010, 2*(10), 2369–2387. https://doi.org/10.3390/RS2102369

Motzer, T., Munz, N., Küppers, M., Schmitt, D., & Anhuf, D. (2005). Stomatal conductance, transpiration and sap flow of tropical montane rain forest trees in the southern Ecuadorian Andes. *Tree Physiology*, *25*(10), 1283–1293. https://doi.org/10.1093/TREEPHYS/25.10.1283

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, *114*(4), 358–371. https://doi.org/10.1016/J.BIOSYSTEMSENG.2012.08.009

Omer, G., Mutanga, O., Abdel-Rahman, E. M., Adam, E., Waser, L. T., Roy, S., & Thenkabail, P. S. (2016). Empirical prediction of leaf area index (LAI) of endangered tree species in intact and fragmented indigenous forests ecosystems using WorldView-2 Data and two robust machine learning algorithms. *Remote Sensing 2016, 8*(4), 324. https://doi.org/10.3390/RS8040324

Onyemaobi, O., Sangma, H., Garg, G., Wallace, X., Kleven, S., Suwanchaikasem, P., Roessner, U., & Dolferus, R. (2021). Reproductive stage drought tolerance in wheat: Importance of stomatal conductance and plant growth regulators. *Genes*, *12*(11), 1742. https://doi.org/10.3390/GENES12111742/S1

Otu-Larbi, F., Conte, A., Fares, S., Wild, O., & Ashworth, K. (2021). FORCAsT-gs: Importance of stomatal conductance parameterization to estimated ozone deposition velocity. *Journal of Advances in Modeling Earth Systems*, *13*(9), e2021MS002581. https://doi.org/https://doi.org/10.1029/2021MS002581

Pagano, M. C., & Miransari, M. (2016). The importance of soybean production worldwide. *Abiotic and Biotic Stresses in Soybean Production: Soybean Production: Volume 1*, *5*, 1–26. https://doi.org/10.1016/B978-0-12-801536-0.00001-3

Panda, S., Amatya, D. M., & Hoogenboom, G. (2014). Stomatal conductance, canopy temperature, and leaf area index estimation using remote sensing and OBIA techniques. *Journal of Spatial Hydrology*, *12*(1): 24 p. https://www.fs.usda.gov/research/treesearch/49339

Pandey, P., Ge, Y., Stoerger, V., & Schnable, J. C. (2017). High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. *Frontiers in Plant Science*, *8,* 1348. https://www.frontiersin.org/articles/10.3389/fpls.2017.01348

Pathan, M., Patel, N., Yagnik, H., & Shah, M. (2020). Artificial cognition for applications in smart agriculture: A comprehensive review. *Artificial Intelligence in Agriculture*, *4*, 81–95. https://doi.org/10.1016/J.AIIA.2020.06.001

Prasad, N. R., Patel, N. R., & Danodia, A. (2021). Crop yield prediction in cotton for regional level using random forest approach. *Spatial Information Research*, *29*(2), 195–206. https://doi.org/10.1007/S41324-020-00346-6

Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing*, *109*, 88–97.

https://doi.org/10.1016/J.ISPRSJPRS.2015.09.003

Rasooli, S. V., Soltani Nazarloo, A., Taghinezahd, E., Veza, I., Szumny, A., & Figiel, A. (2022). Prediction of winter wheat leaf chlorophyll content based on VIS/NIR spectroscopy using ANN and PLSR. *Food Science & Nutrition*. 2022, 1 https://doi.org/10.1002/FSN3.3071

Rebetzke, G. J., Rattey, A. R., Farquhar, G. D., Richards, R. A., & Condon, A. (Tony) G. (2012). Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. *Functional Plant Biology*, *40*(1), 14–33. https://doi.org/10.1071/FP12184

Richardson, F., Brodribb, T. J., & Jordan, G. J. (2017). Amphistomatic leaf surfaces independently regulate gas exchange in response to variations in evaporative demand. *Tree Physiology*, *37*(7), 869–878. https://doi.org/10.1093/TREEPHYS/TPX073

Roelfsema, M. R. G., & Hedrich, R. (2005). In the light of stomatal opening: new insights into 'the Watergate.' *New Phytologist*, *167*(3), 665–691. https://doi.org/10.1111/J.1469-8137.2005.01460.X

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*(2), 95–107. https://doi.org/https://doi.org/10.1016/0034-4257(95)00186-7

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., Reynolds, M., & Singh, R. (2016). Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for

grain yield in wheat. *G3: Genes, Genomes, Genetics*, *6*(9), 2799–2808. https://doi.org/10.1534/G3.116.032888/-/DC1

Saunders, A., Drew, D. M., & Brink, W. (2021). Machine learning models perform better than traditional empirical models for stomatal conductance when applied to multiple tree species across different forest biomes. *Trees, Forests and People*, *6*, 100139. https://doi.org/10.1016/j.tfp.2021.100139

Shakoor, N., Lee, S., & Mockler, T. C. (2017). High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Current Opinion in Plant Biology*, *38*, 184–192. https://doi.org/10.1016/J.PBI.2017.05.006

Sharifi, A. (2021). Yield prediction with machine learning algorithms and satellite images. *Journal of the Science of Food and Agriculture*, *101*(3), 891–896. https://doi.org/10.1002/JSFA.10696

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing 2020, 12*(19), 3136. https://doi.org/10.3390/RS12193136

Sobejano-Paz, V., Mikkelsen, T. N., Baum, A., Mo, X., Liu, S., Köppl, C. J., Johnson, M. S., Gulyas, L., & García, M. (2020). Hyperspectral and thermal sensing of stomatal conductance, transpiration, and photosynthesis for soybean and maize under drought. *Remote Sensing 2020, 12*(19), 3182. https://doi.org/10.3390/RS12193182

Struthers, R., Ivanova, A., Tits, L., Swennen, R., & Coppin, P. (2015). Thermal infrared imaging of the temporal variability in stomatal conductance for fruit

trees. *International Journal of Applied Earth Observation and Geoinformation*, *39*, 9–17. https://doi.org/10.1016/J.JAG.2015.02.006

Su, Y., Shao, W., Vlček, L., & Langhammer, J. (2019). Ecohydrological behaviour of mountain beech forest: Quantification of stomatal conductance using sap flow measurements. *Geosciences 2019, 9*(5), 243. https://doi.org/10.3390/GEOSCIENCES9050243

Sugiura, R., Noguchi, N., & Ishii, K. (2005). Remote-sensing technology for vegetation monitoring using an unmanned helicopter. *Biosystems Engineering*, *90*(4), 369–379. https://doi.org/10.1016/j.biosystemseng.2004.12.011

Takai, T., Yano, M., & Yamamoto, T. (2010). Canopy temperature on clear and cloudy days can be used to estimate varietal differences in stomatal conductance in rice. *Field Crops Research*, *115*(2), 165–170. https://doi.org/10.1016/j.fcr.2009.10.019

Tan, K., Ma, W., Wu, F., & Du, Q. (2019). Random forest–based estimation of heavy metal concentration in agricultural soils with hyperspectral sensor data. *Environmental Monitoring and Assessment*, *191*(7), 1–14. https://doi.org/10.1007/S10661-019-7510-4/FIGURES/6

Thorp, K. R., Thompson, A. L., Harders, S. J., French, A. N., & Ward, R. W. (2018). High-throughput phenotyping of crop water use efficiency via multispectral drone imagery and a daily soil water balance model. *Remote Sensing* 10(11), 1682 https://doi.org/10.3390/rs10111682

Thruppoyil, S. B., & Ksiksi, T. (2020). Time-dependent stomatal conductance and

growth responses of tabernaemontana divaricata to short-term elevated CO2 and water stress at higher than optimal growing temperature. *Current Plant Biology*, *22*, 100127. https://doi.org/10.1016/j.cpb.2019.100127

Tian, Y., Xu, Y. P., & Wang, G. (2018). Agricultural drought prediction using climate indices based on Support Vector Regression in Xiangjiang River basin. *Science of The Total Environment*, *622–623*, 710–720. https://doi.org/10.1016/J.SCITOTENV.2017.12.025

Toro, G., Flexas, J., & Escalona, J. M. (2019). Contrasting leaf porometer and infra-red gas analyser methodologies: an old paradigm about the stomatal conductance measurement. *Theoretical and Experimental Plant Physiology*, *31*(4), 483–492. https://doi.org/10.1007/S40626-019-00161-X

Vadivambal, R., & Jayas, D. S. (2011). Applications of thermal imaging in agriculture and food industry-A review. *Food and Bioprocess Technology*, *4*(2), 186–199. https://doi.org/10.1007/S11947-010-0333-5

van Dijk, A. D. J., Kootstra, G., Kruijer, W., & de Ridder, D. (2021). Machine learning in plant science and plant breeding. *IScience*, *24*(1), 101890. https://doi.org/https://doi.org/10.1016/j.isci.2020.101890

Vandegehuchte, M. W., & Steppe, K. (2012). Interpreting the Heat Field Deformation method: Erroneous use of thermal diffusivity and improved correlation between temperature ratio and sap flux density. *Agricultural and Forest Meteorology*, *162–163*, 91–97. https://doi.org/10.1016/J.AGRFORMET.2012.04.013

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., Hawkesford, M. J., (2016). Field

Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. *Functional Plant Biology*, *44*(1), 143–153. https://doi.org/10.1071/FP16163

Virnodkar, S. S., Pachghare, V. K., Patil, V. C., & Jha, S. K. (2020). Remote sensing and machine learning for crop water stress determination in various crops: a critical review. *Precision Agriculture*, *21*(5), 1121–1155. https://doi.org/10.1007/S11119-020-09711-9

Vitrack-Tamam, S., Holtzman, L., Dagan, R., Levi, S., Tadmor, Y., Azizi, T., Rabinovitz, O., Naor, A., & Liran, O. (2020). Random forest algorithm improves detection of physiological activity embedded within reflectance spectra using stomatal conductance as a test case. *Remote Sensing 2020, 12*(14), 2213. https://doi.org/10.3390/RS12142213

Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal*, *4*(3), 212–219. https://doi.org/10.1016/J.CJ.2016.01.008

Wehner, M., Easterling, D. R., Lawrimore, J. H., Heim, R. R., Vose, R. S., & Santer, B. D. (2011). Projections of future drought in the continental United States and Mexico. *Journal of Hydrometeorology*, *12*(6), 1359–1377. https://doi.org/10.1175/2011JHM1351.1

Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, *236*, 111402. https://doi.org/10.1016/J.RSE.2019.111402

Wijewardane, N. K., Ge, Y., Wills, S., & Libohova, Z. (2018). Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library. *Soil Science Society of America Journal*, *82*(3), 722–731. https://doi.org/10.2136/sssaj2017.10.0361

Wijewardane, N. K., Ge, Y., Wills, S., & Loecke, T. (2016). Prediction of soil carbon in the conterminous United States: visible and near infrared reflectance spectroscopy analysis of the rapid carbon assessment project. *Soil Science Society of America Journal*, *80*(4), 973–982. https://doi.org/10.2136/sssaj2016.02.0052

Wijewardane, N. K., Zhang, H., Yang, J., Schnable, J. C., Schachtman, D. P., & Ge, Y. (2023). A leaf-level spectral library to support high throughput plant phenotyping: Predictive accuracy and model transfer. *Journal of Experimental Botany*, erad129. https://doi.org/10.1093/jxb/erad129

Xu, J., Wu, B., Ryu, D., Yan, N., Zhu, W., & Ma, Z. (2021). A canopy conductance model with temporal physiological and environmental factors. *Science of The Total Environment*, *791*, 148283. https://doi.org/10.1016/J.SCITOTENV.2021.148283

Xu, R., Li, C., & Paterson, A. H. (2019). Multispectral imaging and unmanned aerial systems for cotton plant phenotyping. *PLOS ONE*, *14*(2), e0205083. https://doi.org/10.1371/JOURNAL.PONE.0205083

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Z., Li, H., & Yang, H. (2017).

Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science*, *8*, 1111. https://doi.org/10.3389/FPLS.2017.01111/BIBTEX

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., & Yan, J. (2020). Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Molecular Plant*, *13*(2), 187–214. https://doi.org/10.1016/J.MOLP.2020.01.008

Ye, Z. P., Ling, Y., Yu, Q., Duan, H. L., Kang, H. J., Huang, G. M., Duan, S. H., Chen, X. M., Liu, Y. G., & Zhou, S. X. (2020). Quantifying light response of leaf-scale water-use efficiency and its interrelationships with photosynthesis and stomatal conductance in C3 and C4 species. *Frontiers in Plant Science*, *11*, 374. https://doi.org/10.3389/FPLS.2020.00374

Yu, L., Wang, W., Zhang, X., & Zheng, W. (2016). Review on leaf temperature sensor: Measurement methods and application. *IFIP Advances in Information and Communication Technology*, *478*, 216–230. https://doi.org/10.1007/978-3-319-48357-3_21

Zhang, C., Marzougui, A., & Sankaran, S. (2020). High-resolution satellite imagery applications in crop phenotyping: An overview. *Computers and Electronics in Agriculture*, *175*, 105584. https://doi.org/10.1016/J.COMPAG.2020.105584

Zhang, J., Guan, K., Peng, B., Pan, M., Zhou, W., Jiang, C., Kimm, H., Franz, T. E., Grant, R. F., Yang, Y., Rudnick, D. R., Heeren, D. M., Suyker, A. E., Bauerle, W. L., & Miner, G. L. (2021). Sustainable irrigation based on co-regulation of soil

water supply and atmospheric evaporative demand. *Nature Communications 2021, 12*(1), 1–10. https://doi.org/10.1038/s41467-021-25254-7

Zhang, Q., Peng, S., & Li, Y. (2019). Increase rate of light-induced stomatal conductance is related to stomatal size in the genus Oryza. *Journal of Experimental Botany*, *70*(19), 5259–5269. https://doi.org/10.1093/JXB/ERZ267

Zhang, Y., Wu, J., & Wang, A. (2022). Comparison of various approaches for estimating leaf water content and stomatal conductance in different plant species using hyperspectral data. *Ecological Indicators*, *142*, 109278. https://doi.org/10.1016/J.ECOLIND.2022.109278

Zhao, L., Wang, L., Li, J., Bai, G., Shi, Y., & Ge, Y. (2021). Toward accurate estimating of crop leaf stomatal conductance combining thermal IR imaging, weather variables, and machine learning. Proc. SPIE 11747, Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping VI, 117470L (12 April 2021); https://doi.org/10.1117/12.2587577

Zhou, Z., Majeed, Y., Diverres Naranjo, G., & Gambacorta, E. M. T. (2021). Assessment for crop water stress with infrared thermal imagery in precision agriculture: A review and future prospects for deep learning applications. *Computers and Electronics in Agriculture*, *182*, 106019. https://doi.org/10.1016/J.COMPAG.2021.106019

Zhu, K., Sun, Z., Zhao, F., Yang, T., Tian, Z., Lai, J., Long, B., & Li, S. (2020). Remotely sensed canopy resistance model for analyzing the stomatal behavior of environmentally-stressed winter wheat. *ISPRS Journal of Photogrammetry and*

Remote Sensing, *168*, 197–207. https://doi.org/10.1016/J.ISPRSJPRS.2020.08.012

Zia, S., Romano, G., Spreer, W., Sanchez, C., Cairns, J., Araus, J. L., & Müller, J. (2013). Infrared Thermal Imaging as a Rapid Tool for Identifying Water-Stress Tolerant Maize Genotypes of Different Phenology. *Journal of Agronomy and Crop Science*, *199*(2), 75–84. https://doi.org/10.1111/J.1439-037X.2012.00537.X