

**AMELIORATING ENVIRONMENTAL EFFECTS ON HYPERSPECTRAL
IMAGES FOR IMPROVED PHENOTYPING IN GREENHOUSE AND
FIELD CONDITIONS**

by

Dongdong Ma

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy



School of Agricultural and Biological Engineering

West Lafayette, Indiana

August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Jian Jin, Chair

Department of Agricultural and Biological Engineering

Dr. Dennis R. Buckmaster

Department of Agricultural and Biological Engineering

Dr. Dharmendra Saraswat

Department of Agricultural and Biological Engineering

Dr. Mitchell R. Tuinstra

Department of Agronomy

Approved by:

Dr. Nathan S. Mosier

To Linzi
Always love.

ACKNOWLEDGMENTS

Ph.D. refers to Doctor of Philosophy. Philosophy relates to the love of learning. Looking back at my four-year Ph.D. life, I feel so grateful to have you all making me enjoy every moment along this incredible journey.

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Jian Jin, for his guidance through each stage of the process, from platform construction, research project, conference presentation, journal publication, to dissertation writing. I appreciate all the opportunities he provides to me, which allows me to explore my real potential in every challenging project. His eagerness, inspiration, and determination in plant phenotyping always encourage me to keep innovative in the research and never give up. Besides being a great academic advisor, Dr. Jin is also a close friend to me, who always gives me the most reliable support whenever I need it. Without his continuous encouragement and patient teaching, I wouldn't have made it this far.

I would also like to thank my committee members, Dr. Dennis Buckmaster, Dr. Dharmendra Saraswat, and Dr. Mitchell Tuinstra, for their valuable comments and insights in my research. Each of the above professors is a giant in the field, and I appreciate them for allowing me to stand on their shoulders. Special thanks to Dr. Tuinstra for his professional advice regarding the experimental design in plant studies. Also, many thanks for the generous help from his lab regarding plant growing and ground truth data collection, which have been extremely important to my research.

I also appreciate the great help from my colleagues in the plant phenotyping lab. It has been an incredible and lucky journey for me to work with them and learn from them. Special thanks to Liangju, Libo, Hideki, and Tanzeel for their support and encouragement. I am also grateful to my friends, Xin, Tong, Jiang, Bai, JK, Xiaowei, Yongkai, Chen, Jiafei, Dongqi, Ding, for making my life so colorful and beautiful. I will always remember every moment at Baywater and Amber3884.

Finally, I would like to express my gratitude to my parents, Jingbo Ma and Liping Zhang, and my wife, Linzi, for making me what I am today. Thank you so much for supporting me unconditionally. Special thanks to my wife, Linzi, for your love, care, support, and every little effort you are doing for me is simply amazing.

TABLE OF CONTENTS

LIST OF TABLES	9
LIST OF FIGURES	11
NOMENCLATURE.....	15
ABSTRACT	17
CHAPTER 1. GENERAL INTRODUCTION.....	18
1.1 Overview.....	18
1.2 Dissertation Outline.....	19
1.3 Literature review	20
1.3.1 Hyperspectral imaging-based plant phenotyping	20
1.3.2 Microclimate heterogeneity in greenhouse	23
1.3.3 Sources of variability in plant reflectance characteristics	25
CHAPTER 2. GREENHOUSE ENVIRONMENT MODELING AND SIMULATION FOR MICROCLIMATE CONTROL.....	27
2.1 Introduction.....	27
2.2 Materials and methods.....	29
2.2.1 Greenhouse geometry and plants layout.....	29
2.2.2 Model Development	30
2.2.3 Computation.....	33
2.2.4 Optimization for greenhouse conveyor movement based on simulation result	34
2.2.5 Measurements and validation.....	36
2.3 Results and discussion	37
2.3.1 Simulated temperature and radiation.....	37
2.3.2 Model validation.....	41
2.3.3 Use model to drive movement of the conveyor	43
2.4 Conclusions.....	44
CHAPTER 3. REMOVAL OF GREENHOUSE MICROCLIMATE HETEROGENEITY WITH A CONVEYOR SYSTEM FOR INDOOR PHENOTYPING	46
3.1 Introduction.....	46
3.2 Materials and methods.....	49

3.2.1	High throughput imaging greenhouse with conveyor system.....	49
3.2.2	Microclimate simulation model to optimize conveyor movement.....	50
3.2.3	Comparison experiment design.....	52
3.2.4	Plant imaging collection and analysis.....	53
3.2.5	Ground truth measurements.....	54
3.2.6	Statistical analysis.....	55
3.2.7	Software and computation.....	55
3.3	Results and discussions.....	56
3.3.1	Projected Leaf Area, RWC and SPAD measurements.....	56
3.3.2	t-test for projected leaf area, RWC, SPAD measurements and plant indices.....	59
3.3.3	Variance analysis for Projected Leaf Area, RWC, SPAD measurements and plant indices.....	59
3.4	Conclusions.....	63
CHAPTER 4. MODELING OF DIURNAL CHANGES IN AERIAL CROP IMAGES.....		64
4.1	Introduction.....	64
4.2	Methods.....	67
4.2.1	High-throughput field imaging acquisition system.....	67
4.2.2	Experiment design and data collection.....	70
4.2.3	Image segmentation and feature extraction.....	71
4.2.4	Data quality check.....	73
4.2.5	Evaluating the impacts from treatments, stages and genotypes to diurnal changing patterns.....	74
4.2.6	Diurnal patterns calculation by time series signal decomposition.....	75
4.2.7	Diurnal pattern fitting.....	75
4.2.8	Model performance evaluation.....	76
4.2.9	Diurnal models' applications.....	76
4.3	Results.....	76
4.3.1	The NDVI diurnal fluctuations.....	77
4.3.2	The impacts of nutrient treatments, genotypes and leaf stages on diurnal variation..	79
4.3.3	Diurnal changing pattern.....	81
4.3.3.1	Decomposition.....	82

4.3.3.2	Pattern fitting.....	83
4.3.3.3	Applications of the model.....	84
4.3.4	Other image-derived phenotyping features.....	85
4.4	Discussions.....	88
4.4.1	Strengths	88
4.4.2	Limitations and future work.....	88
4.5	Conclusions.....	89
CHAPTER 5. MODELING OF ENVIRONMENTAL EFFECTS ON AERIAL HYPERSPSPECTRAL IMAGES FOR CORN PHENOTYPING.....		90
5.1	Introduction.....	91
5.2	Materials and methods.....	93
5.2.1	Experiment design and data collection	93
5.2.2	Time series decomposition for environment-induced variation.....	95
5.2.3	Environmental data transformation and selection	96
5.2.4	Data quality check	97
5.2.5	Artificial Neural network (ANN) model	97
5.2.5.1	Architecture.....	97
5.2.5.2	Training and optimization.....	98
5.2.6	Performance evaluation	99
5.2.6.1	Evaluation metrics	99
5.2.6.2	Multi-model comparison analysis across genotypes and nitrogen treatments	99
5.2.6.3	Phenotyping features for testing the model and workflow	100
5.2.7	Software and computation	100
5.3	Results and discussion	100
5.3.1	Time series decomposition result.....	100
5.3.2	Environmental feature selection and range.....	101
5.3.3	Performance of the ANN models	103
5.3.3.1	Overall performance	103
5.3.3.2	Multi-model comparison analysis across genotypes and nitrogen treatments	105
5.3.4	Modeling of environmentally induced variation in predicted RWC	106
5.4	Conclusions.....	108

CHAPTER 6. SUMMARY AND FUTURE WORK.....	110
6.1 Summary.....	110
6.2 Future work.....	111
REFERENCES	113

LIST OF TABLES

Table 1.1. Distinct wavelength regions in the visible and near-infrared spectral range with relevant functions.	21
Table 2.1. Properties of materials used in the greenhouse microclimate simulation model.	32
Table 2.2. Summary of running and break time for conveyor movement optimization.....	35
Table 2.3. The temperature ranges of the plants from the simulation model.	41
Table 2.4. The inputs of the simulation model for the validation test.	42
Table 3.1. The mean and standard deviation of projected leaf area, RWC and SPAD measurements	57
Table 3.2. $-\text{Log}_{10}$ (<i>P-value</i>) of Projected Leaf Area, RWC and vegetative indices for plant status estimation between the well-watered and drought plant groups in the conveyor and traditional greenhouses.	61
Table 3.3. Standard deviation of Projected Leaf Area, RWC and vegetative indices for plant status estimation of the well-watered and drought plant groups in the conveyor and traditional greenhouses.	62
Table 4.1. Parameters for the hyperspectral imaging sensor.....	68
Table 4.2. Number of image samples after the quality check for six plant plots.	74
Table 4.3. Statistical analysis of the NDVI at Day 1 for six plant plots.....	79
Table 4.4. The DTW distance scores between NDVI's RDR curves of the plant plots with different nitrogen treatments and genotypes.....	81
Table 4.5. The DTW distance scores between NDVI's RDR curves of the plant plots with different plant stages	81
Table 4.6. Fit and error of the piecewise NDVI diurnal changing pattern model (Eq 4.8) on assorted experiment plots.....	84
Table 4.7. Proper imaging windows for different thresholds of the NDVI changes caused by diurnal factors (when solar noon is the standard imaging time).....	85
Table 4.8. The fitted results of piecewise diurnal changing pattern models for other plant phenotyping features including predicted RWC, Red and NIR bands in the spectra.	86
Table 5.1. Abbreviations of plant plots with different nitrogen treatments and genotypes.	94
Table 5.2. Hyperspectral images and environmental data collection.	95
Table 5.3. Data pool after data quality check.	97
Table 5.4. The ranges in environmental conditions experienced by the modeling data during the 31 days.....	103

Table 5.5. The results of the two-sample *t*-test between the variance of daily NDVI before and after ANN model correction.105

LIST OF FIGURES

Fig. 1.1. The spectral reflectance curve of vegetation. The major absorption and reflectance features are indicated (Roman and Ursu, 2016).	21
Fig. 2.1. Overall greenhouse geometry and plants layout. 3D coordinates represent x as north, y as west, and z as zenith.....	30
Fig. 2.2. Purdue Lily 13-4 Greenhouse with automated, high-throughput imaging system and a belt conveyor-based plant transportation system.....	30
Fig. 2.3. The air flow inside and around house related with temperature and radiation.....	33
Fig. 2.4. Greenhouse is divided into sub faces by meshing, with facets as triangles.	34
Fig. 2.5. Layout of simulated plants in the greenhouse.....	35
Fig. 2.6. (a). The layout of the Xiaomi Flower Care Sensor including temperature and radiation sensors; (b). The sensor was inserted into the soil of the pot during data collection.....	37
Fig. 2.7. Simulated radiation (W/m ²) at four different time points during a day: (a). 8am; (b). 12pm; (c). 18pm; (d). 20pm. The radiation value is shown in grey-scale color legend.	38
Fig. 2.8. (a) Positions of five samples labeled from top view of greenhouse, (b) simulated radiation (W) of plant with the time (hour). The simulated radiation is calculated for the total integrated plant body surface.	39
Fig. 2.9. Simulated temperature (°C) at four different time points during the day, (a) 8am, (b) 12pm, (c) 18pm and (d) 20pm. The temperature value is shown in grey-scale color legend.....	40
Fig. 2.10. (a). Six sample positions labeled from top view of greenhouse; (b). Simulated temperature of plant with the time (hour).	41
Fig. 2.11. Accumulated radiation results: actual versus simulated measurements.	42
Fig. 2.12. Accumulated temperature results: actual versus simulated measurements.	43
Fig. 2.13. Coefficient of variation (CV) of the simulated radiation received by plants from different running and break time combinations for conveyor movement optimization.....	44
Fig. 3.1. Purdue Lily Greenhouse with an automated, high-throughput imaging system and a belt-conveyor-based plant transportation system.....	49
Fig. 3.2. High-throughput hyperspectral imaging system with belt conveyor system: (a). The layout of the conveyor system and the direction of plant movement; (b). A high-throughput hyperspectral imager tower.	50
Fig. 3.3. : Overall greenhouse geometry: the air flow inside and around house related with temperature and radiation.	51
Fig. 3.4. The growth of maize plants in two greenhouses at different leaf stages: (a). Traditional greenhouse; (b). Conveyor greenhouse.	53

Fig. 3.5. The top-view images of maize in the hyperspectral imaging tower and the segmented binary images: (a). Well-watered; (b). Drought.	54
Fig. 3.6. NDVI of plants from different water treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.	56
Fig. 3.7. The probability density of Projected Leaf Area of maize from both treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.	58
Fig. 3.8. The probability density of measured relative water content (RWC) of maize from both treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.	58
Fig. 3.9. The probability density of SPAD measurements of maize from both treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.	58
Fig. 4.1. Field VNIR hyperspectral platform at Purdue University. (a). VNIR hyperspectral imaging sensor; (b). Local weather station (Ambient Weather, Chandler, AZ, USA); (c) and (d). Xiaomi flower care sensor (Xiaomi Inc., Beijing, China); (e). The layout of the East-West orientated imaging system; (f). Image sample of the whole field; (g). Enlarged image sample of part of the field; (h). Binary image after segmentation; (i). Layout of the plots with three nitrogen treatments and two genotypes. The green boxes were for Genotype P1105AM, and the blue boxes were for Genotype B73 x Mo17. Nitrogen treatments of high nitrogen (HN), low nitrogen (LW) and medium nitrogen (MN) were also labeled; (j). NDVI heatmap; (k). The spectra from different genotypes and N treatments.	69
Fig. 4.2. The NDVI heatmaps for the whole field at three leaf stages with different accumulated days after planting (DAP). (a). 31 DAP, leaf stage V4. The red boxes indicate the regions for 30 distinct plots; (b). 38 DAP, leaf stage V6; (c). 49 DAP, leaf stage V9.	70
Fig. 4.3. The growth of corn plants in the Purdue ACRE field during the experiment at different accumulated days after planting (DAP).	71
Fig. 4.4. The Relative Water Content (RWC) prediction model based on Partial Least Square Regression Relative (PLSR): measurement vs. prediction.	73
Fig. 4.5. (a). The NDVI of HN and Genotype B73 x Mo17 plot from V4 stage to the R1 stage (obtained from the hyperspectral images); (b). The NDVI measurements across Day 1 (black dots). The red curve is the predicted NDVI diurnal variance by PROSAIL model. Parameters for the PROSAIL model for the corn canopies followed the work of Ishihara et al. in 2015.	78
Fig. 4.6. The mean NDVI relative difference ratio (RDR) curves of six plant plots versus time (H): (a). HN & B73 x Mo17; (b). HN & P1105AM; (c). MN & B73 x Mo17; (d). MN & P1105AM; (e). LN & B73 x Mo17; (f). LN & P1105AM.	80
Fig. 4.7. The NDVI of HN and Genotype B73 x Mo17 plot from V4 stage to the R1 stage. The raw NDVI plot was decomposed into the day-to-day trend (Tt) and diurnal pattern (Dt). The red boxes are the days with incomplete data measurements due to the extreme weather conditions, which are 36, 37, 58, 59 and 60 DAP.	82
Fig. 4.8. The diurnal changes of NDVI summarized from 31 days and six experimental plots. The x-axis is the diurnal time in a unit of hour. The y-axis is the diurnal NDVI changes. The black line	

is the mean diurnal NDVI adjustment value. The shaded area is the 95% confidence interval. The red line is the first-order piecewise fitted result for the mean diurnal NDVI adjustment value. ..83

Fig. 4.9. The intersections between the adjusted NDVI’s diurnal changes and three different thresholds for proper imaging windows. The x-axis is the diurnal time in a unit of hour. The y-axis is the adjusted NDVI’s diurnal changes when adjustment at solar noon to be 0.....85

Fig. 4.10. The diurnal changing patterns and fitted results of other plant phenotyping features including: (a). predicted RWC; (b). a red band; (c). a NIR band. The black line in each plot is the mean diurnal adjustment value. The shaded area is the 95% confidence interval. The red line is the piecewise fitted result for the mean diurnal adjustment value.87

Fig. 5.1. Field VNIR hyperspectral platform at Purdue University. It consists of a VNIR hyperspectral imaging sensor (MSV-101-W, Middleton Spectral Vision, Middleton, WI, USA) and a local weather station (Ambient Weather, Chandler, AZ, USA). The gantry platform is seven meters high capable of scanning all or part of a 50-by-5 meter strip field under a wide range of weather conditions.95

Fig. 5.2. The ANN architecture: Input layer (15 neurons) - Hidden layer 1 (100 neurons, followed with Leaky ReLU) - Hidden layer 2 (1000 neurons, followed with Leaky ReLU) - Output layer (1 neuron).....98

Fig. 5.3. The NDVI of the sample dataset (G1H) from the V4 stage to the R1 stage. The raw NDVI plot was decomposed into the day-to-day trend (Tt) and the periodic change (Dt). The red boxes are the days with incomplete data measurements due to the extreme weather conditions, which are DAP 36, 37, 58, 59 and 60. 101

Fig. 5.4. Single-factor correlation analysis for NDVI. The x-axis is the environment factors and their squared or root squared formats. The y-axis is the adjusted R^2 between each x variable and the calculated environment-induced variation in NDVI. Larger adjusted R^2 value means the variable is more correlated to the NDVI variation..... 102

Fig. 5.5. The five-fold cross-validated prediction results of environment-induced variation in NDVI for the sample dataset (G1H). The ANN prediction values show a significant correlation with $R^2=0.823$ and $RMSE=0.006$ 103

Fig. 5.6. Box plots for the five-fold cross-validated correction result of the sample dataset (G1H). (1). The raw NDVI with huge daily variances across 31 days. (2). The ANN model corrected NDVI with much more condensed boxes. 104

Fig. 5.7. Two-sample *t*-test between the variance of daily NDVI for the sample dataset (G1H) before and after ANN model correction..... 105

Fig. 5.8. Accuracy heatmaps of R^2 and RMSE of ANN models for NDVI. Red boxes: the region with relatively poor predictive results..... 106

Fig. 5.9. The five-fold cross-validated prediction results of environment-induced variation in predicted RWC for the sample dataset (G1H). The ANN prediction values show a significant performance with $R^2=0.791$ and $RMSE=0.722\%$ 107

Fig. 5.10. Box plots for the 5-fold cross-validated correction result of sample dataset (G1H). (1). The raw predicted RWC showed huge daily variances across 31 days. (2). The ANN model corrected predicted RWC has much more condensed boxes..... 108

NOMENCLATURE

ACRE	Agronomy Center for Research and Education
ANN	Artificial Neural Network
BDRF	Bidirectional Reflectance
Carte1	Vegetation Index named after Carte
CV	Coefficient of Variation
DAP	Days after Planting
Datt1	Vegetation Index named after Datt
DTW	Dynamic Time Warping
EVI	Enhanced Vegetation Index
GNDVI	Green Normalized Difference Vegetation Index
GPU	Graphic Processing Unit
GSD	Ground Sample Distance
IQR	Interquartile Range
KDE	Kernel Density Estimate
MCARI	Modified Chlorophyll Absorption in Reflectance Index
mNDVI705	Modified Red Edge Normalized Difference Vegetation Index
MSV	Middleton Spectral Vision
MTVI1	Modified Triangulation Vegetation Index-1
NDCI	Normalized Difference Chlorophyll Index
NDVI	Normalized Difference Vegetation Index
NDVI705	Red Edge Normalized Difference Vegetation Index
NIR	Near Infrared
NVI	New Vegetation Index
OSAVI	Optimized Soil Adjusted Vegetation Index
PCA	Principle Component Analysis
PLA	Projected Leaf Area
PLSR	Partial Least Squares Regression
PRI	Photochemical Reflectance Index

RDR	Relative Difference Ratio
RDVI	Renormalized Difference Vegetation Index
REP	Red Edge Position Index
RGB	Red-Green-Blue
RI1dB	Ratios Corresponding to 1dB
RMSE	Root Mean Square Error
RWC	Relative Water Content
SPAD	Soil Plant Analysis Development
SPVI	Spectral Polygon Vegetation Index
SR [675,700]	Simple Ratio between 675nm and 700nm
SWIR	Short Wave Infrared
TCARI	Transformed Chlorophyll Absorption Reflectance Index
TVI	Transformed Vegetation Index
UAV	Unmanned Aerial Vehicles
VNIR	Visible to Near Infrared
VOG1	Vogelmann Red Edge Index 1

ABSTRACT

Hyperspectral imaging has become one of the most popular technologies in plant phenotyping because it can efficiently and accurately predict numerous plant physiological features such as plant biomass, leaf moisture content, and chlorophyll content. Various hyperspectral imaging systems have been deployed in both greenhouse and field phenotyping activities. However, the hyperspectral imaging quality is severely affected by the continuously changing environmental conditions such as cloud cover, temperature and wind speed that induce noise in plant spectral data. Eliminating these environmental effects to improve imaging quality is critically important. In this thesis, two approaches were taken to address the imaging noise issue in greenhouse and field separately. First, a computational simulation model was built to simulate the greenhouse microclimate changes (such as the temperature and radiation distributions) through a 24-hour cycle in a research greenhouse. The simulated results were used to optimize the movement of an automated conveyor in the greenhouse: the plants were shuffled with the conveyor system with optimized frequency and distance to provide uniform growing conditions such as temperature and lighting intensity for each individual plant. The results showed the variance of the plants' phenotyping feature measurements decreased significantly (i.e., by up to 83% in plant canopy size) in this conveyor greenhouse. Secondly, the environmental effects (i.e., sun radiation) on aerial hyperspectral images in field plant phenotyping were investigated and modeled. An artificial neural network (ANN) method was proposed to model the relationship between the image variation and environmental changes. Before the 2019 field test, a gantry system was designed and constructed to repeatedly collect time-series hyperspectral images with 2.5 minutes intervals of the corn plants under varying environmental conditions, which included sun radiation, solar zenith angle, diurnal time, humidity, temperature and wind speed. Over 8,000 hyperspectral images of corn (*Zea mays* L.) were collected with synchronized environmental data throughout the 2019 growing season. The models trained with the proposed ANN method were able to accurately predict the variations in imaging results (i.e., 82.3% for NDVI) caused by the changing environments. Thus, the ANN method can be used by remote sensing professionals to adjust or correct raw imaging data for changing environments to improve plant characterization.

CHAPTER 1. GENERAL INTRODUCTION

1.1 Overview

With the human population that is expected to reach 9.7 billion by the end of 2050, the global crop production should be doubled to meet the rising demands (Ehrlich and Harte, 2016; Ray et al., 2013). Boosting crop yields is a critical solution (Conceição and Mendoza, 2009). To select high-yield plants faster and more efficiently, modern phenotyping is the key. Hyperspectral imaging has become one of the most important phenotyping technologies in the recent decades (Li et al., 2014a). Compared to the traditional imaging technologies like greyscale and red-green-blue (RGB) imaging, hyperspectral imaging captures hundreds of bands, which enable researchers and scientists to efficiently and accurately measure many plant physiological features such as plant biomass, leaf moisture content, and chlorophyll content (Clevers et al., 2008; Di Gennaro et al., 2018; Garaba and Dierssen, 2018; Li et al., 2014). Currently, various hyperspectral imaging platforms exist. Imaging towers have been built for scanning the plants grown in the greenhouse (Ma et al., 2019a). Hyperspectral sensors are also mounted on the ground-based vehicles to image the plants in the field (Li et al., 2014). The unmanned aerial vehicle (UAV) is widely used to image the plants in the field and it has become one of the most popular imaging platforms in field hyperspectral studies (Garaba and Dierssen, 2018). Finally, satellites are also used to collect hyperspectral images for regional plant growth measurements (Ishida et al., 2017). Hyperspectral imaging has advanced plant phenotyping with more accurate measurements of plant features and has largely accelerated the progress of plant breeding studies (Di Gennaro et al., 2018). However, current hyperspectral imaging data still suffers from severe noise caused by the changing environment during the imaging, which greatly affects its quality (Maji et al., 2014).

Plant hyperspectral phenotyping suffers from several major challenges. Firstly, plants in the experiment are treated with a non-uniform environment, especially in the greenhouse. The fans, cooling wall and temperature managing system are used to create more desirable environment for the growth of plants (Choab et al., 2019). However, these systems create non-uniform microclimates which make the plants in the greenhouse look different from each other (Aaslyng et al., 2007). For example, some plants grow directly under the lamps, while others are away from the lights. Furthermore, the distances between the plants and the fans and cooling walls vary. All

these generate a significant gradient in temperature and lighting profiles in the same room (Ahonen et al., 2008). These non-uniformities will cause unwanted noise in plant phenotyping. The second challenge mostly arises in the field, where the environment changes frequently and is uncontrollable. For example, the plant imaging data is greatly affected by the changing intensity and color of the sun light. Besides, the plant reflectance spectrum is affected by other environmental factors such as changing cloud cover, wind speed, and temperature (Gamon et al., 2015). Moreover, the plants themselves change over time, which is known as circadian behavior. For example, the Normalized Difference Vegetation Index (NDVI) of crops measured in the morning differs by over 10% from one assessed at noon (Beneduzzi et al., 2017; Stickse et al., 2004). These variations in environmental conditions and time cause significant changes in the spectra data, which further reduce the accuracy of plant analysis.

It is important to discern between plant signal and environmental noise in plant imaging. Unfortunately, plant phenotyping researchers are usually looking for weak signals (e.g., nitrogen improvement that varies by less than 3%) in severely inconsistent environmental conditions, such as non-uniform environments in the greenhouse, different imaging times, or various sun radiations (Oliveira and Scharf, 2014). Therefore, reducing the impact of environmental noise on the hyperspectral imaging data is critically needed by phenotyping researchers. This dissertation develops methods to analyze, model, and eliminate the effects of environmental noise on the crop hyperspectral phenotyping data in both greenhouse and field conditions.

1.2 Dissertation Outline

The overall objective of the research is to analyze and ameliorate the environmental effects on hyperspectral images for improved phenotyping quality in greenhouse and field conditions.

The organization of this dissertation is as follows. In chapter 1, the overview of this dissertation is presented, followed by a review of hyperspectral imaging in plant phenotyping, and the challenges from the environmental impacts in greenhouse and field conditions.

Chapter 2 and Chapter 3 introduce the research work conducted in greenhouse phenotyping, aiming to develop a model-driven approach for the removal of greenhouse microclimates heterogeneity with an automated conveyor system. Chapter 2 presents the study on greenhouse environment modeling and simulation for microclimates estimation. A computational greenhouse

environment model is introduced to quantitatively simulate the 3-D temperature and radiation distributions over the 24-hour cycle in a local greenhouse. The simulation results of temperature and radiation distributions are then utilized to optimize pot movement (distance and frequency) in a greenhouse equipped with an automated conveyor system. Chapter 3 introduces a newly constructed microclimates-impact-alleviated automatic greenhouse by applying the methodology in the Chapter 2 to minimize microclimate heterogeneity. The performance of model-optimized conveyor movement solution is evaluated with a comparison test assay for plant variances between the conveyor greenhouse and a neighboring traditional greenhouse.

Chapter 4 and Chapter 5 document the work in field phenotyping, focusing on correcting the aerial remote sensing results considering environment variations. Chapter 4 describes the research on the changing patterns of aerial crop phenotyping features along the diurnal timeline. A field experiment is conducted with the new field imaging gantry at Purdue University to collect a substantial number of time-series hyperspectral images on corn at a high sampling frequency throughout the growing season. The collected images are then used to analyze and model the changes in crop phenotyping features (e.g., NDVI) along the diurnal time. Chapter 5 extends the exploration of Chapter 4. It introduces a new ANN model trained with synchronized hyperspectral imaging data and environmental data (e.g., sun radiation, temperature and wind speed) for predicting the environment-induced variations in crop phenotyping features. The performance of the ANN model is evaluated by comparing the variations in the phenotyping result before and after the model correction.

In Chapter 6, the results from this dissertation are summarized. Some limitations of the proposed methods are discussed, and potential solutions are proposed for the future work.

1.3 Literature review

1.3.1 Hyperspectral imaging-based plant phenotyping

Modern plant breeding takes advantage of the combination of genotypic and phenotypic data. With the rapid development of advanced phenotyping technologies, plant phenotypic properties can be more efficiently captured in the imaging data in a non-destructive way. The captured images are then processed with imaging process algorithms and statistical tools to access plant traits such as drought tolerance (Causse et al., 2018), disease resistance (Miklas et al., 2019; Stone et al.,

2010), seed dormancy (Torada et al., 2005), and photosynthesis (Herve et al., 2001), which largely facilitate the period of plant breeding (Li et al., 2014a). Among all the modern phenotyping technologies such as RGB, multispectral, and thermal imaging, hyperspectral imaging represents one of the most popular.

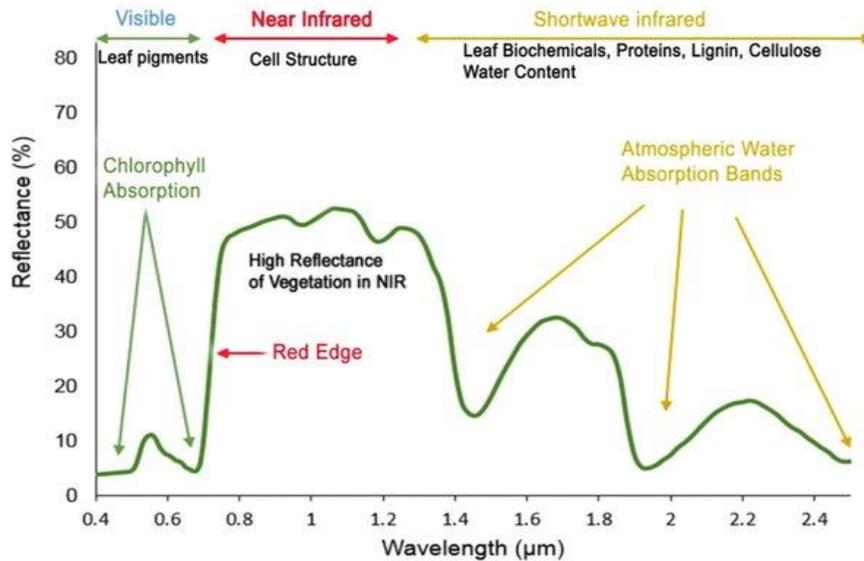


Fig. 1.1. The spectral reflectance curve of vegetation. The major absorption and reflectance features are indicated (Roman and Ursu, 2016).

Table 1.1. Distinct wavelength regions in the visible and near-infrared spectral range with relevant functions.

Name	Wavelength (nm)	Traits measured	References
Visible (VIS)	400–700	Photosynthetic pigments such as chlorophyll a, chlorophyll b, carotenoids, and xanthophyll.	(Broge and Leblanc, 2001)
Near-infrared (NIR)	700–1,100	Chlorophylls, water, or macronutrients (e.g., nitrogen), various vegetative indices (e.g., NDVI).	(Ma et al., 2019a; Ni et al., 2015; Zhang et al., 2019a).
Short-wave-infrared (SWIR)	1,100–2,500	Biochemical components such as protein, phosphorus, cellulose, and hemicellulose. Several water relevant bands (e.g., 1160 nm, 1440 nm, and 1930 nm).	(Tian and Philpot, 2015)

In plant hyperspectral imaging, the visible and near infrared spectral region (400–2500 nm) is usually measured (Fig. 1.1), which has proven effective for the identification of various plant

phenotypic characteristics such as plant biomass, leaf moisture content, and chlorophyll content (Pandey et al., 2017). The relationship between spectral reflectance and plant phenotypic components has been studied for decades (Gewali et al., 2018; Reflectance, 1973). Typically, based on the functions, the spectral region (400–2500 nm) is usually partitioned into three distinct wavelength regions: the visible (VIS, 400–700 nm), near-infrared (NIR, 700–1,100 nm) and short-wave-infrared (SWIR, 1,100–2,500 nm) spectral regions (Pandey et al., 2017). As shown in Table 1.1, the VIS region is the one most responsible for photosynthetic pigments such as chlorophyll a, chlorophyll b, carotenoids, and xanthophyll (Broge and Leblanc, 2001). The NIR region is determined by the structural reflection of the turgid plant cell (Pandey et al., 2017), and it is widely used for predicting essential leaf constituents such as chlorophylls, water, or macronutrients (e.g., nitrogen) (Ma et al., 2019a; Ni et al., 2015; Zhang et al., 2019a). Healthy plant tissue typically has unique characteristics that it absorbs in the red and blue wavelengths, reflects in the green wavelength, and strongly reflects in the near-infrared (NIR) wavelength. Therefore, the combined VIS and NIR wavelength is called VNIR, which represents the most commonly used range for plant analysis. For example, several vegetation indices are based upon the VNIR range, with the most popular represented by the Normalized Difference Vegetation Index (NDVI) (Thenkabail et al., 1997) and Carotenoid Reflectance Index 1 (CRI1) (Stylinski et al., 2002). Besides that, the SWIR region is used to quantify a wider range of biochemical components such as protein, phosphorus, cellulose, and hemicellulose. Moreover, there are several water relevant bands in the SWIR range, including 1160 nm, 1440 nm, and 1930 nm (Tian and Philpot, 2015).

With the development of the advanced imaging technologies and platforms, hyperspectral imaging on plant phenotyping is more capable and accurate in analyzing plant features, which largely accelerates the progress of plant breeding studies (Di Gennaro et al., 2018b). However, the quality of plant hyperspectral imaging data still suffers from severe environmental effects in both greenhouse and field conditions (Maji et al., 2014). In greenhouse, the plants are not exposed to the uniform environment. The environment within the same greenhouse is still heterogeneous, where diverging regions are referred to as microclimates. Greenhouse plant science assays have been impacted by heterogeneous microclimates which causes significant level of noise to hyperspectral imaging measurements (Ahonen et al., 2008). In field, the plants are not imaged under the same environmental condition. The changing environmental conditions have been reported to have significant impacts on the imaging result (Gamon et al., 2015). The intensity in

remotely sensed images is severely impacted based on the phenomenon of when and where the image is captured (Beneduzzi et al., 2017; Maji et al., 2014; Padilla et al., 2019). One source of the variation is due to the complicated interactions between camera's sensitivity, camera's view angle, canopy geometry, solar zenith angle, solar azimuth angle, and shadows (Danner et al., 2019; Ishihara et al., 2015; Jackson et al., 1979; Ranson et al., 1985). Another source of variation results from plants' endogenous stress responses to the environmental conditions with complicated interactions between their genetic backgrounds, external environments, and treatments (An et al., 2017; Ranson et al., 1985). All of these, collectively regarded as the environment-induced variation in phenotyping features, affect plants' final reflectance characteristics.

Although the spectral reflectance signature of the plants contains the rich information about their biochemical, physiological and physical characteristics, the noise from the changing environment, in either greenhouse or field condition, still limit the accuracy of spectral phenotyping results. It is critical to understand the cause of these environmental noise and develop relevant solutions to eliminate the environmental effects on hyperspectral images for improved phenotyping in greenhouse and field conditions.

1.3.2 Microclimate heterogeneity in greenhouse

Greenhouses use solar irradiance to create a favorable microclimate for plant growth (Choab et al., 2019). Due to the advantages of generating relatively more stable environments and efficiently achieving controlled plant growth compared to field tests (Vásquez et al., 2015; Zhang et al., 2019a), greenhouse facilities have proven more attractive than ever as crop research tools. However, the microclimate within a greenhouse varies considerably (Alain, 1989). Various factors lead to the greenhouse microclimate heterogeneity including ventilation (Norton et al., 2007), cooling wall (McCartney et al., 2018), greenhouse architecture (Taki et al., 2016), greenhouse location (Çakir and Sahin, 2015), internal lighting position (LI et al., 2003), and ambient control system (Kläring et al., 2007). Controlling the microclimate while accounting for these features, which are intrinsic to a greenhouse, is never easy.

The microclimate heterogeneity has received considerable attention by greenhouse researchers (Alain, 1989; Baille et al., 2001; Hartung et al., 2019; Körner et al., 2007). Environmental factors such as temperature, airflow, relative humidity, radiation, and carbon dioxide concentration were found to be non-uniformly distributed in the greenhouse, and these

impact the growth, production and quality of the crops (Ahemd et al., 2016; Alain, 1989; Baille et al., 2001; Castro et al., 1991; Körner et al., 2007). These heterogeneities resulted in heterogeneous stress signals on plants, eventually resulting in inconsistent greenhouse experiments (Ahonen et al., 2008; Brien et al., 2013).

Researchers have continued to explore methods to eliminate the heterogeneity of the greenhouse microclimates, both directly and indirectly. The most common and relatively simple way involves the direct improvement of the greenhouse either in coating the surface or redesigning the construction shape (Baille et al., 2001; Taki et al., 2016). By whitening the greenhouse roof located in the coastal area of eastern Greece, the nonuniform environment issue was well addressed (Baille et al., 2001). Moreover, to indirectly reduce microclimate heterogeneity, people use a randomized complete block design (RCBD) method to statistically randomize the positions of crop samples in the greenhouse, which randomly redistributes the variance instead of removing it (Hartung et al., 2019). However, the RCBD method just helps to redistribute the environmental noise on plants but does not remove the noise. Moreover, this method is not suitable when complete block contains considerable variability.

Recently, benefiting from the rapid advancements in the indoor plant transport facility, an innovative solution of relocating plants during the experiment has arisen. One study detailed a greenhouse containing plants on conveyor systems to account for the variation of the microclimates (Brien et al., 2013). The strategy of cycling plants through greenhouse locations relies on the assumption that plants experience uniform growth conditions when spending an equal amount of time in each microclimate (Brien et al., 2013). However, this shuffle of plants was done at the lane level. Groups of plants sitting on the same lane were moved around together, but the microclimate impacts within the lane remained. For example, plants in the middle of the lane are higher than the plants at the edge. Moreover, their conveyor facility was under simple mechanical movements with empirical operation scenarios. The optimization of the conveyor movement for efficiently relocating plants with minimized conveyor movement cost have rarely been explored.

To ensure the success of this plant relocation scheme, plants must spend an equal amount of time in each microclimate, preferably during equivalent growth stages. To date, methods to efficiently relocate individual plant while minimizing conveyor movement cost have not been addressed.

1.3.3 Sources of variability in plant reflectance characteristics

Plant canopy reflectance equals the ratio of the amount of light leaving the canopy to the amount of incoming light (Kitchen et al., 2001). Current remote sensing techniques rely heavily on this ratio to obtain information on photosynthetic pigment compounds, vegetation indices, water content, or other biochemical contents in crops (Meyer and Neto, 2008; Neuwirthová et al., 2017; Quemada et al., 2014). Plant canopy reflectance actively changes over time, especially in the field, which introduces a great deal of variation in plant phenotyping results (Beneduzzi et al., 2017; Maji et al., 2014; Padilla et al., 2019). The variability in canopy reflection is complex due to the diversity in the canopy size, shape, composition, and arrangement of cells, leaves, stems, and plants within ecosystems (Ollinger, 2011). However, the sources for nearly all such variability can be summarized into two categories: changes in leaf optical properties and sun-leaf-sensor geometry.

The interaction of spectral radiation and plant canopy, including reflectance, transmission, and absorption, depends on optically relevant constituents such as pigments, liquid water, and several other biochemical components (Blackburn, 1998; Ceccato et al., 2001; Kokaly et al., 2009). During growth, the plants' chemical constituents as the environment does (Gamon et al., 2015; Ishihara et al., 2015; Jackson et al., 1979; Ranson et al., 1985). Environmental factors such as air temperature, solar zenith angle, solar azimuth angle, shadows, humidity, and wind speed were found to be highly correlated with the change of chemical constituents in plants (An et al., 2017; Gamon et al., 2015; Oliveira and Scharf, 2014; Zhao et al., 2018; Zhou et al., 2017). In addition to their environments, plants' chemical constitutions also change over time. Plant physiological properties usually follow periodical behavior, which is commonly referred to as circadian rhythmicity (Greenham et al., 2015; Tindall et al., 2015). The circadian rhythms in plants represent the subset of biological rhythms that occur over a period of time (Ast and Dunlap, 2004), which have been observed in a diverse range of plant species including corn, potato, rice, wheat, barley and soybean during the last decades (Dunford et al., 2002; Kloosterman et al., 2013; Preuss et al., 2012; Tindall et al., 2015; Turner, 2007).

Furthermore, the incidence of radiation on the leaf and different angles of observation (sun-leaf-sensor geometry) impacts the plant canopy's reflectance (Atrashevskii et al., 1999). During the reflectance measurement, the sensor viewing direction is usually fixed (Gamon et al., 2015; Ranson et al., 1985). Therefore, most changes in the sun-leaf-sensor geometry arise from the

changing sun position and plant leaf angle during the day. These changes affect the scattering on the plant canopy, which is eventually reflected in the change in observed reflectance characteristics captured by the sensor. For example, one study found the reflectance factor and view angle strongly impacted the reflectance characteristics observed for all of the plant canopies considered (Ranson et al., 1985). The reference change from sun-leaf-sensor geometry is also theoretically explained by the Bidirectional Reflectance Distribution Function (BRDF) (Honkavaara et al., 2009). In recent decades, BRDF has been widely used for many remote sensing studies to correct the distortions caused by the view and illumination angle effects (Gatebe and King, 2016; Liang and Strahler, 1994; Qi et al., 2000).

The combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model (PROLSAIL) was developed (Jacquemoud and Baret, 1990) to more comprehensively analyze the variability in plant canopy reflectance characteristics (Duan et al., 2014; Jacquemoud et al., 2009; Ni et al., 2015). As previously mentioned, canopy reflectance represents the joint product of several spectrally relevant factors (leaf optical properties and sun-leaf-sensor geometry), which affect the electromagnetic radiation transfer in spectral regions. PROSAIL allows the decomposition of the effects of these factors on the canopy reflectance (Kattenborn, 2019). Furthermore, sensitivity analyses performed in some studies advanced our understanding of how biophysical variables and geometric factors (e.g., incident radiation on the leaf) contribute to the variability in canopy reflectance (Berger et al., 2018a; Kattenborn, 2019).

However, the PROLSAIL model usually does not meet the accuracy requirement in plant phenotyping remote sensing. For example, Berger et al. (Berger et al., 2018b) compared PROSAIL's simulation result with the spectra data collected from the field and found that PROSAIL's predicted spectra showed severe drifts from the field measurements especially at the green and red edge (~700nm) wavelengths (Berger et al., 2018b). Besides, the PROSAIL prediction theoretically requires three input variables including leaf structure parameter, photosynthetic pigment concentration, and water content, while they are difficult and costly to measure in remote sensing practices (Jacquemoud et al., 2009). Therefore, a more accurate environmental impact analysis for plant reflectance characteristics is still critically needed to improve the quality of agricultural remote sensing.

CHAPTER 2. GREENHOUSE ENVIRONMENT MODELING AND SIMULATION FOR MICROCLIMATE CONTROL

Greenhouse plant science assays have been impacted by heterogeneous microclimates that cause significant noise in plant growth measurement data. Researchers and scientists have randomized pots locations in greenhouses, which helps to redistribute, but does not remove the noise. The varying impacts from microclimates can be eliminated by shuffling plants, but no studies have been conducted on how such shuffling can be optimized, such as by adjusting the frequency and distance of pot movements. The quantitative study of the microclimates in the greenhouse is important to properly optimize the shuffling pattern. This study proposes a computer modeling approach to simulate a greenhouse's microclimates, then employs these results to optimize pot movement distance and frequency. A computational greenhouse microclimate simulation model was developed using inputs from the actual design, materials and location of a Purdue Lily greenhouse in West Lafayette, Indiana. This model predicted the microclimate variables, including ambient temperature and lighting radiation over a 24-hour, seven-day period. Thermometers and lighting sensors were also distributed in the greenhouse to measure the ground truth over a seven-day period. A comparison of the microclimate variables measured with those predicted by the model, including temperature and radiation, demonstrated that the simulation could precisely predict temperatures and light radiation at any time and in different positions in the greenhouse. The simulation results were then used to determine the optimal pot movement frequency and distance. The resulting shuffling pattern of the conveyor movement can remove over 90% of the variance in microclimate radiation as well as eliminate more than 95% of the conveyor motion compared with non-stop movement.

2.1 Introduction

Currently, greenhouse facilities serve as prevalent crop research tools whose advantages includes the ability to easily achieve optimal plant growth and, provides a relatively more uniform environment than the field tests do (Vásquez et al., 2015; Zhang et al., 2019a). However, due to intrinsic greenhouse features, the climate within the same greenhouse retains heterogeneous and can be considered microclimates (Alain, 1989). The nonuniform distributions of temperature,

relative humidity, carbon dioxide concentration and radiation reflected in these microclimates have recently received attention from several researchers (Ahemd et al., 2016; Alain, 1989; Baille et al., 2001; Castro et al., 1991; Körner et al., 2007), as they unevenly impact on the growth, production, and quality of crops. In particular, temperature and radiation represent two major factors that cause heterogeneous microclimatic effects (Gonzalez-Real and Baille, 2000; Lindquist et al., 2005; Stone et al., 1998).

Researchers have continued to explore methods to eliminate the uneven impact of the greenhouse microclimate. Baille et al. (2001) introduced the greenhouse whitening method. They studied how whitening a greenhouse roof influenced its microclimate and canopy behavior during the summer in a greenhouse located in the coastal area of eastern Greece. Another innovative study was conducted by Brien et al. (2013), who designed a greenhouse that contained a plant conveyor system to account for microclimate variation. To ensure the success of the plant relocation methods, plants needed to spend an equal amount of time in each microclimate, preferably during comparable growth stages. In addition to the redesign of greenhouses, studies have also examined how to optimally control the greenhouse environment.

To more precisely control the local crop temperature, a simulation model developed by Körner et al. (2007) modeled the time-dependent crop microclimate from a greenhouse macroclimate. By employing temperature measurements from the area close to the leaf surfaces collected periodically by PT100 thermometer located above the crop, the model predicted the greenhouse microclimate. Teitel et al. (2008) performed a similar study examining the temperature distributions within a greenhouse. They investigated the ambient temperature distribution inside a greenhouse equipped with vertical roof openings. The computational methods they employed proved useful in the prediction of the local temperature distribution for greenhouse climate control purposes. However, these models mainly focused upon temperature and failed to include radiation data, equally distributed within the greenhouse environment. Moreover, they rarely discussed how their simulation model could aid in solving the real greenhouse microclimate problem.

Greenhouse generates more controlled environments for plant growth. However, as discussed above, the microclimate impacts such as the non-uniform temperature and radiation distributions are still unsolved due to the lack of understanding of the microclimates. With the goal of helping greenhouse researchers to improve plant phenotyping quality, this chapter describes the

research work aiming at developing greenhouse microclimate impacts removal solutions, with three major objectives as follows:

1. Develop a greenhouse microclimate simulation model to quantitatively analyze the temperature and radiation distributions.
2. Validate the simulation results with measurement data from regular radiation and temperature sensors.
3. Utilize the simulation model to optimize the conveyor system settings in the greenhouse to save conveyor motion cost while upholding the standard of efficiency in eliminating microclimate impacts.

2.2 Materials and methods

2.2.1 Greenhouse geometry and plants layout

The simulation study was designed and carried out in the Lily 13-4 greenhouse (latitude 40.4259° N, longitude 86.9081° W) at Purdue University, West Lafayette, Indiana (Fig. 2.1). The N-S orientated greenhouse facility is covered with transparent glass along with an automated high throughput imaging and a belt conveyor-based plant transferring system (Fig. 2.2). In addition to the existing microclimatic effects, the installation of a large imaging box across the northern wall further escalates these effects. The temperature of the greenhouse is regulated using an evaporative cooling system consisting of two ventilation fans (installed on the southern wall of the greenhouse) and rectangular evaporative cooling pads (installed on the northern wall) (Fig. 2.2). The greenhouse is equipped with 17 sodium lights, each processing a power rating of 1000 Watts. As the girders underneath the glass roof of the greenhouse cast shadows over plant canopies, they were therefore also introduced into the simulation model.

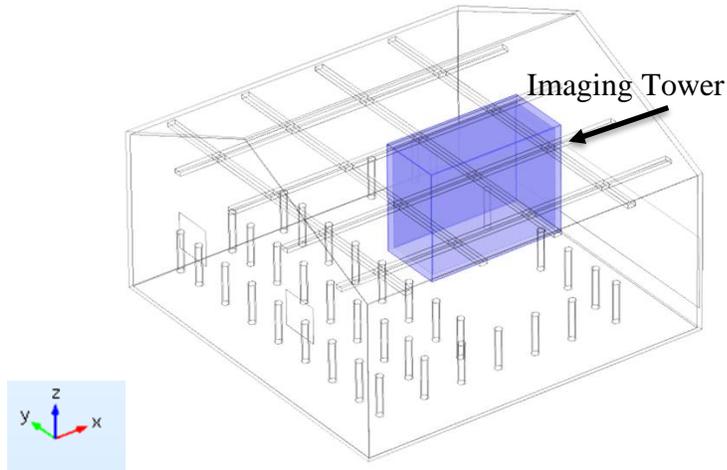


Fig. 2.1. Overall greenhouse geometry and plants layout. 3D coordinates represent x as north, y as west, and z as zenith.



Fig. 2.2. Purdue Lily 13-4 Greenhouse with automated, high-throughput imaging system and a belt conveyor-based plant transportation system.

2.2.2 Model Development

The change in the greenhouse temperature distribution and radiation absorption on the surface of the plant was modeled with representations of heat transfer and radiation energy. The entire model was computed with the explicit finite element method, which leads itself particularly well for the numerical simulation of heat transfer problems (Korioth and Versluis, 1997). More specifically, the radiative heat exchange between the greenhouse's interior surfaces such as the

convective heat exchange within indoor air as well as the heat exhaust from the wet wall cooling system and the sunlight radiation effect correspond well with elements of this method.

The heat transfer between different objects can be expressed by the conduction, convection, and radiation phenomena (Kimball, 1973). Assuming the simulation model contains constant thermal properties, the heat conducted by a solid medium can be described by the classic Fourier's heat conduction (Eq. 2.1) equation (Eck et al., 2016):

$$\rho C_p \frac{\partial T}{\partial t} + \nabla \cdot q = Q \quad (2.1)$$

Where ρ is the density of the solid material (kgm^{-3}), C_p ($\text{Wm}^{-1} \text{K}^{-1}$) represents the specific heat capacity, $\frac{\partial T}{\partial t}$ is time-based temperature derivative (K/s), Q (Wm^{-3}) denotes the amount of heat generated and q is the conduction heat flux (Wm^{-2}) which can be represented by Eq. 2.2:

$$q = -k \nabla T \quad (2.2)$$

Where k ($\text{Wm}^{-1} \cdot \text{K}^{-1}$) is the thermal conductivity and ∇T (Km^{-1}) denotes three-dimensional temperature distribution.

In addition to heat conduction within solids, the surface heat transfer, which is affected by heat convection and radiation, also plays an important role (Baxevanou et al., 2010). As the greenhouse in the study is equipped with an evaporative cooling system, the heat dissipated from all surfaces exposed to the air through forced heat convection can be modeled using the heat transfer equation (Eq. 2.3).

$$Q = A * h(T - T_{\text{ambient}}) \quad (2.3)$$

Where A denotes the surface area (m^2), h represents the heat transfer coefficient ($\text{Wm}^{-2} \text{K}^{-1}$), T is the object temperature (K), and T_{ambient} denotes the ambient air temperature (K).

In the greenhouse system, the radiative exchange between two objects can represent a major factor behind the temperature and plant radiation absorption variation in the greenhouse environment (Eck et al., 2016). A plant's temperature under direct sunlight is significantly higher than it would be elsewhere. However, because solar radiation represents the main source of energy in the plant photosynthesis process, the nonuniform spread of solar radiation has always represented the main part of microclimate impact. The overall radiosity exchange (Eq. 2.4) between two objects (Eck et al., 2016) can be defined as

$$R_{ij} = A_i F_{ij} (J_i - J_j) \quad (2.4)$$

Where R_{ij} is the power transmitted (W) from body i to body j , A_i represents the surface area (m^2) of body i , J_i denotes the total radiative flux (Wm^{-2}) leaving surface i , J_j is the total radiative flux (Wm^{-2}) leaving surface j and F_{ij} is the view factor from body i to body j and can be expressed as (Eq. 2.5)

$$F_{ij} = \frac{\text{Radiation leaving } A_i \text{ and hitting } A_j}{\text{Total radiation leaving } A_i} \quad (2.5)$$

For a certain temperature difference between two surfaces the radiation transmitted from one surface to another can be defined as (Eq. 2.6)

$$R_{ij} = \frac{A_i \varepsilon_i (\sigma T_i^4 - J_i)}{1 - \varepsilon_i} \quad (2.6)$$

Where A_i is the thermal energy (W) leaving surface i , ε_i represents the emissivity (m^2) of surface i , and σ is the Stefan Boltzmann constant with a value of 5.67×10^{-8} ($W m^{-2} K^{-4}$).

This means the radiation flux between two objects can be defined by the Stefan-Boltzmann Law (Eq. 2.7).

$$R_{ij} = A_i \varepsilon_i \sigma (T_i^4 - T_j^4) \quad (2.7)$$

Where T_i represents the temperature (K) of surface i , and T_j denotes the temperature (K) of surface j .

As the problem is governed by the heat transfer equation (Eq. 2.1), the difference between the energy moving in and out of the system therefore determines the final temperature change. By combining all equations and boundary conditions, Eq. 2.1 solving for the overall energy balance can be rewritten as (Eq. 2.8)

$$-k \nabla T = h(T - T_{\text{ambient}}) + \varepsilon_i \sigma (T_i^4 - T_j^4) \quad (2.8)$$

The physical properties of the materials used in this study are summarized in Table 2.1, which is taken from the COMSOL material library.

Table 2.1. Properties of materials used in the greenhouse microclimate simulation model.

Model inputs	Heat Capacity $Jkg^{-1}K^{-1}$	Thermal conductivity $Wm^{-1}K^{-1}$	Surface emissivity
Glass of greenhouse	703	1.38	0.93
Imager tower	1470	0.18	0.6
Plant	3000	5.48	0.85
Roof girders (aluminum)	900	201	0.3
Ground (concrete)	880	1.8	0.94

2.2.3 Computation

The simulation model was programmed in the COMSOL 5.2b (COMSOL Inc., Burlington, MA, USA), which is a cross-platform finite element analysis, solver and Multiphysics simulation software. COMSOL 5.2b includes four radiation modules; this study used the “Heat Transfer with Surface-to-Surface Radiation” module. In the application area, three equations (Eqs. 2.1, 2.2, 2.7) were applied to account for the heat transfer in the greenhouse including heat transfer in solids, heat transfer in fluid, and surface to face radiation (Eck et al., 2016). Referring to Fig. 2.3, the initial boundary for the greenhouse was set to measured ambient air temperature (Frei, 2016). Eq. 2.3 was used to set the flux boundary condition for both the inside and outside of the greenhouse.

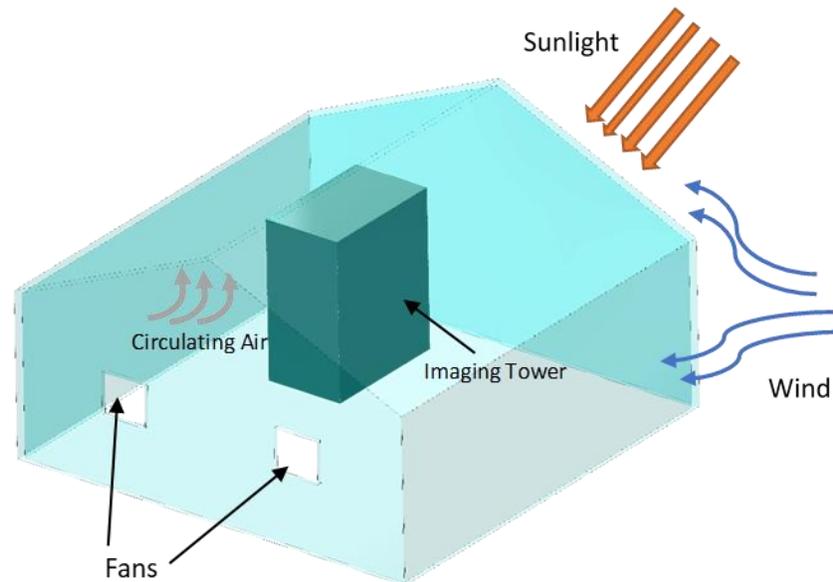


Fig. 2.3. The air flow inside and around house related with temperature and radiation.

The domain was meshed with Lagrange triangular quadratic elements. Triangular elements were selected to allow for local mesh refinement (Fig. 2.4). With integration of elements, COMSOL computed the view factor with the finite element method as the backend. The simulations were all executed on a ThinkPad workstation P300 (Lenovo PC international, Morrisville, Morrisville, NC, USA) equipped with 16-gigabytes (GB) of random-access memory (RAM) and a 3.70 GHz Intel[®] Xeon[™] E1270 processor.

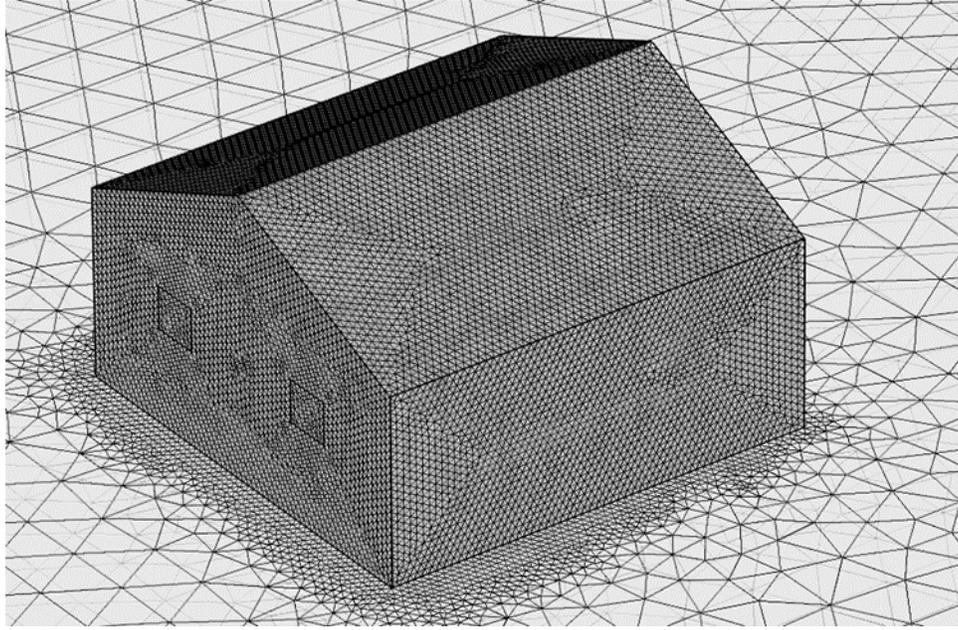


Fig. 2.4. Greenhouse is divided into sub faces by meshing, with facets as triangles.

2.2.4 Optimization for greenhouse conveyor movement based on simulation result

Once the simulation results were obtained, the model data was used to optimize the movement of the conveyor. Initially, the simulation model calculated the temperature and radiation by assuming that all target positions were fixed. However, to incorporate the effect of the conveyor movement, the changing temperature and radiation over the movement were then calculated based on assumption on conveyor running time, break time, speed, and each plant's initial position.

Thirty-eight plant samples were distributed in the greenhouse (Fig. 2.5). Each plant was labeled and its position along the conveyor was marked. For instance, Plant 1 was set at 0 m, Plant 2 was at 0.75 m, Plant 3 was at 1.5 m, and so on. The values of all coordinates were determined by their real positions in the model shown in Fig. 2.5. The conveyor running speed was set to be 12 m/min; the time required to start and stop the conveyor was ignored in this study as they made up a small relative share of the total movement time. Since the initial position (X_0) of each plant, speed of conveyor (v), and the direction of the conveyor were known, the time-based position (X_t) of the plant was calculated by the linear position function (Eq. 2.9).

$$X_t = vt + X_0 \quad (2.9)$$

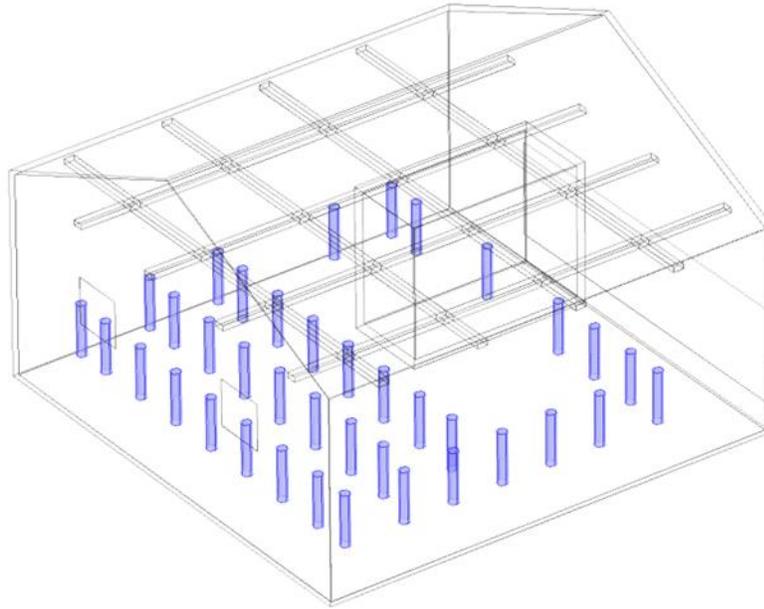


Fig. 2.5. Layout of simulated plants in the greenhouse.

The microclimate information regarding both temperature and radiation varied in the spatio-temporal domain. The expressions of temperature and radiation could be written as temperature (X_t, t) and radiation (X_t, t) , respectively. Once the target plant and running time were determined, the real-time position (X_i) of the plant was calculated using Eq. 2.9. The two plants closest to X_i , for each of 38 fixed positions were marked as X_1 and X_2 . At the same time, the simulated temperature and radiation results were also obtained as $T_1(X_1, t)$, $T_2(X_2, t)$, $R_1(X_1, t)$, and $R_2(X_2, t)$. The real time temperature $T(X_t, t)$ and radiation $R(X_t, t)$ of plant could then be calculated by the linear interpolation method (Eq. 2.10 and 2.11).

$$(X_t, t) = R_1(X_1, t) + (X_t - X_1) * \frac{R_2(X_2, t) - R_1(X_1, t)}{X_2 - X_1} \quad (2.10)$$

$$T(X_t, t) = T_1(X_1, t) + (X_t - X_1) * \frac{T_2(X_2, t) - T_1(X_1, t)}{X_2 - X_1} \quad (2.11)$$

Table 2.2. Summary of running and break time for conveyor movement optimization

Item	1	2	3	4	5	6	7	8
------	---	---	---	---	---	---	---	---

Running time (second)	10	20	40	80	160	320	640	1280
Break time (minute)	0	1	2	4	8	16	32	64

For both the running and break phases, eight different levels of movement frequencies and distances existed. For example, if a combination of 40 s running and 4 min break phase is selected, then the conveyor moves for 40 s followed by a 4 min rest period; the process is then repeated. This resulted in a total of 64 final combinations of the 8 x 8 step levels. For each combination, the accumulated radiation energy received by 38 plants was calculated by summing up $R(Xt, t)$ from Eq. 2.10. Different possible combinations were compared to optimize energy efficiency, as expressed by the lowest coefficient of variation (Eq. 2.12) and running time ratio.

$$\text{Coefficient of variation (CV)} = \frac{\text{Standard Deviation}}{\text{Average}} \times 100 \quad (2.12)$$

2.2.5 Measurements and validation

To validate the simulated results, temperature and radiation data was collected from each pot with Xiaomi flower care sensors (Xiaomi Inc., Beijing, China). With embedded temperature and radiation sensors (Fig. 2.6a), the Xiaomi Flower Care Sensor (Model: HHCCJCY01HHCC) provided real-time temperature and light intensity recordings. As shown in Fig. 2.6b, the sensor was inserted into the soil of the pot, with actual plants when collecting temperature and radiation. To monitor the real microclimates distribution, there were 20 sensors randomly distributed in the greenhouse. All sensors were calibrated before the validation test by recording temperature and radiation measurements under the same environmental condition and adjusting the measurement biases between the sensors. The data was continuously collected for seven days from Aug 09, 2018 to Sep 04, 2018 and logged at 1/60 Hz using a single-board Raspberry pi 3 computer (Raspberry Pi foundation, Cambridge, United Kingdom).



Fig. 2.6. (a). The layout of the Xiaomi Flower Care Sensor including temperature and radiation sensors; (b). The sensor was inserted into the soil of the pot during data collection.

2.3 Results and discussion

By solving the combined heat transfer equation for energy balance, the simulation model developed in Section 2.2 enables a detailed analysis of the temperature and radiation change, and distributions caused by microclimates in the greenhouse. The analysis reveals significant differences between the environmental conditions that exist at various locations within the greenhouse.

2.3.1 Simulated temperature and radiation

The 3-D simulated radiation profiles for the greenhouse illustrates the spatio-temporal variation of the radiation (Fig. 2.7). As the sun rises from the east, Fig. 2.7a demonstrates that the east-facing parts of the greenhouse are much brighter, indicating they receive more radiation energy in the morning. Furthermore, Fig. 2.7a clearly shows a shadow oriented towards the west. At noon, due to the high quantity of solar radiation, the overall color of Fig. 2.7b is brighter. Fig. 2.7c shows the radiation distribution at 6pm (just before sunset) with a shadow oriented towards the east due to the west-oriented sky. At 8pm, at the approach of night, the radiation in the greenhouse has become very weak.

For more specific radiation distribution information, we selected five simulated plants (Fig. 2.8a). These five samples were drawn from different locations in the greenhouse covering, solar radiation levels ranging from low to high. The temporal distribution of the radiation for the selected five plants is plotted in Fig 2.8b, indicating that the plants under shadow received

relatively less radiation energy. For example, at 8am, Plant 1 was located in the northwestern corner in this greenhouse. Due to the imaging tower, it did not receive enough radiation energy (Fig. 2.8b). Conversely, at 6pm, Plant 4 fell under the shadow of the imaging tower. The complex structure of this greenhouse thus leads to nonuniform levels of radiation received by plants.

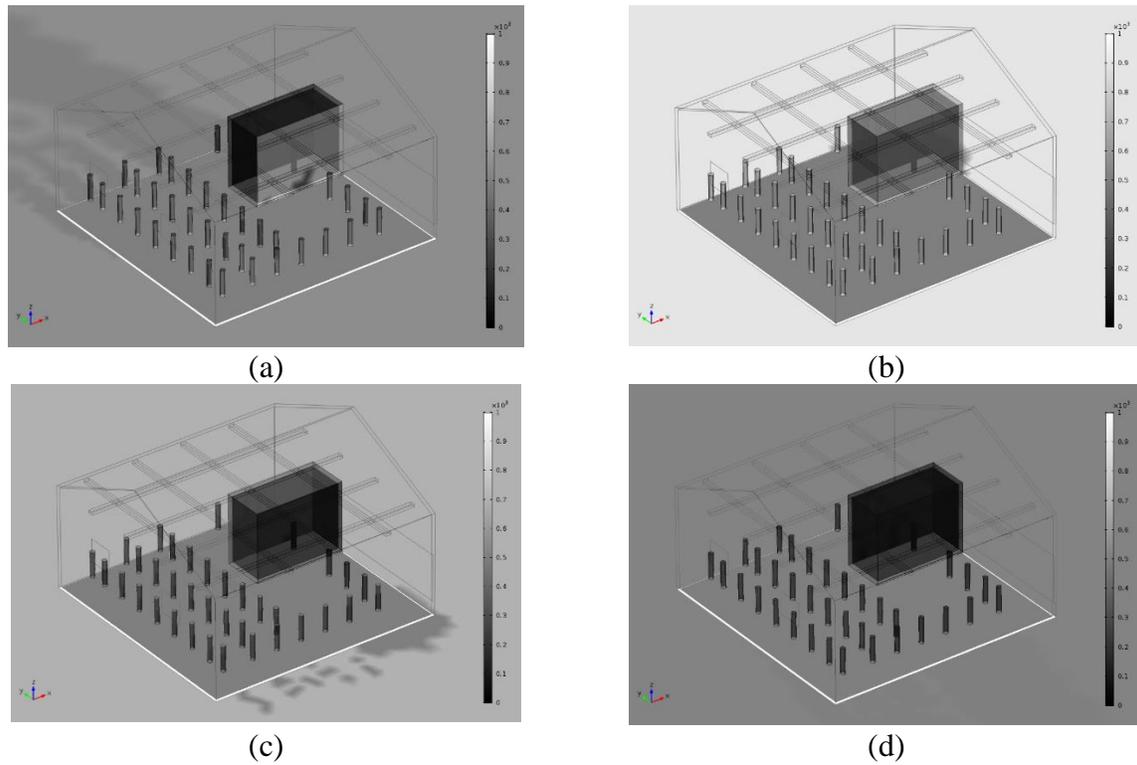


Fig. 2.7. Simulated radiation (W/m^2) at four different time points during a day: (a). 8am; (b). 12pm; (c). 18pm; (d). 20pm. The radiation value is shown in grey-scale color legend.

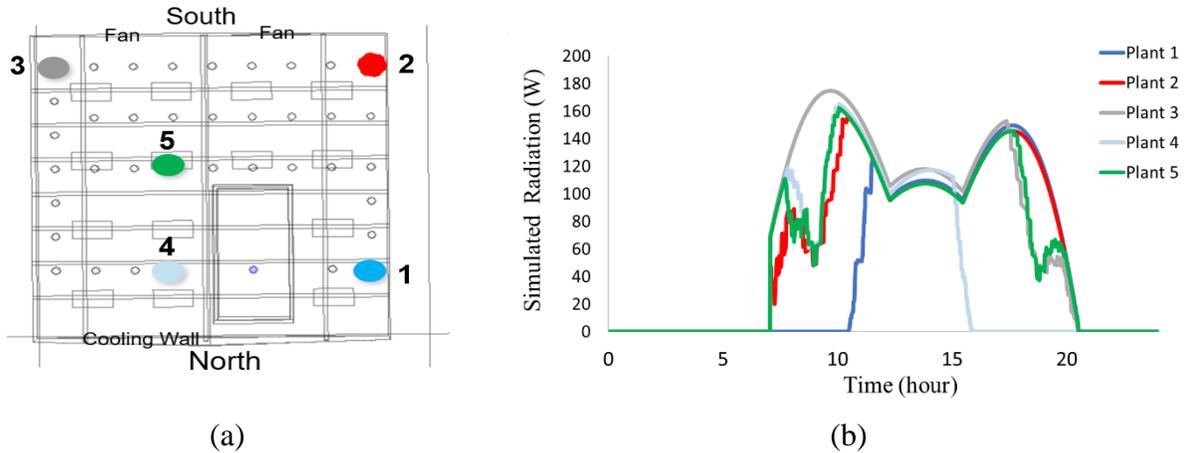


Fig. 2.8. (a) Positions of five samples labeled from top view of greenhouse, (b) simulated radiation (W) of plant with the time (hour). The simulated radiation is calculated for the total integrated plant body surface.

The 3-D simulated temperature profiles for the greenhouse are shown in Fig. 2.9, clearly indicating changes in the temperature changes over both time and space. In the morning, the overall temperature was relatively low (Fig. 2.9a). The east face of the greenhouse demonstrated a higher temperature due to the direct incident solar radiation in the morning. At noon (Fig. 2.9b) and in the afternoon (Fig. 2.9c), the temperature of the greenhouse was high. In some areas, like on the cement floor within the greenhouse, the temperature reached up to 45 °C due to its continuous exposure to the sunlight and the small heat capacity properties intrinsic to cement.

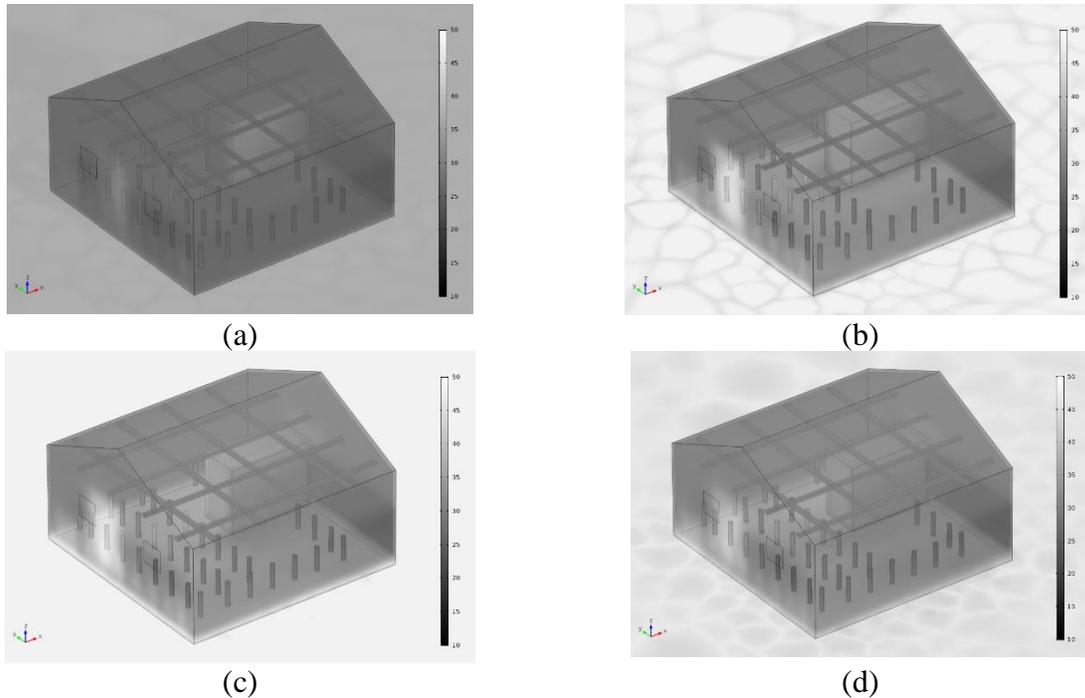


Fig. 2.9. Simulated temperature ($^{\circ}\text{C}$) at four different time points during the day, (a) 8am, (b) 12pm, (c) 18pm and (d) 20pm. The temperature value is shown in grey-scale color legend.

Six simulated plants were chosen as an example from different locations covering temperatures ranging from low to high. As shown in Fig. 2.10a, six plants were chosen, whose temperature changes over time were extracted and plotted in Fig. 2.10b. In this simulation, the plant's temperature was heavily affected by position, even within the same greenhouse. As shown in Table 2.3, plants displayed lower temperatures when they were exposed to the cooling wall from the north side of the greenhouse (Plant 3, 5 and 6). Plants 2 and 4, which were blocked from the cooling wall by the imaging tower, obviously demonstrated higher temperatures. All the plants in Fig. 2.10b also showed a change in temperature profiles over time. They achieved their highest temperature points between 3 to 4 pm in the afternoon, as September is still in the summer and the hottest part of the day in summer occurs typically between 3:00 p.m. and 4:30 p.m. local time, depending on cloud cover and wind speed. The complex structure of this greenhouse led to nonuniform temperature distribution. This simulation did not consider plant evapotranspiration because the impact from the plant was assumed to be minor compared to the other factors such as sun radiation and ambient temperature. Meanwhile, the thermal properties of "wood" as parameter were used as inputs for the plant in the simulation model (taken from the COMSOL 5.2b material library), explaining why the simulated temperature reached as high as 45°C . We intended to

quantity the variation from the heterogenous environment in the greenhouse via this simulation, assuming its results successfully represented the effects of the environmental variances.

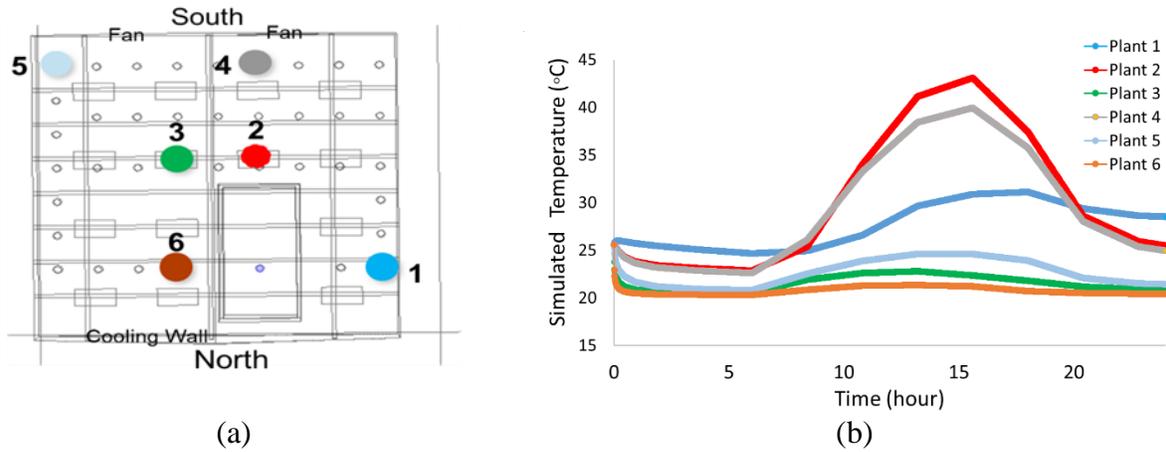


Fig. 2.10. (a). Six sample positions labeled from top view of greenhouse; (b). Simulated temperature of plant with the time (hour).

Table 2.3. The temperature ranges of the plants from the simulation model.

Plants	Range of simulated temperature (°C) within a day
Plant 1	24.65 – 31.10
Plant 2	22.85 – 43.08
Plant 3	20.55 – 25.56
Plant 4	22.59 – 39.94
Plant 5	20.80 – 25.59
Plant 6	20.30 – 25.59

2.3.2 Model validation

The greenhouse model was simulated using the assumptions of a sunny day and steady wind speed. The inputs of the simulation model for the validation test was illustrated in Table 2.4. For validation purposes, we chose seven days from Sep 03, 2018 for the ground truth data collection, as the weather reflected these environmental assumptions. As described in the methods, the data was logged at 1/60 Hz by Xiaomi Flower Care Sensors resulting in one measurement per minute. In total, 1440 instantaneous data points were collected per day, accumulated from midnight-midnight of each day.

Table 2.4. The inputs of the simulation model for the validation test.

Inputs for the simulation model	Values
Ambient temperature	Hourly data from the local weather reports
Sun radiation at solar noon	1000W/m ²
Fan ventilation speed	5m/s
Cooling wall temperature	20°C
Atmosphere pressure	1 atm
Humidity	50%

For radiation, based on the collected data and simulated result, we calculated the diurnal accumulated radiation value and compared the simulated versus the measured radiance. From this point of view, the “accumulated radiation” and “accumulated temperature” represented suitable choices, that were also easily compared. As shown in Fig. 2.11, 20 samples were employed that covered radiation ranges ranging from low to high. The x -axis represents the simulated radiation from software COMSOL 5.2b, in which the original unit is J. Therefore, simulated accumulated radiation was measured in J/day. Conversely, the y -axis denotes the radiation measured by Xiaomi Flower Care sensor. The results of the linear regression demonstrated a significant correlation between measured and simulated radiation, with a R^2 of 0.9053.

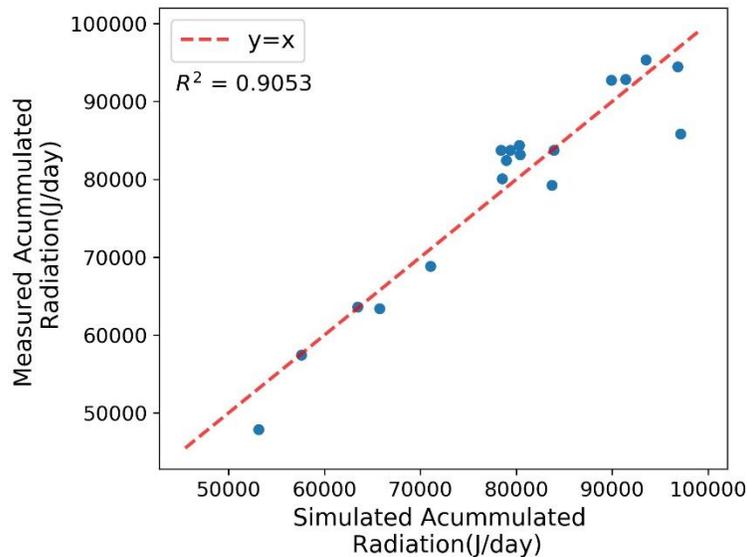


Fig. 2.11. Accumulated radiation results: actual versus simulated measurements.

The “accumulated temperature” represents the “integration of temperature over time”. This study used this variable to evaluate the heat treatment modeled after “the accumulate of radiation”.

This simulation model mainly served as a way to predict the heterogenous environment (temperature and radiation) in the greenhouse. Therefore, we measured “accumulated temperature”, or the accumulated impact from the temperature on the plants, by summing the measured temperature data points. The accumulated heat based on both the prediction as well as ground truth measurement from the 20 samples was determined for the same time period. The R^2 between the measured and predicted data was 0.8819 (Fig. 2.12). This result shows that the accumulated heat based on temperature can be quantified using the simulation model, including the variation in different spots.

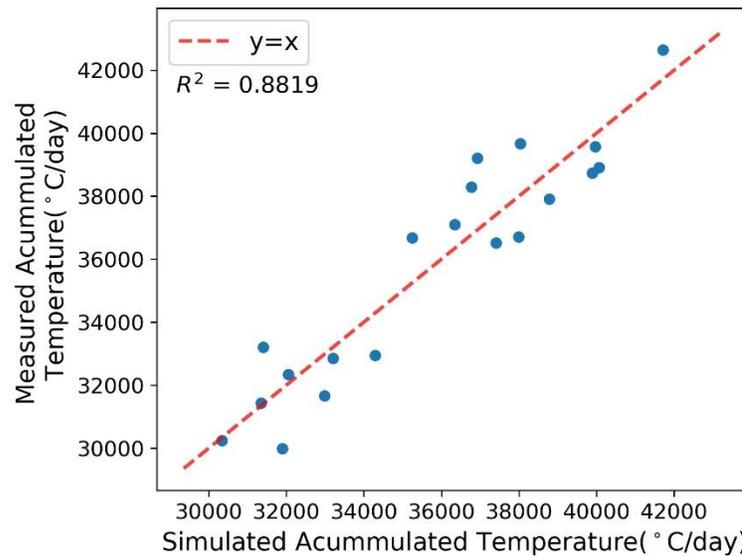


Fig. 2.12. Accumulated temperature results: actual versus simulated measurements.

2.3.3 Use model to drive movement of the conveyor

The simulation result data was used to compare 64 different combinations of running and break time settings. The goal was to find “smart” conveyor “run” and “stop” time intervals, that could mostly eliminate the microclimate effect while also significantly reducing the conveyor running cost. Fig. 2.13 shows the microclimate variance in each case as the coefficient of variation (CV) of accumulated radiation energy. The table is colored as a gray scale heatmap to clearly illustrate the different CVs for each of the 64 setting combinations. Generally, plants in the greenhouse receive more uniform radiation energy with a longer running time and shorter break time; for the same break time, CV decreases when running time increases, since a longer moving distance should help eliminate the microclimate effect; conversely, under the condition of the same

running time, CVs increase when break time do as well, as lower moving frequency result in a larger impact from microclimates on the plants. It should be noted that 640s running time worked fairly well regardless of break time. It might be because 640s of running time nearly aligned with some near symmetrical movement of pots; a full loop of the conveyor movement took around 640s.

To indicate the uniformity of radiations, we assumed CVs larger than 1 to be high variance whereas those with CVs lower than 1 were considered low variance. After filtering all low-variance cases, the optimal decision was selected by choosing the combination that expended the least energy. For example, shifting from the original setting of continuous operation to a new setting of “10s on, 8 min off” would result in the use of 2% of the original energy (98% savings) while still keeping the CV low at 0.91. This choice would thus allow the avoidance of unnecessary conveyor movement.

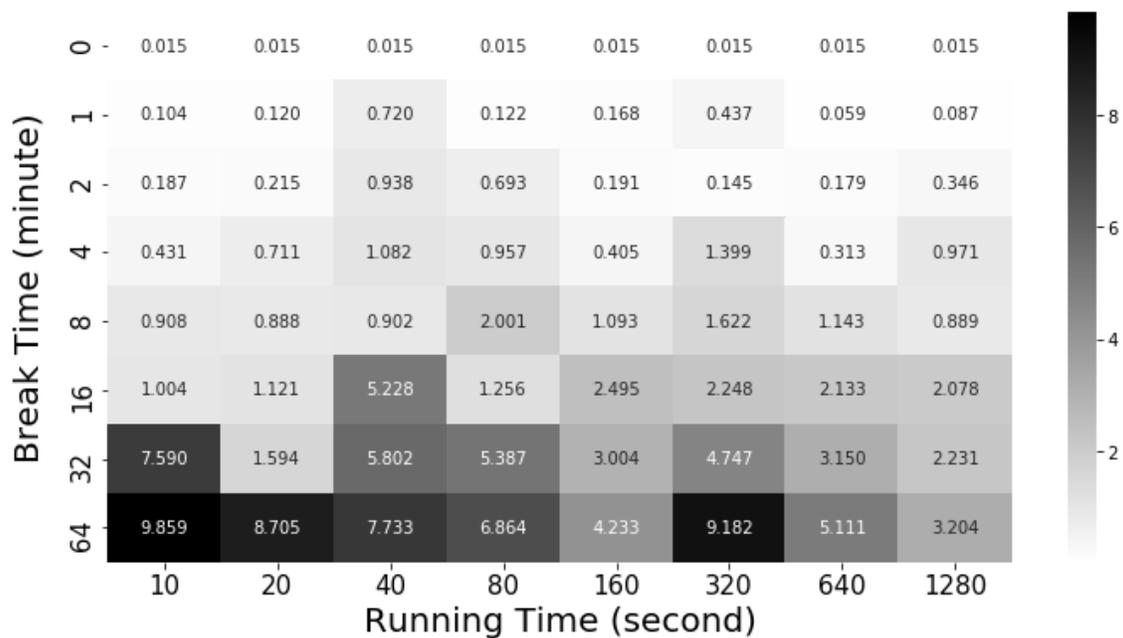


Fig. 2.13. Coefficient of variation (CV) of the simulated radiation received by plants from different running and break time combinations for conveyor movement optimization

2.4 Conclusions

A numerical greenhouse temperature and radiation simulation model was developed for the purpose of greenhouse microclimate control. The model successfully predicted temperature and radiation distributions over time and space in Purdue University’s Lily 13-4 greenhouse. By

applying the simulation results in the optimization of conveyor system movement in Lilly 13-4, this study showed new insights into how to quantitatively use simulation data to minimize the greenhouse microclimates. The proposed model was validated by comparing the simulated radiation and heat result with ground truth measurements from temperature and light sensors distributed in the greenhouse. The squared correlation coefficient between prediction and measurement was 0.9053 for accumulated light radiation and 0.8819 for accumulated heat.

As an application, the model's results demonstrated its potential for optimizing greenhouse pot movements by choosing the lowest cost with the same ability to mitigate the microclimate problem. It will greatly benefit the greenhouses that currently possess automated conveyor systems, which have been developed and applied in many plant breeding studies (Ge et al., 2016; Gehan et al., 2017; Golzarian et al., 2011). In Purdue's Lilly 13-4 greenhouse case, shifting from the original setting with continuous running to the optimal setting "10s on and 8 min off" suggested by the model could save 98% of the original energy required for the conveyor system. In situations where pot movements are impractical, this modeling method can still aid researchers in quantitatively estimating the impact of microclimates, to distinguish signal from noise when analyzing the plant growth measurement data.

CHAPTER 3. REMOVAL OF GREENHOUSE MICROCLIMATE HETEROGENEITY WITH A CONVEYOR SYSTEM FOR INDOOR PHENOTYPING

Greenhouse facilities represent a prevalent tool in contemporary crop research, attractive for their ability to achieve controlled plant growth in all seasons and with more uniform environments than field experiments. However, the environment within the same greenhouse retains heterogeneous; such divergent regions are referred to as microclimates. The nonuniform distributions of temperature and radiation reflected by microclimates create problems that impact the growth, production, and quality of the crops. To address these issues, this study equipped a traditional greenhouse with a conveyor system controlled by a computer algorithm to expose plants to more equal amounts of heat and radiation. Using a simulation analysis of the greenhouse microclimates, the algorithm minimized the conveyor motion while exposing all plants to a nearly equivalent accumulation of heat and radiation. In a traditional greenhouse as well as another outfitted with the novel conveyor system, one maize genotype was replicated 50 times in each facility and subjected to simultaneous drought and well-watered treatments. The plants were scanned with VNIR hyperspectral cameras from both the top and side views. The results showed that the measured plant phenotyping features within both treatment groups demonstrated a significantly lower variance in the experiment performed in the greenhouse equipped with the conveyor system ($p < .05$). The findings also indicated the greenhouse equipped with the conveyor system displayed a better experimental performance than the traditional one; the well-watered and drought stressed plants showed statistically stronger differences ($p < .05$) in projected leaf area, Relative Water Content (RWC), Soil Plant Analysis Development (SPAD), as well as 21 other commonly used plant indices. These findings reveal that the novel conveyor greenhouse phenotyping mode demonstrated a better performance with lower plant variances compared to the conventional greenhouses.

3.1 Introduction

Modern plant breeding efforts depend on a combination of genotypic and phenotypic data. Developments in sequencing technology have reduced the time and cost required to generate large amounts of genotypic data. However, the volume of plant phenotypic data has not increased at the

same rate due to the high cost of labor associated with collecting it, thus becoming the limiting factor in genetic improvement efforts (Fiorani and Schurr, 2013; Houle et al., 2010; Ribaut and Ragot, 2019). Meanwhile, conventional field phenotyping is labor-intensive and time-consuming, and therefore automated greenhouse platforms demonstrate the potential to increase the accuracy and quality of plant phenotyping data (Li et al., 2014; Peirone et al., 2018).

Greenhouse facilities are more attractive than ever as a crop research method due to the greater control they possess over the generation of more uniform environments and therefore the success of controlled plant growth compared to field tests (Vásquez et al., 2015). Equipped with advanced imaging systems for high-throughput plant phenotyping, greenhouses have been used to assess many traits. Greenhouses are often used to appraise abiotic stresses such as drought (Causse et al., 2018) and salt tolerance (Lee et al., 2004), biotic stresses such as disease resistance (Miklas et al., 2019; Stone et al., 2010), qualitative traits such as seed dormancy (Torada et al., 2005), and biological processes such as photosynthesis (Herve et al., 2001). Greenhouses have also been used to study advancements in crop improvement techniques, including varietal selection methods (Hickey et al., 2017; Jannink et al., 2010). Moreover, greenhouses are employed for trait introgression in conjunction with marker assisted selection (MAS) (Prasad et al., 2006) in place of off-season nurseries (Bassi et al., 2015), rapid-generation advancement (Tanaka et al., 2016), and early-generation selections (Glenn et al., 2017). However, the within-greenhouse environment, often referred to as a “microclimate”, remains heterogeneous (Alain, 1989). The nonuniform distributions of temperature, relative humidity, carbon dioxide concentration, and radiation caused by the differences between microclimates create disparate impacts on the growth, production, and quality of the crops.

Modern greenhouses are equipped with ambient control systems to optimize plants’ growing environments. However, these control systems remain unsatisfactory. For example, the temperature in greenhouse air is monitored at a singular point that may not be representative of all points within the environment (Körner et al., 2007). Moreover, because heat transfer disperses within environments at variable rates, crop temperature frequently does not equal greenhouse air temperature. To enable uniform growth conditions for the plants, scientists have designed different greenhouses to diminish the microclimate impacts. One study examined the influence of a whitened greenhouse roof on microclimates during the summer in a greenhouse located on the eastern coast of Greece (Baille et al., 2001). It successfully reduced the nonuniformity from the

outside environment, but failed to account for the internal heterogenous environmental impacts within greenhouses on the plants.

Brien (2013) described a greenhouse containing plants on conveyor systems to account for variations in the microclimate. The strategy of cycling plants through greenhouse locations relies on the assumption that plants experience uniform growth conditions when spending an equal amount of time in each microclimate (Brien et al., 2013). However, this shuffle of plants was done at the lane level. Groups of plants sitting on the same lane were moved around together, but the microclimate impacts within the lane remained. For example, plants in the middle of the lane are higher than the plants at the edge. Similar automatic greenhouses with conveyor systems have been developed and applied in many plant breeding studies (Ge et al., 2016; Gehan et al., 2017; Goltzarian et al., 2011). No greenhouse facility has continuously moved the plants on an individual plant level to different locations to address the microclimate issues during their entire period of growth.

Recent advances in conveyor-based automatic phenotyping greenhouses enable relocating the plants during the indoor experiment. In these facilities, the conveyor is used for two purposes simultaneously: (1) shuffling plants to remove microclimate effects and (2) transferring plants to the imager. In all the previously existing facilities of this type, plants are moved with simple empirical operation scenarios (Ge et al., 2016; Gehan et al., 2017; Goltzarian et al., 2011). It has rarely been explored to optimize the conveyor movement for efficiently relocating plants with minimum conveyor movement cost. This chapter introduces the development of a new phenotyping greenhouse at Purdue University with a smartly controlled conveyor system to shuffle the plants for equal exposure to temperature and radiation conditions while minimizing the cost of conveyor movement. There were three major objectives of this work:

1. Design and construct a greenhouse with an automated, high-throughput imaging system and automatic conveyor system.
2. Apply environmental modeling approach to optimize the conveyor movement to equally expose the plants to heat and radiation conditions while minimizing the cost of movement.
3. Evaluate the performance of optimized plants movement solution with a comparison plant assay test.

3.2 Materials and methods

3.2.1 High throughput imaging greenhouse with conveyor system

To control the microclimate issues, a conveyor system was constructed in a high-throughput imaging greenhouse at Purdue University shown in Fig. 3.1. The facility consists of two major parts: (a) the conveyor system and (b) the hyperspectral imaging system. The conveyor system addresses issues of microclimate variation. Plants are kept on the conveyor system for the entire growing cycle, which should, in theory, expose them to similar amounts of heat and radiation. The hyperspectral imaging system (Fig. 3.2) includes an imaging tower with Middleton Spectral Vision MSV 500 cameras that accommodates plants up to 1.5m tall. These cameras are push-broom sensors, which scan top-view and side-view hyperspectral images from 370 to 1030 nanometers with a spectral resolution of 1.3 nanometers. Inside the imaging tower, eight studio halogen lamps provide illumination. The imaging tower automatically opens and closes its doors to accept sample plants and exclude ambient light. The imaging time ranges from approximately 1-2 mins depending on plant size. The plants are automatically rotated on the imaging platform, so the widest plane of the plant faces the side-view camera. The whole system is fully automated in regard to plant imaging, watering, and fertilizing.



Fig. 3.1. Purdue Lily Greenhouse with an automated, high-throughput imaging system and a belt-conveyor-based plant transportation system.

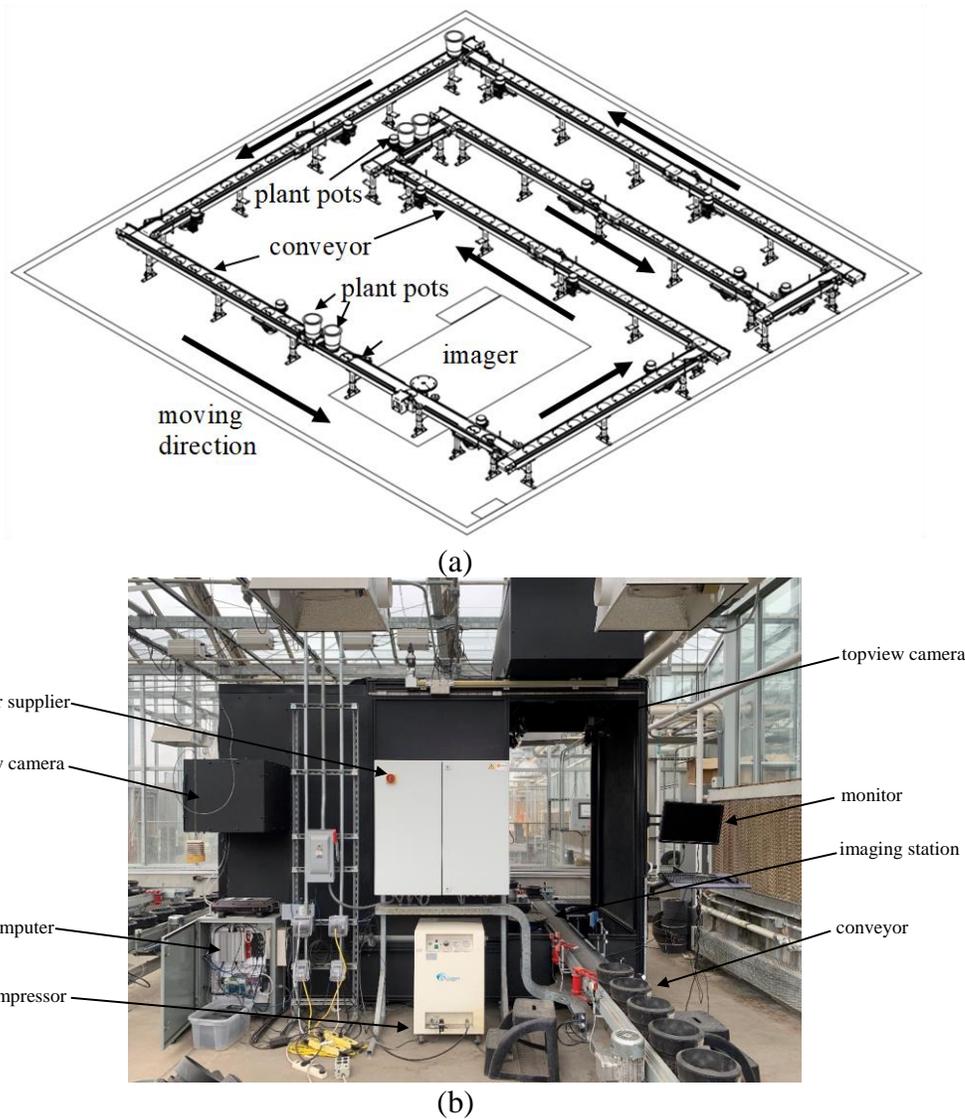


Fig. 3.2. High-throughput hyperspectral imaging system with belt conveyor system: (a). The layout of the conveyor system and the direction of plant movement; (b). A high-throughput hyperspectral imager tower.

3.2.2 Microclimate simulation model to optimize conveyor movement

To optimize the environmental distribution on plants while minimizing the conveyor operation cost, a numerical greenhouse temperature and radiation simulation model was developed for greenhouse microclimate control (Ma et al., 2019b). With an explicit finite element method (Korioth and Versluis, 1997), the radiative heat exchange between greenhouse interior surfaces (such as convective heat exchange within indoor air, heat exhaust resulting from the wet wall

cooling system, and the effect of sunlight radiation) can be readily associated with its corresponding elements. The change in the greenhouse temperature distribution and radiation absorption on the surface of the plant can be modeled with equations for heat transfer and radiation energy (Fig. 3.3).

$$-k \nabla T = h(T - T_{ambient}) + \varepsilon_i \sigma (T_i^4 - T_j^4) \quad (3.1)$$

Where T_i is the temperature (K) of surface i , T_j denotes the temperature (K) of surface j , ε_i represents the emissivity (m^2) of surface i , and σ is the Stefan Boltzmann constant with a value of $5.67 \times 10^{-8} \text{ (W m}^{-2} \text{ K}^{-4}\text{)}$.

By solving the overall energy balance in the greenhouse (Eq. 3.1), the simulation was explored to find the optimized conveyor “run” and “break” time intervals, so that the microclimate effect is still mostly eliminated while the conveyor running cost is significantly reduced.

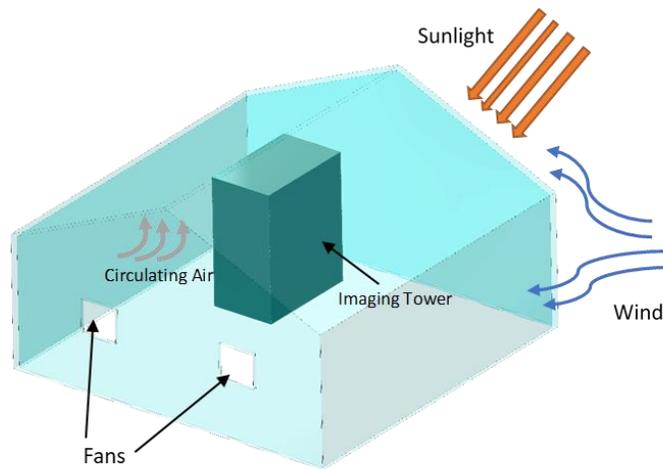


Fig. 3.3. : Overall greenhouse geometry: the air flow inside and around house related with temperature and radiation.

The numerical simulation results from the greenhouse microclimate control model determined the optimal pot movement frequency and distance (Ma et al., 2019b). With appropriate assumptions, the model output suggests shifting from the original, continuously running setting to the optimal setting of “30s running time and 5min break time”, which implies a great deal of energy from the conveyor system can be saved. The model’s suggested shuffling pattern can remove most microclimate variance while reducing huge efforts (~90%) to shuffle plants compared with non-

stop movement. Though the simulation result results from a sample calculation, it provides a preliminary suggestion about how to move the plant on the conveyor system.

3.2.3 Comparison experiment design

To test if the conveyor system could eliminate the microclimate and enable a more uniform plant growth environment, a comparison experiment occurred at Purdue University (latitude 40.4259° N, longitude 86.9081° W) in the summer of 2017. Maize hybrid B73xMo17 was grown in a Lily greenhouse with a traditional setup and in a room that was outfitted with a high-throughput imager with conveyor system (Fig. 3.4). In the comparison experiment, each room housed 100 pots seeded with a custom soil media mix (Fig. 3.4). The custom soil media mix was composed of equal amount of topsoil, sand, and Turface Athletics MVP. Topsoil and sand were combined to provide low water holding capacities and low nutrient content to the substrate. The Turface is a calcined, non-swelling illite clay with 60% minimum amorphous silica that also includes 5% or less of iron oxide, aluminum oxide, calcium oxide, magnesium oxide, potassium oxide, sodium oxide, and titanium oxide with strong cation exchange capacity. In the traditional greenhouse, the pots were arranged in three separate rows with 33-34 replicates each running in a north to south orientation. These plants were allocated in a randomized block design (Sharma et al., 1999). Each pot was irrigated with 200ml of water every morning. Inside the conveyor system, the plants were grown on the conveyor, which ran for 30s, followed by a five-minute break. The automatic water station on the conveyor distributed 200ml water to each plant every morning. The air temperature of both greenhouses was set to remain within 23-29°C, and supplemental lighting was on 12 hours a day.

Beginning at plant stage V6, two water treatments were given in both greenhouses: 50 pots of maize were kept well-watered, while the other 50 pots of maize were placed under drought stress conditions. Water treatments were applied for three days before the final sampling (imaging collection and ground truth measurements), at which point the drought stress response became apparent.

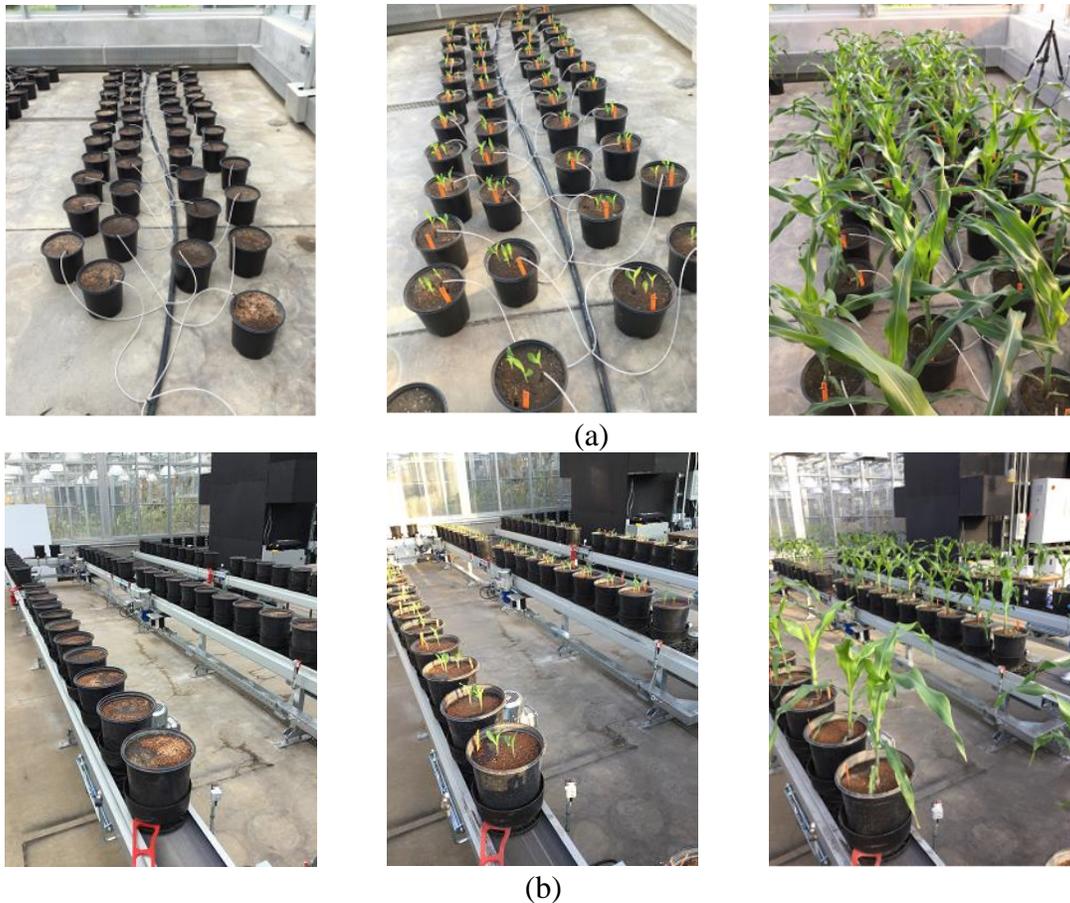


Fig. 3.4. The growth of maize plants in two greenhouses at different leaf stages: (a). Traditional greenhouse; (b). Conveyor greenhouse.

3.2.4 Plant imaging collection and analysis

All the hyperspectral images of the plants were collected with the hyperspectral imager in the conveyor greenhouse. Once collected, the hyperspectral images were processed using a segmentation procedure with convolution methodology (Zhang et al., 2019a). A vector of sequential integers ranging from -20 to 20 was multiplied by the reflectance intensity vector from the red-edge region ($680\text{--}720\text{ nm}$). By choosing threshold 7 as the boundary between plant tissue and the background, the maize was successfully segmented. Fig. 3.5 illustrates the top-view of two maize plants that each experienced three days of one of the two the water treatments, whose results were obviously reflected by their morphology. The well-watered plant in Fig. 3.5a is clearly larger with flat leaves, whereas the drought-stressed plant in Fig. 3.5b has withered, curling leaves.

Using hyperspectral spectra, 21 different plant indices were calculated including NDVI, Carte1, Datt1, GNDVI, MCARI, MTVI1, NDCI, NDVI705, NVI, OSAVI, PRI, RDVI, REP,

RI1dB, SPVI, SPVI2, SR_675_700, TCARI, TVI, VOG1 and mNDVI705 (Liang et al., 2015). Using the segmented binary images, morphological features were also obtained, including projected leaf area, perimeter, major axis length, minor axis length, and eccentricity.

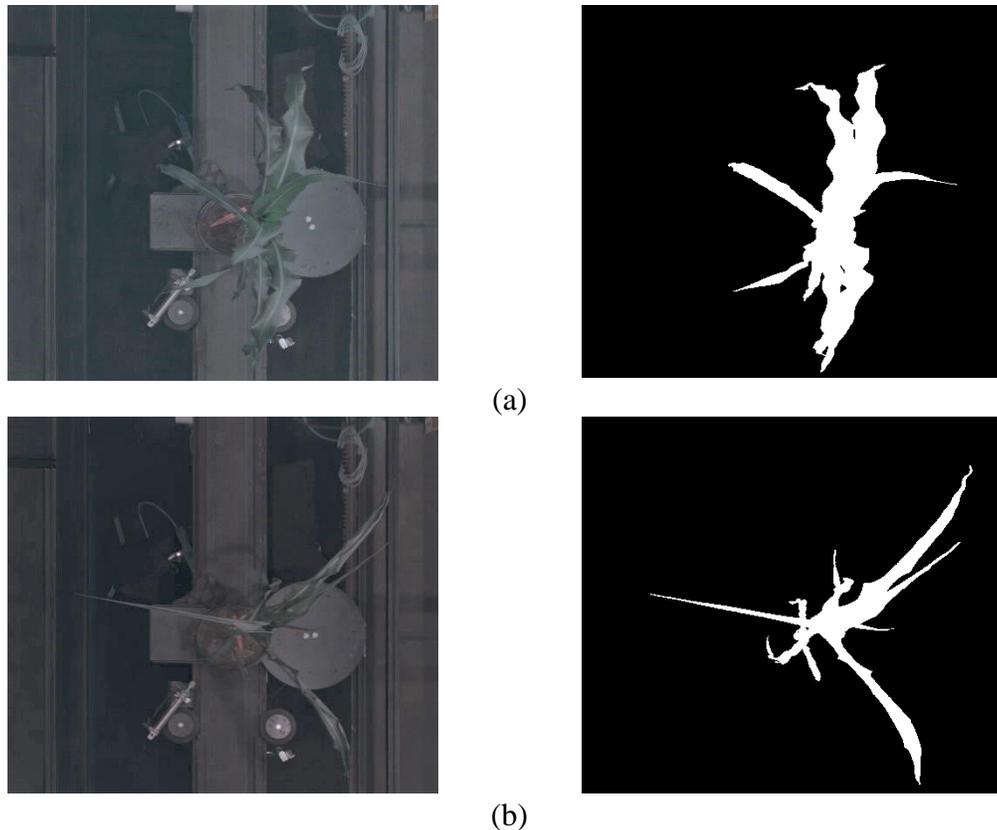


Fig. 3.5. The top-view images of maize in the hyperspectral imaging tower and the segmented binary images: (a). Well-watered; (b). Drought.

3.2.5 Ground truth measurements

Ground truth measurements, including SPAD and RWC of plants, were also collected. All plants were sampled for three SPAD measurements on the top-collared leaf by using the SPAD-502Plus meter (Konica Minolta Sensing Americas, Inc., USA); the average of the three provided the final measurement. The plant was then imaged in the hyperspectral imaging tower including both top-view and side-view camera angles. After imaging, a small section (2.5 cm x 5.0 cm) of the top-collared leaf was taken to measure relative water content (Turner, 1981). On the day of sampling, the fresh weight of the leaf sample was measured, and the leaf tissue was added to vials full of deionized water to obtain turgid weights. These samples were then dried over a three-day

period in a 90 °F dryer; their dry weights were then measured and their RWC computed using the Eq. 3.2:

$$RWC = [(FW - DW) / (TW - DW)] * 100\%. \quad (3.2)$$

Where *FW* denotes fresh weight, *DW* represents dry weight, and *TW* is the turgid weight.

3.2.6 Statistical analysis

A two-tailed *t-test* was conducted to evaluate the differences in projected leaf area, RWC, SPAD measurements, and results from 21 other plant indices between the well-watered and the drought groups in both greenhouses. As performed in *P-Value Precision and Reproducibility* (Boos and Stefanski, 2011), the $-\log_{10}(p\text{-value})$ generated from comparing drought-stressed and well-watered plant indices within experiment were used to compare against the experiment performed in their greenhouse counterpart.

Levene's test was used to analyze the homogeneity of variance across groups to determine whether variability significantly changed between treatment groups. Levene's Test for Equality of Variances is commonly employed to measure shared variance between populations (Schultz, 1985). Levene's test was performed using the statistical programming language R (R Core Team, 2016) with the "cars" package.

3.2.7 Software and computation

In this study, the greenhouse simulation model was programmed in the COMSOL 5.2b (COMSOL Inc., Burlington, MA, USA), a cross-platform finite element analysis solver and multiphysics simulation software. All the simulations were executed on a ThinkPad workstation P300 (Lenovo PC international, Morrisville, Morrisville, NC, USA) equipped with 16-gigabytes (GB) of random-access memory (RAM) and a 3.70 GHz Intel® Xeon™ E1270 processor. In addition, Matlab R2016a software (The MathWorks Inc., Natick, MA, USA) was used to develop the image processing algorithms. The vegetative indices were computed in the Python version 3.7.2 software environment (Python Software Foundation, 2018). All the Matlab and Python computations were run on a HP 17 G3 Mobile Workstation (Hewlett-Packard, Palo Alto, CA, USA) equipped with 64-gigabytes (GB) of random-access memory (RAM) and a 2.70 GHz Intel® Core™ i7-6820HQ processor.

3.3 Results and discussions

3.3.1 Projected Leaf Area, RWC and SPAD measurements

This study first utilized pixel-based projected leaf area and ground truth measurements including RWC and SPAD as the major indicators. For example, Fig. 3.6 illustrates a straightforward improvement of the conveyor greenhouse on the tradition one: the NDVI measurements of plants within each treatment group showed much better uniformity and lower variance. Furthermore, Table 3.1 also summarizes pixel-based projected leaf area and ground truth measurements. The well-watered and drought-stress treatments evidently succeeded in both greenhouses. Compared with the drought group, the well-watered group showed a larger Projected Leaf Area and higher RWC and SPAD measurements. Meanwhile, the results from the traditional greenhouse demonstrated a higher variance within both treatment groups, with larger standard deviations. This held particularly true for the drought-stressed group, in which this variance was relatively large, which would reduce the signal-to-noise ratio during analysis. The traditional greenhouse environment was heterogeneous, and the non-uniform distributions of temperature and radiation led to disparate impacts on the growth, production, and quality of the plants. Conversely, in the conveyor greenhouse, the plants were relocated to different positions throughout the whole experiment. By exposing plants to more equal environments, both treatment groups demonstrated lower variances.

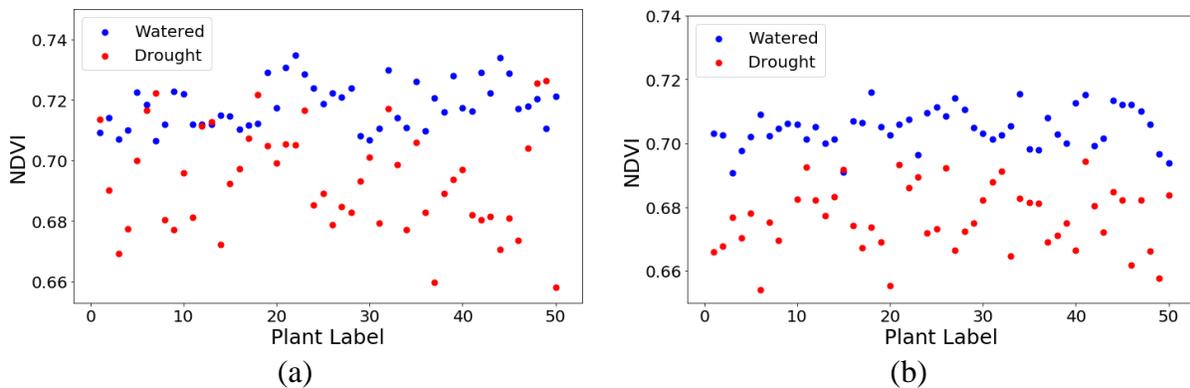


Fig. 3.6. NDVI of plants from different water treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.

Table 3.1. The mean and standard deviation of projected leaf area, RWC and SPAD measurements

Measurements	Traditional greenhouse	Conveyor greenhouse
Projected Leaf Area (Well-watered) (pixel)	$2.55 \times 10^5 \pm 3.89 \times 10^4$	$1.65 \times 10^5 \pm 1.54 \times 10^4$
RWC (Well-watered) (%)	96.15±1.07	94.36±1.70
SPAD (Well-watered)	48.99 ±6.41	42.35 ±3.78
Projected Leaf Area (Drought) (pixel)	$1.82 \times 10^5 \pm 4.57 \times 10^4$	$9.96 \times 10^4 \pm 8.25 \times 10^3$
RWC (Drought) (%)	78.43±10.48	70.22±4.80
SPAD (Drought)	43.29 ± 5.51	33.93±3.73

To further illustrate the experimental improvement from a uniform growth environment, Fig. 3.7, 3.8, and 3.9 illustrate the estimated probability density distribution of Projected Leaf Area, RWC and SPAD measurements from both traditional and conveyor greenhouses, respectively. The estimated probability density distribution was calculated with the kernel density estimate (KDE) method (Kristan et al., 2011). For the traditional greenhouse, even though the results showed that the well-water group resulted in larger Projected Leaf Area and higher RWC and SPAD measurements on average, the measured data of two groups still overlapped significantly. Due to the significant overlap between the two groups (Fig 3.7a, 3.8a, and 3.9a), determining a cutoff line to separate the well-watered and drought groups would prove difficult, leading to severe Type 1 and Type 2 errors in classifying the maize plants. Conversely, the data measured in the conveyor greenhouse displayed lower variance, with a more enhanced signal from water treatments. Clear gaps between existed between the well-watered and drought groups, determining a cutoff threshold to separate the two groups was easy (Fig 3.7b, 3.8b, and 3.9b).

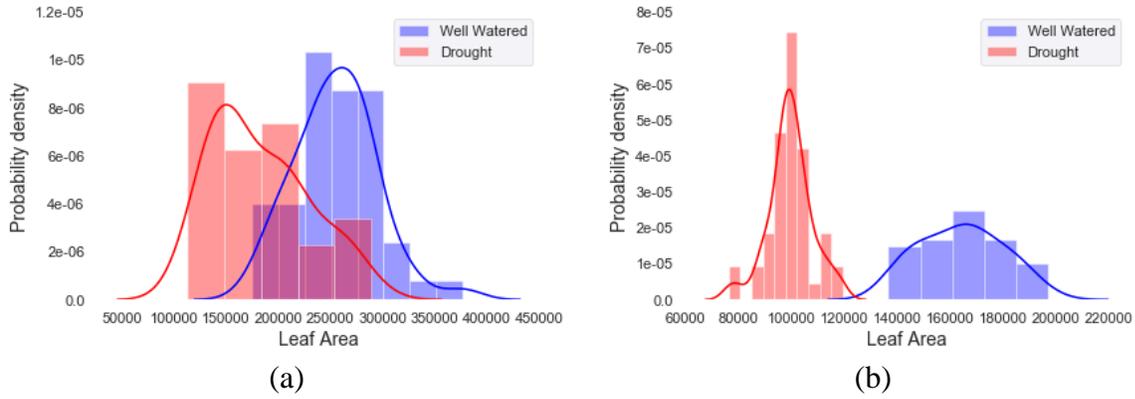


Fig. 3.7. The probability density of Projected Leaf Area of maize from both treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.

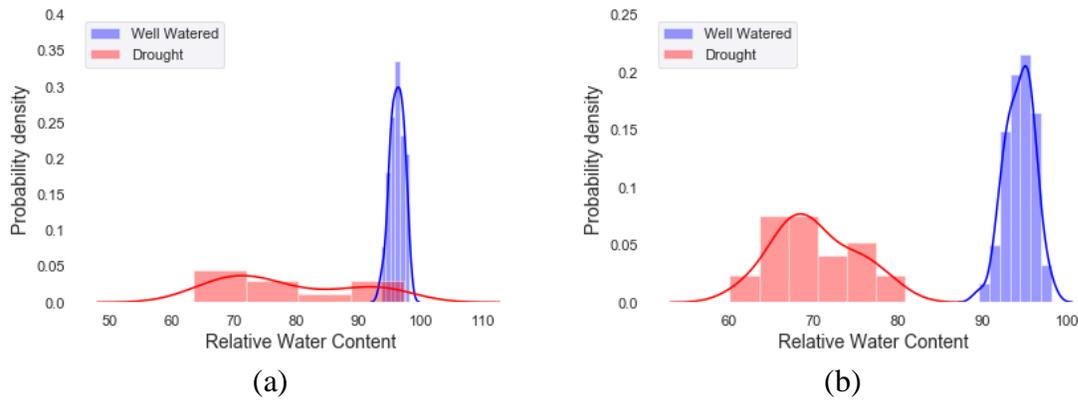


Fig. 3.8. The probability density of measured relative water content (RWC) of maize from both treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.

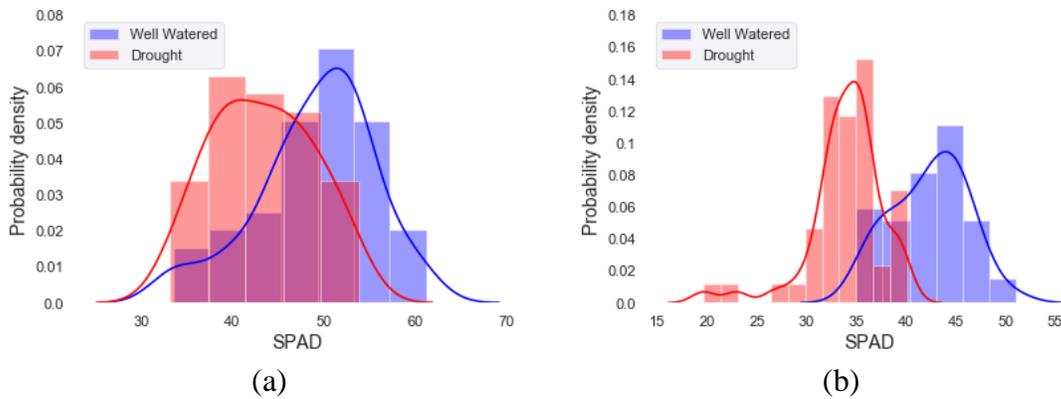


Fig. 3.9. The probability density of SPAD measurements of maize from both treatments: (a). Traditional greenhouse; (b). Conveyor greenhouse.

3.3.2 t-test for projected leaf area, RWC, SPAD measurements and plant indices

To quantify the differences between the well-watered and drought-stressed groups, we conducted two-tailed t -tests and computed $-\log_{10}(p\text{-values})$ for the projected leaf area (pixel), RWC, SPAD measurements, and 21 plant indices, respectively. The statistical results are summarized in Table 3.2. Larger $-\log_{10}(p\text{-values})$ result in smaller p -values, indicating a larger significance in the difference between the two groups. Meanwhile, we selected $-\log(p\text{-value})$ 1.3, which equals a p -value of 0.05, as the critical value. Therefore, any plant indices' p -values that equaled less than 1.3 after the negative log transformation for both traditional and conveyor greenhouses were not considered in the comparison test. This transformation illustrated that differences between the two groups were more significant in the conveyor greenhouse, particularly for the GNDVI, NDCI and PRI measurements. These measurements did not significantly differ between treatments in the traditional greenhouse ($-\log_{10}(p\text{-value}) = 12.27, 2.28$ and 23.76). In summary, compared with the traditional greenhouse, the greenhouse equipped with the conveyor system displayed a stronger experimental performance through lower p -values for predicting measurements like RWC and SPAD and hyperspectral imaging results (projected leaf area and plant indices).

3.3.3 Variance analysis for Projected Leaf Area, RWC, SPAD measurements and plant indices

To determine whether variability significantly changed between treatment groups, Levene's Test was used to analyze the homogeneity of variance across groups. Its results showed that plant variance was significantly higher in the traditional greenhouse compared to the conveyor greenhouse ($p < .05$) within both well-watered and drought treatments.

For the well-watered treatment measurements, projected leaf area (pixel), RWC, SPAD, and all plant indices excluding NDVI705 were found to display statistically different ($p < .05$) variances when the traditional greenhouse and the greenhouse equipped with the conveyor system were compared. For the drought treatment measurements, projected leaf area (pixel), RWC, SPAD, and all plant indices excluding EVI were found to demonstrate statistically different variances ($p < .05$) when comparing the traditional greenhouse and the greenhouse equipped with the conveyor system.

In addition to the homogeneity of variance tests, more detailed results examining the variance of each variable are shown in Table 3.3. For the well-watered group, all variables except

for RWC and NDVI displayed lower variances in the conveyor greenhouse. For the drought group, all the variables demonstrated lower variances in the conveyor greenhouse. Once the variance in plant phenotyping feature was reduced, the two groups were more easily separated for better experimental performance. To conclude, this conveyor greenhouse phenotyping mode showed stronger hyperspectral phenotyping experimental performance with lower plant variances within both the well-watered and drought treatment groups compared to the conventional greenhouses.

Table 3.2. $-\log_{10}$ (P -value) of Projected Leaf Area, RWC and vegetative indices for plant status estimation between the well-watered and drought plant groups in the conveyor and traditional greenhouses.

Phenotyping features	Formulation	References	With conveyor	Traditional
<i>Projected Leaf Area (pixel)</i>	<i>Top-view projected leaf area</i>	(Ge et al., 2016)	44.71	12.60
RWC (%)	$[(FW - DW) / (TW - DW)] * 100$	(Turner, 1981)	53.69	19.42
SPAD	<i>SPAD-502Plus meter</i>	Konica Minolta Sensing Americas, Inc.	18.03	5.08
NDVI	$(R_{800}-R_{650})/(R_{800}+R_{650})$	(Schafleitner et al., 2007)	25.86	14.25
Carte1	R_{695}/R_{420}	(Carter, 1994)	0.33	0.66
Datt1	$(R_{850}-R_{710})/(R_{850}-R_{680})$	(Datt, 1999)	7.57	2.80
EVI	$2.5[(R_{800}-R_{670})/(R_{800}-6R_{670}-7.5R_{475}+1)]$	(Huete et al., 1994)	23.70	7.03
GNDVI	$(R_{750}-R_{550})/(R_{750}+R_{550})$	(Gitelson et al., 1996)	12.27	0.79
MCARI	$[(R_{700}-R_{670})-0.2(R_{750}-R_{550})] (R_{700}/R_{670})$	(Daughtry et al., 2000)	60.50	12.36
MTVI1	$1.2[1.2(R_{800}-R_{550})-2.5(R_{670}-R_{550})]$	(Haboudane, 2004)	35.56	4.19
NDCI	$(R_{762}-R_{527})/(R_{762}+R_{527})$	(Marshak et al., 2000)	2.28	0.35
NDVI₇₀₅	$(R_{750}-R_{705})/(R_{750}+R_{705})$	(Gitelson and Merzlyak, 1994)	1.29	0.24
NVI	$(R_{777}-R_{747})/R_{673}$	(Gupta, 2001)	20.46	9.33
OSAVI	$(1+0.16) (R_{800}-R_{670})/(R_{800}+R_{670}+0.16)$	(Rondeaux et al., 1996)	33.82	12.01
PRI	$(R_{531}-R_{570})/(R_{531}+R_{570})$	(Gamon et al, 1992)	23.76	0.69
RDVI	$(R_{800} - R_{670})/\sqrt{(R_{800} + R_{670})}$	(Roujean and Breon, 1995)	32.60	4.62
REP	$700+40[(R_{670}+R_{780})/2-R_{700}]/(R_{740}-R_{700})$	(Clevers, 1994)	60.66	17.32
R_{1dB}	R_{735}/R_{720}	(Gupta et al., 2003)	36.08	10.72
SPVI	$0.4*3.7(R_{800}-R_{670})-1.2 R_{550}-R_{670} $	(Cho et al., 2011)	16.07	2.24
SR_[675,700]	R_{675}/R_{700}	(Chappelle et al., 1992)	49.79	13.36
TCARI	$3[(R_{700}-R_{670})-0.2(R_{700}-R_{550})] (R_{700}/R_{670})$	(Haboudane, Miller, 2002)	59.69	9.85
TVI	$0.5[120(R_{750}-R_{550})]-2.5(R_{670}-R_{550})$	(Broge and Leblanc, 2001)	18.18	2.39
VOG1	R_{740}/R_{720}	(Vogelmann et al., 1993)	41.24	12.34
mNDVI₇₀₅	$(R_{750}-R_{705})/(R_{750}+R_{705}-2R_{445})$	(Datt, 1999)	36.64	5.18

Table 3.3. Standard deviation of Projected Leaf Area, RWC and vegetative indices for plant status estimation of the well-watered and drought plant groups in the conveyor and traditional greenhouses.

Phenotyping features	Well-watered		Drought	
	With conveyor	Traditional	With conveyor	Traditional
<i>Projected Leaf Area (pixel)</i>	1.54E+04	3.89E+04	8.25E+03	4.57E+04
<i>RWC (%)</i>	1.70E+00	1.07E+00	4.80E00	1.05E+01
<i>SPAD</i>	3.80E+00	6.41E+00	3.73E+00	5.51E+00
<i>NDVI</i>	5.87E-03	7.6E-03	1.26E-02	1.70E-02
<i>Carte1</i>	2.15E-04	1.04E-02	4.71E-04	9.12E-03
<i>Datt1</i>	4.45E-05	4.14E-04	1.01E-04	4.03E-04
<i>EVI</i>	2.49E-01	3.84E-01	1.99E-01	4.28E-01
<i>GNDVI</i>	3.05E-05	1.61E-04	8.18E-05	2.21E-04
<i>MCARI</i>	1.65E-04	2.25E-03	8.57E-05	2.34E-03
<i>MTVI1</i>	5.73E-03	1.97E-01	3.22E-03	1.49E-01
<i>NDCI</i>	3.47E-05	2.03E-04	8.26E-05	2.49E-04
<i>NDVI705</i>	3.11E-05	1.07E-04	9.93E-05	1.47E-04
<i>NVI</i>	5.90E-04	3.25E-03	8.59E-04	5.40E-03
<i>OSAVI</i>	6.34E-05	1.30E-04	1.59E-04	3.40E-04
<i>PRI</i>	3.13E-06	1.35E-04	2.05E-05	1.53E-04
<i>RDVI</i>	3.98E-04	7.20E-03	2.97E-04	6.19E-03
<i>REP</i>	2.06E-02	7.32E-02	7.29E-02	6.73E-01
<i>RI1dB</i>	7.92E-05	3.44E-04	2.41E-04	9.33E-04
<i>SPVI</i>	3.33E-03	1.32E-01	2.27E-03	1.08E-01
<i>SR_675_700</i>	9.43E-05	3.11E-04	2.80E-04	9.53E-04
<i>TCARI</i>	4.17E-04	9.08E-03	2.08E-04	7.91E-03
<i>TVI</i>	4.07E+00	2.09E+02	2.95E+00	1.71E+02
<i>VOG1</i>	1.20E-04	5.24E-04	3.64E-04	1.74E-03
<i>mNDVI705</i>	3.38E-04	2.01E-03	4.46E-04	3.05E-03

3.4 Conclusions

In this study, a novel phenotyping greenhouse facility was designed and built with an automated, high-throughput imaging system and a belt-based plant conveyor. With the conveyor system controlled by the computer algorithm, plants were relocated in the greenhouse throughout the experiment. This system enabled uniform growth conditions because plants spent an equal amount of time in each microclimate, at a low cost of conveyor movement. A comparison test was performed in which 50 maize plants were subjected to drought and well-watered treatments respectively in both a neighboring traditional greenhouse and the novel greenhouse with the conveyor system. The results showed that variance of the phenotyping feature measurements among the plants in the conveyor greenhouse fell significantly (by up to 83% in Projected Leaf Area, for example) when compared with the traditional greenhouse. The findings also indicated the greenhouse equipped with the conveyor system demonstrated a better experimental performance with more significant differences between the two treatments (drought and well-watered) for Projected Leaf Area, RWC, SPAD, and plant indices.

CHAPTER 4. MODELING OF DIURNAL CHANGES IN AERIAL CROP IMAGES

Aerial remote sensing technologies have been widely applied in field crop phenotyping. However, the quality of current remote sensing data suffers from significant diurnal variances. The severity of the diurnal issue has been reported in various plant phenotyping studies over the last four decades, but limited studies exist that model the diurnal changing patterns so people can precisely predict the level of diurnal effects. To comprehensively investigate the diurnal variability, collecting time-series field images with a very high sampling frequency (normally one sample per 15 minutes) was necessary. In 2019, Purdue agricultural engineers deployed their first field visible-to-near-infrared (VNIR) hyperspectral gantry platform, which can repetitively image the same field plots every 2.5 minutes. This gantry and image equipment collected 8631 hyperspectral images of two genotypes of corn plants from vegetative stage V4 to reproductive stage R1 in the 2019 growing season. The analysis of these images showed that although almost all the image-derived phenotyping features vary significantly throughout the day, these changes follow stable patterns. This allows the prediction of the imaging drifts by modeling the changing patterns. This paper reports detailed diurnal changing patterns for several selected plant phenotyping features such as Normalized Difference Vegetation Index (NDVI), Relative Water Content (RWC), and single spectrum bands. For example, NDVI showed a repeatable V-shaped diurnal pattern, which linearly drops by 0.012 per hour before the highest sun angle and increases thereafter by 0.010 per hour. The different diurnal changing patterns in various nitrogen stress treatments, genotypes, and leaf stages were also compared and discussed. With the modeling results of this work, remote sensing users will be able to more precisely estimate the deviation or change in crop feature predictions caused by the specific imaging time of the day. This will aid researchers in deciding upon an acceptable imaging time window during a day. It can also be used to correct/compensate the remote sensing results considering the time effect.

4.1 Introduction

Modern plant breeding efforts depend on phenotypic data to select high-yield and stress-tolerant plants quickly and efficiently (Fiorani and Schurr, 2013; Houle et al., 2010). Ag remote

sensing technologies have been developing rapidly for many years. Currently, field phenotyping activities have been performed with satellites, airborne platforms (manned and unmanned), and ground-based vehicles (Li et al., 2014; Rehman et al., 2019). Various sensors such as RGB (Red-Green-Blue), hyperspectral and thermal cameras are carried by these platforms to take images of the crop field. These technologies have proven effective in various Ag remote sensing projects (Li et al., 2014; Wang et al., 2020). However, the quality of Ag remote sensing data is still limited by various sources of noise, such as changes in daylight, wind speed, temperature, sun angle, etc. (de Souza et al., 2010; Fiorani et al., 2012; Gamon et al., 2015; Maji et al., 2014). Among these noises, diurnal variability represents one of the major factors that causes significant quality issues in Ag remote sensing.

The diurnal impact on plants' reflectance characteristics is a complicated process. It introduces strong noise in plant phenotyping result (Beneduzzi et al., 2017; Maji et al., 2014; Padilla et al., 2019). For example, the reflectance characteristics of the same plants at noon can differ greatly from those in the afternoon. These variations are caused by interactions between camera sensitivity, camera view angle, canopy geometry, solar zenith angle, solar azimuth angle, and shadows (Jackson et al., 1979; Ranson et al., 1985). Meanwhile, the plant itself displays an endogenous sensitivity to the environmental conditions that result from complicated interactions between the genetic backgrounds, the external environments, and the treatments (G(Genotype)*E(Environment)*T(Treatment)). All these affect the final reflectance characteristics of plants throughout the day, causing the diurnal variabilities.

The diurnal variance on phenotyping data has resulted in unignorable impacts in many plant studies. Gardener (Gardner, 1985) stated that this variance represented a major unresolved noise issue in using reflectance measurements for estimating leaf area, plant biomass, or phenology as all were affected by diurnal changes. In the most recent decades, diurnal variability has been documented in various plant phenotyping studies for corn (Oliveira and Scharf, 2014), soybean, wheat (Beneduzzi et al., 2017) with both passive (de Souza et al., 2010; Sticksel et al., 2004) and active sensors (Oliveira and Scharf, 2014). These variances are retained in the captured images, weakening the signal power of the data. Sometimes, the generated variances are even larger than the plant differences caused by biotic or abiotic stresses, which severely limits the accuracy in phenotyping data. However, in most current remote sensing studies, people rarely consider the impacts from diurnal variabilities, which introduces much noise into the results of the final analysis.

For example, the NDVI values demonstrated a difference of over 10% over time on the same plant from the raw remote sensing measurements (Beneduzzi et al., 2017; Sticksel et al., 2004). Similar diurnal variances are also found in many other plant features such as plant temperature, spectra features, chlorophyll content, and so on (Gamon et al., 2015; Zhao et al., 2018). Therefore, the reduction of diurnal variabilities in the data is important for improved remote sensing quality.

With the development of plant phenotyping, researchers are better aware of the diurnal variabilities although limited studies exist on how to deal with diurnal variability in Ag remote sensing. To reduce impacts of diurnal variability, remote sensing technologies, such as unmanned aerial vehicles (UAV), usually have strict rules on the sampling time window and weather condition requirement (Barbedo, 2019; Di Gennaro et al., 2018; Gracia-Romero et al., 2019). According to Bellvert et al. (Bellvert and Girona, 2012), the proper time of the day to acquire thermal and multispectral images is around noon, due to the almost complete absence of shadow effects. Meanwhile, plant physiology changes on a cyclical diurnal basis due to photosynthetic activity and processes dependent on incident solar radiation (Bellvert and Girona, 2012). Consequently, UAV data collections are required to operate during a certain time period to accurately monitor crop physiology (Barbedo, 2019).

Quantitatively modeling the diurnal variability of crop phenotyping features is important for improved Ag remote sensing quality. The combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model (PROSAIL) has been widely used to predict the change in plant canopy spectral reflectance influenced by the changing environmental conditions such as solar angle (Ishihara et al., 2015; Jacquemoud and Baret, 1990). However, the model does not meet the accuracy requirement in plant phenotyping remote sensing. For example, Berger et al. (Berger et al., 2018b) compared PROSAIL's simulation result with the spectra data collected from the field and found that PROSAIL's predicted spectra differed drastically from the field measurements, especially at the green and red edge (~700nm) wavelengths (Berger et al., 2018b). This finding was also confirmed in the field remote sensing experiment in 2019 at Purdue University, in which the PROSAIL model only predicted less than 5% of the diurnal variance of the NDVI observed with a hyperspectral camera. More accurate diurnal variability models are thus still urgently needed.

Modeling diurnal variability in plant phenotyping data is difficult due to the complicated interactions between plants and real-time environment conditions such as cloud coverage, wind,

and so on. Collecting hundreds or thousands of time series images of the same field in order to model the diurnal changes is necessary but has proven difficult with the current airborne platforms. For instance, the UAV platforms rarely imaged the same field more than once a day due to various environmental and logistical limitations (Krishna, 2018). To solve this problem, the Ag engineers at Purdue University deployed a fixed field VNIR hyperspectral gantry platform as a “mock drone” system in Purdue’s research farm in 2019. With this field gantry imaging system, hyperspectral images of the same field can be continuously collected under various weather conditions throughout the daytime on a minute-by-minute schedule. A local weather station and distributed soil sensors in the field provide environmental condition data that synchronize with the images.

This chapter introduced the research work to analyze the diurnal variations in plant phenotypic data by collecting time-series crop images with the unprecedented high imaging frequency of the new field imaging gantry. With the overarching goal of enabling remote-sensing users to understand crop feature prediction changes caused by imaging time of day due to diurnal factors, this work aimed for the following objectives:

1. Collect time-series hyperspectral images of two varieties of corn plants with three nitrogen treatments from V4 to R1 extensively (every 2.5 minutes) throughout the daylight hours.
2. Apply imaging processing algorithms to characterize phenotyping features such as NDVI, predicted RWC, and single spectral bands from canopy spectra.
3. Propose a diurnal pattern modeling approach to analyze the diurnal variation in plant phenotyping results.

4.2 Methods

4.2.1 High-throughput field imaging acquisition system

The field VNIR hyperspectral gantry platform at Purdue University’s Agronomy Center for Research and Education (ACRE) was used to collect imaging data in this study. A weatherproof VNIR push-broom hyperspectral camera (MSV-101-W, Middleton Spectral Vision, Middleton, WI, USA) was carried by the 7-meter-high gantry platform to scan a 50-by-5 meter strip field under a wide range of weather conditions. The VNIR images contained a spectral range from 376 nm to 1044 nm with a spectral resolution of 1.22 nm and spatial resolution of 0.5 cm/pixel ground sample distance (GSD). The system could be configured to automatically scan the crops in the

field repeatedly. It takes 6.5 minutes to scan the 250-square meters field, although the scanning frequency can be higher if a sub-portion of the field needs to be scanned (Fig. 4.1).

This single-sided imaging gantry stood at the north side of the field and the length of the camera structure on the top was restricted by a certain ratio of the gantry’s height. This unique design prevented any of the shadow from falling on the crops, which enabled the system to more realistically simulate drone remote sensing in the field. The hyperspectral camera utilized sunlight as the lighting source, so the gantry system could function any time after sunrise until sunset on each day. A white reference panel was installed 0.5 meters underneath the hyperspectral camera and moves along with it. A local mini weather station and distributed soil sensors were installed in the same field to collect real-time environmental condition data such as temperature, solar radiation, wind speed, and soil moisture when each image was taken.

Table 4.1. Parameters for the hyperspectral imaging sensor

Parameters	MSV-101-W
Camera model	acA780-75gm
Spectrograph	SpecIM V10H
Frame rate	30 FPS
Exposure time	6 ms
Spectral resolution	1.22 nm
Ground sample distance (GSD)	0.5 cm/pixel
Spectral range	376 nm -1044 nm

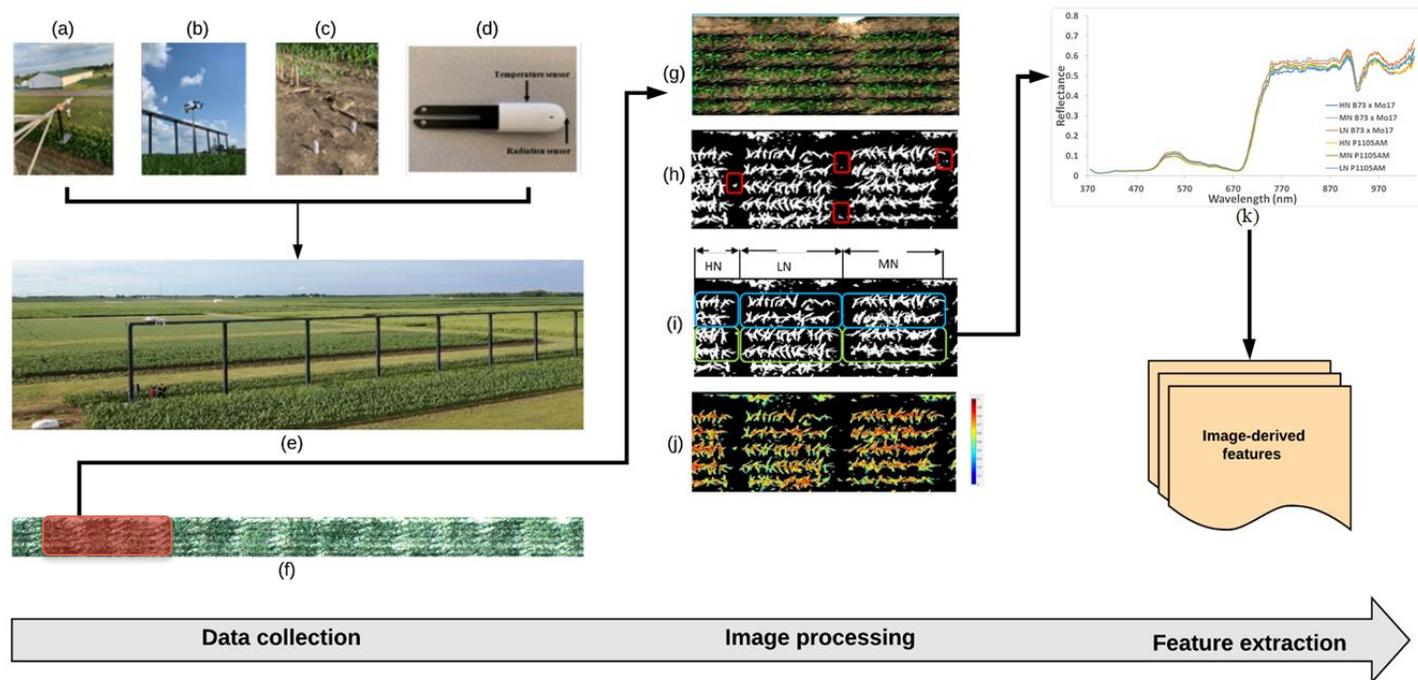


Fig. 4.1. Field VNIR hyperspectral platform at Purdue University. (a). VNIR hyperspectral imaging sensor; (b). Local weather station (Ambient Weather, Chandler, AZ, USA); (c) and (d). Xiaomi flower care sensor (Xiaomi Inc., Beijing, China); (e). The layout of the East-West orientated imaging system; (f). Image sample of the whole field; (g). Enlarged image sample of part of the field; (h). Binary image after segmentation; (i). Layout of the plots with three nitrogen treatments and two genotypes. The green boxes were for Genotype P1105AM, and the blue boxes were for Genotype B73 x Mo17. Nitrogen treatments of high nitrogen (HN), low nitrogen (LN) and medium nitrogen (MN) were also labeled; (j). NDVI heatmap; (k). The spectra from different genotypes and N treatments.

4.2.2 Experiment design and data collection

To study the diurnal variation in hyperspectral image data, two genotypes of corn plants, including genotype B73 x Mo17 and P1105AM, were grown in the field underneath the gantry in 2019. Each genotype was treated with three different nitrogen solutions: High Nitrogen (HN) with 56 kg/ha (32mL 28-0-0 in 1L water), Medium Nitrogen (MN) with 28 kg/ha (16mL 28-0-0 in 1L water), and Low Nitrogen (LN) with 0 kg/ha (water). Each of the genotype-by-nitrogen-solution treatment combination is repeated in five 2-rows-by-3-meter mini-plot replicates so 30 plots in total exist in the field.

To capture the instant effects to the images from changes in cloud coverage, wind speed and other environmental conditions, the team decided to image the field every 2.5 minutes. This allowed us to scan up to six different plots, once the extra time needed for data transfer, real-time image processing, and homing the gantry cart is taken into consideration. The six plots were selected to cover all three nitrogen levels and both genotypes.

The continuous imaging started when the corn plants reached V4 stage and lasted 31 days until the plants are at R1 stage on average. On each of the 31 days, imaging started at 7:00am and ended at 7:30pm. Around 280 hyperspectral images were collected with a repetitive imaging frequency of every 2.5 minutes. The gantry was only turned off during extreme weather conditions such as thunderstorms to protect the equipment. By the end of the 31 days, a total of 8,631 hyperspectral images were collected.

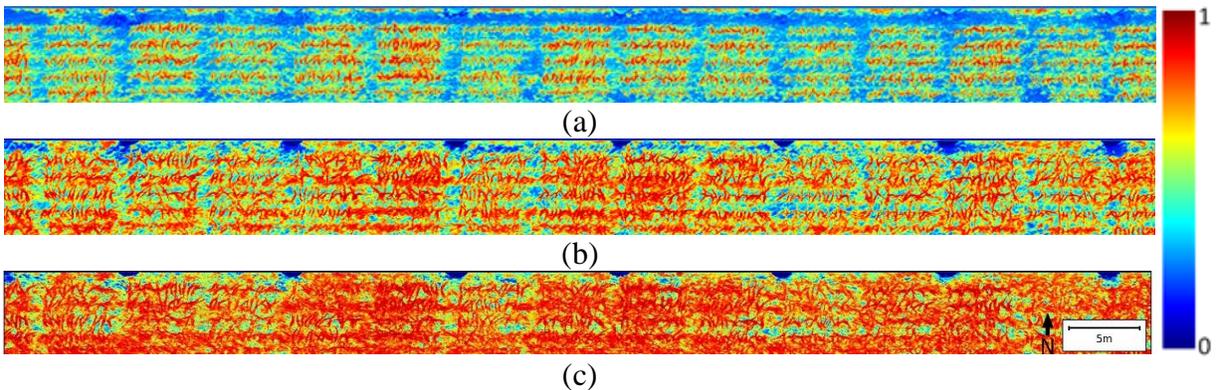


Fig. 4.2. The NDVI heatmaps for the whole field at three leaf stages with different accumulated days after planting (DAP). (a). 31 DAP, leaf stage V4. The red boxes indicate the regions for 30 distinct plots; (b). 38 DAP, leaf stage V6; (c). 49 DAP, leaf stage V9.

In the middle of the project when the plants were at V9, we collected ground truth measurements such as nitrogen content, RWC, and plant fresh weight. Two plants were randomly sampled from each plot. The plant shoot was cut to measure the fresh weight. A small section (2.5 cm × 5.0 cm) of the top-collared leaf was taken to measure the RWC using Eq. 4.1 (Turner, 1981). The remaining part of the top-collared leaf was sent to the Great Lakes A&L laboratories (A & L Great Lakes Laboratories, Inc., Fort Wayne, IN, USA) for measuring the nitrogen percentage.

$$Relative\ Water\ Content\ (RWC) = \left[\frac{(FW - DW)}{(TW - DW)} \right] \times 100 \quad (4.1)$$

Where FW represents fresh weight, DW denotes dry weight, and TW is the turgid weight.

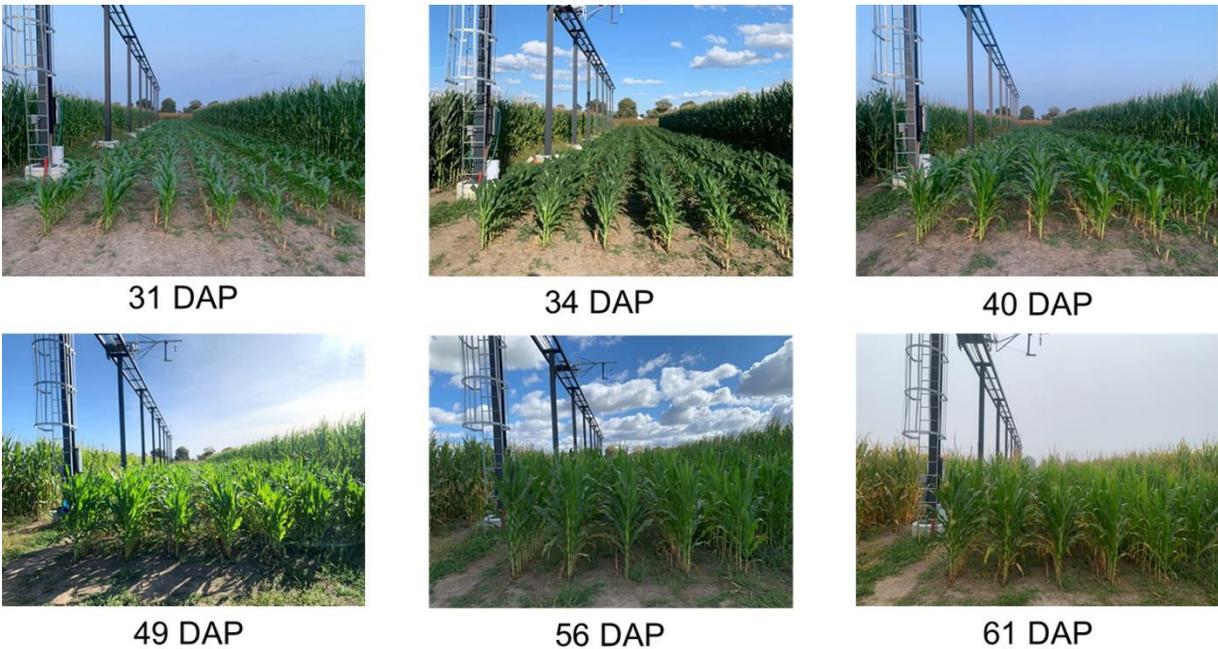


Fig. 4.3. The growth of corn plants in the Purdue ACRE field during the experiment at different accumulated days after planting (DAP).

4.2.3 Image segmentation and feature extraction

After data collection, standard imaging processing protocols were performed to extract the plant phenotyping features of interest. The raw hyperspectral images were firstly calibrated with the real-time white reference, which ensured each scanning line from this push-broom sensor was

calibrated with a reference that used the same lighting conditions. The image calibration was performed by the following equation:

$$R_{cali} = \frac{R_{raw} - R_{dark}}{R_{white} + R_{dark}} \quad (4.2)$$

Where R_{cali} is the calibrated image, R_{raw} denotes the raw hyperspectral image, R_{dark} represents the dark reference image and R_{white} is the hyperspectral image of the white reference. The calibrated images were then processed using a segmentation procedure with convolution methodology (Ma et al., 2019a; Zhang et al., 2019a). A vector of sequential integers from -20 to 20 was multiplied by the reflectance intensity vector from the red-edge region (680 – 720 nm). By choosing threshold 7 as the boundary, the plant tissue was successfully segmented from the background (Fig. 4.1h). Some images contained weeds whose size was irrelevant compared to the corn plants (See the red boxes in Fig. 4.1h).

The average reflectance spectrum from each plot was calculated. In total, 51,786 spectra (8631 images* 6 plots/image) were calculated for the plots with different genotypes and nitrogen treatments. These spectra were used to calculate the crop remote sensing results such as NDVI and predicted relative water content (RWC). The formula below was used for calculating NDVI.

$$NDVI = \frac{R_{800nm} - R_{650nm}}{R_{800nm} + R_{650nm}} \quad (4.3)$$

Where R_{800nm} and R_{650nm} represent the reflectance values of wavelength 800nm and 650nm respectively (Daughtry et al., 2000; Schafleitner et al., 2007). The partial least squares regression (PLSR) model was used to predict RWC from the spectra. In order to avoid prediction drifts between facilities (Alamar et al., 2007; Ji et al., 2015; Li et al., 2015), instead of using an existing RWC prediction model developed in previous studies, the team decided to build a new PLSR model with the spectra and RWC ground truth data collected in the same project. The new model predicted RWC with the cross-validation coefficient of determination (R^2) of 0.722 and root mean square error (RMSE) of 6.22% (Fig. 4). This RWC model was then applied to each of the 51,786 spectra to predict RWC in those plots over the 31 diurnal cycles.

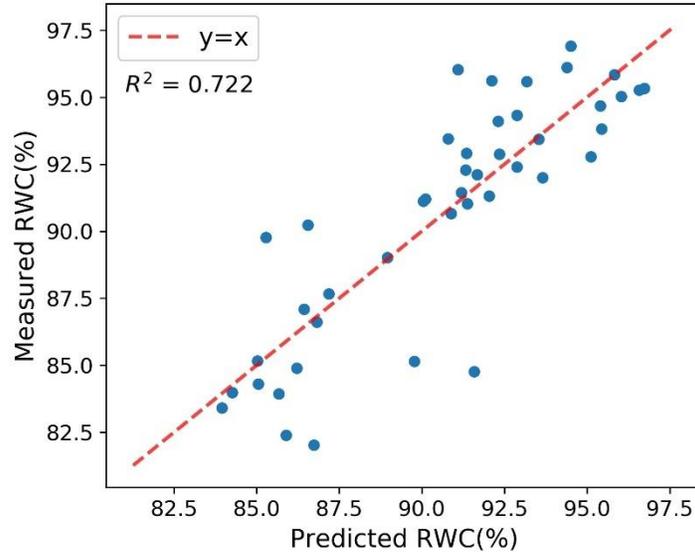


Fig. 4.4. The Relative Water Content (RWC) prediction model based on Partial Least Square Regression Relative (PLSR): measurement vs. prediction.

4.2.4 Data quality check

After calculating the remote sensing result data such as spectra, NDVI and predicted RWC, the data quality was checked, and outlier data removed. For each feature from each day, the measurements between the upper inner fence ($Q3+1.5IQR$) and lower inner fence ($Q1-1.5IQR$) were noted (Schwertman and de Silva, 2007). IQR is the interquartile range, equaling the difference between 75th (Q3) and 25th (Q1) percentiles. This quality filtering removed the blunders and gross errors in the image acquisition setup and process. For example, some of the removed outlier images were taken under very high wind-speed conditions, causing the plant stems to bend severely and significantly altering the images. The data before 10:00 am and after 5:30 pm also showed extreme variances and noises. This could be caused by the dews on the surface of leaves and dim lighting conditions (Manea and Calin, 2015). Only the data collected between 10 am and 5:30 pm was used in the diurnal pattern analysis since very few airborne remote sensing activities occur outside this time range. We ended up using this selected data pool (Table 4.2) for the diurnal pattern analysis. Please note the time was in the Eastern Standard Time (EST) Zone Daylight Saving Time (DST) and this selected time window remained centered at the solar noon in West Lafayette, IN.

Table 4.2. Number of image samples after the quality check for six plant plots.

Plant Plots	Number of samples after the quality check
HN & B73 x Mo17	5,070
MN & B73 x Mo17	5,092
LN & B73 x Mo17	5,083
HN & P1105AM	5,108
MN & P1105AM	5,084
LN & P1105AM	5,093

4.2.5 Evaluating the impacts from treatments, stages and genotypes to diurnal changing patterns

Before modeling the diurnal changing patterns, the diurnal patterns between different nitrogen treatments, growth stages and genotypes were compared to decide if any of these factors significantly impacted the diurnal pattern. If not, the data from different treatments, stages and genotypes could be combined for the diurnal pattern modeling. Otherwise, the modeling should be performed separately for each different case.

The changing patterns were compared by applying the dynamic time warping (DTW) method to calculate the similarity between the relative difference ratio (RDR) curves from the different plants plots (Berndt and Clifford, 1994). An RDR curve was calculated to describe the diurnal changing pattern each day as the percentage of the change of the phenotyping feature value relative to the feature's value at the reference time point (Eq. 4.4). For the reference time, we selected solar noon since this represents the center point of the daytime, when the lowest NDVI value was observed every day.

$$Relative\ difference\ ratio = \frac{Feature_t - Feature_{solar\ noon}}{Feature_{solar\ noon}} \quad (4.4)$$

The DTW method was selected as it is an algorithm more commonly used to measure the similarity between two time-series data (Berndt and Clifford, 1994). More specifically, DTW represents a time series alignment algorithm developed to align two sequences of feature vectors by warping the time axis iteratively until an optimal match occurs (Kate, 2016). Thus, a distance score is generated during the process of alignment, which can be used as the difference between two curves. For example, a small distance score means higher similarity between two curves. DTW also allows non-linear mapping, which was appropriate for the purpose of pattern matching. With

the distance scores from DTW, the similarity of RDR curves of different plots were quantitatively compared and discussed.

4.2.6 Diurnal patterns calculation by time series signal decomposition

Inspired by the idea of the time series decomposition method (Cleveland et al., 1990), we decomposed the changing signal of each feature into two major parts: the day-to-day trend (T_t) and the diurnal pattern (D_t). T_t is calculated with the LOESS (locally estimated scatterplot smoothing) method (Rojo et al., 2017). By fitting a non-parametric regression curve on the scattered plot of the data, the day-to-day change can be clearly extracted from the raw signal (Cleveland et al., 1990). This trend primarily reflects changes in plant growth and general weather conditions over the 31 days of imaging. The diurnal component (D_t) was calculated by subtracting the day-to-day trend (T_t) from the raw signal. D_t is also called the detrended data. D_t contains the higher-frequency-variance components mainly caused by short-term events such as plants' circadian behavior, sun angle, solar radiation and temperature changes during the day.

$$\text{Raw Time Series Signal} = T_t + D_t \quad (4.5)$$

The mean curve of D_t with all the data combined for 31 days was calculated for the diurnal pattern fitting. Meanwhile, the 95% confidence interval was also calculated to evaluate the consistency of the diurnal patterns across the days.

4.2.7 Diurnal pattern fitting

Both first-order (Eq. 4.6) and second-order (Eq. 4.7) piecewise models were tested to fit the diurnal changing patterns of the crop phenotyping features including NDVI, the RWC prediction as well as two single spectral band. By calculating the first-order derivative of the mean curve, the critical lowest point t_a (which perfectly matched the time with the highest sun angle) was derived. The performance of developed models was evaluated and compared with the coefficient of determination (R^2) and root mean square error (RMSE) between the fitted results and the original measurements. For each feature, when both displayed excellent regression performances with no significant difference between the two models, we adopted the simpler first-order model.

$$f(t) = \begin{cases} a_1t + b_1 & \text{if } t \leq t_a \\ a_2t + b_1 & \text{if } t > t_a \end{cases} \quad (4.6)$$

$$f(t) = \begin{cases} a_1t^2 + b_1t + c_1 & \text{if } t \leq t_a \\ a_2t^2 + b_2t + c_2 & \text{if } t > t_a \end{cases} \quad (4.7)$$

Where t is the time offset (in hours) from solar noon. (i.e., t at 12:15 pm is -1.5, on a day and at a location where solar noon is 13:45.)

4.2.8 Model performance evaluation

The performance of the developed diurnal model was evaluated and compared with the R^2 and RMSE between the prediction results and the calculated diurnal changes from time series decomposition. Moreover, to further assess the impacts from nitrogen treatments and genotypes, the general diurnal model was tested on and the R^2 and RSME obtained for each of the six plant plots.

4.2.9 Diurnal models' applications

These regression models of diurnal changes can be used to calculate suitable imaging windows based on allowable diurnal variances. The models can also be used to remove the diurnal effect. For example, the NDVI measured at any other time point of the day can be converted to the NDVI at the highest sun angle time (a “solar noon equivalent”) with the fitted model. However, the data used in the model was drawn from one single field test whose imaging data was collected from Purdue’s field gantry system, which might induce systematic bias in the model. The camera sensor on this gantry is seven meters above the ground, while the normal UAV images above 60 meters (Sadeghi and Sohrabi, 2019). External validation data from the other remote sensing platforms such as UAVs are still needed.

4.3 Results

In this study, the diurnal changing patterns models were built for various plant phenotyping features (including NDVI, predicted RWC, Band670 (Red) and Band760 (NIR)). The detailed procedures and corresponding results for modeling NDVI’s diurnal changing pattern was demonstrated and discussed first, as NDVI is one of the most common plant features in remote

sensing (Cabrera-Bosquet et al., 2011). The summarized modeling results are reported for the other plant phenotyping features including predicted RWC, Red and NIR, respectively.

4.3.1 The NDVI diurnal fluctuations

The measured NDVI values was plotted to show the diurnal fluctuations from the raw imaging data (Fig. 4.5). Fig. 4.5a displays the results for all 31 days. Each gathering group represents one day, and the gaps between groups denote the time between 5:30 in the afternoon until 10:00 next day without imaging data. In this figure, NDVI shows obvious diurnal variabilities with repeatable V-shaped patterns. Fig. 4.5b displays the zoomed in diurnal fluctuation in Day 1. It shows that NDVI continues to decrease until solar noon at 1:39pm on that day. This confirmed similar findings that had been reported in previous research papers that the NDVI of wheat and soybean were higher at the beginning and end of the day (Sagan et al., 2019; Sticksel et al., 2004). This diurnal change could occur due to the combination of the imaging lighting condition change and the plant's physiological changes over a diurnal cycle, which is central in regulating the susceptibility and responses to biotic or abiotic stresses (Dunford et al., 2002; Kloosterman et al., 2013; Turner, 2007).

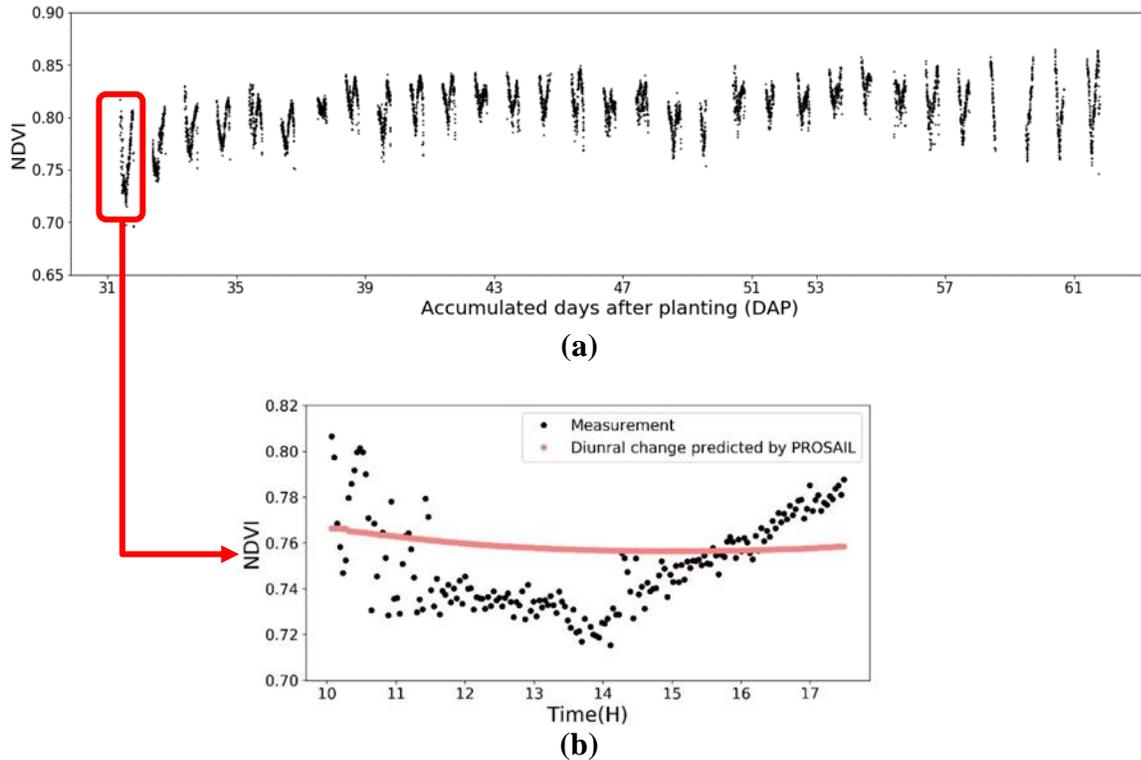


Fig. 4.5. (a). The NDVI of HN and Genotype B73 x Mo17 plot from V4 stage to the R1 stage (obtained from the hyperspectral images); (b). The NDVI measurements across Day 1 (black dots). The red curve is the predicted NDVI diurnal variance by PROSAIL model. Parameters for the PROSAIL model for the corn canopies followed the work of Ishihara et al. in 2015.

The significance of NDVI's diurnal variation was assessed by calculating the difference ratio between the highest and lowest NDVI during the same day (Table 4.3). The HN and B73 x Mo17 genotype plot demonstrated the highest diurnal variation in NDVI, with a 15.71% change over the day. This high diurnal variation is too extreme to be accepted in most plant phenotyping studies (Ni et al., 2015; Rahman et al., 2015). This result therefore justifies the need to model the diurnal pattern and calibrate the diurnal changes.

Table 4.3. Statistical analysis of the NDVI at Day 1 for six plant plots

Plant Plots	Mean	Highest	Lowest	Diurnal variation (%)
HN & B73 x Mo17	0.756	0.808	0.698	15.71
MN & B73 x Mo17	0.732	0.780	0.689	13.33
LN & B73 x Mo17	0.686	0.741	0.645	14.95
HN & P1105AM	0.759	0.817	0.717	13.96
MN & P1105AM	0.738	0.797	0.689	15.68
LN & P1105AM	0.687	0.733	0.648	13.20

4.3.2 The impacts of nutrient treatments, genotypes and leaf stages on diurnal variation

The relative difference ratio (RDR) curves were compared to study the effects of treatments, genotypes and growth stages on diurnal variation patterns. For each plant plot, the mean RDR curve for NDVI was computed from all 31 RDR curves (one RDR curve per day). In Fig. 4.6, the mean NDVI RDR curves from different treatment and genotype plots showed very similar trends.

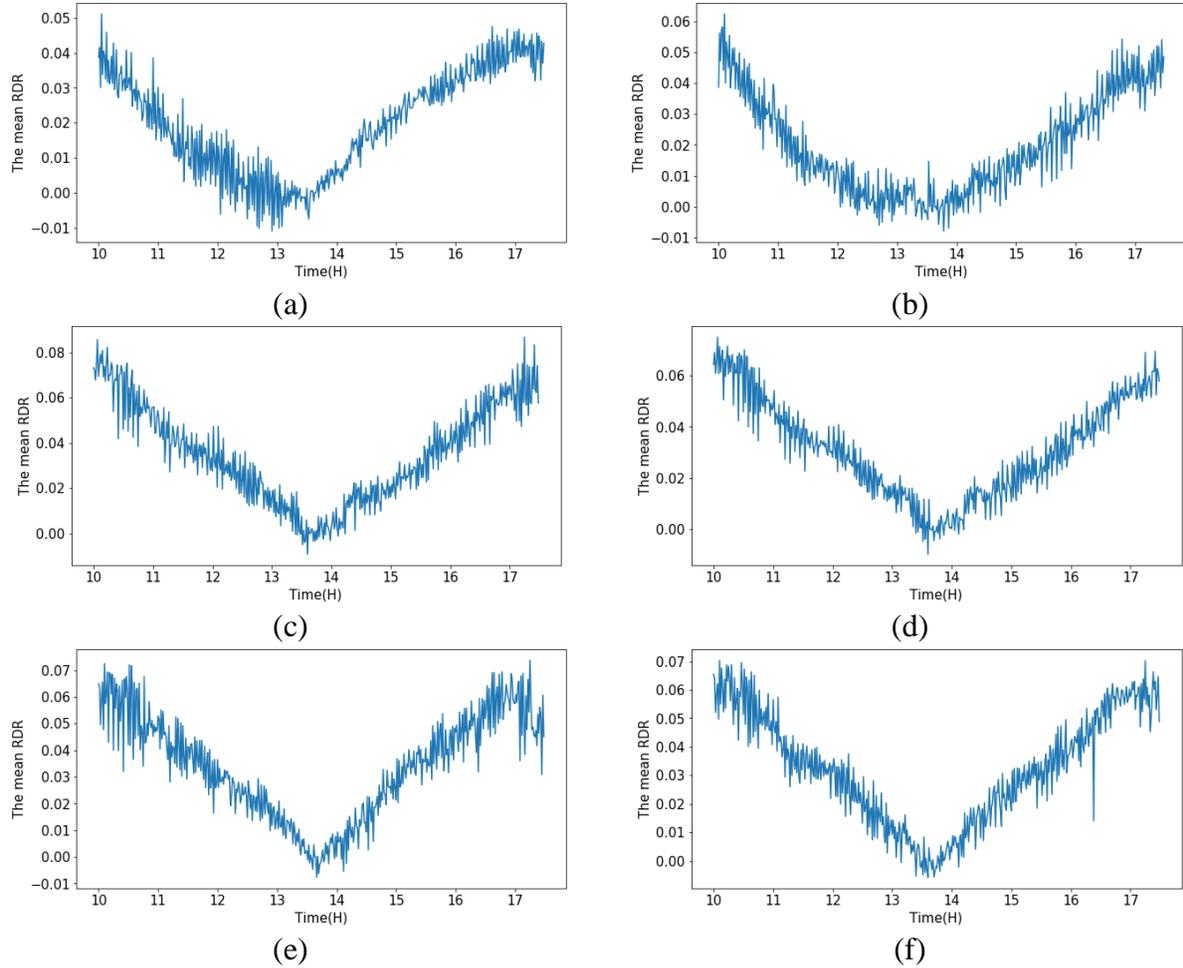


Fig. 4.6. The mean NDVI relative difference ratio (RDR) curves of six plant plots versus time (H): (a). HN & B73 x Mo17; (b). HN & P1105AM; (c). MN & B73 x Mo17; (d). MN & P1105AM; (e). LN & B73 x Mo17; (f). LN & P1105AM.

For a more quantitative comparison, the DTW distance scores were calculated to measure the similarity between RDR curves (Table 4.4). A small distance score means greater similarity between two curves. To indicate similarity of two RDR curves, we assumed DTW larger than 5 to be not similar whereas those with a DTW lower than 5 are considered similar. The distance scores between plots with the same genotype but different nitrogen treatments range between 1.37-3.75. Meanwhile, the distance scores between different genotype plots are 0.33, 0.20 and 0.30 for HN, MN and LN, respectively. Table 4.5 shows the DTW distance scores between different leaf stages. All these distance scores lie below 5. To summarize, comparatively, the nitrogen treatment impacted the variation in diurnal trends more than genotype and plant stage did. This finding confirmed the observation in previous study (Atkin et al., 2000). However, all these factors did not

much impact the diurnal trend ($DTW < 5$), which allowed the combination of data to build one general diurnal changing model.

Table 4.4. The DTW distance scores between NDVI's RDR curves of the plant plots with different nitrogen treatments and genotypes.

Plant Plots	HN & B73 x Mo17	HN & P1105AM	MN & B73 x Mo17	MN & P1105AM	LN & B73 x Mo17	LN & P1105AM
HN & B73 x Mo17	0.00	0.33	1.37	1.48	1.40	2.31
HN & P1105AM	0.33	0.00	2.67	2.95	2.59	3.75
MN & B73 x Mo17	1.37	2.67	0.00	0.20	0.14	0.60
MN & P1105AM	1.48	2.95	0.20	0.00	0.30	0.82
LN & B73 x Mo17	1.40	2.59	0.14	0.30	0.00	0.30
LN & P1105AM	2.31	3.75	0.60	0.82	0.30	0.00

Table 4.5. The DTW distance scores between NDVI's RDR curves of the plant plots with different plant stages

Plant Plots	Distance score between early stage and late stage
HN & B73 x Mo17	0.90
MN & B73 x Mo17	1.26
LN & B73 x Mo17	1.15
HN & P1105AM	0.25
MN & P1105AM	0.74
LN & P1105AM	0.63

Note: Early stage contains the first 15 days, and late stage contains the later 16 days.

4.3.3 Diurnal changing pattern

Based on the result in Section 4.3.2, the data of all six plant plots were combined to model the diurnal changing pattern since the nitrogen treatments, genotypes and leaf stages only had minor effects on the diurnal variation ($DTW < 5$); a general diurnal pattern model is thus preferred for its relative ease of use.

4.3.3.1 Decomposition

The changing signal in predicted phenotyping features were decomposed into two major parts: day-to-day trend (T_t) and diurnal pattern (D_t). Fig. 4.7 shows the decomposition result for NDVI. As seen in the chart of trend T_t , NDVI values increased overall from the early leaf stage to the late stage. This confirmed the previous findings that as plants mature, NDVI increased until the reproductive stage (Wang et al., 2016). In our experiment, there were two big dips in NDVI along the timeline, which correspond to two severe temperature drops in the West Lafayette area.

The diurnal changes illustrated in Fig. 4.7 showed clear repetition of V-shaped NDVI measures over time. The NDVI value consistently decreased until it reached solar noon. We were not able to collect the complete data for the whole day on 36, 37, 58, 59 and 60 DAP due to the extreme weather conditions. In this paper, we only focus on modeling the diurnal changing patterns under “normal” weather conditions. The modeling of the weather condition impacts will be reported in another paper.

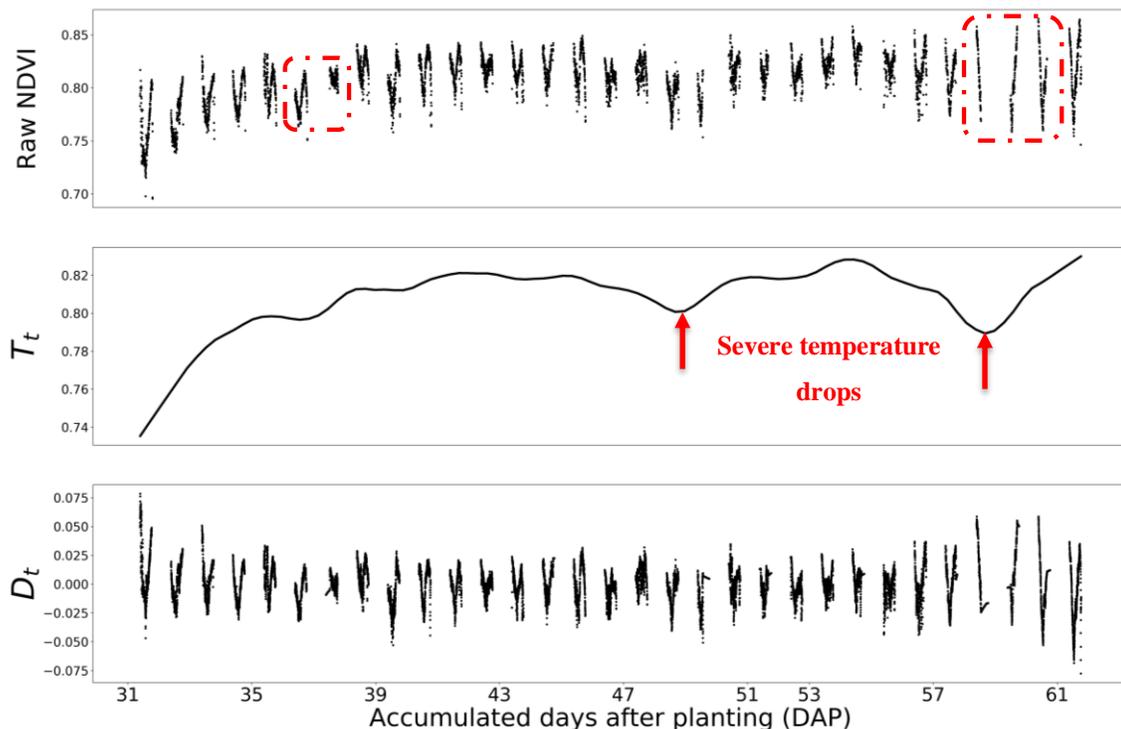


Fig. 4.7. The NDVI of HN and Genotype B73 x Mo17 plot from V4 stage to the R1 stage. The raw NDVI plot was decomposed into the day-to-day trend (T_t) and diurnal pattern (D_t). The red boxes are the days with incomplete data measurements due to the extreme weather conditions, which are 36, 37, 58, 59 and 60 DAP.

To model the general diurnal changes, the NDVI diurnal changes for all 31 days were combined to calculate the mean curve and 95% confidence interval (Fig. 4.8). The narrow 95% confidence interval also provided strong evidence that the diurnal pattern remains consistent through the days. As discussed in methodology, since the confidence interval is much lower compared with the mean curve's change over the diurnal cycle, the mean curve (black line) is used to model the diurnal pattern. This also helps to “average out” the impacts from abnormal weather conditions on a few days.

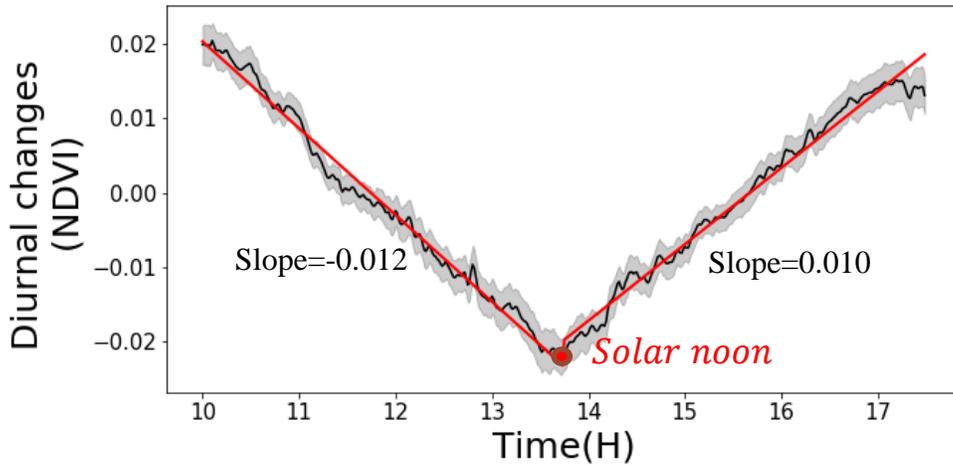


Fig. 4.8. The diurnal changes of NDVI summarized from 31 days and six experimental plots. The x-axis is the diurnal time in a unit of hour. The y-axis is the diurnal NDVI changes. The black line is the mean diurnal NDVI adjustment value. The shaded area is the 95% confidence interval. The red line is the first-order piecewise fitted result for the mean diurnal NDVI adjustment value.

4.3.3.2 Pattern fitting

Both first-order and second-order piecewise polynomial models were tested to fit the diurnal changes of NDVI. Both displayed excellent regression performances without significant difference between the two models. We therefore adopted the simpler first-order model, concluding that the NDVI diurnal changes can be described by a linear model. The fitted first-order model in Eq. 4.8 can then be used to describe the NDVI diurnal changing pattern based on the time offset from solar noon.

$$NDVI(t) = \begin{cases} -0.012 \times t + NDVI_{\text{solar noon}} & \text{Before solar noon} \\ 0.010 \times t + NDVI_{\text{solar noon}} & \text{After solar noon} \end{cases} \quad (4.8)$$

Where t is the time offset (in hours) from solar noon. (i.e., t at 12:15 pm is -1.5, on a day and at a location where solar noon is 13:45.) $NDVI_{\text{solar noon}}$ is the NDVI measured at solar noon.

Equations Eq. 4.8 fit the observed diurnal pattern with an R^2 of 0.99 and RMSE of 0.0012. Fig. 4.8 illustrates the detailed fitted result. To test the model for different genotypes and nitrogen treatments, the fitted equation was applied on each different field plot. The model maintained an accurate performance in all cases (Table 4.6), proving this general diurnal pattern model remains valid for plants from different nutrient treatments and genotypes.

Table 4.6. Fit and error of the piecewise NDVI diurnal changing pattern model (Eq 4.8) on assorted experiment plots.

Plant Plots	R²	RMSE
HN & B73 x Mo17	0.77	0.0058
MN & B73 x Mo17	0.96	0.0036
LN & B73 x Mo17	0.95	0.0032
HN & P1105AM	0.91	0.0038
MN & P1105AM	0.94	0.0034
LN & P1105AM	0.98	0.0022

4.3.3.3 Applications of the model

The diurnal changing pattern model can be utilized to determine the proper imaging window based on specific quality requirements. In Table 4.7, the proper imaging windows were calculated based on the different tolerance thresholds for variations of NDVI. For example, when solar noon was the standard imaging time, to limit the NDVI's diurnal changes within ± 0.03 on the same plant, the imaging needs to occur between 10:55 and 16:30.

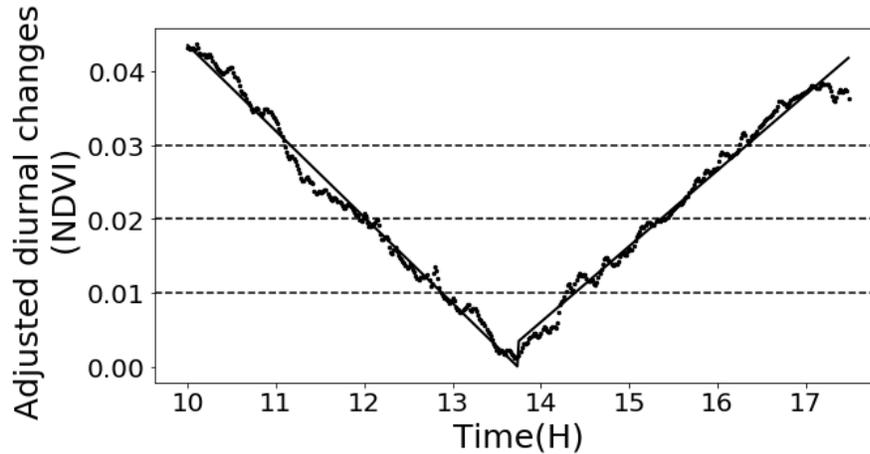


Fig. 4.9. The intersections between the adjusted NDVI's diurnal changes and three different thresholds for proper imaging windows. The x-axis is the diurnal time in a unit of hour. The y-axis is the adjusted NDVI's diurnal changes when adjustment at solar noon to be 0.

Table 4.7. Proper imaging windows for different thresholds of the NDVI changes caused by diurnal factors (when solar noon is the standard imaging time).

Thresholds	Suggested imaging time	Range
0.01	12:45 – 14:40	1h 55min
0.02	11:55 – 15:45	3h 50min
0.03	10:55 – 16:30	5h 35min

Besides uncovering an acceptable imaging time window, the diurnal changing model can also be used to easily calibrate the diurnal variances. For example, an NDVI measured two hours before solar noon should be decreased by 0.024 (0.012×2), so the adjusted NDVI can be considered “solar noon equivalent”. This ameliorates the significant diurnal variance and enable phenotyping researchers to do fairer comparisons and analyses.

4.3.4 Other image-derived phenotyping features

Along with NDVI, the diurnal changing patterns of other plant phenotype features, including RWC prediction, as well as Red and NIR bands in the spectra, were modeled in a similar way. Due to the time-consuming modeling process, we chose the Red and NIR bands to be analyzed first because these are important bands for calculating many plant features such as chlorophyll content. Both first-order and second-order piecewise models were tested to fit the diurnal changes of predicted RWC, Red and NIR. The fitting results were shown in Table 4.8. For

the predicted RWC and NIR, both first-order and second-order piecewise models had excellent regression performances and there was no significant difference. Therefore, we adopted the simpler first-order model. We therefore concluded the RWC and NIR diurnal changing patterns could be described by linear models. On the other hand, the Red diurnal changing pattern displayed a non-linear pattern, so we selected the second-order model.

Table 4.8. The fitted results of piecewise diurnal changing pattern models for other plant phenotyping features including predicted RWC, Red and NIR bands in the spectra.

Features	R^2		RMSE	
	First-order model	Second-order model	First-order model	Second-order model
Predicted RWC	0.96	0.97	0.30	0.16
A red band	0.82	0.91	0.0013	0.00094
A NIR band	0.97	0.98	0.013	0.0071

The diurnal changes of predicted RWC, as well as Red and NIR in the spectra were fitted and plotted in Fig. 4.10a, b, and c. Equations 4.9, 4.10 and 4.11 represent the diurnal changing pattern models. All models showed an excellent description of the observed diurnal patterns. Similarly, these models demonstrated the potential to help improve the efficiency and data quality in phenotyping practices. Moreover, successful application of the same proposed diurnal pattern modelling method on the predicted RWC, Red and NIR indicated that this method could potentially generalize well to other phenotypic features, which could be further explored in future research.

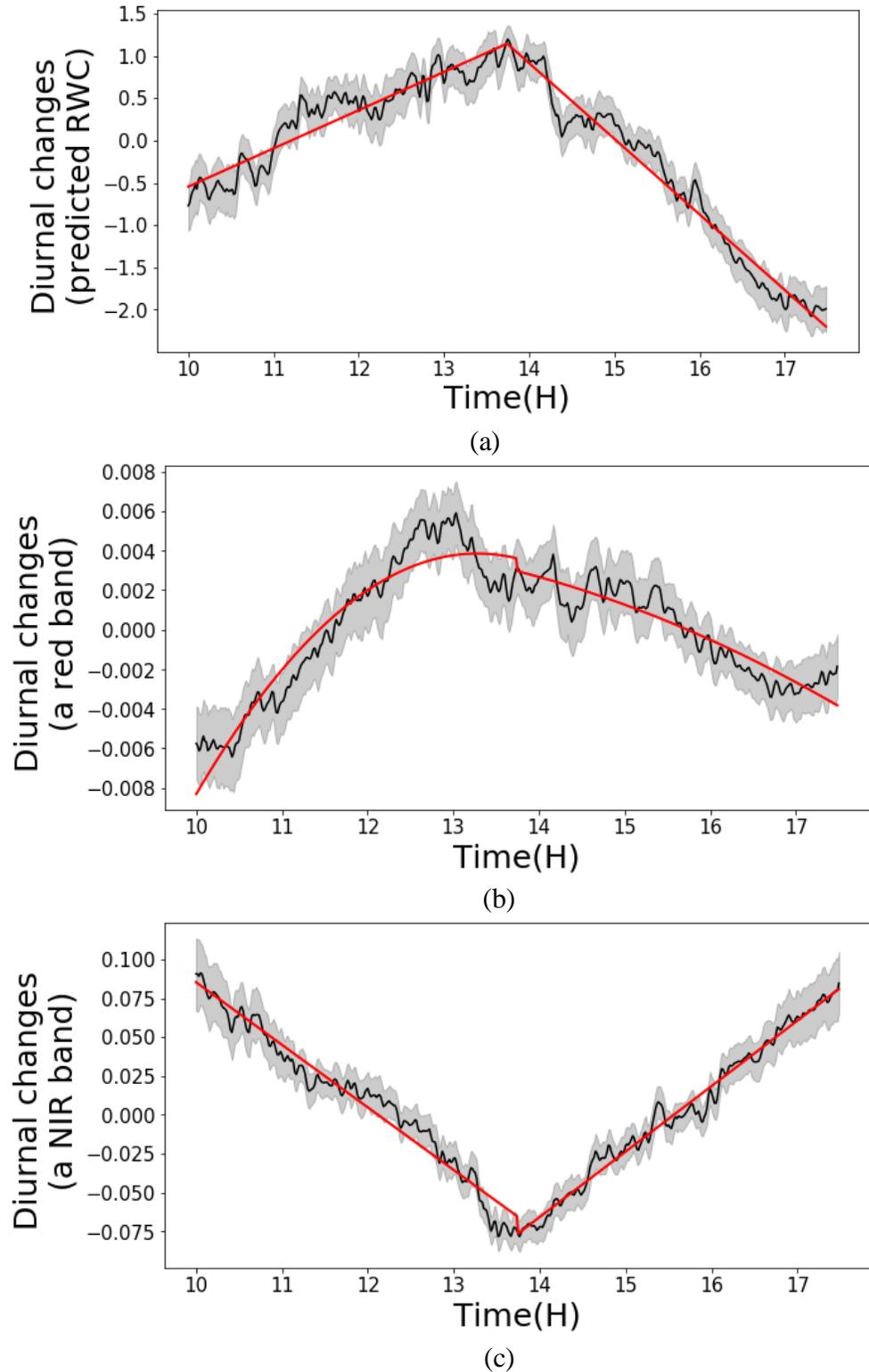


Fig. 4.10. The diurnal changing patterns and fitted results of other plant phenotyping features including: (a). predicted RWC; (b). a red band; (c). a NIR band. The black line in each plot is the mean diurnal adjustment value. The shaded area is the 95% confidence interval. The red line is the piecewise fitted result for the mean diurnal adjustment value.

$$RWC(t) = \begin{cases} 0.45 \times t + RWC_{\text{solar noon}} & \text{Before solar noon} \\ -0.89 \times t + RWC_{\text{solar noon}} & \text{After solar noon} \end{cases} \quad (4.9)$$

$$Red(t) = \begin{cases} -0.0011 \times t^2 - 0.0011 \times t + Red_{\text{solar noon}} & \text{Before solar noon} \\ -0.00019 \times t^2 - 0.0011 \times t + Red_{\text{solar noon}} & \text{After solar noon} \end{cases} \quad (4.10)$$

$$NIR(t) = \begin{cases} -0.040 \times t - NIR_{\text{solar noon}} & \text{Before solar noon} \\ 0.042 \times t - NIR_{\text{solar noon}} & \text{After solar noon} \end{cases} \quad (4.11)$$

4.4 Discussions

4.4.1 Strengths

The diurnal patterns of plant phenotyping features, including NDVI, predicted RWC, and two individual spectral bands (Red and NIR), were quantitatively described and modeled in detail for the first time. The unprecedented high-imaging frequency of the new field hyperspectral-imaging gantry at Purdue, with over 8,000 repeated images during a single growing season. Furthermore, most previous studies reported the diurnal changes directly from the raw phenotyping features (e.g., NDVI, chlorophyll content) (Beneduzzi et al., 2017; Ishihara et al., 2015; Sticksel et al., 2004). However, these results remain inconsistent due to changes in plant growth stage and weather condition. To address this issue, our proposed approach incorporates a modified time-series decomposition method to filter out the signal variance caused by the changing plant stage and weather conditions, so the later modeling process can exclusively focus on the higher-frequency diurnal changes.

4.4.2 Limitations and future work

In this experiment, all the imaging data was collected from the same field imaging gantry facility. Additional testing is needed to validate the other remote sensing platforms such as UAV. In addition, the diurnal patterns are reported based on corn plants, which may limit the scope of application. More tests should be conducted on diverse plant species (e.g., soybean, wheat, and rice).

In the 2020 season, the developed diurnal models will be externally validated using the field gantry's images. The model will also be validated on a UAV hyperspectral imaging platform in

the same field. Finally, the diurnal changing patterns on other plant species, such as soybean wheat and rice will be also explored and compared.

Besides the diurnal effect, the changing environmental conditions in the field such as cloud coverage, temperature, wind speed, and so on all showed strong impacts on the phenotyping results. The data has shown that these effects could also be precisely modeled. The environmental calibration models for plant remote sensing will be reported in future papers.

4.5 Conclusions

In this study, diurnal changing patterns in crop remote sensing images were quantitatively investigated with the proposed novel modeling approach. Over 8000 repeated hyperspectral images of the corn field were taken by the field imaging gantry at Purdue University throughout the 2019 growing season. Images of corn plants from stages V4 to R1 were captured. Hyperspectral image processing algorithms were applied to calculate the reflective canopy spectra and predict plant physiological features such as NDVI, RWC, and two individual spectral bands (Red and NIR). The diurnal patterns for these phenotyping features were successfully modeled. The results showed that the NDVI presents a repeatable V-shaped diurnal changing pattern: it linearly decreases by 0.012/h before the solar noon and increases by 0.010/h thereafter. Besides NDVI, predicted RWC and the NIR band display linear diurnal changing pattern as well, where predicted RWC changes in an inversed V-shaped pattern and NIR changes in a normal V-shaped pattern. The red band shows a quadratic diurnal changing pattern in an inversed V shape. With the result of this work, Ag remote sensing users will be able to more precisely understand the deviation or change in crop feature predictions caused by the specific imaging time of the day. The reported diurnal models can also be used to correct the remote sensing result so to remove the diurnal effects.

CHAPTER 5. MODELING OF ENVIRONMENTAL EFFECTS ON AERIAL HYPERSPECTRAL IMAGES FOR CORN PHENOTYPING

Aerial imaging technologies have been widely applied in agricultural plant remote sensing. However, an as-yet unexplored challenge with field imaging is that the environmental conditions such as sun angle, cloud coverage, temperature, and so on. can significantly alter plant appearance and thus affect the imaging sensor's accuracy toward extracting plant feature measurements. These image alterations result from the complicated interaction between the real-time environments and plants. Analysis of these impacts requires continuous monitoring of the changes through various environmental conditions which has been difficult with current airborne remote sensing systems. This paper aimed to propose a modeling method to comprehensively understand and model the environmental influences on hyperspectral imaging data. In 2019, a fixed hyperspectral imaging gantry was constructed in Purdue University's research farm, and over 8,000 repetitive images of the same corn field were taken with a 2.5 minutes interval for 31 days. Time-tagged local environment data including solar zenith angle, solar irradiation, temperature, wind speed, and so on. were also recorded during the imaging time. The images were processed for phenotyping data, and the time series decomposition method was applied to extract the phenotyping data variation caused by the changing environments. An artificial neural network (ANN) was then built to model the relationship between the phenotyping data variation and environmental changes. The ANN model was able to accurately predict the environmental effects in remote sensing results, and thus could be used to effectively eliminate the environment-induced variation in the phenotyping features. The test of the Normalized Difference Vegetation Index (NDVI) calculated from the hyperspectral images showed that variance in NDVI was significantly reduced by 79%. A similar performance was confirmed with the Relative Water Content (RWC) predictions. Therefore, this modeling method showed great potential to be applied in airborne remote sensing applications in agriculture to significantly improve the imaging quality by effectively eliminating the effects from the changing environmental conditions.

5.1 Introduction

Recent years have seen the rapid growth of remote sensing applications in the field of agriculture (Gracia-Romero et al., 2019; Li et al., 2014; Wang et al., 2016). The advent and advances of low-weight, low-cost imaging platforms, and smart imaging devices resulted in the improved capability of the agricultural data collection. With various sensors such as red-green-blue (RGB), hyperspectral and thermal cameras carried by these platforms, plant phenotypic properties are captured in the images that largely facilitate the process of crop analyses such as accessing plant biomass, nutrient level, diseases stresses and so on (Li et al., 2014; Ma et al., 2019b; Wang et al., 2020; Zhang et al., 2019b). However, the changing environmental conditions have been reported to significantly impact the imaging result (Gamon et al., 2015). The intensity of remotely sensed images changes greatly based on when and where the image is captured (Beneduzzi et al., 2017; Maji et al., 2014; Padilla et al., 2019). One source of the variation arises from the complicated interactions between the camera's sensitivity, camera's view angle, plant canopy geometry, solar zenith angle, solar azimuth angle, and shadows (Danner et al., 2019; Ishihara et al., 2015; Jackson et al., 1979; Ranson et al., 1985). Another source of variation results from plants' endogenous stress responses to the environmental conditions with complicated interactions between their genetic backgrounds, external environments, and treatments (An et al., 2017; Ranson et al., 1985). All of these, collectively regarded as the environment-induced variation in phenotyping features, affect plants' final reflectance characteristics.

To reduce the impacts from environment variation, a relatively simple method involves standardizing or fixing the sampling time of the day and restricting imaging to clear conditions without cloud coverage (Barbedo, 2019; Di Gennaro et al., 2018; Gracia-Romero et al., 2019). Bellvert and Girona (2012) suggested that the field phenotypic data should be collected around noon under clear weather conditions. Similarly, unmanned aerial vehicles (UAV) imaging is preferably performed at midday to ensure consistent data collection and analysis (Krishna, 2018). These restrictions could reduce the environmental impacts on the data, but they also significantly inhibit the imaging window and flexibilities. Practically speaking, performing the imaging at a fixed time is difficult, as many procedures, such as equipment setup, need to be completed before imaging, while the field environment is naturally uncontrollable and unpredictable. These challenges often result in the collection of remote sensing data at different times of the day under varying environmental conditions. Therefore, the correction of the impact of different imaging

time and varying environmental conditions to improve the quality of agricultural remote sensing is critically important to study.

Most airborne remote sensing systems require imaging white reference panels beside the target plants so that the sensing data are calibrated against reference values to remove the illumination variation between images (Miura and Huete, 2009). White reference calibration is effective in compensating different lighting conditions, but these reference images do not precisely reflect the bidirectional reflectance (BRDF) on the leaf surface. Variations from the changes in leaf angles and the plant's endogenous responses remain. An existing calibration method, the combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model (PROSAIL) (Jacquemoud and Baret, 1990), enables the prediction of the plant canopy spectral reflectance changes caused by the changing environmental conditions (Ishihara et al., 2015; Jacquemoud and Baret, 1990). However, the model usually does not meet the accuracy requirement in plant phenotyping remote sensing (Berger et al., 2018b). Furthermore, the PROSAIL prediction theoretically requires three input variables including leaf structure parameter, photosynthetic pigment concentration, and water content, which are difficult and costly to measure in remote sensing practices (Jacquemoud et al., 2009).

Another potential solution arises from the use of different regression methods to predict and compensate the environmental effects. For example, a correction model with the polynomial regression method was developed to predict the crop reflectance as a function of solar zenith angle, time of day, and ICI (instantaneous clearness index). The capability of the model in reducing the diurnal variation with Green Normalized Difference Vegetation Index (GNDVI) and some individual bands (de Souza et al., 2010) was tested. However, that model only calibrates the imaging time and ambient lighting factors, while many more environmental condition changes such as temperature and wind speed also impact the imaging result. Moreover, the plant data was collected on a small portion of the leaf by a handheld radiometer with four bands, which may not properly simulate airborne remote sensing platforms carrying hyperspectral or multispectral cameras. The simple polynomial regression model could successfully describe the changes in data over three consecutive days. However, it may fail to represent the general pattern on other days when the plants are at different stages of their growth cycle. Therefore, a comprehensive environmental impact analysis for general aerial remote sensing images is still critically needed. This analysis requires the continuous collection of crop images at various plant stages through

different environmental conditions, a task that has proven challenging with existing airborne remote sensing systems.

On the modeling method side, the artificial neural network (ANN) models, as opposed to conventional regression models, has received considerable attention because of its ability to learn the features directly from the raw data without prior knowledge and human effort in feature design (Wang et al., 2019). Due to their better data utilization capacity, ANN models have outperformed conventional methods for solving regression problems in many ways (Abiodun et al., 2018). For example, researchers have achieved high accuracies and efficiencies on modeling multivariable and time-series datasets (Mokarram and Bijanzadeh, 2016; Yilmaz and Kaynar, 2011; Zhang et al., 2017). Given the previous successful applications, an ANN model can prove a reliable and efficient alternative for modeling the environment-induced effects in remote sensing data.

This chapter introduced the research work of correcting the aerial remote sensing results by modeling and analyzing the effects from the changing field environmental conditions such as sun radiation, solar zenith angle, humidity, temperature and wind speed. There are three major objectives in the work of this chapter:

1. Collect time-series hyperspectral images of two varieties of corn plants with three nitrogen treatments from V4 to R1 every 2.5 minutes throughout the whole growing season, along with synchronized environmental condition data.
2. Build a prediction model for the environment-induced variation in each of the measured phenotyping features (e.g., NDVI and RWC) with time-series decomposition and ANN method.
3. Evaluate the performance of the trained ANN models and their effects in removing the environmental noise by comparing the variances in the phenotyping features (e.g., NDVI and RWC) before and after the model correction.

5.2 Materials and methods

5.2.1 Experiment design and data collection

To analyze the environmentally induced variation in phenotyping data, hyperspectral images of the crops and environmental data were collected from a corn field in the Purdue University Agronomy Center for Research and Education (ACRE). Two genotypes (B73 x Mo17

and P1105AM) of corn (*Zea mays* L.) were grown in the summer of 2019. Each genotype was treated with three different nitrogen (N) solutions: high N with 56 kg/ha (32mL, 28-0-0 in 1L water), medium N with 28 kg/ha (16mL, 28-0-0 in 1L water), and low N with 0 kg/ha (water). In total, six experimental plots existed, with one of two corn genotypes treated with one of three nitrogen level; each plot had around 25 plant replicates. The abbreviation for each experimental plot is listed in Table 5.1.

Table 5.1. Abbreviations of plant plots with different nitrogen treatments and genotypes.

Plant groups	Genotypes	N Treatments	Abbrev
1	B73 x Mo17 (Genotype 1)	High	G1H
2	B73 x Mo17 (Genotype 1)	Medium	G1M
3	B73 x Mo17(Genotype 1)	Low	G1L
4	P1105AM (Genotype 2)	High	G2H
5	P1105AM (Genotype 2)	Medium	G2M
6	P1105AM (Genotype 2)	Low	G2L
1-6 combined	All combined	All combined	All

Hyperspectral images of corn plants were continuously acquired using the Purdue field VNIR hyperspectral imaging gantry system (Ma et al. 2020). To capture the instant environmental effects on the images, imaging frequency was set at 2.5 minutes. Starting from the vegetative growth stage V4, the continuous imaging occurred for 31 consecutive days until the plants reached the reproductive stage R1. Every day, imaging started at 8:00 am and ended at 7:30 pm. In total, we collected 8631 hyperspectral images of the same crop field (Table 5.2) for this study. After data collection, the hyperspectral images were further processed to measure the plant phenotyping features of interest, including the reflectance spectrum, NDVI, and predicted RWC for each experimental plot. The reflectance spectrum was obtained from the averaged data of plant tissues using the image segmentation algorithm highlighted in (Ma et al., 2019a). The NDVI was calculated from the spectrum by following Eq. 5.1 (Daughtry et al., 2000; Schafleitner et al., 2007). The plant's RWC was predicted with the pretrained partial least squares regression (PLSR) model (Ma et al., 2020).

$$NDVI = \frac{R_{800nm} - R_{650nm}}{R_{800nm} + R_{650nm}} \quad (5.1)$$

Table 5.2. Hyperspectral images and environmental data collection.

Data collection	Sampling days	# Samples	Variables
Hyperspectral images	31	8,631	VNIR Spectra: 376-1044nm with 1.22nm interval.
Environmental data	31	8,631	Air temperature (°C) Sun radiation (W/m ²) Wind speed (m/s) Solar zenith angle (degree) Humidity (%) Diurnal time (min)

In addition to the hyperspectral imaging data, a local miniature weather station (Ambient Weather, Chandler, AZ, USA) was installed within the experimental plots to collect real-time time-tagged environmental data. The environmental data included air temperature (°C), solar radiation (W/m²), wind speed (m/s), sun zenith angle (degree), humidity (%), and diurnal time (min) (Table 5.2).



Fig. 5.1. Field VNIR hyperspectral platform at Purdue University. It consists of a VNIR hyperspectral imaging sensor (MSV-101-W, Middleton Spectral Vision, Middleton, WI, USA) and a local weather station (Ambient Weather, Chandler, AZ, USA). The gantry platform is seven meters high capable of scanning all or part of a 50-by-5 meter strip field under a wide range of weather conditions.

5.2.2 Time series decomposition for environment-induced variation

The phenotyping data from 31 days were collected to provide enough images under various environmental conditions. However, besides the instant environmental effects, the plant growth

change and other day-by-day gradual weather changes also contributed to the variation among the images. These different components of variation need to be clearly separated before we can focus on modeling the instant environmental effects. As most of the field environment factors fluctuate over the course of a single day (Weatherley, 1951; Zhou et al., 2017), we hypothesized that the higher frequency environment-induced variation could be identified by removing the lower frequency variation as the day-to-day trend. Thus, a time series decomposition method was applied, decomposing the original time-series phenotyping signal into two major parts: day-to-day trend (T_t) and daily instant changes (D_t) (Eq. 5.2). More specifically, T_t is calculated with the locally estimated scatterplot smoothing (LOESS) method (Rojo et al., 2017). By fitting a non-parametric regression curve on the scattered plot of the data, the day-to-day changing trend can be extracted from the raw signal (Cleveland et al., 1990). This trend mostly reflects the changes in the plant growth stage and general weather conditions over the 31 days of imaging. The daily instant changes (D_t) were calculated by subtracting the day-to-day trend (T_t) from the raw signal. D_t contains the higher frequency variation components mostly caused by the plant's circadian behavior and environmental condition changes such as sun angle, solar radiation, and temperature changes during the day. In this study, D_t was used as the output of the proposed model.

$$\text{Raw time series measurements} = T_t + D_t \quad (5.2)$$

5.2.3 Environmental data transformation and selection

To generate more discriminatory power in higher-dimensional feature spaces besides the original environmental variables (temperature, solar zenith angle, wind speed, etc.) for improved model accuracy, a feature transformation was performed by taking the square and square root of the measurements of the original environment factors (Kusiak, 2001). These new non-linear variables have proven more effective in modeling environment variation (de Souza et al., 2010). Finally, the transformed variables were combined with the original variables for further processing.

After transformation, the features were selected to remove the irrelevant input of some environmental variables to reduce overfitting. This also lowered the difficulty of future applications, with fewer measurements required. A single-factor correlation analysis was performed. Each of the original or transformed environment variables was fitted with the calculated environment-induced variation in phenotyping data (daily instant changes D_t in Eq. 5.2)

in a linear regression model. The adjusted R^2 , which indicates the relevance of each feature to the estimated environmental variation, was calculated. A higher adjusted R^2 means a variable is more relevant (Helland, 1987). By comparing the adjusted R^2 of each variable, we determined the final list of input environmental variables for the model.

5.2.4 Data quality check

The training data quality is critically important for a supervised machine learning model. The data quality was checked before training the model, and the outlier data was removed (Schwertman and de Silva, 2007). For each phenotyping feature (NDVI and predicted RWC, etc.), the daily measurements between the upper inner fence ($Q3+1.5IQR$) and lower inner fence ($Q1-1.5IQR$) were kept (Schwertman and de Silva, 2007). IQR is the interquartile range, which equals the difference between 75th ($Q3$) and 25th ($Q1$) percentiles. Meanwhile, image data before 10:00 am and after 5:30 pm was also removed as it demonstrated extreme variance and noise (Ma et al., 2020). Using the NDVI as an example, we employed the training data sizes shown in Table 5.3 for training the ANN model.

Table 5.3. Data pool after data quality check.

Datasets	Number of samples before the quality check	Number of samples after the quality check
G1H	8,631	5,070
G1M	8,631	5,092
G1L	8,631	5,083
G2H	8,631	5,108
G2M	8,631	5,084
G2L	8,631	5,093
All	51,789	30,530

5.2.5 Artificial Neural network (ANN) model

5.2.5.1 Architecture

The architecture of the proposed model is based on a feed-forward multi-layer perceptron (MLP) network, a class of ANN models (Fig. 5.2). Due to their adjustable architecture, MLP models are particularly flexible. This flexibility increases the suitability of MLP for regression

prediction problems where a real-valued quantity is predicted given a set of inputs such as time-series data (Mellit and Pavan, 2010). In this study, after some speed-accuracy tradeoff pretests on model performance, the proposed ANN architecture is configured with a four-layer model containing an input layer, two hidden layers, and an output layer. After each hidden layer, the Leaky ReLU activation is performed to add non-linear properties to the function (Sharma V. Avinash, 2017). The selected environmental variables serve as input for the model, whereas the D_t environment-induced variation of selected phenotyping features is the output. To accelerate learning and lead to a faster convergence, both input and output data were normalized for modeling purposes (Zhang et al., 2018), while the final prediction results were denormalized back to the original scale of the phenotyping feature.

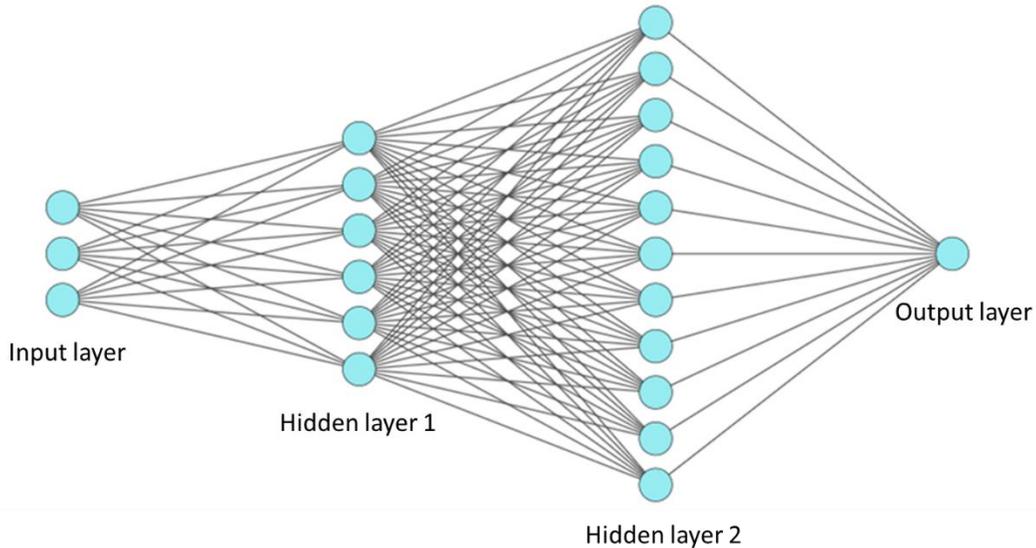


Fig. 5.2. The ANN architecture: Input layer (15 neurons) - Hidden layer 1 (100 neurons, followed with Leaky ReLU) - Hidden layer 2 (1000 neurons, followed with Leaky ReLU) - Output layer (1 neuron).

5.2.5.2 *Training and optimization*

To train the network with minimum overfitting, the training process followed a five-fold cross-validation scenario. We randomly divided the whole dataset into five roughly equal subsets. In the first iteration, the first subset was used to test the model and the rest aided in training the model. This process was repeated until each subset has been used as the testing set. During training, the loss function was Mean Square Error (MSE). The network was initialized with the Kaiming

weights (Tonylins, 2019). All the ANN models were trained using the Adam optimizer (Le et al., 2011).

The accuracy of the model was optimized by adjusting the learning rate, batch size, and the number of epochs. The learning rate controls the rate or speed at which the model learns (Jacobs, 1988). The batch size establishes the accuracy of the error gradient estimate when training neural networks (Gholamrezaei and Ghorbanian, 2007; Livingstone, 2008). The number of epochs impacts the ability of the model to be generalized by determining how many times the model trains on the same data. Finally, by comparing the accuracy (R^2 and RMSE) of models with different combinations of the learning rate, batch size, and number of epochs, the model parameters with a batch size of 600 for 120 epochs with learning rate at $1e-3$ were chosen for this study.

5.2.6 Performance evaluation

5.2.6.1 Evaluation metrics

The performance of all the developed models was evaluated and compared with the coefficient of determination (R^2) and root mean square error (RMSE) between the prediction results and the original measurements. Meanwhile, we also compared the daily variances of the selected phenotyping features (e.g., NDVI) before and after the model correction. A two-sample *t*-test was performed to check if daily variance in features fell significantly.

5.2.6.2 Multi-model comparison analysis across genotypes and nitrogen treatments

The impacts of nitrogen treatments and genotypes on the ANN modeling were also investigated to determine whether a separate ANN model was needed for each case or if one general ANN model could fit the different treatments and genotypes. The ANN model of each experimental plot was trained separately (Table 5.3) and was tested on the other treatments and genotypes. We also built a general ANN model containing the entire sample data to provide a unified and general “all-in-one” correction approach. With the group-to-group cross-validation on each of the datasets, the R^2 and RMSE of each model’s performance in the other datasets were examined to evaluate the generalization of models across genotypes and nitrogen treatments. For example, if the ANN model (ALL) results in similar outcomes as the individual plot models (G1H, G1M, G1L, G2H, G2M, G2L) for each plant plot, this unified model would be adopted. Otherwise,

different models should be adopted separately for each different case. The aim was to find the most appropriate correction solution as the best balance between ease of use and accuracy.

5.2.6.3 Phenotyping features for testing the model and workflow

To demonstrate the detailed modeling procedures and performance evaluation, NDVI was chosen as the first example as it represents one of the most common plant features in remote sensing (Cabrera-Bosquet et al., 2011). We also then tested the same ANN architecture and workflow on the RWC to validate the generalization of the proposed method.

5.2.7 Software and computation

All the imaging processing work was implemented with Matlab R2018a software (MATLAB, 2018). The modeling work was performed in the Python version 3.7.2 software environment (Van Rossum and Drake, 2009). The ANN model was implemented in PyTorch 0.4.1 (Paszke et al., 2017). The time-series data was analyzed and manipulated using Pandas (McKinney, 2010) and Numpy (Oliphant, 2006). The figures were drawn with Seaborn (Waskom et al., 2017) and Matplotlib (Hunter, 2007). The Matlab and Python computations were all executed on a ThinkPad workstation P300 (Lenovo PC international, Morrisville, Morrisville, NC, USA) equipped with 16-gigabytes (GB) of random-access memory (RAM), a 3.70 GHz Intel® Xeon™ E1270 processor and Nvidia GTX 1070 GPU.

5.3 Results and discussion

5.3.1 Time series decomposition result

The time-series data of raw NDVI was successfully decomposed into the day-to-day trend (T_t) and daily instant changes (D_t) (Fig. 5.3). The raw NDVI plot (row 1 in Fig. 5.3) captured the variation in NDVI over the daytime period with gaps indicating the time between 5:30 in the afternoon until 10:00 next day without imaging data. The raw NDVI plot showed a clear and repetitive V-shaped pattern for each day, which was caused by environment variation during imaging. The day-to-day trend T_t (row 2 in Fig. 5.3) represented the changes of plant growth stage and plant health conditions. As plants mature, the NDVI was expected to increase until the reproductive stage. Meanwhile, the two big dips along the T_t curve precisely captured the impacts

from two severe temperature drops in the West Lafayette area. This kind of long-term environmental impact was not included in the proposed analysis.

On the other hand, the daily instant changes (D_t) (row 3 in Fig. 5.3) showed clear periodical changes with V shapes. Due to the extreme weather conditions, parts of the data were missing on DAP 36, 37, 58, 59, and 60. Overall, D_t remained substantially consistent through 31 days, while amplitude and minor skewness differences existed among the D_t from different imaging days. For example, the D_t of DAP 42 demonstrates a smaller amplitude than that of DAP 56. These differences were caused by different environmental conditions, which would be addressed by the environmental correction model in this study.

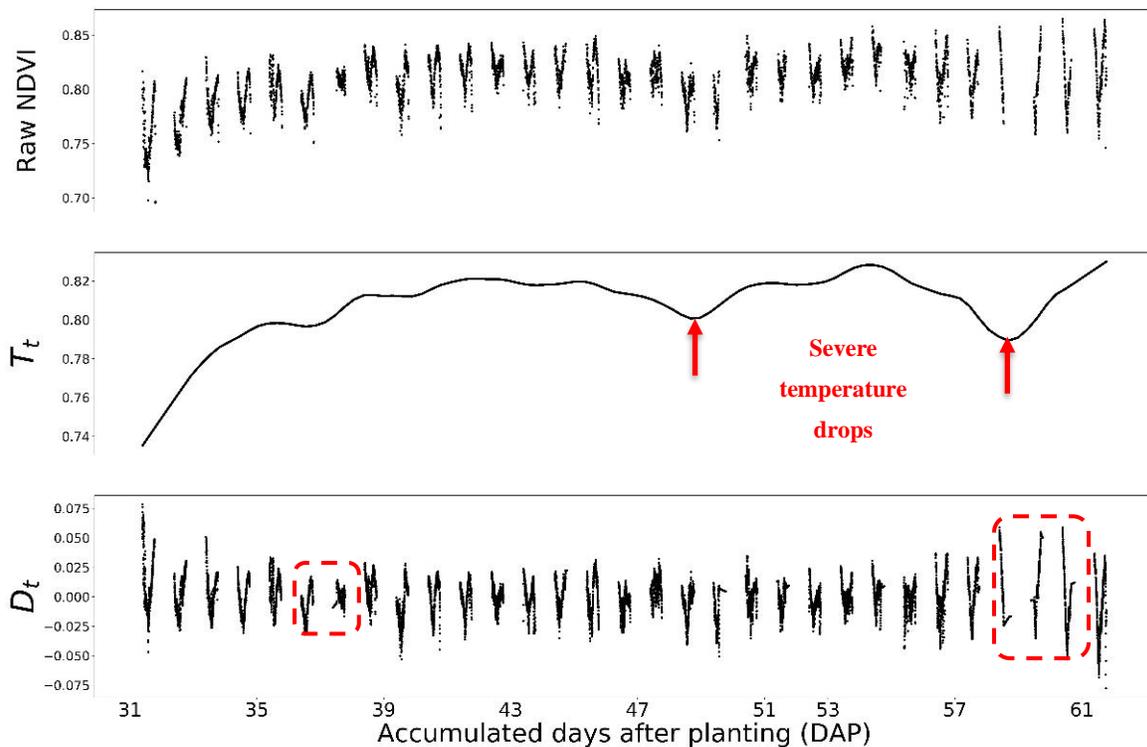


Fig. 5.3. The NDVI of the sample dataset (G1H) from the V4 stage to the R1 stage. The raw NDVI plot was decomposed into the day-to-day trend (T_t) and the periodic change (D_t). The red boxes are the days with incomplete data measurements due to the extreme weather conditions, which are DAP 36, 37, 58, 59 and 60.

5.3.2 Environmental feature selection and range

The results of the single-factor correlation analysis for NDVI are shown in Fig. 5.4. The environmental variables were all correlated with the environment-induced variation in NDVI,

except for humidity and its derivatives. The adjusted R^2 values for humidity were almost 0, indicating no correlation found between humidity and NDVI changes. This confirmed the findings from the previous literature that while plant-sensing data were strongly impacted by environment factors such as air temperature, solar radiation, sun zenith angle, and diurnal time (de Souza et al., 2010; Ishihara et al., 2015; Jones et al., 2009), humidity was rarely reported to demonstrate a similar impact. Thus, we removed humidity and its derivatives in the model. Finally, the input feature for each modeling sample is a 1-by-15 vector consisting of air temperature, solar radiation, wind speed, solar zenith angle, diurnal time and their square or square root values. Besides, the range in environmental conditions experienced by the modeling data was shown in Table 5.4. These ranges illustrate the appropriate environmental condition to apply the model.

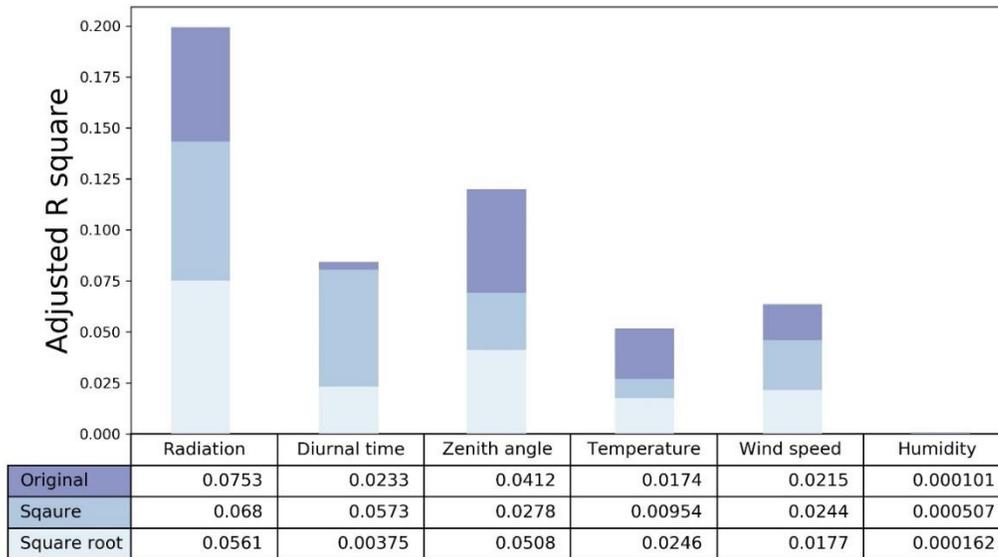


Fig. 5.4. Single-factor correlation analysis for NDVI. The x-axis is the environment factors and their squared or root squared formats. The y-axis is the adjusted R^2 between each x variable and the calculated environment-induced variation in NDVI. Larger adjusted R^2 value means the variable is more correlated to the NDVI variation.

Table 5.4. The ranges in environmental conditions experienced by the modeling data during the 31 days.

Environmental variables	Min	Max
Sun radiation (W/m ²)	85.76	954.23
Diurnal time (min)	600 (at 10am)	1050 (at 5:30pm)
Solar zenith angle (degree)	35.2	78.26
Air temperature (°C)	11.79	33.27
Wind speed (m/s)	0	8.3
Humidity (%)	26.52	97.06

Note: Diurnal time counts from midnight, so the value at midnight is 0 min.

5.3.3 Performance of the ANN models

5.3.3.1 Overall performance

The R^2 and RMSE measure the precision of the predicted environment-induced variation in NDVI. The environment-induced variation predicted by the ANN model for the sample dataset (G1H) showed a fairly accurate linear relationship with the coefficient of determination (R^2) equal to 0.823 (Fig. 5.5). The RMSE also demonstrated a relatively low value of 0.00611. The prediction result was five-fold cross-validated.

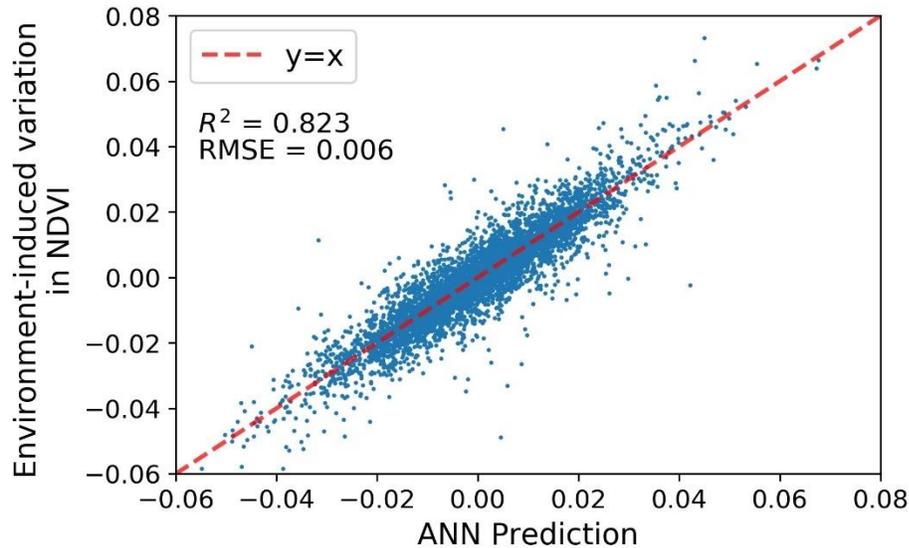


Fig. 5.5. The five-fold cross-validated prediction results of environment-induced variation in NDVI for the sample dataset (G1H). The ANN prediction values show a significant correlation with $R^2=0.823$ and $RMSE=0.006$.

The predicted environmentally induced variation was further used to correct the noise caused by environmental effects in the raw NDVI signal. Fig. 5.6 shows the NDVI corrected by subtracting the predicted variation (Fig. 5.5) from the raw NDVI. In Fig. 5.6b, each box represents the NDVI changes within a day. The trained ANN model largely eliminated the daily variance in the NDVI, so the boxes of the corrected NDVI (Fig. 5.6b) were much more condensed compared with the original NDVI (Fig. 5.6a). To facilitate the comparison, we compared the variances of NDVI before and after model correction with a two-sample *t*-test (Fig. 5.7 and Table 5.5). The result confirmed that the daily variances in NDVI were significantly reduced (p -value < 0.01) by 79% on average, thereby confirming the ability of the proposed ANN model to effectively eliminate the environmentally induced effects on the raw signal.

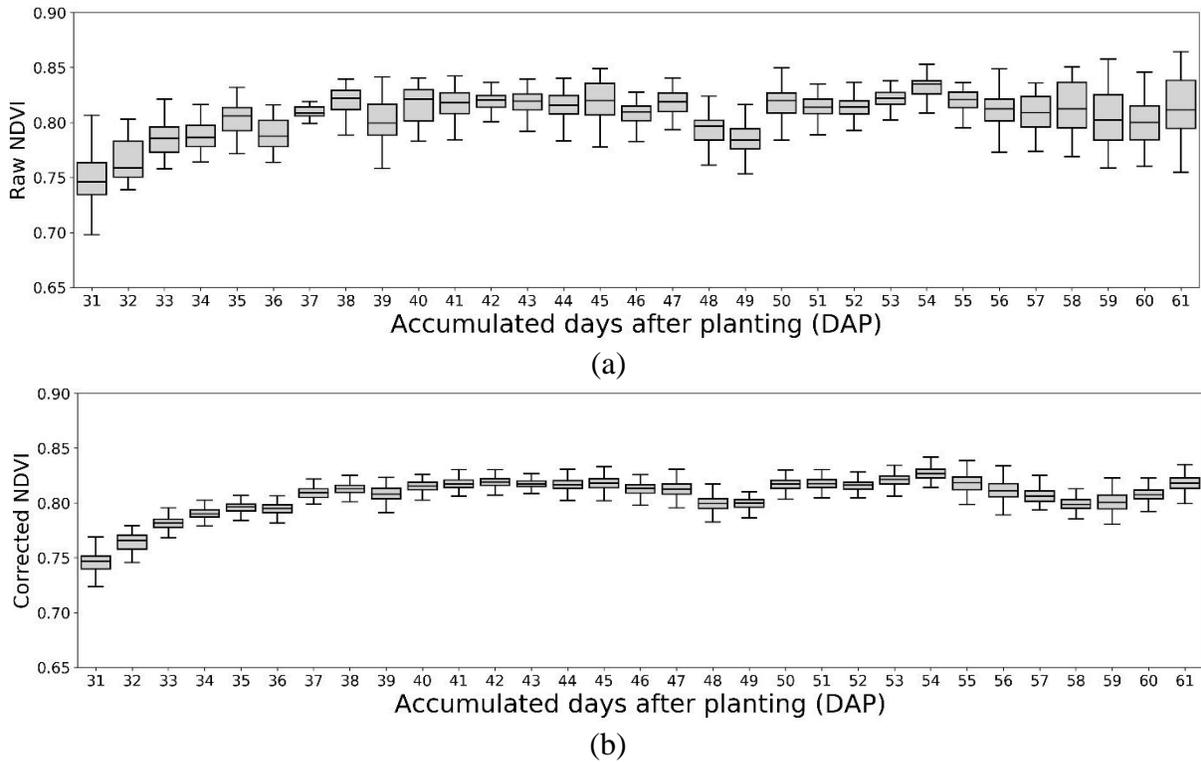


Fig. 5.6. Box plots for the five-fold cross-validated correction result of the sample dataset (G1H). (1). The raw NDVI with huge daily variances across 31 days. (2). The ANN model corrected NDVI with much more condensed boxes.

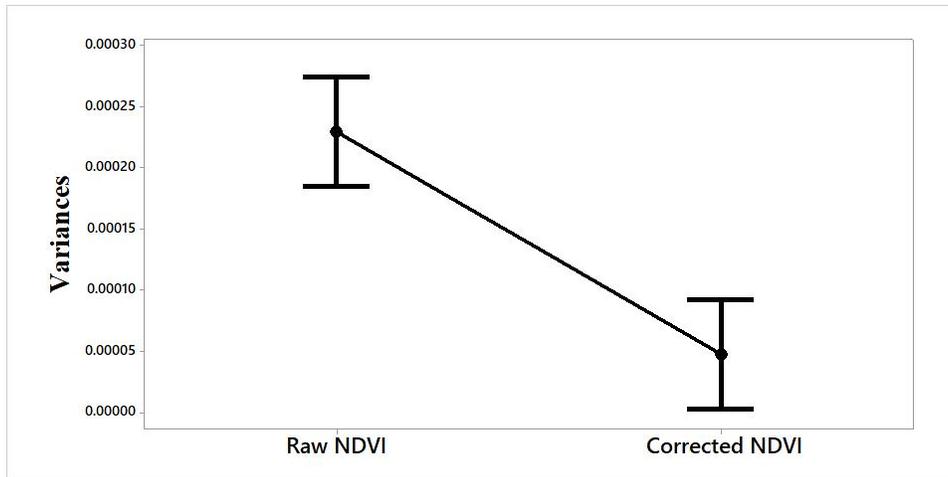


Fig. 5.7. Two-sample t -test between the variance of daily NDVI for the sample dataset (G1H) before and after ANN model correction.

Table 5.5. The results of the two-sample t -test between the variance of daily NDVI before and after ANN model correction.

Groups	N	Mean	StDev	SE Mean	T-value	P-value
Raw NDVI	31	0.000230	0.000174	0.000031	5.78	<0.01
Corrected NDVI	31	0.0000472	0.0000248	0.0000045		

5.3.3.2 Multi-model comparison analysis across genotypes and nitrogen treatments

The ANN models built for each dataset were tested on the other datasets to evaluate the drifts between different genotypes and nutrient treatments. For datasets from a different genotype or treatment, the ANN model demonstrated a weaker prediction performance compared to the results on dataset it had been trained with (Fig. 5.8). Notably, the predictions were least accurate when the ANN models trained with nitrogen-stressed plots (G1M, G2M, G1L, and G2L) were applied to the high-nitrogen groups (G1H and G2H), as shown within the red boxes in the Fig. 5.8. The results of the multi-model comparison indicate that the nitrogen stress levels on plants should be considered when modeling the environment-induced variation in phenotyping features. Compared to the nitrogen treatments, the genotype difference demonstrated a minor impact. The R^2 between plots with the same nitrogen treatment but different genotypes were between 0.59-0.79 with RMSE between 0.009-0.013. The general ANN model (ALL) trained with the entire sample data performed well across the different genotypes and treatments with substantially high R^2 (0.617-0.843) and low RMSE values (0.008-0.0010). This allowed us to

apply the same one ANN model (ALL) for diverse corn stages (already included in the modeling), genotypes, and treatments (validated in Fig 5.8).

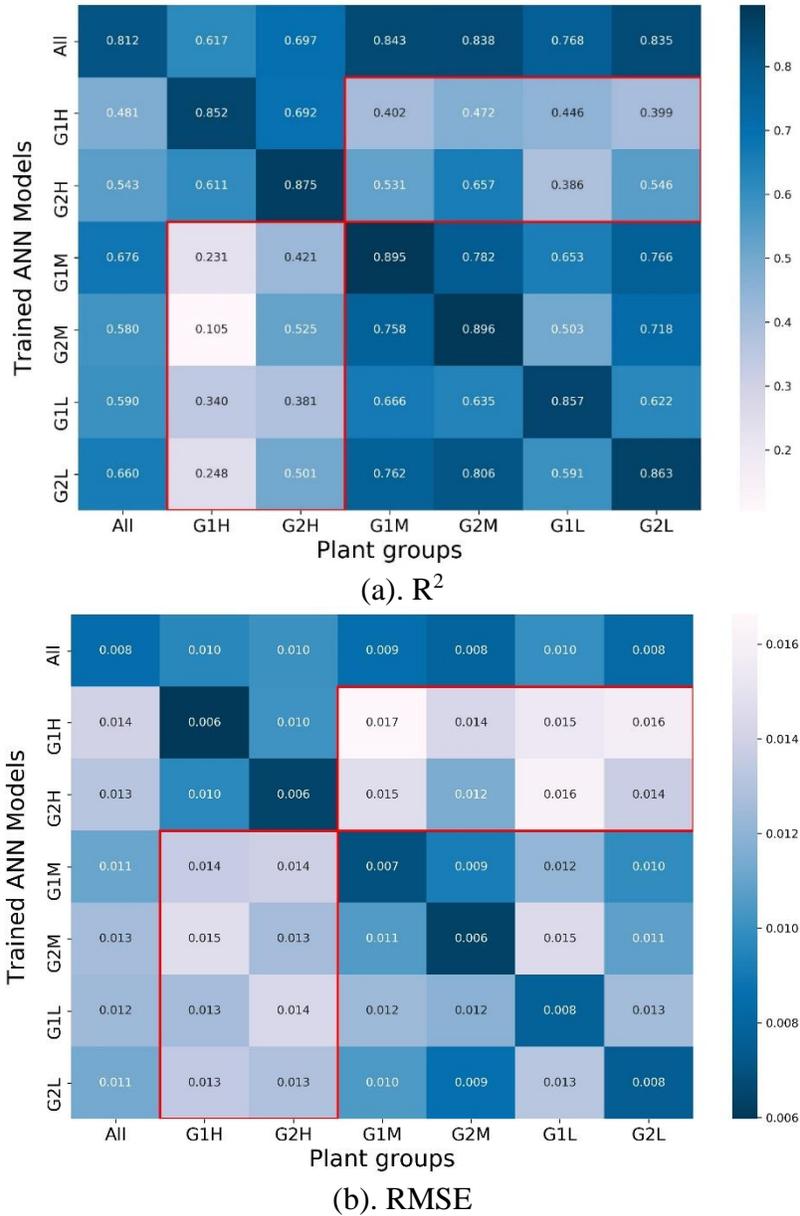


Fig. 5.8. Accuracy heatmaps of R^2 and RMSE of ANN models for NDVI. Red boxes: the region with relatively poor predictive results.

5.3.4 Modeling of environmentally induced variation in predicted RWC

Besides NDVI, the environmentally induced variation in the predicted RWC was also modeled and predicted. In Fig. 5.9, the predicted and the measured variation were strongly

correlated, with a R^2 of 0.791 and RMSE of 0.722%. With this model, the variance of the corrected RWC was significantly reduced (p -value < 0.01) by 72% on average compared to the raw predicted RWC data (Fig. 5.10). The successful application of the same proposed ANN architecture and decomposition method on the predicted RWC and NDVI indicated that this method has the potential to be generally applied on other phenotyping features that can be further explored. Moreover, the corrected predicted RWC plot (Fig. 5.10b) demonstrates a more obvious day-to-day trend than the raw predicted RWC (Fig. 5.10a). Therefore, with the environmental effects removed, plant remote-sensing researchers can more accurately track the plant growth signals.

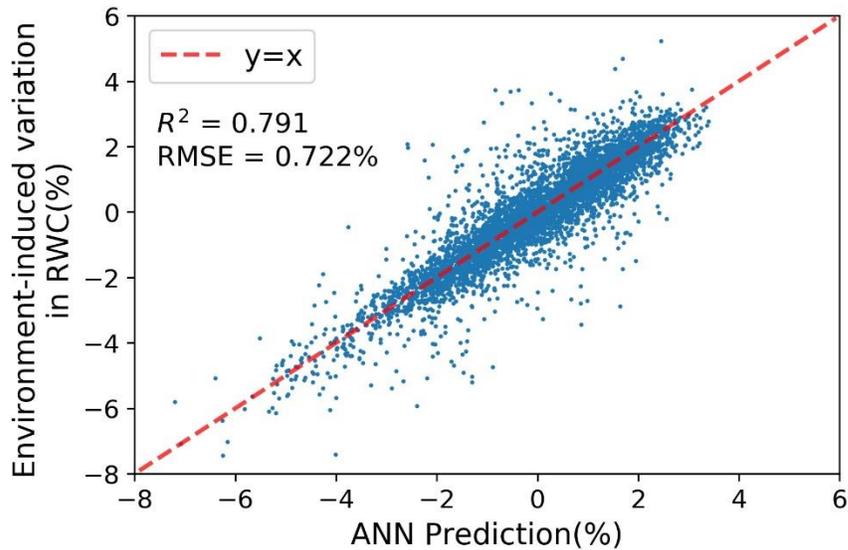


Fig. 5.9. The five-fold cross-validated prediction results of environment-induced variation in predicted RWC for the sample dataset (G1H). The ANN prediction values show a significant performance with $R^2=0.791$ and $RMSE=0.722\%$.

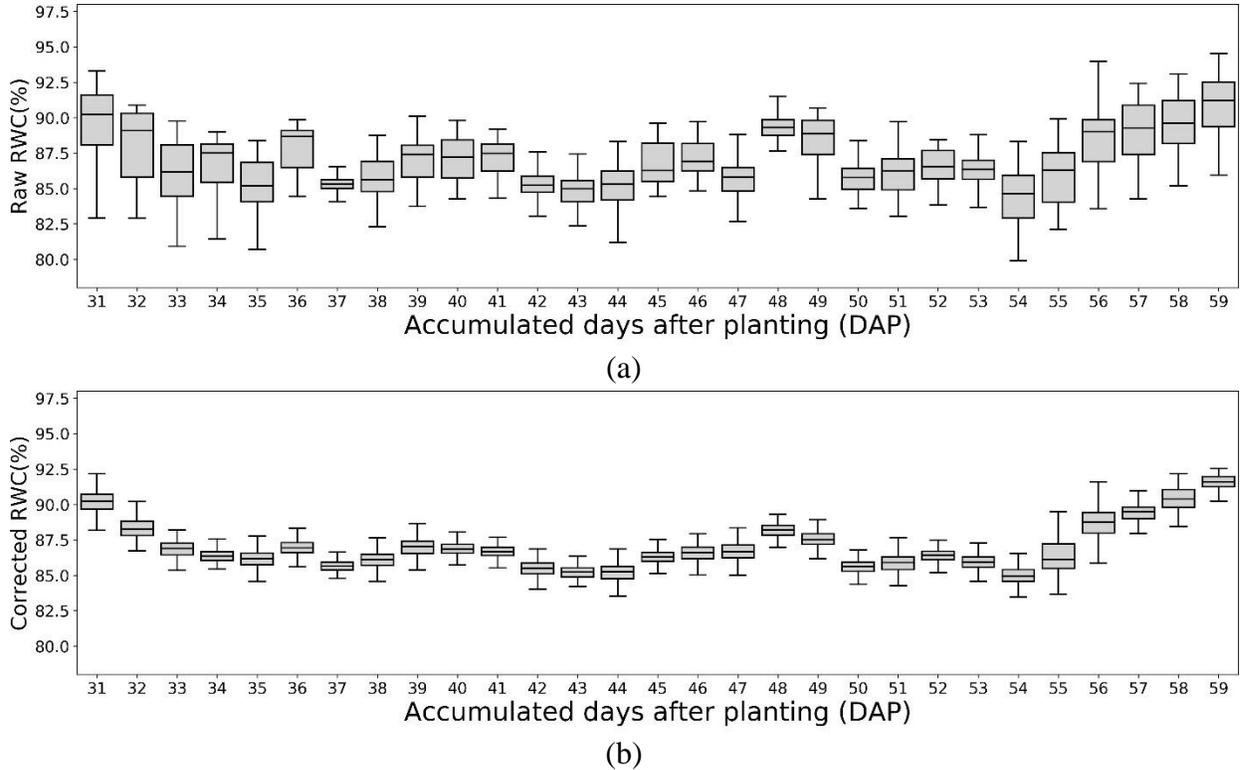


Fig. 5.10. Box plots for the 5-fold cross-validated correction result of sample dataset (G1H). (1). The raw predicted RWC showed huge daily variances across 31 days. (2). The ANN model corrected predicted RWC has much more condensed boxes.

5.4 Conclusions

In this paper, a new modeling method was successfully proposed to precisely predict the environmental effects on the hyperspectral imaging results (such as NDVI and predicted RWC) in airborne crop remote sensing. Over 8,000 hyperspectral images, together with synchronized environment data were collected over 31 days for field corn plants with different nitrogen treatments and genotypes. Experimental results demonstrated that the proposed ANN method could accurately predict the environment-induced variations in the selected phenotyping features. For example, the trained model for NDVI achieved promising predictive results for the sample dataset with an R^2 of 0.822 and an RMSE of 0.00611 compared with the measured variation. The predicted values were used to correct the raw phenotyping data, and the daily variance of NDVI was significantly reduced by 79%. The proposed method also achieved satisfactory results when tested on predicted RWC (daily variance reduced by 72%). The applicability of the proposed method on two different features highlighted its potential to correct the other phenotyping features

of interest. Based on these results, this proposed modeling method can help agricultural remote sensing researchers to effectively eliminate the signal drifts caused by the environment variation, which will drastically increase the accuracy of field plant sensing.

CHAPTER 6. SUMMARY AND FUTURE WORK

6.1 Summary

This thesis proposed methods for analyzing and ameliorating the environmental effects on hyperspectral images for improved phenotyping in greenhouse and field conditions. The proposed methods effectively minimized the unwanted noise in imaging results caused by the environment variation, which drastically improved the quality of plant phenotyping.

A computational simulation model was built to study the greenhouse microclimate changes (such as the temperature and radiation distributions) through a 24-hour cycle in a research greenhouse. With the simulated temperature and radiation profiles over time and space, the heterogenous microclimate in the greenhouse was precisely estimated. The simulation results were validated by comparing them to the ground truth temperature and radiation measurements with the distributed environmental sensors in the greenhouse ($R^2=0.88$ and 0.91 , respectively). The simulation results of temperature and radiation distributions were then utilized to optimize the distance and frequency of pot movement in a greenhouse equipped with an automated conveyor system. The conveyor movement was changed from its original setting of continuous movement to the optimal setting of “10s running phase followed by 8min break in each cycle”, which saved 98% of the conveyor motion while having the same reduction in the heterogenous microclimates.

The details of the design and development of the automated phenotyping greenhouse were introduced. The novel design enables the hyperspectral imager in the greenhouse to scan each individual plant 15 times/day. This greenhouse was specially designed with a conveyor system to automatically shuffle the plants to reduce the heterogenous microclimate impacts. The comparison test between this conveyor greenhouse and a neighboring traditional greenhouse validated the effectiveness of the optimized conveyor movement solution. Plant feature (e.g., canopy size, NDVI, RWC) variances were significantly reduced by 28-83% in the conveyor greenhouse.

Diurnal changing patterns in crop aerial remote sensing images were quantitatively investigated with the proposed novel modeling approach. Over 8000 hyperspectral images of two varieties of corn with three nitrogen treatments were taken by the field imaging gantry at Purdue University over the 2019 growing season. The imaging covered the plant stages from V4 to R1. Crop phenotyping features such as NDVI, RWC, and two individual spectral bands (Red and NIR)

were calculated from the imaging data. The proposed diurnal pattern modeling method successfully described the diurnal changes along the timeline for these phenotyping features. For example, the modeled diurnal changing pattern shows that the NDVI presents a repeatable V-shaped diurnal changing pattern: it linearly decreases by 0.012/h before solar noon and increases by 0.010/h thereafter. Besides NDVI, predicted RWC and the NIR band display linear diurnal changing patterns as well, where predicted RWC changes in an inversed V-shaped pattern and NIR changes in a normal V-shaped pattern. The red band shows a quadratic diurnal changing pattern with an inversed V shape. With the modeling results of this work, remote sensing users can more precisely estimate the deviation or change in crop feature predictions at different imaging time of the day. The diurnal pattern model helps researchers in deciding upon an acceptable imaging time window, and it can also be used to correct/compensate the remote sensing results considering the time effect.

To extend the study of environmental effects on aerial hyperspectral images, an ANN model was trained with synchronized hyperspectral imaging data and environmental data (including sun radiation, solar zenith angle, diurnal time, temperature and wind speed) to understand the correlation of the variations between the two datasets. A time-series decomposition method was applied to extract the phenotyping data variation caused by the changing environments. By learning the relationship between the phenotyping data variation and environmental changes, the developed ANN model was able to precisely predict the environmental effects on remote sensing results (i.e., 82.3% for NDVI), and thus could be used to effectively eliminate the environment-induced variation in the phenotyping features. The two-sample *t*-tests on the NDVI and predicted RWC of corn plants showed that the daily variances in NDVI and predicted RWC were significantly reduced by 79% and 72%, respectively.

6.2 Future work

Though the conveyor greenhouse has proven effective in eliminating the impact of microclimates on plants, further study is still needed to explore the conveyor movement's disturbance on the plant's growth. Future studies could investigate the level of disturbance by comparing the plants who are and are not moved by the conveyor system. Three groups of plants could be grown: one on the conveyor system, one on the ground beside the conveyor, and the last in the conventional greenhouse. Ground truth measurements, including leaf area, RWC, and

SPAD, could be collected and compared to estimate the influence of conveyor movement on plant growth.

The applications of simulation results could also be further explored in the traditional greenhouse when pot movement is impractical. The simulation results could quantitatively estimate the severity of microclimate impact in a conventional greenhouse. Thus, researchers could avoid the locations with extreme environmental impacts on the plants.

The proposed ANN-based method showed a promising performance in modeling the environment-induced variations in different plant phenotyping features. However, the data used in the model was drawn from one single field test whose imaging data was collected from Purdue's field gantry system, which might induce systematic bias in the model. External validation data from the other remote sensing platforms such as UAVs are needed. In the 2020 growing season, the proposed method will be validated with the images from the field UAV system as well as the RGBN camera-based imaging sensor (Ncam) (Wang et al., 2020a). Furthermore, this modeling method was developed based on corn images, which might limit the scope of application. It is necessary to conduct more tests on more diverse plant species (e.g., soybean, wheat, and rice). This will help further validate the developed models, as well as improve the robustness of the prospective models.

In the future, I will also continue exploring training models for all the single spectral bands to adjust/correct the whole spectrum data considering environmental variations. Remote sensing users could benefit from the spectrum calibration model to correct the prediction results from any plant feature prediction models.

REFERENCES

- Aaslyng, J.M., Andreassen, A.U., Allee, H., Holst, N., 2007. Microclimate Prediction for Dynamic Greenhouse Climate Control 42, 272–279.
- Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A.E., Arshad, H., 2018. State-of-the-art in artificial neural network applications: A survey. *Heliyon* 4, e00938. <https://doi.org/10.1016/j.heliyon.2018.e00938>
- Ahemd, H.A., Al-Faraj, A.A., Abdel-Ghany, A.M., 2016. Shading greenhouses to improve the microclimate, energy and water saving in hot regions: A review. *Sci. Hortic. (Amsterdam)*. 201, 36–45. <https://doi.org/10.1016/j.scienta.2016.01.030>
- Ahonen, T., Virrankoski, R., Elmusrati, M., Box, P.O., 2008. Greenhouse Monitoring with Wireless Sensor Network. *IEEE/ASME Int. Conf. Mechtronic Embed. Syst. Appl.* 403–408. <https://doi.org/10.1109/MESA.2008.4735744>
- Alain, B., 1989. Greenhouse microclimate and its management in mild winter climates. *Acta Hortic.*
- Alamar, M.C., Bobelyn, E., Lammertyn, J., Nicolai, B.M., Moltó, E., 2007. Calibration transfer between NIR diode array and FT-NIR spectrophotometers for measuring the soluble solids contents of apple. *Postharvest Biol. Technol.* 45, 38–45. <https://doi.org/10.1016/j.postharvbio.2007.01.008>
- An, N., Welch, S.M., Markelz, R.J.C., Baker, R.L., Palmer, C.M., Ta, J., Maloof, J.N., Weinig, C., 2017. Quantifying time-series of leaf morphology using 2D and 3D photogrammetry methods for high-throughput plant phenotyping. *Comput. Electron. Agric.* 135, 222–232. <https://doi.org/10.1016/j.compag.2017.02.001>
- Ast, J.C., Dunlap, P. V., 2004. Phylogenetic analysis of the lux operon distinguishes two evolutionarily distinct clades of *Photobacterium leiognathi*. *Arch. Microbiol.* 181, 352–361.
- Atkin, O.K., Holly, C., Ball, M.C., 2000. Acclimation of snow gum (*Eucalyptus pauciflora*) leaf respiration to seasonal and diurnal variations in temperature: The importance of changes in the capacity and temperature sensitivity of respiration. *Plant, Cell Environ.* 23, 15–26. <https://doi.org/10.1046/j.1365-3040.2000.00511.x>
- Atrashevskii, Y.I., Sikorskii, A. V, Sikorskii, V. V, Stel'makh, G.F., 1999. The reflection and scattering of light by a plant leaf. *J. Appl. Spectrosc.* 66, 105–114.

- Baille, A., Kittas, C., Katsoulas, N., 2001. Influence of whitening on greenhouse microclimate and crop energy partitioning. *Agric. For. Meteorol.* 107, 293–306. [https://doi.org/10.1016/S0168-1923\(01\)00216-7](https://doi.org/10.1016/S0168-1923(01)00216-7)
- Barbedo, J.G.A., 2019. A Review on the Use of Unmanned Aerial Vehicles and Imaging Sensors for Monitoring and Assessing Plant Stresses. *Drones* 3, 40. <https://doi.org/10.3390/drones3020040>
- Bassi, F.M., Bentley, A.R., Charmet, G., Ortiz, R., Crossa, J., 2015. Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.). *Plant Sci.* 242, 23–36. <https://doi.org/10.1016/j.plantsci.2015.08.021>
- Baxevanou, C., Fidaros, D., Bartzanas, T., Kittas, C., 2010. Numerical simulation of solar radiation, air flow and temperature distribution in a naturally ventilated tunnel greenhouse. *Agric. Eng. Int. CIGR J.* 12, 48–67.
- Bellvert, J., Girona, J., 2012. The use of multispectral and thermal images as a tool for irrigation scheduling in vineyards. *Om.Ciheam.Org* 137, 131–137.
- Beneduzzi, H.M., Souza, E.G., Bazzi, C.L., Schenatto, K., 2017. Temporal variability in active reflectance sensor-measured NDVI in soybean and wheat crops. *Eng. Agric.* 37, 771–781. <https://doi.org/10.1590/1809-4430-Eng.Agric.v37n4p771-781/2017>
- Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F., Hank, T., 2018a. Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: A review study. *Remote Sens.* 10. <https://doi.org/10.3390/rs10010085>
- Berger, K., Atzberger, C., Danner, M., Wocher, M., Mauser, W., Hank, T., 2018b. Model-based optimization of spectral sampling for the retrieval of crop variables with the PROSAIL model. *Remote Sens.* 10. <https://doi.org/10.3390/rs10122063>
- Berndt, D., Clifford, J., 1994. Using dynamic time warping to find patterns in time series. *Work. Knowl. Knowl. Discov. Databases* 398, 359–370.
- Boos, D.D., Stefanski, L.A., 2011. P-value precision and reproducibility. *Am. Stat.* 65, 213–221.
- Brien, C.J., Berger, B., Rabie, H., Tester, M., 2013. Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems. *Plant Methods* 9. <https://doi.org/10.1186/1746-4811-9-5>

- Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* 76, 156–172. [https://doi.org/10.1016/S0034-4257\(00\)00197-8](https://doi.org/10.1016/S0034-4257(00)00197-8)
- Cabrera-Bosquet, L., Molero, G., Stellacci, A., Bort, J., Nogués, S., Araus, J., 2011. NDVI as a potential tool for predicting biomass, plant nitrogen content and growth in wheat genotypes subjected to different water and nitrogen conditions. *Cereal Res. Commun.* 39, 147–159. <https://doi.org/10.1556/CRC.39.2011.1.15>
- Çakir, U., Sahin, E., 2015. Using solar greenhouses in cold climates and evaluating optimum type according to sizing, position and location: A case study. *Comput. Electron. Agric.* 117, 245–257.
- Carter, G.A., 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Int. J. Remote Sens.* 15, 517–520. <https://doi.org/10.1080/01431169408954109>
- Castro, V., Isard, S.A., Irwin, M.E., 1991. The microclimate of maize and bean crops in tropical America: a comparison between monocultures and polycultures planted at high and low density. *Agric. For. Meteorol.* 57, 49–67. [https://doi.org/10.1016/0168-1923\(91\)90078-5](https://doi.org/10.1016/0168-1923(91)90078-5)
- Causse, M., Derivot, L., Pascual, L., Diouf, I.A., Bitton, F., 2018. Water Deficit and Salinity Stress Reveal Many Specific QTL for Plant Growth and Fruit Quality Traits in Tomato. *Front. Plant Sci.* 9, 1–13. <https://doi.org/10.3389/fpls.2018.00279>
- Chappelle, E.W., Kim, M.S., McMurtrey, J.E., 1992. Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. *Remote Sens. Environ.* 39, 239–247. [https://doi.org/10.1016/0034-4257\(92\)90089-3](https://doi.org/10.1016/0034-4257(92)90089-3)
- Cho, M.A., Ramoelo, A., Koch, S., Main, R., Mathieu, R., O’Kennedy, M.M., 2011. An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS J. Photogramm. Remote Sens.* 66, 751–761. <https://doi.org/10.1016/j.isprsjprs.2011.08.001>
- Choab, N., Allouhi, A., El Maakoul, A., Kousksou, T., Saadeddine, S., Jamil, A., 2019. Review on greenhouse microclimate and application: Design parameters, thermal modeling and simulation, climate controlling technologies. *Sol. Energy* 191, 109–137. <https://doi.org/10.1016/j.solener.2019.08.042>

- Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I., 1990. STL: A seasonal-trend decomposition. *J. Off. Stat.* 6, 3–73.
- Clevers, J.G.P.W., Kooistra, L., Schaepman, M.E., 2008. Using spectral information from the NIR water absorption features for the retrieval of canopy water content. *Int. J. Appl. Earth Obs. Geoinf.* 10, 388–397. <https://doi.org/10.1016/j.jag.2008.03.003>
- Conceição, P., Mendoza, R.U., 2009. Anatomy of the global food crisis. *Third World Q.* 30, 1159–1182. <https://doi.org/10.1080/01436590903037473>
- Danner, M., Berger, K., Woche, M., Mauser, W., Hank, T., 2019. Fitted PROSAIL parameterization of leaf inclinations, water content and brown pigment content for winter wheat and maize canopies. *Remote Sens.* 11. <https://doi.org/10.3390/rs11101150>
- Datt, B., 1999. A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using Eucalyptus leaves. *J. Plant Physiol.* 154, 30–36. [https://doi.org/10.1016/S0176-1617\(99\)80314-9](https://doi.org/10.1016/S0176-1617(99)80314-9)
- Daughtry, C.S.T., Walthall, C.L., Kim, M.S., De Colstoun, E.B., McMurtrey Iii, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239.
- de Souza, E.G., Scharf, P.C., Sudduth, K.A., 2010. Sun position and cloud effects on reflectance and vegetation indices of corn. *Agron. J.* 102, 734–744. <https://doi.org/10.2134/agronj2009.0206>
- Di Gennaro, S.F., Rizza, F., Badeck, F.W., Berton, A., Delbono, S., Gioli, B., Toscano, P., Zaldei, A., Matese, A., 2018. UAV-based high-throughput phenotyping to discriminate barley vigour with visible and near-infrared vegetation indices. *Int. J. Remote Sens.* 39, 5330–5344. <https://doi.org/10.1080/01431161.2017.1395974>
- Duan, S.B., Li, Z.L., Wu, H., Tang, B.H., Ma, L., Zhao, E., Li, C., 2014. Inversion of the PROSAIL model to estimate leaf area index of maize, potato, and sunflower fields from unmanned aerial vehicle hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* 26, 12–20. <https://doi.org/10.1016/j.jag.2013.05.007>
- Dunford, R.P., Yano, M., Kurata, N., Sasaki, T., Huestis, G., Rocheford, T., Laurie, D.A., 2002. Comparative mapping of the barley Ppd-H1 photoperiod response gene region, which lies close to a junction between two rice linkage segments. *Genetics* 161, 825–834.

- Eck, R. Van, Klep, M., Schijndel, J. Van, 2016. Surface to Surface Radiation Benchmarks. Proc. 2016 COMSOL Conf. Munich.
- Ehrlich, P.R., Harte, J., 2016. Opinion: To feed the world in 2050 will require a global revolution. Proc. Natl. Acad. Sci. 112, 14743–14744. <https://doi.org/10.1073/pnas.1519841112>
- Fiorani, F., Rascher, U., Jahnke, S., Schurr, U., 2012. Imaging plants dynamics in heterogenic environments. Curr. Opin. Biotechnol. <https://doi.org/10.1016/j.copbio.2011.12.010>
- Fiorani, F., Schurr, U., 2013. Future Scenarios for Plant Phenotyping. Annu. Rev. Plant Biol. 64, 267–291. <https://doi.org/10.1146/annurev-arplant-050312-120137>
- Frei, W., 2016. Thermal Modeling of the Air Flow Inside and Around Your House [WWW Document]. URL <https://www.comsol.com/blogs/thermal-modeling-of-the-air-flow-inside-and-around-your-house/>
- Gamon, J.A., Kovalchuck, O., Wong, C.Y.S., Harris, A., Garrity, S.R., 2015. Monitoring seasonal and diurnal changes in photosynthetic pigments with automated PRI and NDVI sensors. Biogeosciences 12, 4149–4159. <https://doi.org/10.5194/bg-12-4149-2015>
- Garaba, S.P., Dierssen, H.M., 2018. An airborne remote sensing case study of synthetic hydrocarbon detection using short wave infrared absorption features identified from marine-harvested macro- and microplastics. Remote Sens. Environ. 205, 224–235. <https://doi.org/10.1016/j.rse.2017.11.023>
- Gardner, B.R., 1985. TECHNIQUES FOR REMOTELY MONITORING CANOPY DEVELOPMENT AND ESTIMATING GRAIN YIELD OF MOISTURE STRESSED CORN (LANDSAT, SENSED).
- Gatebe, C.K., King, M.D., 2016. Airborne spectral BRDF of various surface types (ocean, vegetation, snow, desert, wetlands, cloud decks, smoke layers) for remote sensing applications. Remote Sens. Environ. 179, 131–148.
- 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. Comput. Electron. Agric. 127, 625–632. <https://doi.org/10.1016/j.compag.2016.07.028>

- Gehan, M.A., Fahlgren, N., Abbasi, A., Berry, J.C., Callen, S.T., Chavez, L., Doust, A.N., Feldman, M.J., Gilbert, K.B., Hodge, J.G., Hoyer, J.S., Lin, A., Liu, S., Lizárraga, C., Lorence, A., Miller, M., Platon, E., Tessman, M., Sax, T., 2017. PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ* 5, e4088. <https://doi.org/10.7717/peerj.4088>
- Gewali, U.B., Monteiro, S.T., Saber, E., 2018. Machine learning based hyperspectral image analysis: A survey.
- Gholamrezaei, M., Ghorbanian, K., 2007. Rotated general regression neural network. *IEEE Int. Conf. Neural Networks - Conf. Proc.* 2, 1959–1964. <https://doi.org/10.1109/IJCNN.2007.4371258>
- Gitelson, A., Merzlyak, M.N., 1994. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* 22, 247–252. [https://doi.org/10.1016/1011-1344\(93\)06963-4](https://doi.org/10.1016/1011-1344(93)06963-4)
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS- MODIS. *Remote Sens. Environ.* 58, 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)
- Glenn, K.C., Ward, J.M., Goley, M., Vicini, J.L., Bell, E., Parrott, W., Alsop, B., Martin, C., Sparks, O., Liu, B., Souder, C., Urquhart, W., Jenkinson, J., 2017. Bringing New Plant Varieties to Market: Plant Breeding and Selection Practices Advance Beneficial Characteristics while Minimizing Unintended Changes. *Crop Sci.* 57, 2906. <https://doi.org/10.2135/cropsci2017.03.0199>
- Golzarian, M.R., Frick, R.A., Rajendran, K., Berger, B., Roy, S., Tester, M., Lun, D.S., 2011. Accurate inference of shoot biomass from high-throughput images of cereal plants 1–11.
- Gonzalez-Real, M.M., Baille, A., 2000. Changes in leaf photosynthetic parameters with leaf position and nitrogen content within a rose plant canopy (*Rosa hybrida*). *Plant Cell Environ.* 23, 351–363.
- Gracia-Romero, A., Kefauver, S.C., Fernandez-Gallego, J.A., Vergara-Díaz, O., Nieto-Taladriz, M.T., Araus, J.L., 2019. UAV and ground image-based phenotyping: A proof of concept with durum wheat. *Remote Sens.* 11. <https://doi.org/10.3390/rs11101244>
- Greenham, K., Lou, P., Remsen, S.E., Farid, H., McClung, C.R., 2015. TRiP: Tracking Rhythms in Plants, an automated leaf movement analysis program for circadian period estimation. *Plant Methods* 11, 1–11. <https://doi.org/10.1186/s13007-015-0075-5>

- Gupta, R.K., 2001. New Hyperspectral Vegetation. Ratio 201–206.
- Gupta, R.K., Vijayan, D., Prasad, T.S., 2003. Comparative analysis of red-edge hyperspectral indices. *Adv. Sp. Res.* 32, 2217–2222. [https://doi.org/10.1016/S0273-1177\(03\)90545-X](https://doi.org/10.1016/S0273-1177(03)90545-X)
- 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 Sens. Environ.* 90, 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426. [https://doi.org/10.1016/S0034-4257\(02\)00018-4](https://doi.org/10.1016/S0034-4257(02)00018-4)
- Hartung, J., Wagener, J., Ruser, R., Piepho, H.P., 2019. Blocking and re-arrangement of pots in greenhouse experiments: which approach is more effective? *Plant Methods* 15, 1–11. <https://doi.org/10.1186/s13007-019-0527-4>
- Helland, I.S., 1987. On the interpretation and use of R² in regression analysis. *Biometrics* 61–69.
- Herve, D., Berrios, E., Leroux, N., Chaarani, G. Al, Planchon, C., Sarrafi, A., Gentzbittel, L., 2001. QTL analysis of photosynthesis and water status traits in sunflower (*Helianthus annuus* L.) under greenhouse conditions. *J. Exp. Bot.* 52, 1857–1864.
- Hickey, J.M., Chiurugwi, T., Mackay, I., Powell, W., 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nat. Genet.* 49, 1297–1303. <https://doi.org/10.1038/ng.3920>
- Honkavaara, E., Arbiol, R., Markelin, L., Martinez, L., Cramer, M., Bovet, S., Chandelier, L., Ilves, R., Klonus, S., Marshal, P., Schläpfer, D., Tabor, M., Thom, C., Veje, N., 2009. Digital airborne photogrammetry-a new tool for quantitative remote sensing?-a state-of-the-art review on radiometric aspects of digital photogrammetric images. *Remote Sens.* 1, 577–605. <https://doi.org/10.3390/rs1030577>
- Houle, D., Govindaraju, D.R., Omholt, S., 2010. Phenomics: The next challenge. *Nat. Rev. Genet.* 11, 855–866. <https://doi.org/10.1038/nrg2897>
- Huete, A., Justice, C., Liu, H., 1994. Developmeclassification and soil indices for MODIS-EOS. *Remote Sens. Environ.* 49, 224–234. [https://doi.org/Doi.10.1016/0034-4257\(94\)90018-3](https://doi.org/Doi.10.1016/0034-4257(94)90018-3)
- Hunter, J.D., 2007. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95.

- Ishida, H., Oishi, Y., Morita, K., Moriwaki, K., Nakajima, T.Y., 2017. Remote Sensing of Environment Development of a support vector machine based cloud detection method for MODIS with the adjustability to various conditions.
- Ishihara, M., Inoue, Y., Ono, K., Shimizu, M., Matsuura, S., 2015. The impact of sunlight conditions on the consistency of vegetation indices in croplands-Effective usage of vegetation indices from continuous ground-based spectral measurements. *Remote Sens.* 7, 14079–14098. <https://doi.org/10.3390/rs71014079>
- Jackson, R.D., Pinter, P.J., Jr., S.B.I., R.H.Reginato, 1979. Wheat spectral reflectance: Interactions between crop configuration, sun elevation, and azimuth angle. *Appl. Opt.* 18, 3730–3732.
- Jacobs, R.A., 1988. Increased rates of convergence through learning rate adaptation. *Neural networks* 1, 295–307.
- Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. *Remote Sens. Environ.* 34, 75–91. [https://doi.org/10.1016/0034-4257\(90\)90100-Z](https://doi.org/10.1016/0034-4257(90)90100-Z)
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sens. Environ.* 113, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>
- Jannink, J.L., Sorrells, M.E., Lorenz, A.J., Heffner, E.L., 2010. Plant breeding with genomic selection: Gain per unit time and cost. *Crop Sci.* <https://doi.org/10.2135/cropsci2009.11.0662>
- Ji, W., Viscarra Rossel, R.A., Shi, Z., 2015. Improved estimates of organic carbon using proximally sensed vis-NIR spectra corrected by piecewise direct standardization. *Eur. J. Soil Sci.* 66, 670–678. <https://doi.org/10.1111/ejss.12271>
- Jones, H.G., Serraj, R., Loveys, B.R., Xiong, L., Wheaton, A., Price, A.H., 2009. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. *Funct. Plant Biol.* 36, 978–989.
- Kate, R.J., 2016. Using dynamic time warping distances as features for improved time series classification. *Data Min. Knowl. Discov.* 30, 283–312.
- Kattenborn, T., 2019. Linking Canopy Reflectance and Plant Functioning through Radiative Transfer Models.
- Kimball, B.A., 1973. Simulation of the energy balance of a greenhouse. *Agric. Meteorol.* 11, 243–260. [https://doi.org/10.1016/0002-1571\(73\)90067-8](https://doi.org/10.1016/0002-1571(73)90067-8)

- Kitchen, N.R., Goulding, K.W.T., Follett, R.F., Hatfield, J.L., 2001. On-farm technologies and practices to improve nitrogen use efficiency. *Nitrogen Environ. Sources, Probl. Manag.* Elsevier, Amsterdam, Netherlands.
- Kläring, H.-P., Hauschild, C., Heißner, A., Bar-Yosef, B., 2007. Model-based control of CO₂ concentration in greenhouses at ambient levels increases cucumber yield. *Agric. For. Meteorol.* 143, 208–216.
- Kloosterman, B., Abelenda, J.A., Gomez, M.D.M.C., Oortwijn, M., De Boer, J.M., Kowitzanich, K., Horvath, B.M., Van Eck, H.J., Smaczniak, C., Prat, S., Visser, R.G.F., Bachem, C.W.B., 2013. Naturally occurring allele diversity allows potato cultivation in northern latitudes. *Nature* 495, 246–250. <https://doi.org/10.1038/nature11912>
- Korioth, T.W.P., Versluis, A., 1997. Modeling the Mechanical Behavior of the Jaws and Their Related Structures By Finite Element (Fe) Analysis. *Crit. Rev. Oral Biol. Med.* 8, 90–104. <https://doi.org/10.1177/10454411970080010501>
- Körner, O., Aaslyng, J.M., Andreassen, A.U., Holst, N., 2007. Microclimate prediction for dynamic greenhouse climate control. *HortScience*.
- Krishna, K.R., 2018. *Agricultural drones: a peaceful pursuit.* Taylor & Francis.
- Kristan, M., Leonardis, A., Skočaj, D., 2011. Multivariate online kernel density estimation with Gaussian kernels. *Pattern Recognit.* 44, 2630–2642. <https://doi.org/10.1016/j.patcog.2011.03.019>
- Kusiak, A., 2001. Feature transformation methods in data mining. *IEEE Trans. Electron. Packag. Manuf.* 24, 214–221.
- Le, Q. V, Ngiam, J., Coates, A., Lahiri, A., Prochnow, B., Ng, A.Y., 2011. On optimization methods for deep learning.
- Lee, G.J., Boerma, H.R., Villagarcia, M.R., Zhou, X., Carter, T.E., Li, Z., Gibbs, M.O., 2004. A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. *Theor. Appl. Genet.* 109, 1610–1619. <https://doi.org/10.1007/s00122-004-1783-9>
- LI, J., ZOU, Z., YANG, X., CHEN, X., 2003. Analysis and discussion and application on solar greenhouse about azimuth angle and lighting surface angles. *Acta Agric. Boreali-occidentalis Sin.* 2.
- Li, L., Zhang, Q., Huang, D., 2014. A Review of Imaging Techniques for Plant Phenotyping 20078–20111. <https://doi.org/10.3390/s141120078>

- Li, X., Cai, W., Shao, X., 2015. Correcting multivariate calibration model for near infrared spectral analysis without using standard samples. *J. Near Infrared Spectrosc.* 23, 285–291. <https://doi.org/10.1255/jnirs.1165>
- Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S., Lin, H., 2015. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. *Remote Sens. Environ.* 165, 123–134. <https://doi.org/10.1016/j.rse.2015.04.032>
- Liang, S., Strahler, A.H., 1994. Retrieval of surface BRDF from multiangle remotely sensed data. *Remote Sens. Environ.* 50, 18–30.
- Lindquist, J.L., Arkebauer, T.J., Walters, D.T., Cassman, K.G., Dobermann, A., 2005. Maize radiation use efficiency under optimal growth conditions. *Agron. J.* 97, 72–78. <https://doi.org/10.2134/agronj2005.0072>
- Livingstone, D.J., 2008. Artificial neural networks: methods and applications. Springer.
- Ma, D., Carpenter, N., Amatya, S., Maki, H., Wang, L., Zhang, L., Neeno, S., Tuinstra, M.R., Jin, J., 2019a. Removal of greenhouse microclimate heterogeneity with conveyor system for indoor phenotyping. *Comput. Electron. Agric.* 166, 104979. <https://doi.org/10.1016/j.compag.2019.104979>
- Ma, D., Carpenter, N., Maki, H., Rehman, T.U., Tuinstra, M.R., Jin, J., 2019b. Greenhouse environment modeling and simulation for microclimate control. *Comput. Electron. Agric.* 162, 134–142. <https://doi.org/10.1016/j.compag.2019.04.013>
- Ma, D., Rehman, T.U., Zhang, L., Jin, J., 2020. Modeling of diurnal changing patterns in airborne crop remote sensing images. *Remote S* 1–18.
- Maji, S., Chandra, B., Viswavidyalaya, K., 2014. Diurnal Variation in Spectral Properties of Potato under Different Dates of Planting and N-Doses Diurnal Variation in Spectral Properties of Potato under Different Dates of Planting and N-Doses.
- Manea, D., Calin, M.A., 2015. Hyperspectral imaging in different light conditions. *Imaging Sci. J.* 63, 214–219. <https://doi.org/10.1179/1743131X15Y.0000000001>
- Marshak, A., Knyazikhin, Y., Davis, A.B., Wiscombe, W.J., Pilewskie, P., 2000. Cloud - vegetation interaction: Use of normalized difference cloud index for estimation of cloud optical thickness. *Geophys. Res. Lett.* 27, 1695–1698. <https://doi.org/10.1029/1999GL010993>
- MATLAB, 2018. version 9.4.0.813654 (R2018a). The MathWorks Inc., Natick, Massachusetts.

- McCartney, L., Orsat, V., Lefsrud, M.G., 2018. An experimental study of the cooling performance and airflow patterns in a model Natural Ventilation Augmented Cooling (NVAC) greenhouse. *Biosyst. Eng.* 174, 173–189.
- McKinney, W., 2010. Data structures for statistical computing in python, in: *Proceedings of the 9th Python in Science Conference*. pp. 51–56.
- Mellit, A., Pavan, A.M., 2010. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Sol. Energy* 84, 807–821. <https://doi.org/10.1016/j.solener.2010.02.006>
- Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63, 282–293. <https://doi.org/10.1016/j.compag.2008.03.009>
- Miklas, P.N., Delorme, R., Riley, R., 2019. Identification of QTL Conditioning Resistance to White Mold in Snap Bean. *J. Am. Soc. Hortic. Sci.* <https://doi.org/10.21273/jashs.128.4.0564>
- Miura, T., Huete, A.R., 2009. Performance of three reflectance calibration methods for airborne hyperspectral spectrometer data. *Sensors* 9, 794–813. <https://doi.org/10.3390/s90200794>
- Mokarram, M., Bijanzadeh, E., 2016. Prediction of biological and grain yield of barley using multiple regression and artificial neural network models. *Aust. J. Crop Sci.* 10, 895–903. <https://doi.org/10.21475/ajcs.2016.10.06.p7634>
- Neuwirthová, E., Lhotáková, Z., Albrechtová, J., 2017. The effect of leaf stacking on leaf reflectance and vegetation indices measured by contact probe during the season. *Sensors (Switzerland)* 17. <https://doi.org/10.3390/s17061202>
- Ni, Z., Liu, Z., Huo, H., Li, Z.L., Nerry, F., Wang, Q., Li, X., 2015. Early water stress detection using leaf-level measurements of chlorophyll fluorescence and temperature data. *Remote Sens.* 7, 3232–3249. <https://doi.org/10.3390/rs70303232>
- Norton, T., Sun, D.-W., Grant, J., Fallon, R., Dodd, V., 2007. Applications of computational fluid dynamics (CFD) in the modelling and design of ventilation systems in the agricultural industry: A review. *Bioresour. Technol.* 98, 2386–2414.
- Oliphant, T.E., 2006. *A guide to NumPy*. Trelgol Publishing USA.
- Oliveira, L.F., Scharf, P.C., 2014. Diurnal variability in reflectance measurements from cotton. *Crop Sci.* 54, 1769–1781. <https://doi.org/10.2135/cropsci2013.04.0217>

- Ollinger, S. V., 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.* 189, 375–394. <https://doi.org/10.1111/j.1469-8137.2010.03536.x>
- Padilla, F.M., de Souza, R., Peña-Fleitas, M.T., Grasso, R., Gallardo, M., Thompson, R.B., 2019. Influence of time of day on measurement with chlorophyll meters and canopy reflectance sensors of different crop N status. *Precis. Agric.* 20, 1087–1106. <https://doi.org/10.1007/s11119-019-09641-1>
- Pandey, P., Ge, Y., Stoerger, V., Schnable, J.C., 2017. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. *Front. Plant Sci.* 8, 1–12. <https://doi.org/10.3389/fpls.2017.01348>
- Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A., 2017. Automatic differentiation in PyTorch, in: NIPS-W.
- Peirone, L.S., Pereyra Irujo, G.A., Bolton, A., Erreguerena, I., Aguirrezábal, L.A.N., 2018. Assessing the efficiency of phenotyping early traits in a greenhouse automated platform for predicting drought tolerance of soybean in the field. *Front. Plant Sci.* 9, 1–14. <https://doi.org/10.3389/fpls.2018.00587>
- Prasad, A.K., Chai, L., Singh, R.P., Kafatos, M., 2006. Crop yield estimation model for Iowa using remote sensing and surface parameters. *Int. J. Appl. Earth Obs. Geoinf.* 8, 26–33. <https://doi.org/10.1016/j.jag.2005.06.002>
- Preuss, S.B., Meister, R., Xu, Q., Urwin, C.P., Tripodi, F.A., Screen, S.E., Anil, V.S., Zhu, S., Morrell, J.A., Liu, G., Ratcliffe, O.J., Reuber, T.L., Khanna, R., Goldman, B.S., Bell, E., Ziegler, T.E., McClerren, A.L., Ruff, T.G., Petracek, M.E., 2012. Expression of the *Arabidopsis thaliana* BBX32 gene in soybean increases grain yield. *PLoS One* 7. <https://doi.org/10.1371/journal.pone.0030717>
- Qi, J., Kerr, Y.H., Moran, M.S., Weltz, M., Huete, A.R., Sorooshian, S., Bryant, R., 2000. Leaf area index estimates using remotely sensed data and BRDF models in a semiarid region. *Remote Sens. Environ.* 73, 18–30.
- Quemada, M., Gabriel, J.L., Zarco-Tejada, P., 2014. Airborne hyperspectral images and ground-level optical sensors as assessment tools for maize nitrogen fertilization. *Remote Sens.* 6, 2940–2962. <https://doi.org/10.3390/rs6042940>

- Rahman, M.M., Lamb, D.W., Stanley, J.N., 2015. The impact of solar illumination angle when using active optical sensing of NDVI to infer fAPAR in a pasture canopy. *Agric. For. Meteorol.* 202, 39–43. <https://doi.org/10.1016/j.agrformet.2014.12.001>
- Ranson, K.J., Daughtry, C.S.T., Biehl, L.L., Bauer, M.E., 1985. Sun-view angle effects on reflectance factors of corn canopies. *Remote Sens. Environ.* [https://doi.org/10.1016/0034-4257\(85\)90045-8](https://doi.org/10.1016/0034-4257(85)90045-8)
- Ray, D.K., Mueller, N.D., West, P.C., Foley, J.A., 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0066428>
- Reflectance, S., 1973. *Materials*, 2: Application and Results 12.
- Rehman, T.U., Mahmud, M.S., Chang, Y.K., Jin, J., Shin, J., 2019. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Comput. Electron. Agric.* 156, 585–605. <https://doi.org/10.1016/j.compag.2018.12.006>
- Ribaut, J.-M., Ragot, M., 2019. Modernising breeding for orphan crops: tools, methodologies, and beyond. *Planta* 250, 971–977. <https://doi.org/10.1007/s00425-019-03200-8>
- Rojo, J., Rivero, R., Romero-Morte, J., Fernández-González, F., Pérez-Badia, R., 2017. Modeling pollen time series using seasonal-trend decomposition procedure based on LOESS smoothing. *Int. J. Biometeorol.* 61, 335–348. <https://doi.org/10.1007/s00484-016-1215-y>
- Roman, A., Ursu, T.-M., 2016. Multispectral Satellite Imagery and Airborne Laser Scanning Techniques for the Detection of Archaeological Vegetation Marks. *Landsc. Archaeol. North. Front. Rom. Emp. Porolissum* 141–152.
- Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7)
- Roujean, J.L., Breon, F.M., 1995. Estimating Par Absorbed By Vegetation From Bidirectional Reflectance Measurements. *Remote Sens. Environ.* 51, 375–384. [https://doi.org/10.1016/0034-4257\(94\)00114-3](https://doi.org/10.1016/0034-4257(94)00114-3)
- Sadeghi, S., Sohrabi, H., 2019. the Effect of Uav Flight Altitude on the Accuracy of Individual Tree Height Extraction in a Broad-Leaved Forest. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-4/W18*, 1168–1173. <https://doi.org/10.5194/isprs-archives-xlii-4-w18-1168-2019>

- Sagan, V., Maimaitijiang, M., Sidike, P., Maimaitiyiming, M., Erkbol, H., Hartling, S., Peterson, K.T., Peterson, J., Burken, J., Fritschi, F., 2019. Uav/satellite multiscale data fusion for crop monitoring and early stress detection. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 42, 715–722. <https://doi.org/10.5194/isprs-archives-XLII-2-W13-715-2019>
- Schafleitner, R., Gutierrez, R., Espino, R., Gaudin, A., Pérez, J., Martínez, M., Domínguez, A., Tincopa, L., Alvarado, C., Numberto, G., Bonierbale, M., 2007. Field screening for variation of drought tolerance in *Solanum tuberosum* L. by agronomical, physiological and genetic analysis. *Potato Res.* 50, 71–85. <https://doi.org/10.1007/s11540-007-9030-9>
- Schultz, B.B., Schultz, B.B., 1985. Levene's test for relative variation. *Syst. Zool.* 34, 449–456.
- Schwertman, N.C., de Silva, R., 2007. Identifying outliers with sequential fences. *Comput. Stat. Data Anal.* <https://doi.org/10.1016/j.csda.2006.01.019>
- Sharma V. Avinash, 2017. Understanding Activation Functions in Neural Networks. *Medium* 4, 1–10.
- Sticksel, E., Schächtl, J., Huber, G., Liebler, J., Maidl, F.X., 2004. Diurnal variation in hyperspectral vegetation indices related to winter wheat biomass formation. *Precis. Agric.* 5, 509–520. <https://doi.org/10.1007/s11119-004-5322-0>
- Stone, P.J., Sorensen, I.B., Wilson, D.R., 1998. Radiation interception accounts for the effects of plant population on maize yield. *Proc. 28th Annu. Conf. Agron. Soc. New Zeal.* 28, 9–10.
- Stone, V., Montoya, C., Beaver, J.S., Johnson, E., Miklas, P.N., Zapata, M., 2010. Selective Mapping of QTL Conditioning Disease Resistance in Common Bean. *Crop Sci.* <https://doi.org/10.2135/cropsci1996.0011183x003600050044x>
- Stylinski, C., Gamon, J., Oechel, W., 2002. Seasonal patterns of reflectance indices, carotenoid pigments and photosynthesis of evergreen chaparral species. *Oecologia* 131, 366–374.
- Taki, M., Ajabshirchi, Y., Ranjbar, S.F., Rohani, A., Matloobi, M., 2016. Modeling and experimental validation of heat transfer and energy consumption in an innovative greenhouse structure. *Inf. Process. Agric.* 3, 157–174. <https://doi.org/10.1016/j.inpa.2016.06.002>
- Tanaka, J., Hayashi, T., Iwata, H., 2016. A practical, rapid generation-advancement system for rice breeding using simplified biotron breeding system. *Breed. Sci.* 66, 542–551. <https://doi.org/10.1270/jsbbs.15038>

- Thenkabail, P.S., Smith, R.B., Pauw, E. De, De Pauw, E., 1997. Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sens. Environ.* 4257, 158–182. [https://doi.org/10.1016/S0034-4257\(99\)00067-X](https://doi.org/10.1016/S0034-4257(99)00067-X)
- Tian, J., Philpot, W.D., 2015. Relationship between surface soil water content, evaporation rate, and water absorption band depths in SWIR reflectance spectra. *Remote Sens. Environ.* 169, 280–289. <https://doi.org/10.1016/j.rse.2015.08.007>
- Tindall, A.J., Waller, J., Greenwood, M., Gould, P.D., Hartwell, J., Hall, A., 2015. A comparison of high-throughput techniques for assaying circadian rhythms in plants. *Plant Methods* 11, 1–7. <https://doi.org/10.1186/s13007-015-0071-9>
- Tonylins, 2019. A PyTorch implementation of MobileNetV2 1–9.
- Torada, A., Ikeguchi, S., Koike, M., 2005. Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. *Euphytica* 143, 251–255. <https://doi.org/10.1007/s10681-005-7872-2>
- Turner, A., 2007. The Pseudo-Response Regulator Ppd-H1 Provides The Pseudo-Response Regulator Ppd-H1 Provides Adaptation to Photoperiod in Barley. *Science (80-.)*. 1031, 1031–1035. <https://doi.org/10.1126/science.1117619>
- Turner, N.C., 1981. Techniques and experimental approaches for the measurement of plant water status. *Plant Soil* 58, 339–366.
- Van Rossum, G., Drake, F.L., 2009. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.
- Vásquez, L., Iriarte, A., Almeida, M., Villalobos, P., 2015. Evaluation of greenhouse gas emissions and proposals for their reduction at a university campus in Chile. *J. Clean. Prod.* 108, 924–930. <https://doi.org/10.1016/j.jclepro.2015.06.073>
- Vogelmann, J.E., Rock, B.N., Moss, D.M., 1993. Red edge spectral measurements from sugar maple leaves. *Int. J. Remote Sens.* 14, 1563–1575. <https://doi.org/10.1080/01431169308953986>
- Wang, L., Jin, J., Song, Z., Wang, J., Zhang, L., Rehman, T.U., Ma, D., Carpenter, N.R., Tuinstra, M.R., 2020. LeafSpec: An accurate and portable hyperspectral corn leaf imager. *Comput. Electron. Agric.* 169, 105209. <https://doi.org/10.1016/j.compag.2019.105209>
- Wang, R., Cherkauer, K., Bowling, L., 2016. Corn Response to Climate Stress Detected with Satellite-Based NDVI Time Series. <https://doi.org/10.3390/rs8040269>

- Wang, T., Rostamza, M., Song, Z., Wang, L., McNickle, G., Iyer-Pascuzzi, A.S., Qiu, Z., Jin, J., 2019. SegRoot: A high throughput segmentation method for root image analysis. *Comput. Electron. Agric.* 162, 845–854. <https://doi.org/10.1016/j.compag.2019.05.017>
- Waskom, M., Botvinnik, O., O’Kane, D., Hobson, P., Lukauskas, S., Gemperline, D.C., Augspurger, T., Halchenko, Y., Cole, J.B., Warmenhoven, J., de Ruiter, J., Pye, C., Hoyer, S., Vanderplas, J., Villalba, S., Kunter, G., Quintero, E., Bachant, P., Martin, M., Meyer, K., Miles, A., Ram, Y., Yarkoni, T., Williams, M.L., Evans, C., Fitzgerald, C., Brian, Fannesbeck, C., Lee, A., Qalieh, A., 2017. mwaskom/seaborn: v0.8.1 (September 2017). <https://doi.org/10.5281/zenodo.883859>
- Weatherley, P.E., 1951. Studies in the Water Relations of the Cotton Plant. II. Diurnal and Seasonal Variations in Relative Turgidity and Environmental Factors. *New Phytol.* 50, 36–51.
- Yilmaz, I., Kaynar, O., 2011. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Syst. Appl.* 38, 5958–5966. <https://doi.org/10.1016/j.eswa.2010.11.027>
- Zhang, L., Maki, H., Ma, D., Sánchez-Gallego, J.A., Mickelbart, M. V., Wang, L., Rehman, T.U., Jin, J., 2019a. Optimized angles of the swing hyperspectral imaging system for single corn plant. *Comput. Electron. Agric.* 156, 349–359. <https://doi.org/10.1016/j.compag.2018.11.030>
- Zhang, L., Wang, L., Wang, J., Song, Z., Rehman, T.U., Bureetes, T., Ma, D., Chen, Z., Neeno, S., Jin, J., 2019b. Leaf Scanner: A portable and low-cost multispectral corn leaf scanning device for precise phenotyping. *Comput. Electron. Agric.* 167, 105069. <https://doi.org/10.1016/j.compag.2019.105069>
- Zhang, X., Sugano, Y., Bulling, A., 2018. Revisiting data normalization for appearance-based gaze estimation. *Eye Track. Res. Appl. Symp.* <https://doi.org/10.1145/3204493.3204548>
- Zhang, Z., Masjedi, A., Zhao, J., Crawford, M.M., 2017. Prediction of sorghum biomass based on image based features derived from time series of UAV images. *Int. Geosci. Remote Sens. Symp.* 2017-July, 6154–6157. <https://doi.org/10.1109/IGARSS.2017.8128413>
- Zhao, L., Liu, Z., Xu, S., He, X., Ni, Z., Zhao, H., Ren, S., 2018. Retrieving the diurnal FPAR of a maize canopy from the jointing stage to the tasseling stage with vegetation indices under different water stresses and light conditions. *Sensors (Switzerland)* 18. <https://doi.org/10.3390/s18113965>

Zhou, X., Xu, Y., Zhang, F., 2017. Evaluation of effect of diurnal ambient temperature range on solar chimney power plant performance. *Int. J. Heat Mass Transf.* 115, 398–405. <https://doi.org/10.1016/j.ijheatmasstransfer.2017.07.051>