

SOYBEAN LEAF CHLOROPHYLL ESTIMATION AND IRON DEFICIENCY FIELD RATING  
DETERMINATION AT PLOT AND FIELD SCALES THROUGH IMAGE PROCESSING AND  
MACHINE LEARNING

A Dissertation  
Submitted to the Graduate Faculty  
of the  
North Dakota State University  
of Agriculture and Applied Science

By

Oveis Hassanijalilian

In Partial Fulfillment of the Requirements  
for the Degree of  
DOCTOR OF PHILOSOPHY

Major Program:  
Agricultural and Biosystems Engineering

July 2020

Fargo, North Dakota

# NORTH DAKOTA STATE UNIVERSITY

Graduate School

---

## Title

SOYBEAN LEAF CHLOROPHYLL ESTIMATION AND IRON DEFICIENCY FIELD  
RATING DETERMINATION AT PLOT AND FIELD SCALES THROUGH IMAGE  
PROCESSING AND MACHINE LEARNING

---

## By

Oveis Hassanijalilian

---

The supervisory committee certifies that this dissertation complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

## SUPERVISORY COMMITTEE:

Dr. Igathinathane Cannayen

Chair

Dr. Hans Kandel

Dr. Stephanie Day

Mr. John Nowatzki

Approved:

07/09/2020

Date

Dr. Kenneth Hellevang

Department Chair

## ABSTRACT

Iron deficiency chlorosis (IDC) is the most common reason for chlorosis in soybean (*Glycine max* (L.) Merrill) and causes an average yield loss of 120 million dollars per year across  $1.8 \times 10^6$  ha in the North Central US alone. As the most effective way to avoid IDC is the use of tolerant cultivars, they are visually rated for IDC by experts; however, this method is subjective and not feasible for a larger scale. An alternate more objective image processing method can be implemented in various platforms and fields. This approach relies on a color vegetation index (CVI) that can quantify chlorophyll, as chlorophyll content is a good IDC indicator. Therefore, this research is aimed at developing image processing methods at leaf, plot, and field scales with machine learning methods for chlorophyll and IDC measurement. This study also reviewed and synthesized the IDC measurement and management methods. Smartphone digital images with machine learning models successfully estimated the chlorophyll content of soybean leaves infield. Dark green color index (DGCI) was the best-correlated CVI with chlorophyll. The pixel count of four different ranges of DGCI (RPC) was used as input features for different models, and the support vector machine produced the highest performance. Handheld camera images of soybean plots extracted DGCI, which mimicked visual rating, and canopy size that were used as inputs to decision-tree based models for IDC classification. The AdaBoost model was the best model in classifying IDC severity. Unmanned aerial vehicle soybean IDC cultivar trial fields images extracted DGCI, canopy size, and their product (CDP) for IDC severity monitoring and yield prediction. The area under the curve (AUC) was employed to aggregate the data into a single value through

time, and the correlation between all the features and yield was good. Although CDP at latest growth stage had the highest correlation with yield, AUC of CDP was the most consistent index for soybean yield prediction. This research demonstrated that digital image processing along with the machine learning methods can be successfully applied to the soybean IDC measurement and the various soybean related stakeholders can benefit from this research.

## ACKNOWLEDGEMENTS

I would like to express my great appreciation to my major advisor Dr. Igathinathane Cannayen, without whom I could not finish this PhD dissertation and research publications thereof. He accepted me midway in my program, for which I am highly grateful, and taught me several things related to research especially how to write scientifically sound articles, and to create dazzling visualization to convey research results. I learned to do my research with scrutiny not because others will judge our work, but because this should be our standard for conducting research. I watched his interactions with other graduate students while teaching and advising, and I learned how to be a better teacher myself. I still need a lot to learn, but I am in a better place now after working with him. I am thankful for all the times during the night and cold weather that he chose to spend on my dissertation. Words are incapable of conveying my gratitude towards him, and I wish Dr. Cannayen the best in his personal and professional career.

I would love to thank all of my committee members who had a significant effect on my research and dissertation. I would like to thank Dr. Stephanie Day who taught me GIS, one of the most useful skills in my PhD career and beyond. I used GIS through this dissertation and could work on some great projects in DuPont Pioneer. She was always there to help me even with my lack of knowledge for job interviews. I would like to thank Dr. Hans Kandel who I always enjoyed talking to. The individual study course that I had with him helped me to better understand my research area. Dr. Kandel always pushed me to finish my PhD and not let go, and I am grateful to have him as a mentor. I would love to thank Mr. John Nowatzki,

who secured the funding for my projects and hired me as a PhD student. He supported me through all my PhD with all the questions and favors that I asked him. He was always there to listen to my challenges and give me insight and hope on how to move forward.

Dr. Ted Helms taught me valuable lessons regarding my dissertation, and provided me with necessary data to complete this research. He was always welcoming and dedicated time to teach me complex topics in both agriculture and statistics. I appreciate all he did for me and wish him the best in his retirement.

My friends are undoubtedly a big part of my life in the US who became a family away from home. I would love to thank Paul and Liz Bervick, Brent and Liz Kemp, Nick and Grace Braaksma, as well as Thomas and Katie Banley who trusted me and let me into their lives and treated me as part of their family. They provide me with a place that I could call a "Free Country" where I felt loved and I could share my sadness and happiness with them. I would also love to show my gratitude towards Clay, Nina, Jason, and Allison Pettys who accepted me as their relatives and had me over for different occasions and celebrations. I would also like to thank Renae, Jeff, and Val Lange who helped me and supported me during my trips to Bismarck and treated me like one of their family members.

My other friends including Abbas Atefi, Hamed Sajadikia, Seyedali Hosseinirad, Fereshteh Shahoveisi, Mohsen Tahmasebi Nasab, Mojtaba Ahmadi, Mohsen Azimi, Seyedmojtaba Mansouripour, Nakisa Alinejhad, Seyed Ali Haji Esmaeili, Farnaz Namayandeh, Babak Mamnoon, Parinaz Ghanbari, Sunny Prashant, and Katie Olson who all made memorable memories for me and supported me through my marriage and my PhD. I would like to express my thanks to Greg and Stephanie Haugen who helped me in times of need.

Additionally, Jan Nowatzki who made my wife and I felt welcomed all the time with her lovely and welcoming attitude.

Special thanks to Dr. Sunoj Shajahan and Ms. Subhashree N. Srinivasagan, PhD student, of our research group for their help and encouraging words. Thanks are also due to Mr. Brock Zenzen, Dr. Alimohammad Shirzadifar, Dr. Mohammadmehdi Maharlooei, Mr. Austin Eidi, and Dr. Saravanan Sivarajan for their friendship and moral support.

I would like to thank Dr. Sreekala Bajwa who gave me the opportunity to start my PhD in the US. I also would like to thank Melanie Ziegler, Julie Bietz, Debra Baer, Sara Ogundolani, Sheldon Tuscherer, James Moos, Dr. Dean Steele, Dr. Thomas Scherer, and other ABEN faculty and staff for helping me and listening to my personal and professional problems. Additionally, this research could not even start without the great opportunity and funding by the North Dakota Soybean Council and NDSU Agricultural Experiment Station.

At last, I would like to show my appreciation to my parents who supported me throughout my life and supported my US education. My lovely wife, Mrs. Samira Nadali, who patiently supported me and motivated me through my hard times during my entire research journey. She made it easy to move to different states as we found new opportunities for work and internship. She is calm and she made my stressful life calmer by her presence. Samira will be one happy person and much relieved after the successful completion of my PhD, and I am grateful for her continued support with love.

# DEDICATION

To my wife

Mrs. Samira Nadali



# TABLE OF CONTENTS

ABSTRACT . . . . .	iii
ACKNOWLEDGEMENTS . . . . .	v
DEDICATION . . . . .	viii
LIST OF TABLES . . . . .	xiii
LIST OF FIGURES . . . . .	xiv
LIST OF ABBREVIATIONS . . . . .	xvi
LIST OF SYMBOLS . . . . .	xix
LIST OF APPENDIX FIGURES . . . . .	xxi
1. GENERAL INTRODUCTION . . . . .	1
1.1. Significance of the Problem . . . . .	6
1.2. Statement of Hypothesis . . . . .	7
1.3. Statement of Objectives . . . . .	7
1.4. Dissertation Organization . . . . .	8
2. IRON DEFICIENCY CHLOROSIS MEASUREMENT IN SOYBEAN — A REVIEW * . . . .	11
2.1. Abstract . . . . .	11
2.2. Introduction . . . . .	12
2.3. Number of Publications on Soybean IDC and Nature of Work . . . . .	14
2.4. Measurement of IDC . . . . .	14
2.4.1. Visual Rating . . . . .	14
2.4.2. Atomic Absorption Spectrometry . . . . .	17
2.4.3. Wet Chemistry — Indirect through Chlorophyll . . . . .	19
2.4.4. Sensors — SPAD — Indirect through Chlorophyll . . . . .	21
2.4.5. Image Processing — Digital Images and Rating Iron Deficiency . . . . .	23
2.4.6. Other Imaging Platforms . . . . .	29
2.5. Analysis of IDC Data . . . . .	33

2.5.1. Statistical Methods . . . . .	33
2.5.2. Advanced Machine Learning Methods . . . . .	34
2.6. Management of IDC . . . . .	35
2.7. Concluding Remarks and Future Perspective . . . . .	36
3. CHLOROPHYLL ESTIMATION IN SOYBEAN LEAVES INFIELD WITH SMARTPHONE DIGITAL IMAGING AND MACHINE LEARNING * . . . . .	38
3.1. Abstract . . . . .	38
3.2. Introduction . . . . .	39
3.3. Materials and Methods . . . . .	42
3.3.1. Field Experimental Plots . . . . .	42
3.3.2. Field Data Collection — SPAD Meter and Digital Images . . . . .	43
3.3.3. Overall Methodology and Analysis . . . . .	44
3.3.4. Image Preprocessing for Objects Recognition and Indices Determination	47
3.3.5. Color Vegetation Indices . . . . .	51
3.3.6. Machine Learning Models . . . . .	54
3.4. Results and Discussion . . . . .	58
3.4.1. Field Measurement Results . . . . .	58
3.4.2. DGCI for Chlorophyll Measurement . . . . .	59
3.4.3. Machine Learning Prediction Models Results . . . . .	63
3.5. Conclusions . . . . .	73
4. RATING IRON DEFICIENCY IN SOYBEAN USING IMAGE PROCESSING AND DECISION-TREE BASED MODELS * . . . . .	75
4.1. Abstract . . . . .	75
4.2. Introduction . . . . .	76
4.3. Materials and Methods . . . . .	79
4.3.1. Field Experimental Plots . . . . .	79
4.3.2. Overall Methodology . . . . .	79
4.3.3. IDC Visual Rating . . . . .	81
4.3.4. Image Acquisition . . . . .	82

4.3.5.	Object Recognition in Field Image . . . . .	84
4.3.6.	Machine Learning Models . . . . .	91
4.3.7.	Imbalance Dataset Oversampling Technique . . . . .	94
4.3.8.	Performance Assessment of Machine Learning Models . . . . .	95
4.4.	Results and Discussion . . . . .	97
4.4.1.	Decision Tree . . . . .	100
4.4.2.	Random Forest . . . . .	103
4.4.3.	AdaBoost Model Performance . . . . .	104
4.4.4.	Models Input Features and their Importance . . . . .	105
4.5.	Conclusions . . . . .	108
5.	MEASURING SOYBEAN IRON DEFICIENCY CHLOROSIS PROGRESSION AND YIELD PREDICTION WITH UNMANNED AERIAL VEHICLE * . . . . .	110
5.1.	Abstract . . . . .	110
5.2.	Introduction . . . . .	111
5.3.	Materials and Methods . . . . .	115
5.3.1.	Field Experimental Plots . . . . .	115
5.3.2.	Image Acquisition . . . . .	115
5.3.3.	Orthomosaic Creation . . . . .	119
5.3.4.	Image Processing for Color Calibration and Plots Extraction . . . . .	120
5.3.5.	Object Recognition and Dataset Creation . . . . .	123
5.3.6.	Feature Engineering . . . . .	125
5.4.	Results and Discussion . . . . .	127
5.4.1.	Attributes of Soybean IDC Trial Plots . . . . .	127
5.4.2.	IDC Progress Through DGCI Trends for Different Cultivars . . . . .	128
5.4.3.	DGCI Variation on Yield . . . . .	131
5.4.4.	CS Variation on Yield . . . . .	133
5.4.5.	CDP Variation on Yield . . . . .	135
5.4.6.	AUC Variation on Yield . . . . .	135
5.5.	Conclusions . . . . .	139

6. GENERAL CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK . . . . . 141

    6.1. General Conclusions . . . . . 141

    6.2. Suggestions for Future Work . . . . . 143

REFERENCES . . . . . 144

APPENDIX A. FUNCTION TO DETECT THE LEAF, GREEN AND YELLOW DISK . . . . . 161

APPENDIX B. FUNCTIONS FOR NESTED CROSS VALIDATION AND RANDOMIZED  
GRID SEARCH . . . . . 163

APPENDIX C. FUNCTION TO DETECT TWO MIDDLE ROWS . . . . . 169

APPENDIX D. AERIAL IMAGE FEATURES RELATIONSHIP WITH YIELD FOR DIFFERENT  
DATES . . . . . 173

APPENDIX E. SCRIPTS FOR DATA VISUALIZATION THROUGH R AND PYTHON . . . . . 176

## LIST OF TABLES

Table	Page
2.1. Developed visual rating score guidelines for IDC symptoms <sup>1</sup> . . . . .	18
2.2. Correlation of color vegetation indices with SPAD index using RGB and chromatic coordinates (HSB). . . . .	30
3.1. Distribution of the RGB values collected with smartphone camera and SPAD meter readings for both sampling dates ( $n = 330$ ). . . . .	59
3.2. Correlation of color vegetation indices with SPAD index using RGB and chromatic coordinates (HSB). . . . .	60
3.3. Parameter estimates for fixed effects of random coefficients model and comparison. . . . .	63
3.4. Comparison of models and inputs with using Duncan's mean separation analysis. . . . .	71
3.5. Recommended machine learning models to predict the SPAD meter reading from the field image-based inputs for chlorophyll measurement. . . . .	72
4.1. IDC symptoms and the developed visual rating score guidelines <sup>1</sup> . . . . .	82
4.2. Distribution of the input features extracted from images for both sampling dates ( $n = 320$ total). . . . .	98
4.3. Performance of all machine learning models for both dates. . . . .	102
4.4. Feature importance of the machine learning models through mean decrease impurity factor. . . . .	106
5.1. Distribution of the measured and calculated attributes of soybean IDC trial plots at different ND locations and years. . . . .	128
5.2. Distribution of the DGCI for each flight on different locations and dates. . . . .	130
5.3. Soybean yield prediction models to predict yield based on different variables and combination. . . . .	139

## LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
2.1. Number of publications in Web of Science academic database for different keywords used in the advanced search in the website. . . . .	15
2.2. Iron deficiency chlorosis (IDC) visual score for the scale of 1-5. . . . .	17
2.3. Soybean plots and their iron deficiency chlorosis visual scores. The middle plot shows a high amount of error in its score comparing to the plots on its both sides. . .	18
2.4. Histogram of $a^*$ channel of an aerial image of soybean research plot. . . . .	27
3.1. Soybean experimental field plots for various cultivars and replications showing various levels of IDC symptoms. . . . .	43
3.2. Flowchart of the proposed image processing and machine learning model approach developed to predict chlorophyll content of soybean leaves. . . . .	45
3.3. Soybean leaf placed on the standard calibration board. . . . .	46
3.4. Image preprocessing operations of object recognition and isolation of the standard calibration board, and the yellow and the green discs for extraction of color information. . . . .	48
3.5. Detecting the pink portion of the standard calibration board through red and green histograms. . . . .	49
3.6. The extracted leaf sample (top-left) from the standard calibration board image (fig. 3.4) and a few others showing different levels of chlorophyll in the V3 to V5 stages of the study. . . . .	52
3.7. Scatterplot of corrected DGCI and SPAD meter readings using RCM for first (left) and second (right) sampling dates. . . . .	64
3.8. Developed machine learning models comparison based on the coefficient of determination ( $R^2$ ). . . . .	68
3.9. Developed machine learning models comparison based on RMSE (SPAD units). . . .	69
4.1. Portion of the experimental plots showing the plot design and the existing variation in soybean crop IDC. . . . .	80
4.2. Process flowchart of field image processing and machine learning models development and validation with manual rating for IDC classification. . . . .	81
4.3. A sample image of soybean plot showing the two middle rows with standard color board at two different growing stages (Left: V2-V3 and right: V5-V6). . . . .	83

4.4. Histogram of $a^*$ component of $L^*a^*b^*$ colorspace showing different objects in the image. . . . .	85
4.5. Object recognition processes from field plot images of detecting standard color board and its green and yellow disks. . . . .	87
4.6. Binary image creation and artifacts removal using $a^*$ component of $L^*a^*b^*$ colorspace. . . . .	89
4.7. Identification and isolation of the plot middle rows from border rows and other plots.	90
4.8. Soybean plots iron deficiency chlorosis (IDC) progression as difference between both dates based on visual scores and image-based on average dark green color index (DGCI*) scores. . . . .	99
4.9. Distribution of image-based average dark green color index (DGCI*) among different MR for both dates. . . . .	100
4.10. Decision criteria of the decision tree model. . . . .	101
4.11. Comparison of machine learning models tested showing the f1-score of Date-2 (July 17, 2014), as this date is more representative of the variation in iron deficiency chlorosis (IDC) classes. . . . .	105
5.1. Overall flow diagram of the data collection and analysis followed in this study. . . . .	116
5.2. Aerial image of Hunter, ND, USA trial site (July 14, 2017).. . . . .	117
5.3. Checkerboard that was used for color correction. . . . .	120
5.4. Histogram of plants and background with Otsu threshold to extract plants of the plot. . . . .	122
5.5. Polygonal regions of interest arrangement over soybean plot, and each color is representative of a different replication (40 plots $\times$ 4 replications). . . . .	124
5.6. ArcMap ModelBuilder flowchart to derive dark green color index (DGCI) values from each plot. . . . .	125
5.7. The dark green color index (DGCI) trend among common cultivars in 2016 and 2017 at different locations. . . . .	129
5.8. Unmanned aerial vehicle orthomosaic images of Leonard, ND. . . . .	132
5.9. Unmanned aerial vehicle image extracted parameters correlation with yield in 2016 and 2017 at different experiment locations. . . . .	133
5.10. Yield correlations with area under the curve (AUC) of parameters at different locations and years. . . . .	136
5.11. Unmanned aerial vehicle image combined data extracted parameters variation with yield and models for both years and locations. . . . .	138

## LIST OF ABBREVIATIONS

AAS	Atomic Absorption Spectrometry
ACDP	AUC of Canopy $\times$ DGCI Product
ACS	AUC of Canopy Size
AD	AUC of DGCI
AdaBoost	Adaptive Boosting
AGL	Above Ground Level
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
AUC	Area Under the Curve
CC	Chlorophyll Content
CDP	Canopy $\times$ DGCI Product
CDT	Central Daylight Time
CIVE	Color Index of Vegetation Extraction
CMOS	Complementary Metal-Oxide-Semiconductor
CS	Canopy Size
CV	Cross Validation
CVI	Color Vegetation Index
DGCI	Dark Green Color Index
DN	Digital Number
DS	Date of Sampling
EC	Electrical Conductivity
EGI	Excess Green Index
ERI	Green Minus Red
ESRI	Environmental Systems Research Institute
EVI	Enhanced Vegetation Index



ExG . . . . .	Excess Green Index
ExGR . . . . .	Excess Green Minus Excess Red Index
FCDPC . . . . .	Four Classes of DGCI Pixel Counts
FN . . . . .	False Negative
FP . . . . .	False Positive
GB . . . . .	Green Minus Blue
GBRG . . . . .	GB by RG Ratio
GNDVI . . . . .	Green Normalized Difference Vegetation Index
GR . . . . .	Green Minus Red
HSB . . . . .	Hue, Saturation, and Brightness
HSV . . . . .	Hue, Saturation, and Value
IDC . . . . .	Iron Deficiency Chlorosis
JPEG . . . . .	Joint Photographic Experts Group
KNN . . . . .	K-Nearest Neighborhood
LAI . . . . .	Leaf Area Index
LiDAR . . . . .	Light Detection and Ranging
MexG . . . . .	Modified Excess Green Index
MLR . . . . .	Multiple Linear Regression
MP . . . . .	Mega Pixel
MTVI . . . . .	Modified Triangular Vegetation Index
ND . . . . .	North Dakota
NDI . . . . .	Normalized Difference Index
NDSU . . . . .	North Dakota State University
NDVI . . . . .	Normalized Difference Vegetation Index
NIR . . . . .	Near-Infrared
OSAVI . . . . .	Optimized Soil Adjusted Vegetation Index
PR . . . . .	Polynomial Regression
RCM . . . . .	Random Coefficients Model
RF . . . . .	Random Forest

RG . . . . .	Red Minus Green
RGB . . . . .	Red, Green, and Blue
RMSE . . . . .	Root Mean Square Error
ROI . . . . .	Region of Interest
SLR . . . . .	Statistical Linear Regression
SMOTE . . . . .	Synthetic Minority Oversampling Technique
SMR . . . . .	SPAD Meter Reading
SPAD . . . . .	Soil Plant Analysis Development
SVM . . . . .	Support Vector Machine
TN . . . . .	True Negative
TP . . . . .	True Positive
UAV . . . . .	Unmanned Aerial Vehicle
USDA . . . . .	United States Department of Agriculture
VEG . . . . .	Vegetative Index

## LIST OF SYMBOLS

$C$ . . . . .	Penalty parameter of support vector machine model
$\text{coef0}$ . . . . .	Independent term in kernel of support vector machine model
$B'$ . . . . .	Brightness of HSB color space
$B^*$ . . . . .	Normalized blue from RGB color space
$c_i$ . . . . .	Pixel count of $i^{\text{th}}$ range of DGCI
$C_{i,j}$ . . . . .	Number of observations known to be in group $i$ but predicted to be in group $j$
$\text{Chl}_a$ . . . . .	Chlorophyll 'a' content
$\text{Chl}_b$ . . . . .	Chlorophyll 'b' content
$\text{Chl}_{a+b}$ . . . . .	Total chlorophyll
$E^X$ . . . . .	Extinction of mixed chlorophyll at the wavelength of $X$
$G$ . . . . .	Gini index o calculate purity index in decision tree (eq. 4.5)
$G^*$ . . . . .	Normalized green from RGB color space
$H$ . . . . .	Entropy to calculate purity index in decision tree (eq. 4.6)
$H'$ . . . . .	Hue of HSB color space
$k$ . . . . .	Number of band in an image
$n$ . . . . .	Number of classes related to IDC severity (chapter 4)
$n_f$ . . . . .	Number of flights (eq. 5.2)
$p_k$ . . . . .	Ratio of class $k$ instances among the training instances

$r$ . . . . .	Pearson correlation coefficient
$R^*$ . . . . .	Normalized red from RGB color space
$R^2$ . . . . .	Coefficient of determination
$S'$ . . . . .	Saturation of HSB color space
$T$ . . . . .	Time
$t_i$ . . . . .	Time of the $i^{th}$ flight (day)
$V_i$ . . . . .	Soybean's $i^{th}$ growth stage
$X$ . . . . .	Input or independent variable
$Y$ . . . . .	Output or dependent variable
$y_i$ . . . . .	Extracted feature from the aerial image of the $i^{th}$ flight

Greek symbols

$\alpha$ . . . . .	Slope (gain) of two point calibration (eqs. 2.11 and 2.12)
$\beta$ . . . . .	Intercept (offset) of two point calibration (eqs. 2.11 and 2.12)
$\beta_i$ . . . . .	Slope of regression line (eqs. 3.8–10)
$\epsilon$ . . . . .	Random error term of linear models
$\rho_{\text{green}}$ . . . . .	Spectral reflectance of green wavelength
$\rho_{\text{NIR}}$ . . . . .	Spectral reflectance of near infrared wavelength
$\rho_{\text{red}}$ . . . . .	Spectral reflectance of red wavelength

## LIST OF APPENDIX FIGURES

<u>Figure</u>	<u>Page</u>
B1. An illustration of nested cross-validation applicable to SVM and RF models. . . . .	164
D1. An illustration of relationship of different features from aerial images with yield for 2016. . . . .	174
D2. An illustration of relationship of different features from aerial images with yield for 2017. . . . .	175
E1. Bar graph of $R^2$ with error bars produced by ggplot2 in R to compare performance of different machine learning algorithms. . . . .	178
E2. Bar graph of RMSE with error bars produced by ggplot2 in R to compare performance of different machine learning algorithms. . . . .	179
E3. Bar graphs with with facets showing three independent variables produced by ggplot2 in R to compare the correlation between parameters and yield. . . . .	181
E4. Bar graph on the left and histogram on the right produced by Plotly in Python. . . . .	183
E5. Bar graph produced by Plotly in Python to compare classification metric among different machine learning algorithms. . . . .	184

# 1. GENERAL INTRODUCTION

From an agriculture standpoint, catering to the food, feed, and fiber needs of the growing population, which is expected to grow with an additional 2.4 billion and increasing by 2050, is the major challenge for the farmers, producers, breeders, researchers, industries, planners, and policymakers. Therefore, it is imperative to develop the best management practices and crop cultivars by efficient research practices that increase yield and sustain the whole agriculture enterprise. Among the major field crops that hold the potential to feed the population, soybean (*Glycine max* (L.) Merrill) enjoys global importance. Soybean provides the plant-based protein as soymeal, serves as a good source of edible oil, as well as being an ingredient of several processed foods. But this important crop is facing major concern affecting its productivity and sustainability from the soil and weather conditions in which the crop is grown.

Due to the characteristics of soil and conditions of weather, soybean can experience iron deficiency chlorosis (IDC), which can affect the performance of the crop. The IDC causes \$120 000 000 yield loss every year in parts of the Midwest, let alone the rest of the world. There are several symptoms associated with IDC in soybeans, such as stunted growth and yellowing of leaves due to the reduction of chlorophyll. In addition, the amount of time that iron deficiency is present in soybean significantly affects the final yield. Any methods dealing with iron deficiency should consider measuring the chlorophyll level in the first place to quantify the extent of the damage. Therefore, efficient, timely, and repeatable measurement of chlorophyll is critical in soybean IDC management.

There are different methods available to quantify IDC both directly by measuring Fe or chlorophyll in the leaves and indirectly by measuring the symptoms including the reflectance of soybean canopy or from the leaf samples. Direct methods are atomic absorption spectroscopy (AAS) (Ferreira et al., 2019b) and chlorophyll extraction in the lab (Santos et al., 2019). The AAS method measures the amount of iron in the leaves or even from the roots of soybean. Indirect methods include visual rating (Helms et al., 2010), digital imaging (Bai et al., 2018), chlorophyll meter reading (Bin et al., 2016), and satellite (Peng et al., 2017) or unmanned aerial vehicle (UAV) imaging-based analysis (Dobbels and Lorenz, 2019). Imaging technologies are good to capture larger areas and are capable of recording data frequently and archive historical data. Image-based analysis methods also could be used to delineate the IDC regions (Rogovska and Blackmer, 2009) to plant tolerant and susceptible cultivars in their respective regions. However, breeders usually use the visual rating method. They rate soybean yield or hill plots for the IDC symptoms mainly the yellowness of the canopy or leaves. The yellowness or chlorosis is a direct indicator of IDC severity, but this visual rating method is time-consuming, subjective, sometimes inaccurate, less frequent, and not feasible for large scales.

Another major activity regarding IDC is to alleviate the symptoms of IDC that has the potential to recover the crop or improve the crop yield. Of the several methods of alleviating the IDC symptoms, the major methods considered in general practice include planting a companion crop (Bloom et al., 2011), planting in wider rows (Wiersma, 2007), applying iron chelate in furrow (Kaiser et al., 2014) and on leaves (Franzen et al., 2003), and planting a tolerant cultivar (Naeve and Rehm, 2006). Most of these methods require extra management steps. For instance, when planting a companion crop, this crop should be killed at a certain

point before it starts competing for water or other nutrients. Another example is a seed or foliar treatment which gives a better result when the planing is performed in wider rows or higher seeding density (Goos and Johnson, 2001). However, among all these methods, planting a tolerant cultivar is the most efficient way to avoid IDC (Hansen et al., 2003). The process of selecting a tolerant cultivar involves planting many different cultivars in different locations usually with a history of IDC. In such selection field trials, trained people will walk among these cultivars and visually rate them for IDC symptoms mainly the color of the canopy. However, several factors such as fatigue, gender, and training affect the human perception of colors (Van Den Broek et al., 2002). Furthermore, this process is performed by manual means, it is prone to error and subjectivity.

To make the measurements more objective, technologies like UAV and digital cameras can be used to monitor soybeans and assess their health status through chlorophyll content. Leaf primary colors — red, green, and blue (RGB) can be extracted from the soybean leaf images, which were obtained in the field conditions, through image processing. Several studies measured chlorophyll from images in the laboratory setting (Rigon et al., 2016; Vollmann et al., 2011), in which all lighting is controlled, through primary colors of the leaves and other color vegetation indexes (CVIs). Instead of primary colors, the dark green color index (DGCI) is used in some studies to estimate chlorophyll or nitrogen content. It has been shown that DGCI can correct the errors in the images due to different light conditions or different cameras (Rorie et al., 2011).

From digital images, CVIs can be extracted for IDC rating purposes. The quality of the different turfgrass was rated and compared with visual ratings of experts (Karcher and Richardson, 2003). Although in most cases raters ranked cultivars with equal tolerance



similarly, there was a difference among the ratings of individual plots. In a similar study (Bai et al., 2018), a platform was used to take nadir images of soybean plots and rated them for IDC through machine learning models. The raw RGB values to hue, saturation, and brightness (HSB) were converted and divided the available hue for leaves into five different groups. All the groups were used as independent variables, and office score IDC visual rating was used as a dependent variable to train the machine learning models.

Another factor that greatly affects the final yield in soybean is the amount of time that IDC was present in soybean. The longer IDC is present in soybean the greater the yield loss. It is known that some of the cultivars recover from IDC and become green, but some cultivars do not recover or some even become worse. With the current rating method which rates soybean cultivars twice during the growing season, it is almost impossible to quantify the amount of time that IDC was present in each soybean cultivar. On the other hand, increasing the number of ratings throughout the growing season is not pragmatic because it takes a long time to rate cultivars especially if the number of cultivars is high. However, UAVs can be used to fly the whole cultivar trial in a much shorter time, so they can be flown more frequently to monitor changes in soybean cultivars. For example, Vega et al. (2015) used UAV to monitor sunflowers (*Helianthus annuus* L.) throughout the growing season with a multispectral sensor. They observed that normalized difference vegetation index (NDVI) calculated from the images was correlated with yield, except the values from the early stages of growth. They also observed that the time of day did not influence the NDVI values, so the flexibility of flying the field is higher during the day. Monitoring crops during the growing season frequently can help identify problems early and address them with precision agriculture technologies.

Machine learning is a great tool for yield prediction and disease rating in crops especially when imagery is involved as input to train the models. Machine learning models helped farmers in yield prediction, seed pricing, seed selection, disease detection, management zone clustering, and several other applications (Rehman et al., 2019). Machine learning is both science and art of programming the computers so they can learn from the data (Géron, 2017). There are two main types of machine learning models, supervised and unsupervised models. Supervised models are the ones that the models have access to the original dependent variable (target feature) such as disease rating for classification problems or final yield for regression problems. Different models use the independent variables (input features) to train themselves to predict as closely as possible to the target feature. On the other hand, in an unsupervised problem, the models do not have access to labeled target features to train the model; the model tries to find a pattern among the input features to organizes the data in different categories or clusters for prediction (Moghadam et al., 2017).

The simplest of all supervised machine learning models are simple linear regression that has been used in many studies such as chlorophyll estimation (Vollmann et al., 2011), assessment of leaf nitrogen in corn (*Zea mays* L.) (Rorie et al., 2011), and yield estimation based on canopy vigor (Sankaran et al., 2015). The more sophisticated models include random forest (RF), support vector machine (SVM) K-nearest neighborhood (KNN), and others, have been used in similar studies such as chlorophyll estimation in fresh-cut rocket leave (*Eruca sativa* Mill.) (Cavallo et al., 2017), wheat (*Triticum aestivum* L.) leaf rust detection at canopy level (Azadbakht et al., 2019), identifying field attributes that predict soybean yield (Smidt et al., 2016).

### **1.1. Significance of the Problem**

Identifying an IDC-tolerant cultivar of soybean is the critical step in managing the IDC. However, the current process of IDC measurement is laborious and expensive, as well as selecting the tolerant cultivars is subjective and hard to perform over a large number of cultivars. Moreover, the current method is incapable of measuring the amount of time that soybean cultivars were affected by IDC. On the contrary, digital camera, and aerial imagery from UAVs can be collected from soybean cultivar trials more frequently to monitor both severity and duration of IDC symptoms. The studies that focused on this area are scarce and have been done with special platforms or in controlled environments (Bai et al., 2018; Naik et al., 2017). Therefore, this research study was proposed to measure and monitor IDC at different scales (leaves, plots, and fields) using image processing and machine learning approaches.

Estimating the chlorophyll content in soybean leaves as an indicator of IDC in the field condition is the foundation for any impactful field-based research that implements digital imaging as a tool to monitor IDC. Vegetation indexes could be investigated for their efficiency in estimating chlorophyll content. The best index then could be used alongside other features such as canopy size and other calculated parameters to represent the trends in rating IDC at the plot level. Smartphones, the most common gadget nowadays, could also be employed in obtaining the pictures of leaves or plots and processing them for IDC rating within the same device. This process potentially could be a substitute for visual rating and requires minimal training. The same vegetation index could be implemented in a pipeline that captures aerial images of cultivar trials to monitor the growth of soybean as well as progress and recovery from IDC. Application of UAV enables researchers to monitor agricultural fields more

frequently for better management of IDC and the images can be kept for future use as historical layers and further advanced analysis.

### **1.2. Statement of Hypothesis**

1. Color vegetation indexes from digital images can be used to estimate chlorophyll content, as measured by SPAD meter, in soybean leaves in the field through machine learning models.
2. Color vegetation indexes derived from digital images with machine learning models can be used to rate IDC in soybean plots.
3. UAV images and image processing methods can be applied to monitor the progress of IDC leading to the selection of tolerant cultivars on a cultivar trial.
4. Various image-based parameters derived from IDC affected soybean cultivars will be correlated with yield and could be used to develop prediction models.

### **1.3. Statement of Objectives**

The specific objectives of this research are presented in individual chapters, while the major objectives of this research work are:

1. To review the various methods of IDC determination and IDC rating in soybean.
2. To estimate chlorophyll in soybean leaves in-situ with digital imaging using a smartphone and machine learning.
3. To rate IDC in soybean using image processing technique using a handheld digital camera and machine learning models at the plot scale.

4. To monitor the IDC progress in soybean cultivar trials using UAV images at the field scale and develop the yield prediction models using various image-based parameters.

#### **1.4. Dissertation Organization**

The thesis consists of chapters such as general introduction (Chapter 1), and the study's objective is presented in the form of four peer-reviewed journal articles (Chapters 2–5), general conclusions (Chapter 6), combined references, and appendices (A–E). The research chapters (2–5) are complete with its own introduction with supporting review of literature, materials and methods, results and discussion, and conclusions. The general introduction (this chapter) provides an overall picture of the available and methods practiced to quantify IDC, methods to alleviate the IDC symptoms, and how digital imaging and machine learning could help researchers in quantifying IDC. It also includes the significance of the problem, statement of hypothesis, and statement of objectives.

Chapter 2, dealing with the review of literature on soybean IDC (Paper 1), entitled: “Iron deficiency chlorosis measurement in soybean a review” is presented in the form of a review paper (Paper 1, Objective 1). This review covers the various methods that both directly and indirectly quantify IDC in soybean. Direct methods include AAS and chemical extraction of chlorophyll. Indirect methods include visual rating, chlorophyll meter reading, digital imaging, and aerial and satellite imaging.

Chapter 3, dealing with smartphone-based chlorophyll estimation (Paper 2), entitled: “Chlorophyll estimation in soybean leaves infield with smartphone digital imaging and machine learning” identifies the best vegetation index that is suitable for infield chlorophyll

estimation through digital imaging. In addition, several machine learning models and different sets of input features were evaluated for chlorophyll estimation (Objective 2).

Chapter 4, dealing with digital camera-based plot scale IDC rating (Paper 3), entitled: “Rating iron deficiency in soybean using image processing and decision-tree based models” studies three machine learning models such as decision tree, random forest, and adaptive boosting and their performance in classifying IDC severity from features that were extracted from digital images from soybean plots. The pictures were taken with a handheld digital camera and closer to a human point of view applied in visual rating (Objective 3).

Chapter 5, dealing with digital UAV-based field scale IDC rating (Paper 4), entitled: “Measuring soybean iron deficiency chlorosis progression and yield prediction with unmanned aerial vehicle” implements the vegetation index that was proven in this research to be effective in quantifying IDC to monitor both severity and progress of IDC for different cultivars through images of UAV that were taken few times during the growing season. Soybean yield prediction models were also developed using the image image-based parameters (Objective 4).

Chapter 6 entitled: “General Conclusions” that summarizes the results derived from the papers (1–4) and also presents the “Suggestions for Future Work.”

The “References” listing is a combined unnumbered chapter of the whole dissertation. The APA reference style, which was followed by the American Society of Agricultural and Biological Engineers (ASABE), was used in the reference listing.

Appendices, A through E present the Python and R code sections for data analysis and visualization of the results. Appendix A is related to the script that was used to extract features from smartphone images to estimate chlorophyll content of soybean leaves

(Chapter 3). Appendix B explains the nested cross-validation script along with an example that was used to tune the parameters of advanced machine learning models (Chapter 3). Appendix C demonstrates the script that was required to detect the two middle rows in the soybean plots (Chapter 4). Appendix D illustrates the scatterplots of yield versus different features that were extracted from aerial images of soybean plots (Chapter 5). Finally, Appendix E demonstrates the scripts that were used for data visualization throughout the dissertation.

## 2. IRON DEFICIENCY CHLOROSIS MEASUREMENT IN SOYBEAN — A REVIEW \*

### 2.1. Abstract

Iron deficiency chlorosis (IDC) causes a great yield loss in soybean (*Glycine max* (L.) Merrill). The main symptom for IDC is the lack of chlorophyll that makes crop foliage yellow resulting in stunted growth leading to lower yields. Several methods were developed to alleviate the symptoms of IDC, but quantifying its symptoms should be done in advance for effective management. Measurement of IDC plays an important role in mapping zones that are susceptible to IDC and aids in selecting tolerant cultivars in experimental trials. However, synthesis of various IDC measurements, especially related to modern and efficient methods, such as the use of image processing, is scarce. Therefore, this paper reviews several aspects of IDC measurement that includes the number of publications available on the topic through search using a combination of relevant keywords; different methods of IDC measurement using visual rating by experts, atomic absorption spectroscopy for estimating iron, wet chemistry of chlorophyll extraction and sensors application for chlorophyll estimation, image processing methodologies for chlorophyll and IDC rating through different imaging platforms; analysis of data through statistical and advanced machine learning methods; and methods of management of IDC. The IDC assessment is also important for the researchers who focus on

---

\* This paper is planned to be submitted as a review article in the *Field Crops Research* journal. Authors: Oveis Hassanijalilian, C. Igathinathane, Hans Kandel, Sreekala Bajwa, and John Nowatzki. Oveis performed the literature survey and wrote the manuscript. Dr. Igathinathane Cannayen is the major advisor and the corresponding author who worked with Oveis throughout the manuscript development. All the co-authors have assisted in the research direction and review of the manuscript.



gene expressions and measure IDC in the lab environment and those developing digital image processing methodologies to estimate chlorophyll and rate IDC. This review explains the capabilities of the various IDC measurement methods applicable in different environments and research scales. Also outlined in this review are the future perspectives of application of remote sensing, advanced hyperspectral sensors, LiDAR, and soil sampling analysis with historical data as inputs to machine learning models to detect and delineate IDC.

## 2.2. Introduction

Iron deficiency chlorosis (IDC) is one of the reasons that causes a great yield loss in soybean (*Glycine max* (L.) Merrill), and that is alarming for the farmers, producers, breeders, and researchers, while there is a growing interest in this crop in North America. In 2004, soybean was planted on  $1.45 \times 10^6$  ha (USDA-NASS, 2005), but in the past three years, it was planted on an average of  $2.70 \times 10^6$  ha in North Dakota (ND) almost twice the land area that was planted for corn (*Zea mays* L.). The amount of land that was planted with soybean was almost doubled in 14 years (USDA-NASS, 2020), but the land in ND is prone to IDC which was responsible for \$120 000 000 yield loss in  $1.8 \times 10^6$  ha in just North Central America (Hansen et al., 2004). The soil that is susceptible to IDC is usually characterized by high PH, calcareous, excess moisture, excess nitrate, and high levels of bicarbonate (Hansen et al., 2003; Roriz et al., 2014).

Soybeans that are affected by IDC show stunt growth and interveinal chlorosis — loss of green color and becoming yellow, due to lack of chlorophyll, and they might become necrotic — brown and die in severe cases. Iron deficiency happens in younger leaves because iron is immobile after the first trifoliolate emerges, and it might be observed in soybean for

eight weeks, but in some cases, IDC might persist throughout the entire season (Hansen et al., 2003).

There are several ways to minimize the expression of IDC such as planting a companion crop (Bloom et al., 2011; Naeve, 2006), applying iron chelate at planting (Kaiser et al., 2014), or foliar application (Franzen et al., 2003), and increasing seed density at the time of planting (Goos and Johnson, 2000; Wiersma, 2007). However, the most effective way to mitigate IDC symptoms is to plant an IDC tolerant soybean cultivar (Naeve and Rehm, 2006). In a survey that was done in 2002 in Minnesota regarding IDC in soybean and best management practices, the majority of farmers stated that the cultivar selection is the best way to avoid IDC (Hansen et al., 2003).

Measuring IDC through its symptoms could be beneficial in detecting the severity of IDC. Additionally, it could help farmers and researchers to delineate the areas that IDC is prevalent through aerial platforms such as an unmanned aerial vehicle (UAV) and satellites. With recent technologies, cultivars with different IDC tolerance levels could be planted in the field. Since a tolerant cultivar does not necessarily produce the highest yield, a susceptible cultivar could be planted in non-IDC regions with the goal of improving the yield of total field. Information regarding IDC is scattered and scarce, therefore there is a need to synthesize the literature and update the information with new research developments for the benefit of the users.

Our objective with this review paper is to discuss different methods of measuring IDC. Some of these methods directly measure the iron content in soybean like atomic absorption spectroscopy, while the others work with the symptoms such as visual rating and imagery (different imaging platforms) for measuring chlorophyll as the main indicator of IDC.

### **2.3. Number of Publications on Soybean IDC and Nature of Work**

A snapshot of the number of publications on the subject area of IDC with different combination of keywords in “Web of Science” database from 1900 to 2020 is presented in figure 2.1. This snapshot provides an idea of the research landscape of soybean IDC. The keywords are related to research topics (TS) about soybean, IDC, chlorophyll measurement, and different remote sensing platforms. The legend of the generated results was sorted from lowest to highest frequency.

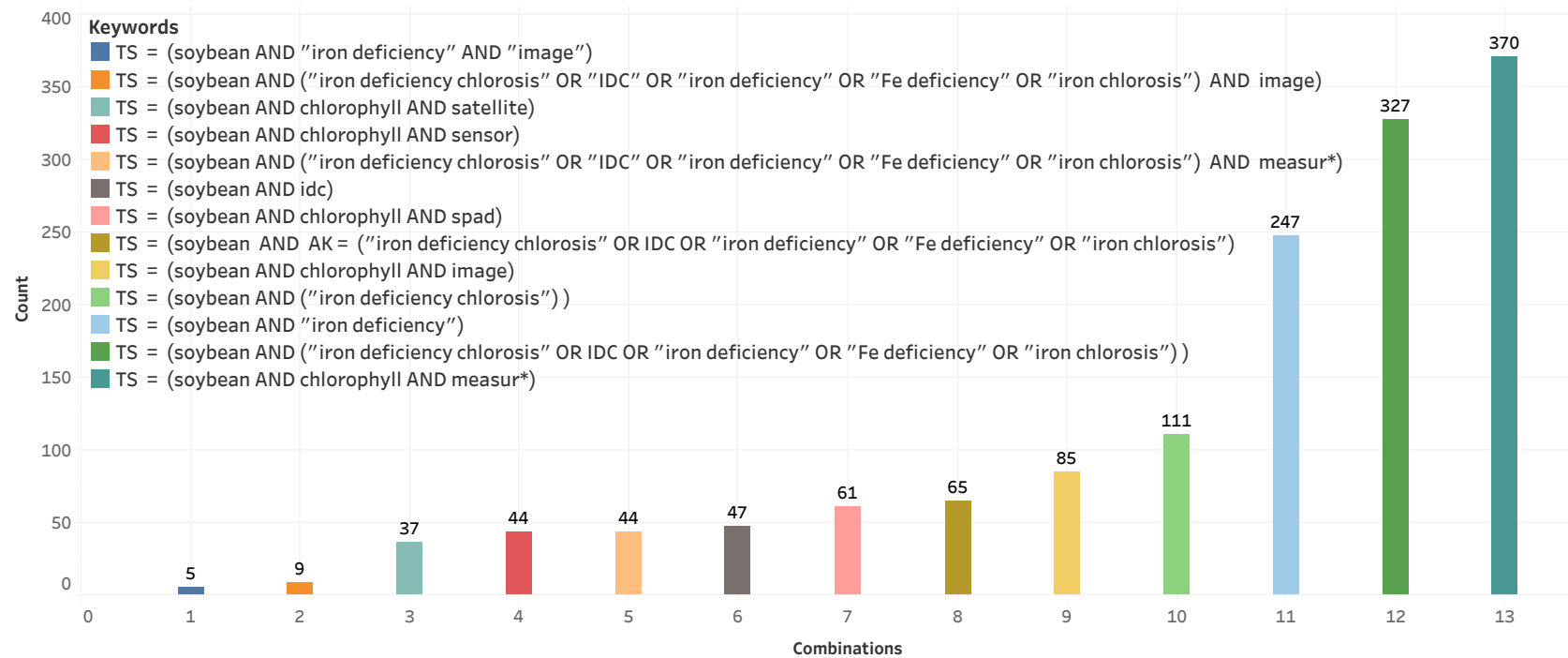
The total number of publications on chlorophyll measurements are high, but chlorophyll measurement through images shows a 75 % reduction in the number of publications. As chlorophyll is one of the main indicators of different nutrient deficiency in soybean, the number of publications therefore on chlorophyll measurement is significantly higher than measuring IDC in soybean. However, the lack of chlorophyll is the main symptom of IDC, and measuring chlorophyll is useful in assessing IDC severity.

Direct studies on IDC are mostly related to practices to avoid IDC and on crop genetics, but less attention was given towards measuring and quantifying IDC spatially. Therefore, measuring chlorophyll through remote sensing platforms such as satellite is significantly less. The number of publication on measuring IDC through images or studies that implement the modern use of sensors on different platforms such as digital cameras, UAVs, and satellites are highly limited.

### **2.4. Measurement of IDC**

#### **2.4.1. Visual Rating**

Breeding programs in academic institutions and seed companies are constantly trying to produce soybean seeds that are more tolerant to IDC. Iron deficiency varies by location and



**Figure 2.1. Number of publications in Web of Science academic database for different keywords used in the advanced search in the website. TS stands for topic and different combinations of keywords presented. The legend is sorted from the lowest frequency to the highest.**

over time such that soybean plants within a meter distance show significantly different IDC symptoms. The main reason is the heterogeneity of soil which forces the researcher to consider different environments in their research designs (Naeve and Rehm, 2006).

Breeders plant many different cultivars among different environments for the purpose of cultivar selection. For cultivar selection, soybean is planted either as hill-plots (Helms et al., 2010) or multiple-row plots to be harvested for yield (Nowling and Cai, 2018). Helms (2015) planted 7264 hill-plots within 4 locations and 4 replications, a total of 454 different cultivars, to be able to compare them for IDC symptoms side by side. The other method includes planting multiple rows of soybean for each plot which usually is 4.6 to 6.1 m long with 3 or 4 replications (USDA Soybean Test North Region). The rating is done only on the center rows to avoid the “border” effect. Similarly, 3.7 to 4.9 m of the center rows are harvested to avoid the “end of the row” effect.

There are several methods of rating IDC in soybean, and the most widely used one is 1 to 5 scale with one being completely healthy and 5 is severe chlorosis with some necrosis (Helms et al., 2010; Naeve and Rehm, 2006). A visual scorecard of the range of IDC severity and its related score are presented in figure 2.2 . The IDC rating scores that were used in the experiments and the description of the symptoms developed by Helms et al. (2010) are also given in table 2.1.

Although this rating method is widely used, it is a subjective measure and human bias introduces significant errors in this process. For example, fig. 2.3 shows three different plots that were rated by one rater (expert performing the visual rating), and clearly the middle plot rated very high for IDC while the symptoms are not that different compared to the healthier plot. Furthermore, Karcher and Richardson (2003) used raters and digital images to quantify



**Figure 2.2. Iron deficiency chlorosis (IDC) visual score for the scale of 1-5. *Image source: Visual representation of NDSU IDC scoring scale, photos by T. Helms, NDSU soybean breeder.***

the color of turfgrass, and they observed differences in absolute rating values among different raters. Based on the analysis, they concluded that a more consistent rating can be obtained across researchers, years, and locations through digital images.

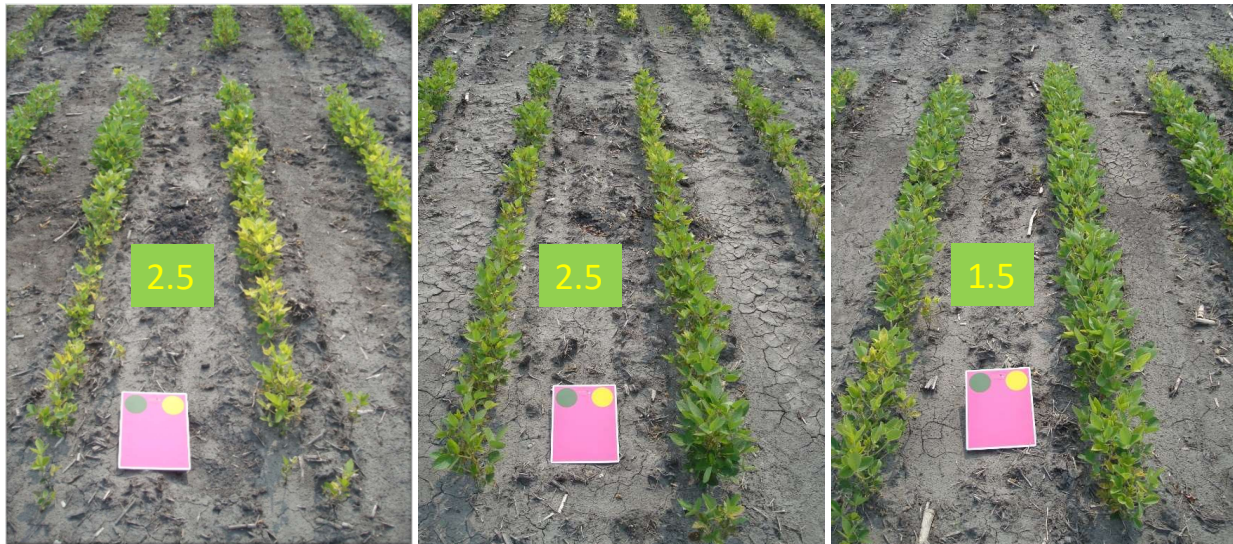
#### **2.4.2. Atomic Absorption Spectrometry**

Atomic absorption spectrometry (AAS) is one of the widely used techniques for multi-element determination within a sample. It is based on the measurement of absorption of energy through the excitation process of free atoms or emission of energy through the

**Table 2.1. Developed visual rating score guidelines for IDC symptoms<sup>1</sup>.**

IDC symptoms description	IDC Score
No chlorosis	1
Slight yellowing of the upper leaves	2
Upper leaves without necrosis or stunting and with interveinal chlorosis	3
Upper leaves with reduced growth or beginning of necrosis with interveinal chlorosis	4
Severe stunting, damaged to growing point and chlorosis	5

Note: IDC - Iron deficiency chlorosis; and <sup>1</sup> IDC scores scale was developed by Helms et al. (2010).



**Figure 2.3. Soybean plots and their iron deficiency chlorosis visual scores. The middle plot shows a high amount of error in its score comparing to the plots on its both sides.**

decay process of atoms that were excited by light source (Beaty and Kerber, 1978). The amount of material corresponds to the amount of light that was either absorbed or emitted from the cloud of atoms. Rodríguez-Lucena et al. (2010) used AAS (Perkin-Elmer AAnalyst 800 Spectrophotometer) to measure the amount of iron in different parts of soybean plants in order to compare the effectiveness of different iron chelates and complexes in correcting IDC. The AAS proved to produce the same result in measuring iron in soybean as inductively coupled plasma mass spectroscopy (ICP-MS) (Jajda et al., 2015). The range of iron for AAS

was 78.6–210.6 ppm (average = 123.23 ppm) and for ICP-MS was 71.8–200.4 (average = 123.73) with a correlation of 0.98 between two methods.

This method is capable of analyzing other parts of soybean such as root and stem for Fe concentration. The concentrations of Fe did not change over time and found significantly lower than those found in the leaves; this indicated the transitory behavior of Fe in the soybean's stem (Ferreira et al., 2019b).

All parts of plants including leaves, stem, and roots should be washed and rinsed appropriately to remove dirt and inorganic contaminants. After separating different parts of the plant, they could be washed with 0.1 % HCL and 0.01 % non-ionic detergent solution (Tween 80, Probus, Barcelona, Spain) in 0.1 M HCl for 30 s. Then distilled water should be used twice to rinse the parts. To determine the dry weights, the parts then should be placed in a force air oven to get dry in 60 °C–65 °C for 72 h. To crush the plant samples, a titanium mill (Retsch ZM200) could be used followed by muffle furnace at 480 °C for 2 h to turn them into ash. Another way is to calcine them at 480 °C for 4 h. Acid digestion should be performed with 1:1 diluted HCl (30 % Suprapur, Merck) or HNO<sub>3</sub> (Suprapur, Merck) for ash solubilisation (Jones Jr, 2001). The total Fe concentration in the digested samples should be filtered through 20–25 µm paper filters (FilterLab 1238) and could be determined by flame atomic absorption spectroscopy (AAS).

#### **2.4.3. Wet Chemistry — Indirect through Chlorophyll**

The reduction of chlorophyll below normal chlorophyll content in the leaf is the main symptom of IDC in soybean. Therefore, chlorophyll measurement could be an indicator of IDC severity. There are several methods to measure the chlorophyll in soybean leaves, but the most accurate one is to extract chlorophyll pigments and measure their concentration. With



this measurement, the concentration of  $\text{Chl}_a$ ,  $\text{Chl}_b$ , and total chlorophyll could be determined by measuring the extinction of the extract at  $\sim 664$  nm for  $\text{Chl}_a$  and  $\sim 647$  nm for  $\text{Chl}_b$ . In this process, several solvents could be used such as acetone, N,N'-dimethylformamide (DMF), dimethylsulfoxide (DMSO), and methanol. Aqueous acetone (80:20 v:v pH = 7.8) is the common solvent, which exhibits a sharp peak at the aforementioned wavelengths. The DMSO and DMF exhibit a sharp peak at those wavelengths, but they are more toxic. On the other hand, methanol exhibit a broader peak at those wavelengths, but better degrades chlorophyll in alkaline conditions (Porra, 1991). Cutting and grinding the sample shows a better result in measuring chlorophyll.

Researchers have been using equations 2.1–3 to determine the concentration of  $\text{Chl}_a$ ,  $\text{Chl}_b$ , and total chlorophyll ( $\text{Chl}_{a+b}$ ), respectively in aqueous 80 % acetone (Arnon, 1949). The followings are the equations to calculate  $\text{Chl}_a$ ,  $\text{Chl}_b$  and total chlorophyll in  $\mu\text{g} \cdot \text{ml}^{-1}$  (Arnon, 1949).

$$\text{Chl}_a = 12.70E^{663} - 2.69E^{645} \quad (2.1)$$

$$\text{Chl}_b = 22.90E^{645} - 4.68E^{663} \quad (2.2)$$

$$\text{Chl}_{a+b} = 20.21E^{645} + 8.02E^{663} \quad (2.3)$$

where  $\text{Chl}_a$  is chlorophyll *a* content,  $E$  is the extinction of mixed chlorophyll at 663 and 645 nm,  $\text{Chl}_b$  is chlorophyll *b* content, and  $\text{Chl}_{a+b}$  is total chlorophyll content.

However, Porra et al. (1989) argues that these equations produce an average of 10 % error compared to their method, which produced only an error of 1 % with AAS. New equations by Porra that uses buffered aqueous 80 % acetone (pH 7.8) can be found below (for other solvents read Porra (2002)).

$$\text{Chl}_a = 12.25E^{663.6} - 2.55E^{646.6} \quad (2.4)$$

$$\text{Chl}_b = 20.31E^{646.6} - 4.91E^{663.6} \quad (2.5)$$

$$\text{Chls}_{a+b} = 17.76E^{646.6} + 7.34E^{663.6} \quad (2.6)$$

#### 2.4.4. Sensors — SPAD — Indirect through Chlorophyll

Several sensors can be used in quantifying IDC in soybean by estimating the chlorophyll content when an expert detects that the lack of chlorophyll or discoloration is the cause of IDC. Soil plant analysis development (SPAD) meter (Minolta, 1989) is one of the sensors that is widely used especially among researches. The SPAD meters estimate the chlorophyll content by measuring the absorbance of a leaf in two different wavelengths, namely 650 nm (red) and 940 nm (near-infrared). At 650 nm both chlorophyll ‘a’ and ‘b’ have the maximum absorbance, whereas 940 nm is used to adjust for individual leaf thickness (Vollmann et al., 2011). The adjusted coefficient of determination of a second-degree polynomial model between SPAD and total chlorophyll content found to be high ( $R^2 = 0.96$ ) (Rigon et al., 2016)

Actual chlorophyll content could be derived from SPAD meter readings (Markwell et al., 1995) for soybean and corn using the following conversion relationship (eq. 2.7):

$$\text{CC} = 10^{(\text{SMR}^{0.265})} \quad (2.7)$$

where CC is chlorophyll content ( $\mu\text{mol} \cdot \text{m}^{-2}$ ), and SMR is the SPAD meter reading. However, SPAD meters do not necessarily produce identical data. Therefore, the SPAD meter values

from different devices should not be directly compared unless calibration procedures are standardized (Markwell et al., 1995).

To properly use the device, the SPAD meter has to be used without any sample in the sample area. This helps the SPAD meter to record and save the values in the memory for different wavelengths. Then it has to be used with the provided reference and the measurements have to be with a defined range. Finally, the SPAD meter could calculate the transmission of the different wavelengths for other samples relative to the saved values. (Markwell et al., 1995).

Naeve and Rehm (2006) measured relative leaf chlorophyll concentration with a SPAD meter and found a high correlation with IDC visual rating ( $r = -0.93$ ). Chatterjee et al. (2017) studied the effect of different foliar application of Fe on re-greening and yield of soybean. They found that the SPAD meter readings of treated soybean plots indicated higher chlorophyll content over the growing season compared to the control treatment. Moreover, the application of two different iron chelates that were applied to the soil with chlorotic soybean plants showed an increase in SPAD measurements in soybean leaves (Ferreira et al., 2019b).

Symptoms of IDC appear between the veins on the leaf, so measurements over the veins result in higher reading because veins contain more chlorophyll (Bin et al., 2016; Coste et al., 2010; Uddling et al., 2007). Additionally, the sampling size of this device is really small, and to deal with this issue, users measure three to five different spots on a leaf and average the values to get a more representative estimation for the leaf (Gamble et al., 2014). Since the IDC is not uniform throughout the plant and some leaves have more IDC than

others, images from a digital camera can capture the whole leaf and they are more accessible than SPAD meter (Rigon et al., 2016; Vollmann et al., 2011).

#### 2.4.5. Image Processing — Digital Images and Rating Iron Deficiency

Reduction of chlorophyll and chlorosis is the main symptom of IDC in soybean and shows itself as the yellow color in the leaves (Ferreira et al., 2019b). Similar to human eyes, digital cameras are capable of capturing colors. Through images processing software different features can be extracted from the images and the correlation between these features can be studied with real features of the soybean such as chlorophyll content (Rigon et al., 2016; Vollmann et al., 2011), leaf area index (Liu et al., 2013), canopy area (Yu et al., 2016), among others. The cameras that are widely used for the aforementioned applications are capable of recording the reflection of the RGB color in three different layers. The reflected values mostly get recorded as 8-bit numbers, which can be anywhere between 0 and 255 ( $2^8 = 256$ ) for every pixel.

The final image is the product of combining these values for every pixel. For instance, the RGB values of 255, 255, 0 will produce yellow, and RGB values of 255, 0, 255 will produce purple. The RGB color space might not be intuitive, but the hue, saturation, and value (HSV) or the hue, saturation, and brightness (HSB) color space is closer to human perception of color and can be calculated from RGB values (range 0–1) as follows (Smith, 1978):

$$H' = \begin{cases} \text{If } \max(R, G, B) = R; & 60 \times \left[ \frac{G - B}{\max(R, G, B) - \min(R, G, B)} \right] \\ \text{If } \max(R, G, B) = G; & 60 \times \left[ 2 + \frac{B - R}{\max(R, G, B) - \min(R, G, B)} \right] \\ \text{If } \max(R, G, B) = B; & 60 \times \left[ 4 + \frac{R - G}{\max(R, G, B) - \min(R, G, B)} \right] \end{cases} \quad (2.8)$$

$$S' = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (2.9)$$

$$B' = \max(R, G, B) \quad (2.10)$$

where  $H'$  is hue,  $S'$  is saturation, and  $B'$  is the brightness of HSB color space.

Another color space that is being widely used is CIE  $L^*a^*b^*$ , where  $L^*$  is lightness from black to white,  $a^*$  changes from green(−) to red(+), and  $b^*$  changes between blue(−) to yellow(+).

Bai et al. (2018) converted RGB values of soybean plots into HSV, and extracted all the pixels with hue between  $40^\circ$  to  $170^\circ$  which includes green, yellow, and orange for initial plant segmentation. They used the available hue values to train a machine learning model to score soybean plots for IDC. Naik et al. (2017) calculated the proportion of yellow and brown colors on the soybean leaves base on their respective hues ( $21^\circ$ – $50^\circ$  and  $51^\circ$ – $80^\circ$ , respectively). They used these values to train several machine learning models to predict IDC visual scores.

#### **2.4.5.1. Image acquisition**

The first step in using image processing is acquiring images that provides helpful information about the health status of the soybean. Images could be captured in different formats such as raster, spectral reflectance, and point cloud through different platforms and sensors. Multispectral, RGB, and thermal cameras produce raster formats that can be captured in different spatial resolution from sub-centimeter of smartphone cameras (Hassanijalilian et al., 2020b) to 30 m resolution of satellite images (Peng et al., 2017). Hyperspectral cameras can record the amount of reflectance of soybean in finer resolutions

compared to multipsectral cameras. However, not all the information is useful in detecting a disease or deficiency, and techniques such as principal component analysis (PCA) should be used to extract the relevant wavelength information. Sensors such as LiDAR and Kinect (Ma et al., 2019) can produce 3D point clouds that can be used to measure canopy height as one of the main symptoms of IDC.

The spectral and spatial resolution of the sensor plays an important role in the quality of the information. The RGB cameras record three wide ranges of the electromagnetic spectrum, but hyperspectral sensors capture many narrow ranges of the spectrum. The RGB data is easier to store, process, and understand, but hyperspectral that has a better spectral resolution gives more information about the crop. With respect to spatial resolution, image registration through georeferencing is necessary for multitemporal studies specially for the ones that have finer resolution such as UAV images (Dobbels and Lorenz, 2019).

Georeferencing enables the researchers to compare the data between different dates.

#### **2.4.5.2. Image registration — georeference, radiometric correction, color correction**

IDC is known for lack of chlorophyll in soybean leaves and canopy, which make them appear in yellowish color. However, to be able to compare the images of different leaves, canopies, and fields that were taken in a different time, images need to be corrected for different lighting conditions; these methods usually are performed for images taken in the field because the lighting condition in the controlled environment is constant between images.

In studies with UAV usually a calibration board is placed in the imagery area to correct the images for different lighting conditions. This board is paint-sand mixed in checkerboard pattern with different shades of gray to create a near-lambertian reflection surface. Depends on the area of the study, several calibration board could be made. The relationship between

digital numbers (DN) and radiance that reaches the sensors is linear. Therefore relative correction could be performed with linear equations for multi-temporal studies as follows:

$$\text{Black}_{T1,k} = \alpha_k \times \text{Black}_{T2,k} + \beta_k \quad (2.11)$$

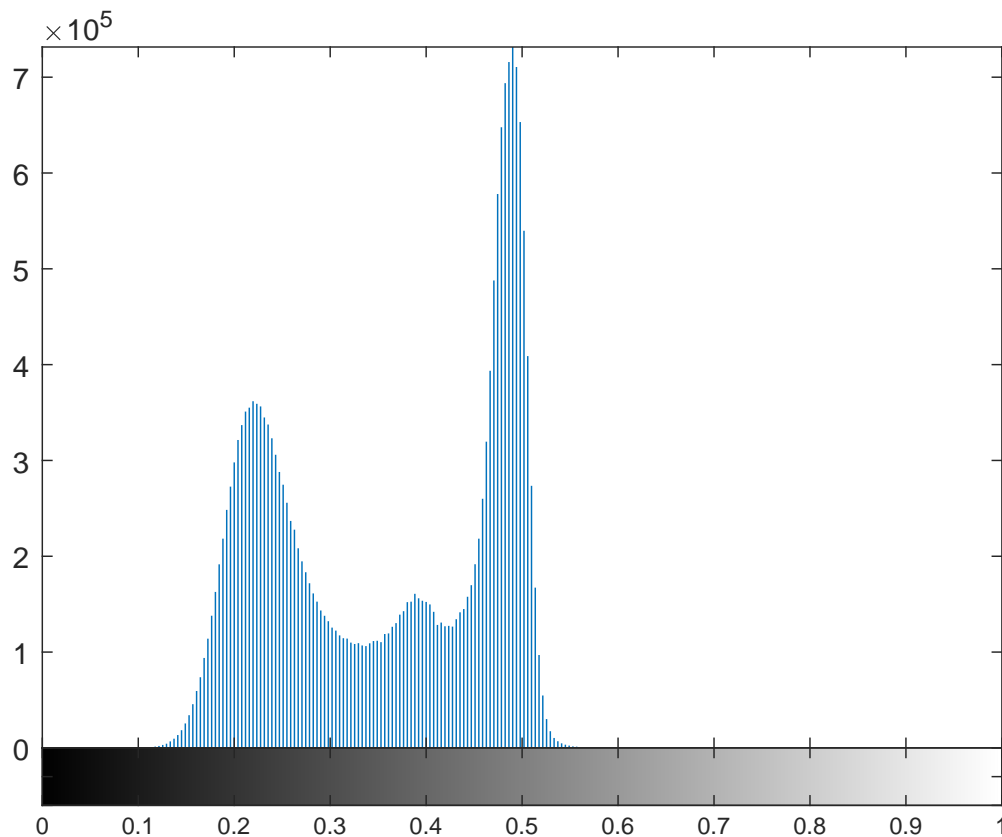
$$\text{White}_{T1,k} = \alpha_k \times \text{White}_{T2,k} + \beta_k \quad (2.12)$$

where Black and White are DN at time T of band  $k$  in the images,  $\alpha$  is the slope (gain), and  $\beta$  is the intercept (offset). The average value of black and white shades of those board could be used for eqs. 2.11 and 2.12. In case of over exposure of white shade, other objects in the study area could be use for calibration (Yu et al., 2016). The same method is used for calibrating satellite imagery with empirical line method (Collings et al., 2011; Smith and Milton, 1999).

#### **2.4.5.3. Image processing — image segmentation, removing background**

Images of soybean could be taken in different light settings and through different platforms; pictures can be taken in a controlled environment and in the field. Regardless of the environment, crop segmentation is necessary to isolate soybean leaves or canopy from the surrounding background, which is soil when pictures are taken in the field. One method to isolate soybean leaves is to set a manual threshold. This threshold captures the range of values that is related to soybean in different health conditions (green to yellow and some cases orange). Different color spaces could be used for this purpose, but the HSV color space is the common one. Hue settings of  $33^\circ$ – $110^\circ$  and saturations of  $6^\circ$ – $100^\circ$  were used in the controlled environment (Vollmann et al., 2011), while  $40^\circ$ – $170^\circ$  of hue was used to isolate soybean crops in the field condition (Bai et al., 2018).

Another method is to use an automatic threshold, so this method could adjust itself with different conditions. Otsu method is the famous method that assumes the values in the image are in two classes, and it tries to maximize the inter-class variance between the classes. Usually, one out of three channels from different color spaces or any other grayscale image that is a result of color vegetation index (CVI) is used with Otsu (Liu et al., 2017). Figure 2.4 shows a bimodal histogram of  $a^*$  channel of an aerial image of the soybean research plot. The peak in the lower end of the histogram relates to soybean canopies, and the other one relates to the soil.



**Figure 2.4. Histogram of  $a^*$  channel of an aerial image of soybean research plot. The peak on the left are values that are related to soybean plots, and the peak on the right is related to the soil background.**



Using special CVI could be beneficial in segmenting crops; CVI such as excess green index (ExG) and excess red (ExR), or sometimes a combination of CVI could be used for crop segmentation. one of the unique combination is ExG+ExR that was used in a couple of studies for crop segmentation (Jin et al., 2017; Meyer and Neto, 2008); the threshold of 0 was used to remove the negative (soil) values. Finally, other sensors such as LiDAR and thermal camera (Deery et al., 2014) and other advanced methods such as supervised and unsupervised machine learning could be used for crop segmentation (Dyson et al., 2019; Sadeghi-Tehran et al., 2017).

#### **2.4.5.4. Color vegetation indices**

From the RGB images of digital cameras, several other color spaces can be derived, and individual components of these color spaces can be used as a color vegetation index or they can be combined to make more sophisticated ones (Bai et al., 2018; Karcher and Richardson, 2003; Vollmann et al., 2011). Some of the cameras have sensors that are capable of recording near infrared (NIR) and infrared (IR) wavelength as well. The reflection of soil, stressed plants and healthy plants are very different for these wavelengths (Thomas et al., 2018), so they can be used in image segmentation and feature extraction (da Silva Junior et al., 2018; Rehman et al., 2019).

Individual components of different color spaces were used in several studies to estimate the amount of chlorophyll, which is an indicator of IDC in soybean leaves. Vollmann et al. (2011) found that the green component of RGB has the highest correlation with chlorophyll content of soybean leaves at the R3 stage in a controlled environment. Since the red, green, and blue components of RGB and hue, saturation, and brightness of HSB color space are closely correlated, they observed a similar relationship between green and

chlorophyll content of soybean leaves. In addition to RGB and HSB color space, Rigon et al. (2016) used CIE L\*a\*b components to estimate the chlorophyll content of soybean leaves. They also used the dark green color index (DGCI), which is a combination of HSB components, and found a strong correlation with total chlorophyll ( $r = 0.84$ ). Table 2.2 shows the correlation between different CVIs and chlorophyll measurement related to different studies; one of these studies were performed in the field condition, while the other ones were in the controlled environments. Based on the results (table 2.2), DGCI shows a great correlation in both field and controlled environment, where single channels like green show inconsistent results.

Normalized difference vegetation index (NDVI) is the most popular vegetation index that has been used in agricultural research. It can be recorded through custom made digital cameras (Lelong et al., 2008), crops sensors (Gamble et al., 2014), and multispectral sensors (Maimaitijiang et al., 2017). Gamble et al. (2014) used GreenSeeker crop sensor (NTech Industries, Inc., Ukiah, CA) over the soybean canopy to record the reflection to determine NDVI as a measure of IDC. They mounted the sensor on a hand-pulling bicycle and recorded the reflection from 80 cm above the canopy of the center rows. They found that NDVI declined with visual chlorosis score at both V3 (Fehr et al., 1971) and V5 growing stage ( $r = -0.63$  and  $r = -0.84$ , respectively).

## **2.4.6. Other Imaging Platforms**

### **2.4.6.1. Satellite imagery**

Normal digital cameras, which produce RGB imagery, are ubiquitous and available for both hand-held and UAV applications, but their capability of recording stress in the plants is limited. On the other hand, multispectral and hyperspectral cameras/sensors can record

**Table 2.2. Correlation of color vegetation indices with SPAD index using RGB and chromatic coordinates (HSB).**

Color vegetation index	Definition	<i>r</i> for field	<i>r</i> for lab	<i>r</i> for lab
Red	-	-0.84	-0.68	-
Green	-	-0.77	-0.82	-0.94
Blue	-	0.22	0.58	-
Hue	-	0.79	0.69	-
Saturation	-	-0.56	-0.88	-
Brightness	-	-0.84	-0.81	-
L*	-	-	-0.81	-
a*	-	-	0.62	-
b*	-	-	-0.90	-
DGCI (Karcher and Richardson, 2003)	$[(\text{Hue} - 60)/60 + (1 - \text{Saturation}) + (1 - \text{Brightness})]/3$	0.90	0.84	-
ExG (Woebbecke et al., 1995)	$2 \times \text{green} - \text{red} - \text{blue}$	0.05	-	-
ExR (Meyer et al., 1998)	$1.3 \times \text{red} - \text{green}$	-0.82	-	-
CIVE (Kataoka et al., 2003)	$0.441 \times \text{red} - 0.811 \times \text{green} + 0.385 \times \text{blue} + 18.78745$	0.25	-	-
ExGR (Meyer and Neto, 2008)	$\text{ExG} - \text{ExR}$	0.45	-	-
Gray	$0.2898 \times \text{red} + 0.5870 \times \text{green} + 0.1140 \times \text{blue}$	-0.79	-	-
GB (Woebbecke et al., 1995)	$\text{green} - \text{blue}$	-0.82	-	-
ERI (Golzarian and Frick, 2011)	$(\text{red} - \text{green}) \times (\text{red} - \text{blue})$	-0.81	-	-
EGI (Golzarian and Frick, 2011)	$(\text{green} - \text{red}) \times (\text{green} - \text{blue})$	-0.36	-	-
RG (Woebbecke et al., 1995)	$\text{red} - \text{green}$	-0.76	-	-
GR	$\text{green} - \text{red}$	0.62	-	-
GBRG (Woebbecke et al., 1995)	$(\text{green} - \text{blue})/(\text{red} - \text{green})$	0.78	-	-
COM1 (Guijarro et al., 2011)	$\text{ExG} + \text{CIVE} + \text{ExGR}$	0.43	-	-
MexG (Burgos-Artizzu et al., 2011)	$1.262 \times \text{green} - 0.884 \times \text{red} - 0.311 \times \text{blue}$	0.41	-	-

RGB - Chromatic coordinates; RGB\* - Combined combined single RGB input (grayscale); DGCI - Dark green color index; ExG - Excess green; ExR - Excess red; CIVE - Color index of vegetation extraction; ExGR - Excess green minus excess red index; COM1- Combined index 1; and MexG - Modified excess green index; GB - Green minus blue; ERI - Excess red feature; EGI - Excess green feature; RG - red minus green; GR - Green minus red; and GBRG - GB by RG ratio.

more information from canopies and have the potential to distinguish stress plants. However, these cameras are more expensive, and usually, if they want to be used with a UAV platform, both the UAV and the sensor must be customized for installation. These sensors are mostly used on high altitude platforms such as UAVs and satellites. Peng et al. (2017) used the bands of multispectral sensors aboard Sentinel-2 to assess canopy chlorophyll content of both soybean and corn, which were chemically extracted in the lab. They used several vegetation indices in both visible and NIR range to study the correlation between canopy reflection and

canopy chlorophyll content. They found that the vegetation indices that consisted of NIR such as NDVI were highly correlated with chlorophyll content ( $r = 0.92$ ) and were minimally affected by crop phenology so they could be used for both crops without algorithm re-parameterization.

To monitor the dynamics of green LAI, different bands from LANDSAT could be used to derive vegetation indices such as NDVI, the optimized soil adjusted vegetation index (OSAVI), the two band enhanced vegetation index (EVI2) and the modified triangular vegetation index (MTVI2)(Liu et al., 2012). The formulas can be found below:

$$\text{NDVI} = (\rho_{\text{NIR}} - \rho_{\text{red}}) / (\rho_{\text{NIR}} + \rho_{\text{red}}) \quad (2.13)$$

$$\text{OSAVI} = 1.16(\rho_{\text{NIR}} - \rho_{\text{red}}) / (\rho_{\text{NIR}} + \rho_{\text{red}} + 0.16) \quad (2.14)$$

$$\text{EVI2} = 2.5(\rho_{\text{NIR}} - \rho_{\text{red}}) / (\rho_{\text{NIR}} + 2.4\rho_{\text{red}} + 1) \quad (2.15)$$

$$\text{MTVI2} = \frac{1.5|1.2(\rho_{\text{NIR}} - \rho_{\text{green}}) - 2.5(\rho_{\text{red}} - \rho_{\text{green}})|}{\sqrt{(2\rho_{\text{NIR}} + 1)^2 - (6\rho_{\text{NIR}} - 5\sqrt{\rho_{\text{red}}}) - 0.5}} \quad (2.16)$$

where  $\rho$  is spectral reflectance at different wavelength such as near infrared (NIR), red, and green. Although at high LAI, the uncertainty of NDVI is the smallest, due to saturation the uncertainty propagated to the LAI estimation is the largest. However, NDVI is the only one that is the most influenced by the leaf chlorophyll. The rest of the indices with comparable results were good indicators of LAI ( $R^2 = 0.83$ ).

In a similar study, Gitelson et al. (2005) developed “green and red edge” models to estimate chlorophyll content based on spectral reflectance measurement at a canopy level. To test the models with spectral bands of satellite sensors, they used green and NIR bands from MODIS System (onboard NASA’s Terra and Aqua Satellite), and the red edge and NIR bands

of the MERIS system (onboard the polar orbiting Envisat Earth Observation Satellite). Both models could produce an accurate estimation of total chlorophyll ( $0.92 \leq R^2 \leq 0.94$  for both crops and models). However, the calibration coefficient for both models remains species-specific for these specific spectral bands from satellites.

#### **2.4.6.2. Aerial imagery**

Rating IDC is a time consuming and labor-intensive task, and that is why the breeders only perform this task only once or twice during the growing season. However, UAV could capture the variation in canopy colors among different cultivars in research plots more efficiently. A different algorithm such as random forest and the neural network could be used to predict the severity of IDC to mimic the breeders' rating. The rating that is done later in the season (V6 growth stage) found to be more accurate due to the bigger size of canopies in the aerial images (Dobbels and Lorenz, 2019).

Kyaw et al. (2008) used aerial imagery and soil electrical conductivity (EC) to delineate chlorosis management zones. They found that the combination of NDVI and EC is the best in predicting soybean yield. Moreover, this combination helped them to delineate the chlorosis zones and associate the yield loss with IDC. Similarly, Rogovska and Blackmer (2009) used aerial images of soybean canopy to map high pH and calcareous soils at field scale as a tool for farmers to identify IDC-susceptible areas in their field. The aerial imagery consisted of both RGB and NIR, and instead of NDVI, they derived green NDVI (GNDVI) as the vegetation index. They used GNDVI to select 10–18 points for plant and soil sampling during the growing season. They stated that soybean yield and GNDVI had a positive correlation among all the fields ranging from 0.69 to 0.91 in early July. They found a negative correlation between GNDVI and soil pH in the range  $5.4 < \text{pH} < 8.4$ , which could be partially

explained by IDC symptoms in the canopy. They also observed a higher variation in GNDVI for soil pH > 7.9, which indicates that factors other than pH cause variability in GNDVI.

## **2.5. Analysis of IDC Data**

Different methods can be used to estimate or predict IDC in soybean. Most methods use features from digital images because they provide useful information on the visual symptoms of IDC. Traditional statistical methods that were usually used to estimate chlorophyll could be great indicators of IDC (Rigon et al., 2016; Vollmann et al., 2011) while more sophisticated machine learning models could classify IDC severity using more features from different sources (Bai et al., 2018; Hassanijalilian et al., 2020b; Naik et al., 2017).

### **2.5.1. Statistical Methods**

Linear regression is the model that has been heavily used in estimating chlorophyll or nutrients such as iron and nitrogen in different crops. Since the model only can handle one input feature, it cannot use all the measured or captured features from the plants (Vollmann et al., 2011). Different studies used RGB images to estimate chlorophyll in a controlled environment. The presented models could use one of the three provided channels or a mixture of all of them as a color vegetation index (Rigon et al., 2016). However, the indices that were calculated in the controlled environment are not necessarily good predictors of soybean chlorophyll content measurement in the field condition (Hassanijalilian et al., 2020b).

Another method of implementing chlorophyll estimation for IDC measurement is to study the trendlines of chlorophyll estimations for different iron treatments (Bin et al., 2016) or different soil types (Schenkeveld et al., 2008) through different growth stages. The

differences between treatments could be identified through Dunnett's *t* test (Ferreira et al., 2019a).

### **2.5.2. Advanced Machine Learning Methods**

In contrast to common statistical models, machine learning models are capable of handling and incorporating many input features. Machine learning models are mostly categorized into two main types: supervised and unsupervised models. (Géron, 2017). Supervised models are the ones that target feature is already known and has labels. The target feature is used in training the models and some portion of it will be used for assessment of the model. Regression and classification models are among the supervised models. On the other hand, unsupervised models do not have pre-labeled target feature, and the model tries to put the dataset in different clusters based on the patterns in the input features.

Classification models such as support vector machine (SVM), decision tree, random forest, adaptive boosting (AdaBoost) have been used in classifying different severities of IDC based on extracted features from images of handheld cameras (Bai et al., 2018; Naik et al., 2017; Zhang et al., 2017) or from aerial images of UAV (Dobbels and Lorenz, 2019). The purpose of these studies was to build machine learning models to predict visual rating that was done by a human.

The input features that were used in most of these studies to train the models were the ratio of different ranges of hue compared to total pixels that were occupied by soybean. Using only the hue might not give a complete picture of soybean health status as Hassanijalilian et al. (2020b) observed that adding other channels improved the performance of the random forest model in predicting the chlorophyll content of soybean leaves.

Moreover, other features such as canopy size could be used as input features because stunt growth is one of the main symptoms of IDC (Hassanijalilian et al., 2020a).

The IDC dataset is often imbalanced and it can affect both training and assessment methods. Different techniques such as defining proper weight for each class or synthetic minority oversampling technique (SMOTE) could be used to deal with the imbalance nature of these kind of datasets. In assessing the models that were trained on imbalance datasets, different metrics such as precision, recall, f1-score, and especially false positive rate could be used instead of accuracy which misleads the assessments (Géron, 2017; Hassanijalilian et al., 2020a)

## **2.6. Management of IDC**

Even though there are different management practices to alleviate IDC symptoms, some practices require farmers to extra steps in their managements such as planting a small grain companion crop (Kaiser and Bloom, 2018). Oats (*Avena sativa* L.) or wheat (*Triticum aestivum* L.) can reduce both moisture and nitrate in the soil, which both exacerbate the IDC symptoms for soybean. However, cereal crops should be killed before they start competing with soybean over resources such as soil moisture (Bloom et al., 2011). Overall it has been observed that the companion crops are not providing consistent results and it is not always profitable (Kaiser et al., 2014; Naeve, 2006).

Another management that requires an extra step is to apply iron chelates either to the soil or as foliar application (Bin et al., 2016; Gamble et al., 2014; Rodríguez-Lucena et al., 2010; Schenkeveld et al., 2008). Iron chelates provide Fe in a usable form for soybean roots, however, not all chelates have the same performance. Some release the iron at a slower rate



and create a higher availability for soybean root (Polyakov et al., 2019), and some foliar applications only improves the visual symptoms and not the final yield (Liesch et al., 2011).

Increasing the seeding rate or planting soybean in wider rows is another practice that potentially can reduce IDC in the field (Goos and Johnson, 2000; Naeve, 2006). This could be because of reduced soil moisture under the row, and the increase in acidity of the root zone that provides a more friendly environment for the root to take up soluble iron (Goos and Johnson, 2001). Moreover, planting soybean in wider rows provides farmers with more choices to add other IDC prevention practices such as adding iron chelate, which increase the concentration of chelates in the furrow that aid in the uptake of soluble iron and improve the chlorophyll content of the crop (Goos and Johnson, 2000).

Finally, the most efficient method to avoid IDC is to plant a tolerant cultivar (Helms et al., 2010; Kaiser et al., 2011; Naeve, 2006). Tolerant cultivars could produce the same amount of yield compared to when susceptible cultivars are planted with companion crops or iron chelate (Kaiser et al., 2014). Minnesota farmers revealed in the survey that the majority of them use tolerant cultivar as the main practice to avoid IDC (Hansen et al., 2003).

## **2.7. Concluding Remarks and Future Perspective**

Different methods of IDC assessments and management were synthesized and presented. Some of the methods, such as atomic absorption spectroscopy, chlorophyll extraction, and crop sensors are more relevant to lab environments and studies that focus on gene expressions. Other methods, such as visual rating and the use of remote sensing in different platforms are more relevant in field settings of larger scales and studies that focus more on selecting tolerant cultivars.

Remote sensing is one of the tools that enables researchers to monitor soybean in larger scales and in a repeatable manner. However, the majority of studies were dedicated to normal RGB or multispectral cameras. The capability of these sensors is limited regarding the spectral signature of the plant. On the other hand, studies with hyperspectral sensors regarding IDC detection in the field are scarce. Similarly, LiDAR as an emerging tool can create a 3-D point cloud of soybean canopy and this could be used along with other sensors to detect the chlorosis in different trifoliolates. This could be helpful since iron is immobile in soybean and younger leaves are affected by IDC. Additionally, finer resolution soil grid sampling could be used as input features to machine learning models along with aerial images and historical data to study the different factors affecting IDC severity.

# 3. CHLOROPHYLL ESTIMATION IN SOYBEAN LEAVES INFIELD WITH SMARTPHONE DIGITAL IMAGING AND MACHINE LEARNING \*

## 3.1. Abstract

Soybean (*Glycine max* (L.) Merrill) leaf chlorophyll content is indicative of the plant growth and health issues. However, chlorophyll measurement using the standard chemical procedure is laborious, while the sensor-based electronic options, such as soil plant analysis development (SPAD) meter tend to be highly expensive and made only spot measurements. Therefore, a simpler and less expensive infield method of chlorophyll measurement in soybeans using smartphone camera with image processing and machine learning models was developed. Soybean leaf images (720 images) and SPAD readings were collected from different cultivars (4), with replications (3) and sampling dates (2) from experimental plots. Of the several color vegetation indices (CVIs) tested, the dark green color index (DGCI) had the best correlation with SPAD meter readings ( $r = 0.90$ ), which was further improved by color calibration ( $r = 0.93$ ). The results of the random coefficients model showed that both cultivars and sampling dates had no significant effect ( $0.06 \leq P \leq 0.96$ ), hence data were combined for the analysis. The simpler statistical linear regression (SLR) and polynomial

---

\* This paper has been published in the *Computers and Electronics in Agriculture* journal. Authors: Oveis Hassanijalilian, C. Igathinathane, Curt Doetkott, Sreekala Bajwa, John Nowatzki, and Seyed Ali Haji Esmaeili. Title: Chlorophyll estimation in soybean leaves in field with digital imaging and machine learning. Year: 2020. Volume: 174. Manuscript No.: 105433. DOI: <https://doi.org/10.1016/j.compag.2020.105433>. Oveis performed the literature survey and wrote the manuscript. Dr. Igathinathane Cannayen is the major advisor and the corresponding author who worked with Oveis throughout the manuscript development. All the co-authors have assisted in the research direction and review of the manuscript.

regression (PR), multiple linear regression as well as the advanced machine learning models (support vector machine (SVM), random forest (RF)) tested with color scheme inputs (RGB, DGCI, range pixel count (RPC) of DGCI, and 'Both' (RPC+RGB)) produced the best chlorophyll prediction with DGCI, RPC, and 'Both' inputs ( $0.87 < R^2 < 0.89$ ;  $2.90 \leq \text{RMSE} \leq 3.41$  SPAD units). Overall, these models were not significantly different, but the SVM model found to be the best ( $R^2 = 0.89$  and  $\text{RMSE} = 2.90$  SPAD units). The simpler SLR and PR models with DGCI input ( $R^2 \geq 0.87$  and  $\text{RMSE} \leq 3.1$  SPAD units) performed as good as the advanced SVM and RF models. The SVM model had the potential of predicting the chlorophyll directly with the raw RGB input ( $R^2 = 0.86$  and  $\text{RMSE} = 3.20$  SPAD units) without the need of using the standard calibration board. The developed methodology of image processing with machine learning modeling and conversion relationship of measuring infield soybean leaf chlorophyll is efficient, inexpensive, not requiring the standard calibration board, and can be easily extended to other large-scale aerial imaging platforms and field crops.

### 3.2. Introduction

Monitoring chlorophyll in soybean (*Glycine max* (L.) Merrill) crops can be of great use in assessing growth and detecting important issues such as iron deficiency chlorosis (IDC). The IDC causes significant soybean yield loss on the order of \$120 million per annum across  $1.8 \times 10^6$  hectares in the Midwestern United States (Hansen et al., 2004). The IDC can reduce plant height and leaf area, and is characterized by a reduction in the leaf chlorophyll content which turns the color from green to yellowish (Vasconcelos and Grusak, 2014). One of the standard methods to measure chlorophyll is the laboratory chemical extraction and calculation based on spectrometer reading using the Von Wettstein (1957) equation.

However, this method is time-consuming, labor-intensive, and difficult to perform on a large number of samples. On the other hand, a sensor-based chlorophyll meter method provides a quick and easy measurement of infield chlorophyll content (Uddling et al., 2007), but tend to be expensive. Therefore, it is necessary to develop a practical and cost-effective method for soybean infield chlorophyll measurements.

A commonly used for chlorophyll measurement device is the soil and plant analysis development (SPAD) meter (Minolta, 1989) and is calibrated for several field crops. Unfortunately, SPAD meters are relatively expensive, and based on soybean trifoliolate measurements, the sample area for an individual measurement is small ( $6 \text{ mm}^2$ ) compared to the average entire leaf size of about  $2470 \text{ mm}^2$  (Page et al., 2018). Thus, the very small SPAD sample area will not be necessarily a representative of the whole leaf or the plant. Moreover, users can introduce error by using the chlorophyll meter on veins (typically green under early IDC), potentially leading to a false reading (McClendon and Fukshansky, 1990).

Digital imaging combined with image processing is another technique that has been proven to be useful in measuring chlorophyll (Vollmann et al., 2011). In another study, images from smartphone were used in soybean chlorophyll estimation (Rigon et al., 2016); however, the measurements were performed in the controlled conditions of a greenhouse.

From the digital images, many color vegetation indices (CVIs) can be calculated to assess plant health (Karcher and Richardson, 2003). It was found that with soybean leaves in a controlled environment at the R3 growing stage (Fehr et al., 1971) green and hue CVIs had a strong correlation with chlorophyll meter readings (Vollmann et al., 2011). Similarly, in a controlled environment at different growing stages of a single cultivar of soybean, the hue, saturation, and brightness CVIs showed a strong correlation with chlorophyll meter readings

(Rigon et al., 2016). However, processing the pictures of soybean in a controlled environment may not be representative of images that are taken in field conditions, where lighting may vary at different sampling times. This could have a direct effect on the appearance of the colors of the leaves (Bascle et al., 2006), and leads into potentially inaccurate measurements. A standard calibration board was used in indoors and in the field conditions to account for differences in artificial and natural lighting conditions and the variability decreased significantly ( $R^2 = 0.85$ ) (Rorie et al., 2011).

The extracted CVI from images can be used as input in machine learning models to predict the chlorophyll content. Machine learning has been used to predict the chlorophyll content of fresh-cut rocket (*Eruca sativa* Mill.) leaves (Cavallo et al., 2017), and to identify the field attributes for soybean yield prediction (Smidt et al., 2016). A simple linear regression (SLR) and a random forest (RF) models were the two machine learning methods used in these studies. Compared to SLR, the RF model can handle large datasets and high-dimensional data without the risk of overfitting. Similarly, support vector machine (SVM) has been used widely alongside machine vision for agricultural applications, and are very effective in field conditions (Rehman et al., 2019). Along with SLR and RF, we propose to use polynomial regression (PR), multiple linear regression (MLR), and SVM to study their performance in chlorophyll estimation through image processing.

To the best of our knowledge, no studies on soybean using smartphones have analyzed the effect of different cultivars and different sampling times on the relationship between CVI and chlorophyll meter readings in field conditions in assessing the chlorophyll content of leaves. Furthermore, the use of the smartphone as a contactless imaging device, and the standard calibration board to correct the images for lighting variations in the field,

along with the developed machine learning models will improve the affordability and accuracy of measurements, hence worth an investigation. The results generated even though is applicable to single leaf chlorophyll measurement, the proposed overall methodology can be extended to plot- and field-scales using aerial images through appropriate modifications.

Therefore, the main objectives of this research study were to (i) evaluate the relationship between the SPAD meter readings and CVIs derived from smartphone images of soybean plants for chlorophyll estimation in field condition, (ii) evaluate the effect of different cultivars and sampling time on chlorophyll estimation, and (iii) compare statistically the performance of the selected five machine learning models in chlorophyll prediction and make recommendations.

### **3.3. Materials and Methods**

#### **3.3.1. Field Experimental Plots**

Four different cultivars of soybean (W3018R2, Mycogen 5N050R2, Asgrow AG0333, and Dahlman 5311NRR2Y) were planted in Leonard, ND, USA (46.671 783° N, 97.245 939° W) for the experiments. The average temperature for the month of May, June and July, 2015 were 15.0 °C, 19.4 °C and 22.2 °C, respectively, based on the nearest weather station (46.729 381° N, 97.283 470° W). Furthermore, total rainfall for these months were 134.0, 108.1, and 36.6 mm, respectively (NDAWN, 2015).

The experimental plots with three replications each (4 cultivars × 3 replications = 12 plots) were seeded at a density of 30 seeds/m in each row on May 21, 2015 (fig. 3.1). Each plot was 3.4 m long and 3.0 m wide, and had four rows with a row spacing of 0.76 m. Since the research sites were not large, weed management was performed manually to avoid potential negative symptoms likely to be caused by the herbicides.



**Figure 3.1. Soybean experimental field plots for various cultivars and replications showing various levels of IDC symptoms.**

### **3.3.2. Field Data Collection — SPAD Meter and Digital Images**

Field data, in the form of SPAD meter readings and digital color images, were collected from each plot between 10:00 to 14:00 CDT to maintain consistent light conditions on the two dates, namely July 1st and July 15th, 2015. When data were collected, the soybean plants were at V3 (Fehr et al., 1971) and V5 growth stages. These stages were selected because the breeders rate the crop twice at these stages where the IDC severity usually changes, hence displaying different levels of chlorophyll for measurements (Helms



et al., 2010). From each plot having four rows, leaf samples only from the middle rows (2nd and 3rd) were collected, avoiding the edge effects of the border rows (1st and 4th). The selected leaf was clipped from the plant and the SPAD meter reading was taken first followed by the digital image of the same leaf arranged on the standard calibration board. With 30 selected leaves per plot with 12 plots in total resulted in 360 observations at one sampling date (720 observations for both dates). As the IDC issues were common in the region, several plants in the experimental field displayed IDC symptoms (fig. 3.1).

### **3.3.3. Overall Methodology and Analysis**

The various methods involved in the experiments and the analysis for the chlorophyll estimation in soybean leaves using the digital image and machine learning models with SPAD meter reading as reference are outlined in figure 3.2 and described subsequently.

#### **3.3.3.1. Chlorophyll measurement**

A SPAD meter (Model: SPAD-502; Konica Minolta Sensing, Osaka, Japan) was used to directly estimate the chlorophyll content of soybean leaves. In each plot, the uppermost fully developed leaves of 30 different plants were used for the SPAD meter readings. Since the sampling sensor area of the SPAD meter is small, the measurements were taken from three different spots of every leaf, for representative measurements, and the average was used as the SPAD meter readings (Roriz et al., 2014). As the plots displayed the IDC symptoms, a range of leaf samples with different chlorophyll contents was purposefully selected from a wide spectrum of colors from yellow to green. While making the measurements, the veins of the leaf were cautiously avoided as IDC will have the least effect on them. Finally, the SPAD meter readings were converted to actual chlorophyll content (Markwell et al., 1995) using

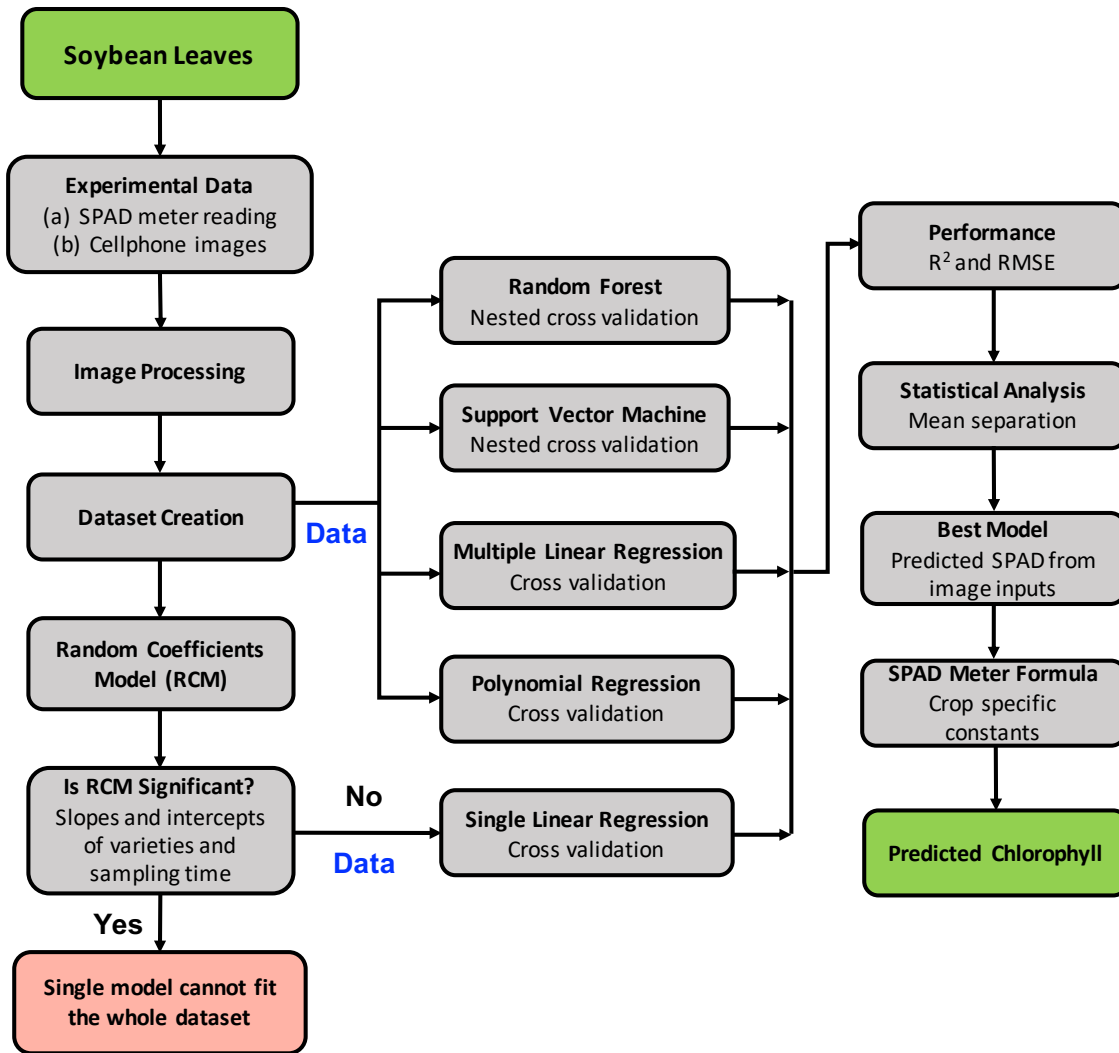


Figure 3.2. Flowchart of the proposed image processing and machine learning model approach developed to predict chlorophyll content of soybean leaves.

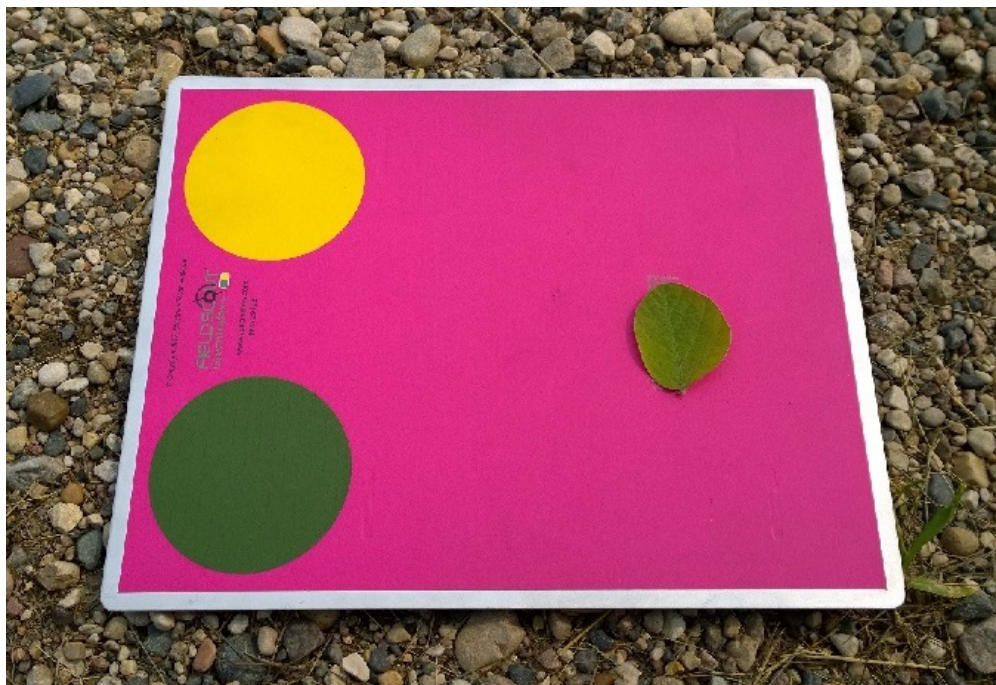
the following conversion relationship (eq. 3.1):

$$CC = 10^{(SMR^{0.265})} \quad (3.1)$$

where CC is chlorophyll content ( $\mu\text{mol m}^{-2}$ ), and SMR is the SPAD meter reading. It should be noted that the conversion relation type and coefficients (eq. 3.1) will vary among different makes of SPAD meters and they also require their calibration before measurements.

### 3.3.3.2. Image acquisition for chlorophyll measurement

Digital color images of the clipped soybean leaves were acquired with a 38 MP smartphone camera (Nokia Lumia 1020) for the same 30 leaves that were also used for the SPAD meter readings from each plot. Double-sided tape was used to attach and flatten the soybean leaves to the middle of a standard calibration board (FieldScout, Aurora, IL; fig. 3.3). The standard calibration board was a 230 mm × 290 mm pink color board with two 90 mm diameter discs, one yellow and one green. The leaves were attached to the pink area of the board, while the green and yellow discs were used to calibrate the images to account for changes in ambient light.



**Figure 3.3. Soybean leaf placed on the standard calibration board.**

Munsell color values of the yellow disc were 5 Y 8/11.1 (#e7c71f hex; 50, 87, 91 HSB; and 231, 199, 31 RGB values) and for green disc were 6.7 GY 4.2/4.1 (#576c43 hex;

91,38,42 HSB; and 87, 108, 67 RGB values). A similar standard calibration board was also used to account for different cameras and illumination scenarios (Rorie et al., 2011).

The smartphone was set to operate in automatic mode, so the phone could set the shutter speed and aperture values to maintain a constant amount of incoming light to the camera. The resolution of the images was on the highest setting of  $3072 \times 1728$  pixels, and the images were saved in a JPEG format. All images (330 usable images per sampling period) were taken so the whole standard calibration board was in the middle of the frame, and at an oblique angle to avoid glare or shadow on the board.

### **3.3.4. Image Preprocessing for Objects Recognition and Indices Determination**

The aim of image preprocessing (objects recognition) is to automatically identify the objects in the standard calibration board with leaf (the calibration board itself, yellow disc, green disc, and leaf) and extract the pixel color information for various CVIs, for example, the dark green color index (DGCI). Images were preprocessed with MATLAB (2015a) Image Processing Toolbox. Suitable codes were written to perform all the stages of the analysis. The various image preprocessing stages involved in the object recognition for the extraction of the color information are depicted in figure 3.4.

#### **3.3.4.1. Object recognition - standard calibration board**

Images were imported to Matlab as a  $3072 \times 1728 \times 3$  (layers) unsigned 8-bit integer matrix, which has layers of red, green and blue (RGB). The RGB components of the dominant pink of the board (fig. 3.4A1) fall in the approximate range of  $R > 220$ ,  $G < 60$ , and  $B > 125$ . Therefore, it is just enough to use red and green bands in the histogram (fig. 3.5) and their thresholds were determined to extract the standard calibration board. Using these red and

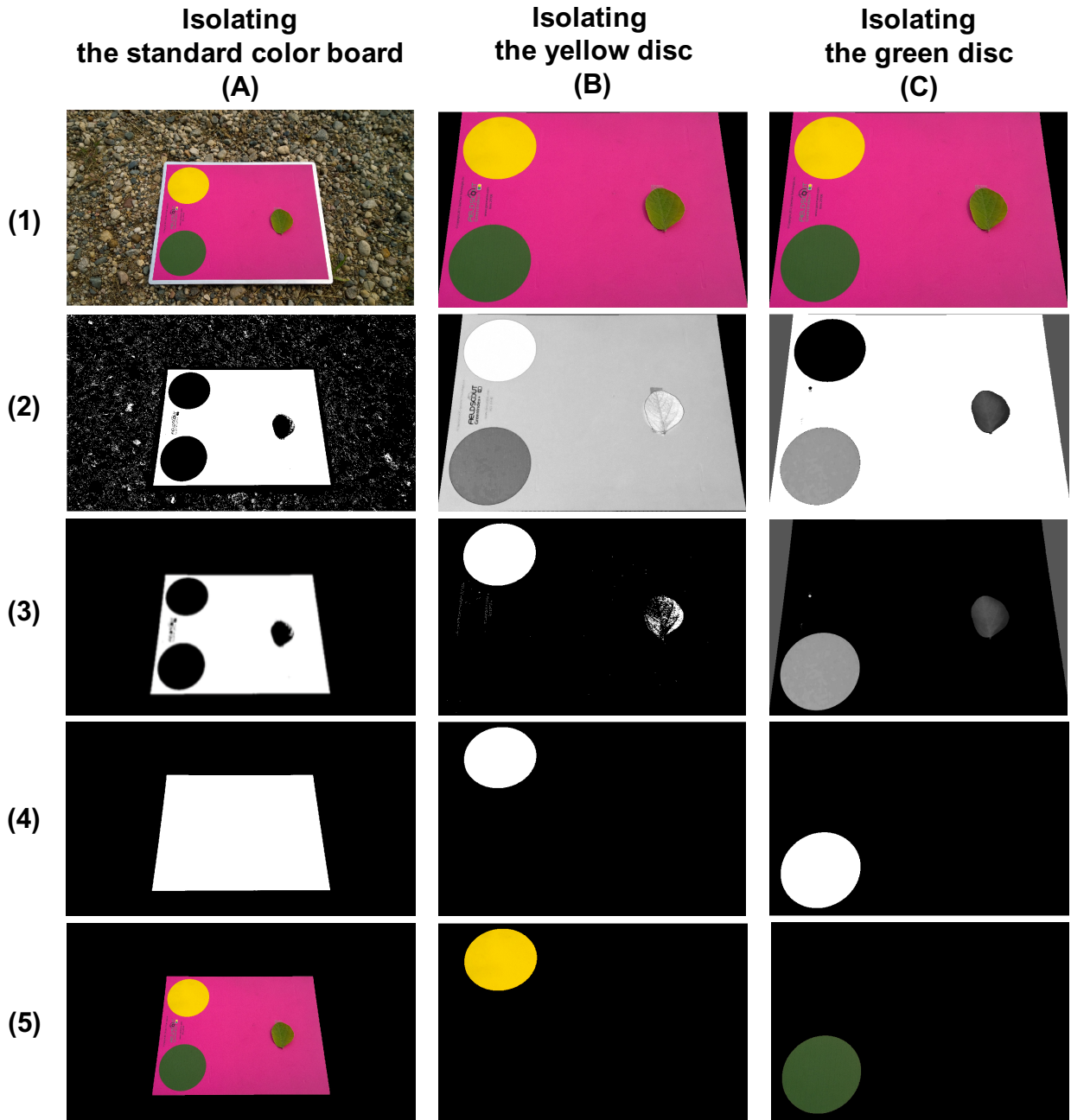
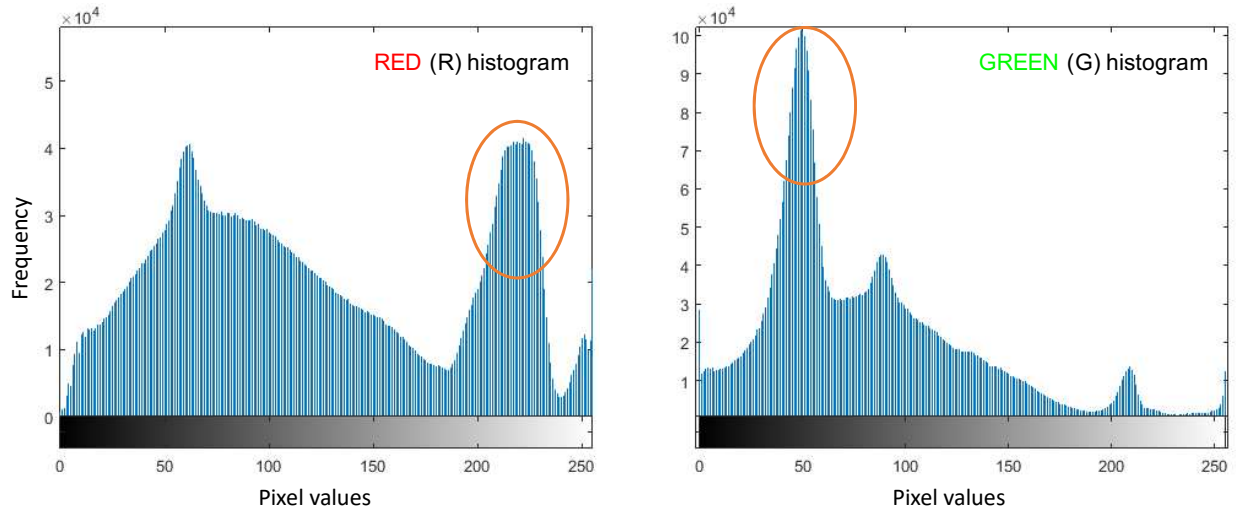


Figure 3.4. Image preprocessing operations of object recognition and isolation of the standard calibration board, and the yellow and the green discs for extraction of color information.

green thresholds, a binary image was created that was used to create a mask for extracting the standard calibration board (fig. 3.4A2).



**Figure 3.5. Detecting the pink portion of the standard calibration board through red and green histograms. The highlighted peaks regions of both histograms are selected the pink portion of the standard calibration board.**

The small white speckles (fig. 3.4A2) representing the background soil particles were removed next by a pixel area threshold  $<500\,000$  pixels by the `bwareaopen()` function of Matlab (fig. 3.4A3). Then the remaining larger objects, such as the green and yellow discs, text logo, and the leaf were filled with white using the `imfill()` function (fig. 3.4A4). This filled binary image was used as the mask to isolate the standard calibration board from the input image (fig. 3.4A5). For the extraction for the reference color discs and the leaf, this identified standard calibration board image serves as the input (fig. 3.4B1 and fig. 3.4C1).

### 3.3.4.2. Object recognition - yellow disc

For the extraction of the yellow disc, as the yellow is the most saturated color in the image, the RGB color space was converted to hue ( $H'$ ), saturation ( $S'$ ), and brightness ( $B'$ )

HSB color space. The conversion of RGB to  $H'S'B'$  can be obtained using standard expressions (Karcher and Richardson, 2003) or through Matlab function `rgb2hsv()`.

From the S channel image (fig. 3.4B2), by defining the threshold of  $S > 0.9$  the binary image was created (fig. 3.4B3). In the binary image after applying erosion followed by dilation using `imopen()` function removed small artifacts and broken up the leaf. Again with `bwareaopen()` using 50 000 pixels threshold the remaining parts of the leaf were removed and the mask was created (fig. 3.4B4). Then from the input (fig. 3.4B1) the created mask the yellow disk was extracted (fig. 3.4B5).

### 3.3.4.3. Object recognition - green disc

For the extraction of the green disc, the DGCI (Karcher and Richardson, 2003) was used as a basis and was calculated using the  $H'$ ,  $S'$ , and  $B'$  values as follows (eq. 3.2):

$$\text{DGCI} = \frac{(H' - 60)/60 + (1 - S') + (1 - B')}{3} \quad (3.2)$$

Applying the DGCI scheme on the input image (fig. 3.4C1), the grayscale image was obtained that made the pink portion of the board white (fig. 3.4C2). To isolate the green disk, pixels with a grayscale value greater than 0.9 were removed, which corresponded to the pink portion of the board. This step's output image consisted of the green disk, the leaf, and other small objects (fig. 3.4C3). To further isolate the green disk and create the mask, the `bwareaopen()` function was used to remove objects with an area smaller than 50 000 pixels that removed the leaf and minor artifacts (fig. 3.4C4). This mask, when applied to the input image, extracted the green disc (fig. 3.4C5).

#### 3.3.4.4. Object recognition - leaf

In the DGCI image (fig. 3.4C3), the leaf was the second-largest object besides the green disc portion. By applying the created mask green disc (fig. 3.4C4) to the DGCI image (fig. 3.4C3), the mask of the leaf was created. Further processing such as `bwareaopen()` (4000 pixels threshold) followed by the `imopen()` with a disk shape structural element removed the leaf stem (if present) and preserved the leaf shape and created the final leaf mask. This leaf mask, when applied to the input image (fig. 3.4C1), extracted the leaf from the standard calibration board, and along with this, a few other leaf samples are shown in Figure 3.6.

#### 3.3.5. Color Vegetation Indices

From the recognized leaf object in the image (fig. 3.6), all three channels of RGB were extracted. The RGB channels of the whole leaf extracted were converted to chromatic coordinates (eq. 3.3), which were used to calculate the excess green index (Hamuda et al., 2016) as follows:

$$r = \frac{R^*}{R^* + G^* + B^*} \quad g = \frac{G^*}{R^* + G^* + B^*} \quad b = \frac{B^*}{R^* + G^* + B^*} \quad (3.3)$$

where  $r$ ,  $g$ , and  $b$  are the chromatic coordinates.  $R^*$ ,  $G^*$ , and  $B^*$  are the normalized RGB values ranging from 0 to 1 (eq. 3.4), and are expressed as:

$$R^* = \frac{R}{R_{\max}} \quad G^* = \frac{G}{G_{\max}} \quad B^* = \frac{B}{B_{\max}} \quad (3.4)$$

where  $R_{\max}$ ,  $G_{\max}$ , and  $B_{\max}$  are the maximum pixel values in the extracted image object with the respective color channels. Since the amount of red and blue influences how the green





**Figure 3.6. The extracted leaf sample (top-left) from the standard calibration board image (fig. 3.4) and a few others showing different levels of chlorophyll in the V3 to V5 stages of the study.**

color appears, the RGB color space was converted to HSB color space as suggested by Karcher and Richardson (2003). A further benefit of HSB color space is that it is closer to the human perception of color and is device-independent. The DGCI (eq. 3.2) is designed to work best with the hues present in the leaves (max: 120 and min: 60); however, the average RGB values were also used for simplicity and model development.

The various CVIs that have been generally used in plant health estimation and plant detection by various researchers (table 3.2) were employed in this study to select the CVI that

correlates best with the chlorophyll measured using SPAD meter. The Pearson correlation coefficients ( $r$ ) were calculated for all leaves between each mean CVI and SPAD meter readings for the selection of CVI. Based on the results presented later (section 3.4.2), a comparison of the  $r$  between CVIs and SPAD meter readings indicated that the DGCI had the best correlation. Therefore, DGCI values were further processed by calibrating the images for illumination with a two-point linear calibration using the color disks of the standard calibration board. Since RGB had lower  $r$  values compared to DGCI, they were not further calibrated for ambient light variation. However, to study their application, as RGB is the most common and simplest values extracted from the images, they were considered as one of the basic input color schemes.

For calculating slope of the calibration line, the difference between the known and observed values of both color disks were required. Yellow and green disks had known DGCI values of 0.0178 and 0.5722, respectively. The average DGCI values of both disks were derived from the images and used as observed values to calculate the calibration line slope and  $y$ -intercept as described in equations 3.5 and 3.6:

$$\text{Slope} = \frac{0.5722 - 0.0178}{\text{Observed green DGCI} - \text{Observed yellow DGCI}} \quad (3.5)$$

$$y\text{-intercept} = 0.5722 - (\text{Slope} \times \text{Observed green DGCI}) \quad (3.6)$$

The illumination-corrected value for DGCI of the leaf were calculated with the calculated slope and  $y$ -intercept, as shown in equation 3.7:

$$\text{Corrected leaf DGCI} = (\text{Slope} \times \text{Observed leaf DGCI}) + y\text{-intercept} \quad (3.7)$$

The corrected leaf DGCI values are expected to compensate for the differences in the lighting conditions (Rorie et al., 2011).

#### **3.3.5.1. Determining the effect of cultivar and sampling date using random coefficients model**

The random coefficients model (RCM) was used to assess the impact of plots, cultivars, and sampling time on the relationship between corrected leaf DGCI values (eq. 3.7) and SPAD meter readings. The RCM, also called multi-level or hierarchical model (Hsiao, 1975; LaHuis and Ferguson, 2009), was developed using SAS software `proc mixed()` method (SAS, 2015). This model could specify variance components for individual factors to look at levels represented as a random effect rather than a fixed effect.

Factors typically treated as random effects include blocks and plots. For this study plots within dates (11) were identified as random effects, while dates (2) and cultivars (4) were identified as fixed effects in the RCM.

#### **3.3.6. Machine Learning Models**

Five machine learning models of varying complexity were studied to fit the image-based data to the chlorophyll measurement through SPAD meter. Three statistical (SLR, PR, and MLR) and two advanced models (SVM and RF) were selected for the study. It should be noted that the application of more advanced machine learning methods, such as neural network and deep learning are expected to perform better when the size of the dataset is big. Since, in this study, the dataset was relatively small, the above simpler machine learning algorithms were employed. The “Scikit-Learn” module in Python (Pedregosa et al., 2011) was used for models development as well as for validation (cross- and nested

cross-validation), and R (R Core Team, 2018) for model performance testing and results visualization.

The simpler statistical models (SLR and PR) used two sets of color schemes as independent variables, such as a combined average RGB into one grayscale (RGB\*), which was calculated by `rgb2gray()` function of Matlab, and the average value of DGCI for all the pixels of each leaf was used as inputs to these models. But the other models used the three sets of independent variables such as (i) RGB — average of each three bands individually, (ii) Range pixel count (RPC) — pixel counts of different ranges of DGCI (0.0–0.6 with 0.15 interval) that were normalized by the size of each leaf, and (iii) ‘Both’ — RGB+RPC. For the combined color scheme (‘Both’), the RGB values and RPC were combined to form the seven input variables. Since the scale of input variables affect the performance of some of the machine learning models, all the input values were normalized with their respective mean and standard deviation values.

Assessments of the performance of the models and parameters were based on the model coefficient of determination ( $R^2$ ) and the root mean square error (RMSE, expressed in SPAD units) of the predictions. The best models are those that produce the highest  $R^2$  and the lowest RMSE. Further details of the models’ development, validation, and performance and statistical analysis are presented subsequently.

For validation, a nested cross-validation (CV) was used to tune the parameters of the advanced models (SVM and RF), and tested the models with the held-out dataset (Cavallo et al., 2017). Since the rest of the statistical models (SLR, PR, and MLR) do not have any parameters to be tuned, only the outer CV of the nested CV was used to assess their performance with the same held-out dataset. In this study, 10 folds were selected for both

inner and outer CVs. The  $R^2$  was selected as the criteria to pass the best parameters from the inner CV to the outer CV.

### 3.3.6.1. Simple, polynomial, and multiple linear regression

The SLR utilizes ordinary least square method to generate a regression line by minimizing the sum of squares of the differences between the predicted and the observed values. The SLR model with usual notations (eq. 3.8; Neter (1974) ) is:

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (3.8)$$

where  $Y$  is the output or dependent variable (SPAD meter readings),  $\beta_0$  is the intercept and  $\beta_1$  is the slope of the regression line,  $X$  is input or independent variable (RGB\* or DGCI values), and  $\epsilon$  is a random error term.

Following the SLR (eq. 3.8) the PR also called a second-order model with one independent variable  $X$  (eq. 3.9), and MLR models with additional independent variables (eq. 3.10) are defined as (Neter, 1974):

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \epsilon \quad (3.9)$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \epsilon \quad (3.10)$$

where  $\beta_0, \beta_1, \dots, \beta_i$  are the parameters,  $X_1, \dots, X_i$  are the independent variables, and  $i = 1, 2, \dots, n$  are the indices.

The ordinary least square method was used to solve SLR (eq. 3.8) with LinearRegression module of Scikit-Learn. The same module was used to solve the PR (eq. 3.9), but PolynomialFeatures were used to add the second-degree features and

interactions to the independent variables. In MLR, not all variables (eq. 3.10) increase the performance of the model, so a backward elimination method, using Statsmodels module in Python, with the exit  $\alpha$  level of 0.05 was chosen to eliminate variables that were insignificant to the model.

### **3.3.6.2. Support vector machine**

An SVM is a powerful model that can be used for both classification and regression, which was found to be an effective model in field conditions (Rehman et al., 2019). The procedure for classification can be thought of fitting the widest “street” possible among the data, and the objective of SVM regression is to fit as much data possible in the street. Several parameters, such as the kernel, degree, and gamma influence the performance of SVM models and they should be properly selected for better performance (Géron, 2017).

These SVM model parameters were tuned through nested cross validation with the following values, as initial guess ranges, used in this study are: (i) C (the penalty parameter — 0.1 to 2.0 with 0.2 interval), (ii)  $\text{coef0}$  (independent term in kernel function — 0 to 10 with 0.2 interval), (iii) degree (2 to 6) and gamma (0.01, 0.1, and 1 ), and (iv) kernel (poly and rbf).

### **3.3.6.3. Random forest**

The RF model is a group of decision trees used for either classification or regression. Each decision tree predicts the outcome individually, and the RF model either votes among the outcomes for classification or averages the outcomes for regression. In RF models different subsets of input variables can be utilized in each tree, which makes it more useful in prediction for datasets with higher dimensions (Géron, 2017). Each set of independent and dependent variables was used to train the random forest models with Scikit-Learn module in

Python (Pedregosa et al., 2011), and the trained random forest models were used to predict the SPAD meter readings.

In order to have a better prediction, parameters of the random forest model were tuned with a nested cross-validation method. The parameters and their defined values were: (i) maximum depth of trees (4 to 10), (ii) the number of trees that composed the forest (25 to 50), and (iii) the minimum number of samples required to split an internal node (2 to 10).

#### **3.3.6.4. Models performance and statistical analysis**

The performance of all models was compared statistically ( $\alpha = 0.05$ ) based on the performance parameters ( $R^2$  and RMSE) using Duncan's mean separation analysis (De Mendiburu, 2019) using the R programming environment. The mean separation analysis was performed individually with the combination of models and color scheme inputs as well as pooled data to compare the performance of models and color scheme inputs individually based on performance parameters.



















### **3.4. Results and Discussion**

#### **3.4.1. Field Measurement Results**

The mean and the statistical distribution of the image-based (RGB) and SPAD meter readings corresponding to the two sampling dates are presented in table 3.1. The generated color patches based on the observed RGB data help in visual comparison. The different colors of the patches can be found in the various regions of the original leaf image (fig. 3.6). For comparison, the reference digital colors (hexcolor and fixed RGB shown) that matched the three colors of the standard calibration board are also shown (table 3.1).

The results and color patches indicate that the leaves RGB is more close to the green disc than the yellow disc (table 3.1). The minimum of the RGB data tended towards

**Table 3.1. Distribution of the RGB values collected with smartphone camera and SPAD meter readings for both sampling dates ( $n = 330$ ).**

Data (count = 330)	July 1st, 2015					July 15th, 2015				
	Red <sup>1</sup>	Green <sup>1</sup>	Blue <sup>1</sup>	Color	SPAD <sup>2</sup>	Red <sup>1</sup>	Green <sup>1</sup>	Blue <sup>1</sup>	Color	SPAD <sup>2</sup>
Mean	115.45	112.01	23.30		16.27	120.82	118.27	34.15		18.24
Standard deviation	26.90	14.49	10.02	–	8.29	30.72	18.73	9.50	–	9.18
Minimum	58.10	75.66	9.15		1.80	64.72	71.29	13.62		2.00
First quartile	92.58	100.80	15.67		9.20	95.69	104.20	26.96		10.40
Median	115.41	112.86	20.80		15.45	118.13	117.80	33.17		17.80
Third quartile	138.76	123.50	30.14		23.18	143.21	131.25	40.64		25.80
Maximum	172.03	145.28	54.61		37.60	210.94	169.02	60.38		40.50
Yellow disc* (#E7C71F)	231	199	31		–	231	199	31		–
Green disc* (#576C43)	87	108	67		–	87	108	67		–
Pink board* (#EA328A)	234	50	138		–	234	50	138		–

<sup>1</sup> Values vary between 0 and 255; <sup>2</sup> Higher values correspond to higher chlorophyll contents; \* Standard values of the color calibration board; and color patches were derived from the corresponding RGB data presented for information and may not refer to the same leaf as the distribution derived from 330 observations.

reference green and the maximum towards reference yellow, and the mean data clearly belonged to the range of these reference colors. For both sampling dates the mean and median coinciding, with quartiles falling symmetrical on both sides was indicative of the normal distribution of the data. Overall, the passage of time (14 days) had only a slight darkening effect during this short period. It should also be noted that these results are based on the distribution of 330 test samples, and the minimum and maximum are based on different samples; however, the results offer some indications of the leaf color variations.

### 3.4.2. DGCI for Chlorophyll Measurement

The various CVIs and their correlation with SPAD meter reading were determined to identify the best CVI for prediction model development. The Pearson correlation coefficients ( $r$ ) between different CVIs and the SPAD meter readings indicated that DGCI had the highest correlation ( $r = 0.9$ ); this correlation was even greater than several green-based indices (table 3.2). Although Vollmann et al. (2011) reported that in soybean the correlation



between direct green channel and SPAD values had the highest absolute value ( $r = -0.94$ ) in comparison to a number of CVIs in a controlled environment, our study found that under the field conditions the correlation was lower ( $r = -0.77$ ) for green (table 3.2). In contrast, in another study with corn (*Zea mays* L.) Vesali et al. (2015) used a contact imaging method and found no relationship between green and the SPAD values. Given the inconsistencies, using green channel alone would not be sufficient to estimate chlorophyll especially in field conditions.

**Table 3.2. Correlation of color vegetation indices with SPAD index using RGB and chromatic coordinates (HSB).**

Color vegetation index	Definition	$r$	$r$
		RGB	chromatic coordinates
Red	-	-0.84	-0.87
Green	-	-0.77	0.33
Blue	-	0.22	0.57
Hue	-	0.79	N/A
Saturation	-	-0.56	N/A
Brightness	-	-0.84	N/A
DGCI (Karcher and Richardson, 2003)	$[(\text{Hue} - 60)/60 + (1 - \text{Saturation}) + (1 - \text{Brightness})]/3$	0.90	N/A
Gray	$0.2898 \times \text{red} + 0.5870 \times \text{green} + 0.1140 \times \text{blue}$	-0.79	-0.09
ExG (Woebbecke et al., 1995)	$2 \times \text{green} - \text{red} - \text{blue}$	0.05	0.33
ExR (Meyer et al., 1998)	$1.3 \times \text{red} - \text{green}$	-0.82	-0.82
CIVE (Kataoka et al., 2003)	$0.441 \times \text{red} - 0.811 \times \text{green} + 0.385 \times \text{blue} + 18.78745$	0.25	-0.37
ExGR (Meyer and Neto, 2008)	$\text{ExG} - \text{ExR}$	0.45	0.59
GB (Woebbecke et al., 1995)	$\text{green} - \text{blue}$	-0.82	-0.23
ERI (Golzarian and Frick, 2011)	$(\text{red} - \text{green}) \times (\text{red} - \text{blue})$	-0.81	-0.77
EGI (Golzarian and Frick, 2011)	$(\text{green} - \text{red}) \times (\text{green} - \text{blue})$	-0.36	0.77
RG (Woebbecke et al., 1995)	$\text{red} - \text{green}$	-0.76	-0.79
GR	$\text{green} - \text{red}$	0.62	0.79
GBRG (Woebbecke et al., 1995)	$(\text{green} - \text{blue}) / (\text{red} - \text{green})$	0.78	0.33
COM1 (Guijarro et al., 2011)	$\text{ExG} + \text{CIVE} + \text{ExGR}$	0.43	0.51
MexG (Burgos-Artizzu et al., 2011)	$1.262 \times \text{green} - 0.884 \times \text{red} - 0.311 \times \text{blue}$	0.41	0.61

DGCI - Dark green color index; ExG - Excess green; ExR - Excess red; CIVE - Color index of vegetation extraction; ExGR - Excess green minus excess red index; COM1- Combined index 1; and MexG - Modified excess green index; GB - Green minus blue; ERI - Excess red feature; EGI - Excess green feature; RG - red minus green; GR - Green minus red; and GBRG - GB by RG ratio.

Some components of RGB and HSB color space had higher absolute correlation values than green (red:  $-0.857$ , hue:  $0.886$ , and brightness:  $-0.847$ ), whereas these components had lower correlations than that of green (Vollmann et al., 2011). The blue component had a low correlation with the SPAD meter readings ( $r = 0.218$ ); which was supported by some (Vesali et al., 2015; Yadav et al., 2010) but not others (Vollmann et al., 2011).

Notwithstanding these inconsistencies, using individual color channels of images in controlled lighting could help estimate the chlorophyll content. But, these channels are less useful when they have to apply in a field setting, where illumination cannot be controlled. However, the usage of HSB color space-based DGCI, which incorporates all the channels of the image, adjusts the image using the range of hues that are available in the leaf for measurements. Furthermore, the adjusted image is calibrated for changes in ambient light using the standard calibration board. The correlation of the uncalibrated DGCI values with SPAD meter readings ( $r = 0.90$ ) was improved after calibrating the values using standard calibration board ( $r = 0.93$ ). Thus, the DGCI, with possible calibration, is likely to be a more reliable index to estimate chlorophyll content. Therefore, DGCI along with the basic RGB were selected for prediction models development and were recommended for infield image analysis for chlorophyll measurement.

#### **3.4.2.1. RCM results for corrected DGCI**

A preliminary assessment of the proportion of variance accounted for by date, cultivar, and plot through intraclass correlations indicated that each variable was associated with  $<5\%$  of the total variation in the data. Of these three possible sources of variation, the plot accounted for roughly twice as much variation as either cultivar or date. Therefore, the final

model was simplified to use date and cultivar as fixed effects with plot within date as a random effect.

Underlying assumptions of normality and homogeneity of variance were assessed based on the residuals and were well met by this model. The results for the Type III error tests  $F$ -value and corresponding probability ( $\text{Pr} > F$ ) at  $\alpha = 0.05$  for DGCI was 1752.80 and  $<0.0001$ , cultivar was 1.72 and 0.202, and date was 2.65 and 0.122, respectively. This result indicated that the correlation between DGCI and SPAD meter readings is significant; additionally, cultivar and date had no significant effect on this relationship.

The SPAD meter reading of Date-2 was 12 % higher than Date-1 (table 3.1). Similarly, in a foliar magnesium treatment application study, a decline of total chlorophyll at the time of the first application and a raise at the time of second application was observed (Teklić et al., 2009). Therefore, the amount of chlorophyll can change through time due to different factors, and it might not have any effect on the chlorophyll estimation through DGCI.

Parameter estimates for fixed effects of the RCM based on  $t$  values indicated their significance (table 3.3). While comparing the multiple levels of a parameter, a set-to-zero approach obtained the solutions, so the last level of categorical factors (e.g., cultivar-4 and Date-2) was essentially used as a reference level, and the parameter estimates for other levels within a factor were adjustments to the dependent variable compared to the reference level. The adjustment for SPAD values was for about a 4.3 % increase across the dates. This value is based on the average SPAD value of about 20 at the corrected DGCI midpoint of about 0.35 as the estimate of 0.8596 divided by 20 was  $\approx 4.3\%$ . The results (table 3.3) indicate that the intercept and slope of the general linear regression were significantly different than zero, while the effect of levels of cultivars and dates were not significantly different in their groups.

**Table 3.3. Parameter estimates for fixed effects of random coefficients model and comparison.**

Effect	Estimate	<i>t</i> -value	Pr >   <i>t</i>
Intercept*	−5.6493	−7.24	<.0001
Slope*	87.6383	41.87	<.0001
Cultivar-1	−0.0338	−0.05	0.9630
Cultivar-2	−0.5705	−0.80	0.4369
Cultivar-3	−1.6320	−2.03	0.0578
Cultivar-4	0.000	−	−
Date-1	−0.8596	−1.63	0.1220
Date-2	0.000	−	−

\* Parameters (slope and intercept) of the general linear regression in the random coefficient model.

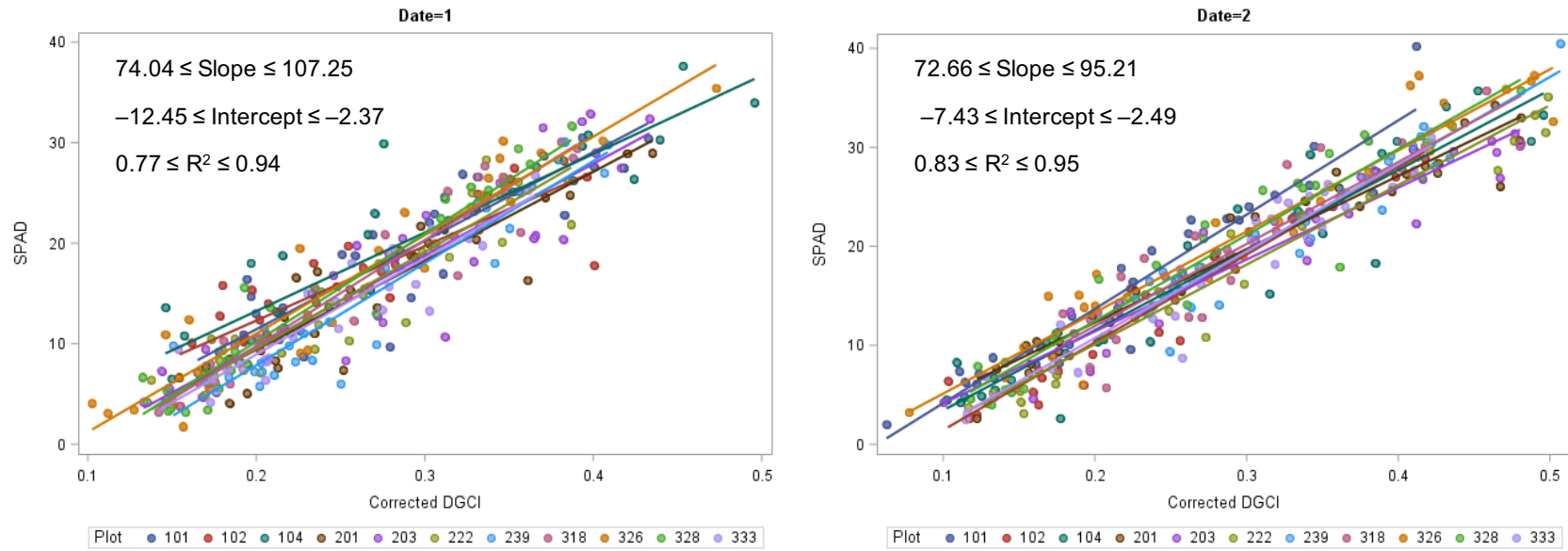
The variability in the regression parameters across the plots within dates is represented in the estimates for the RCM intercepts and slopes. The adjustments to the slope and intercept in this case are relative to the fixed effects estimates for “Intercept” and “Slope” in table 3.3. Most of the adjustments for individual plots are not statistically significant ( $\alpha = 0.05$ ; data not shown).

The plot of regression lines of RCM grouped based on dates showed similar variation based on fit slopes and intercepts (fig. 3.7). Overall, the regression lines fitted well across individual plots and dates ( $0.77 < R^2 < 0.95$ ). Also, the grouping of the regression lines showed an insignificant difference between plots for both slope and intercept, and which indicated that the relationship between DGCI and SPAD meter readings among the plots is not significantly different (fig. 3.7). As a result, all the observations can be combined and analyzed for chlorophyll estimation through the selected machine learning algorithms.

### 3.4.3. Machine Learning Prediction Models Results

#### 3.4.3.1. Single linear regression

While using DGCI, the statistics of the 10 folds validation produced an average  $R^2$  of  $0.87 \pm 0.02$  for the test scores while the training scores average was  $R^2$  of  $0.88 \pm 0.002$  that



**Figure 3.7. Scatterplot of corrected DGCI and SPAD meter readings using RCM for first (left) and second (right) sampling dates. Each line represents a regression line for individual plots.**

were assessed with only the external part of the nested cross-validation. Similarly, the average value for RMSE for the SLR model is  $3.10 \pm 0.26$  SPAD units. Furthermore, from the 10 folds, slopes range from 85.25 to 87.08 and intercepts range from  $-6.36$  to  $-5.91$ .

The training and test scores for all 10 folds of the SLR models are close to each other, which means that the SLR models did not over- or under-fit the data. The fact that the DGCI equation is based on leaf color variations could help the linear model to be a more suitable model for field conditions. However, average DGCI as one independent variable does not provide the model with enough information to have a more accurate prediction compared to more advanced models. On the other hand, using the combined RGB as independent variable did not give a good performance ( $R^2 < 0.56$ ,  $RMSE > 5.88$  SPAD units; fig. 3.8).

#### **3.4.3.2. Polynomial regression**

The DGCI was used as the only independent variable in this model to predict the SPAD meter readings in PR, similar to SLR. The average  $R^2$  is  $0.87 \pm 0.02$  and RMSE is  $3.09 \pm 0.27$  SPAD units, which is the best among all models when the average RGB was used to train the other models. Furthermore, it performed similarly to SVM when RPC was used as independent variables. The PR performed almost the same as the SLR and hence not significantly different from SLR, even though the PR fitted the data non-linearity. Similar to SLR, this model did not have a good performance when it used the combined RGB ( $R^2 < 0.56$ ,  $RMSE > 5.83$  SPAD units; fig. 3.8). This could be because of losing information while aggregating three RGB channels into one.

#### **3.4.3.3. Multiple linear regression**

Since the green and red were highly correlated ( $r = 0.93$ ), and the green had a lower correlation with the SPAD meter readings ( $r = -0.77$ ) compared to red ( $r = -0.84$ ), the

green was removed from the models due to multicollinearity, which should be avoided in MLR modeling (Neter, 1974). Using the remaining average red and blue values, the MLR performance was the lowest ( $R^2 = 0.83$  and RMSE = 3.55 SPAD units) compared to the advanced models ( $R^2 > 0.84$ ; RMSE < 3.4 SPAD units), but when it used the RPC input, it performed significantly better ( $R^2 = 0.87$  and RMSE = 3.11 SPAD units) that is similar to using 'Both' as input. The performance of MLR using RPC or 'Both' was not significantly different than the advance machine learning models.

#### **3.4.3.4. Support vector machine**

When using RGB input with the SVM model, it performed almost the same as simpler models (SLR and PR) with DGCI as input, with an only 3.3 % increase in the RMSE. The SVM model, however, performed the best when 'Both' was used as independent variables. The average  $R^2$  among all 10 folds is 0.89 and RMSE was 2.90 SPAD units, which was a high score for an uncontrolled environment. The better performance of the raw RGB input in combination with SVM model ( $R^2 = 0.86$ ) compared to the other two inputs ( $R^2 = 0.87$ – $0.89$ ), allowed to derive an interesting and useful conclusion of a field chlorophyll measurement using image processing without the standard calibration board. Based on the range of the data (table 3.1), the RMSE values can be normalized to a minimum error of 5.6 %, an average error of 7.5 %, and a maximum error of 8.2 %. These values were the results of models that were built with the selected parameters of the inner CV.

#### **3.4.3.5. Random forest**

In RF models predicting SPAD meter reading from RGB values, although the blue values had a low correlation with SPAD meter readings, the blue values significantly improved the performance of the random forest models. The average  $R^2$  among all 10 folds

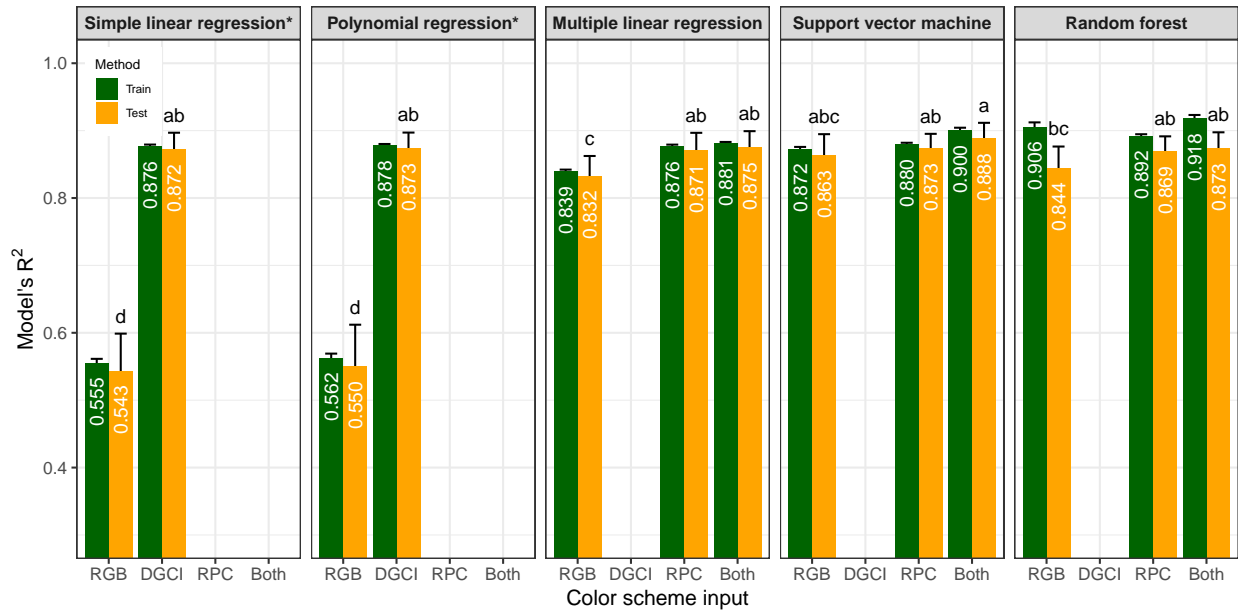
of nested cross-validation was  $0.87 \pm 0.02$  and RMSE of  $3.09 \pm 0.28$  SPAD units, which was again a high score for an uncontrolled environment similar to SVM. The RMSE values can be normalized (based on data table 3.1) to minimum error of 6 %, average error of 8 %, and maximum error of 8.6 %. Even though overfitting occurred in the RF model, observed from the differences between train and test scores, the test scores on unseen data were almost equal to or higher than the SLR model.

#### **3.4.3.6. Performance of the machine learning models**

The performance bar plots present the average and one standard deviation of  $R^2$  (fig. 3.8) and RMSE (fig. 3.9), as the positive component of the error bar, of all the 10 folds of cross-validation related to each of the five machine learning models used in this study. The closeness of  $R^2$  for both train and test datasets indicated that the models did not over- or under-fit the datasets (fig. 3.8). The Duncan mean separation analysis, however, was performed only on the test dataset (figs. 3.8 and 3.9) as it resembles the performance of the model on the unseen datasets.

The difference between  $R^2$  for RGB\* and DGCI for the SLR and PR was large and significantly different (fig. 3.8). This result shows that RGB\* applicable to SLR and PR cannot serve as a good color input to the models. Even though the RGB has a lower  $R^2$  compared to RPC, a combination of these ('Both') was not in the middle but better than these. This observation shows that the contribution of RGB (individual channel inputs) along with RPC is required to produce the best modeling performance. The variation in chlorophyll can be accounted through RPC and RGB. However, as these inputs are not significantly different, and for the sake of simplicity, the DGCI for SLR and PR and RPC input for the rest of the models can be used directly.

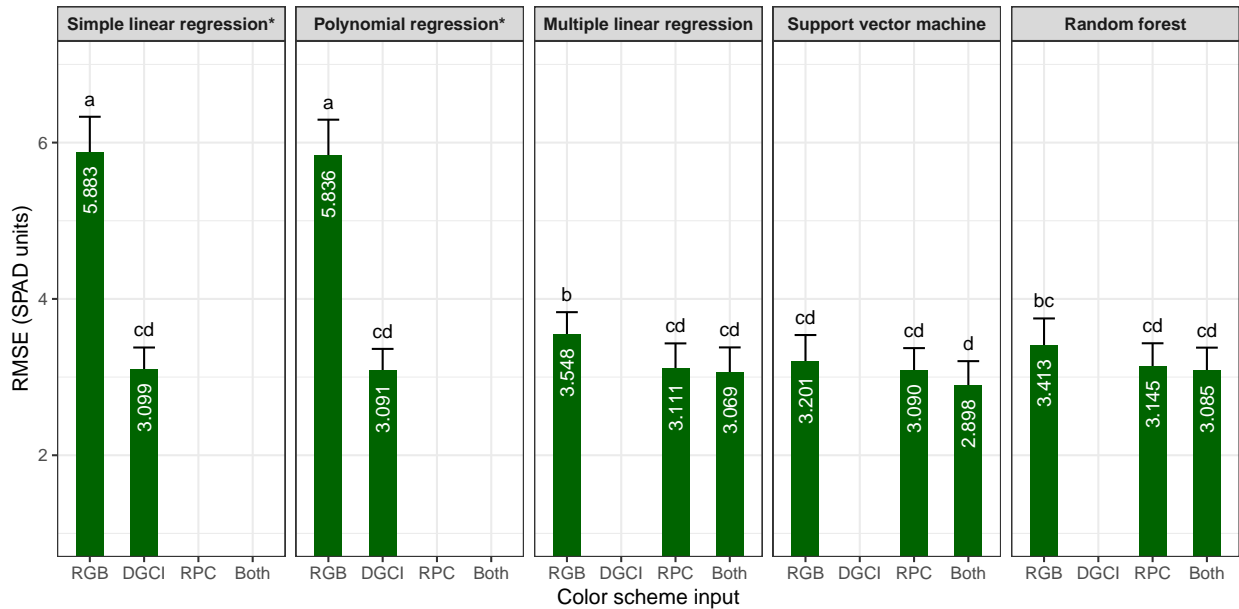




**Figure 3.8. Developed machine learning models comparison based on the coefficient of determination ( $R^2$ ). Note: Models with \* are simpler statistical models and they accept only one independent variable such as RGB (also equals RGB\*) or DGCI, while the other advanced models accept more variables such as individual color channels (R, G, B), RPC of four DGCI, or ‘Both’. Error bars showing only the positive component correspond to one standard deviation.**

Among all models with test datasets, SVM had the best  $R^2$  values (0.86–0.89) across different sets of input variables (fig. 3.8). When ‘Both’ was used, the model performance for SVM was the greatest ( $R^2 = 0.89$ ) among all models and input combinations. However, with the direct input of RPC producing no significant difference among models, it is worth noting that the most simple SLR model with DGCI as input is producing comparable performance ( $R^2 = 0.87$ ).

The other model performance parameter, namely RMSE had the obvious inverse trend (lower the better; fig. 3.9) with  $R^2$  (fig. 3.8) and had similar significance results. For the best SVM model, the minimum RMSE that was obtained among 10 folds of CV was 2.15 SPAD units (and corresponding  $R^2 = 0.93$ , data not shown) using the ‘Both’ input variables. The



**Figure 3.9. Developed machine learning models comparison based on RMSE (SPAD units).** *Note:* Models with \* are simpler statistical models and they accept only one independent variable such as RGB (also equals RGB\*) or DGCI, while the other advanced models accept more variables such as individual color channels (R, G, B), RPC of four DGCI, or ‘Both’. Error bars showing only the positive component correspond to one standard deviation.

RMSE for ‘Both’ was smaller by 6.2 % compared to RPC for SVM, but adding the RGB values to RPC for other models (MLR and RF) did not decrease the RMSE (1.35 %–1.91 %) as much as it did for SVM (fig. 3.9).

### 3.4.3.7. Statistical comparison of the developed models and recommendations

The Duncan’s mean separation statistical analysis, after appropriate grouping (consolidation) of the data, on both  $R^2$  and RMSE results (table 3.4), showed that models, such as SLR with RGB\* and PR with RGB\* were statistically similar but they were significantly different ( $\alpha = 0.05$ ) from the rest of the five models with their corresponding inputs. Similarly, inputs such as ‘Both’, DGCI, and RPC were similar, and RGB\* and RGB were statistically different from each other and the other groups.

The SLR and PR models with DGCI input were not significantly different from each other but performed ( $R^2 \geq 0.87$ ; RMSE < 3.10 SPAD units) almost as good as the advanced models ( $0.84 \leq R^2 \leq 0.89$ ; RMSE < 3.21 SPAD units; table 3.4). However, the advanced models (SVM and RF) have the potential of using the uncorrected RGB values and produce good results ( $R^2 > 0.86$ ; RMSE < 3.21 SPAD units). With regard to input variables, the performance of RGB and RGB\* ( $R^2 < 0.85$ ; RMSE > 3.38 SPAD units) as exclusive inputs was significantly lower than the others (DGCI, RPC, and 'Both'), hence may not be suitable for prediction model exclusive inputs.

Observing the performance of the models and based on statistical analysis, five selected models with their inputs, from simple to advanced, are presented in table 3.5. These models are recommended to predict the SPAD meter readings with their listed inputs. These models were of high performance ( $R^2 \geq 0.87$ ; RMSE  $\leq$  3.12 SPAD units). It is interesting to observe the simple family of models such as SLR ( $R^2 = 0.875$ ) and PR ( $R^2 = 0.886$ ) with DGCI can be highly efficient in predicting the SPAD meter readings, while the advanced SVM produced the best performance ( $R^2 = 0.888$ ). Based on the presented statistical analysis (fig. 3.8 and table 3.4) all these recommended models (table 3.5) will be statistically similar and any of these can be used equally.

Using the predicted SPAD meter values, from the developed models along with suitable conversion relationships, such as equation 3.1, the chlorophyll contents of the leaf sample from the digital image can be determined. It is expected that similar crop-specific conversion function or factors will be available or for new crops can be developed in combination with laboratory procedure of chlorophyll estimation. Application of image processing and associated models also makes it easier to record and review the collected

**Table 3.4. Comparison of models and inputs with using Duncan's mean separation analysis.**

Performance based on $R^2$						Performance based on RMSE (SPAD units)					
Variation due to models			Variation due to inputs			Variation due to models			Variation due to inputs		
Model	Value	Groups <sup>†</sup>	Inputs	Value	Groups <sup>†</sup>	Model	Value	Group <sup>†</sup>	Inputs	Value	Group <sup>†</sup>
Support vector machine	0.875	a	'Both'	0.879	a	Simple linear regression* & RGB*	5.883	a	RGB*	5.859	a
Polynomial regression* & DGCI	0.873	a	DGCI	0.873	a	Polynomial regression* & RGB*	5.836	a	RGB	3.387	b
Simple linear regression* & DGCI	0.872	a	RPC	0.871	a	Multiple linear regression	3.243	b	RPC	3.116	c
Random forest	0.862	a	RGB	0.847	b	Random forest	3.214	b	DGCI	3.095	c
Multiple linear regression	0.859	a	RGB*	0.546	c	Simple linear regression* & DGCI	3.099	b	'Both'	3.017	c
Polynomial regression & RGB*	0.550	b				Polynomial regression* & DGCI	3.091	b			
Simple linear regression & RGB*	0.543	b				Support vector machine	3.063	b			

<sup>†</sup>Dissimilar mean separation letter groups represent statistical significant ( $\alpha = 0.05$ ) difference; RGB is the individual color channel intensity values; DGCI is the dark green color index derived from HSB color space (eq. 3.2) that is applicable to SLR and PR; models with \* are simpler statistical models and they accept only one independent variable while others accept more variables (figs. 3.8 and 3.9); RGB\* is the combined single RGB input (grayscale) and RPC is pixel count of different ranges of DGCI.

**Table 3.5. Recommended machine learning models to predict the SPAD meter reading from the field image-based inputs for chlorophyll measurement.**

Model name	Input	Recommended final model	$R^2$	RMSE (SPAD units)
Simple linear regression	DGCI	$Y = -6.17 + 85.93X$ ; eq. 3.8	0.875	3.12
Polynomial regression	DGCI	$Y = -8.09 + 100.99X - 26.15X^2$ ; eq. 3.9	0.886	3.11
Multiple linear regression	R, B, RPC (1–3)	$Y = 17.29 - 1.88X_R + 0.52X_B - 5.66X_{RPC1} - 5.56X_{RPC2} - 2.32X_{RPC3}$ ; eq. 3.10	0.881	3.07
Support vector machine	‘Both’	$C = 0.3$ , $\text{coef0} = 1.8$ , $\text{degree} = 4$ , $\text{gamma} = 0.1$ , $\text{kernel} = \text{“poly”}$	0.888	2.90
Random forest	‘Both’	Maximum depth of trees = 5, Number of trees = 48, Minimum sample leaf = 6	0.873	3.09

DGCI is the average dark green color index;  $Y$  is the predicted SPAD meter reading;  $X$  is the input variable used; R and B are the raw red and blue channel intensity values; RPC (1–3) is the normalized pixel count of first 3 out of four ranges of DGCI; ‘Both’ corresponds to raw R, G, B and RPC. The  $R^2$  and the RMSE belong to the best model from the 10-fold cross-validation. It should be noted that these models recommendations were derived from data and will be influenced by the different makes of SPAD meters and their calibration.

data, as the image and other information can be stored digitally. The application of image processing and machine learning methodology described herein, to be used infield that too without the standard calibration board, is expected to have a positive impact among farmers/producers and can also be readily extended to other field crops.

### 3.5. Conclusions

This paper presented the use of a digital image processing technique to estimate the chlorophyll content of the soybean leaves infield along with the SPAD meter conversion relationship. The dark green color index (DGCI) had the highest correlation with SPAD meter readings among several color vegetation indices, and the correlation was further improved by color calibration using the standard calibration board. The effect of cultivar, sampling date, and plot on chlorophyll estimation using RCM with DGCI indicated that these factors did not significantly affect the relationship between DGCI and the SPAD meter readings, hence allowed to combine all data and model them with simple linear regression for chlorophyll estimation. Five machine learning models (two simpler statistical: simple linear regression and polynomial regression; multiple linear regression; and two advanced: support vector machine (SVM) and random forest (RF)) tested in combination with image-based inputs, such as combined RGB\*, individual RGB, average DGCI of all pixels in the leaves, pixel count of four different ranges of DGCI (RPC), and RGB+RPC ('Both') showed only SLR and PR models with RGB\* were statistically different. Among the inputs, only RGB\* and RGB were statistically different from each other and other inputs.

The simple statistical models such as SLR and PR with DGCI input ( $R^2 \geq 0.87$ ;  $RMSE \leq 3.1$  SPAD units) performed almost as good as the advanced models such as SVM and RF ( $0.84 \leq R^2 \leq 0.89$ ;  $2.90 \leq RMSE \leq 3.41$  SPAD units). The SVM was found to be the best

model with the input of 'Both' ( $R^2 = 0.89$ ; RMSE = 2.90 SPAD units), and SVM model had the potential of predicting the chlorophyll directly with the raw RGB input ( $R^2 = 0.86$ ; RMSE = 3.20 SPAD units) without the use of the standard calibration board. Therefore, image processing methodology can be used to measure the chlorophyll content of soybean leaves infield using the standard calibration board with simple statistical models using DGCI input, or as an alternate inexpensive method without the standard calibration board, through advanced models such as SVM with raw RGB input. The methodology, even though developed for individual leaves, can be suitably modified to plot- and field-scale applications by processing images from unmanned aerial vehicles/systems. This developed image processing methodology can also be readily extended to other field crops and is expected to be a simpler and inexpensive chlorophyll measurement tool for the agricultural community.

## 4. RATING IRON DEFICIENCY IN SOYBEAN USING IMAGE PROCESSING AND DECISION-TREE BASED MODELS \*

### 4.1. Abstract

The most efficient way of soybean (*Glycine max* (L.) Merrill) iron deficiency chlorosis (IDC) management is to select a tolerant cultivar suitable for the specific growing condition. These cultivars are selected by field experts based on IDC visual ratings. However, this visual rating method is laborious, expensive, time-consuming, subjective, and impractical in larger scales. Therefore, a modern digital image-based method using tree-based machine learning classifier models for rating soybean IDC in plot-scale was developed. Data were collected from a soybean IDC cultivar trial plots. Images were processed with MATLAB and corrected for light intensity by using a standard color board in the image. The three models used in this study were decision tree (DT), random forest (RF), and adaptive boosting (AdaBoost). Calculated indices from images, such as dark green color index (DGCI), canopy size, and pixel counts into DGCI ranges and IDC visual scoring were used as input and target variables to training these models. Metrics such as precision, recall, and f1-score were used to assess the performance of the classifier models. Among all three models, AdaBoost had the greatest performance (average f1-score = 0.75) followed by RF and DT the least. The developed method can be easily scaled-up using images from aerial platforms.

---

\* This paper is submitted to the *Remote Sensing* journal and the second revision is in progress following the review comments. Authors: Oveis Hassanijalilian, C. Igathinathane, Sreekala Bajwa, John Nowatzki. Title: Rating Iron Deficiency in Soybean using Image Processing and Decision-Tree based Models. Oveis performed the literature survey and wrote the manuscript. Dr. Igathinathane Cannayen is the major advisor and the corresponding author who worked with Oveis throughout the manuscript development. All the co-authors have assisted in the research direction and review of the manuscript.



## 4.2. Introduction

The US is the second exporter of soybean (*Glycine max* (L.) Merrill) and its product in the world with a crop value of over \$39 billion in 2018 (ASA, 2019), and the Midwest is one of the biggest production regions. However, the soybean production in general and in the Midwest specifically can be declined by iron deficiency chlorosis (IDC). The IDC is characterized by a reduction in the chlorophyll of the leaves, which makes the leaves turn from green to yellowish that in consequence interferes with photosynthesis, and causing reduced plant height and leaf area (Vasconcelos and Grusak, 2014) which negatively affect the production. For efficient management of soybean IDC, measurement and assessment of the extent of the damage is the key step. The most common and current method employed IDC assessment is the manual visual scoring system by the field experts, where a higher score means increased incidence. This method, however, is laborious, expensive, time-consuming, as well as subjective. Furthermore, it is impractical to use this method in larger scales. Therefore, a modern method of image processing from the actual field images was proposed, tested, and compared with manual rating in this study.

Among different methods available to combat IDC, such as planting a companion crop (Bloom et al., 2011; Naeve, 2006), applying iron chelate (Lucena, 2003; Nadal et al., 2012), and increasing seeding rate and planting in wider rows (Goos and Johnson, 2000), plating tolerant cultivars was proposed as the most effective way to avoid IDC (Hansen et al., 2003; Kaiser et al., 2014; Naeve and Rehm, 2006). Although planting IDC-tolerant cultivar reduces IDC incidence, it will not increase the yield throughout the entire field all the time (Helms et al., 2010). For IDC assessment, the commonly followed method is IDC visual score, but is prone to human inconsistency and is a subjective measurement, which means that different

experts may rate plots with different scores. A relatively poor correlation among experts ( $r < 0.68$ ) was observed when experts visually rated turf plots for leaf spot, density and color (Horst et al., 1984). Light conditions in the field also affect the appearance of color and the human perception of color, which in turn will influence the score (Van Den Broek et al., 2002). The field rating will also be inconsistent because an expert may sometimes rate the same plot with different scores. Raters' variance of visual rating scores was found significantly higher than the variance of dark green color index (DGCI) from digital images after evaluating the same turf plots for several times (Karcher and Richardson, 2003).

As an alternative to manual visual rating, digital image processing can be used for efficient IDC assessment in field conditions. Digital image processing captures the reflection of light similar to human eyes and is much more objective and expected to be consistent compared to visual rating. These images were used for phenotyping of different crops by extracting different color vegetation indices (CVI) from them (Atefi et al., 2019; Vollmann et al., 2011). Yellow and green discs, components of the standard color board, of known DGCI values, were used to compensate for different lighting conditions and the effect of different sensitivity to colors among various cameras on corn (*Zea mays* L.) (Rorie et al., 2011). A close relationship between the amount of nitrogen in corn leaf and DGCI value was reported from that study. Image processing was demonstrated to estimate chlorophyll from soybean leaf images using the DGCI (Hassanijalilian et al., 2020b).

The CVI values that are extracted from digital images for IDC are continuous, but that from visual ratings are discrete. Advanced machine learning classification algorithms are required to predict discrete classes, similar to visual scores, based on continuous CVI values. These machine learning methods have been used in many aspects of agriculture, such as yield

prediction (Smidt et al., 2016), weed detection (Bakhshipour and Jafari, 2018), iron deficiency rating (Naik et al., 2017), disease detection in wheat (*Triticum aestivum* L.) (Azadbakht et al., 2019), classifying broad agricultural land cover types (Yang et al., 2019), to mention a few.

Of the several machine learning methods, identification of the most important field attributes in soybean yield prediction was performed using random forest regression (Smidt et al., 2016). The random forest model was also used to improve soybean yield estimation and predict soybean maturity through aerial imagery (Yu et al., 2016). Among several machine learning classification models, such as decision tree, random forest, and hierarchical model that were used to rate IDC in soybean five seeds hill plots, under controlled lighting conditions with images captured from above, the hierarchical classifier resulted in the best accuracy (Naik et al., 2017). In a similar study, Bai et al. (2018) used a platform to capture images of soybean plots from above, without ambient light compensation, and used linear discriminant analysis (LDA) and support vector machine (SVM) to classify the severity of IDC, and found the SVM produced the increased accuracy. They also rated soybean for IDC “in-office” by observing the pictures called office score and found that office score was more consistent in the IDC scoring across different site-years.

Research on soybean IDC rating using image processing in an uncontrolled condition at actual field plots with lighting compensation is not found in the reported literature. Therefore, a research investigation on the application of image processing in determining the IDC and develop prediction models from the image data at the field plot scale was proposed. The specific objectives of this research are to (i) evaluate the relationship between IDC visual scoring and DGCI image information obtained infield along with a standard color board, (ii)

assess several machine learning models on IDC rating, and (iii) to develop ready-to-use prediction methodology from DGCI input. Outputs of the research work will have broader applications and expected to impact the soybean farmers/producers, breeders, crop consultants, agricultural extension agents, and other users.

### **4.3. Materials and Methods**

#### **4.3.1. Field Experimental Plots**

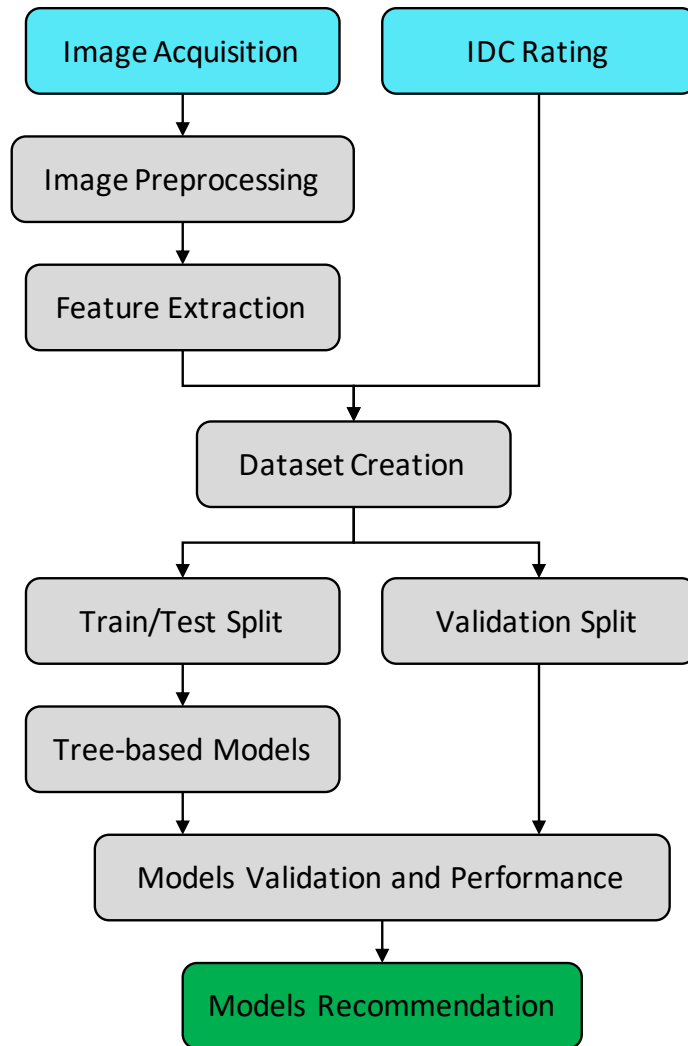
Forty different cultivars of soybean were planted in Leonard, ND, USA (46.671 783° N, 97.245 939° W) for the IDC tolerant cultivar experiment. The experimental plots with four replications each (40 cultivars × 4 replications = 160 plots; fig. 4.1) were seeded at a density of 30 seeds/m in each row on mid-May, 2014. Each plot was 3.4 m long and 3.0 m wide, and had four rows with a row spacing of 0.76 m. The 1st and 4th rows were the border rows and to avoid the edge effect were not considered, while 2nd and 3rd rows were only considered for the analysis.

#### **4.3.2. Overall Methodology**

The research aims to apply image processing methodology using field plot images and develop prediction models to determine the IDC classes of soybean. The overall methodology showing various processes involved in image processing and machine learning model development and validation for IDC classification is illustrated in figure 4.2, and described subsequently. All the plot images were processed to extract features, such as the two-colored disks and two middle soybean rows. From the middle rows, new features were calculated to serve as input features to develop and train the tree-based models. These models used visual expert's ratings as the target variable and were tested using classification metrics on a portion of the dataset that was not used in the training of the models.



**Figure 4.1. Portion of the experimental plots showing the plot design and the existing variation in soybean crop IDC.**



**Figure 4.2. Process flowchart of field image processing and machine learning models development and validation with manual rating for IDC classification.**

### 4.3.3. IDC Visual Rating

In the experiments, soybeans were visually rated for IDC by experts in 2014 first at the V2–V3 (Fehr et al., 1971) growing stage (July 3) and the second time at the V5–V6 growing stage (July 17). Table 4.1 shows the developed IDC rating scores used in the experiments (Helms et al., 2010). For each rating, visual IDC scores were recorded to the nearest one-half (0.5) unit.

**Table 4.1. IDC symptoms and the developed visual rating score guidelines<sup>1</sup>.**

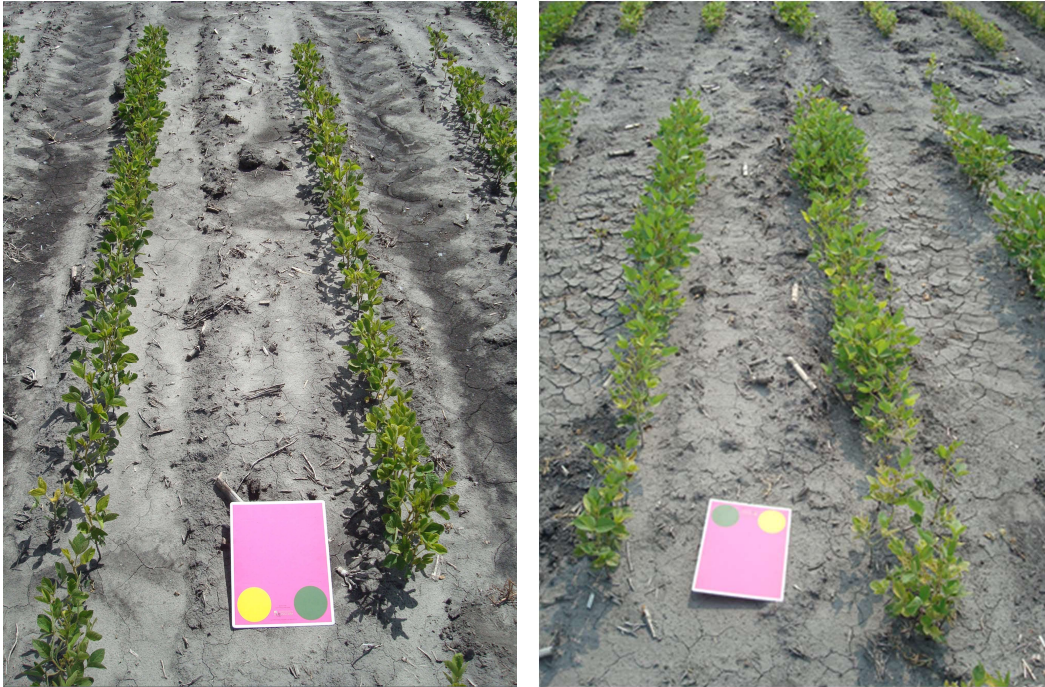
IDC symptoms description	IDC Score
No chlorosis	1
Slight yellowing of the upper leaves	2
Upper leaves without necrosis or stunting and with interveinal chlorosis	3
Upper leaves with reduced growth or beginning of necrosis with interveinal chlorosis	4
Severe stunting, damaged to growing point and chlorosis	5

Note: IDC - Iron deficiency chlorosis; and <sup>1</sup> IDC scores scale was developed by Helms et al. (2010).

#### **4.3.4. Image Acquisition**

Images of soybean plots were captured on the same days that the expert rated soybean for IDC (July 3 and 17, 2014). To account for differences in ambient lighting and the effect of different cameras on final images, a standard color board was placed between 2nd and 3rd rows in each image that will be later processed (fig. 4.3). The middle rows (2nd and 3rd) with the standard color board were captured and analyzed for the image-based IDC scoring. A SONY™ DSC-W80 digital camera with an image size of 2304 × 3072 pixels was set to an ISO of 125 in the “Automatic” mode to ensure of equal brightness for all images captured. Images were saved as JPEG (Joint Photographic Experts Group) files and numbered based on cameras’ internal numbering sequence. Images were processed in MATLAB (MATLAB, 2015b) to extract the desirable features (e.g., vegetation, standard color board, color discs) and perform various analyses.

Standard colors consisted of yellow and green 90 mm diameter circular discs set on a pink (#eA328A hex; 234, 50, 138 RGB)board. The standard HSB values of yellow disk are 50, 87, and 91 (#e7c71f hex; 231, 199, 31 RGB) and for green disc are 91, 38, and 42 (#576c43 hex; 87, 108, 67 RGB), which produced the DGCI (eq. 4.1) values of 0.0178 for yellow and



**Figure 4.3. A sample image of soybean plot showing the two middle rows with standard color board at two different growing stages (Left: V2-V3 and right: V5-V6).**

0.5722 for green disc (Rorie et al., 2011). These known DGCI values of disks were used along with the observed values of these disks from the image to calibrate DGCI values of soybeans.

The expression of DGCI values for soybean from the HSB color space channels (Karcher and Richardson, 2003) used in the analysis is:

$$DGCI = \frac{(\text{Hue} - 60)/60 + (1 - \text{Saturation}) + (1 - \text{Brightness})}{3} \quad (4.1)$$

The raw DGCI values (eq. 4.1) of the extracted vegetation component of the image needs correction based on the lighting conditions, and this correction was performed using the extracted colors of the discs of the standard color board. For this correction, it is assumed that the known and observed values of DGCI from the camera follow a simple linear relation.



The DGCI correction (eqs. 4.2–4) was performed based on the slope and intercept of the assumed linear response as:

$$\text{Slope} = \frac{0.5722 - 0.0178}{\text{Observed green DGCI} - \text{Observed yellow DGCI}} \quad (4.2)$$

$$y\text{-intercept} = 0.5722 - (\text{Slope} \times \text{Observed green DGCI}) \quad (4.3)$$

$$\text{Corrected leaf DGCI} = (\text{Slope} \times \text{Observed leaf DGCI}) + y\text{-intercept} \quad (4.4)$$

#### 4.3.5. Object Recognition in Field Image

The purpose of object recognition is to detect multiple features from the images (fig. 4.3) such as middle rows, standard color board, and the colored disks on the board for features extraction required in the analysis. The feature detection and extraction were performed with MATLAB (2015b) image processing toolbox, which was used in the calculation of DGCI and other input features. Brief description of object recognition is presented subsequently (figs. 4.5 and 4.7).

Images were imported into MATLAB as a 3D matrix with 3 layers of  $2304 \times 3072$  2D unsigned 8-bit integer matrix, which has layers of red, green, and blue (RGB). In the image preprocessing, grayscale images were created by slicing the original color images to its component layers of RGB (e.g., Red = `image(:, :, 1)`). The pixel values of the grayscale images vary between 0 to 255, where 0 represents absences of the color and is showed by black. Similarly, 255 represents the pure color and is showed by white; everything in between represents different shades of that color (R,G,B), but are showed with different shades of gray. The brightness of the loaded images was calculated through hue, saturation, brightness (HSB) color space, which is closer to human perception of color and therefore helpful to

recognize objects in the images. This image preprocessing is essential for features identification and extraction.

#### 4.3.5.1. Extraction of standard color board

To recognize and extract the two-color disks of the standard color board that had the reference color information, the color board had to be detected first. The color board can be easily recognized through  $a^*$  component (red/green chromatics) of  $L^*a^*b^*$  colorspace as the histogram clearly showing three distinct peaks (fig. 4.4). The  $a^*$  component of the original image is a grayscale image (fig. 4.5a) that can be used to process several specific binary images.

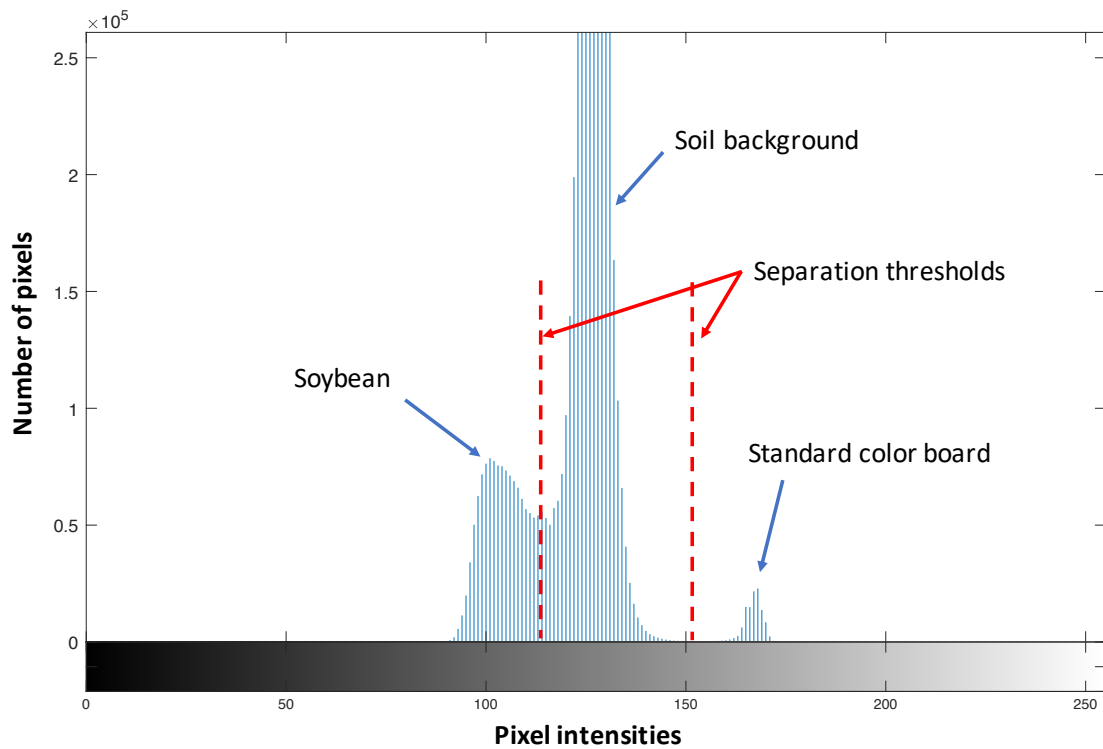


Figure 4.4. Histogram of  $a^*$  component of  $L^*a^*b^*$  colorspace showing different objects in the image. The small right peak represents the color board, the larger middle peak represents the soil, and the left mean peak represents the soybean rows based on the number of pixels in each object. The values were converted to 8-bit unsigned integer (0-255).

Using `imbinarize()` command and proper thresholds (based on the peaks) binary images were created recognizing different features. This binary image was used to create a mask for extracting the standard color board using the object's area property as the board's area dominates over other artifacts created by leaves. The artifacts were removed using the `imopen()` command with a "square" structure to both remove the small objects and preserve the shape of the board. In the binary image, it can be seen that the colored disks left holes (fig. 4.5b) and using `bwconvhull()` command the holes got filled.

At this stage, the color board was the only object in the binary image (fig. 4.5c), and its corresponding major and minor axes lengths were extracted using `regionprops()` command. These lengths were used to estimate the board area to be used for thresholds in later morphological operations. Furthermore, the bounding box of the board was extracted and used to "crop" the board from the image for easier extraction of green and yellow disks.

#### **4.3.5.2. Extraction of yellow and green disks of color board**

As  $L^*a^*b^*$  colorspace separated the standard color board from the background and plant rows, HSB colorspace was suitable to recognize color disks of the color board. The DGCI values of yellow and green disks were required to calculate the slope and  $y$ -intercept (eqs. 4.2 and 4.3). Since the hue (H of HSB; a grayscale image, fig. 4.5d) for yellow and pink are significantly different, a value of  $H > 0.3$  was used as the threshold to remove non-yellow pixels from the cropped image (fig. 4.5d). However, hue for green and yellow are close to each other, but the saturation (S of HSB) of the green disk was the lowest of the board. Therefore, using a range of  $S < 0.1$  ensured that the biggest object in the cropped image is the yellow disk. The rest of the cropped board was converted to a binary image using `imbinarize()` command and Otsu threshold method (fig. 4.5e). Otsu method separates the

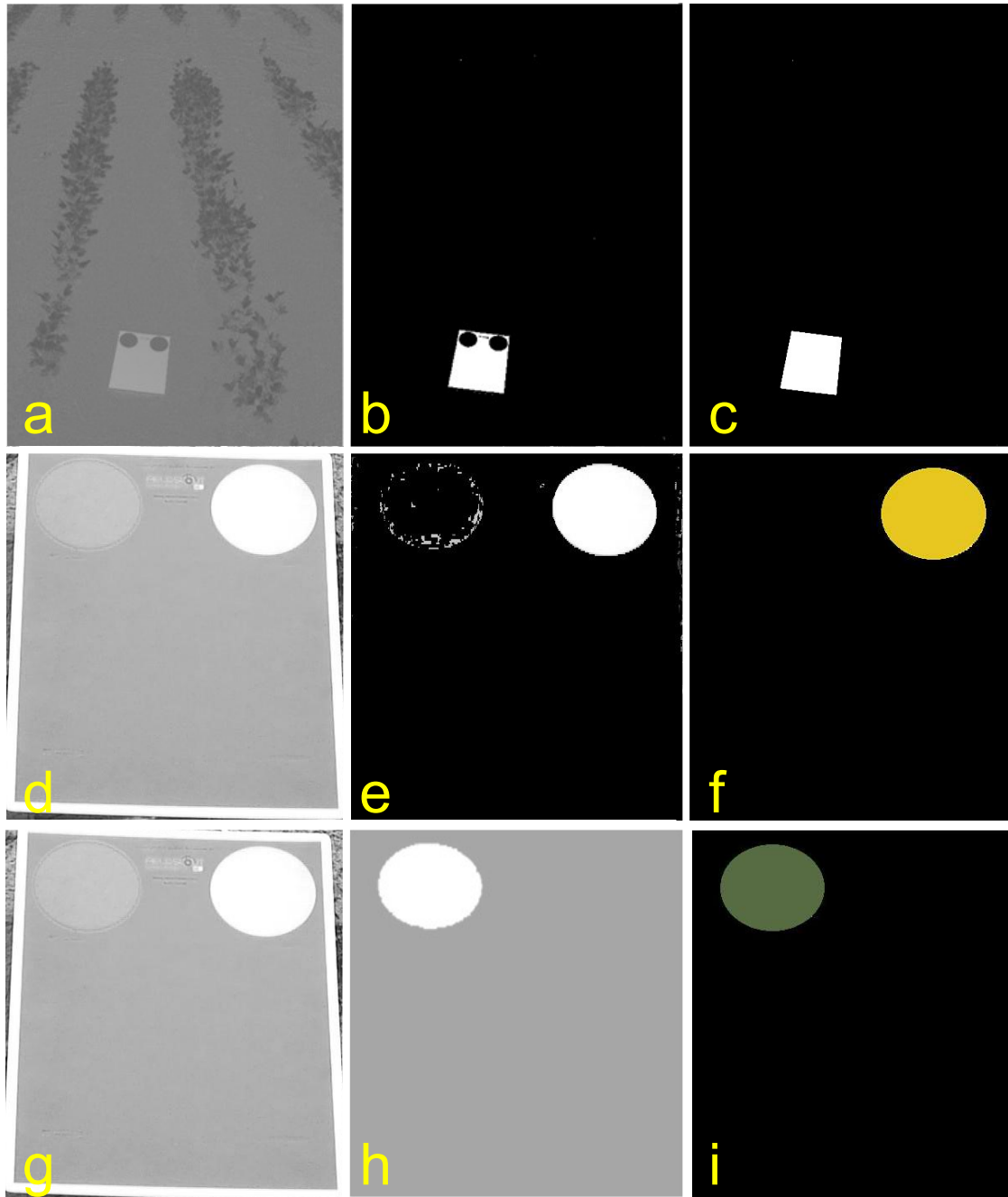


Figure 4.5. Object recognition processes from field plot images of detecting standard color board and its green and yellow disks. (a) Grayscale image of  $a^*$  channel of  $L^*a^*b^*$ ; (b) binary image after using a proper threshold on previous image; (c) filling the holes in the board using `bwconvhull()` and using `imopen()` to remove the small artifacts and create a mask to crop the board; (d) grayscale image of green channel of RGB related to cropped board; (e) binary image after removing the pixels with  $H > 0.3$  or  $S < 0.1$ ; (f) the recognized yellow disk; (g) grayscale image of green channel of RGB related to cropped board; (h) only keeping pixels with  $0.2 < \text{hue} < 0.5$ ; and (i) the recognized green disk.

foreground and background in a grayscale image by maximizing the inter-class variance. Finally, to remove the small particles from the cropped binary image and to preserve the shape of the disk, a circular structure was used within `imerode()` function using the area threshold from the previous step automatically extracted the yellow disk (fig. 4.5f).

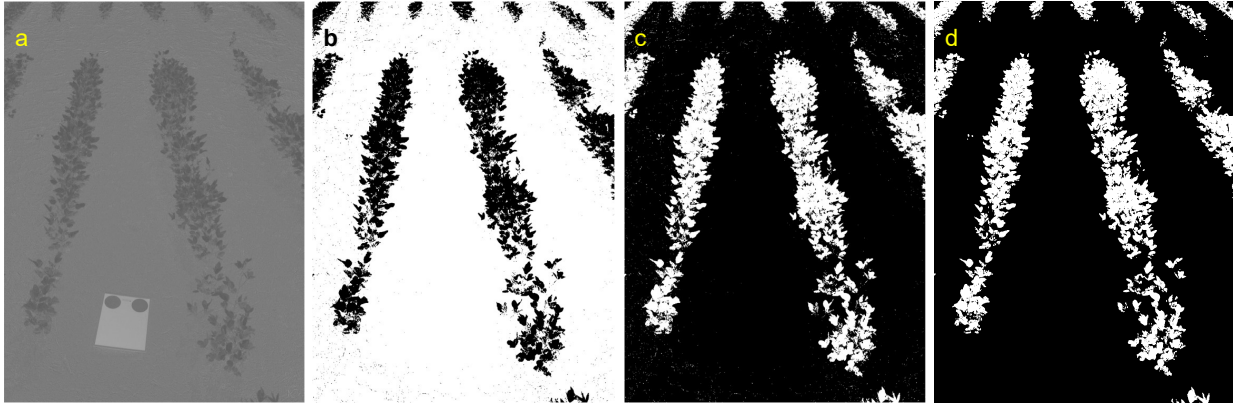
Similarly, the green component of RGB (fig. 4.5g) was used for green disk extraction. Pixels with hue values outside the range of 0.2–0.5, which do not belong to green, were removed (fig. 4.5h). Since the shape and size of the disks were equal, the same morphological operations (structure and area threshold) were performed to extract the green disk (fig. 4.5i).

#### **4.3.5.3. Removing soil background for plant rows extraction**

To identify the soil background from the images and extract the soybean rows, the values from  $a^*$  were increased by 127 to have all values in a positive range. Next, the whole channel was converted to 8-bit unsigned integer values for easier calculation using `uint8()` resulting in a grayscale image (fig. 4.6a). Finally, the Otsu method was used to separate the soybean rows in the foreground from the soil in the background. Since the soil pixels have higher values in  $a^*$  channel (fig. 4.6a), the binary image removed the soybean rows as the background (fig. 4.6b). Therefore, the values of the binary image were inverted so soybean rows can be used as the mask (fig. 4.6c). The small artifacts were removed with `imopen()` (fig. 4.6d).

#### **4.3.5.4. Detecting plot middle rows**

After extracting the plant rows, it is necessary to identify the two middle rows of the plot for further processing. The first step to extract the middle rows was to detect their centerlines, explained subsequently, which is necessary to create buffers in order to contain the plant pixels between them. Slope and intercept were required to derive the equation for



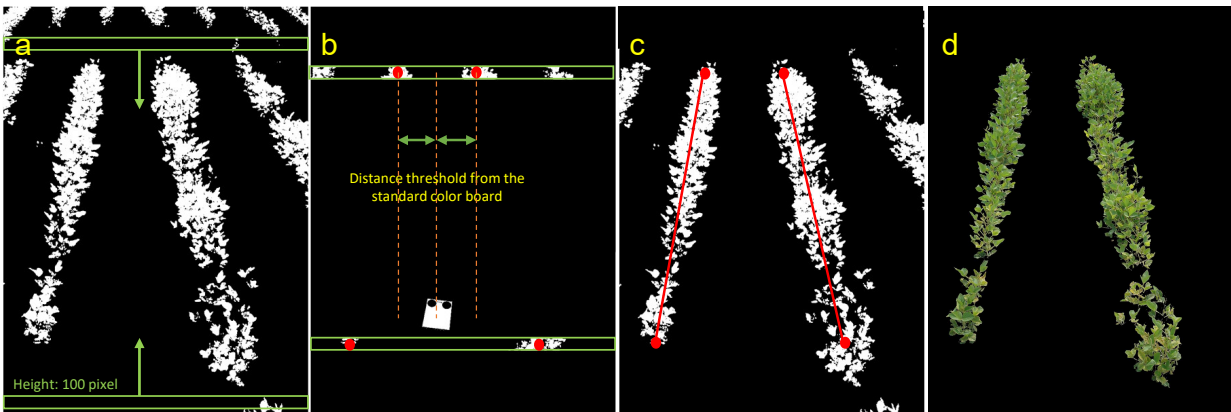
**Figure 4.6. Binary image creation and artifacts removal using  $a^*$  component of  $L^*a^*b^*$  colorspace. (a) The  $a^*$  channel grayscale image of  $L^*a^*b^*$ ; (b) binary image after applying Otsu method; (c) inverted binary image of the previous step showing artifacts (very small objects); and (d) cleaned image after artifacts removal using `imopen()`.**

both center lines. To extract these values, the cleaned images obtained earlier (fig. 4.6d) were used and the top 10% of the images were ignored because other plots were visible in the images.

These preprocessed images were searched for two or more objects with using “search windows” of 100-pixel high and image width (e.g., 2304 pixels) wide from top and bottom (fig. 4.7a). The bottom search window scanned the image first from the bottom proceeding upwards and stopped when it found the objects, and the top window performed the search from the opposite side for objects (fig. 4.7a).

Because of the perspective nature of the images and consistent framing of the image of each plot, the two objects in the bottom were always related to the middle rows. However, a proper distant threshold from the standard color board was defined to select the two appropriate objects on the top (fig. 4.7b). The center points (centroids) of the objects were detected and the equations of the lines were derived for each row (fig. 4.7c). Pixels that fell in close proximity of the lines (a fixed buffer) were kept as the mask to extract the middle

rows (fig. 4.7d). The pixels corresponding to these two rows were used to calculate input features (e.g., DGCI) of machine learning models that are explained subsequently.



**Figure 4.7. Identification and isolation of the plot middle rows from border rows and other plots. (a) Cleaned image showing only the plants with 100-pixel high and image width wide “search windows” from top and bottom working inward for identifying plant pixels; (b) top edge portion representing the adjacent plots were removed, the two objects that were found from the middle from the bottom and two objects from the top search windows fixed from the standard color board, and centroids were marked (red dots); (c) identified top and bottom centroids were connected to create the rows centerline and a mask was created using a buffer along the centerlines; and (d) extracted two middle rows of the image after applying the mask for image-based analysis.**

#### 4.3.5.5. Color vegetation index

To measure the greenness of the object in the digital image, it is not practical to use only the green channel of RGB because the amount of red and blue also affects the appearance of green in the image. To make the interpretation of digital colors easier and similar to human perception, Karcher and Richardson (2003) suggested converting the RGB values to HSB. They derived DGCI (eq. 4.1) from HSB color space to estimate the amount of nitrogen in corn leaves. The DGCI was found suitable than other CVIs from different studies, such as average red, average green, average blue, and indices that are a combination of RGB

values; and applied based on a fitted model ( $R^2 = 0.89$ ) to estimate the amount of chlorophyll in individual soybean leaves (Hassanijalilian et al., 2020b). Since chlorophyll is one of the major indicators of iron deficiency in soybean leaves, the DGCI was selected as the best CVI to study the relationship between IDC measured through image processing and visual rating based on the reported studies and the other CVIs with low correlations were not considered for analysis.

#### **4.3.6. Machine Learning Models**

A few advanced machine learning models that were successful in related applications (Liakos et al., 2018) were used to rate the soybean IDC based on DGCI and compare with visual ratings. In this study, decision tree (DT), random forest (RF) and adaptive boosting (AdaBoost) were tested because they are suitable for classification purposes and they range from simple (DT) to more advanced and computationally intensive model (AdaBoost). Algorithms such as RF and AdaBoost are considered as “black-box” and hard to interpret because of the models’ complexity, compared to DT and logistic regression. A brief description of these selected models is presented subsequently. Inputs to these models were the different features extracted from the images and the outputs were the prediction for IDC severity merged rating, which is an integer between 1 and 4.

##### **4.3.6.1. Decision tree classifier**

The DT is a machine learning model that can be used for both classification and regression and is referred to as “classification and regression tree” or CART (Géron, 2017). The DT is an inverse tree-like graph with root on top and decision nodes at the bottom. When a node splits further down, it is called a decision node; and when a node cannot be split, it is called a leaf or terminal node. Each decision node assesses a condition on a single feature to



split into a sub-node. To find that condition and feature, DT uses the purity index such as Gini (eq. 4.5) and entropy (eq. 4.6).

$$G = 1 - \sum_{k=1}^n p_k^2 \quad (4.5)$$

$$H = - \sum_{k=1}^n p_k \log_2(p_k) \quad (4.6)$$

where  $G$  is the Gini index (dimensionless),  $n$  is the number of classes,  $p_k$  is the ratio of class  $k$  instances among the training instances in an individual node, and  $H$  is the entropy (dimensionless).

Furthermore, the DT has some parameters that define how and when each node can be split. Minimum sample leaf and maximum depth are two of the parameters that have a significant effect on the DT's behavior. The former defines the minimum number of samples in each node before that node can be split into sub-nodes, and the latter defines the maximum depth of the tree which defines how many levels can be between the root and the leaf nodes.

#### 4.3.6.2. Random forest classifier

RF is one of the popular ensemble models that use DTs as base estimators. Ensemble classifiers in machine learning consist of several weak classifiers, and voting among these individual classifiers determines the output class of the ensemble model (Géron, 2017). In addition to DT's parameters, RF has other parameters that can be tuned, such as the number of estimators and bootstrap. The number of estimators defines how many decision trees are used to build an RF, and bootstrap determines if each tree uses all or a portion of the dataset. Having a bootstrap option active in the model helps the model to generalize better on future datasets.

#### 4.3.6.3. Adaptive boosting (AdaBoost) classifier

In the ensemble methods, the individual classifiers are independent of each other and train in parallel, but when they are dependent on each other and train sequentially and these are known as boosting algorithms. AdaBoost is one of the boosting algorithms that passes a new weight for training instances from a classifier to the next, so the next one can focus more on the instances that were difficult to predict with the current classifier (Géron, 2017).

Different models can act as the base estimator for the AdaBoost model, but the most common as well as a default is decision tree (Pedregosa et al., 2011). AdaBoost has some parameters that are not part of the base estimators and can be set within AdaBoost. “Learning rate” is one of the parameters that define the boosting for misclassified instances. There is a tradeoff between the number of estimators and learning rate; the higher the number of estimators, the smaller the learning rate can be set.

#### 4.3.6.4. Models’ input features

The DGCI values of the middle rows of each plot were used for feature extraction (fig. 4.7e). The range of available DGCI (0–1) was divided into five equal range classes, named  $c_1$  to  $c_5$ , with the lowest range ( $c_1$ ) signifying the least dark green meaning the most severe IDC, and highest range ( $c_5$ ) signifying the darkest green meaning the most healthy crop. This image-based DGCI ranges follow the IDC symptoms visual rating scores (table 4.1; Helms et al. (2010)), but the rating scores in reverse order with least score means healthy crop, and consists of the number of pixels that falls within each range. These ranges were selected to study the distribution of IDC (yellowness) in each plot.

Since the canopy size (CS) of each plot was different, based on the number of pixels, the classes observed in each image were normalized by using the minimum, maximum, and

the total number of pixels in the plots, so all the canopy sizes fall between 0 % and 100 % (eq. 4.7).

$$\text{Normalized CS (\%)} = \frac{\text{CS} - \min(\text{CS})}{\max(\text{CS}) - \min(\text{CS})} \times 100 \quad (4.7)$$

The mean and standard deviation of DGCI of both rows in each image were also used to train the models. The target feature was IDC visual scores, which were between 1 to 5 with 0.5 increments, but the highest in all datasets from the visual rating was 4. As the most severe IDC (score = 5) cases were not observed in the study trials, due to lack of such data in each class, and to increase the instance per class ratio, the observed ratings were merged into 4 classes and named merged rating (MR). Since visual rating 1 is the best score for a soybean plot related to IDC tolerance, it was assigned to class 1. Visual ratings of 1.5 and 2.0 were merged as class 2, and so on. The same procedure was performed on both dates; however, the highest MR for the Date-1 was only a rating of 3.0 as the plants were at the young stage (V2–V3; fig. 4.3), unlike Date-2, before the full expression of IDC. It should be noted that the visual IDC scores (table 4.1) or the MR have an inverse relationship with DGCI range classes (high IDC score = high IDC; high DGCI range class = low IDC) regarding IDC classification.

#### **4.3.7. Imbalance Dataset Oversampling Technique**

Often the number of instances that are labeled across classes are remarkably different, and this can cause issues in the outcome of some machine learning models. The selected models fit better on the classes with the “majority” of instances and might not perform well with the rest (“minority”), which in an overall sense is not a good approach. The IDC studies are mostly produce imbalance datasets and the distribution is not equal among all severities (Bai et al., 2018). There are two ways to deal with an imbalance dataset such as adjusting the weight of training instances (Domingos, 1999) and resample the original dataset by

either oversampling or undersampling. Sometimes, as observed in this study, the application of undersampling is not practical due to the lack of data in both minority and majority class to fit the machine learning models, hence oversampling was followed.

One of the widely used oversampling methods is synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002) which has been used in credit card fraud detection (Fiore et al., 2019), cancer identification (Geetha et al., 2019), agricultural applications (Aiken et al., 2019; Espejo-Garcia et al., 2018; Shahriar et al., 2014). In this method, the minority class is oversampled by creating synthetic instances from the “feature space” rather than “data space.” Synthetic instances are created along the line between each instance in the minority class and k-nearest neighbors (KNN algorithm). The amount of oversampling determines how many of the nearest neighbors are used in the process.

#### **4.3.8. Performance Assessment of Machine Learning Models**

In machine learning model development, it is common practice to split the dataset between training and test datasets. In this study, a 10% of the dataset was randomly extracted and used as the validation set. The other 90% was used in 10-fold cross-validation to train and test the models. The 10-fold cross-validation function splits the data into 10 and uses 9 of them for training and the remaining 1 for testing the model, and this process is repeated for 9 more times until all the splits were used as the test set (Géron, 2017). The 10-fold cross-validation was used each time after changing the parameters to evaluate the performance of the models. After obtaining the highest performance, the validation set was used to assess the performance of the models with unseen test data.

The performance of cross-validation for a classifier model can be assessed through different scores such as accuracy, precision, recall, and f1-score. These scores are based on a

confusion matrix. “A confusion matrix  $C$  is such that  $C_{i,j}$  is equal to the number of observations known to be in the group  $i$  but predicted to be in the group  $j$ ” (Pedregosa et al., 2011). In a binary classification  $C_{0,0}$  is considered as true negative (TN),  $C_{0,1}$  is false positive (FP),  $C_{1,0}$  is false negative (FN), and  $C_{1,1}$  is true positive (TP). In a multi-class confusion matrix, the value in matrix  $C$  is true when  $i = j$ , and the rest  $i \neq j$  are considered as false. It is desirable to have increased values of TP and reduced values of FN, which represents a better performance of the model.

Precision is the ratio of correct predictions for one class to all instances that are predicted belonging to that class (i.e., what fraction of predictions of each class is true) as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4.8)$$

where TP is the number of true positive cases, and FP is the number of false positive cases.

The recall is the ratio of correct prediction of one class to all instances belonging to that class (i.e., what fraction of instances of each class predicted is true) as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4.9)$$

where FN is the number of false positive cases.

The f1-score is a harmonic mean between precision and recall which considers both FP and FN and is defined as:

$$\text{f1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.10)$$

Another common model performance parameter is “accuracy,” and this with imbalanced data as seen in the study dataset, sometimes misleads the performance, hence was not used in this study.

#### 4.4. Results and Discussion

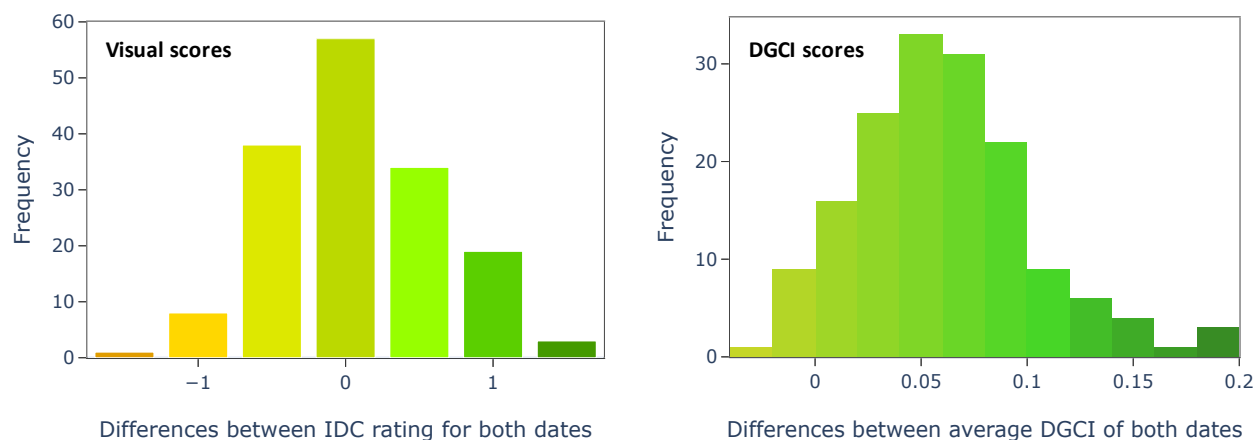
With the progress of time (Date-1 to Date-2) and crop growth, it was observed that the average canopy size for the Date-2 was increased, but the minimum canopy size was smaller compared to Date-1; this suggests that some plots had stunted growth or lost some of their leaves. As expected, with IDC development during crop growth, the average visual IDC score in the Date-2 was higher than the Date-1 (table 4.2). However, the image-based average DGCI score (DGCI\*) increased as well, which suggests that the crop foliage was greener, and in other words, the IDC decreased in the plots. This inconsistency between visual IDC score and image-based DGCI\* could be because of both human inconsistency in visual scoring and ignoring the variation of IDC in the plots by using only the average of DGCI (DGCI\*; table 4.2). The worst chlorosis in each plot, which is shown by DGCI  $c_1$ , has higher average values for Date-1 compared to Date-2, whereas  $c_2$  and  $c_3$  had higher values in Date-2 that meant more green portions of the crop in each plot. All of these values corroborate with field observations and the DGCI\* values of both dates (table 4.2).

The changes in the greenness of the plots between the dates were observed through the differences between visual score as well as average DGCI values ( $DGCI^*_{Date2} - DGCI^*_{Date1}$ ) of individual plots (160 data points each) (fig. 4.8). A positive difference value means the greening of plant canopy and negative the yellowing. Based on visual scores, almost 30 % of the plots become worse compared to the Date-1, and which is not the actual case. Human

**Table 4.2. Distribution of the input features extracted from images for both sampling dates ( $n = 320$  total).**

Data ( $n = 160$ each)	July 03, 2014 (Date-1)									July 17, 2014 (Date-2)								
	Normalized canopy size	IDC score	DGCI*	SD DGCI <sup>†</sup>	DGCI ranges					Normalized canopy size	IDC score	DGCI*	SD DGCI <sup>†</sup>	DGCI ranges				
					$c_1$	$c_2$	$c_3$	$c_4$	$c_5$					$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
Average	31.34	1.98	0.23	0.11	41.52	50.07	8.15	0.26	0.0	52.12	2.06	0.29	0.11	20.93	62.64	15.08	1.28	0.07
SD	8.98	0.54	0.04	0.01	15.28	11.97	3.86	0.23	0.0	14.99	0.85	0.03	0.02	9.00	7.97	7.16	1.91	0.27
Minimum	14.23	1.00	0.14	0.09	8.03	21.24	1.92	0.01	0.0	9.26	1.00	0.22	0.07	4.21	41.31	3.63	0.01	0.00
First Quartile	24.80	1.50	0.20	0.10	30.86	40.17	5.35	0.12	0.0	41.65	1.00	0.27	0.09	14.38	56.91	9.79	0.26	0.00
Median	30.93	2.00	0.23	0.11	41.56	50.98	7.43	0.20	0.0	52.10	2.00	0.29	0.10	19.96	63.36	13.98	0.60	0.00
Third Quartile	35.47	2.50	0.26	0.11	53.58	59.18	9.92	0.33	0.0	61.63	3.00	0.31	0.12	27.07	69.23	19.49	1.22	0.00
Max	56.58	3.00	0.33	0.13	76.81	73.92	23.92	1.29	0.0	100.00	4.00	0.41	0.18	44.56	79.06	39.01	10.06	2.03

Note: \* represents the average DGCI values of all the pixel within each plot, <sup>†</sup> represents the standard deviation of DGCI values of all the pixel within each plot, SD is the standard deviation, and  $c_1 - c_5$  are the five DCGI image-based IDC classes (low value = less green and high IDC; high value = dark green and low IDC) analyzed.



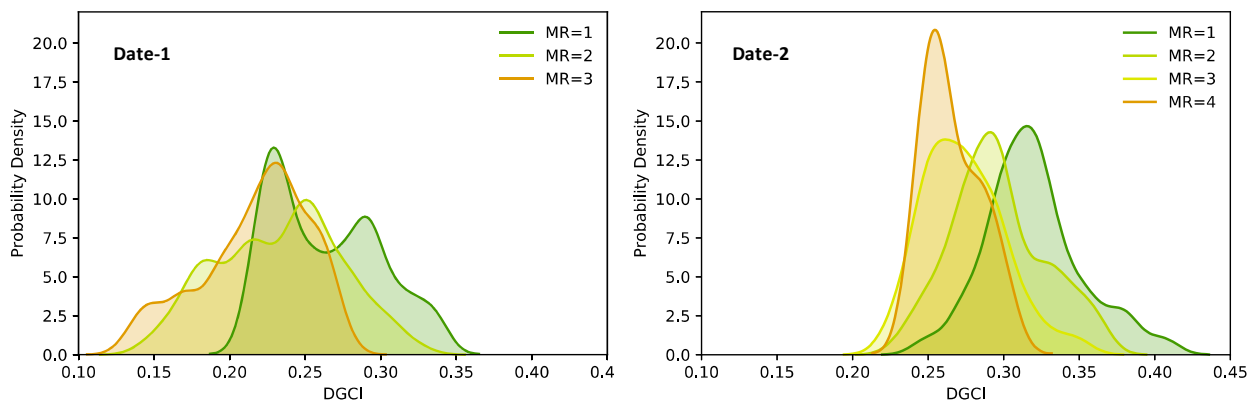
**Figure 4.8. Soybean plots iron deficiency chlorosis (IDC) progression as difference between both dates based on visual scores and image-based on average dark green color index (DGCI\*) scores.**

inconsistency can cause this difference (Karcher and Richardson, 2003; Van Den Broek et al., 2002), and it was shown that the IDC rating that was done in the “office” or “through computer” is more consistent than the field rating (Bai et al., 2018). Based on average DGCI values, most of the plots showed an increase in the average DGCI (table 4.2), which is an indicator of improvement in their health status. This can be seen from the difference frequency plot (fig. 4.8) that DGCI had less negative values compared to visual rating scores.

The MR for the Date-1 consisted of 3 classes, and class 2 had the most instances (61%), and class 1 was the least instances (11%). Similarly, in the Date-2 class 2 had the majority of instances (41%), but class 4 was the minority (7%). The distribution of classes of Date-2 is similar to other studies that classes with severe IDC have a lower number of instances (Bai et al., 2018), and that is the imbalance nature of these kinds of studies (Naik et al., 2017). To create a balanced dataset, the SMOTE method oversampled the minority classes in a way that they have an equal amount of instances in their classes compared to the majority class.



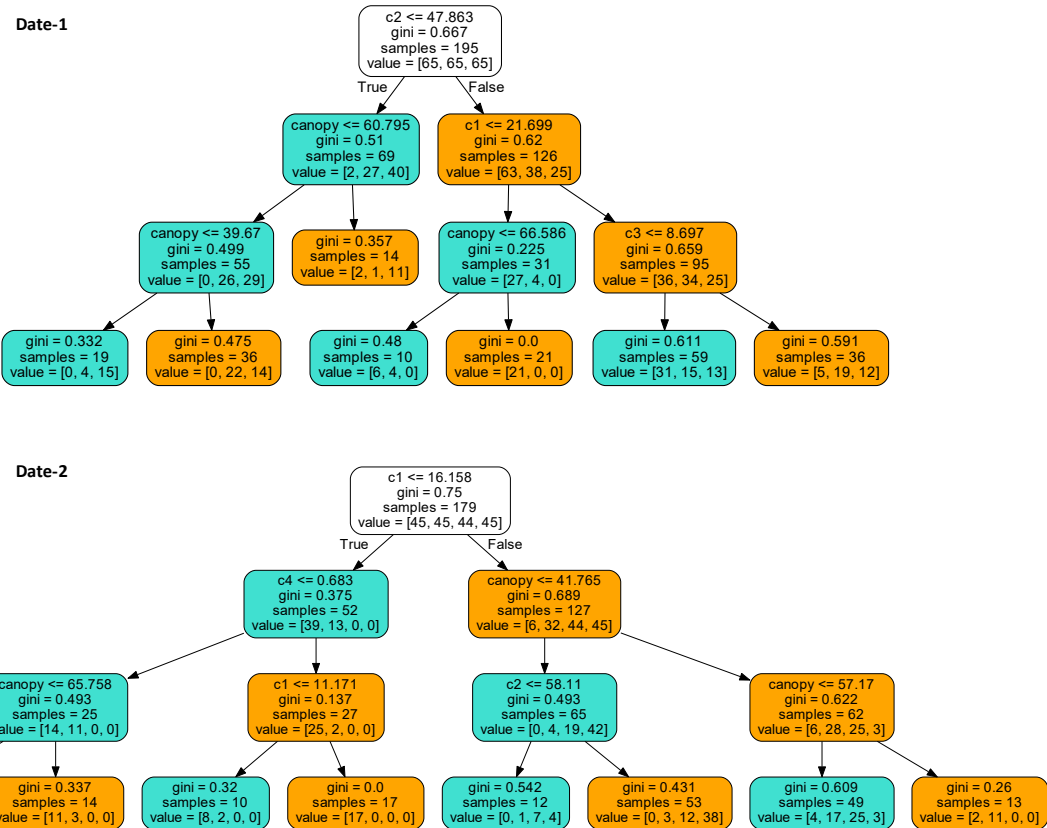
Distributions of the average DGCI values for different MR show that plots with lower MR are scattered around higher average DGCI values, but still, there were a notable amount of overlaps among adjacent classes (fig. 4.9). Field expert rater's error in detecting the right class, and having an average of DGCI for each plot representing IDC severity resulted in having these increased standard deviations in each class. Therefore, using different classes of DGCI to study the distribution of chlorosis in each plot could alleviate some of the problems for machine learning models in the future. It should also be noted that both methods performing average by some means will result in overlaps of various degrees leading to misclassification and cannot be eliminated.



**Figure 4.9. Distribution of image-based average dark green color index (DGCI\*) among different MR for both dates.**

#### 4.4.1. Decision Tree

The DT output as graphs for both dates are shown in figure 4.10 for depth = 3 with splitting conditions and Gini index for each node. The nodes with a lower Gini index have less impurity. The leaf nodes which are at the bottom of the trees show the final guess of each tree depending on which class in the class vector has the highest number. The Gini index for



**Figure 4.10. Decision criteria of the decision tree model. Two sub-nodes are derived from each node, and the one that is derived to the left of the node represents the “true” evaluation of the parent node’s condition. The higher number in the value vector is the predicted class in the final nodes. Only features with least impurities will be utilized, hence all features and classes will not feature in the DT models.**

the leaf nodes for the Date-2 is relatively lower compared to the Date-1.

The significant overlap of distributions among different MR for the Date-1 is evidence that the distinction between different classes is a harder task for the DT (fig. 4.9). Further, the DT had a harder time predicting classes 2 and 3, but it significantly performed better for class 1 and 4. Classes 1 and 4 are both extreme cases of IDC and much easier to detect by the visual rating experts (class 1 is the greenest and class 4 is yellow and almost dying).

However, classes 2 and 3 are in the middle of the crop health status and the chance of getting

them mistaken by each other is much higher than the extreme cases. In another study (Bai et al., 2018), a similar misclassification was experienced for the plots that were in the middle of the health status range. Moreover, the overall average scores of performance for Date-2 are higher than the Date-1 (table 4.3), which suggests that the DT had a harder time distinguishing IDC between classes in the Date-1.

**Table 4.3. Performance of all machine learning models for both dates.**

Model	MR	Visual Score	Date-1 (July 03, 2014)			Date-2 (July 17, 2014)		
			Precision	Recall	f1-score	Precision	Recall	f1-score
Decision Tree	1	1.0	0.67	0.74	0.70	0.79	0.85	0.81
	2	1.5–2.0	0.38	0.26	0.31	0.46	0.46	0.46
	3	2.5–3.0	0.58	0.70	0.64	0.50	0.43	0.46
	4	3.5–4.0	-	-	-	0.86	0.92	0.89
	Average		0.54	0.57	0.55	0.65	0.66	0.66
Random Forest	1	1.0	0.71	0.79	0.75	0.79	0.85	0.81
	2	1.5–2.0	0.80	0.42	0.55	0.50	0.38	0.43
	3	2.5–3.0	0.70	0.95	0.81	0.54	0.50	0.52
	4	3.5–4.0	-	-	-	0.81	1.00	0.90
	Average		0.74	0.72	0.70	0.66	0.68	0.67
AdaBoost	1	1.0	0.78	0.95	0.86	0.80	0.92	0.86
	2	1.5–2.0	0.73	0.58	0.65	0.56	0.38	0.45
	3	2.5–3.0	0.85	0.85	0.85	0.60	0.64	0.62
	4	3.5–4.0	-	-	-	0.93	1.00	0.96
	Average		0.79	0.79	0.78	0.72	0.74	0.72

Note: MR is the merged rating and merges the half step iron deficiency chlorosis (IDC) scores to next integer (e.g., visual IDC score of 2.5 and 3 were merged to MR 3; Refer defined equations for precision (eq. 4.8), recall (eq. 4.9), and f1-score (eq. 4.10).

Therefore, the training data for the border classes 1 and 4 had higher quality and the DT performed better in those cases. Recall for Date-2 shows that 85 % and 92 % of all actual instances of classes 1 and 4 were predicted correctly. Similarly for precision, 79 % and 86 % of

predicted classes were actually belong to classes 1 and 4, respectively. The f1-score indicates that DT tends to score lower than the ensemble methods (fig. 4.11). On the contrary, Naik et al. (2017) found that the performance of DT is higher compared to other ensemble methods in classifying IDC severity. However, they only used two features such as %yellow and %brown pixels, which both were based only on the hue channel. Using a single channel to identify chlorosis is less effective than using DGCI (Hassanijalilian et al., 2020b). Moreover, the parameters that were used to train the models were not discussed in the study, and the train-test split was done randomly disregarding the imbalanced nature of their dataset.

#### **4.4.2. Random Forest**

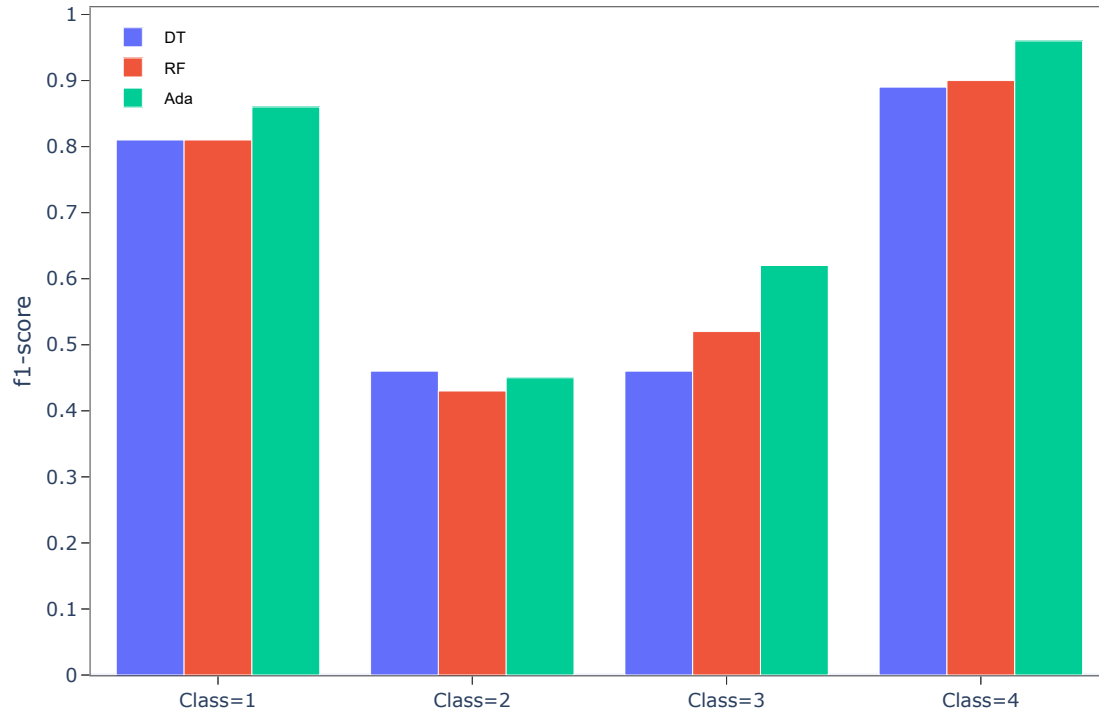
Performance of RF for both dates when the `bootstrap = false`, which determined if RF uses all or portion of the data for training individual trees, is shown in table 4.3. Usually having the `bootstrap` as “true” helps the RF model to generalize better for unseen data, but in this study as the amount of data was limited it reduced the performance, hence this setting was not used. At Date-1, all performance metrics of RF model were significantly higher to those of DT model; however, at Date-2 the metrics were almost similar for both DT and RF models. However, recall for class 4 had a slight increase in RF model. In this study, the recall has an advantage because its purpose is to reduce FN and have the highest number of right predictions of available instances in each class. The effect of FP will be reduced when plots are analyzed for replication because the chance of having the correct classification is higher based on higher recall. Usually, ensemble models like RF tend to perform better than models like DT because they could be trained with different sub-samples and subset of input features so that they can be generalized and perform better regarding the unseen datasets. However, RF was shown to have lower performance metrics compare to DT (Naik et al., 2017). The

speculation of why this has happened have been mentioned in the last section. In this study, the f1-scores for RF are slightly higher in predicting most of the IDC classes (fig. 4.11). The performance of RF in Date-1 is slightly better than Date-2 because there were only 3 classes available in Date-1 and the RF had a lower chance of miss classification compare to 4 classes of Date-2 (table 4.3).

#### **4.4.3. AdaBoost Model Performance**

The AdaBoost model showed an improvement compared to both DT and RF models (fig. 4.11). The average f1-score compared to RF had 8 % and 5 % increase for Date-1 and -2, respectively. The recall is of great importance in this study especially for class 1 because it is the healthiest class and will help farmers to understand and manage IDC symptoms.

AdaBoost increased the overall f1-score of classes 1 and 3 for Date-1 compared to RF, but the performance was significantly higher compared to DT model (16 % and 21 % increase for classes 1 and 3, respectively). With regard to date-2, AdaBoost showed a 7 % increase in recall for class 1 compared to both DT and RF models. The overall f1-score for AdaBoost was at least 5 % higher than other models (72 %). AdaBoost also struggled with classifying the middle MRs (2 and 3), but it performed better than the DT and RF. For instance, f1-score at Date-1 for class 2 is 65 %, which is still lower than those of class 1 and 3 ( $\approx 85\%$ ), but it is much greater than that of DT (31 %) and RF (55 %). Similar to RF, the performance of AdaBoost at Date-1 is significantly better than Date-2 because there were only 3 classes available in Date-1 and the RF had a lower chance of miss classification compare to 4 classes of Date-2 (table 4.3).



**Figure 4.11. Comparison of machine learning models tested showing the f1-score of Date-2 (July 17, 2014), as this date is more representative of the variation in iron deficiency chlorosis (IDC) classes. DT - decision tree, RF - random forest, and Ada - AdaBoost.**

#### 4.4.4. Models Input Features and their Importance

The distribution of the models input features (canopy size, DGCI\*, and DGCI ranges  $c_1 - c_5$ ) that were extracted from the images are presented in table 4.2. Among different classes of DGCI, the ones with lower DGCI ranges ( $c_1$  and  $c_2$ ) had the higher pixel counts. The average value for  $c_1$  for the Date-1 was higher than the Date-2, which corroborates with average DGCI values (DGCI\*) of both dates ( $0.23 < 0.29$ ). The DGCI\* value for the Date-2 is higher than the Date-1 and this suggests that the plants were slightly healthier, so the proportion of  $c_1$  was lower. Similarly, the values for  $c_2$  and  $c_3$  for the Date-2 are higher compared to Date-1 that suggests these plants were less chlorotic. About canopy size, the Date-1 had the smaller average canopy size and with a lower variation which is an indicator

of more uniform emergence and less dead plants. However, the Date-2 had an increased canopy size due to its growth but had a higher variation because IDC stunted growth of the canopy and the number of dead plants was higher.

On the influence of model performance, not all features affect the output with the same weight. Some of these features, such as canopy size and DGCI  $c_1$  contributed the most to train the models. The feature importance, in training the model, with the tree-based models can be explained using a factor termed “mean decrease impurity” (Louppe et al., 2013), evaluated through available function in Python and presented in table 4.4.

**Table 4.4. Feature importance of the machine learning models through mean decrease impurity factor.**

Feature	Date-1 (July 03, 2014)			Date-2 (July 17, 2014)			Mean*
	DT	RF	AdaBoost	DT	RF	AdaBoost	
$c_1$	0.41	0.28	0.14	0.51	0.28	0.22	0.31
$c_2$	0.11	0.15	0.14	0.00	0.01	0.03	0.07
$c_3$	0.03	0.11	0.20	0.22	0.16	0.24	0.16
$c_4$	0.09	0.09	0.08	0.00	0.09	0.05	0.07
$c_5$	0.00	0.00	0.00	0.00	0.01	0.01	0.00
Canopy Size	0.35	0.16	0.26	0.26	0.22	0.32	0.26
Average DGCI	0.00	0.14	0.05	0.00	0.17	0.09	0.08
SD DGCI	0.00	0.08	0.13	0.01	0.06	0.03	0.05

Note: Mean\* represents the average performance of each feature among all models and both dates; Average DGCI is the average DGCI of the pixels of the two middle rows (DGCI\*); SD DGCI is the standard deviation of the DGCI values of the pixels in the two middle rows; DT is decision tree; RF is random forest; and AdaBoost is adaptive boosting.

Among all the features that were used in decision-tree-based models, the DCGI  $c_1$  had the highest importance (table 4.4). The DT, unlike the ensemble models, only uses one tree for prediction, and  $c_1$  shows the highest importance especially for the Date-2, where the average canopy size and variation among canopy sizes were dominant. Canopy size is the

second important feature as IDC causes stunt growth in soybean plots. Middle range DGCI range  $c_3$  is in the third place and it showed significant importance in all models, except DT for the Date-1. The average value and variation of  $c_3$  in the Date-1 were much lower than previous classes and lower than the Date-2, so DT could not use this class effectively for prediction; however, the ensemble methods utilized this class more effectively.

The DGCI\* showed good feature importance only in RF model because RF model is the only model in this study that was set to use a subset of both features and samples to train individual trees. The DGCI\* was more effective when some other features were absent in a tree. Therefore, RF could utilize DGCI\* better than the other models for training and the importance of this feature is relatively higher with RF.

The other features' contributions are significantly low and could be removed from the model. If these features are not removed from the model, they would not decrease the performance of the model in both training and prediction phases because the total number of features is low. Otherwise, these features should be removed, or dimensionality reduction techniques such as principal component analysis (PCA) should be used to improve the performance of the model.

Conversely, simply removing features that have lower correlation with the target variable, sometimes have a potentially negative effect on the performance of the models because ML algorithms could work in higher dimensions that are difficult to illustrate. For example, Bai et al. (2018) removed the ratio\_GC feature, which was observed to be small and changed little across the different IDC scores that could conversely affect the model performance. Hassanijalilian et al. (2020b) found a low correlation between the blue index



and chlorophyll content of soybean leaves, but it significantly increased the performance of the RF model, illustrating against the elimination of low correlation features.

#### **4.5. Conclusions**

Image processing in combination with machine learning was successful in rating iron deficiency chlorosis (IDC) as an alternative to visual rating by experts. The use of the dark green color index (DGCI), previously showed to have a high correlation with chlorophyll content in individual soybean leaves, was also found efficient in plot-scale image application to be used in machine learning algorithm development for IDC rating.

Advanced tree-based machine learning classification algorithms such as decision tree (DT), random forest (RF), and adaptive boosting (AdaBoost) were found suitable in the classification of severity of IDC in soybean plots. AdaBoost and RF models, which are ensembles of decision tree, performed better than a single DT. Even though some performance deviation was observed between the datasets (two dates) studied, the AdaBoost model performance was better than the RF model on both dates and with all classes (f1-score  $\approx 75\%$ ). The MR classes 2 and 3, which are in the middle of the health spectrum and easier to get misclassified even by the field expert rater, represent lower quality training target variable and are the hardest to classify by all models.

To ensure the higher quality of the input data, ratings should be performed by several experts and/or several times by one expert. Testing with a large number of soybean cultivars for IDC measurement with replications providing a sufficiently large number of images will improve the performance of the selected machine learning as well as other advanced methods, such as convolutional neural network and deep learning. As the most effective way to avoid IDC is to use tolerant cultivars, which have the lowest visual ratings, the

image-based objective technique of the study will be of great use in soybean cultivar selection. This developed plot-scale image processing technique could be extended with suitable modification to field-scale operations through aerial imaging platforms (unmanned aerial vehicles, UAV) implementation.

# 5. MEASURING SOYBEAN IRON DEFICIENCY CHLOROSIS PROGRESSION AND YIELD PREDICTION WITH UNMANNED AERIAL VEHICLE \*

## 5.1. Abstract

Iron deficiency chlorosis (IDC) in soybean (*Glycine max* (L.) Merrill), with symptoms of reduction in chlorophyll and stunted growth, causes a great yield loss every year in the Midwest, and the most efficient method to manage IDC is to plant tolerant cultivars. The assessment of cultivars tolerance is traditionally performed by visually rating the IDC symptoms based on leaves discoloration twice during the growing season. However, the visual rating method is time-consuming, subjective, not suitable at large scales, labor-intensive, and unaffordable for frequent observation. Therefore, in this study, we used unmanned aerial vehicle (UAV) as a tool to monitor the soybean cultivars more frequently and more efficiently through image processing approach of the whole field. Images were taken with a DJI Phantom 4 and orthomosaicked in Agisoft Photoscan. A 40-cultivar soybean experimental plots (3000 m<sup>2</sup>; 1 ac) at 5 locations of North Dakota, USA (Amenia, Cofax, Leonard (2), and Hunter) for 2 years (2016 and 2017) were used in the study. The stitched images were processed in MATLAB to calculate the dark green color index (DGCI), which is a

---

\* This paper will be submitted to the *Computers and Electronics in Agriculture* journal. Authors: Oveis Hassanijalilian, C. Igathinathane, Stephanie Day, Sreekala Bajwa, John Nowatzki. Oveis performed the literature survey and wrote the manuscript. Dr. Igathinathane Cannayen is the major advisor and the corresponding author who worked with Oveis throughout the manuscript development. All the co-authors have assisted in the research direction and review of the manuscript.

good indicator of chlorophyll in soybean leaves. The grayscale DGCI images were then processed in ArcGIS to extract average DGCI and canopy size (CS) for each plot for each flight. The area under the curve (AUC) was calculated for DGCI, CS, and CS  $\times$  DGCI product (CDP) to aggregate the values of all flights within each year. The correlation of AUC of CDP and yield was the more consistent among both years and the better predictor of yield ( $R^2 = 0.74$  and  $R^2 = 0.79$ ). The latest growth stage (more representative of yield) values of both years were combined to build models for yield prediction and the CDP produced the lowest error (11.72 %). Future studies should look into IDC progress measurement involving more cultivars, geographical locations, frequent imaging, as well as methods applied to regular soybean production sites to evaluate various image-based parameters and their interaction for yield predictions.

## 5.2. Introduction

In 2004 soybean was planted on  $1.45 \times 10^6$  ha (USDA-NASS, 2005), but in the past three years, it was planted on an average of  $2.7 \times 10^6$  ha in North Dakota (ND), which is almost double the area of the corn (*Zea mays* L.) — a leading field crop of the US. Even though the area of the land that was planted with soybean in ND was almost doubled in 14 years (USDA-NASS, 2020) showing the farmers' interest in the crop, some of the lands in ND are prone to iron deficiency chlorosis (IDC), a symptom of yellowing of plant leaves caused by iron deficiency that affects photosynthesis thereby affecting the soybean yields. The IDC was responsible for \$120 000 000 yield loss in  $1.8 \times 10^6$  ha in just North Central America alone (Hansen et al., 2004). The kind of soil that exacerbates the IDC symptoms in soybean is usually characterized by high PH, calcareous, excess moisture, excess nitrate, and high levels of bicarbonate among other factors (Hansen et al., 2003; Roriz et al., 2014).

Soybeans that are affected by IDC show stunt growth and interveinal chlorosis — loss of green color and becoming yellow, due to lack of chlorophyll, and they might become necrosis — brown and die in severe cases. The IDC happens in younger leaves because iron is immobile after the first trifoliolate emerges, and it might stay in soybean for 8 weeks, but in some cases, IDC might persist throughout the entire season (Hansen et al., 2003).

There are several ways to alleviate the impact of soybean IDC, such as planting a companion crop (Bloom et al., 2011; Naeve, 2006), applying iron chelate (Kaiser et al., 2014), foliar application (Franzen et al., 2003), and increasing seed density at the time of planting (Goos and Johnson, 2000; Wiersma, 2007). However, researches indicate that the most effective way to mitigate IDC symptoms is to plant a tolerant cultivar (Naeve and Rehm, 2006). For example, in a 2002 survey in Minnesota regarding IDC in soybean and best management practices, the majority of farmers stated that the cultivar selection is the best way to avoid IDC (Hansen et al., 2003).

Breeding programs in academic institutions and seed companies are constantly trying to produce soybean cultivars that are more tolerant to IDC. They plant several cultivars in locations that have a history of IDC and rate each cultivar by visually scoring the plants for IDC symptoms, usually two times during the growing season.

This visual method of rating, however, is both labor-intensive and impractical for larger-scale trials. Moreover, the involvement of human experts in this procedure makes it subjective. Karcher and Richardson (2003) used raters (experts performing the visual rating) and digital images to quantify the color of turfgrass, and they observed differences in absolute rating values among different raters. Based on the analysis, they concluded that a more consistent rating can be obtained through digital images across researchers, years, and

locations. Moreover, lighting conditions also affect the appearance of color and the human perception of color and these, in turn, will influence the score (Van Den Broek et al., 2002).

In contrast to visual rating, the application of digital images is expected to be more objective and repeatable. Also, there are methods to compensate for the effect of different light conditions, which affects the visual rating, on digital images through image processing techniques. Furthermore, the digital image based methods can be used in larger-scale trials where the use of human raters is not pragmatic.

At present, platforms that can capture digital images are ubiquitous such as digital single lens reflex (DSLR) camera, smartphone, unmanned aerial vehicle (UAV), manned aircraft, and satellites. All of these devices are equipped with sensors that can capture the reflection of the electromagnetic (EM) spectrum that is visible to human eyes as well as that record other portions of the EM spectrum. The images from these devices can be processed to assess the IDC symptoms through image processing and machine learning techniques.

Hassanijalilian et al. (2020b) used DSLR to rate IDC in soybean plots; the ratings related to plots with the lowest and highest scores were quite successful. Kyaw et al. (2008) used aerial imagery and soil electrical conductivity (EC) to delineate chlorosis management zones. They found that the combination of normalized vegetation index (NDVI) and electrical conductivity (EC) is the best in predicting soybean yield. This combination helped them to delineate the chlorosis zones and associate the yield loss from IDC. Similarly, Rogovska and Blackmer (2009) used manned aircraft to capture aerial images of soybean canopy to map high pH and calcareous soils at field scale as a tool for farmers to identify IDC-susceptible areas in their field.

Given the versatility and affordability, the UAV can be used as an imaging platform to rate IDC and to monitor its progress. In the traditional method, IDC is rated only a couple of times during the growing season (Helms et al., 2010), which is not enough to measure when and the extent of severity of IDC changes through time. However, UAV flights can cover the field in a drastically shorter time compared to a human rater, and this makes it possible to monitor, after proper analysis, the soybean trial plots for as many times as required ( $> 2$ ) to study the changes in IDC severity. Consequently, monitoring the IDC changes at necessary intervals can lead to a better cultivar selection and yield prediction. Moreover, the images can be archived for further or advanced analysis at a later time.

Research on monitoring IDC progress in soybean using UAV and image processing at actual field plot scale and in an uncontrolled condition is not found in the literature, even though this method holds great potential. Therefore, a research investigation on the application of UAV and image processing in monitoring the changes in IDC severity and study its relationship with yield and cultivar tolerance to IDC was proposed. The specific objectives of this research are to (i) evaluate the correlation between image-based crop vegetation index, canopy size, and their interaction with yield in different growth stages, and (ii) develop the prediction models based on progress of soybean IDC and yield.

The methodologies and recommendations from this research will provide a tool to the users to assess the IDC status of their field themselves repeatably with UAV and other limited resources. Outputs of the research work are expected to have a wider application and impact various stakeholders, such as farmers/producers, breeders, crop consultants, agricultural extension agents, and other users.

### 5.3. Materials and Methods

A brief overview of different hardware, and software, and details of the techniques used and developed are presented subsequently. The main piece of hardware is the UAV (DJI Phantom 4; <https://www.dji.com/phantom-4>) and others include the necessary tools for color correction. Software includes Agisoft (Agisoft, 2016), MATLAB (MATLAB, 2015b), and ArcMap (ArcMap, 2017). Tableau (Tableau, 2020) and Python (Van Rossum and Drake, 2009) were used for data visualization and exploratory data analysis, respectively. An overview of all the processes involved in the analysis as a flow diagram is illustrated in figure 5.1.

#### 5.3.1. Field Experimental Plots

Forty different cultivars of soybean were planted in Amenia, ND (47.049 000° N, 97.329 833° W), Colfax, ND, (46.444 170° N, 97.084 731° W), and Leonard, ND, USA (46.671 783° N, 97.245 939° W) mid-May, 2016. In 2017, three locations were flown with UAV, where only two of them were harvested for yield. The locations were Hunter, ND, (47.167 486° N, 97.314 839° W) and Leonard, ND, USA like the previous year. The experimental plots with four replications each (40 cultivars × 4 replications = 160 plots) were seeded at a density of 30 seeds/m in each row (fig. 5.2). Each plot was 4.0 m long, and had four rows with a row spacing of 0.76 m (fig. 5.2). The overall area of the field that contained the experimental plots was about 3000 m<sup>2</sup> (≈ 1 ac).

#### 5.3.2. Image Acquisition

To monitor the growth of soybean and severity of IDC in different growth stages, aerial images were captured for four and five times during the growing season in 2016 and 2017, respectively. The durations between each flight were tried to be close to a week and were



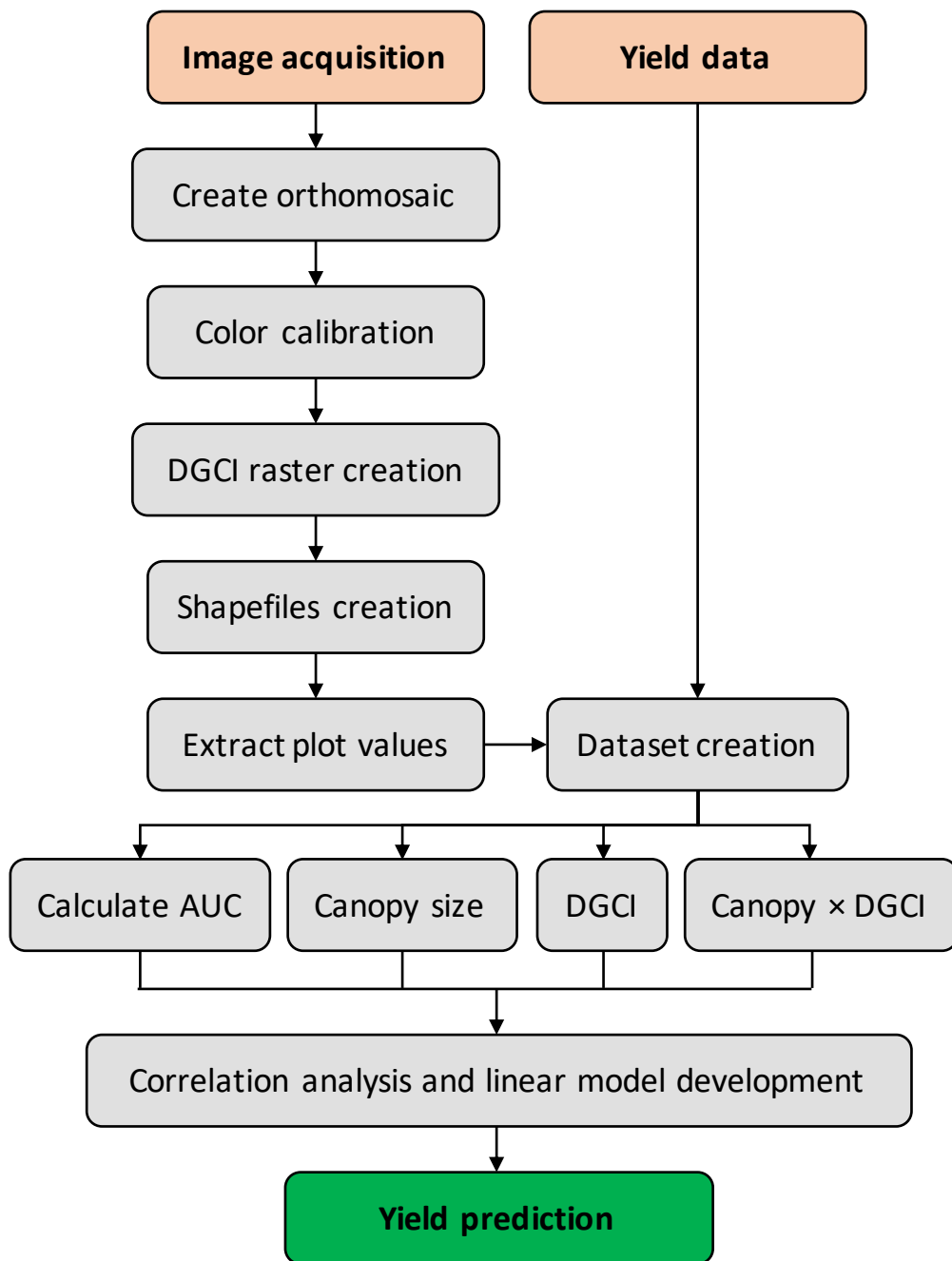


Figure 5.1. Overall flow diagram of the data collection and analysis followed in this study.

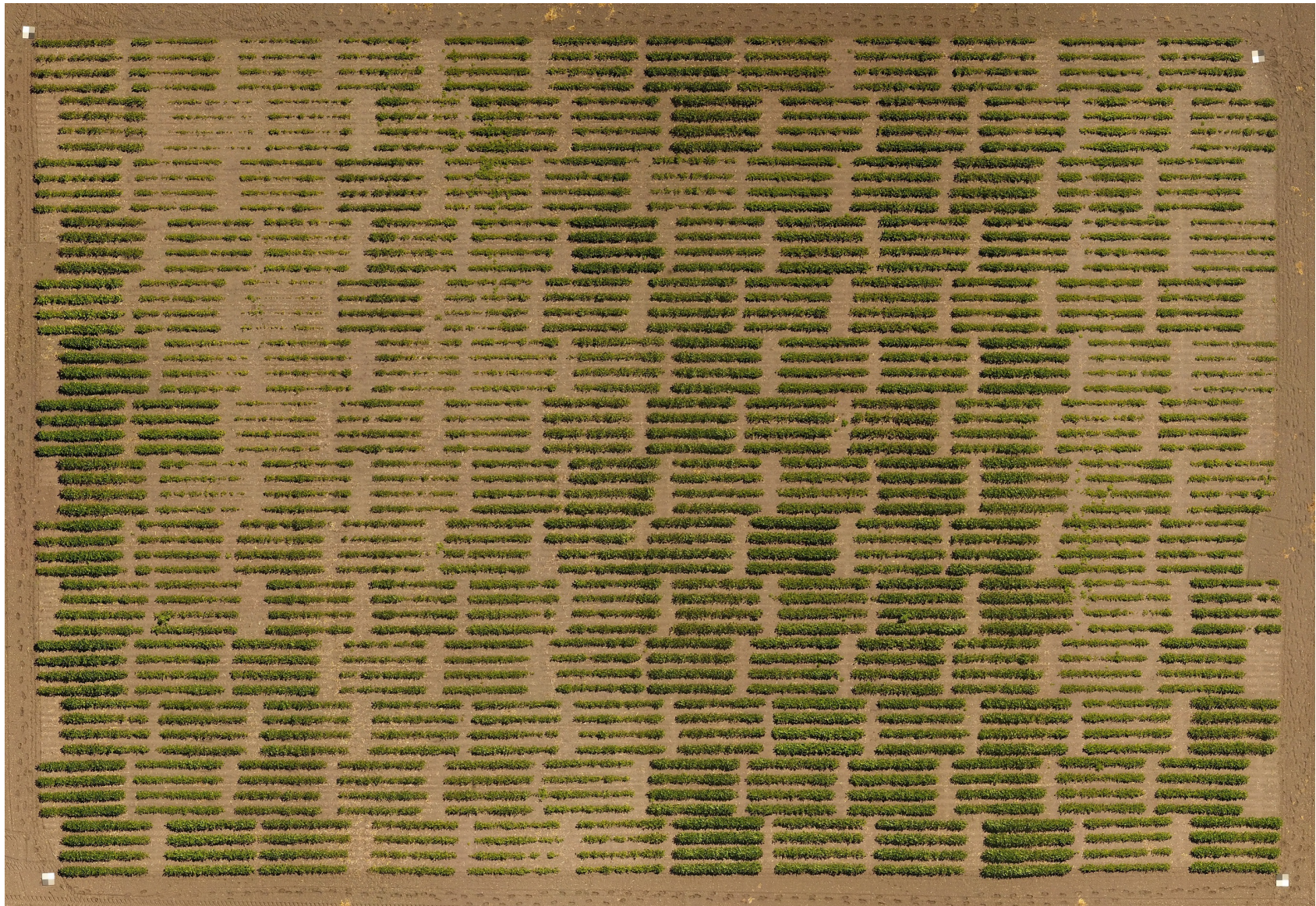


Figure 5.2. Aerial image of Hunter, ND, USA trial site (July 14, 2017). Plots that consist of four rows are easily recognizable because they were planted in an offset pattern. All experimental fields had four replications of 40 soybean cultivars arranged in a quadrant layout. The overall field area of the plots was about  $3000 \text{ m}^2$  ( $\approx 1 \text{ ac}$ ).

scheduled considering the weather condition. Images of all locations were captured within the same day for each growth stage around 10:00–14:00 to have consistent lighting for all locations. Details of image acquisition and image processing are presented subsequently.

#### **5.3.2.1. UAV information**

There are main types of UAV that have a wide range of applications are fixed wings and rotary wings (Sankaran et al., 2015). The fixed wings are good for larger fields and they can withstand higher wind speeds, but rotary wings are better at hovering and flying the smaller fields. A DJI Phantom 4 was used in this study to fly the soybean cultivar trials. Phantom 4 is a quad-copter that weighs 1380 g and was used with the original camera. The camera sensor is 1/2.3" CMOS with 12.4 million effective pixels. The image size is  $4000 \times 3000$ , and the whole camera was stabilized through a gimbal that could help the camera stay nadir against all 3-axis movements of the UAV.

Regarding the lighting of the pictures, the camera settings such as white balance and exposure settings were set to auto to ensure the same amount and quality of light was recorded for each picture throughout the growing season. Images were recorded as JPEG, which is a compressed format, and each image recorded latitude and longitude (geotag) through the on-board GPS. The GPS tags helped Agisoft to align the pictures and recognize their positions.

#### **5.3.2.2. Flight control**

To record the aerial pictures in a consistent way, DroneDeploy (<https://www.dronedeploy.com/solutions/agriculture/>) mobile application was used. DroneDeploy helps a pilot to create a boundary to a field and it creates a serpentine flight path, so the UAV can take pictures of the whole field in a consistent manner. To make the

flight path, there are several settings that some of them affect the accuracy of Agisoft orthomosaic output. The most important of these settings is the amount of front-overlap and side-overlap. The amount of overlap in both directions was set to 70 %. The other settings were flight altitude, flight speed, and exposure control that were set to 100 ft, 5 m/s, and 'Auto,' respectively. The remote control was set to full GPS control mode, so DroneDeploy can function properly.

### **5.3.2.3. Color calibration checkerboards**

Color calibration was performed to account for variation in illumination between images of different dates. In order to correct images for illumination, four checkerboards were placed in each corner around the cultivar trials field (fig. 5.3). The checkerboards consisted of the colors black, white, and two shades of gray — one light and another dark (Haghighattalab et al., 2016; Yu et al., 2016).

Each color on the board had a known (standard) RGB values, and average values of each color on the checkerboard was measured through images and used as observed values. The standard and observed values were used in an “empirical line method” to derive a prediction equation. The equation was calculated separately for each channel and each date to correct the images individually (Yu et al., 2016).

### **5.3.3. Orthomosaic Creation**

Agisoft Photoscan was used to attach individual pictures from the UAV together and create an orthomosaic of each field. Agisoft could recognize the geotag of each image and be able to align them and locate their position. Locating the pictures in Agisoft depends on the accuracy of the on-board GPS. Next, for each image, Agisoft searches for similar pixels due to



**Figure 5.3. Checkerboard that was used for color correction.**

70 % overlap to improve the alignment accuracy. The product was saved as a 'tiff' file and exported to be used in MATLAB, ArcMap, or other image processing software.

#### **5.3.4. Image Processing for Color Calibration and Plots Extraction**

The orthomosaic image covered an area 2-3 times bigger than the cultivar trials, and in order to reduce the size of the file and further calculations, the mosaic files were cropped through MATLAB to only include the cultivar trials.

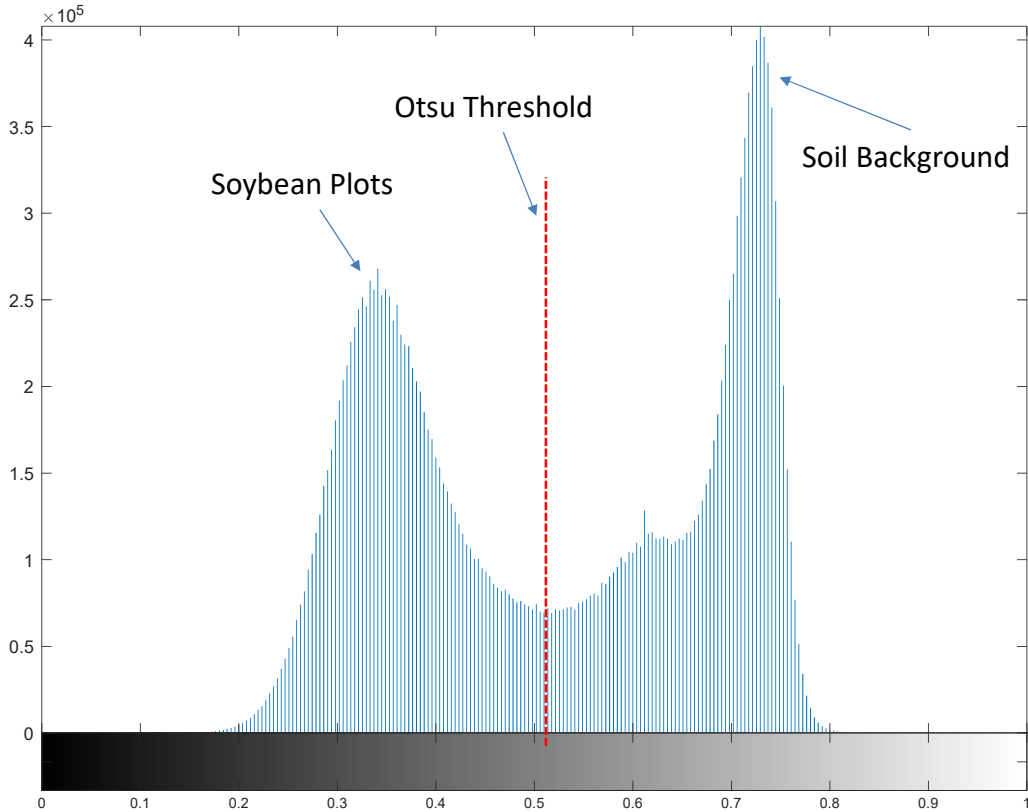
For each field-date average values of the checkerboards were measured through MATLAB, and used alongside the standard values (white section: red=238, green=239, blue=234; black section: red=47, green=47, blue=48; and two shades of gray that were not

used in the color calibration) to calculate the prediction equation. The center of the checkerboard with four colors is could get identified easily. The prediction equation was then used to correct the images of corresponding dates on individual RGB levels.

To isolate the soybean plots from the soil in the background, the RGB color space was converted to  $L^*a^*b^*$ . The  $L^*$  channel represents the lightness of colors from black to white,  $a^*$  represents colors from green to red, and  $b^*$  represents colors from blue to yellow. Since the green component has a stronger presence in soybean canopy, the  $a^*$  channel was used for crop and soil segmentation. To make it easier to work with  $a^*$  channel, and to create a mask the absolute value of the minimum of  $a^*$  was added to all the pixel values of  $a^*$ . Finally, the values were divided by the maximum value of  $a^*$  to convert all the pixel values between 0 and 1 (fig. 5.4). To remove the noise from the image, a median filter was used (`medfilt2()`). Each pixel in the output image contains the median value of a 3-by-3 neighborhood surrounding the corresponding input pixel.

As it appears in figure 5.4,  $a^*$  has a bimodal distribution. The smaller values on the left correspond to soybean plots and the higher values correspond to the soil in the background. Otsu method (`graythresh()` function of MATLAB) finds a threshold between these two classes by maximizing the inter-class variance. This threshold was used to create a binary image through `imbinarize()` function. The values above the threshold were assigned 1 and the rest 0. Since the soil had higher values, the values of the mask were inverted, so the soybean plots could be extracted through the mask.

The dark green color index (DGCI), a crop vegetation index producing good performance, of the image were calculated from the images (eq. 5.1) to monitor the progress and severity of IDC in the soybean plots. Karcher and Richardson (2003) used DGCI to



**Figure 5.4. Histogram of plants and background with Otsu threshold to extract plants of the plot.**

quantify the color of turfgrass using image processing. They observed that DGCI was more consistent than the ratings of the experts. The DGCI has also been used in measuring chlorophyll of soybean leaves, and rating IDC in soybean plots (Hassanijalilian et al., 2020b). Another advantage of DGCI is that it is derived from HSB color space and it uses all the available channels in the image, so it gives a more accurate measurement of green color because the amount of red and blue can affect the appearance of green in the images (Karcher and Richardson, 2003).

$$DGCI = \frac{(\text{Hue} - 60)/60 + (1 - \text{Saturation}) + (1 - \text{Brightness})}{3} \quad (5.1)$$

For easier preprocessing of the color image, the equation 5.1 can be used to generate a DGCi image by mapping all pixels. The DGCi image is a grayscale image with pixels values that were transformed to vary between 0 and 1 with higher values represent greener pixels. Therefore, from DGCi images, the soil is removed between soybean rows using the higher DGCi pixel values, an indicator of greener soybean, with a suitable threshold. These DGCi images were then used in ArcMap to extract individual plots within each field.

### **5.3.5. Object Recognition and Dataset Creation**

The DGCi images were imported into ArcMap to manually digitize polygons over each plot. The first polygon was drawn over the two middle rows of the first plot leaving some margin on the bottom and top portion (fig. 5.5). This area represented by the “rectangular” polygon region of interest (ROI) in each plot was the same area that the visual experts use to rate the IDC in each plot. This polygon ROI was copied over all other plots to maintain a consistent sampling size. All the polygon ROIs were included in a single Environmental Systems Research Institute (ESRI) shapefile. A “PlotNumber” field was added to the shapefile, so it could store the plot numbers for later use in processing and yield analysis.

To automate the process the ModelBuilder of ArcMap was used to design a tool that could get the shapefile and DGCi image as inputs and produce an Excel table with mean DGCi and corresponding plot numbers (fig. 5.6). MATLAB converted the DGCi images to unsigned 8-bit integer values so that all values were between 0 and 255. In order to fix this, ‘Raster Calculator’ was used in ModelBuilder. The ‘float( )’ function was used to convert the integers to float, then when all the pixel values were divided by 255 the DGCi image values fell between 0 and 1.



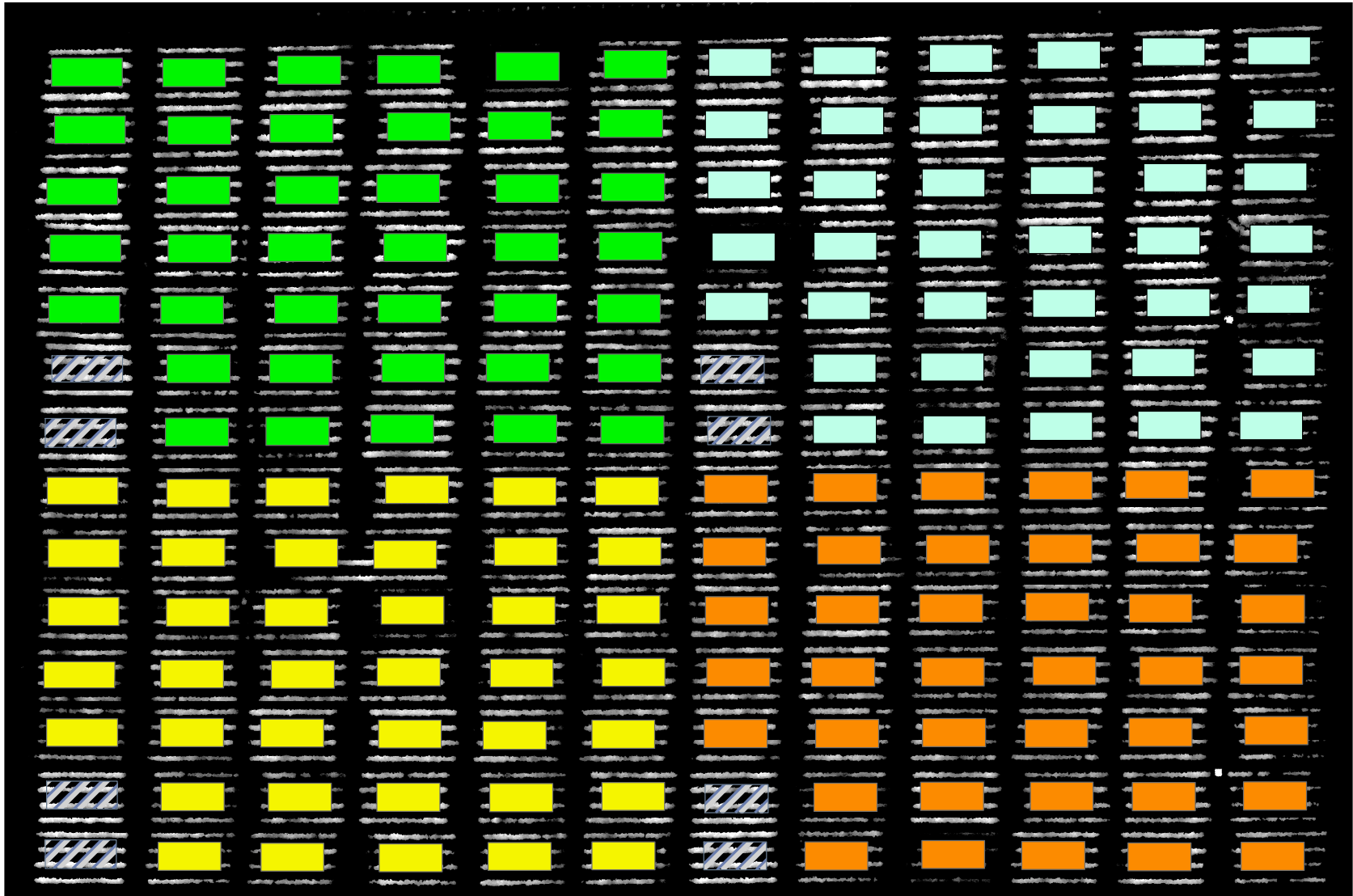
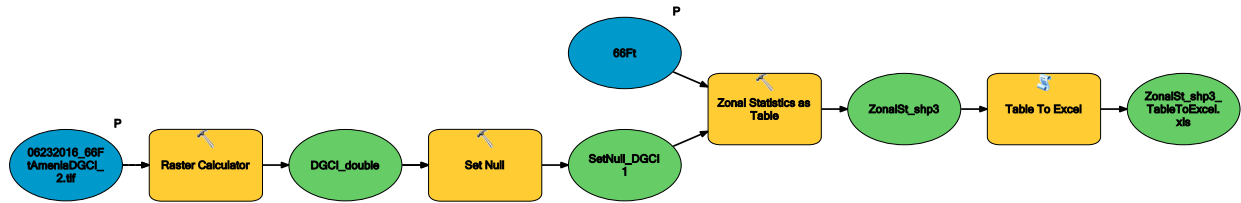


Figure 5.5. Polygonal regions of interest arrangement over soybean plot, and each color is representative of a different replication (40 plots  $\times$  4 replications). Hatched plots were not used in the analysis.



**Figure 5.6. ArcMap ModelBuilder flowchart to derive dark green color index (DGCI) values from each plot.**

Polygonal ROIs that were placed over each plot comprised of both soybean canopies and the soil background in between and around them. For ArcMap to not count the soil pixels into consideration, ‘Set Null’ tool was used to assign Null to pixels with 0 value. Then, ‘Zonal Statistics as Table’ tool was used to aggregate the values related to DGCI pixels that fell within each polygon. The values that were calculated and added to the shapefile were: (i) the average of DGCI pixels, and (ii) the total number of pixels as a representative of the canopy size. The product of the last step was an ArcMap table, which was converted to an Excel table through ‘Table To Excel’ tool. A python script was written to merge the datasets based on plot numbers (Appendix-D). The merged datasets contained the yield data and extracted tables from ArcMap for individual flights.

### 5.3.6. Feature Engineering

Direct (DGCI and canopy size) and combined features (combination of direct and other parameters) derived from the data set to monitor the health status of each soybean plot. All the features were also studied against yield to determine their relationship.

#### 5.3.6.1. DGCI and canopy size

The DGCI and canopy size are good indicators of soybean yield if selected at an appropriate point of time during the crop growth. The continuous variation of these

parameters will be more useful to study the progression of the IDC and vegetative growth of the crop. Sometimes, based on the performance of the cultivar, soybean crop recovers from IDC and the DGCI and canopy size measurements at a later stage (end of peak vegetative stage) will predict the yield better than any other stage or an average.

### 5.3.6.2. Combined DGCI and canopy size

Although DGCI as well as canopy size is informative about IDC severity, it does not provide information on the growth status of soybean plots. Since stunted growth is one of the main indicators of IDC, combining DGCI and canopy size as “canopy × DGCI product” (CDP) is expected to provide more valuable information regarding IDC severity.

### 5.3.6.3. Area under the parameter curve

The area under the curve (AUC) of DGCI and dates, and CDP and dates were used for each plot to combine all the DGCI values from all flights to a single value (Simko and Piepho, 2012). This aggregation provides a better input for yield prediction models as opposed to single-point data such as DGCI of the latest growth stage. The data of the latest growth stage even though expected to correlate well with the final yield performance of the crop, the limited data might not have enough variations to build a reliable model, but AUC can take into account all the previous growth stages (trend). The higher the AUC is for a plot the healthier it is and the less time it was affected by IDC. The AUC can be calculated as:

$$\text{AUC} = \sum_{i=1}^{n_f-1} \frac{y_i + y_{i+1}}{2} \times (t_{i+1} - t_i) \quad (5.2)$$

where AUC is the area under the curve,  $y_i$  is the DGCI for the  $i$ th flight (dimensionless),  $t_i$  is the time in day at the  $i$ th flight, and  $n$  is the number of flights. The AUC was also calculated again where  $y_i$  is CDP at the time of each flight.

## 5.4. Results and Discussion

### 5.4.1. Attributes of Soybean IDC Trial Plots

Summary of different attributes measured or calculated for soybean plots for both years and all locations are presented in table 5.1. Colfax, ND in 2016 had the highest average yield among all sites in both years. Furthermore, it had one of the lowest standard deviations among the sites in 2016, which showed that high yield was more consistent among all plots within this location. Similarly, Colfax had the highest average canopy size with the lowest standard deviation. The canopy size of Colfax was close to Leonard, ND in 2016, but the average yield was significantly greater than Leonard. This can be due to IDC leading the plots to stunt growth and show high variability in the canopy size in Leonard. The DGCI for Colfax 2016 is the highest among all locations and years, which suggests that the soybean plots were in better health compare to the other sites, and it led to the highest yield. However, Colfax in 2016 did not have the lowest average IDC score, but it had the highest variation. On the other hand, Leonard had the lowest yield in 2016 along with the highest IDC score and lowest average DGCI value. Leonard's location had a history of severe IDC and the same pattern was observed in 2017. In 2017, the average canopy size was smaller than that of 2016, and the standard deviation of canopy size for both locations were double the amount for all locations in 2016. This shows a great variability in canopy growth and was one of the indicators of lower yield in 2017. Both locations had lower yield compared to 2016, but Leonard produced a significantly lower yield.

All attributes regarding 2017 show significant differences among both locations (table 5.1); despite the fact that the cultivars were the same for both locations, IDC symptoms were more prevalent in Leonard, which caused a much lower yield. Although

**Table 5.1. Distribution of the measured and calculated attributes of soybean IDC trial plots at different ND locations and years.**

Attribute	2016			2017	
	Amenia	Colfax	Leonard	Hunter	Leonard
Yield (kg/ha)	3487(990)	5267(634)	3470(580)	2419(556)	625(455)
Canopy size (m <sup>2</sup> )	3.92(0.73)	4.95(0.34)	4.67(0.39)	3.91(0.80)	1.61(0.84)
IDC score	1.24(0.32)	2.06(0.54)	2.15(0.51)	2.35(0.79)	3.28(0.42)
DGCI *	0.31(0.03)	0.35(0.04)	0.27(0.03)	0.30(0.04)	0.23(0.02)

Iron deficiency chlorosis (IDC) score visually rated (range: 1–5) with higher values represent severe IDC symptoms and vice versa. \* DGCI values (range: 0–1) represent average value of all plots among all flights with higher values represent healthier fields.

Hunter, ND had a higher yield than Leonard in 2017, it yielded less than all the locations including Leonard in 2016. All the attributes of Hunter in 2017 are close to those of Leonard 2016; the yield of these to sites are also close to each other.

#### **5.4.2. IDC Progress Through DGCI Trends for Different Cultivars**

Breeders usually rate IDC twice during the growing season to monitor the severity of IDC and to assess if a cultivar recovers from the IDC. However, different cultivars recover at different rates and some may recover even after the second visual rating. Aerial images make it possible to monitor the plots more frequently and in a more efficient way. Figure 5.7 shows different patterns of IDC recovery for the common cultivars in both years. Each line represents a plot and each color represents each location, the four replications of each cultivar in a location are shown in one color.

For some cultivars, plots (replications) within each location follow one another closely, which is an indicator that they show the same symptoms of IDC in different spots in a field such as Dahlman 6703XN in 2016 (fig. 5.7). On the other hand, some plots show different symptoms of IDC among a location, which could be an indicator of different severity of IDC

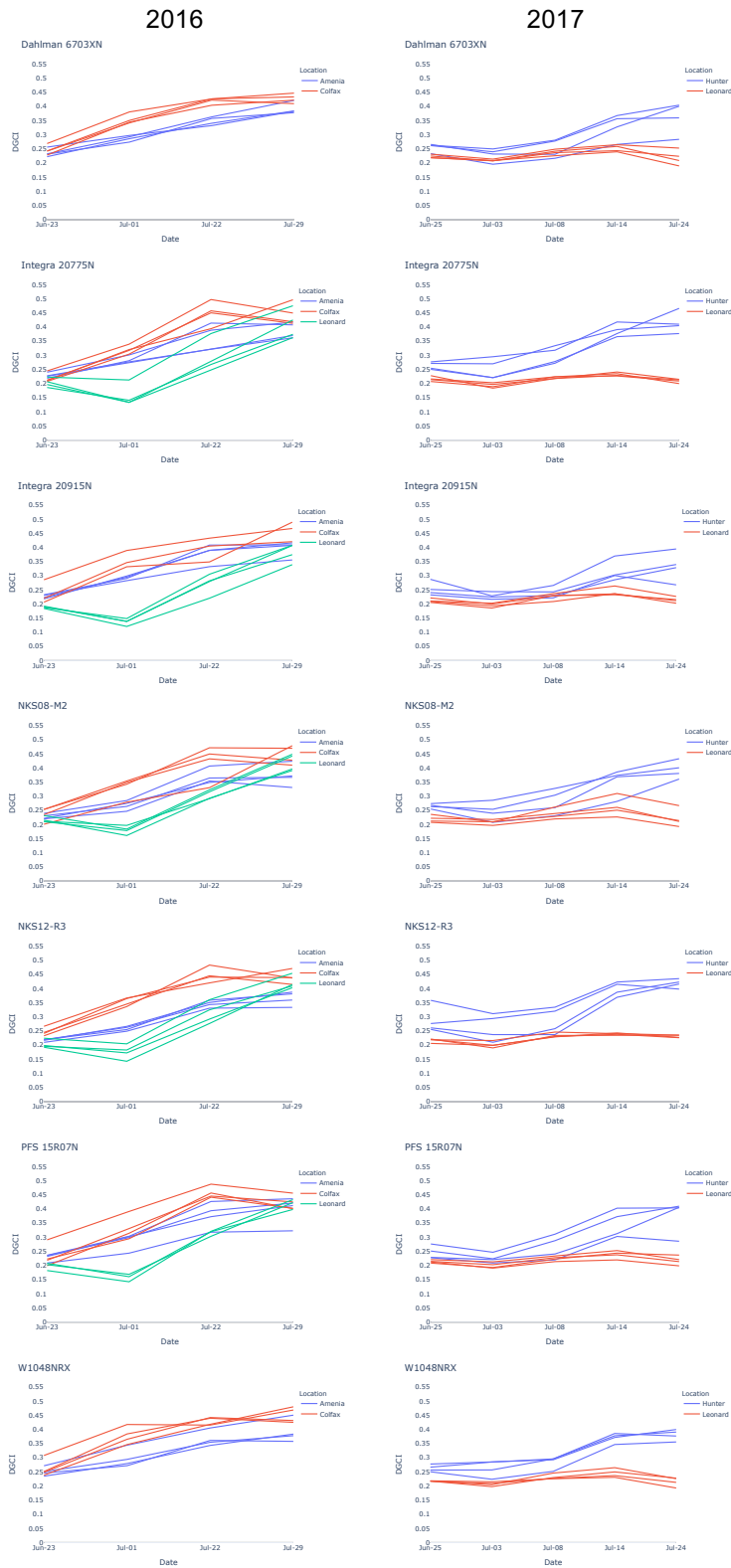


Figure 5.7. The dark green color index (DGCI) trend among common cultivars in 2016 and 2017 at different locations.

(NKS12-R3 in 2017 and Integra 20775N in 2016 fig. 5.7). Furthermore, some cultivars have close responses to IDC for all plots and locations such as Dahlman 6703XN (2016) and W1048NRX (2016), which shows a consistent response to IDC. However, the DGCI values for both cultivars were removed from the analysis for Leonard where the IDC had the highest severity. This could be an indicator that these two cultivars are not tolerant to IDC, but show great DGCI values under no IDC.

The average and standard deviation of DGCI for all the trial plots in each location with time (different UAV flights) are shown in table 5.2. All the locations have increasing DGCI as the crops grow except Leonard in both years. In 2016, the DGCI at Leonard for the first flight was close to those of other locations, but it decreased in the second flight, while the increase in DGCI values was significantly high with other sites. In 2017, Hunter had the increasing DGCI for all dates except a slight drop in the second flight. The same pattern was observed for Leonard, but with smaller DGCI values for all dates, and the last flight had a drop of DGCI as well.

**Table 5.2. Distribution of the DGCI for each flight on different locations and dates.**

Date	2016			Date	2017	
	Amenia	Colfax	Leonard		Hunter	Leonard
June 23	0.23(0.02)	0.23(0.03)	0.20(0.02)	June 25	0.26(0.02)	0.22(0.01)
July 01	0.28(0.02)	0.33(0.04)	0.16(0.03)	July 03	0.24(0.04)	0.21(0.01)
July 22	0.35(0.04)	0.41(0.05)	0.30(0.04)	July 08	0.28(0.04)	0.24(0.02)
July 29	0.37(0.04)	0.43(0.03)	0.40(0.04)	July 14	0.35(0.04)	0.25(0.02)
				July 24	0.38(0.04)	0.22(0.03)

DGCI - dark green color index. Average DGCI values at respective dates and locations are shown along with their respective standard deviation in the parenthesis.

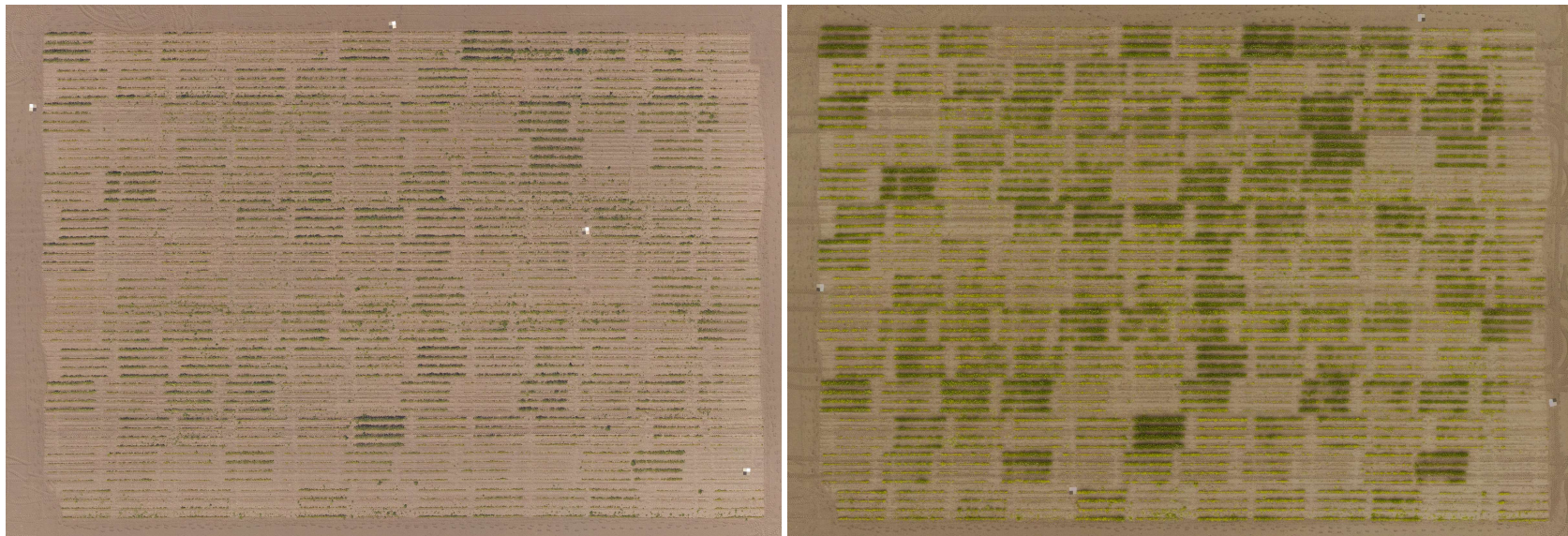
The aerial images of Leonard for both July 14 and 24, 2017 shown side by side, and it is apparent that most of the plots, instead of recovering from IDC, got worse and become yellow (fig. 5.8), even though some natural vegetative growth was observed. Lower DGCI values of all flights ( $\leq 0.43$ ; max = 1.0), as well as a reduction in DGCI in the second and last flight, were symptoms of severe IDC and low yield at the end of the season. In both years Leonard had the lowest DGCI and also the lowest yield. Multiple observation of soybean cultivars (table 5.2) gives a better understanding of each cultivar's IDC progress or recovery rate compared to the visual ratings recorded.

#### **5.4.3. DGCI Variation on Yield**

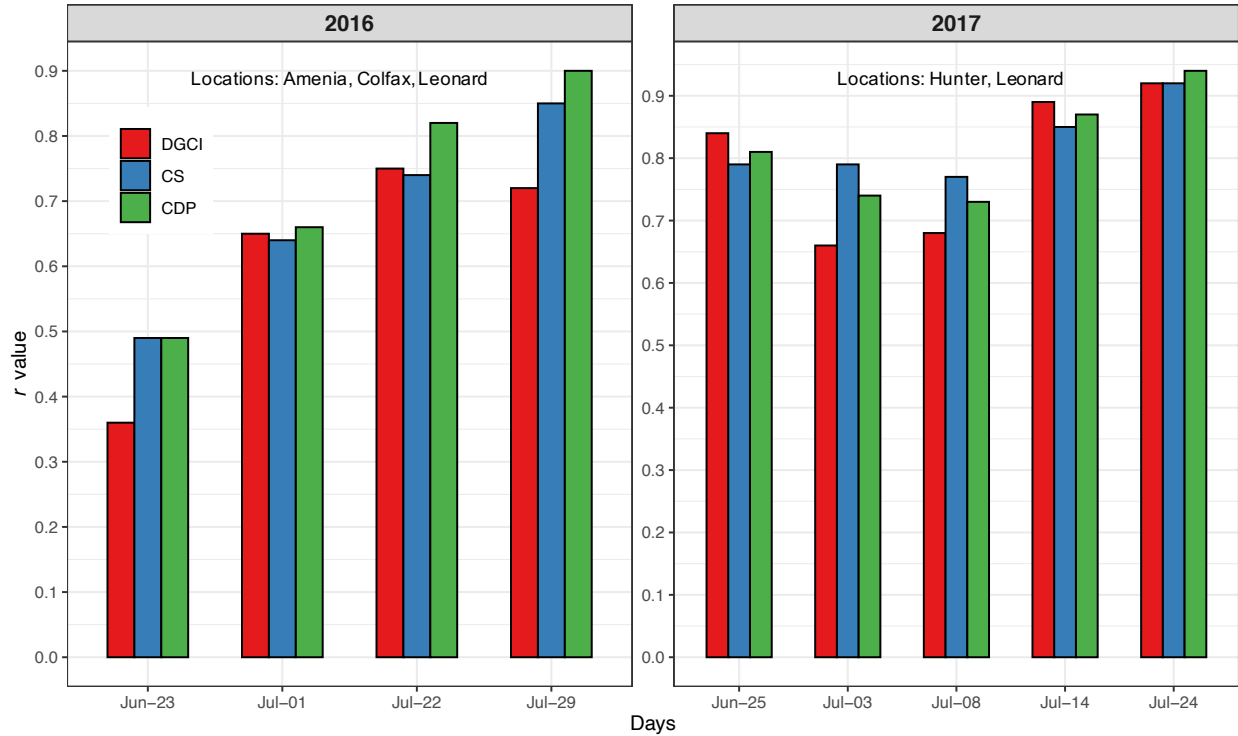
The correlation of DGCI with yield in 2016 and 2017 at different experiment locations (fig. 5.9) showed an overall increase in the correlation ( $r$  value) with the progress of the season. This observation indicated that the DGCI at later stages of crop growth is better correlated with the yield and produced better prediction models. Plots of complete data set for both years and respective locations presented as appendix figures (figs. D1 and D2) showed a clear separation (as different clusters of data) of the DGCI vs yield among different locations, which indicated the growth and yield pattern of the crop cultivar and location were different.

In 2017, because of the yield difference, this trend is more pronounced even towards the two locations considered. As observed in the IDC progress of DGCI (fig. 5.8) a drop in the values of 2016 in 2/3 of locations in the 3rd and 4th flight dates, and in 2017 a total flat trend on 1/2 of locations, reflected on the DGCI correlation (fig. 5.9). Soybean crop recovery from IDC, which can also be observed with DGCI vs yield correlation on the later stages, especially in 2017.





**Figure 5.8. Unmanned aerial vehicle orthomosaic images of Leonard, ND. (Left) Plots of small canopy on July 14, 2017; and (Right) Plots with slightly increased canopy but with many plots the iron deficiency chlorosis (IDC) symptoms got worse compared to July 24, 2017.**



**Figure 5.9. Unmanned aerial vehicle image extracted parameters correlation with yield in 2016 and 2017 at different experiment locations. DGCI - dark green color index, CS - canopy size, and CDP - canopy and DCCI product.**

As expected, higher values of DGCI (dark green of the canopy) are related to higher yields, and the later flights have a better relationship; and the same positive relationship is exhibited for all locations in 2017 (figs. D1 and D2). A similar pattern is also reported in different studies that different vegetation indices including normalized difference vegetation index (NDVI) relationship with the yield for soybean was stronger at later growth stages such as R2 – R5 (Maimaitijiang et al., 2020; Zhang et al., 2019).

#### 5.4.4. CS Variation on Yield

The average of DGCI alone is not a good predictor of IDC because it does not give any information on the size of the canopy. In some cases, only a few plants have survived at the

end of the season but they were healthy and might have produced normal yield per plant, however, have affected the overall yield of the plot that was measured. In such situations, only considering DGCI can mislead the yield prediction. It is interesting to observe that the CS follows a similar correlation variation as DGCI (figs. D1 and D2). A bigger canopy is an indicator of soybean plots' health and results in higher yield. However, in the case of IDC, soybean canopy can stop growing and in some cases shrinks or dies. A similar pattern can be observed as different plots show different patterns of growth between the last two flights in Leonard in 2017 (fig. 5.8).

In 2016, unlike DGCI the CS, as the plants grow, a clear increasing trend of CS vs yield correlation with days was seen; however, in 2017, there was a dip in the correlation of CS is not as pronounced in DGCI because CS most likely will increase towards a peak value, and any stagnation or reduction might be due to stunted growth or plants dying. The increase in the later stages might indicate the crop recovery from IDC as observed both with CS and DGCI and a positive effect on the prediction of these parameters to predict the yield (fig. 5.9).

The canopy size had a positive relationship with yield in both years, but this correlation was stronger for 2017. Both the years had a higher correlation towards the end of the growing season. The variation in canopy size was greater in 2017, so the relationship between canopy size and yield was the greatest ( $r = 0.92$ ). Hoyos-Villegas and Fritschi (2013) observed the same relationship ( $r = 0.89$ ) when soybeans were under moderate or no water stress. However, the relationship was weak when the water stress was severe. In a different study, almost similar correlation ( $r = 0.71$ ) was observed at the R3 growth stage (Yu et al., 2016).

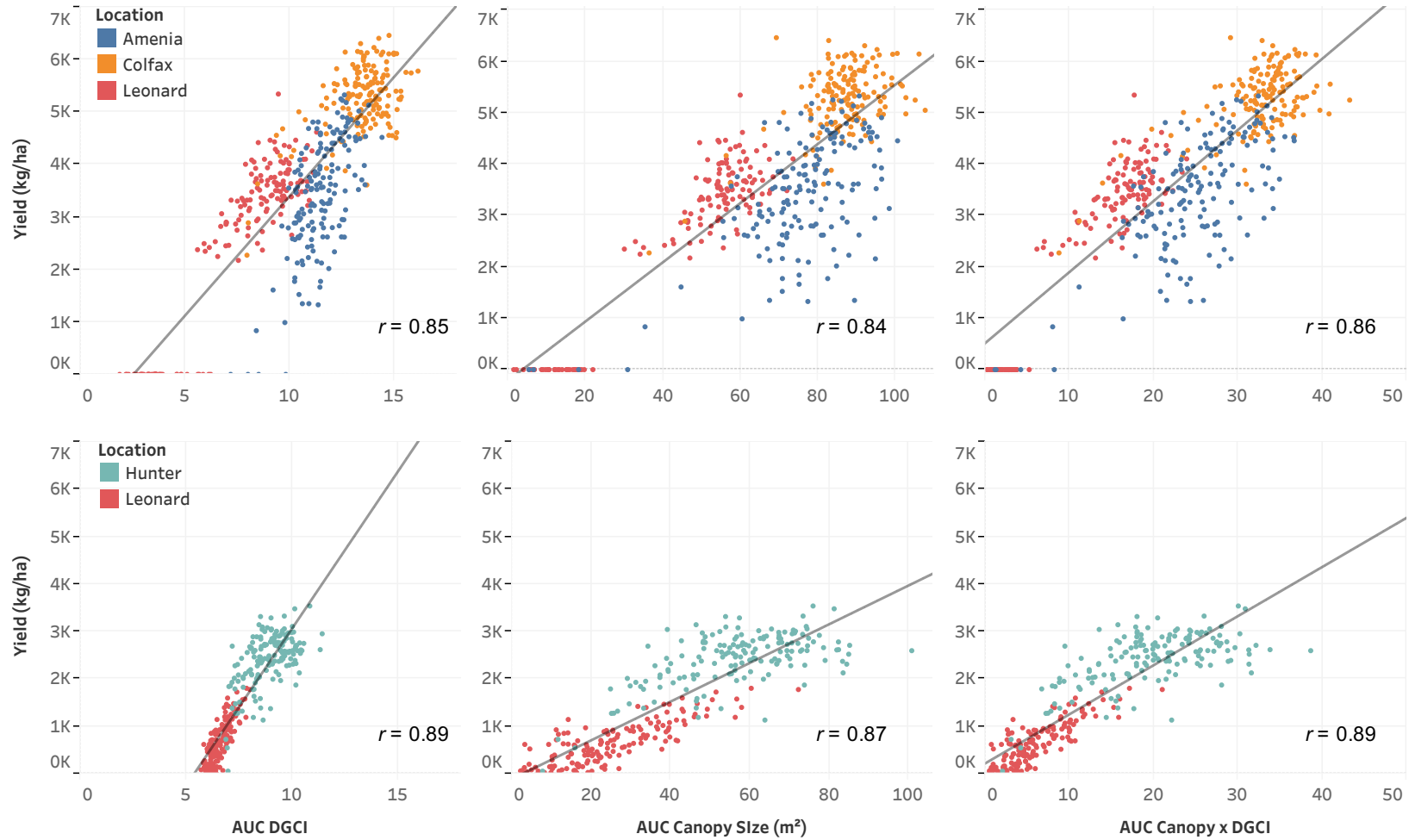
#### 5.4.5. CDP Variation on Yield

As observed before, higher values of DGCI and CS are indicators of healthier soybean plots. Therefore, the combined CDP product can be a better and potential indicator and higher values of this product signify healthier soybeans (fig. 5.9). The CDP in both years showed a positive correlation between yield and CDP (figs. D1 and D2). Similar to DGCI and canopy size, the graphs that are related to flight days toward the end of the growing season have a stronger correlation. In 2016, the last flight shows a strong correlation ( $r = 0.84$ ), and it is as strong as the last flight in 2017 ( $r = 0.94$ ). This correlation for the most flights of both years was greater than those of DGCI or canopy size. Based on these observations, the interaction between canopy size and DGCI (CDP) could be a good predictor of yield.

#### 5.4.6. AUC Variation on Yield

The relationship between AUC of DGCI (ADGCI), AUC of canopy size (ACS), as well as of AUC of CDP (ACDP) and yield is shown in figure 5.10 considering  $y_i$  in equation 5.2. Both trends show a good correlation with yield, however, the ACDP seems to have a stronger correlation with yield in general and is more consistent among both years ( $r = 0.86$  and  $r = 0.89$  for 2016 and 2017, respectively) compared to only DGCI ( $r = 0.85$  and  $r = 0.89$ ).

The AUC aggregates all values for all flights and gives a better indicator compared to when the yield is predicted based on one flight (single data point). For instance, the correlation of ADGCI and yield for 2016 ( $r = 0.85$ ) is greater than that of all flights of that year. This shows that including the history of DGCI can be helpful for predicting yield. Similarly, in 2017, the correlation of ADGCI and yield ( $r = 0.89$ ) is greater than that of all flights that year except the last one which had the highest correlation ( $r = 0.92$ ).

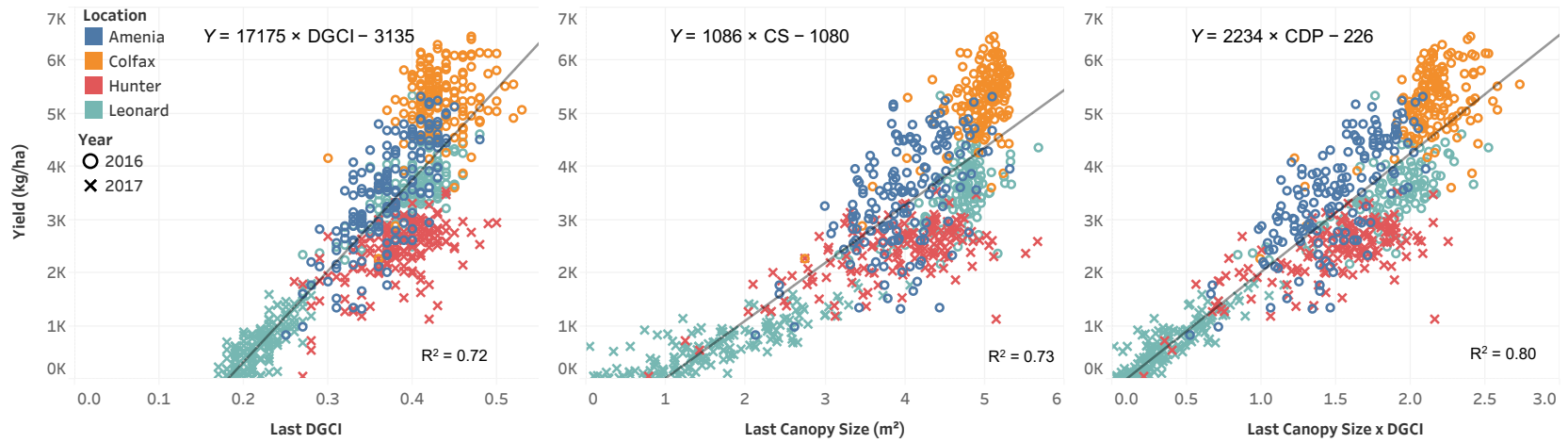


**Figure 5.10.** Yield correlations with area under the curve (AUC) of parameters at different locations and years. (Left) AUC for dark green color index (DGCI); (Middle) AUC of canopy size, and (Right) AUC for Canopy  $\times$  DGCI.

Despite the highest correlation of AUC and yield, these data could not be combined in this study because the sampling dates were different. Moreover, the period of monitoring in 2016 was greater than that of 2017 (36 days and 29 days for 2016 and 2017, respectively). However, by assuming that soybeans are at the same growth stages at last flight, the CS and DGCI could be combined as defined earlier as CDP for both years to build a more representative dataset. The scatterplot (fig. 5.11) shows the individual (DGCI and CS) and combined (CDP) data distribution and yield with corresponding linear models. As expected the CDP had a stronger fit with yield ( $R^2 = 0.80$ ) compared to the direct inputs ( $R^2 \leq 0.73$ ).

The various best fitted linear models of yield with selected independent parameters (ADGCI, ACS, and ACDP) are presented in table 5.3 along with their corresponding coefficient of determination ( $R^2$ ), root mean square error (RMSE), and the error (ratio of the RMSE compared to the respective yield) in percentage. The measured data and the fitted models are also shown in figures 5.10 and 5.11. At first glance, based on RMSE, it seems that the models in 2017 performed better. However, the error was almost the same as other models because the yield was much lower in 2017. Overall the best performance is for the model that covers both years and uses CDP as an independent variable. The error rate is lower than all the other models (11.72 %) and it is more representative because it is based on two years of data.

Even though the ADGCI and canopy DGCI interaction is a good and consistent predictor of yield and a good metric to select tolerant cultivars to IDC. In this study, as the weather restriction, the sampling dates, and the period of study were not the same for both years, the AUC could not be combined with CDP for both years. However, the latest growth stage (subdata), which are more relevant to grain yield than earlier stages, values of each



**Figure 5.11.** Unmanned aerial vehicle image combined data extracted parameters variation with yield and models for both years and locations. (Left) Dark green color index (DGCI) variation with yield; (Middle) Canopy size (CS) variation with yield; and (Right) Canopy and DGCI product (CDP) variation with yield.

**Table 5.3. Soybean yield prediction models to predict yield based on different variables and combination.**

Parameter	Year	Model	$R^2$	RMSE (kg/ha)	Error
Direct	2016	$Y = 19155 \times \text{DGCI} - 3621$	0.52	1304	20.18 %
		$Y = 1006 \times \text{CS} - 405$	0.72	892	13.80 %
	2017	$Y = 10581 \times \text{DGCI} - 1668$	0.85	397	11.22 %
		$Y = 668 \times \text{CS} - 320$	0.84	418	11.82 %
	Both years	$Y = 17175 \times \text{DGCI} - 3135$	0.72	833	12.89 %
		$Y = 1086 \times \text{CS} - 1080$	0.73	880	13.61 %
Combined	2016	$Y = 2303 \times \text{CDP} - 98$	0.81	756	11.70 %
		$Y = 455 \times \text{ADGCI} - 1182$	0.72	888	13.74 %
		$Y = 58 \times \text{ACS} - 237$	0.70	918	14.21 %
		$Y = 138 \times \text{ACDP} + 495$	0.74	850	13.16 %
	2017	$Y = 1472 \times \text{CDP} + 136$	0.88	361	10.21 %
		$Y = 661 \times \text{ADGCI} - 3594$	0.80	456	12.90 %
		$Y = 41 \times \text{ACS} - 130$	0.76	502	14.19 %
		$Y = 103 \times \text{ACDP} + 199$	0.79	477	13.48 %
	Both years	$Y = 2234 \times \text{CDP} - 226$	0.80	757	11.72 %

RMSE - root mean square error, Error - ratio of RMSE to the corresponding yield, Y - soybean yield (kg/ha), DGCI - dark green color index, CS - canopy size ( $\text{m}^2$ ), CDP - canopy and DGCI product extracted from the image, ADGCI - area under curve of DGCI with time, ACS - area under curve of CS with time, ACDP - area under curve of CDP with time. The direct parameter (DGCI and CS) models used the data of the latest growth stage (subset) within the dataset. Maximum yield were 6461 and 3537 kg/ha for 2016 and 2017, respectively.

year were combined and used to build models to predict yield. Frequent and consistent monitoring during all growth stages in the future will help in determining what independent variables will model the soybean multi-year yield data.

### 5.5. Conclusions

Selecting a tolerant cultivar is the most efficient way to avoid iron deficiency chlorosis (IDC), which at present performed by visual rating by experts; and this process of visual rating is subjective, time-consuming, not feasible in larger scales, and not frequent enough to record the recovery of all cultivars. Unmanned aerial vehicle (UAV) was successfully used in



this study to monitor both severity of IDC and recovery rate for each cultivar throughout the growing season using the image processing approach and model the yield based on several image extracted parameters. The dark green color index (DGCI), canopy size (CS), and CS and DGCI product (CDP), extracted using image processing of UAV imagery, were useful in assessing the IDC progress of different soybean cultivars and model crop yields. The best correlation for these features and yield was observed in the latest growth stage, as the earlier stages of crop growth did not influence significantly the crop yield. However, using the single value data (latest growth stage) is not representative of the overall health status of soybean plots, and therefore the area under the curve (AUC) that encompasses the whole trend was found to successfully aggregate different features that were extracted from aerial images across different times.

The correlation between AUC of CDP and yield was more consistent across the years studied and the linear model produced about 13 % error for each year. As the AUC cannot be combined with CDP for both years, due to weather restriction and the difference in sampling dates in both years, however, the latest growth stage values of both year were combined and CDP produced the best yield prediction model (11.72 % error). Future studies should look into IDC progress measurement involving more cultivars, geographical locations, and frequent monitoring flights and image data. Also, along with IDC experiments the regular soybean production sites could be evaluated using AUC of DGCI, canopy size, and their interaction for yield predictions.

## 6. GENERAL CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

### 6.1. General Conclusions

Soybean is one of the important agricultural commodities especially in the Midwest. However, iron deficiency chlorosis (IDC) can significantly reduce its yield. To alleviate IDC from soybean fields, it is known that planting a tolerant cultivar is the most efficient method. Digital imaging on different platforms was used in this study to reduce the subjectivity that is in the process of selecting tolerant cultivars. Moreover, these process could help breeders to both monitor the progresses of IDC and perform studies in larger scales. Machine learning models were used along with the digital imaging platforms, and these models were able to detect the pattern that was not feasible for humans to detect among the input features with high dimensions. The findings from each objective are presented here:

- Objective 1 (*IDC — review*): Different methods of measuring IDC were explained through this objective that are suitable for research projects in small and large scales. Small scale projects that can be done in a controlled environment can benefit from methods such as atomic absorption spectroscopy, chlorophyll extraction, and crop sensors. Visual rating and image processing through different platforms such as digital cameras, unmanned aerial vehicle (UAV), and satellites could be helpful in quantifying IDC for research and production. The lack of more sophisticated sensors such as hyperspectral cameras, LiDAR for IDC assessment is noticeable among research

publications. Soil grid sampling with higher resolution could be used along with many of the aforementioned methods' output for the purpose of building machine learning models to gain a better understanding of IDC.

- Objective 2 (*Chlorophyll measurement — smartphone*): Several machine learning models were used to estimate the chlorophyll content of soybean leaves infield condition from smartphone images. The models were simple linear regression (SLR), polynomial regression (PR), multiple linear regression (MLR), random forest (RF), and support vector machine (SVM). Different sets of input features were used to train these models such as average DGCI of all pixels of a leaf, average RGB value, separate averages of RGB, pixel count of different ranges of DGCI (RPC) and separate RGB + RPC. The SVM model had the highest  $R^2$  of 0.89 and lowest root mean square (RMSE) of 2.90 in predicting the chlorophyll content. It even performed better than other models where it used separate averages of RGB without color calibration ( $R^2 = 0.86$  and  $RMSE = 3.20$ ).
- Objective 3 (*Plot IDC rating — handheld camera*): Decision tree based models such as decision tree, random forest, and adaptive boosting were used to classify different severity of IDC. These models are famous supervised machine learning models that can be trained using a dataset with predefined target feature labels. Features such as canopy size and pixel count of different classes of DGCI were extracted from the images that were captured close to the human point of view with a DSLR. The adaptive boosting model performed better than other models especially in classifying classes that are in both extreme ends of the health status. All the performance metrics were

significantly lower for two classes in the middle of health status. That was due to a higher error in the visual rating of these classes that were mistaken for each other.

- Objective 4 (*IDC cultivar trials — UAV*): Aerial images of soybean cultivar trials were captured 4-5 times per growing season in two years to monitor the growth of soybean as well as the recovery rate of IDC among different cultivars. It was found that DGCI that were extracted from images toward the end of the growing season had a higher correlation with yield. The area under the curve (AUC) of DGCI for different flights were extracted for individual plots. It was found that the AUC that is calculated from  $DGCI \times \text{canopy size}$  is highly correlated to the yield and more consistent among both years ( $R^2 = 0.79$ ).

## 6.2. Suggestions for Future Work

The future research works suggested here pertain to possible work that can be extended from each of the objectives considered in this dissertation.

1. Develop convolutional neural networks from different cultivars of soybean infield conditions.
2. Develop models that could benefit from multispectral and hyperspectral sensors for early detection of IDC.
3. Develop models that incorporate LiDAR and RGB technology to rate IDC.
4. Develop machine learning models that incorporate high-resolution soil grid sampling to rate IDC in commercial fields.
5. Develop a methodology that can benefit from historical yield maps and other sensors' output to study the effect of multi-cultivar planting.

## REFERENCES

- Agisoft (2016). Agisoft PhotoScan user manual: Professional Edition, Ver. 1.2. St. Petersburg, Russia; [https://www.agisoft.com/pdf/photoscan-pro\\_1\\_2\\_en.pdf](https://www.agisoft.com/pdf/photoscan-pro_1_2_en.pdf).
- Aiken, V. C. F., Dórea, J. R. R., Acedo, J. S., de Sousa, F. G., Dias, F. G., & de Magalhães Rosa, G. J. (2019). Record linkage for farm-level data analytics: Comparison of deterministic, stochastic and machine learning methods. *Computers and Electronics in Agriculture*, 163, 104857.
- ArcMap, E. (2017). 10.6. *Environmental System Research Institute*.
- Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. *Plant Physiology*, 24(1), 1.
- ASA (2019). 2019 SOYSTATS a reference guide to soybean facts and figures. American Soybean Association; URL: SoyGrowers.com.
- Atefi, A., Ge, Y., Pitla, S., & Schnable, J. (2019). In vivo human-like robotic phenotyping of leaf traits in maize and sorghum in greenhouse. *Computers and Electronics in Agriculture*, 163, 104854.
- Azadbakht, M., Ashourloo, D., Aghighi, H., Radiom, S., & Alimohammadi, A. (2019). Wheat leaf rust detection at canopy scale under different lai levels using machine learning techniques. *Computers and Electronics in Agriculture*, 156, 119–128.
- Bai, G., Jenkins, S., Yuan, W., Graef, G. L., & Ge, Y. (2018). Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. *Frontiers in Plant Science*, 9.

- Bakhsipour, A. & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Computers and Electronics in Agriculture*, 145, 153–160.
- Bascle, B., Bernier, O., & Lemaire, V. (2006). Illumination-invariant color image correction. In *International Workshop on Intelligent Computing in Pattern Analysis and Synthesis* (pp. 359–368).: Springer.
- Beaty, R. D. & Kerber, J. D. (1978). *Concepts, instrumentation and techniques in atomic absorption spectrophotometry*. Perkin-Elmer USA.
- Bin, L. M., Weng, L., & Bugter, M. H. (2016). Effectiveness of FeEDDHA, FeEDDHMA, and FeHBED in preventing iron-deficiency chlorosis in soybean. *Journal of Agricultural and Food Chemistry*, 64(44), 8273–8281.
- Bloom, P. R., Rehm, G. W., Lamb, J. A., & Scobbie, A. J. (2011). Soil nitrate is a causative factor in iron deficiency chlorosis in soybeans. *Soil Science Society of America Journal*, 75(6), 2233–2241.
- Burgos-Artizzu, X. P., Ribeiro, A., Guijarro, M., & Pajares, G. (2011). Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*, 75(2), 337–346.
- Cavallo, D. P., Cefola, M., Pace, B., Logrieco, A. F., & Attolico, G. (2017). Contactless and non-destructive chlorophyll content prediction by random forest regression: A case study on fresh-cut rocket leaves. *Computers and Electronics in Agriculture*, 140, 303–310.

- Chatterjee, A., Lovas, S., Rasmussena, H., & Goos, R. (2017). Foliar application of iron fertilizers to control iron deficiency chlorosis of soybean. *Crop, Forage & Turfgrass Management*, 3(1).
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Collings, S., Caccetta, P., Campbell, N., & Wu, X. (2011). Empirical models for radiometric calibration of digital aerial frame mosaics. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7), 2573–2588.
- Coste, S., Baraloto, C., Leroy, C., Marcon, É., Renaud, A., Richardson, A. D., Roggy, J.-C., Schimann, H., Uddling, J., & Hérault, B. (2010). Assessing foliar chlorophyll contents with the SPAD-502 chlorophyll meter: A calibration test with thirteen tree species of tropical rainforest in French Guiana. *Annals of Forest Science*, 67(6), 607–607.
- da Silva Junior, C. A., Nanni, M. R., Shakir, M., Teodoro, P. E., de Oliveira-Júnior, J. F., Cezar, E., de Gois, G., Lima, M., Wojciechowski, J. C., & Shiratsuchi, L. S. (2018). Soybean varieties discrimination using non-imaging hyperspectral sensor. *Infrared Physics & Technology*, 89, 338–350.
- De Mendiburu, F. (2019). *Agricolae: Statistical Procedures for Agricultural Research*. R package version 1.3-0, CRAN.
- Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., & Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy*, 4(3), 349–379.

- Dobbels, A. A. & Lorenz, A. J. (2019). Soybean iron deficiency chlorosis high throughput phenotyping using an unmanned aircraft system. *Plant Methods*, 15(1), 97.
- Domingos, P. (1999). Metacost: A general method for making classifiers cost-sensitive. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 155–164). Volume 99.
- Dyson, J., Mancini, A., Frontoni, E., & Zingaretti, P. (2019). Deep learning for soil and crop segmentation from remotely sensed data. *Remote Sensing*, 11(16), 1859.
- Espejo-Garcia, B., Martinez-Guanter, J., Pérez-Ruiz, M., Lopez-Pellicer, F. J., & Zarazaga-Soria, F. J. (2018). Machine learning for automatic rule classification of agricultural regulations: A case study in Spain. *Computers and Electronics in Agriculture*, 150, 343–352.
- Fehr, W., Caviness, C., Burmood, D., & Pennington, J. (1971). Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill 1. *Crop Science*, 11(6), 929–931.
- Ferreira, C. M., López-Rayó, S., Lucena, J. J., Soares, E. V., & Soares, H. M. (2019a). Evaluation of the efficacy of two new biotechnological-based freeze-dried fertilizers for sustainable Fe deficiency correction of soybean plants grown in calcareous soils. *Frontiers in Plant Science*, 10, 1335.
- Ferreira, C. M., Sousa, C. A., Sanchis-Pérez, I., López-Rayó, S., Barros, M. T., Soares, H. M., & Lucena, J. J. (2019b). Calcareous soil interactions of the iron (iii) chelates of dph and azotochelin and its application on amending iron chlorosis in soybean (*Glycine max*). *Science of the Total Environment*, 647, 1586–1593.



- Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 479, 448–455.
- Franzen, D., O'Barr, J., & Zollinger, R. (2003). Interaction of a foliar application of iron heda and three postemergence broadleaf herbicides with soybeans stressed from chlorosis. *Journal of Plant Nutrition*, 26(12), 2365–2374.
- Gamble, A., Howe, J., Delaney, D., Van Santen, E., & Yates, R. (2014). Iron chelates alleviate iron chlorosis in soybean on high pH soils. *Agronomy Journal*, 106(4), 1251–1257.
- Geetha, R., Sivasubramanian, S., Kaliappan, M., Vimal, S., & Annamalai, S. (2019). Cervical cancer identification with synthetic minority oversampling technique and PCA analysis using random forest classifier. *Journal of Medical Systems*, 43(9), 286.
- Géron, A. (2017). *Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc.
- Gitelson, A. A., Vina, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, 32(8).
- Golzarian, M. R. & Frick, R. A. (2011). Classification of images of wheat, ryegrass and brome grass species at early growth stages using principal component analysis. *Plant Methods*, 7(1), 28.
- Goos, R. J. & Johnson, B. (2001). Seed treatment, seeding rate, and cultivar effects on iron deficiency chlorosis of soybean. *Journal of Plant Nutrition*, 24(8), 1255–1268.
- Goos, R. J. & Johnson, B. E. (2000). A comparison of three methods for reducing iron-deficiency chlorosis in soybean. *Agronomy Journal*, 92(6), 1135–1139.

- Guijarro, M., Pajares, G., Riomoros, I., Herrera, P., Burgos-Artizzu, X., & Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1), 75–83.
- Haghighattalab, A., Pérez, L. G., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., Ortiz-Monasterio, I., Singh, R. P., Goodin, D., & Poland, J. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods*, 12(1), 35.
- Hamuda, E., Glavin, M., & Jones, E. (2016). A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture*, 125, 184–199.
- Hansen, N., Jolley, V., Naeve, S., & Goos, R. (2004). Iron deficiency of soybean in the North Central US and associated soil properties. *Soil Science and Plant Nutrition*, 50(7), 983–987.
- Hansen, N., Schmitt, M. A., Anderson, J., & Strock, J. (2003). Iron deficiency of soybean in the upper midwest and associated soil properties. *Agronomy Journal*, 95(6), 1595–1601.
- Hassanijalilian, O., Igathinathane, C., Bajwa, S., & Nowatzki, J. (2020a). Rating iron deficiency in soybean using image processing and decision-tree based models. *Submitted to Remote Sensing, MDPI*.
- Hassanijalilian, O., Igathinathane, C., Doetkott, C., Bajwa, S., Nowatzki, J., & Haji Esmaeili, S. A. (2020b). Chlorophyll estimation in soybean leaves infield with smartphone digital imaging and machine learning. *Computers and Electronics in Agriculture*, 174, 105433.

- Helms, T., Scott, R., Schapaugh, W., Goos, R., Franzen, D., & Schlegel, A. (2010). Soybean iron-deficiency chlorosis tolerance and yield decrease on calcareous soils. *Agronomy Journal*, 102(2), 492–498.
- Horst, G., Engelke, M., & Meyers, W. (1984). Assessment of visual evaluation techniques 1. *Agronomy Journal*, 76(4), 619–622.
- Hoyos-Villegas, V. & Fritschi, F. (2013). Relationships among vegetation indices derived from aerial photographs and soybean growth and yield. *Crop Science*, 53(6), 2631–2642.
- Hsiao, C. (1975). Some estimation methods for a random coefficient model. *Econometrica: Journal of the Econometric Society*, (pp. 305–325).
- Jajda, H., Patel, K., Patel, S., Solanki, V., Patel, K., & Singh, S. (2015). Comparative efficacy of two standard methods for determination of iron and zinc in fruits, pulses and cereals. *Journal of Food Science and Technology*, 52(2), 1096–1102.
- Jin, X., Liu, S., Baret, F., Hemerlé, M., & Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sensing of Environment*, 198, 105–114.
- Jones Jr, J. B. (2001). *Laboratory Guide for Conducting Soil Tests and Plant Analysis*. CRC press.
- Kaiser, D. & Bloom, P. (2018). Managing iron deficiency chlorosis in soybean. University of Minnesota Extension - Online Publication: Saint Paul, MN, USA.
- Kaiser, D., Lamb, J., & Bloom, P. (2011). Managing iron deficiency chlorosis in soybean. *University of Minnesota Extension*.

- Kaiser, D., Lamb, J., Bloom, P., & Hernandez, J. (2014). Comparison of field management strategies for preventing iron deficiency chlorosis in soybean. *Agronomy Journal*, 106(6), 1963–1974.
- Karcher, D. E. & Richardson, M. D. (2003). Quantifying turfgrass color using digital image analysis. *Crop Science*, 43(3), 943–951.
- Kataoka, T., Kaneko, T., Okamoto, H., & Hata, S. (2003). Crop growth estimation system using machine vision. In *Advanced Intelligent Mechatronics, 2003. AIM 2003. Proceedings. 2003 IEEE/ASME International Conference on*, volume 2 (pp. b1079–b1083).: IEEE.
- Kyaw, T., Ferguson, R. B., Adamchuk, V. I., Marx, D., Tarkalson, D., & McCallister, D. (2008). Delineating site-specific management zones for pH-induced iron chlorosis. *Precision Agriculture*, 9(1-2), 71–84.
- LaHuis, D. M. & Ferguson, M. W. (2009). The accuracy of significance tests for slope variance components in multilevel random coefficient models. *Organizational Research Methods*, 12(3), 418–435.
- Lelong, C. C., Burger, P., Jubelin, G., Roux, B., Labbé, S., & Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors*, 8(5), 3557–3585.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
- Liesch, A., Diaz, D., Martin, K., Olson, B., Mengel, D., & Roozeboom, K. (2011). Management strategies for increasing soybean yield on soils susceptible to iron deficiency. *Agronomy Journal*, 103(6), 1870–1877.

- Liu, J., Pattey, E., & Admiral, S. (2013). Assessment of in situ crop LAI measurement using unidirectional view digital photography. *Agricultural and Forest Meteorology*, 169, 25–34.
- Liu, J., Pattey, E., & Jégo, G. (2012). Assessment of vegetation indices for regional crop green LAI estimation from Landsat images over multiple growing seasons. *Remote Sensing of Environment*, 123, 347–358.
- Liu, S., Baret, F., Andrieu, B., Burger, P., & Hemmerle, M. (2017). Estimation of wheat plant density at early stages using high resolution imagery. *Frontiers in Plant Science*, 8, 739.
- Louppe, G., Wehenkel, L., Sutura, A., & Geurts, P. (2013). Understanding variable importances in forests of randomized trees. In *Advances in Neural Information Processing Systems* (pp. 431–439).
- Lucena, J. J. (2003). Fe chelates for remediation of fe chlorosis in strategy I plants. *Journal of Plant Nutrition*, 26(10-11), 1969–1984.
- Ma, X., Zhu, K., Guan, H., Feng, J., Yu, S., & Liu, G. (2019). High-throughput phenotyping analysis of potted soybean plants using colorized depth images based on a proximal platform. *Remote Sensing*, 11(9), 1085.
- Maimaitijiang, M., Ghulam, A., Sidike, P., Hartling, S., Maimaitiyiming, M., Peterson, K., Shavers, E., Fishman, J., Peterson, J., Kadam, S., et al. (2017). Unmanned aerial system (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 43–58.
- Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., & Fritschi, F. B. (2020). Soybean yield prediction from uav using multimodal data fusion and deep learning. *Remote Sensing of Environment*, 237, 111599.

- Markwell, J., Osterman, J. C., & Mitchell, J. L. (1995). Calibration of the minolta spad-502 leaf chlorophyll meter. *Photosynthesis Research*, 46(3), 467–472.
- MATLAB (2015a). Matlab Image Processing Toolbox (R2015b). The MathWorks, Inc., Natick, MA, USA.
- MATLAB (2015b). Version 8.6 (R2015b). The MathWorks Inc., Natick, MA, USA.
- McClendon, J. H. & Fukshansky, L. (1990). On the interpretation of absorption spectra of leaves–I. Introduction and the correction of leaf spectra for surface reflection. *Photochemistry and Photobiology*, 51(2), 203–210.
- Meyer, G., Mehta, T., Kocher, M., Mortensen, D., & Samal, A. (1998). Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. *Transactions of the ASAE*, 41(4), 1189.
- Meyer, G. E. & Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, 63(2), 282–293.
- Minolta (1989). Manual for chlorophyll meter SPAD-502, Minolta Co., Ltd. Minolta Radiometric Instruments Divisions, Osaka, Japan.
- Moghadam, P., Ward, D., Goan, E., Jayawardena, S., Sikka, P., & Hernandez, E. (2017). Plant disease detection using hyperspectral imaging. In *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (pp. 1–8).: IEEE.
- Nadal, P., García-Delgado, C., Hernández, D., López-Rayó, S., & Lucena, J. J. (2012). Evaluation of Fe-N, N'-Bis (2-hydroxybenzyl) ethylenediamine-N, N'-diacetate (HBED/Fe<sup>3+</sup>) as Fe carrier for soybean (*Glycine max*) plants grown in calcareous soil. *Plant and Soil*, 360(1-2), 349–362.

- Naeve, S. L. (2006). Iron deficiency chlorosis in soybean: Soybean seeding rate and companion crop effects. *Agronomy Journal*, 98(6), 1575–1581.
- Naeve, S. L. & Rehm, G. W. (2006). Genotype× environment interactions within iron deficiency chlorosis-tolerant soybean genotypes. *Agronomy Journal*, 98(3), 808–814.
- Naik, H. S., Zhang, J., Lofquist, A., Assefa, T., Sarkar, S., Ackerman, D., Singh, A., Singh, A. K., & Ganapathysubramanian, B. (2017). A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. *Plant Methods*, 13(1), 23.
- NDAWN (2015). North Dakota Agricultural Weather Network, NDAWN monthly data for 5 months beginning April 2015.  
[https://ndawn.ndsu.nodak.edu/get-table.html?station=72&variable=mdavt&year=2020&ttype=monthly&quick\\_pick=&begin\\_date=2015-04&count=5](https://ndawn.ndsu.nodak.edu/get-table.html?station=72&variable=mdavt&year=2020&ttype=monthly&quick_pick=&begin_date=2015-04&count=5)  
Accessed: 2020-01-28.
- Neter, J. (1974). *Applied Linear Statistical Models; Regression, Analysis of Variance, and Experimental Designs*. Homewood, Ill: R.D. Irwin.
- Nowling, G. L. & Cai, G. (2018). The uniform soybean tests: Northern region 2018.
- Page, G. F., Liénard, J. F., Pruet, M. J., & Moffett, K. B. (2018). Spatiotemporal dynamics of leaf transpiration quantified with time-series thermal imaging. *Agricultural and Forest Meteorology*, 256, 304–314.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-Learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

- Peng, Y., Nguy-Robertson, A., Arkebauer, T., & Gitelson, A. (2017). Assessment of canopy chlorophyll content retrieval in maize and soybean: implications of hysteresis on the development of generic algorithms. *Remote Sensing*, 9(3), 226.
- Polyakov, A. Y., Lebedev, V. A., Volkov, D. S., Pankratov, D. A., Veligzhanin, A. A., Perminova, I. V., Lucena, J. J., et al. (2019). Eco-friendly iron-humic nanofertilizers synthesis for the prevention of iron chlorosis in soybean (*Glycine max*) grown in calcareous soil. *Frontiers in Plant Science*, 10, 413.
- Porra, R. (1991). Recent advances and re-assessments in chlorophyll extraction and assay procedures for terrestrial, aquatic, and marine organisms, including recalcitrant algae. *Chlorophyll*, (pp. 31–57).
- Porra, R., Thompson, W., & Kriedemann, P. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 975(3), 384–394.
- Porra, R. J. (2002). The chequered history of the development and use of simultaneous equations for the accurate determination of chlorophylls a and b. *Photosynthesis Research*, 73(1-3), 149–156.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, R version 3.5.2., Vienna, Austria.
- 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. *Computers and Electronics in Agriculture*, 156, 585–605.



- Rigon, J., Capuani, S., Fernandes, D., & Guimarães, T. (2016). A novel method for the estimation of soybean chlorophyll content using a smartphone and image analysis. *Photosynthetica*, 54(4), 559–566.
- Rodríguez-Lucena, P., Hernández-Apaolaza, L., & Lucena, J. J. (2010). Comparison of iron chelates and complexes supplied as foliar sprays and in nutrient solution to correct iron chlorosis of soybean. *Journal of Plant Nutrition and Soil Science*, 173(1), 120–126.
- Rogovska, N. & Blackmer, A. M. (2009). Remote sensing of soybean canopy as a tool to map high pH, calcareous soils at field scale. *Precision Agriculture*, 10(2), 175.
- Rorie, R. L., Purcell, L. C., Karcher, D. E., & King, C. A. (2011). The assessment of leaf nitrogen in corn from digital images. *Crop Science*, 51(5), 2174–2180.
- Roriz, M., Carvalho, S. M., & Vasconcelos, M. W. (2014). High relative air humidity influences mineral accumulation and growth in iron deficient soybean plants. *Frontiers in Plant Science*, 44, 189–198.
- Sadeghi-Tehran, P., Virlet, N., Sabermanesh, K., & Hawkesford, M. J. (2017). Multi-feature machine learning model for automatic segmentation of green fractional vegetation cover for high-throughput field phenotyping. *Plant Methods*, 13(1), 103.
- Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., Miklas, P. N., Carter, A. H., Pumphrey, M. O., Knowles, N. R., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *European Journal of Agronomy*, 70, 112–123.
- Santos, C. S., Ozgur, R., Uzilday, B., Turkan, I., Roriz, M., Rangel, A. O., Carvalho, S. M., & Vasconcelos, M. W. (2019). Understanding the role of the antioxidant system and the tetrapyrrole cycle in iron deficiency chlorosis. *Plants*, 8(9), 348.

- SAS (2015). The MIXED procedure - random coefficients. SAS/STAT(R) 14.1 User's Guide. SAS Institute Inc., Cary, NC, USA.
- Schenkeveld, W., Dijcker, R., Reichwein, A., Temminghoff, E., & Van Riemsdijk, W. (2008). The effectiveness of soil-applied feeddha treatments in preventing iron chlorosis in soybean as a function of the o, o-feeddha content. *Plant and Soil*, 303(1-2), 161–176.
- Shahriar, M. S., Rahman, A., & McCulloch, J. (2014). Predicting shellfish farm closures using time series classification for aquaculture decision support. *Computers and Electronics in Agriculture*, 102, 85–97.
- Simko, I. & Piepho, H.-P. (2012). The area under the disease progress stairs: calculation, advantage, and application. *Phytopathology*, 102(4), 381–389.
- Smidt, E. R., Conley, S. P., Zhu, J., & Arriaga, F. J. (2016). Identifying field attributes that predict soybean yield using random forest analysis. *Agronomy Journal*, 108(2), 637–646.
- Smith, A. R. (1978). Color gamut transform pairs. *ACM Siggraph Computer Graphics*, 12(3), 12–19.
- Smith, G. M. & Milton, E. J. (1999). The use of the empirical line method to calibrate remotely sensed data to reflectance. *International Journal of Remote Sensing*, 20(13), 2653–2662.
- Tableau (2020). Tableau Prep Help: 2020 Tableau Software. Seattle, WA, USA; <https://www.tableau.com>.
- Teklić, T., Vratarić, M., Sudarić, A., Kovačević, V., Vukadinović, V., & Bertić, B. (2009). Relationships among chloroplast pigments concentration and chlorophyllmeter

- readings in soybean under influence of foliar magnesium application. *Communications in Soil Science and Plant Analysis*, 40(1-6), 706–725.
- Thomas, S., Kuska, M. T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., Behmann, J., & Mahlein, A.-K. (2018). Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. *Journal of Plant Diseases and Protection*, 125(1), 5–20.
- Uddling, J., Gelang-Alfredsson, J., Piikki, K., & Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research*, 91(1), 37–46.
- USDA-NASS (2005). Crop production 2004 summary.
- USDA-NASS (2020). Crop production 2019 summary.
- Van Den Broek, E., Vuurpijl, L., Kisters, P., & Von Schmid, J. (2002). Content-based image retrieval: Color-selection exploited. In *Proceedings of the Dutch-Belgian Information Retrieval Workshop (DIR2002)*, volume 3 (pp. 37–46).
- Van Rossum, G. & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- Vasconcelos, M. W. & Grusak, M. A. (2014). Morpho-physiological parameters affecting iron deficiency chlorosis in soybean (*Glycine max* L.). *Plant and Soil*, 374(1-2), 161–172.
- Vega, F. A., Ramirez, F. C., Saiz, M. P., & Rosúa, F. O. (2015). Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop. *Biosystems Engineering*, 132, 19–27.

- Vesali, F., Omid, M., Kaleita, A., & Mobli, H. (2015). Development of an android app to estimate chlorophyll content of corn leaves based on contact imaging. *Computers and Electronics in Agriculture*, 116, 211–220.
- Vollmann, J., Walter, H., Sato, T., & Schweiger, P. (2011). Digital image analysis and chlorophyll metering for phenotyping the effects of nodulation in soybean. *Computers and Electronics in Agriculture*, 75(1), 190–195.
- Von Wettstein, D. (1957). Chlorophyll-letale und der submikroskopische formwechsel der plastiden. *Experimental Cell Research*, 12(3), 427–506.
- Wiersma, J. V. (2007). Iron acquisition of three soybean varieties grown at five seeding densities and five rates of fe–eddha. *Agronomy Journal*, 99(4), 1018–1028.
- Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, 38(1), 259–269.
- Yadav, S. P., Ibaraki, Y., & Gupta, S. D. (2010). Estimation of the chlorophyll content of micropropagated potato plants using RGB based image analysis. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 100(2), 183–188.
- Yang, N., Liu, D., Feng, Q., Xiong, Q., Zhang, L., Ren, T., Zhao, Y., Zhu, D., & Huang, J. (2019). Large-scale crop mapping based on machine learning and parallel computation with grids. *Remote Sensing*, 11(12), 1500.
- Yu, N., Li, L., Schmitz, N., Tian, L. F., Greenberg, J. A., & Diers, B. W. (2016). Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sensing of Environment*, 187, 91–101.

Zhang, J., Naik, H. S., Assefa, T., Sarkar, S., Reddy, R. C., Singh, A., Ganapathysubramanian, B., & Singh, A. K. (2017). Computer vision and machine learning for robust phenotyping in genome-wide studies. *Scientific Reports*, 7(1), 1–11.

Zhang, X., Zhao, J., Yang, G., Liu, J., Cao, J., Li, C., Zhao, X., & Gai, J. (2019). Establishment of plot-yield prediction models in soybean breeding programs using UAV-based hyperspectral remote sensing. *Remote Sensing*, 11(23), 2752.

## APPENDIX A. FUNCTION TO DETECT THE LEAF, GREEN AND YELLOW DISK

This appendix provides the pseudocode of the function which detects the standard color board and extract different features such as DGCI value for green and yellow disks and the leaf picture (Chapter 3). Comments are presented subsequently for following the overall algorithm and better understanding of the processes.

---

```
/*
 *          Function to detect the leaf, green and yellow disk
 *          Developed by H. Oveis & C. Igathinathane; ABEN, NDSU
 */
*****/
;
*****/
// input: im the image of the leaf, pink board and background
// output: pic (leaf image)
//          DGCIg (DGCI value of the green disk)
//          DGCIy (DGCI value of the yellow disk)
function [pic,DGCIg,DGCIy]=Leaf(im)
    // detecting the pink board
    r = im(:,:,1);g = im(:,:,2);b = im(:,:,3);
    //create a black background to fill it later with the pink board mask
    bwPink = zeros(size(g));
    //using proper threshold on green and red to extract the board
    bwPink(g<100 & r>100) = 1;

    bwPink = bwareaopen(bwPink,500000); //removing small objects
    bwPink = imfill(bwPink,'holes');
    // applying open morphological operation
    bwPink = imopen(bwPink,strel('square',20));
    //extract the bounding box of the board to crop the board
    Prop = regionprops(bwPink,'BoundingBox');
    r(bwPink==0) = 0;g(bwPink==0) = 0;b(bwPink==0) = 0;
    im = cat(3,r,g,b);
    //cropping the board by the extracted bounding box
    im = imcrop(im,Prop.BoundingBox);
```

```

/***** Extracting Features from Cropped Image *****/
//Here im is a cropped image of the original one.
r = im(:,:,1);g = im(:,:,2);b = im(:,:,3);
hsv = rgb2hsv(im); //converting the im to the HSV color space
h = hsv(:,:,1); s = hsv(:,:,2); v = hsv(:,:,3);
DGCI = (((360*h-60)/60)+(1-s)+(1-v))/3;
DGCIg = DGCI; // creage a copy of DGCI to extract green disk
DGCIg(DGCIg>0.9) = 0; // threshold to remove pink color
/***** Detecting Color Disks *****/

//Detecting Green Disk
bwGreen = im2bw(DGCIg,0.5);
bwGreen = bwareaopen(bwGreen,5000);
DGCIg(bwGreen==0) = 0;
DGCIg = mean(nonzeros(DGCIg));

//Detecting Yellow Disk
bwYellow=zeros(size(s));
bwYellow(s>0.9)=1;
bwYellow=imopen(bwYellow,strel('disk',25));
bwYellow=bwareaopen(bwYellow,5000);
DGCIy=DGCI;
DGCIy(bwYellow==0)=0;
DGCIy=mean(nonzeros(DGCIy));

/***** Extracting the Leaf *****/

rLeaf = r;
rLeaf(s<0.2 | h>0.4)=0;
bwLeaf=im2bw(rLeaf,graythresh(rLeaf));
bwLeaf=imfill(bwLeaf,'holes');
bwLeaf=imerode(bwLeaf,strel('disk',3));
bwLeaf=bwareaopen(bwLeaf,4000);
StatLeaf=regionprops(bwLeaf,'BoundingBox');
r(bwLeaf==0 | h>0.6)=0;g(bwLeaf==0 | h>0.6)=0;b(bwLeaf==0 | h>0.6)=0;
pic=cat(3,r,g,b);
pic=imcrop(pic,StatLeaf.BoundingBox);

```

---

## **APPENDIX B. FUNCTIONS FOR NESTED CROSS VALIDATION AND RANDOMIZED GRID SEARCH**

This appendix provides a brief explanation on nested cross validation along with a visual example (Chapter 3). Following the example, the code that was used in Python can be found. Nested cross validation was used in this study to assess the performance of the machine learning models that have parameters and to find the best parameters for those models. An example of nested cross validation is illustrated in figure B1. In this study, a randomized grid search, which chooses random combinations of parameter values, was performed as the regular grid search trains the models with all possible combinations of the defined parameters that are computationally intensive. Since the study dataset was not very large, the number of combinations was set to 50.

To perform the nested cross-validation, the first dataset was split into 10 folds (external CV). One fold was held-out as a test set and other nine were combined into one dataset and passed to randomized grid search. In the next step, randomized grid search again split the receiving dataset into 10 folds (inner CV). Again, one fold was left out as a validation set and other 9 were combined as a training set to train the models that were built with all 50 random combinations. Random grid search then tests all 50 models with the validation set and records the test scores. This process will repeat 9 more times by keeping the next fold out as a validation set every time.



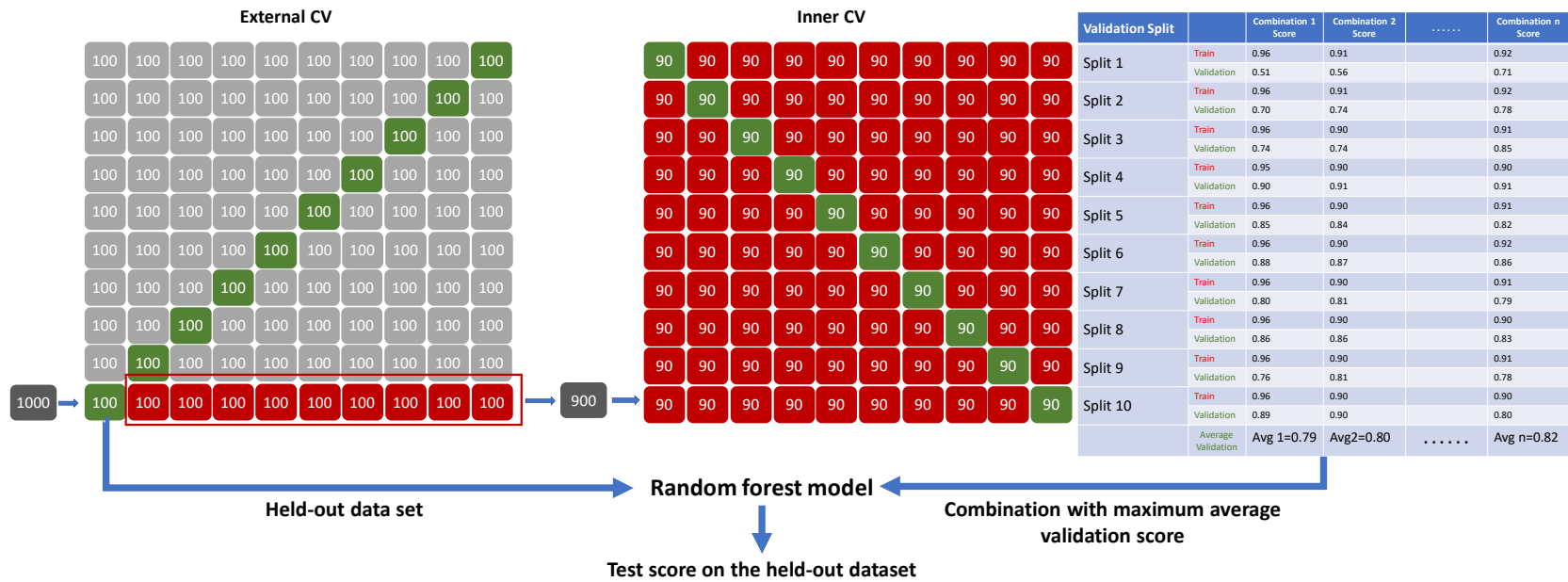


Figure B1. An illustration of nested cross-validation applicable to SVM and RF models. This illustration uses a dataset with 1000 observations showing one external CV and must be repeated for 9 times to complete the cross-validation.

For all the 10 folds, the randomized grid search calculates the average test score for all 50 models and passes the best model with the highest average test score to the external CV. Then this model was trained again with the initial dataset that was passed into the inner CV. The new model predicted values for the held-out dataset to assess its performance on the unseen data. The whole process is repeated 9 more times by keeping another fold as held-out dataset every time.

---

```

/*****
*           Function to to randomized the parameter search for nested cross
*           validation
*           Developed by H. Oveis & C. Igathinathane; ABEN, NDSU
*****/
;
/*****/
//importing the necessary modules into the script
import numpy as np
from sklearn import model_selection
import pandas as pd
from sklearn.metrics import mean_squared_error
from sklearn.utils import shuffle

//input: X_train (the folds of data from features that should be used to train the
          models)
//  X_test  (the fold of data from features that should be used to test the models)
//  y_train (the folds of data from target feature that should be used to train the
          models)
//  y_test  (the fold of data from target feature that should be used to test the
          models)
//  reg     (the type of machine learning model)
//  parameters (a dictionary of parameters of the model with its corresponding
          values)
//  scoring (the value that nested cv uses to assess the performance of each model
          withing inner and outer cv. Typically  $R^2$ )
//  inner_split (number of folds within inner cv, default value is 10)
//  n_iter     (number of samples from parameters dictionary, default value is 2)

//output: cv_results (the cross validation parameters of inner cv)
//  scores (the score for each fold of inner and outer cv)

```

```

def randomized_search(X_train, X_test, y_train, y_test, reg, parameters, scoring,
                    inner_split=10, n_iter=2):

    scores=[]
    regName=str(type(reg)).split('.')[1][:-2]

    randomSearch=model_selection.RandomizedSearchCV(reg,
        param_distributions=parameters, scoring=scoring, cv=inner_split,
        n_iter=n_iter, random_state=42, return_train_score=True)
    randomSearch.fit(X_train, y_train)
    print(randomSearch.best_params_)
    cv_results = pd.DataFrame(randomSearch.cv_results_)

    best_reg=randomSearch.best_estimator_
    best_reg.fit(X_train, y_train)
    y_pred=best_reg.predict(X_test)

    print('*'*20,regName,'*'*20)

    print("R Square for training is {}".format(best_reg.score(X_train,y_train)))
    scores.append(best_reg.score(X_train,y_train))
    print("R Square for test is {}".format(best_reg.score(X_test,y_test)))
    scores.append(best_reg.score(X_test,y_test))
    print("mean square error is
        {}".format(np.sqrt(mean_squared_error(y_test,y_pred))))
    scores.append(np.sqrt(mean_squared_error(y_test,y_pred)))
    print('*****')

    return (cv_results, scores)

/*****
*           Nested cross validation function
*           Developed by H. Oveis & C. Igathinathane; ABEN, NDSU
*****/
;
/*****/
// This function deals with both machine learning model types with/out parameters
//input: X (input features dataset)
// y (target featurer dataset)
// reg (machine learning model)
// inner_split
// outer_split
// n_iter (number of random samples from parameters dictionary that will be used
a input for randomized_search() function)
// scoring

```

```

//output: cv_result_list
//  dgci_coef (list of coefficient for linear regression models)
//  external_cv (list of best parameters selected for outer cv)
//  scores_list (list of scores for both inner and outer cv)

def nested_cross_validation(X, y, reg, inner_split, outter_split, n_iter=10,
    parameters=None ,scoring='neg_mean_squared_error'):

    X, y = shuffle(X,y, random_state=13)

    outter=model_selection.KFold(n_splits=outter_split, shuffle=False,
        random_state=53)

    dgci_coef=[]
    external_cv=[]
    cv_result_list=[]
    scores_list=[]

    for train_index_outer, test_index_outer in outter.split(X):
        X_train, X_test = X[train_index_outer], X[test_index_outer]
        y_train, y_test = y[train_index_outer], y[test_index_outer]
        #X_List2.append(X_test) # to see if the loop produce the same results
            everytime. Just for test

        if str(type(reg)).split('.')[1][:-2] in ['RandomForestRegressor', 'SVR']:

            cv_results, scores = randomized_search(X_train, X_test, y_train, y_test,
                reg, parameters, scoring, inner_split=inner_split, n_iter=n_iter)
            cv_result_list.append(cv_results)
            scores_list.append(scores)

        else:

            scores=[]

            reg.fit(X_train, y_train)
            lin_coef=list(reg.coef_)
            lin_intercept=reg.intercept_
            lin_coef.append(lin_intercept)
            train_score=reg.score(X_train,y_train)
            lin_coef.append(train_score)
            dgci_coef.append(lin_coef)

            y_pred=reg.predict(X_test)
            print("R Square for training is {}".format(reg.score(X_train,y_train)))
            print("R Square for test is {}".format(reg.score(X_test,y_test)))

```

```
print("mean square error is
      {}".format(np.sqrt(mean_squared_error(y_test,y_pred))))

scores.append(reg.score(X_train,y_train))
scores.append(reg.score(X_test,y_test))
scores.append(np.sqrt(mean_squared_error(y_test,y_pred)))
print('*****')
external_cv.append(scores)

return cv_result_list, dgci_coef, external_cv,scores_list
```

---

## APPENDIX C. FUNCTION TO DETECT TWO MIDDLE ROWS

This appendix provides the pseudocode of the function which detects the middle two rows of soybean plots (Chapter 4). Comments are presented subsequently for following the overall algorithm and better understanding of the processes.

---

```

/*****
                Function to detect Two Middle Rows
                Developed by H. Oveis & C. Igathinathane; ABEN, NDSU
*****/
// input: I the image of a 4-row plot
// output: Rows the mask for the two middle rows
function [Rows] = Row_Detection(I)
    // The images were rotated, so rows are close to horizontal
    I = imrotate(I, 90);
/*****/
    //Detecting the standard color board
    //converting the image to L*a*b* color space
    lab = rgb2lab(I);
    a = lab(:,:,2);
    //Since the original values for a* are between -128,127, the values were increased
        by 127. Therefore, all the values are positive integers
    a = uint8(a+127);
    board = a;
    //turning all the pixels off except standard board pixels which have new a* values
        greater than 150
    board(a<150) = 0;
    //converting the image to binary
    board = im2bw(board);
    board = imfill(board, 'holes');
    //removing objects small than 10000 pixels
    board = bwareaopen(board,10000);
    //storing the centroid of the standard board
    statPink = regionprops(board, 'Centroid');
/*****/
    //Removing soil and detecting plots
    bg = a; //assigning a* to background (bg)
    //turing the board pixels to the average values of the a*, so it does not
        interfere with Otsu method for fg/bg separation

```

```

bg(board==1) = mean(mean(a));
//converting the image to binary and inverting it at the same time to extract the
    soybean plots
bg = ~im2bw(bg, graythresh(bg));
//using morphological operation in the next 3 lines to clean the small speckles
bg = imerode(bg, strel('disk', 3));
bg = imdilate(bg, strel('disk', 30));
indBW = bwareaopen(bg, 40000);
statRows = regionprops(indBW, 'Centroid', 'PixelIdxList', 'Orientation');
/*****
section = 25;                //width of each section
for i = 1000:section:2000
    indbw = indBW(:, i:section+i); //creating a section of image
    indbw = imdilate(indbw, strel('disk', 120));
    stat = regionprops(indbw, 'Centroid');
    statSize = size(stat, 1);
    if statSize >= 2
        break
    end
end
findCenter = zeros(1, statSize);

for j = 1:statSize
    findCenter(j) = stat(j).Centroid(2) - statPink(1).Centroid(2);
end
//Finding the minimum of the centroids and switching the values with 999999
    because the loop goes over the findCenter again.
for mid = 1:2
    [~, id] = min(abs(findCenter));
    nStat(mid) = stat(id);
    nStat(mid).Centroid(1) = nStat(mid).Centroid(1) + i;
    findCenter(id) = 99999;
end

/*****
section = 25;                //width of each section
for i = 2500:-1*section:1500
//Plots are almost horizontal because we rotated the image.
    indbw = indBW(:, i-section:i); //creating a section of image
    indbw = imdilate(indbw, strel('disk', 120));
    stat = regionprops(indbw, 'Centroid');
    statSize = size(stat, 1);
    if statSize >= 2
        break
    end
end
end

```

```

findCenter = zeros(1,statSize);
for j = 1:statSize
    findCenter(j) = stat(j).Centroid(2)-statPink(1).Centroid(2);
end
for mid = 3:4
    [~,id] = min(abs(findCenter));
    nStat(mid) = stat(id);
    nStat(mid).Centroid(1) = nStat(mid).Centroid(1)+i-75;
    findCenter(id) = 99999;
end
/*****/
//Altering Matlab because sometimes it put the greater Yvalue of two points in 2nd
place
//nStat(5) is only a placeholder for swaping the values
if nStat(3).Centroid(2) < nStat(4).Centroid(2)
    nStat(5).Centroid(2) = nStat(3).Centroid(2);
    nStat(3).Centroid(2) = nStat(4).Centroid(2);
    nStat(4).Centroid(2) = nStat(5).Centroid(2);
end
if nStat(1).Centroid(2) < nStat(2).Centroid(2)
    nStat(5).Centroid(2) = nStat(1).Centroid(2);
    nStat(1).Centroid(2) = nStat(2).Centroid(2);
    nStat(2).Centroid(2) = nStat(5).Centroid(2);
end
/*****/
//Drawing two centerlines and finding slope of two lines
m1 = (nStat(3).Centroid(2)-nStat(1).Centroid(2))/
    (nStat(3).Centroid(1)-nStat(1).Centroid(1));
m2 = (nStat(4).Centroid(2)-nStat(2).Centroid(2))/
    (nStat(4).Centroid(1)-nStat(2).Centroid(1));
/*****/
//Detecting rows
Rows = false(size(indBW));
for obj = 1:size(statRows,1)
    //y2 = m(x2-x1)+y1
    //Calculating the y of centerpoints of all possible plot rows.
    Yline = round(m1*(statRows(obj).Centroid(1)-nStat(3).Centroid(1))+
        nStat(3).Centroid(2));
    //if the distance of calculated y of centerpoint was less than 200 pixel away
    from its actual y, then it is one of the middle rows.
    if abs(Yline-statRows(obj).Centroid(2)) < 200
        Rows(statRows(obj).PixelIdxList) = true;
    end
end
end
for obj = 1:size(statRows,1)

```



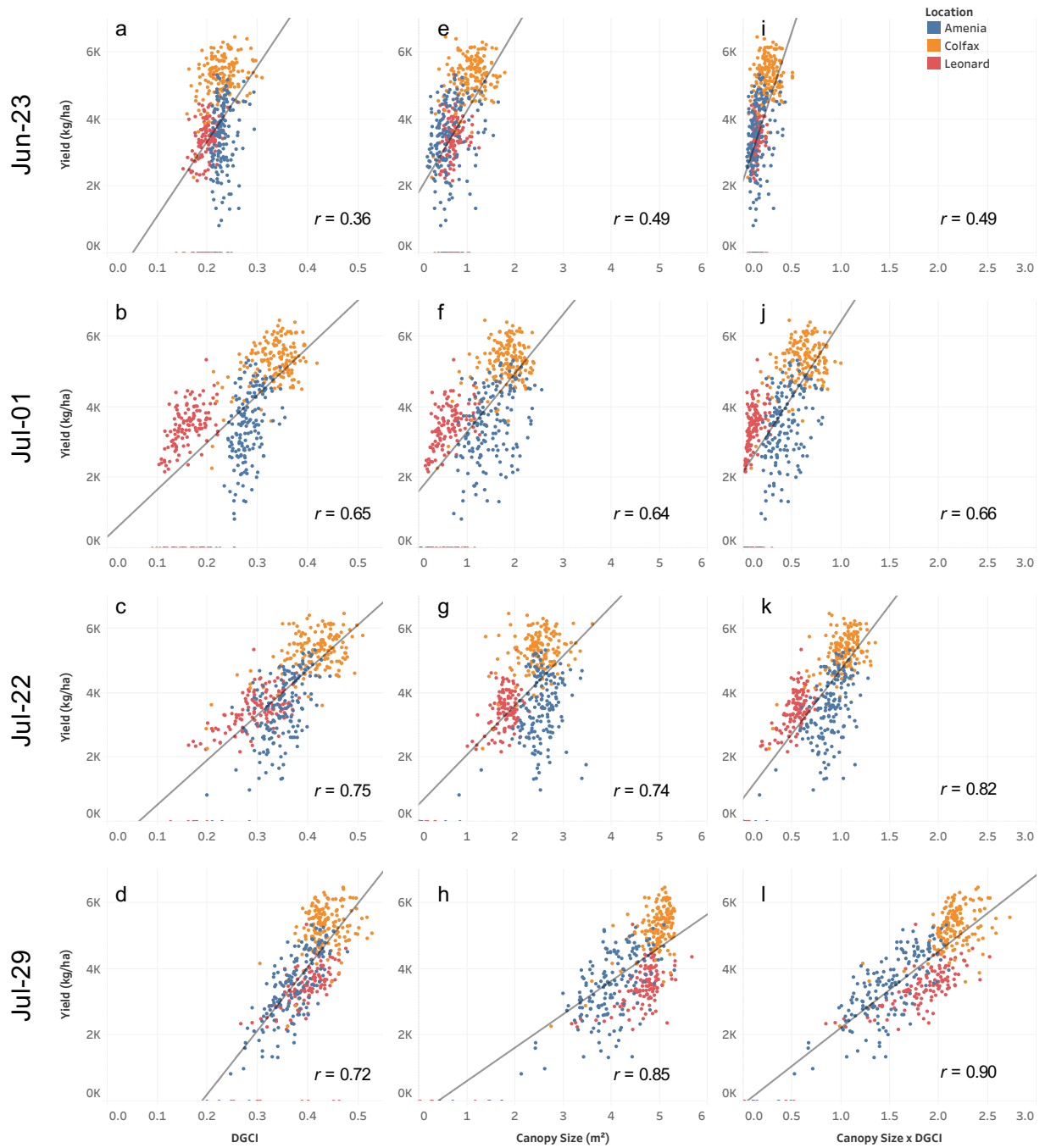
```
Yline = round(m2*(statRows(obj).Centroid(1)-nStat(4).Centroid(1))+
    nStat(4).Centroid(2));
if abs(Yline-statRows(obj).Centroid(2)) < 200
    Rows(statRows(obj).PixelIdxList) = true;
end
end

stat = regionprops(Rows, 'Orientation', 'PixelIdxList');
for z = 1:size(stat,1)
    if abs(stat(z).Orientation) > 65
        Rows(stat(z).PixelIdxList) = false;
    end
end
end
```

---

## **APPENDIX D. AERIAL IMAGE FEATURES RELATIONSHIP WITH YIELD FOR DIFFERENT DATES**

This appendix provides illustrations on the relationship of different features from aerial images (Chapter 5) with yield for 2016 (fig. D1) and 2017 (fig. D2). These features include dark green color index (DGCI), canopy size, and canopy size  $\times$  DGCI product.



**Figure D1.** An illustration of relationship of different features from aerial images with yield for 2016. (Left) DGCI, (Middle) Canopy size, (Right) DGCI × canopy size.

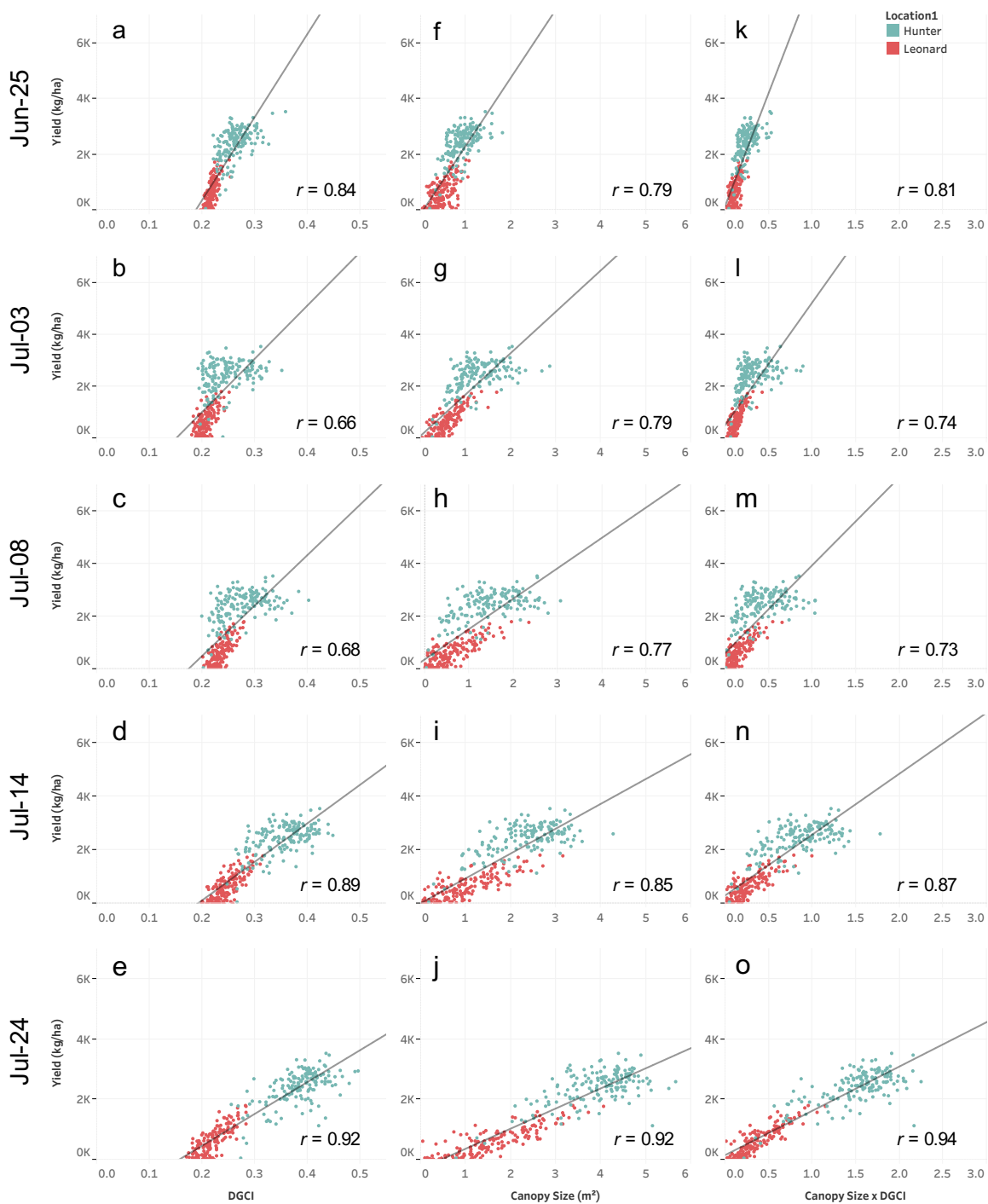


Figure D2. An illustration of relationship of different features from aerial images with yield for 2017. (Left) DGCI, (Middle) Canopy size, (Right) DGCI  $\times$  canopy size.

# APPENDIX E. SCRIPTS FOR DATA VISUALIZATION THROUGH R AND PYTHON

This appendix provides scripts for data visualization in R through ggplot2 (<https://ggplot2-book.org>) and in Plotly through Python (<https://plotly.com/python/>). The original graphs derived from chapters 3, 4, and 5 are reproduced again for ready reference at the end of each code sections.

---

```

/*****
  Script to create bar graphs to compare the performance of different
  machine learning algorithms in ggplot (R)

  Developed by Oveis Hassanijalilian and C. Igathinathane; ABEN, NDSU
*****/
---
title: "ggplot2 model drawing"
author: "Oveis and C. Igathi"
date: "2/23/2019"
output: pdf_document
---

““{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
x = 3
““

#setwd("C:/Users/Oveis/Desktop/dgci_spad/model_scores/AllModels/R_csvFiles")

Models R^2 with color scheme inputs

““{r}
library(ggplot2)
library(scales)
df3 <- read.csv('aggregated_Rsqr.csv')
df3
```

```

df3$models <- factor(df3$models,levels = c("Simple linear regression*",
                                           "Polynomial regression*",
                                           "Multiple linear regression",
                                           "Support vector machine",
                                           "Random forest")) # reorder the levels
df3$Inputs <- factor(df3$Inputs,levels = c("RGB", "DGC", "Both")) # reorder the
  levels
df3$Method <- factor(df3$Method,levels = c("Train", "Test")) # reorder the levels
df3$label <- factor(df3$label)

p <- ggplot(data=df3, aes(x=Inputs, y=Rsqr, fill=Method)) +
  geom_errorbar(aes(ymin = Rsqr - error, ymax = Rsqr + error),
               position = position_dodge2(width = 0.5, padding = 0.5)) +
  geom_bar(stat="identity", position=position_dodge()) +
  scale_fill_manual(values=c('darkgreen', 'orange')) +
  scale_y_continuous(limits = c(0.3, 1.0), oob = rescale_none) +
  facet_grid(. ~ models) +
  geom_col(position = "dodge2") +
  geom_text(aes(label=sprintf("%.3f", round(Rsqr,3))), vjust=0.5, hjust=1.3,
           color="white",
           position = position_dodge(0.9), size=3.5, angle = 90) +
  geom_text(aes(label=label, y=Rsqr + error), vjust=-0.8, hjust=0.5, position =
           position_dodge(0.9), size=3.5) +
  theme_bw() +
  theme(strip.text.x = element_text(size = 9, face = "bold")) # order is important

p + theme(legend.position = c(0.04, 0.9),
         legend.text=element_text(size=6),
         legend.title=element_text(size=7),
         legend.key.size = unit(0.175, "in")) +
  labs(x = "Color scheme input", y = expression("Model's" ~ R^2))

ggsave("modelR2.pdf", width = 9, height = 4.5, units = "in")
'''

```

RMSE with color scheme inputs

```

'''{r}
library(ggplot2)
library(scales)
df1<- read.csv('aggregated_RMSE.csv')
df1

df1$models <- factor(df1$models,levels = c("Simple linear regression*",
                                           "Polynomial regression*",

```

```

                                "Multiple linear regression",
                                "Support vector machine",
                                "Random forest")) # reorder the levels
df1$Inputs <- factor(df1$Inputs,levels = c("RGB", "DGCI", "Both")) # reorder the
  levels
df1$label <- factor(df1$label)

ggplot(data=df1, aes(x=Inputs, y=RMSE)) +
  geom_errorbar(aes(ymin = RMSE - error, ymax = RMSE + error), width=0.4) +
  geom_bar(stat="identity", fill="darkgreen", width = 0.5) +
  facet_grid(. ~ models) + # should follow geom_bar
  scale_y_continuous(limits = c(1, 7.0), oob = rescale_none) +
  geom_text(aes(label=sprintf("%.3f", round(RMSE,3))), vjust=0.5, hjust=1.3,
    color="white",
    position = position_dodge(0.9), size=3.5, angle = 90) +
  geom_text(aes(label=label, y=RMSE + error), vjust=-0.8, hjust=0.5, position =
    position_dodge(0.9), size=3.5) +
  labs(x = "Color scheme input", y = "RMSE") +
  theme_bw() +
  theme(strip.text.x = element_text(size = 9, face = "bold"))

ggsave("modelRMSE.pdf", width = 9, height = 4.5, units = "in")

```

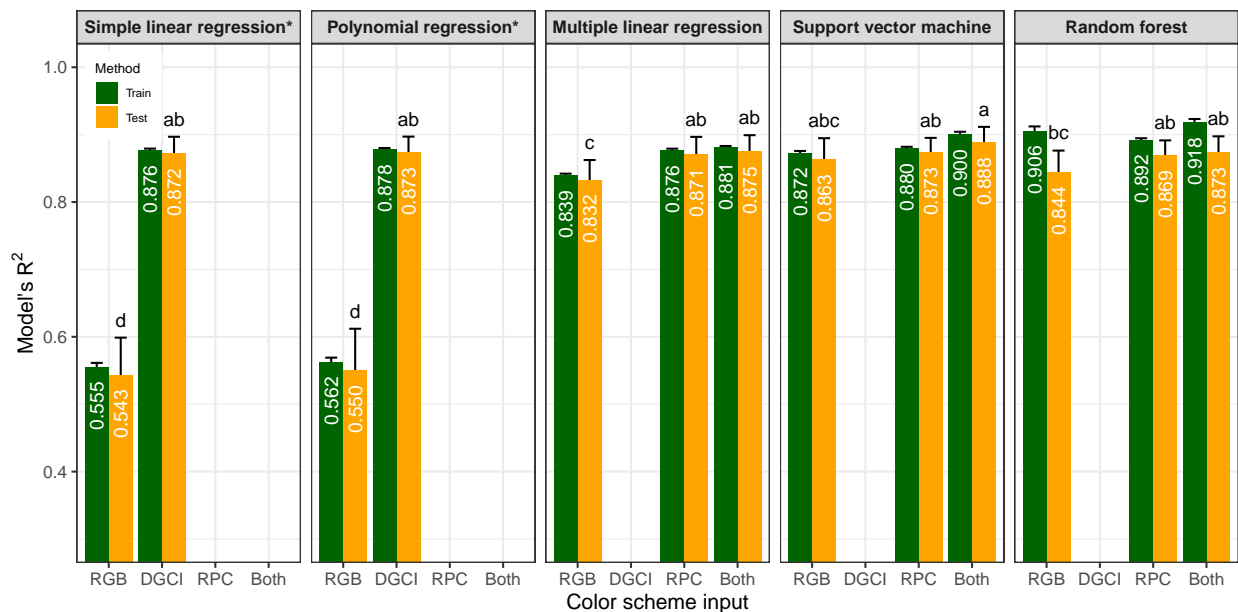
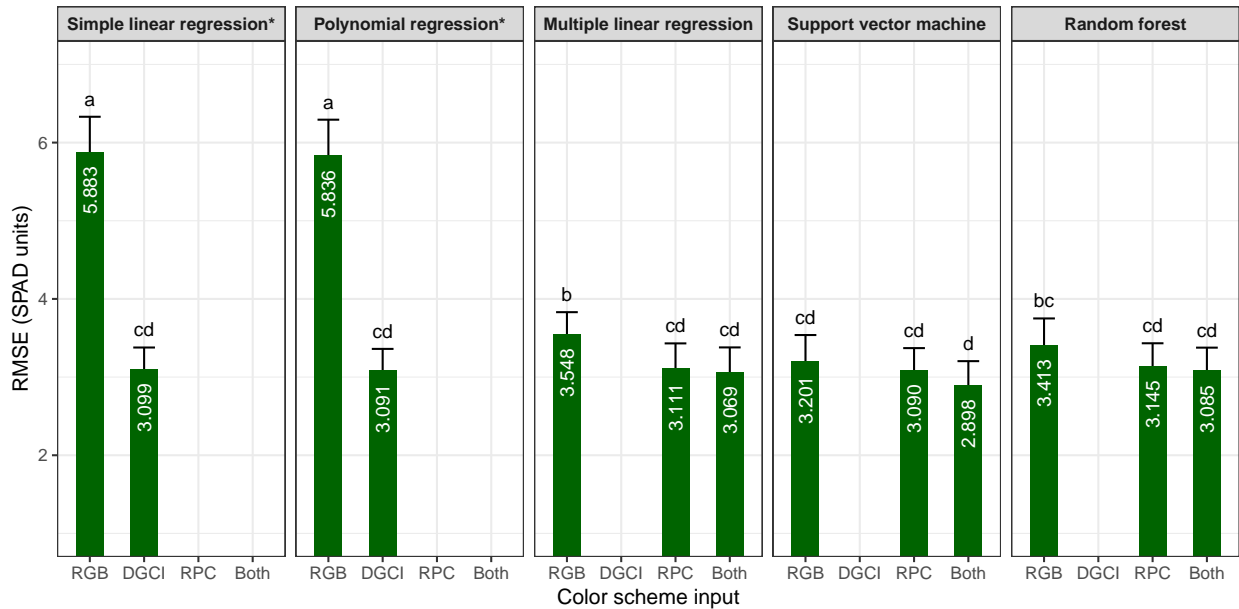


Figure E1. Bar graph of  $R^2$  with error bars produced by ggplot2 in R to compare performance of different machine learning algorithms.



**Figure E2. Bar graph of RMSE with error bars produced by ggplot2 in R to compare performance of different machine learning algorithms.**

```

/*****
Script to create bar graphs to compare correlation of UAV image parameters in
ggplot (R)

```

Developed by Oveis Hassanijalilian and C. Igathinathane; ABEN, NDSU

```

*****/
---
```

```

# Visualization of correlation of UAV image parameters
# with soybean yield - R ggplot2 code
```

```

dat <- read.table(header = T, text = "
year day parameter rvalue
2016 Jun-23 DGCI 0.36
2016 Jul-01 DGCI 0.65
2016 Jul-22 DGCI 0.75
2016 Jul-29 DGCI 0.72
2017 Jun-25 DGCI 0.84
2017 Jul-03 DGCI 0.66
2017 Jul-08 DGCI 0.68
2017 Jul-14 DGCI 0.89
2017 Jul-24 DGCI 0.92

```



```

2016 Jun-23 CS 0.49
2016 Jul-01 CS 0.64
2016 Jul-22 CS 0.74
2016 Jul-29 CS 0.85
2017 Jun-25 CS 0.79
2017 Jul-03 CS 0.79
2017 Jul-08 CS 0.77
2017 Jul-14 CS 0.85
2017 Jul-24 CS 0.92
2016 Jun-23 CDP 0.49
2016 Jul-01 CDP 0.66
2016 Jul-22 CDP 0.82
2016 Jul-29 CDP 0.90
2017 Jun-25 CDP 0.81
2017 Jul-03 CDP 0.74
2017 Jul-08 CDP 0.73
2017 Jul-14 CDP 0.87
2017 Jul-24 CDP 0.94
")

```

```

library(ggplot2)
library(RColorBrewer)

dat$year <- factor(dat$year, levels = c("2016", "2017"))
dat$day <- factor(dat$day, levels =
  c("Jun-23", "Jul-01", "Jul-22", "Jul-29", "Jun-25", "Jul-03", "Jul-08", "Jul-14",
    "Jul-24"))
dat$parameter <- factor(dat$parameter, levels = c("DGCI", "CS", "CDP"))
dat$rvalue <- as.numeric(as.character(dat$rvalue))

plt <- ggplot(dat, aes(x = day, y = rvalue, fill = parameter)) +
  geom_bar(width = 0.6, colour="black", stat="identity", position="dodge")+
  theme_bw() +
  scale_fill_brewer(palette="Set1") +
  theme(legend.position= c(.08, .8), legend.title = element_blank()+
  facet_wrap(year~., scales = "free") +
  scale_y_continuous(breaks=seq(0.0,1.0,0.1)) +
  theme(strip.text.x = element_text(size = 12, face = "bold")) +
  xlab("Days") +
  ylab(expression(italic("r ")~value))

```

```
plt
```

---

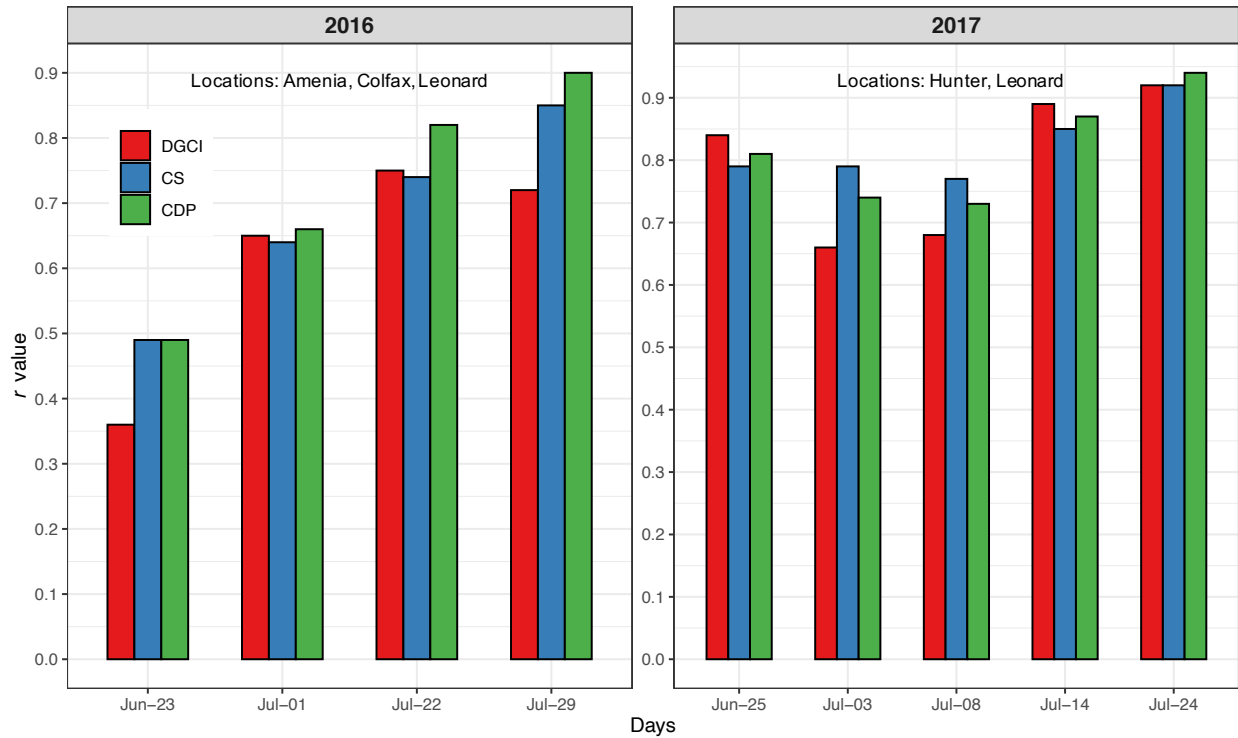


Figure E3. Bar graphs with with facets showing three independent variables produced by ggplot2 in R to compare the correlation between parameters and yield.

```

/*****
  Script to create bar graphs to show frequency of different classes
  in Plotly (Python)

  Developed by Oveis Hassanijalilian; ABEN, NDSU
  *****/

```

```

x = sorted(Diff.diff_rate.unique())
y = Diff.diff_rate.value_counts().sort_index()
y = [i for i in y]

colors = ['#E09C00', '#FFD700', '#DDE800', '#BBD900', '#97FF00', '#5BCF00', '#449A00']

fig = go.Figure(data=[go.Bar(
    x=x, y=y,
    text=y,
    marker_color=colors
)])

```

```

fig.update_layout(
    template="plotly_white",
    autosize=True,
    width=700,
    height=500,
    yaxis=dict(
        title='Frequency'
    ),
    xaxis=dict(
        title='Differences between IDC rating for both dates'
    ),
    xaxis_showgrid=False, yaxis_showgrid=False
)
fig.update_yaxes(showline=True, linewidth=1,
    linecolor='black', ticks='outside', mirror=True)
fig.update_xaxes(showline=True, linewidth=1,
    linecolor='black', ticks='outside', mirror=True)

fig.show()

```

```

/*****
                Script to create a histogram
                in Plotly (Python)

```

Developed by Oveis Hassanijalilian; ABEN, NDSU

```

*****/

```

```

colors= ['#c5d627', '#b3d627', '#9fd627', '#90d627',
        '#7fd627', '#6ad627', '#56d627', '#47d627',
        '#43bd28', '#3fab27', '#3b9c25', '#388c24']
fig = go.Figure(data=[go.Histogram(x=Diff.diff_dgci, marker_color=colors)])
fig.update_layout(
    template="plotly_white",
    autosize=True,
    width=600,
    height=500,
    yaxis=dict(
        title='Frequency'
    ),
    xaxis=dict(
        title='Differences between average DGCI of both dates'
    ),
)

```

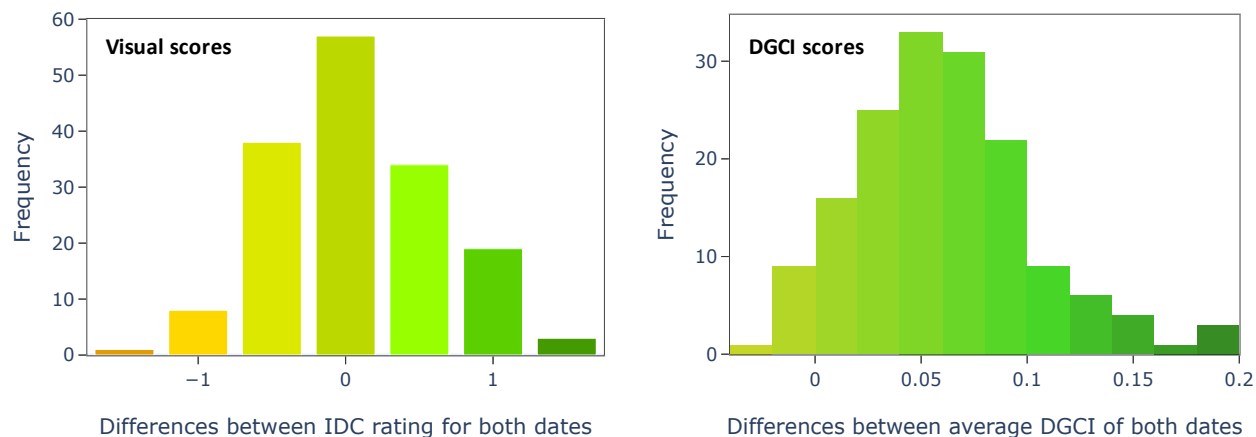
```

axis_showgrid=False, yaxis_showgrid=False)

fig.update_yaxes(showline=True, linewidth=1,
                 linecolor='black', ticks='outside', mirror=True)
fig.update_xaxes(showline=True, linewidth=1,
                 linecolor='black', ticks='outside', mirror=True)

fig.show()

```



**Figure E4. Bar graph on the left and histogram on the right produced by Plotly in Python.**

```

/*****
    Script to create bar graphs to compare the performance of different
    machine learning algorithms in Plotly (Python)

    Developed by Oveis Hassanijalilian; ABEN, NDSU
    *****/
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df = pd.read_csv('./Performance.csv')

classes=['Class=1', 'Class=2', 'Class=3', 'Class=4']

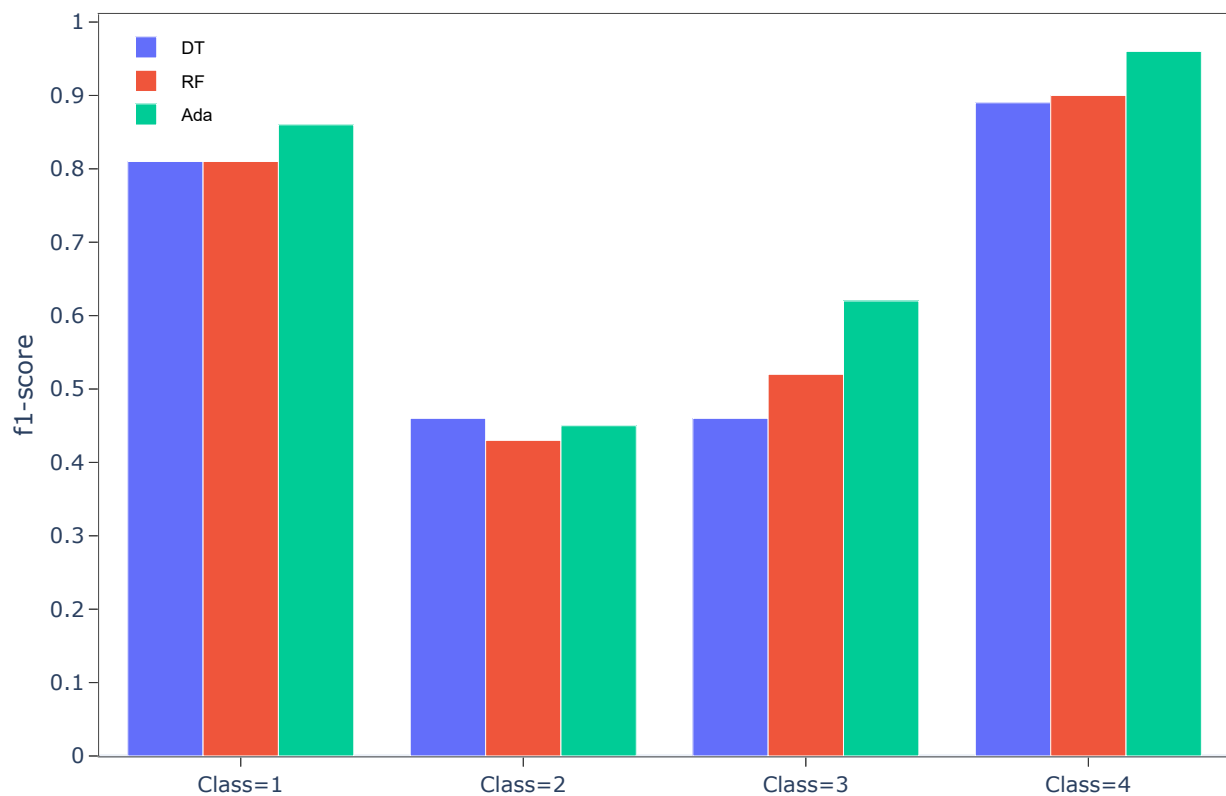
fig = go.Figure(data=[
    go.Bar(name='DT', x=classes, y=df2.loc[df2.Model=='DT', 'Value']),
    go.Bar(name='RF', x=classes, y=df2.loc[df2.Model=='RF', 'Value']),
    go.Bar(name='Ada', x=classes, y=df2.loc[df2.Model=='Ada', 'Value']),
])

```

```
# Change the bar mode
fig.update_layout(template="plotly_white",
                  barmode='group',width=700,yaxis=dict(title='f1-score',tickmode =
                  'linear',tick0 = 0.00,dtick = 0.1),
                  axis_showgrid=False, yaxis_showgrid=False)

fig.update_yaxes(showline=True, linewidth=1,
                 linecolor='black',ticks='outside',mirror=True)
fig.update_xaxes(showline=True, linewidth=1,
                 linecolor='black',ticks='outside',mirror=True)

fig.show()
```



**Figure E5. Bar graph produced by Plotly in Python to compare classification metric among different machine learning algorithms.**