

Implementing sensor technology to evaluate genetic and spatial variability within the Kansas  
State University Wheat Breeding program

by

Byron J Evers

B.S., University of Wisconsin-Stevens Point, 2006  
M.S., Kansas State University, 2012

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy  
College of Agriculture

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

2022

## **Abstract**

Globally wheat is one of the three most important cereal crops globally providing 20% of protein and total calories consumed. In the world as well as the state of Kansas, wheat is planted on more acres than any other crop. Additionally, wheat sales generated \$1.27 billion in revenue in 2021 making wheat an economic driver for the entire state. However, the annual genetic gain in wheat is 0.8-1.2% and is not sufficient to support the increasing global population. Therefore, the adoption of new technology and computational methods are critical to increase genetic gain and increase wheat adaptability both globally and in the Central Plains.

Proper temporal resolution is critical for quality HTP sensor data collection, as collection at key physiological growing points can increase yield prediction and assist with phenotypic selection. However, growth stages are dependent on weather and fluctuate both across locations and years. This makes day of year or day after sowing a poor phenology metric, particularly with winter wheat where the vernalization requirement compounds phenology prediction challenges and significantly shifts developmental stages relative to calendar days. This study was designed to assess the performance of various phenology models to predict heading time of both historically adapted and experimental genotypes of wheat genotypes in Kansas. The results suggest that full season models with multi-phase coefficients can increase phenology prediction over traditional thermal indices. However, using cumulative thermal times after the vernalization requirements also provided phenology predictions that were statistically similar to the full season phase change models.

Genotype by environment interactions is a prominent issue for breeding programs, particularly when performance testing elite lines across multiple locations and years. In addition to macroenvironments, variations in soil properties have shown to develop microenvironments within location years. These soil microenvironments can potentially be quantified through both traditional and precision agriculture tools. Whereas, traditional soil sampling density is limited by cost and time, precision agriculture on-the-go soil sensors have the potential to gather large quantities of data. However, these measurements are often giving only relative measurements. Through this experiment two sensor platforms were evaluated as potential tools to quantify spatial variability within breeding programs. This study showed that soil spatial variability does

impact genotype yield performance and that indirect measurements from both sensor platforms can quantify this impact.

The continued development of high quality, cost effective multi-spectral imaging devices has led to numerous studies to evaluate this technologies ability to predict traits and grain yield. Despite these advancements the widespread implementation of these tools for selection has been slow and most breeders still rely on harvested grain yield and visual selection for cultivar advancement. The intention of this experiment was to evaluate high spatial resolution data from, multi-spectral sensors at multi-temporal collection points to make yield group rank order selections. Additionally, a random forest algorithm was used to evaluate the potential of incorporating machine learning with HTP data as a selection tool. Although the rank order correlations were higher than the correlation to grain yield, the selection accuracies of random forest were not statistically better than the no-information rate. However, this study does lay the groundwork for future similar studies using alternative sensor aided metrics and machine learning algorithms.

Overall, the combined results of these studies show that these precision agriculture tools have to potential to increase genetic gain in plant breeding. However, these studies also show that both sensor and computational limitations still exist. Moving forward it is pivotal that future studies focus on technology combinations that have the potential to easily be implemented within a breeding program.

Implementing sensor technology to evaluate genetic and spatial variability within the Kansas  
State University Wheat Breeding program

by

Byron J Evers

B.S., University of Wisconsin-Stevens Point, 2006  
M.S., Kansas State University, 2012

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy  
College of Agriculture

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

2022

Approved by:

Co-Major  
Jesse A. Poland

Approved by:

Co-Major Professor  
Allan K. Fritz

# Copyright

© Byron Evers 2022.

## **Abstract**

Globally wheat is one of the three most important cereal crops globally providing 20% of protein and total calories consumed. In the world as well as the state of Kansas, wheat is planted on more acres than any other crop. Additionally, wheat sales generated \$1.27 billion in revenue in 2021 making wheat an economic driver for the entire state. However, the annual genetic gain in wheat is 0.8-1.2% and is not sufficient to support the increasing global population. Therefore, the adoption of new technology and computational methods are critical to increase genetic gain and increase wheat adaptability both globally and in the Central Plains.

Proper temporal resolution is critical for quality HTP sensor data collection, as collection at key physiological growing points can increase yield prediction and assist with phenotypic selection. However, growth stages are dependent on weather and fluctuate both across locations and years. This makes day of year or day after sowing a poor phenology metric, particularly with winter wheat where the vernalization requirement compounds phenology prediction challenges and significantly shifts developmental stages relative to calendar days. This study was designed to assess the performance of various phenology models to predict heading time of both historically adapted and experimental genotypes of wheat genotypes in Kansas. The results suggest that full season models with multi-phase coefficients can increase phenology prediction over traditional thermal indices. However, using cumulative thermal times after the vernalization requirements also provided phenology predictions that were statistically similar to the full season phase change models.

Genotype by environment interactions is a prominent issue for breeding programs, particularly when performance testing elite lines across multiple locations and years. In addition to macroenvironments, variations in soil properties have shown to develop microenvironments within location years. These soil microenvironments can potentially be quantified through both traditional and precision agriculture tools. Whereas, traditional soil sampling density is limited by cost and time, precision agriculture on-the-go soil sensors have the potential to gather large quantities of data. However, these measurements are often giving only relative measurements. Through this experiment two sensor platforms were evaluated as potential tools to quantify spatial variability within breeding programs. This study showed that soil spatial variability does

impact genotype yield performance and that indirect measurements from both sensor platforms can quantify this impact.

The continued development of high quality, cost effective multi-spectral imaging devices has led to numerous studies to evaluate this technologies ability to predict traits and grain yield. Despite these advancements the widespread implementation of these tools for selection has been slow and most breeders still rely on harvested grain yield and visual selection for cultivar advancement. The intention of this experiment was to evaluate high spatial resolution data from, multi-spectral sensors at multi-temporal collection points to make yield group rank order selections. Additionally, a random forest algorithm was used to evaluate the potential of incorporating machine learning with HTP data as a selection tool. Although the rank order correlations were higher than the correlation to grain yield, the selection accuracies of random forest were not statistically better than the no-information rate. However, this study does lay the groundwork for future similar studies using alternative sensor aided metrics and machine learning algorithms.

Overall, the combined results of these studies show that these precision agriculture tools have to potential to increase genetic gain in plant breeding. However, these studies also show that both sensor and computational limitations still exist. Moving forward it is pivotal that future studies focus on technology combinations that have the potential to easily be implemented within a breeding program.

# Table of Contents

List of Figures .....	x
List of Tables.....	xvi
Acknowledgements .....	xviii
Dedication .....	xix
Chapter 1 - Literature Review .....	1
Global and local importance of wheat .....	1
Current Breeding Technologies .....	2
High Throughput Phenotyping using Sensor Platforms.....	3
Modeling Phenological Development.....	5
Spatial Variability in Crop Breeding.....	8
References .....	10
Chapter 2 - Integrating Phenological Development from Crop Models into High Throughput Phenotyping .....	17
Introduction .....	17
Materials and Methods .....	20
Plant Material .....	20
Phenotypic Data Collection.....	22
Weather Data.....	23
Thermal Indices.....	23
Agricultural Production Systems Simulator (APSIM) Wheat Model.....	27
Statistical Analyses.....	28
Results and Discussions .....	29
Conclusions .....	36
References .....	52
Chapter 3 - Assessment of Spatial Variability within the KS Wheat Breeding Program .....	56
Introduction .....	56
Materials and Methods .....	60
Experiment locations.....	60
Plant Material .....	61

Soil Sensors .....	62
Soil Analysis .....	63
Spatial Analysis.....	64
Statistical Analysis .....	66
Results and Discussions .....	67
Conclusions .....	75
References .....	106
Chapter 4 - Rank order Phenotypic Selection for Wheat Breeding using UAVs .....	112
Introduction .....	112
Materials and Methods .....	115
Experimental Locations.....	115
Plant Material .....	115
Thermal Time Corrections .....	117
Spatial Corrections .....	118
Statistical Analyses.....	120
Random Forest .....	120
Results and Discussions .....	122
Conclusions .....	127
References .....	145
Appendix A - Supplementary Material Chapter 2 .....	150
Appendix B - Supplementary Material Chapter 3.....	153
Appendix C - Supplementary Material Chapter 4.....	165

## List of Figures

Figure 2.1. Experiment and weather station locations for both the AM Panel and Kansas Wheat Breeding experiments. The shapes represent the weather station locations in relationship to the experiment locations. The green polygon represents winter wheat acres planted in 2021 ..... 37

Figure 2.2. Observed heading date distribution by location year for the AM Panel (A) and the Kansas State University Wheat Breeding Program (B). The x-axis is the harvest year growing season and the y-axis is the observed heading date in day of year (DOY) after Jan 1<sup>st</sup>. Colors represent location for the Kansas State University Wheat Breeding Program data. .... 37

Figure 2.3 Mean monthly temperature by location year for the AM Panel (A) and the KS Wheat Breeding Program (B). The x-axis is the month during the growing season and the y-axis is the recorded mean monthly temperature from the Kansas Mesonet weather stations. Symbol and line combinations represent the location and year the data was recorded..... 38

Figure 2.4. Cumulative Growing Degree Days by year and location fit with a logistic regression model for the AM Panel. The x-axis is day of year (DOY) after March 1<sup>st</sup> (A), January 1<sup>st</sup> (B) and days after sowing (DAS)(C). The y-axis is the cumulative thermal time calculated using GDD1. The shapes denote the individual location years and the red line is the fitted trend for all seasons. The fit lines for both figure A and B are logistic regression models and figure C is a polynomial model. .... 39

Figure 2.5. Cumulative Growing Degree Days by year and location fit with a logistic regression model for the Kansas State University Wheat Breeding program. The x-axis is day of year (DOY) after March 1<sup>st</sup> (A), January 1<sup>st</sup> (B) and days after sowing (DAS)(C). The y-axis is the cumulative thermal time calculated using GDD1. The shapes denote the individual location years and the red line is the fitted trend for all seasons. The fit lines for both figure A and B are logistic regression models and figure C is a polynomial model. .... 40

Figure 2.6. Growing Degree Day accumulation at observed heading date by year for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> accumulation periods within the AM Panel. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the cumulative thermal time calculated using GDD1. .... 41

Figure 2.7. Growing Degree Day accumulation at observed heading date by year and location for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> accumulation periods within the Kansas State University Wheat Breeding program. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the cumulative thermal time calculated using GDD1. .... 42

Figure 2.8. Growing Degree Days (GDD1) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods within the AM Panel. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the GDD1 predicted heading date, DOY. .... 43

Figure 2.9. Growing Degree Days (GDD1) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods within the Kansas State University Wheat Breeding program. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the GDD1 predicted heading date, DOY..... 44

Figure 2.10. APSIM Predicted vs Observed Heading Dates for the AM Panel (A) and the Kansas State University Wheat Breeding program (B). The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the APSIM predicted heading date, DOY. Shape differences represent sample year and the line is the linear model fitted trend line. .... 45

Figure 3.1. 21THD Veris P4000 sample points. Points in pink represent the raw GPS positions obtained from the P4000 and the green points represent RTK corrected points. Ariel image was obtained after sampling and grid marks are sampling paths created by the MSP3 sensor. .... 76

Figure 3.2. Distribution of soil values obtained by the Veris P4000 by depth for EC, Force, Red and IR parameters for 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F). The x-axis is the measured value of the parameter, force is in KPa, apparent electrical conductivity (EC<sub>a</sub>) is in  $\mu\text{S cm}^{-1}$ , and reflectance are 16-bit digital values. Each sensor sample depth is represented by color. .... 77

Figure 3.3. Distribution of soil values obtained by the Veris MSP3 for shallow EC (30cm), deep EC (90cm), Red and IR parameters at 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F). The x-axis is the measured value of the parameter, force is in KPa, apparent electrical conductivity (EC<sub>a</sub>) is in  $\mu\text{S cm}^{-1}$ , and reflectance are 16-bit digital

values. The colors in the EC<sub>a</sub> graphs represent sample depth and represent the spectral band for the reflectance graphs. .... 78

Figure 3.4. Distribution of soil values obtained from lab analysis of physical soil cores by depth for VWC, pH, EC, N, P, K, CEC, OM, and BD at 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F). The x-axis is the measured value of each soil parameter, volumetric water content (VWC), organic matter (OM) and nitrogen (N) are precents. Phosphorous (P) and potassium (K) are ppm and bulk density (BD) is g cc<sup>-1</sup>. Each sample depth is represented by color. .... 79

Figure 3.5. Distribution of soil values obtained form lab analysis of physical soil cores by depth for pH, EC, CEC, Ca, and Mg at 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F) The x-axis is the measured value of each soil parameter, electrical conductivity (EC) is in μS cm<sup>-1</sup>, and cation exchange capacity (CEC) is cmol kg<sup>-1</sup>. Calcium (Ca), and magnesium (Mg) are ppm and pH is unitless. Each sample depth is represented by color..... 80

Figure 3.6. Soil texture distribution by location and year. Soil textural categories are in accordance to the USDA textural soil classification system (García-Gaines & Frankenstein, 2015)..... 81

Figure 3.7. K-means cluster contour map for EC, Force, Red and IR measurements from the P4000 at 21RNS with PYN and YT plot map overlays. Number of color clusters were determined by k-means clusters and the values for the cluster color represent the median value of the cluster range..... 82

Figure 3.8. P4000 k-means cluster contour effects of EC, Force (F), Red (R) and IR on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-YT (F) and 21RNS-PYN (G). The title in the gray bar represents the soil measurement and depth. Depth-1 is 0-40 cm Depth 2 is 40-60cm and depth 3 is 60-100cm. The trait abbreviations are electrical conductivity (EC), force (F), red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is μS/cm, force is kPa and both red and near infrared are digital reflectance values. Symbols denote Wilcoxon t-test significance where "\*\*\*\*\*"=0.001, "\*\*\*\*"=0.001, "\*\*\*"=0.01, "\*\*"=0.05 and ns= not significant. .... 85

Figure 3.9. MSP3 k-means cluster contour effects of EC, Force (F), Red (R) and IR on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-YT (F) and 21RNS-

PYN (G). The title in the gray bar represents the soil measurement and depth. ECSH is 0-30 cm and ECDP is 0-90cm. The trait abbreviations are electrical conductivity (EC) , red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S}/\text{cm}$  and both red and near infrared are digital reflectance values. Symbols denote Wilcoxon t-test significance where “\*\*\*\*\*”=0.001, “\*\*\*\*”=0.001, “\*\*\*”=0.01, “\*\*”=0.05 and ns= not significant. .... 89

Figure 3.10. P4000 Kriged soil property correlation with grain yield of EC, Force (F), Red (R) and IR at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-PYN (F) and 21RNS-YT (G). The title in the gray bar represents the soil measurement and depth. Depth-1 is 0-40 cm Depth 2 is 40-60cm and depth 3 is 60-100cm. The trait abbreviations are electrical conductivity (EC), force (F), red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S}/\text{cm}$ , force is kPa and both red and near infrared are digital reflectance values..... 92

Figure 3.11. MSP3 Kriged soil property correlation with grain yield of EC, Force (F), Red (R) and IR on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-PYN (F) and 21RNS-YT (G). The title in the gray bar represents the soil measurement and depth. ECSH is 0-30 cm and ECDP is 0-90cm. The trait abbreviations are electrical conductivity (EC) , red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S}/\text{cm}$  and both red and near infrared are digital reflectance values. .... 95

Figure 3.12. Soil analysis kriged soil property correlation with grain yield of BD, EC, CEC, K, N, OM, P, pH and VWC on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-PYN (F) and 21RNS-YT (G). The title in the gray bar represents the soil measurement and depth. Depth-1 is 0-40 cm and Depth 2 is 40-60cm. The x-axis is the measured value of each soil parameter, volumetric water content (VWC), organic matter (OM) and nitrogen (N) are precents. Phosphorous (P) and potassium (K) are ppm and bulk density (BD) is  $\text{g cc}^{-1}$ . Electrical conductivity (EC) is in  $\mu\text{S cm}^{-1}$ , and cation exchange capacity (CEC) is  $\text{cmol kg}^{-1}$ . Calcium (Ca), and magnesium (Mg) are ppm and pH is unitless. Each sample depth is represented by color. .... 99

Figure 3.13. Coefficient of variation (CV%) by trial for raw data, data adjusted by the MAD-2 method 3, spatially adjusted yields and spatio-statistical adjusted yield. The x-axis

represents CV% and the y-axis represents the location, year and experiment. The colored bars represent the type of spatial correction.....	100
Figure 4.1. Experiment locations for the Kansas Wheat Breeding experiments over a five-year period. Temperature contours are derived from the 2020 observed mean monthly temperatures. The maps shows the north to south temperature gradient for the locations that impacts the cumulative growing degree units used to normalize UAV flight dates. ....	129
Figure 4.2. Schematic diagram of proposed experimental selection using UAV VI's within the KS Wheat Breeding program across 5 years.....	129
Figure 4.3. Mean monthly temperature (A) and mean monthly precipitation (B) by location year across 18 site years. The x-axis is the month during the growing season and the y-axis is the recorded mean monthly temperature (A) and mean monthly precipitation (B) from the Kansas Mesonet weather stations. Symbol and line combinations represent the location and year the data was recorded. ....	130
Figure 4.4. Grain yield by year and location for the KSWB experiments over 5 years. The x-axis is the harvest year growing season and the y-axis is adjusted grain yield in tons ha <sup>-1</sup> . Colors represent location that data was collected for during each season. ....	131
Figure 4.5. VI distribution by thermal time flight date for GNDVI, NDRE and NDVI at the 2021 Reno County location. The x-axis is the thermal time at the flight date, the y-axis is the calculated VI value and the colors indicate the VI.....	132
Figure 4.6. Broad sense heritability for GNDVI, NDRE and NDVI by flight date. Graphs include 5 years and 9 locations. However, not all locations are represented every year. ....	133
Figure 4.7. 3D Variogram created from 2D splines for NDRE collected 6/15/21 for the 2021 Reno Yield trial. The x-axis is the row effect the y-axis is the column displacement, and the z-axis is the spatial trend. ....	133
Figure 4.8 2D splines spatial correction workflow example. Data is the NDRE, VI collected 6/15/21 from the 2021 Reno Yield trial. The x-axis on the maps are columns and the y-axis are the rows. ....	134
Figure 4.9. Pearson correlation (A) and Spearman rank order correlation (B) values for adjusted grain yield, raw grain yield, and spatially fitted GNDVI, NDRE and NDVI at 2021 Reno County. ....	135

Figure 4.10. Pearson correlation (A) and Spearman rank order correlation (B) values for adjusted grain yield, raw grain yield, and spatially fitted  $\Delta VI$  at 2021 Reno County. .... 136

Figure 4.11. Linear model performance for grain yield rank order and fitted rank order by experiment at the 2021 McPherson location. The x-axis is the rank order of the fitted GNDVI values, and the y-axis is the adjusted grain yield rank order. The gray title box indicates the year location, experiment, and thermal time flight date..... 137

Figure 4.12. Rank order importance for the random forest model incorporating all VI and temporal combinations for 3 distinct yield groupings. The x-axis is the decreased mean accuracy and the y-axis the VI, thermal time combination. Higher decrease accuracy numbers indicate higher importance. .... 138

## List of Tables

Table 2.1-Association Mapping Panel Entries for 2017, 2018, 2019 and 2020 experiments located at the Kansas State University Ashland Bottoms and Rocky Ford Research Stations. .....	46
Table 2.2- Kansas Wheat Breeding Program for 2018, 2019 and 2020 experiments located in Ellsworth, McPherson, Saline, Reno, and Riley Counties. Included established and experimental genotypes.....	46
Table 2.3- Field experimental locations and weather station details for the Association Mapping Panel experiments.....	47
Table 2.4- Field experimental locations and weather station details for the Kansas Wheat Breeding Program experiments. ....	47
Table 2.5- Average heading date, day of year (DOY) by experiment, location, and year for both the AM Panel and the Kansas Wheat Breeding Program.....	48
Table 2.6- Thermal time index performance of AM Panel across all years and locations .....	48
Table 2.7- Thermal time index performance of Kansas State Wheat Breeding program across all years and locations. ....	49
Table 2.8- APSIM thermal time accumulation, photo period sensitivity and vernalization sensitivity coefficients by genotype for the AM Panel .....	50
Table 2.9- APSIM thermal time accumulation, photo period sensitivity and vernalization sensitivity coefficients by genotype for the Kansas Wheat Breeding Program .....	51
Table 3.1- Field experimental locations sample pattern, sensor platform and sample density for spatial soil sampling .....	101
Table 3.2- Field experimental locations and weather station details for the Association Mapping Panel experiments.....	101
Table 3.3- MSP3 K-means clustering, variogram model type, sill and range for all locations and all years.....	102
Table 3.4- P4000 K-means clustering, variogram model type, sill and range for all locations and all years.....	103

Table 3.5- Coefficient of Variation (CV%) of yield data from seven trial across two years.  
 Values are obtained from the raw yields, the experimental design spatial corrections, and spatial corrections from soil core, MSP3 and P4000 data..... 105

Table 4.1- Growing degree accumulation at flight date for all locations. Thermal time categories are designated at the top of the table. These categories were used for all analysis..... 139

Table 4.2- Equations for vegetation indices used to evaluate rank order predictions..... 140

Table 4.3- Pearson’s correlation of spatially fitted VI temporal points to adjusted grain yield by location. .... 141

Table 4.4- Spearman’s correlation of spatially fitted VI temporal points to adjusted grain yield by location. .... 141

Table 4.5- P-values for linear model VI rank order predictions to observed grain yield rank order ..... 142

Table 4.6- Table of percentages of the total population needed to capture the top 10% yield group using single temporal VI’s as a selection criteria. .... 143

Table 4.7- Random forest model performance metrics. All models used a 60% training population and a 40% testing population. Additionally, all models use the GNDVI, NDRE, and NDVI indices. The thermal time column indicates the temporal thermal times used to create and test the model. .... 144

## Acknowledgements

I would like to thank my co-major advisors, Dr. Jesse Poland and Dr. Allan Fritz for providing me with this opportunity. Additionally, I would like to thank my committee members Dr. Ignacio Ciampitti and Dr. Romulo Lollato for their continued support throughout my research. I would also like to recognize Dr. Lucas Haag for his support with UAV operations at Northwest Research-Extension Center in Colby KS. His willingness to preform regular UAV flights saved many hours of travel.

Acknowledgment is made to the APSIM Initiative which takes responsibility for quality assurance and a structured innovation program for APSIM's modelling software, which is provided free for research and development use (see [www.apsim.info](http://www.apsim.info) for details). Additionally, I would like to recognize Dr. Pengcheng Hu, Dr. Bangyou Zheng, and Dr. Scott Chapman for the assistance with the APSIM simulations.

I greatly appreciate Xu Wang, and Rich Brown for their guidance and continuing to push me with precision agriculture technology. Additionally, I would like to thank Mark Lucas for teaching me SQL database management and adding a unique perspective to curating large datasets. I would like to thank Duane Wilson, John Ruapp and Shuangye Wu for their professional support through this process and helping me with multiple projects throughout this process. I would like to recognize Marshall Clinesmith, Shaun Winnie and Andrew Auld for their help with collection and curating the data for the Kansas State Wheat Breeding yield trials. Additionally, I would like to thank all the Poland and WGRC Lab members, particularly Dr. Paula Silva, Dr. Dal-Hoe Koo and Dr. Bernd Friebe for their friendship, support, and input over the last 5 years.

Lastly, I would like to acknowledge my family for their patience, support, guidance, and encouragement throughout this journey. Particularly my wife KC and daughters Ashlyn and Kamryn. They have invested equal time and energy while making sacrifices to ensure the completion of this project.

## **Dedication**

I would like to dedicate this work to all my family members that have influenced my decision to pursue a career in agriculture. Particularly my parents Jim and Kathleen Evers, my late uncle John Evers and my late grandparents Harry and Bernice Evers and Richard and Berneva Stubbe.

# Chapter 1 - Literature Review

## Global and local importance of wheat

Wheat (*Triticum aestivum*) is an important source of calories worldwide. It is estimated that wheat provides 20% of the world's protein and 20% percent of the total calories consumed worldwide come from wheat (Reynolds et al., 2012). This makes it one of the three most important cereal crops globally. Currently, annual genetic gain in wheat is 0.8-1.2%. However, this rate is not sufficient to provide enough food and fuel for projected population increases and it is estimated that genetic gain needs to increase to 2.4% annually to be sustainable (Ray et al., 2013). Therefore, efforts are needed to increase genetic gain amongst breeding programs.

In the United States, Kansas historically has the highest production of hard red winter wheat, earning the nickname "The Wheat State". Concerted efforts to develop cultivars less prone to lodging and more adapted to the climate of the Central Plains resulted in better production for producers (Olmstead & Rhode, 2011). Currently, wheat is planted on more acres in Kansas than any other crop. It was reported that 6,800,000 acres of wheat were planted in 2020, producing 281,250,000 bushels (NASS 2021). Furthermore, wheat sales generated \$1.27 billion in revenue making wheat the third most economically important grain crop grown in Kansas, behind corn (*Zea mays* L.) and soybeans (*Glycine max* L.). Currently, both private and public breeding programs, continue to allocate resources to develop improved cultivars. With the adoption of new technology and the incorporation of diverse germplasm, there is new potential to continue to increase genetic gain and increase wheat adaptability in the Central Plains (Crain et al., 2018; He et al., 2019).

## **Current Breeding Technologies**

Traditional wheat breeding programs develop new varieties through crossing, inbreeding, performance evaluation and selection. All these processes are simultaneously occurring in a continuous cycle, with each stage taking time and resources. Therefore, efforts are being made at all stages to increase efficiency, allowing breeders to increase production on reduced budgets (Bentley & Mackay, 2017).

Using genetic tools allows breeders to make better selections and therefore put more time and financial resources in to evaluating plants with desired traits. Marker assisted selection (MAS) has successfully been used in breeding applications to identify desirable candidate lines (Chhetri et al., 2017; Tanksley & Rick, 1980). Similarly using genomic selection (GS) models holds potential to reduce selection cycles in field crops compared to phenotypic selection alone (Voss-Fels et al., 2019). Furthermore, GS can increase prediction accuracies over other marker platforms in wheat (Poland et al., 2012). However, developing representative training sets and having proper marker coverage can be an issue with this technology. Furthermore, genotype by environment and epistatic effects limit the prediction effectiveness of both MAS and GS (Lorenzana & Bernardo, 2009).

Advancements and implementation of double haploid (DH) production have decreased the time of the inbreeding cycle. Through the wheat-maize DH system, breeders can now produce complete inbred lines of elite crosses in one year. These lines have been used for genetic studies or evaluated for variety release (Berg et al., 2006; da Silva et al., 2019). However, the current DH process has limited throughput and creating these lines adds cost which limits the adoption of this technology in breeding programs. Currently, breeders have adopted speed breeding as an alternative way to increase generation advancement (Ghosh et al., 2018). This

coupled with other technology such as genomic selection can increase trait selection within wheat programs (Watson et al., 2019).

Of these three stages, performance testing is normally the slowest and most expensive step, because large populations of individuals need to be evaluated over many years and multiple environments prior to selection. Currently, research efforts are being focused on non-mechanical ways to evaluate large populations at early selection cycles (Hu et al., 2020). Previous research has shown high-throughput phenotyping (HTP) with unmanned aerial vehicles (UAVs) can be used as a tool to increase the accuracy of selection and evaluation of larger populations, leading to increased genetic gain (Haghighattalab et al., 2016). In addition to increasing selection accuracy, sensor-based HTP technologies also allow for increased temporal evaluation of plants in a non-destructive manner (Yasrab et al., 2021). This may allow researchers to gain useful insight on crop physiological responses throughout multiple crop development stages.

Although each breeding tool is individually effective, it is evident that each one of these is a tool and not a stand-alone solution. As with any tool, the effectiveness is usually increased when several methods are combined. Recently, it has been shown that combining GS and HTP-hyperspectral data was able to increase prediction accuracy in wheat nurseries (Krause et al., 2019). Similarly in canola HTP data was used to identify QTL for early growth-related traits (Knoch et al., 2020). Overall, as breeding programs move forward and technology continues to improve, integration of multiple technologies into a cohesive breeding strategy may lead to higher genetic gains.

### **High Throughput Phenotyping using Sensor Platforms**

The principle of plant interactions with both visible and non-visible areas of the light spectrum have been a part of agronomic research since Gausman (1974) first reported

distinguishable reflectance curves in plants. Currently in agriculture, both production and research, applications of reflectance data are gaining momentum. However, it has been well documented that even advanced sensors have limitations in spatial and temporal resolution causing bottlenecks for implementing HTP data into breeding programs (Hein et al., 2021; Song et al., 2021). As technology and computational power continues to improve, continued research will be needed to understand the relationship between HTP and key crop physiological characteristics.

Often the final goal in plant breeding is to increase production in the form of yield. Therefore, yield prediction using spectral sensors has been evaluated across many crops and production situations (Hassan et al., 2019; Wang et al., 2014). In many cases multiple yield components or vegetative indices are combined to predict yield and evaluate diverse treatments and genotypes. Multispectral satellite data has been used to model maize yield and make in-season corn yield predictions (Joshi et al., 2019; Peralta et al., 2016). However, these examples use large scale production fields, while within breeding applications where yield plots are significantly smaller, satellite imagery may not be adequate for data collection due to limitations in both spatial and temporal resolution (Tattaris et al., 2016).

In contrast to direct yield prediction, HTP sensor platforms can also be used to collect data on physiological characteristics that can be associated with yield and used for variety selection within a breeding program. For example, plant height can impact crop production, therefore plant breeders need to evaluate this trait for selection so that it was substantially modified through breeding over time in Kansas (Maeoka et al., 2020). Multiple methods have been used to demonstrate the feasibility of extracting plant height through HTP sensor platforms (Hu et al., 2018; Wang et al., 2018). In these examples authors extracted height measurements

from orthorectified photos through the development of digital elevation models (DEM). In addition to directly measuring plant height, DEMs can also evaluate lodging severity (Sun et al., 2019). Additionally, pairing HTP lodging scores with genome-wide association studies (GWAS) can identify markers and assist selection in early stages of germplasm development (Singh et al., 2019). These tools are potentially useful for breeders to identify unfavorable genotypes early in the breeding cycle and avoid advancing genotypes that are prone to lodging.

Crop diseases can also greatly reduce crop yield potential in Kansas (Cruppe et al., 2021) and impact variety selection. Therefore, it has been proposed to use spectral sensors as a means for monitoring crop health (Lowe et al., 2017) and aiding fungicide decisions (Cruppe et al., 2017). In many cases, early detection of a disease or infestation can be crucial in reducing the impacts on the crop by allowing producers to potentially treat a field. Both multi-spectral and hyperspectral sensors have been used to identify diseases such as crown rot, root rot powdery mildew, and leaf rust (Franke & Menz, 2007; Hillnhutter et al., 2011). Additionally, combining convolutional neural networks (CNNs) with evolving multispectral imagery acquisition has shown promise for improved disease detection (DeChant et al., 2017).

Through the combination of HTP data and genetic tools such as GWAS and marker assisted selection (MAS) breeders can make better selections for cultivar development. Although this has already been implemented in many crops and many programs (Condorelli et al., 2018; Pauli et al., 2016), the need to further develop and explore the potential of this technology within wheat breeding programs remains.

## **Modeling Phenological Development**

Historically a major limitation to phenological models was access to weather data with the needed temporal resolution and accuracy of parameters within an acceptable spatial

proximity of a trial. However, recent technical advancements have made sensors and dataloggers cheaper, more reliable, and more powerful. This has shifted public weather data collection from human observations to highly technical research-grade instrumentation. As of 2019, there were 62 publicly maintained open access weather stations that were part of the Kansas Mesonet network. These stations have the capability of collecting hourly precipitation, air temperature, relative humidity, barometric pressure, windspeed, solar radiation, soil temperature and soil moisture (Patrignani et al., 2020). Many of these stations are located at Kansas State University Research Stations or on land of cooperative landowners. The distribution and quantity make these weather stations a good source of weather data for field trials, including many of the Kansas State Wheat Breeding program's research trials.

Modeling phenology of crops can provide insight to key physiological and genetic responses to yield components (Dhillon et al., 2020). In HTP, within breeding programs, these models also can provide a link to multi-year reflectance data. Currently, researchers have used several different approaches to combine these multi-year data sets. The simplest way is to use day of year, with January 1<sup>st</sup> being day one (Naser et al., 2020). However, this method has limitations primarily because day of year is subject to management factors such as planting date. Furthermore, this method does not account for environmental factors such as solar radiation, precipitation, or temperature regime. All these factors can impact plant development and can significantly change between years, particularly with a winter crop such as wheat (Han et al., 2019).

Incorporating weather data to create thermal time indices to model phenology has been used in crop research for decades (Hildreth et al., 1941). One of the most popular thermal time indices uses is Growing Degree Days (GDD) (Gilmore Jr & Rogers, 1958). This model accounts

for environmental factors by incorporating daily temperature when predicting phenological growth stages and is commonly used in small grain simulations (McMaster & Wilhelm, 2003). The GDD models use three different temperature measurements a maximum ( $T_{MAX}$ ), minimum ( $T_{MIN}$ ) and a base temperature ( $T_{BASE}$ ). While this method is widely accepted and uses environmental data there are still discrepancies in what base temperatures to use and how the base temperature is incorporated into the equation. McMaster and Wilhelm (1997), reported up to a 9% difference in cumulative thermal units in Colorado when calculating GDD two different ways. This example provides insight into how thermal time indices can be misrepresented through lack of standardization.

In addition to model discrepancies, the vernalization component of winter wheat provides unique challenges for phenology modeling compared to most commercially grown crops. Although photoperiod and vernalization impacts on phenology are genetically controlled, only ~53% of the variation in phenology is due to genetics (Cane et al., 2013). Meaning that accounting for environmental factors such as temperature and water stress can have major implications on predicting phenology. Some studies avoid incorporating this period by modeling the growth from a fixed time, after most of the vernalization period has occurred (Lollato et al., 2020). Even though it has been shown that phenology stage can reasonably be predicted through ignoring vernalization, there are several reports that support incorporating vernalization to improve the phenological prediction. Wang and Engel (1998) proposed calculating total physiological development days after emergence as a two-part function. In this example physiological development days needed for vegetative and reproductive phase are calculated differently. Whereas the reproductive stage is only temperature dependent, the vegetative stage

incorporates a vernalization function that incorporates photoperiod. Variations of this this equation have been tested and adapted for winter wheat crops globally (Xue et al., 2004).

It has been theorized that incorporating other environmental factors such as soil temperature and soil moisture can improve phenology models. Wang and Engel (1998) proposed using soil temperature, if available, to calculate emergence. Additionally, it has been proposed that incorporating soil water measurement into phenology models can increase accuracy (Chauhan et al., 2019). However, this proposes a unique challenge as soil water available is highly dynamic. Spatial variation in soil texture, along with other physical and chemical properties can all greatly impact soil available water capacity (AWC) (Lei et al., 2012). Additionally, crop canopy and root systems can affect soil available water, particularly in variety trials and breeding applications where a wider range of genotypes are being tested (Amiri et al., 2021). Therefore, when evaluating the right model and model parameters for a breeding program multiple environmental factors need to be considered.

### **Spatial Variability in Crop Breeding**

In plant breeding spatial variation effects can increase errors and ultimately genetic gain. In-field micro-environments can impact yield trials particularly when many genotypes are tested within a single field (Wasson et al., 2014). This variability can strongly affect phenotypic response of these genotypes, making it necessary to develop spatial corrections for plant breeding. Many plant breeding programs correct for spatial effects through implementing modified experimental designs. It has even been theorized that spatial corrections are more important than experimental design (Borges et al., 2019). However, it has been documented that combining spatial modeling with the proper experimental design allows for the best allocation of resources within a breeding program (Gonzalez-Barrios et al., 2019; Piepho & Williams, 2010).

In addition to computational modeling advancements, sensing platforms have also recently been used to quantify spatial variation. High throughput phenotyping (HTP) sensors can provide useful data for spatial variability models. Recently, vegetative indices derived from multispectral data were used to evaluate spatial variability in durum wheat and identify yield sub-areas of the fields at harvest stage (Marino & Alvino, 2019). UAV imagery has also been used with spatial models to improve detection of spatial patterns in vineyards, proving to be useful in vineyard management (Matese et al., 2019). Furthermore, UAV data has the ability to increase spatial variability resolution in prediction of biomass production (Insua et al., 2019).

In plant breeding, most of the inherent spatial variability affecting field screening of experimental genotypes is due to plant, soil, and water interactions. However, due to the heterogeneous nature of soil, it is often difficult to quantify spatial variation with limited point data. New advancements in soil sensing, such as soil probes with optical spectroscopy, electrical conductivity, and penetrometer sensors, have shown potential for spatially identifying soil microclimates that can be used in plant breeding prediction models (Pei et al., 2019). Additionally, on-the-go soil electrical conductivity sensors have the ability to predict grid sample site with up to 80% accuracy making them an alternative to grid sampling (Shaner et al., 2008).

Combining HTP spectral data, soil sensor data, and point soil core samples can increase the accuracy of spatial maps over point measurements alone. Current research using data fusion of spatial data has shown promise in spatial models (Ji et al., 2019). Similar methods have also been used on a production scale to cluster satellite remote sensing data and soil sensing data to identify homogeneous parts of agriculture fields (Saifuzzaman et al., 2019). As large data sets continue to be collected using HTP, the importance of integrating data will continue to be an area of needed exploration.

## References

- Amiri, Z., Gheysari, M., Mosaddeghi, M. R., Amiri, S., & Tabatabaei, M. S. (2021). An attempt to find a suitable place for soil moisture sensor in a drip irrigation system. *Information Processing in Agriculture*.
- Bentley, A. R., & Mackay, I. (2017). Advances in wheat breeding techniques. In *Achieving sustainable cultivation of wheat Volume 1* (pp. 109-111). Burleigh Dodds Science Publishing.
- Berg, J. E., Stougaard, R. N., Cook, C. R., Kephart, K. D., Carlson, G. R., Wichman, D. M., Eckhoff, J. L., Riveland, N., Kushnak, G. D., Nash, D. L., & Bruckner, P. L. (2006). Registration of 'MT1159CL' wheat. *Crop Science*, *46*(3), 1395-1396.  
<https://doi.org/10.2135/cropsci2005.11-0405>
- Borges, A., González-Reymundez, A., Ernst, O., Cadenazzi, M., Terra, J., & Gutiérrez, L. (2019). Can Spatial Modeling Substitute for Experimental Design in Agricultural Experiments? *Crop Science*, *59*(1), 44-53.  
<https://doi.org/https://doi.org/10.2135/cropsci2018.03.0177>
- Cane, K., Eagles, H. A., Laurie, D. A., Trevaskis, B., Vallance, N., Eastwood, R. F., Gororo, N. N., Kuchel, H., & Martin, P. J. (2013). Ppd-B1 and Ppd-D1 and their effects in southern Australian wheat. *Crop & Pasture Science*, *64*(2), 100-114.  
<https://doi.org/10.1071/Cp13086>
- Chauhan, Y. S., Ryan, M., Chandra, S., & Sadras, V. O. (2019). Accounting for soil moisture improves prediction of flowering time in chickpea and wheat. *Scientific Reports*, *9*.  
<https://doi.org/ARTN 7510>  
10.1038/s41598-019-43848-6
- Chhetri, M., Bariana, H., Wong, D., Sohail, Y., Hayden, M., & Bansal, U. (2017). Development of robust molecular markers for marker-assisted selection of leaf rust resistance gene Lr23 in common and durum wheat breeding programs. *Molecular Breeding*, *37*(3).  
<https://doi.org/ARTN 21>  
10.1007/s11032-017-0628-6
- Condorelli, G. E., Maccaferri, M., Newcomb, M., Andrade-Sanchez, P., White, J. W., French, A. N., Sciara, G., Ward, R., & Tuberosa, R. (2018). Comparative Aerial and Ground Based High Throughput Phenotyping for the Genetic Dissection of NDVI as a Proxy for Drought Adaptive Traits in Durum Wheat (vol 9, 893, 2018). *Frontiers in Plant Science*, *9*. <https://doi.org/ARTN 1885>  
10.3389/fpls.2018.01885

- Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2018). Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *Plant Genome*, 11(1). <https://doi.org/ARTN 170043>  
10.3835/plantgenome2017.05.0043
- Cruppe, G., DeWolf, E., Jaenisch, B. R., Onofre, K. A., Valent, B., Fritz, A. K., & Lollato, R. P. (2021). Experimental and producer-reported data quantify the value of foliar fungicide to winter wheat and its dependency on genotype and environment in the US central Great Plains. *Field Crops Research*, 273, 108300.
- Cruppe, G., Edwards, J. T., & Lollato, R. P. (2017). In-season canopy reflectance can aid fungicide and late-season nitrogen decisions on winter wheat. *Agronomy Journal*, 109(5), 2072-2086.
- da Silva, C. L., Fritz, A., Clinesmith, M., Poland, J., Dowell, F., & Peiris, K. (2019). QTL mapping of Fusarium head blight resistance and deoxynivalenol accumulation in the Kansas wheat variety 'Everest'. *Molecular Breeding*, 39(3). <https://doi.org/ARTN 35>  
10.1007/s11032-019-0937-z
- DeChant, C., Wiesner-Hanks, T., Chen, S. Y., Stewart, E. L., Yosinski, J., Gore, M. A., Nelson, R. J., & Lipson, H. (2017). Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning. *Phytopathology*, 107(11), 1426-1432. <https://doi.org/10.1094/Phyto-11-16-0417-R>
- Dhillon, J. S., Figueiredo, B. M., Eickhoff, E. M., & Raun, W. R. (2020). Applied use of growing degree days to refine optimum times for nitrogen stress sensing in winter wheat. *Agronomy Journal*, 112(1), 537-549. <https://doi.org/10.1002/agj2.20007>
- Franke, J., & Menz, G. (2007). Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precision Agriculture*, 8(3), 161-172. <https://doi.org/10.1007/s11119-007-9036-y>
- Gausman, H. W. (1974). Leaf Reflectance of near-Infrared. *Photogrammetric Engineering and Remote Sensing*, 40(2), 183-191. <Go to ISI>://WOS:A1974S106600005
- Ghosh, S., Watson, A., Gonzalez-Navarro, O. E., Ramirez-Gonzalez, R. H., Yanes, L., Mendoza-Suárez, M., Simmonds, J., Wells, R., Rayner, T., Green, P., Hafeez, A., Hayta, S., Melton, R. E., Steed, A., Sarkar, A., Carter, J., Perkins, L., Lord, J., Tester, M., Osbourn, A., Moscou, M. J., Nicholson, P., Harwood, W., Martin, C., Domoney, C., Uauy, C., Hazard, B., Wulff, B. B. H., & Hickey, L. T. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nature Protocols*, 13(12), 2944-2963. <https://doi.org/10.1038/s41596-018-0072-z>
- Gilmore Jr, E., & Rogers, J. (1958). Heat units as a method of measuring maturity in corn 1. *Agronomy Journal*, 50(10), 611-615.

- Gonzalez-Barrios, P., Diaz-Garcia, L., & Gutierrez, L. (2019). Mega-Environmental Design: Using Genotype x Environment Interaction to Optimize Resources for Cultivar Testing. *Crop Science*, 59(5), 1899-1915. <https://doi.org/10.2135/cropsci2018.11.0692>
- Haghighattalab, A., Perez, 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. <https://doi.org/ARTN> 35  
10.1186/s13007-016-0134-6
- Han, L., Yang, G. J., Dai, H. Y., Yang, H., Xu, B., Li, H. L., Long, H. L., Li, Z. H., Yang, X. D., & Zhao, C. J. (2019). Combining self-organizing maps and biplot analysis to preselect maize phenotypic components based on UAV high-throughput phenotyping platform. *Plant Methods*, 15. <https://doi.org/ARTN> 57  
10.1186/s13007-019-0444-6
- Hassan, M. A., Yang, M. J., Rasheed, A., Yang, G. J., Reynolds, M., Xia, X. C., Xiao, Y. G., & He, Z. H. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, 282, 95-103. <https://doi.org/10.1016/j.plantsci.2018.10.022>
- He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P., Forrest, K., Fritz, A., Hucl, P., Wiebe, K., Knox, R., Cuthbert, R., Pozniak, C., Akhunova, A., Morrell, P. L., Davies, J. P., Webb, S. R., Spangenberg, G., Hayes, B., Daetwyler, H., Tibbits, J., Hayden, M., & Akhunov, E. (2019). Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome (vol 50, pg 896, 2019). *Nature Genetics*, 51(7), 1194-1194. <https://doi.org/10.1038/s41588-019-0463-2>
- Hein, N. T., Ciampitti, I. A., & Jagadish, S. V. K. (2021). Bottlenecks and opportunities in field-based high-throughput phenotyping for heat and drought stress. *Journal of Experimental Botany*, 72(14), 5102-5116. <https://doi.org/10.1093/jxb/erab021>
- Hildreth, A. C., Magness, J. R., & Mitchell, J. W. (1941). Effects of Climatic Factors on Growing Plants [Article]. *USDA*, p. 292-307. <http://handle.nal.usda.gov/10113/IND43893773>
- Hillnhutter, C., Mahlein, A. K., Sikora, R. A., & Oerke, E. C. (2011). Remote sensing to detect plant stress induced by *Heterodera schachtii* and *Rhizoctonia solani* in sugar beet fields. *Field Crops Research*, 122(1), 70-77. <https://doi.org/10.1016/j.fcr.2011.02.007>
- Hu, P. C., Chapman, S. C., Wang, X. M., Potgieter, A., Duan, T., Jordan, D., Guo, Y., & Zheng, B. Y. (2018). Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding. *European Journal of Agronomy*, 95, 24-32. <https://doi.org/10.1016/j.eja.2018.02.004>
- Hu, Y. C., Knapp, S., & Schmidhalter, U. (2020). Advancing High-Throughput Phenotyping of Wheat in Early Selection Cycles. *Remote Sensing*, 12(3). <https://doi.org/ARTN> 574  
10.3390/rs12030574

- Insua, J. R., Utsumi, S. A., & Basso, B. (2019). Estimation of spatial and temporal variability of pasture growth and digestibility in grazing rotations coupling unmanned aerial vehicle (UAV) with crop simulation models. *Plos One*, *14*(3). <https://doi.org/ARTN e0212773>  
10.1371/journal.pone.0212773
- Ji, W. J., Adamchuk, V. I., Chen, S. C., Su, A. S. M., Ismail, A., Gan, Q. J., Shi, Z., & Biswas, A. (2019). Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study. *Geoderma*, *341*, 111-128.  
<https://doi.org/10.1016/j.geoderma.2019.01.006>
- Joshi, V. R., Thorp, K. R., Coulter, J. A., Johnson, G. A., Porter, P. M., Strock, J. S., & Garcia, A. G. Y. (2019). Improving Site-Specific Maize Yield Estimation by Integrating Satellite Multispectral Data into a Crop Model. *Agronomy-Basel*, *9*(11). <https://doi.org/ARTN 719>  
10.3390/agronomy9110719
- Knoch, D., Abbadi, A., Grandke, F., Meyer, R. C., Samans, B., Werner, C. R., Snowdon, R. J., & Altmann, T. (2020). Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola. *Plant Biotechnology Journal*, *18*(1), 68-82. <https://doi.org/10.1111/pbi.13171>
- Krause, M. R., Gonzalez-Perez, L., Crossa, J., Perez-Rodriguez, P., Montesinos-Lopez, O., Singh, R. P., Dreisigacker, S., Poland, J., Rutkoski, J., Sorrells, M., Gore, M. A., & Mondal, S. (2019). Hyperspectral Reflectance-Derived Relationship Matrices for Genomic Prediction of Grain Yield in Wheat. *G3-Genes Genomes Genetics*, *9*(4), 1231-1247. <https://doi.org/10.1534/g3.118.200856>
- Lei, X., Li, F., Zhou, S., Li, Y., Chen, D., Liu, H., Pan, Y., & Shen, X. (2012). Spatial Variability and Lateral Location of Soil Moisture Monitoring Points on Cotton Mulched Drip Irrigation Field. In D. Li & Y. Chen, *Computer and Computing Technologies in Agriculture V* Berlin, Heidelberg.
- Lollato, R. P., Bavia, G. P., Perin, V., Knapp, M., Santos, E. A., Patrignani, A., & DeWolf, E. D. (2020). Climate-risk assessment for winter wheat using long-term weather data. *Agronomy Journal*, *112*(3), 2132-2151. <https://doi.org/10.1002/agj2.20168>
- Lorenzana, R. E., & Bernardo, R. (2009). Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theoretical and Applied Genetics*, *120*(1), 151-161. <https://doi.org/10.1007/s00122-009-1166-3>
- Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, *13*. <https://doi.org/ARTN 80>  
10.1186/s13007-017-0233-z
- Maeoka, R. E., Sadras, V. O., Ciampitti, I. A., Diaz, D. R., Fritz, A. K., & Lollato, R. P. (2020). Changes in the phenotype of winter wheat varieties released between 1920 and 2016 in response to in-furrow fertilizer: Biomass allocation, yield, and grain protein concentration. *Frontiers in Plant Science*, *10*, 1786.

- Marino, S., & Alvino, A. (2019). Detection of Spatial and Temporal Variability of Wheat Cultivars by High-Resolution Vegetation Indices. *Agronomy-Basel*, 9(5). <https://doi.org/ARTN> 226  
10.3390/agronomy9050226
- Matese, A., Di Gennaro, S. F., & Santesteban, L. G. (2019). Methods to compare the spatial variability of UAV-based spectral and geometric information with ground autocorrelated data. A case of study for precision viticulture. *Computers and Electronics in Agriculture*, 162, 931-940. <https://doi.org/10.1016/j.compag.2019.05.038>
- McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87(4), 291-300. <https://doi.org/Doi> 10.1016/S0168-1923(97)00027-0
- McMaster, G. S., & Wilhelm, W. W. (2003). Phenological responses of wheat and barley to water and temperature: improving simulation models. *Journal of Agricultural Science*, 141, 129-147. <https://doi.org/10.1017/S0021859603003460>
- Naser, M. A., Khosla, R., Longchamps, L., & Dahal, S. (2020). Using NDVI to Differentiate Wheat Genotypes Productivity Under Dryland and Irrigated Conditions. *Remote Sensing*, 12(5). <https://doi.org/ARTN> 824  
10.3390/rs12050824
- Olmstead, A. L., & Rhode, P. W. (2011). Adapting North American wheat production to climatic challenges, 1839-2009. *Proceedings of the National Academy of Sciences of the United States of America*, 108(2), 480-485. <https://doi.org/10.1073/pnas.1008279108>
- Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. *Journal of Atmospheric and Oceanic Technology*, 37(12), 2167-2183. <https://doi.org/10.1175/Jtech-D-19-0214.1>
- Pauli, D., Andrade-Sanchez, P., Carmo-Silva, A. E., Gazave, E., French, A. N., Heun, J., Hunsaker, D. J., Lipka, A. E., Setter, T. L., Strand, R. J., Thorp, K. R., Wang, S., White, J. W., & Gore, M. A. (2016). Field-Based High-Throughput Plant Phenotyping Reveals the Temporal Patterns of Quantitative Trait Loci Associated with Stress-Responsive Traits in Cotton. *G3-Genes Genomes Genetics*, 6(4), 865-879. <https://doi.org/10.1534/g3.115.023515>
- Pei, X. S., Sudduth, K. A., Veum, K. S., & Li, M. Z. (2019). Improving In-Situ Estimation of Soil Profile Properties Using a Multi-Sensor Probe. *Sensors*, 19(5). <https://doi.org/ARTN> 1011  
10.3390/s19051011
- Peralta, N. R., Assefa, Y., Du, J., Barden, C. J., & Ciampitti, I. A. (2016). Mid-Season High-Resolution Satellite Imagery for Forecasting Site-Specific Corn Yield. *Remote Sensing*, 8(10). <https://doi.org/ARTN> 848  
10.3390/rs8100848

- Piepho, H. P., & Williams, E. R. (2010). Linear variance models for plant breeding trials. *Plant Breeding*, 129(1), 1-8. <https://doi.org/10.1111/j.1439-0523.2009.01654.x>
- Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S. Y., Manes, Y., Dreisigacker, S., Crossa, J., Sanchez-Villeda, H., Sorrells, M., & Jannink, J. L. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome*, 5(3), 103-113. <https://doi.org/10.3835/plantgenome2012.06.0006>
- Saifuzzaman, M., Adamchuk, V., Buelvas, R., Biswas, A., Prasher, S., Rabe, N., Aspinall, D., & Ji, W. J. (2019). Clustering Tools for Integration of Satellite Remote Sensing Imagery and Proximal Soil Sensing Data. *Remote Sensing*, 11(9). <https://doi.org/ARTN> 1036 10.3390/rs11091036
- Shaner, D. L., Khosla, R., Brodahl, M. K., Buchleiter, G. W., & Farahani, H. J. (2008). How well does zone sampling based on soil electrical conductivity maps represent soil variability? *Agronomy Journal*, 100(5), 1472-1480. <https://doi.org/10.2134/agronj2008.0060>
- Singh, D., Wang, X., Kumar, U., Gao, L., Noor, M., Imtiaz, M., Singh, R. P., & Poland, J. (2019). High-Throughput Phenotyping Enabled Genetic Dissection of Crop Lodging in Wheat [Original Research]. *Frontiers in Plant Science*, 10(394). <https://doi.org/10.3389/fpls.2019.00394>
- Song, P., Wang, J. L., Guo, X. Y., Yang, W. N., & Zhao, C. J. (2021). High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop Journal*, 9(3), 633-645. <https://doi.org/10.1016/j.cj.2021.03.015>
- Sun, Q., Sun, L., Shu, M., Gu, X., Yang, G., & Zhou, L. (2019). Monitoring Maize Lodging Grades via Unmanned Aerial Vehicle Multispectral Image. *Plant Phenomics*, 2019, 5704154. <https://doi.org/10.34133/2019/5704154>
- Tanksley, S. D., & Rick, C. M. (1980). Isozymic Gene Linkage Map of the Tomato - Applications in Genetics and Breeding. *Theoretical and Applied Genetics*, 57(4), 161-170. <Go to ISI>://WOS:A1980KB26100004
- Tattaris, M., Reynolds, M. P., & Chapman, S. C. (2016). A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. *Frontiers in Plant Science*, 7. <https://doi.org/ARTN> 1131 10.3389/fpls.2016.01131
- Voss-Fels, K. P., Cooper, M., & Hayes, B. (2019). Accelerating crop genetic gains with genomic selection. *Theoretical and Applied Genetics*, 132(3), 669-686. <https://doi.org/10.1007/s00122-018-3270-8>
- Wang, E. L., & Engel, T. (1998). Simulation of phenological development of wheat crops. *Agricultural Systems*, 58(1), 1-24. <https://doi.org/Doi> 10.1016/S0308-521x(98)00028-6

- Wang, L. G., Tian, Y. C., Yao, X., Zhu, Y., & Cao, W. X. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *Field Crops Research*, 164, 178-188. <https://doi.org/10.1016/j.fcr.2014.05.001>
- Wang, X., Singh, D., Marla, S., Morris, G., & Poland, J. (2018). Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods*, 14. <https://doi.org/ARTN> 53  
10.1186/s13007-018-0324-5
- Wasson, A. P., Rebetzke, G. J., Kirkegaard, J. A., Christopher, J., Richards, R. A., & Watt, M. (2014). Soil coring at multiple field environments can directly quantify variation in deep root traits to select wheat genotypes for breeding. *Journal of Experimental Botany*, 65(21), 6231-6249. <https://doi.org/10.1093/jxb/eru250>
- Watson, A., Hickey, L. T., Christopher, J., Rutkoski, J., Poland, J., & Hayes, B. J. (2019). Multivariate Genomic Selection and Potential of Rapid Indirect Selection with Speed Breeding in Spring Wheat. *Crop Science*, 59(5), 1945-1959. <https://doi.org/https://doi.org/10.2135/cropsci2018.12.0757>
- Xue, Q. W., Weiss, A., & Baenziger, P. S. (2004). Predicting phenological development in winter wheat. *Climate Research*, 25(3), 243-252. <https://doi.org/DOI> 10.3354/cr025243
- Yasrab, R., Zhang, J. C., Smyth, P., & Pound, M. P. (2021). Predicting Plant Growth from Time-Series Data Using Deep Learning. *Remote Sensing*, 13(3). <https://doi.org/ARTN> 331  
10.3390/rs13030331

## **Chapter 2 - Integrating Phenological Development from Crop Models into High Throughput Phenotyping**

### **Introduction**

The use of high throughput phenotyping (HTP) with automated data acquisition platforms has the potential to be a transformative selection tool for breeding programs worldwide. Recent studies have shown that collecting sensor data at key physiological stages can increase yield prediction and assist with phenotypic selection (Crain et al., 2017; Hassan et al., 2019). Additionally, HTP sensor data can be used to assess disease response during key growth stages (Silva, 2021). However, growth stages vary from year to year depending on weather and management, which directly impacts the operational use of HTP platforms as well as data efficiency as both biotic and abiotic stressors impacting grain yield are driven by in-season environment trends (Bergkamp et al., 2018; Eversmeyer & Kramer, 2000). Therefore, having the ability to predict growth stages from weather station data is pivotal to deploying HTP data collection in breeding programs. The objective of this study is to evaluate different thermal time indices and crop model systems to predict heading date in winter wheat breeding populations.

In addition to predicting phenology for HTP purposes, these predictions can also be paired with sensor data to predict in-season regional yield and climate adaptation (Parida & Ranjan, 2019). As regional and global climate shifts continue, plant breeders need to develop cultivars that are adapted to the new climate and are resistant to weather extremes of the region. Modern wheat varieties have earlier flower times compared to legacy cultivars (Maeoka et al., 2020; Rezaei et al., 2018), indicating that breeders are selecting genetics that are more adapted to the current climate. However, further advances in flowering time have not shown potential to

improve yields in a modeling exercise (Sciarresi et al., 2019). However, Hu et al. (2005) attributed an earlier flowering date, of 0.8 to 1.2 days per 10 years in the legacy variety Kharkof, to increased spring temperatures in the US Central Plains. These studies confirm that shifting phenology to match changing climate is both genetically and environmentally controlled. This is significant to breeding programs where diverse germplasm needs to be evaluated at multiple locations across multiple years. By understanding phenology through modeling and climate trends, breeders can take a better approach to develop cultivars with the proper maturity to avoid typical high stress weather trends (Lollato et al., 2020).

Although thermal time indices are an important established method to normalize phenology data across years and locations, there are discrepancies on how these indices are calculated. McMaster and Wilhelm (1997) demonstrated that implementing minimum temperature thresholds to either the daily temperature or the daily average temperature can impact thermal indices. Likewise, researchers use different base temperature, maximum temperature, and optimum temperature between and within crops (Salazar-Gutierrez et al., 2013). Similarly, some indices, such as physiological days (Pdays) index (Saiyed et al., 2009), integrates an optimum temperature threshold that weights thermal time accumulation when daily weather is close to the optimum growth temperature. Variations in all calculations and threshold parameters can significantly impact the calculated thermal time accumulation and in-season phenology prediction of the crops. This justifies the exploration of multiple indices for breeding programs and adopting the system that fits best within a region for a given crop.

The vernalization component of winter wheat phenology poses unique additional challenges compared to spring wheat and other summer annual crops. It has been reported that within in Kansas, responses to vernalization, spring green up, and heading date are genotype

specific (Munaro et al., 2020). In an attempt to mitigate vernalization effects and standardized dates between years, researchers often use day of year after January 1<sup>st</sup> to evaluate HTP data both within year and across years (Giri, 2019; Lollato & Edwards, 2015; Silva, 2021). However, accounting for winter wheat early season plant growth can be significant for breeding programs and commercial production, specifically when evaluating stand establishment. Additionally, predicting full season phenology can also be useful for evaluating fall biomass production and estimating grazing thresholds where growth stage is critical to crop viability (Hossain et al., 2003). Thermal indices that calculate thermal time differently for vegetative and reproductive phase have been developed as a solution for phenology development (Streck et al., 2003; Wang & Engel, 1998). These thermal indices may have an advantage over temperature-only models within winter wheat breeding programs.

In addition to temperature, it has been documented that photoperiod can also affect the growing stage in many crops (Yan & Wallace, 1998). Typically, wheat is not as photoperiod dependent, but in crops such as soybeans, it can heavily influence crop development. However, it has been documented that wheat genotypes with *Vrn-A1*, *Vrn-B1*, *Brn-D1*, *Ppd-B1*, *Ppd-D1* are photoperiod and temperature sensitive for phenology (Kiss et al., 2021) and that photoperiod can impact phenological development and spikelet number (Halse & Weir, 1970). Furthermore, Robertson (1968) demonstrated that developing a thermal index that accounts for photoperiod by incorporating day length, can improve phenology estimations in cereal crops. Furthermore, this biometeorological time scale (BMT) contains genotype-specific response coefficients for both temperature and day length, making it adaptable to breeding programs. Additionally, several variations of thermal indices that include either day length or photoperiod coefficients have been widely used in literature (Aslam et al., 2017; Herndl et al., 2008). These reports indicate that

indices that incorporate a photoperiod component could potentially improve phenology prediction of winter wheat in Kansas.

Although thermal indices such as Pdays and BMT have proved to be useful tools individually for predicting phenology in wheat, using these indices within crop models have the potential further improve prediction accuracy of wheat phenology. These models can range from simplistic to robust by quantifying genetics, soil, and other weather variables effects on crop growth and development. Recently, there have been several other models developed that use individual parametrization to account for the genetic control of winter wheat (Aslam et al., 2017; Ceglar et al., 2019; Hodges & Ritchie, 1991). All these models are stage threshold driven meaning that the thermal equation changes once an accumulation limit is reached. Many of these models have 10 or more phase changes depending on crop types. Moreover, these threshold limits are typically genotype-independent, allowing for genotypic variation to be captured. This capability allows for breeders and researchers to evaluate responses to climate changes by genotype (Rezaei et al., 2018).

Overall, the assessment of computational phenology models could provide useful insight when incorporated in to breeding programs. Having the ability couple these simulations with evolving sensor and sequencing techniques further validates the need for established regional performance of the crop models.

## **Materials and Methods**

### **Plant Material**

To evaluate germplasm adapted to the Central Great Plains we evaluated two unique populations. The first population assessed was a subset of a diverse association mapping panel as described by Grogan et al. (2016). The initial panel consisted of 342 pure line winter wheat

varieties that originate from a wide range of breeding programs, time periods and growing regions across the United States. However, due to the diverse nature of this population, only varieties released after 1990 and were at one time in the top 10 varieties grown in Kansas were selected. These criteria resulted in the evaluation of 40 lines, as shown in Table 2.1.

The second population consisted of lines being tested within the Kansas State Wheat Breeding (KSWB) program. Due to the fluid nature of plant breeding, experimental lines are not contained in the same experiment for multiple years. Experimental lines are grown, evaluated and either advanced or discarded at the conclusion of the growing season. For this experiment we evaluated germplasm that was grown in either the Preliminary Yield Nursery (PYN), the Advanced Yield Nursery (AYN) and the Kansas Interstate Nursery (KIN) within the scope of the KSWB program. Although the entries were part of several different experiments over the 3 years, all experiments were in the same field within a location. Each of these trials can contain upwards of 200 lines. However, for this study, we evaluated only lines that were represented across at least 6 of the 9 site years were chosen for evaluation (Table 2.2). In addition to experimental germplasm ten released varieties that were used for checks were also evaluated. However, due to the wide geographic region of testing some checks were only included at locations where the variety was recommended for production.

The lines from the AM panel were grown over three years (2017, 2018 and 2020) at the Ashland Bottoms Experiment Station (39°08'19.07" N, 96°38'21.00" W), and at the Rocky Ford Experiment Station (39°13'48.82" N, 96°34'41.87" W) during the 2019 year. The experiment design was a randomized complete block with 2 replications. Plot dimensions were 2.4 m x 1.4 m (3.36 m<sup>2</sup>) and planted with a Great Plains 3P606NT grain drill that was modified by Kincaid Manufacturing (Haven, KS, USA). Each plot consisted of 6 rows spaced at 20 cm with a seeding

rate of 15 g m<sup>-2</sup>. The plots were planted in October of each year and harvested in late June or early July (Table 2.3).

To assess thermal time models within the KSWB program 5 locations were evaluated over 3 years. Due to location variability, performance, and breeder availability not every location was evaluated for heading date every year. In total 9 location-years were included in this study. Locations include Ellsworth County (2018, 38° 33' 4.13"N, 98° 22' 7.67"W and 2020, 38° 33' 4.13"N, 98° 22' 7.67"W), McPherson County (2020), Reno County (2018, 37° 55' 58.53"N, 98° 01' 53.91"W, 2019, 37° 57' 46.64"N, 98° 07' 14.89"W and 2020, 38° 33' 4.13"N, 98° 22' 7.67"W), Riley County (2018, 39° 7' 37.34"N, 96° 36' 39.93"W and 2019, 39° 7' 38.27"N, 96° 36' 36.56"W) and Saline County (2019, 38° 39' 10.54"N, 97° 36' 4.417"W, Table 2.4). Similarly, to the AM panel the plots were 1.4 m wide and consisted of 6 evenly spaced rows. However, plots were 4.5 m long and planted with a Hege plot drill. The plots were planted in either October or November of each year and harvested in late June or early July (Table 2.3).

### **Phenotypic Data Collection**

Plots in the diversity panel were visually scored for percent heading every two days during the heading period and data was digitally recorded using FieldBook (Rife & Poland, 2014). Percent heading was determined as the percentage of heads within the plot that had emerged from the boot in 10% intervals. Heading date was recorded when a plot reached a score of 50%, meaning 50% of the spike had emerged from 50% of all tillers (Crop Ontology CO\_321:0000840). This determination is consistent with stage 59 on the Zadoks scale (Zadoks et al., 1974). In plots where 50% was not recorded at a single day, the percent heading scores were interpolated to a heading date using a logistic regression model as described in (Wang et al., 2019).

Genotypes in the KSWB yield nurseries were evaluated several times during the heading period similarly to the AM panel. However, instead of using a percent scale to evaluate heading date, the heading date was recorded as the date when 50% of the spike had emerged from 50% of all the stems was observed.

## **Weather Data**

All weather data was obtained from weather stations that are maintained by the Kansas Mesonet network (<http://mesonet.k-state.edu>). All experiments, except for 2019 Saline County, were within 2 km of a Mesonet weather monitoring station (Table 2.4, Fig. 2.1). Hourly weather data was collected during the growing season for precipitation, maximum (Tmax) and minimum (Tmin) air temperature, incident solar radiation (Rs) wind speed at 2-m height (u) and relative humidity (Patrignani et al., 2020).

## **Thermal Indices**

Five base thermal indices, calculated seven different ways, were evaluated in this study. All the indices use daily weather data in combination with key physiological responses and in some cases daily photoperiod. Daily weather values were imported as a Pandas dataframe (Pandas Library, version 1.3.2) through a URL query string in Python programming language (Python, version 3.7.1, Python Software Foundation). Photoperiod, displayed as daily decimal hours with the time of sunrise and sunset on a 24 hour clock, was calculated using the ephemeris Python package (PyEphem, version 4.0.0.2). Both daily and cumulative values were calculated for all thermal indices using custom scripts written in Python. Cumulative thermal indices were calculated on three levels; (i) the entire growing season for planting to harvest designated as day after sowing (DAS); (ii) day of year (DOY) starting from January 1<sup>st</sup> of the harvest year; and (iii) day Mar 1<sup>st</sup> of the harvest year to simulate from the start of green up. The exception to this

method was the BMT thermal index where it was only calculated for the entire growing season due to the reduced scale of the index.

Growing degree days (GDD) is a widely accepted basic thermal index which uses the daily maximum and minimum observed temperature while incorporating a crop specific base ( $T_{base}$ ) temperature. Although the equation is rather simple, there are calculation discrepancies throughout literature and GDD is often calculated two ways as described by McMaster and Wilhelm (1997). One method of calculating GDD integrates the base temperature on the daily average temperature, whereas the other calculation incorporates the base temperature individually on the daily maximum and minimum temperatures before the daily average is calculated. Furthermore, a third way to calculate GDD incorporates an upper temperature threshold ( $T_{upper}$ ), where wheat development has been shown to cease (McMaster et al., 2008). For this experiment all three variations of calculating GDD were explored. The three GDD calculations are designated as GDD method one (GDD1, Eq. 1) and GDD method two (GDD2 Eq. 2) and GDD with a  $T_{max}$  upper threshold ( $T_{upper}$ ) ( $T_{max\_GDD}$  Eq. 3). For all calculations within this experiment the  $T_{base}$  was set at 0C, subsequently the  $T_{upper}$  temperature was set at 30C for the  $T_{max\_GDD}$  equation (Porter & Gawith, 1999).

$$GDD1 = \sum_{Planting}^{Harvest} \left( \frac{T_{max} + T_{min}}{2} \right) - T_{base} \quad [Eq. 1]$$

Where if;

$$\frac{T_{max} + T_{min}}{2} < T_{base} \text{ then } \frac{T_{max} + T_{min}}{2} = T_{base}$$

$$GDD2 = \sum_{Planting}^{Harvest} \left( \frac{T_{max} + T_{min}}{2} \right) - T_{base} \quad [Eq. 2]$$

Where if;

$$T_{max} \text{ or } T_{min} < T_{base} \text{ then } T_{max} \text{ or } T_{min} = T_{base}$$

$$T_{max\_GDD} = \sum_{Planting}^{Harvest} \left( \frac{T_{max} + T_{min}}{2} \right) - T_{base} \quad [Eq. 3]$$

Where if;

$$T_{max\ or\ T_{min}} < T_{base} \text{ then } T_{max\ or\ T_{min}} = T_{base}$$

$$T_{max\ or\ T_{min}} > T_{upper} \text{ then } T_{max\ or\ T_{min}} = T_{upper}$$

In addition to GDD, thermal time was calculated using the method described by Wang and Engel (1998) and is designated as Wang and Engel growing degree days ( $GDD_{WE}$ , Eq. 4). This nonlinear response function uses base, optimum and maximum temperature thresholds designated as  $T_{min}$ ,  $T_{opt}$  and  $T_{max}$  with the daily average temperature. In this experiment the threshold temperatures used were  $T_{min} = 0C$ ,  $T_{opt} = 27.7C$ , and  $T_{max} = 40C$ . In addition to temperature response a photoperiod component was also added to the  $GDD_{WE}$  equation to develop Photo Growing Degree Days (PGDD, Eq. 5) thermal index (Aslam et al., 2017) which uses a photoperiod coefficient and the day length in hours.

$$GDD_{WE} = \frac{2(T_{avg} - T_{min})^{\alpha} (T_{opt} - T_{min})^{\alpha} - (T_{avg} - T_{min})^{2\alpha}}{(T_{opt} - T_{min})^{2\alpha}} \quad [Eq. 4]$$

Where;

$$\alpha = \frac{\ln 2}{\ln \left( \frac{T_{max} - T_{min}}{T_{opt} - T_{min}} \right)}$$

$$PGDD = (WEDD) \left( (1 - 0.002 \times 2) \times (20 - \text{day length}) \right)^2 \quad [Eq. 5]$$

The Physiological days (Pdays) thermal time index was calculated using equation 6. In addition to using maximum and minimum temperature Pdays also incorporates an optimum

temperature. In this study  $T_{min}$  was set at 0C,  $T_{opt}$  at 17 C and  $T_{max}$  at 30C. Pdays implements a weighed scale, where thermal time is accumulated faster when daily temperatures are near the optimum temperature for crop growth (Saiyed et al., 2009).

$$Pdays = \frac{1}{24} (5 \times P(T_1) + 8 \times P(T_2) + 8 \times P(T_3) + 3 \times P(T_4)) \quad [\text{Eq. 6}]$$

Where;

$$T_1 = T_{min}$$

$$T_2 = \frac{(2 \times T_{min}) + T_{max}}{3}$$

$$T_2 = \frac{T_{min} + (2 \times T_{max})}{3}$$

$$T_2 = T_{max}$$

$$P = 0 \quad \text{when } T \leq T_{min}$$

$$P = k \times \left( 1 - \frac{(T - T_{opt}) \times 2}{(T_{max} - T_{min}) \times 2} \right) \quad \text{when } T_{min} \leq T \leq T_{opt}$$

$$P = k \times \left( 1 - \frac{(T - T_{opt}) \times 2}{(T_{max} - T_{opt}) \times 2} \right) \quad \text{when } T_{min} \leq T \leq T_{opt}$$

$$P = 0 \quad \text{when } T \geq T_{max}$$

The thermal time index Biomterological Time (BMT) was first introduced to spring wheat by Robertson (1968). BMT calculation combines temperature and photoperiod with response coefficients (Eq. 7). The day length (L) in hours of sunlight influences a genotypes photoperiod response in reference to the base daylength ( $a_0$ ). Whereas, the observed daily maximum ( $T_{max}$ ) and daily minimum ( $T_{min}$ ) temperatures are subtracted from a base temperature ( $b_0$ ). For this study the base temperature was set at 0°C. The coefficients (used for this equation are developed from Robertson (1968). These coefficients are not constant, and change based on the plants predicted development. The five phenological development stages for the coefficient

response changes are designated as: (i) planting to emergence (PE), (ii) emergence to jointing (EJ), (iii) jointing the heading (JH), (iv) heading to senescence (HS), and (v) senescence to ripening (SR). The development stage specific coefficients allow the model to accumulate thermal units according to the appropriate growth stage.

$$BMT = \sum_{Planting}^{Harvest} V_1 (V_2 + V_3) \quad [Eq. 7]$$

Where;

$$V_1 = a_1(L - a_0) + a_2(L - a_0)^2$$

$$V_2 = b_1(T_{max} - b_0) + b_2(T_{max} - b_0)^2$$

$$V_3 = d_1(T_{min} - b_0) + d_2(T_{min} - b_0)^2$$

### **Agricultural Production Systems Simulator (APSIM) Wheat Model**

The APSIM Wheat Model (Zheng et al.) uses physiological phase changes that are driven by accumulation of thermal time to initiate the phase change. The APSIM model uses the daily recorded maximum and minimum temperatures to calculate the daily maximum ( $T_{cmax}$ ) and minimum crown temperatures ( $T_{cmin}$ ) as shown in equation 8. Furthermore, the APSIM model uses the daily crown mean temperature ( $T_c$ ) to calculate the daily thermal time (TT). This calculation (Eq. 9) uses threshold temperatures of 0°C., 26°C. and 34°C.. The simulation also accounts for photoperiod calculated from experiment location latitude and day of year. The model also simulates vernalization effects (V) from the daily average crown temperature as shown in equation 10. The transition to phenological stages happens when the TT reaches a target accumulation as detailed in Zheng et al.

For this experiment all simulations were executed using the R package ‘rapsim’ (<https://github.com/APSIMInitiative/APSIMWheatPhenology>). Individual genotype coefficients were obtained for three phase changes including end of juvenal, floral initiation and start of grain

fill. Additionally, coefficients were calculated for photoperiod sensitivity and vernalization sensitivity. In total each genotype had a combination of five individual coefficients. In total, 300 iterations were performed to find the optimum combination of the coefficients which was selected by minimizing the RMSE.

$$T_{Cmax} = \begin{cases} 2 + T_{max}(0.4 + 0.0018(H_{snow} - 15)^2) & \text{When } T_{max} < 0 \\ T_{max} & \text{When } T_{max} \geq 0 \end{cases} \quad [\text{Eq. 8}]$$

$$\Delta TT = \begin{cases} T_C & 0 < T_C \leq 26 \\ \frac{26}{8}(34 - T_C) & 26 < T_C \leq 34 \\ 0 & T_C \leq 0 \text{ or } T_C > 34 \end{cases} \quad [\text{Eq. 9}]$$

$$\Delta V = \min (1.4 - 0.0778T_C, 0.5 + 13.44 \frac{T_C}{(T_{max} - T_{min} + 3)^2}) \quad [\text{Eq. 10}]$$

when

$$T_{max} < 30C \text{ and } T_{min} < 15C$$

## Statistical Analyses

Linear models for observed heading date and each thermal index were assessed using the ‘lme4’ package in R (Bates et al., 2007). The model (Eq. 12) creates a genotype specific thermal time coefficient for each genotype.

$$y_i = \mu + G_i + \varepsilon \quad [\text{Eq. 13}]$$

Where,  $y_i$  is the value for the thermal time of interest,  $\mu$  is the overall mean,  $G_i$  is the fixed effect of the  $i^{th}$  entry (genotype) and  $\varepsilon$  is the residual error. This model requires a second

step where the genotype specific thermal time coefficient is used to predict the heading date based on the calculated cumulative thermal time for that growing season.

The accuracy of the models were evaluated using coefficient of determination ( $R^2$ ), root mean square error (RMSE, Eq. 14), Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC). The AIC and BIC scores were calculated using 'AICcmodavg' package in R (Mazerolle & Mazerolle, 2020).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (OBS_i - SIM_i)^2}{n}} \quad [\text{Eq. 14}]$$

## Results and Discussions

In both the AM Panel and the KSWB program both year and location had significant effects on observed heading date. In the AM Panel only 2018 and 2020 had similar calendar dates for heading (Fig. 2.2). Whereas the heading date for 2019 was significantly later and the 2017 heading date was significantly earlier. Heading duration also varied with 2017 having the longest heading period and 2018 the shortest. Similar trends were observed within the KSWB program; however, unlike the AM Panel, all three years were statistically different. In this case collectively 2018 had the earliest whereas 2019 had the latest heading date. We also observed that the 2020 growing season showed the highest variability of heading date by location, with the Ellsworth County site having the latest heading date of all location years. In total, only the 2020 McPherson experiment had similar heading dates with both the 2019 Reno and the 2018 Ellsworth locations (Fig. 2.2). Across all years and experiments, the average heading dates ranged from the 119<sup>th</sup> to 134<sup>th</sup> DOY (Table 2.5). This range is consistent with heading dates reported for other populations within the US Central Great Plains (Grogan et al., 2016).

The difference in observed heading dates could be explained by fluctuations in monthly average temperatures during the growing season. The 2017 growing season for the AM Panel experienced a warmer fall allowing for a higher accumulation thermal time prior to vernalization (Fig. 2.3). Furthermore, the experiment was planted earlier than subsequent years (Table 2.3) allowing for additional development during the fall. Although December of 2017 was cooler than other years, an early green up caused by elevated temperatures in February allowed for an earlier accumulation of thermal time. Inversely, 2019 experienced a warm October but remained cooler than other years through the fall into spring until normal temperatures resumed from April through July. However, the early low temperatures resulted in a slower accumulation of thermal units and could be the reason for the late observed heading date. It is also notable that the 2018 and 2020 temperature patterns were similar with the exception of higher temperatures around anthesis and grain fill in 2018. This increased heat at anthesis paired with low annual precipitation in 2018 could explain the short heading period observed in 2018. Similar effects of elevated temperature decreasing flowering time have been reported by several studies (Sadras & Monzon, 2006; Wang et al., 2015).

The monthly weather trends were similar for the KSWB program where year variation were stronger than location trends. Like the AM Panel, 2019 experienced the coldest fall/winter with the latest green up in spring (Fig. 2.3 B). Likewise, the elevated temperatures around anthesis were observed at all locations in 2018. The greatest location discrepancy within a year was observed in 2020, where Hutchinson location had higher mean temperatures than either the Lorraine or McPherson locations. This pattern was also show in the heading date data as previously discussed. Overall, the high variability in temperature trends between seasons supports using thermal indices to normalize the plant development.

Trends in cumulative thermal indices were comparable to growing season temperatures patterns. This is expected as all the thermal indices explored for this experiment are calculated based on daily maximum and minimum temperatures. This trend was most apparent for the full season curve for the 2017 AM Panel. Fig. 2.4 (C), shows considerable elevated accumulation of growing degree days (GDD1) during the fall right after sowing. This is constant with the early planting date and the warmer observed monthly mean temperature during the fall. In combination with the mild winter and early green up during that growing season, we would expect a higher accumulation of thermal units. Although the 2017 heading date was also early, the large accumulation of GDD early in the growing season potentially skewed the predicted heading date. This trend was observed for all thermal indices. However, the early and elevated accumulation was minimized with the Pdays equation, which could be expected due to its emphasis on optimum temperature compared to the daily average. The 2019 thermal accumulation for GDD1 was the lowest, which follows the trends of observed heading date and monthly temperatures previously reported. While the GDD1 curves collected after January 1<sup>st</sup> and March 1<sup>st</sup> (Fig. 2.4-A and 2.4-B) do minimize the gap in accumulation, 2017 still has the highest total GDD accumulation. It is noteworthy that in GDD1 accumulation after March 1<sup>st</sup> the accumulation trend appear to be more linear and the overall differences are minimized. However, even though the differences are minimized, the rank order of 2017, 2018, 2020 and 2019 is still present.

Like the observed weather data, the accumulation of GDD in the KSWB program was much more dependent on year than location. This is also supported by the observed heading date data previously discussed. Although the thermal accumulation trends for the KSWB program were not as distinguishable as the 2017 AM Panel, differences still existed particularly over the

full season (Fig. 2.5 C). This trend was especially noticeable with the 2019 Riley County data where total accumulation was lower than all other location-years. Contrary, the Reno County location had highest accumulation within years. This location is the most southern location in all years, except for McPherson County in 2020. The higher average accumulation of GDD at Reno County was likely due to its southern geographic location. Like the AM Panel, differences in accumulation were mitigated for both the curve that started January 1<sup>st</sup> (Fig. 2.5 B) and March 1<sup>st</sup> (Fig. 2.5 A).

Cumulative thermal times at heading date had similar discrepancies as the thermal time accumulation figures. Both the AM Panel (Fig. 2.6 A, B and C) and KSWB program (Figure 2.7 A, B and C) demonstrated similar patterns. In both data sets, the year-locations with earlier observed heading dates had a higher accumulation of thermal units at the time of heading. However, these lines converged to a more linear relationship for both the January 1<sup>st</sup> and March 1<sup>st</sup> accumulation start dates. Within the KSWB program, the 2019 Riley County thermal time observations went from a disjunct low group in the full season accumulation (Fig. 2.7 C) to an average accumulation in the March 1<sup>st</sup> accumulation (Fig. 2.7 A). This follows all previously discussed patterns where the early season cool period limited the accumulation of thermal units. The inverse of this is shown in the AM Panel where the 2017 data has elevated thermal accumulation across the entire season (Fig. 2.6 A) but merges closer to the other years as we move to January and March 1<sup>st</sup> start dates (Fig. 2.6 A and B). These observations reinforce that yearly season accumulation, or lack of accumulation during the vernalization period, affects thermal indices' ability accurately predict phenology.

Evaluation of thermal indices in relationship to the observed heading date indicates that full season thermal time accumulation had higher RMSE when compared DOY or models that

incorporate thermal accumulation after January or March 1<sup>st</sup>. As shown in Table 2.6, the RMSE values in the AM Panel improved for all thermal indices with both the January 1<sup>st</sup> and March 1<sup>st</sup> thermal time calculations. This can be explained by both the overall reduction of accumulation days during the vernalization period and their relationship to previously discussed weather patterns. The fitted versus observed results (Fig. 2.8) for the AM Panel further confirm the advantages of thermal accumulation methods with reduced accumulation periods. We observed that with the full season accumulation of thermal time the model underestimates the observed heading date in 2017 and overestimated heading date in 2019. This was reduced but still prevalent when using Jan 1<sup>st</sup> GDD calculations, and greatly improved with March 1<sup>st</sup> based GDD. This could be expected due to the observed weather patterns and total thermal accumulation curves during the 2017 and 2019 seasons. Overall, the superior performance of post-vernalization calculations indicates that the fall conditions are not a major determining factor in the resulting spring development, consistent with the known biology of winter wheat with its vernalization requirements.

Comparable trends were observed in the KSWB program data. The full season GDD index overestimated heading date in 2019 as some estimated exceed 145 days (Figure. 2.9 C). In contrast, heading dates estimated for both 2020 and 2018 were closer to observed heading dates. Once again, these prediction gaps diminish for GDD calculations based on January 1<sup>st</sup> and March 1<sup>st</sup> (Fig. 2.9 A and B). The main difference in the KSWB program dataset is that the RMSE for each index is lowest for the January 1<sup>st</sup> predictions, whereas the lowest RMSE in the AM Panel was observed with the March 1<sup>st</sup> datasets. As previously discussed, in 2018 there were elevated temperatures around heading date. In the case of the KSWB program, which only had 3 years of

data, these elevated temperatures could have had a larger impact than the shorted March 1<sup>st</sup> data set.

Although the main effects of accumulation period are evident, there are specific thermal indices that had better performance. The PGDD and GDD<sub>WE</sub> thermal indices has the lowest RMSE regardless of the accumulation period in both data sets. This could be due to the nonlinear response function that is used to calculate these indices. It is also notable that the introduction of photo period in the PGDD equations tended to reduce RMSE, increase R<sup>2</sup>, and improve AIC and BIC selection criteria. Additionally, differences were observed in all three GDD equations with the most significant RMSE difference of 1.5 days between full season accumulations of GDD1 and GDD2. These results are also supported from previous studies (McMaster & Wilhelm, 1997) and justify the need to standardize calculations within a breeding program.

APSIM model simulations were only made for full growing season with five distinct coefficients observed for each genotype. Genotype-specific coefficient differences were observed for all coefficients for both the AM Panel (Table 2.8) and the KSWB program (Table 2.9). However, for both sets of germplasm, only the end of juvenile thermal time coefficient was statistically significantly different. The end of juvenile coefficient relates to a genotype's response to both early season green up and initiation of transition to reproductive stages. Increased accumulation of thermal units to end this stage would have significant impacts on the genotype's phenology progression throughout a growing season. Within the AM Panel, the genotype Alliance had the highest thermal time requirement to end juvenile phase whereas Fuller had the lowest. These results for all genotypes are supported by published breeder agronomic and target geographic region for production (Fritz, 2007; Watson, 2002). Likewise the KSWB program germplasm showed differences in juvenile coefficients where the breeding line

KS120129M~4 had the highest requirement and four genotypes had the lowest coefficient amongst them was WB4269 which has an early maturity (WestBred, 2018). We did not observe significant differences for photoperiod or vernalization coefficients in the breeding trials. This could be expected as all the released evaluated genotypes were adapted for production in Kansas and the experimental genotypes are being developed for release, thus being well-tuned to the needed vernalization period for Kansas environments.

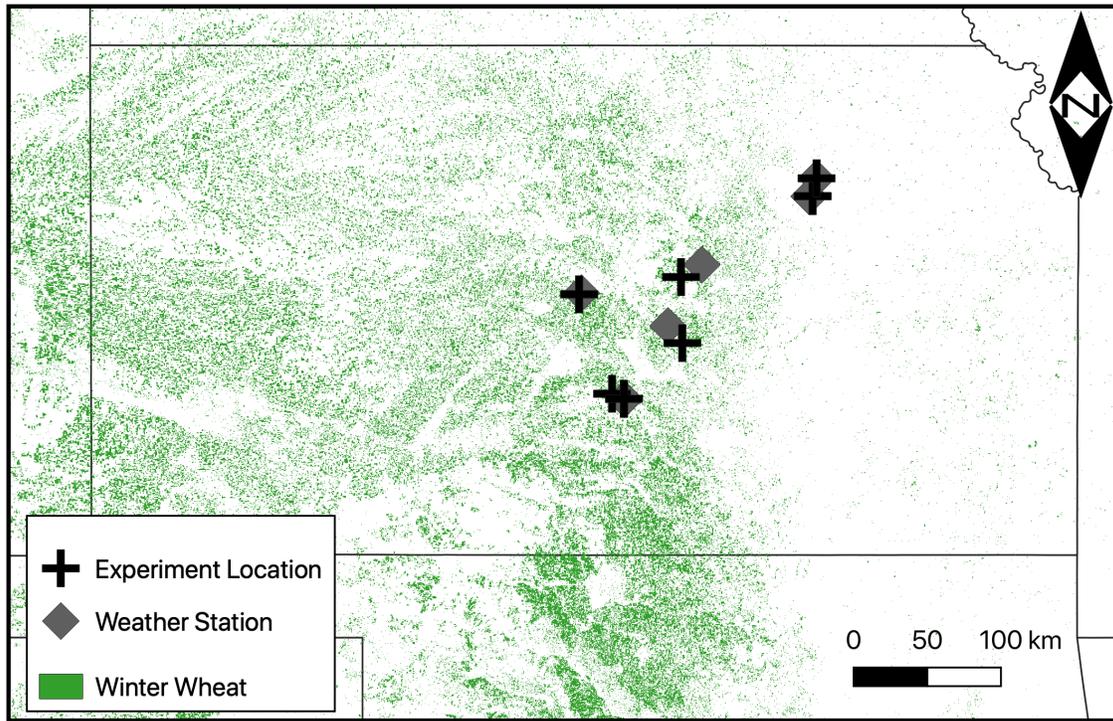
Since APSIM simulations were across the entire growing season, all predictions were made in DAS then normalized to DOY accounting for sowing dates. Thus, the predictions would be consistent with full season thermal time indices. With both the AM panel and KSWB datasets, the APSIM simulations reduced RMSE while increasing  $R^2$  values (Fig. 2.10). Overall, the AM Panel RMSE of 1.812 days and the KSWB program RMSE of 2.405 days outperformed any of the full season individual thermal time indices. However, the full season APSIM for the KSWB program did not outperform the Jan 1 and Mar 1 condensed thermal indices for GDD, PGDD and  $GDD_{WE}$  (Table 2.7). Similarly to the full season thermal indices, the APSIM model over-predicted phenology in 2019 due to the cool winter, which accounted for most of the reduced accuracy and why the condensed temporal indices performed better. However, considering that APSIM allows for full season prediction, the relative performance of the model is still better. The higher observed RMSE observed in the KSWB maybe caused by compounding both location and years.

Overall, the previously discussed weather pattern effects during the vernalization period were not observed with the APSIM model, which can likely be attributed to the phase change calculations within the APSIM model. The model's ability to combine genotype specific thermal time coefficients with thermal times accumulation phase changes likely provides the advantage

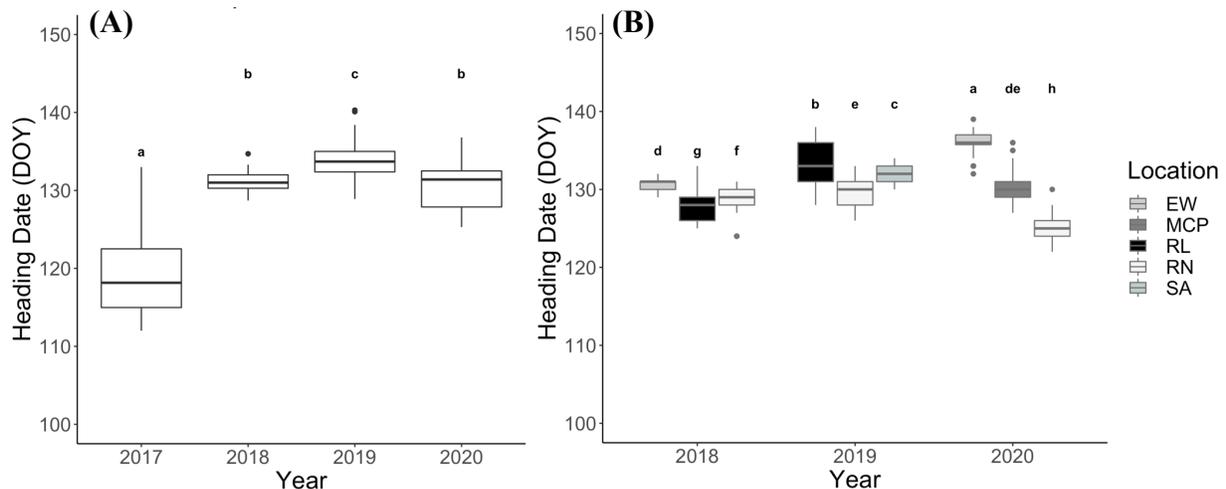
over other thermal indices. Additionally, as yield prediction models are developed, there may be a need to incorporate early season plant vigor phenotyping, which would limit the applications of the condensed temporal models (Kipp et al., 2014).

## **Conclusions**

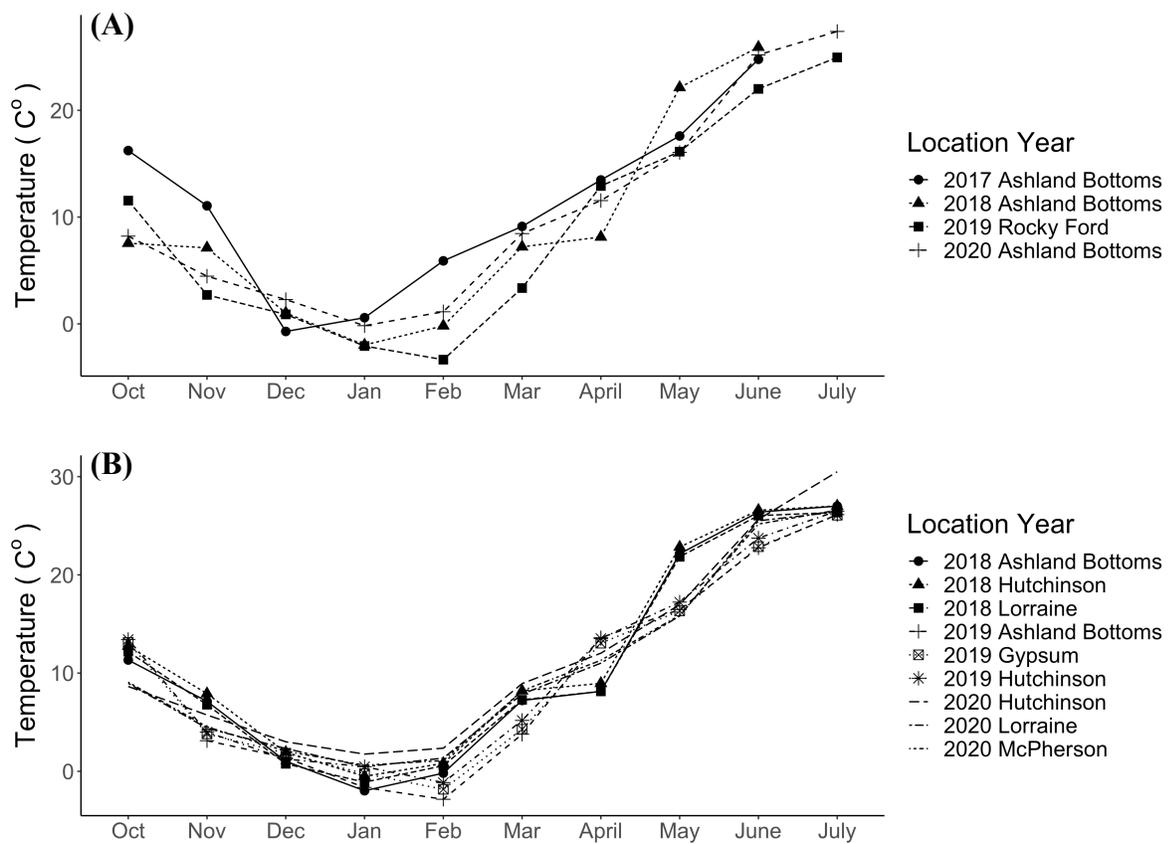
This study supports the use of thermal time to predict phenology within a breeding program where there is more genetic diversity than in a production setting. However, there are several challenges of using thermal indices alone, particularly when using them to predict phenology for an entire growing season. While models that used start dates after the initiation of vernalization had stronger performance, there still is some novelty for breeding selection in predicting the entire growing season particularly as technologies and models advance and the potential need of full season data. In this study, it was demonstrated that multi-phase models, such as the APSIM model, have distinct advantages over thermal indices alone. This in combination with genotype-specific coefficients in season phenology can reasonably be predicted with these models. Future research in phenology modeling is warranted, especially as sensor technology continues to evolve and new analytical techniques such as machine learned are paired with this HTP data.



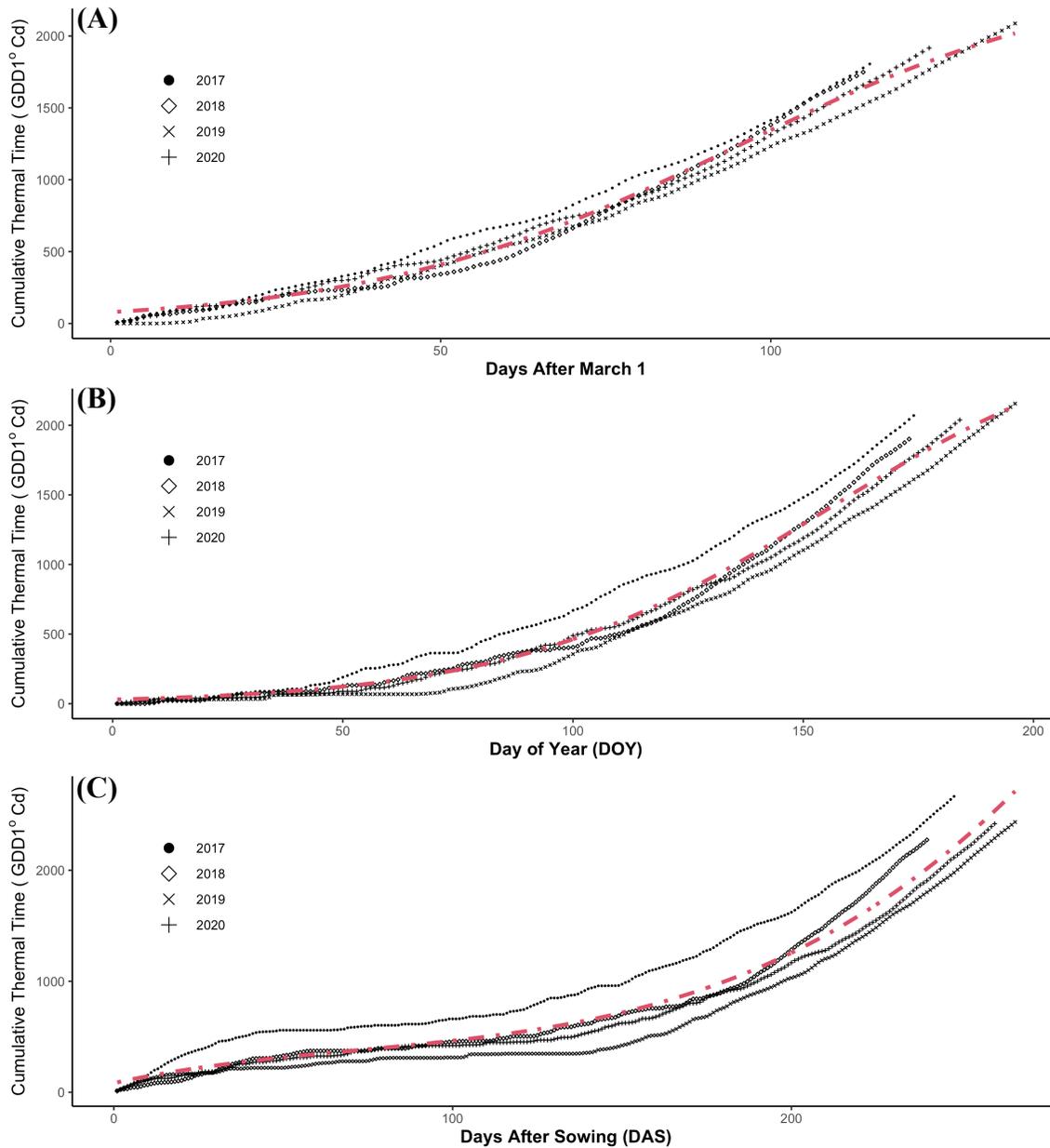
**Figure 2.1.** Experiment and weather station locations for both the AM Panel and Kansas Wheat Breeding experiments. The shapes represent the weather station locations in relationship to the experiment locations. The green polygon represents winter wheat acres planted in 2021



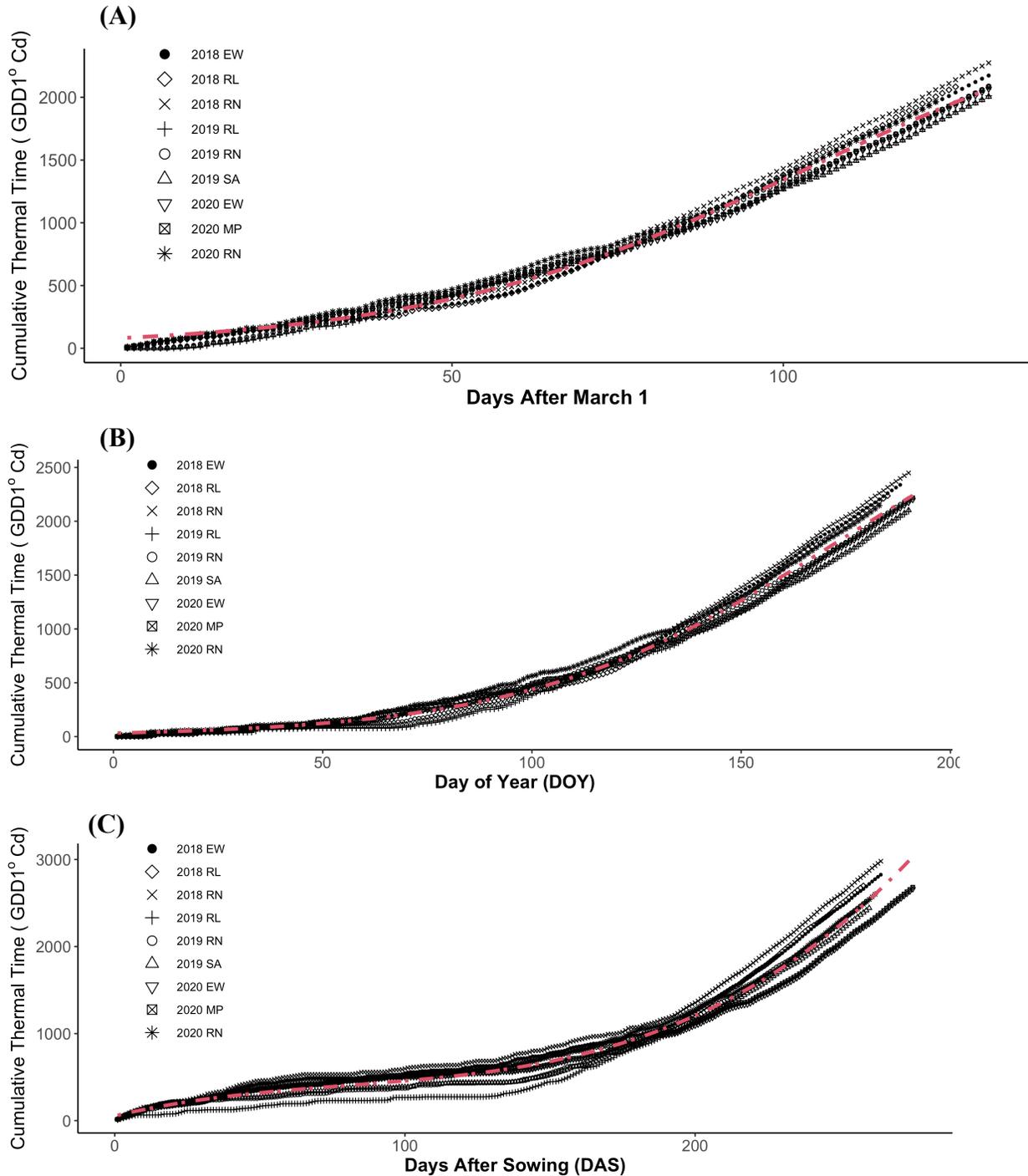
**Figure 2.2.** Observed heading date distribution by location year for the AM Panel (A) and the Kansas State University Wheat Breeding Program (B). The x-axis is the harvest year growing season and the y-axis is the observed heading date in day of year (DOY) after Jan 1<sup>st</sup>. Colors represent location for the Kansas State University Wheat Breeding Program data.



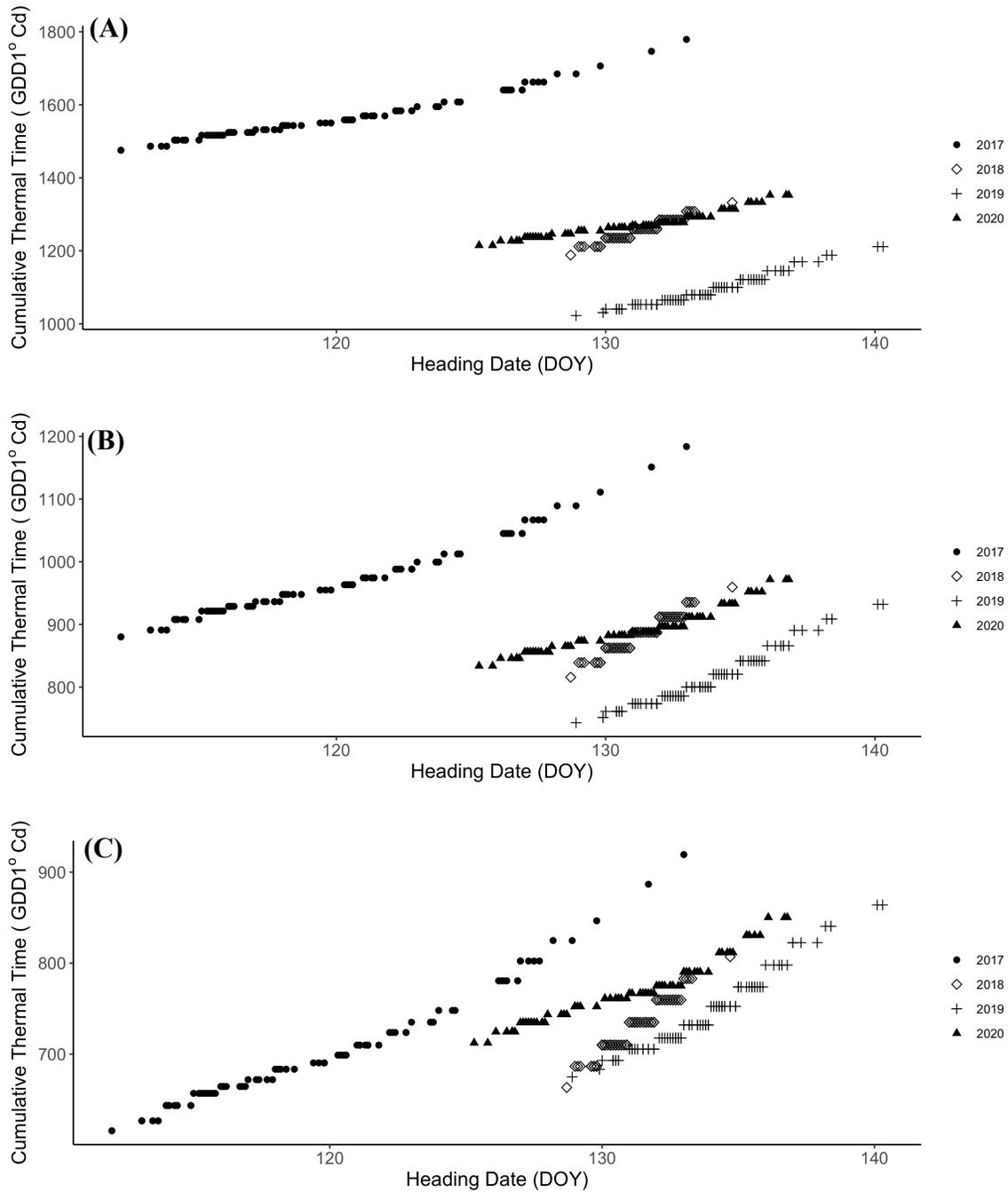
**Figure 2.3** Mean monthly temperature by location year for the AM Panel (A) and the KS Wheat Breeding Program (B). The x-axis is the month during the growing season and the y-axis is the recorded mean monthly temperature from the Kansas Mesonet weather stations. Symbol and line combinations represent the location and year the data was recorded.



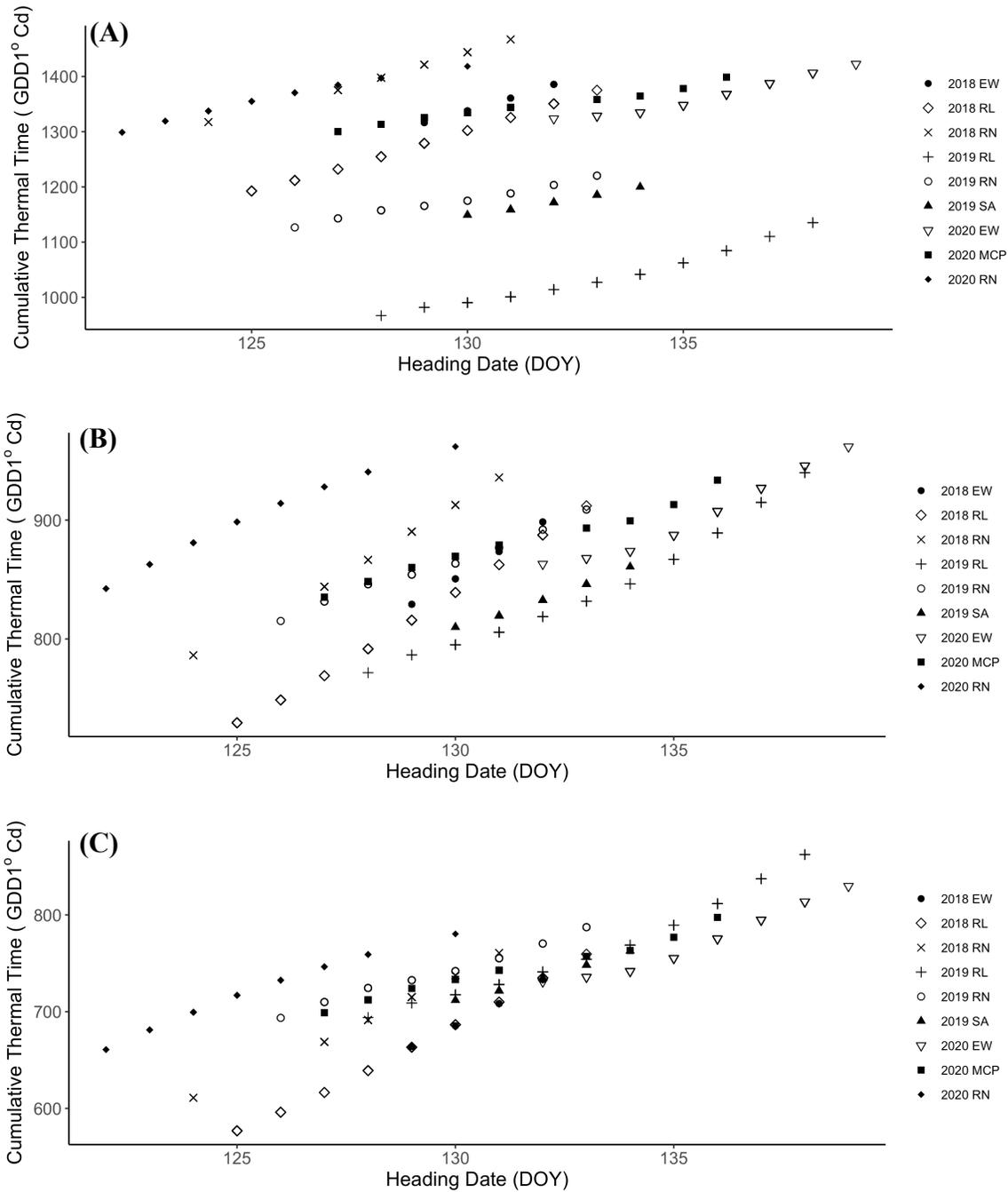
**Figure 2.4.** Cumulative Growing Degree Days by year and location fit with a logistic regression model for the AM Panel. The x-axis is day of year (DOY) after March 1<sup>st</sup> (A), January 1<sup>st</sup> (B) and days after sowing (DAS)(C). The y-axis is the cumulative thermal time calculated using GDD1. The shapes denote the individual location years and the red line is the fitted trend for all seasons. The fit lines for both figure A and B are logistic regression models and figure C is a polynomial model.



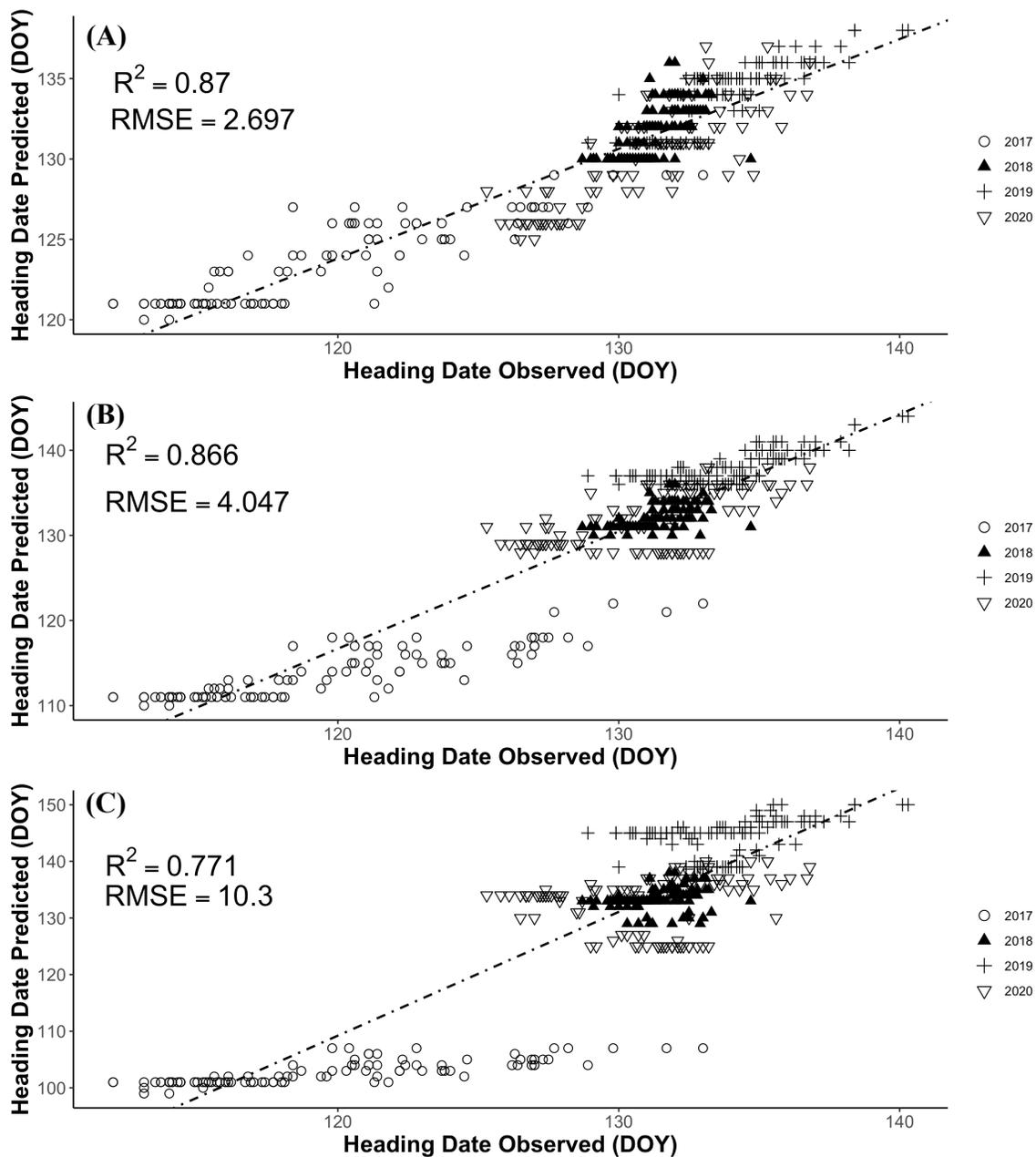
**Figure 2.5.** Cumulative Growing Degree Days by year and location fit with a logistic regression model for the Kansas State University Wheat Breeding program. The x-axis is day of year (DOY) after March 1<sup>st</sup> (A), January 1<sup>st</sup> (B) and days after sowing (DAS)(C). The y-axis is the cumulative thermal time calculated using GDD1. The shapes denote the individual location years and the red line is the fitted trend for all seasons. The fit lines for both figure A and B are logistic regression models and figure C is a polynomial model.



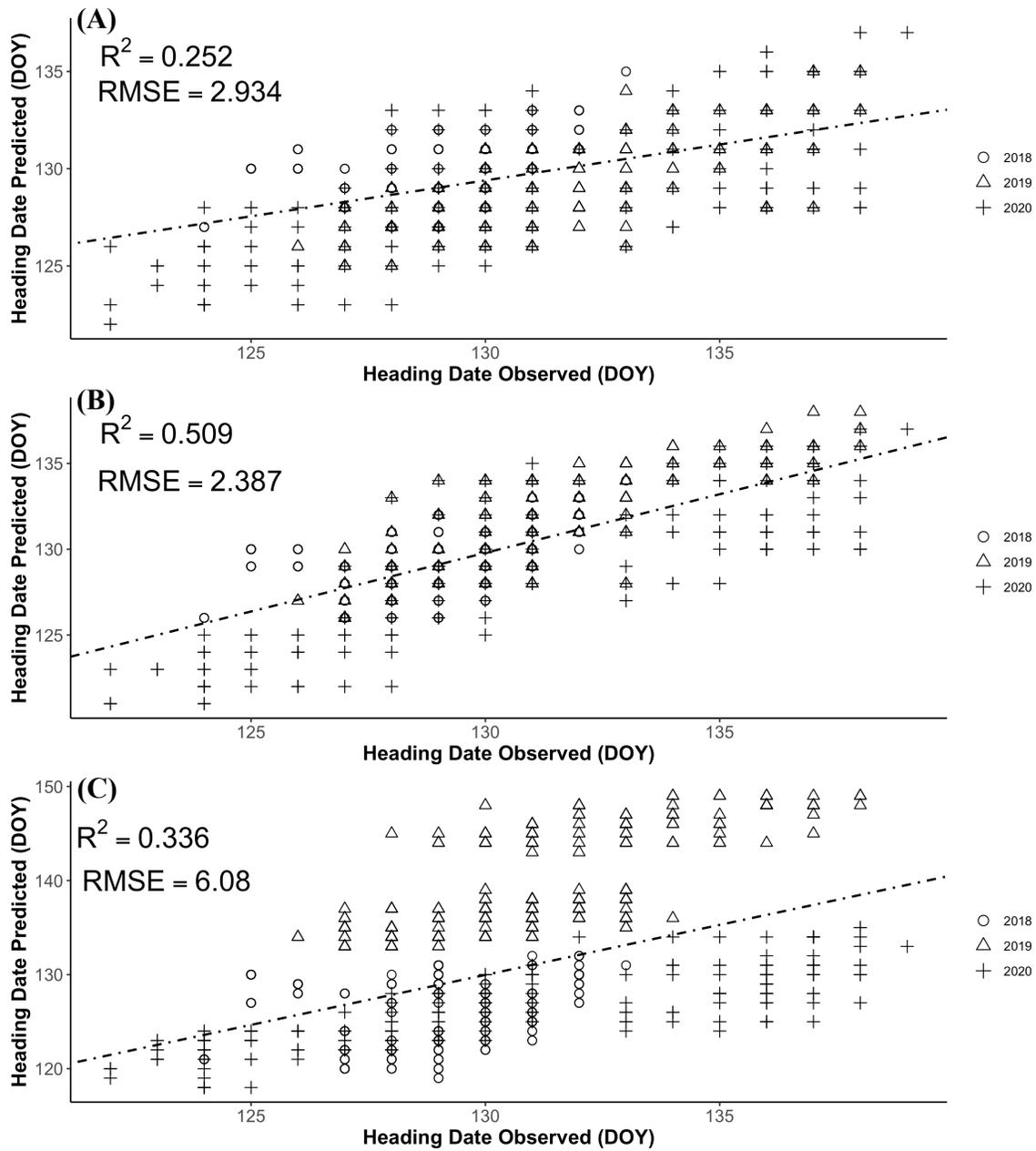
**Figure 2.6.** Growing Degree Day accumulation at observed heading date by year for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> accumulation periods within the AM Panel. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the cumulative thermal time calculated using GDD1.



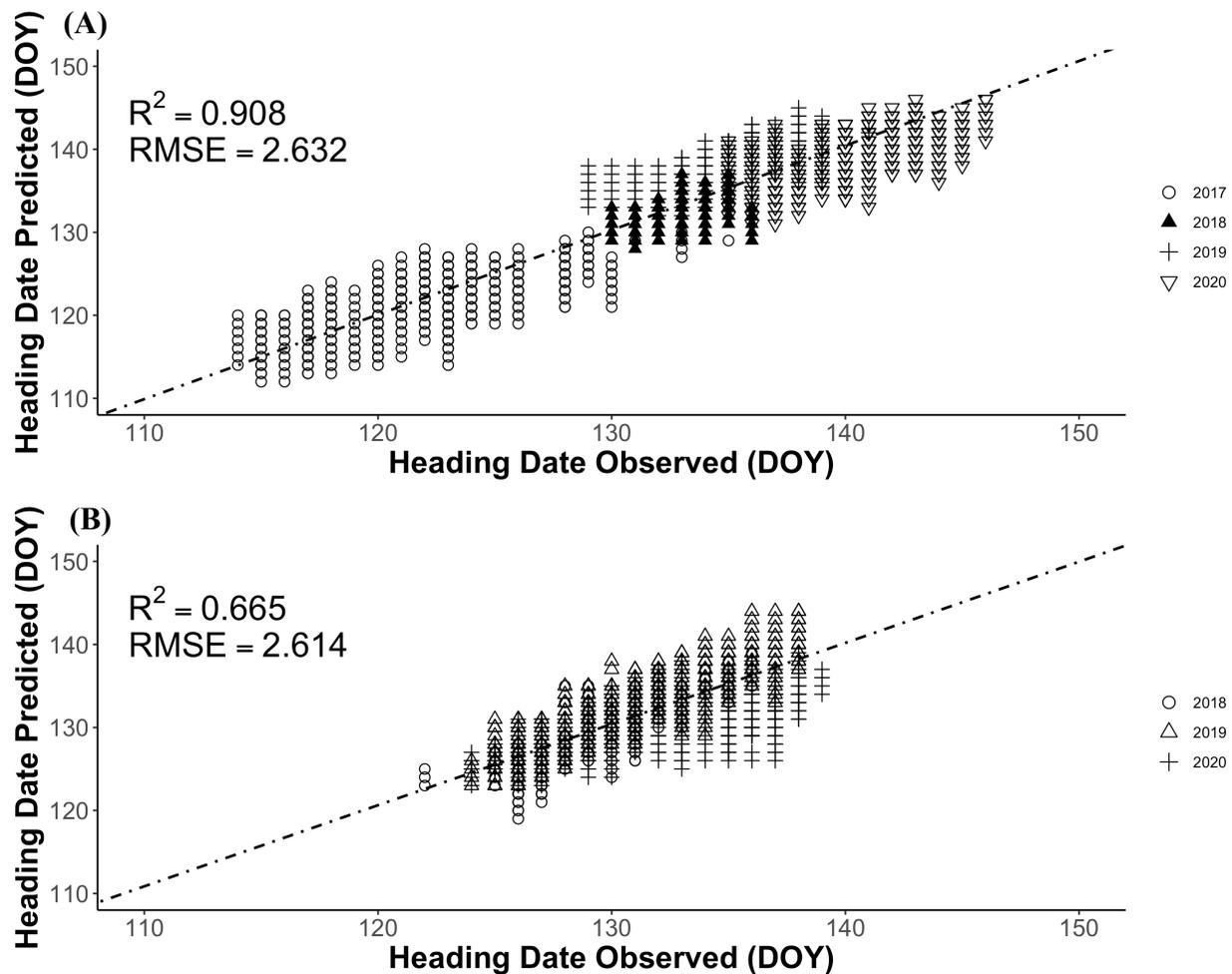
**Figure 2.7.** Growing Degree Day accumulation at observed heading date by year and location for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> accumulation periods within the Kansas State University Wheat Breeding program. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the cumulative thermal time calculated using GDD1.



**Figure 2.8.** Growing Degree Days (GDD1) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods within the AM Panel. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the GDD1 predicted heading date, DOY.



**Figure 2.9.** Growing Degree Days (GDD1) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods within the Kansas State University Wheat Breeding program. The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the GDD1 predicted heading date, DOY.



**Figure 2.10.** APSIM Predicted vs Observed Heading Dates for the AM Panel (A) and the Kansas State University Wheat Breeding program (B). The x-axis is the observed heading date in day of year (DOY) after January 1<sup>st</sup> and the y-axis is the APSIM predicted heading date, DOY. Shape differences represent sample year and the line is the linear model fitted trend line.

**Table 2.1-**Association Mapping Panel Entries for 2017, 2018, 2019 and 2020 experiments located at the Kansas State University Ashland Bottoms and Rocky Ford Research Stations.

Variety	Year of Release	Variety	Year of Release	Variety	Year of Release
1863	2012	Heyne	1998	PostRock	2006
2137	1995	Ike	1993	RonL	2006
2145	2001	Jagalene	2001	Santa Fe	2003
AG Icon	2017	Jagger	1994	Stanton	2000
Alliance	1994	Joe	2015	SY GOLD	2010
Arlin	1992	KanMark	2014	SY Wolf	2011
Armour	2008	Karl 92	1992	Tatanka	2017
Art	2007	Lakin	2000	Trego	1999
Bob Dole	2018	Larned	1976	WB-Cedar	2011
Clara CL	2011	Larry	2016	WB-Grainfield	2012
Danby	2005	LCS Chrome	2016	WB4458	2013
Everest	2009	LCS Mint	2012	Winterhawk	2007
Fuller	2006	Longhorn	1990	Zenda	2017
		Overley	2003		

**Table 2.2-** Kansas Wheat Breeding Program for 2018, 2019 and 2020 experiments located in Ellsworth, McPherson, Saline, Reno, and Riley Counties. Included established and experimental genotypes.

	Variety	
Bob Dole	KS100060K-19	KS13DH0008-23
Everest	KS100509K-2	KS13DH0030-28
Gallagher	KS120081M~5	KS13DH0035-66
Joe	KS120125M~9	KS13DH0041-35
KS090049K-8	KS120129M~4	KS14HW106-6-6
KS090387K-20	KS120252M~14	Larry
KS090413K-4	KS120506M~7	NUSAKA15-3
KS090438K-9	KS120513M~5	SY Monument
KS090616K-1	KS120559M~5	WB4269
KS100028K-10	KS120648M~5	WB4458
KS100028K-11	KS12DH0090-172	Zenda
	KS12DH0156-88	

**Table 2.3-** Field experimental locations and weather station details for the Association Mapping Panel experiments.

Year	Planting Location	Location Coordinates	Planting Date	Harvest Date	Weather Station ID	Weather Station Coordinates
2017	Ashland Bottoms Experiment Station	39°08'19.07" N, 96°38'21.00" W	10/19/16	6/23/17	ASBK1	39° 07' 32.78"N, 96° 38' 11.50"W
2018	Ashland Bottoms Experiment Station	39°08'19.07" N, 96°38'21.00" W	10/26/17	6/27/18	ASBK1	39° 07' 32.78"N, 96° 38' 11.50"W
2019	Rocky Ford Experiment Station	39°13'48.82" N, 96°34'41.87" W	10/23/18	7/15/19	RKFK1	39° 07' 32.78"N, 96° 38' 11.50"W
2020	Ashland Bottoms Experiment Station	39°08'19.07" N, 96°38'21.00" W	10/24/18	7/2/20	ASBK1	39° 07' 32.78"N, 96° 38' 11.50"W

**Table 2.4-** Field experimental locations and weather station details for the Kansas Wheat Breeding Program experiments.

Year	Planting Location	Location Coordinates	Planting Date	Harvest Date	Weather Station ID	Weather Station Coordinates
2018	Ellsworth County (EW)	38° 33' 4.13"N, 98° 22' 7.67"W	10/17/17	6/27/18	LORK1	38° 33' 28.72"N, 98° 21' 07.30"W
2018	Riley County (RL)	39° 7' 37.34"N, 96° 36' 39.93"W	10/20/17	6/23/18	ASBK1	39° 7' 32.7828"N, 96° 38' 11.508"W
2018	Reno County (RN)	37° 55' 58.53"N, 98° 01' 53.91"W	10/18/17	6/28/18	PRGK1	37° 55' 51.56"N, 98° 01' 37.49"W
2019	Riley County (RL)	39° 7' 38.27"N, 96° 36' 36.56"W	11/1/18	7/16/19	ASBK1	39° 7' 32.7828"N, 96° 38' 11.508"W
2019	Reno County (RN)	37° 57' 46.64"N, 98° 07' 14.89"W	10/24/18	6/30/19	PRGK1	37° 55' 51.56"N, 98° 01' 37.49"W
2019	Saline County (SA)	38° 39' 10.54"N, 97° 36' 4.417"W	10/23/18	6/18/18	GYMK1	38° 43' 30.79"N, 97° 26' 38.94"W
2020	Ellsworth County (EW)	38° 33' 4.13"N, 98° 22' 7.67"W	10/9/19	6/30/20	LORK1	38° 33' 28.72"N, 98° 21' 07.30"W
2020	McPherson (MP)	38° 15' 48.18"N, 97° 35' 30.01"W	10/14/19	6/26/20	MPRK1	38° 21' 51.33"N, 97° 41' 59.99"W
2020	Reno County (RN)	37° 55' 58.40"N, 98° 01' 49.41"W	10/14/19	6/22/20	PRGK1	37° 55' 51.56"N, 98° 01' 37.49"W

**Table 2.5-** Average heading date, day of year (DOY) by experiment, location, and year for both the AM Panel and the Kansas Wheat Breeding Program.

Year	Location	Experiment	Average Heading Date (DOY)	SD (DOY)
2017	Ashland Bottoms	AM Panel	119.3	5.0
2018	Ashland Bottoms	AM Panel	131.1	1.1
2019	Rocky Ford	AM Panel	133.8	2.2
2020	Ashland Bottoms	AM Panel	130.7	2.8
2018	Ellsworth County (EW)	KSWB	130.6	2.0
2020	Ellsworth County (EW)	KSWB	136.1	4.7
2020	McPherson (MP)	KSWB	130.0	4.7
2018	Riley County (RL)	KSWB	127.9	2.0
2019	Riley County (RL)	KSWB	133.4	2.6
2018	Reno County (RN)	KSWB	128.8	2.0
2019	Reno County (RN)	KSWB	129.6	2.6
2020	Reno County (RN)	KSWB	125.0	4.7
2019	Saline County (SA)	KSWB	131.9	2.6

**Table 2.6-** Thermal time index performance of AM Panel across all years and locations

Thermal Index	RMSE	R <sup>2</sup>	BIC	BIC Rank	AIC	AIC Rank
BMT	5.41	0.625	2827.36	19	2814.97	19
DOY	5.41	0.221	2926.32	20	2913.92	20
GDD	10.3	0.771	2102.65	2	2090.26	2
GDD Jan1	4.05	0.866	2090.26	1	2077.87	1
GDD Mar1	2.7	0.870	2352.87	12	2340.47	12
GDD Tmax	9.57	0.771	2125.61	4	2113.22	4
GDD Tmax Jan1	3.88	0.859	2139.72	6	2127.32	6
GDD Tmax Mar1	2.97	0.855	2354.42	13	2342.02	13
GDD2	9.57	0.771	2344.91	11	2332.51	11
GDD2 Jan1	3.9	0.859	2743.49	18	2731.1	18
GDD2 Mar1	2.97	0.855	2379.16	16	2366.76	16
Pday	11.1	0.758	2199.32	8	2186.93	8
Pday Jan1	5.92	0.775	2206.04	9	2193.65	9
Pday Mar1	4.95	0.473	2355.26	15	2342.87	15
PGDD	8.33	0.770	2125.26	3	2112.86	3
PGDD Jan1	3.35	0.835	2139.96	7	2127.56	7
PGDD Mar1	3.28	0.833	2354.42	14	2342.02	14
GDD <sub>WE</sub>	11	0.754	2240.82	10	2228.43	10
GDD <sub>WE</sub> Jan1	4.35	0.820	2125.63	5	2113.23	5
GDD <sub>WE</sub> Mar1	2.76	0.859	2387.15	17	2374.76	17

**Table 2.7-** Thermal time index performance of Kansas State Wheat Breeding program across all years and locations.

<b>Thermal Index</b>	<b>RMSE</b>	<b>R<sup>2</sup></b>	<b>BIC</b>	<b>BIC Rank</b>	<b>AIC</b>	<b>AIC Rank</b>
BMT	3.86	0.222	2827.36	19	2814.97	19
DOY	2.91	0.153	2926.32	20	2913.92	20
GDD	6.08	0.336	2352.87	12	2340.47	12
GDD Jan1	2.39	0.509	2102.65	2	2090.26	2
GDD Mar1	2.93	0.252	2090.26	1	2077.87	1
GDD Tmax	6.98	0.259	2354.42	13	2342.02	13
GDD Tmax Jan1	2.89	0.403	2125.61	4	2113.22	4
GDD Tmax Mar1	2.61	0.336	2139.72	6	2127.32	6
GDD2	6.99	0.257	2379.16	16	2366.76	16
GDD2 Jan1	2.91	0.402	2344.91	11	2332.51	11
GDD2 Mar1	2.56	0.359	2743.49	18	2731.1	18
Pday	9.82	0.234	2355.26	15	2342.87	15
Pday Jan1	5.04	0.186	2199.32	8	2186.93	8
Pday Mar1	5.24	0.169	2206.04	9	2193.65	9
PGDD	3.23	0.477	2354.42	14	2342.02	14
PGDD Jan1	2.18	0.574	2125.26	3	2112.86	3
PGDD Mar1	2.37	0.471	2139.96	7	2127.56	7
GDD <sub>WE</sub>	4.26	0.449	2387.15	17	2374.76	17
GDD <sub>WE</sub> Jan1	2.14	0.602	2240.82	10	2228.43	10
GDD <sub>WE</sub> Mar1	2.34	0.491	2125.63	5	2113.23	5

**Table 2.8-** APSIM thermal time accumulation, photo period sensitivity and vernalization sensitivity coefficients by genotype for the AM Panel

<b>Genotype</b>	<b>End of Juvenile TT</b>	<b>Floral Initiation TT</b>	<b>Start of Grain Fill TT</b>	<b>Photoperiod Sensitivity</b>	<b>Vernalization Sensitivity</b>
2137	320	580	420	3.9	4.0
2145	440	360	300	4.0	4.0
AGIcon	480	320	720	4.0	3.6
ALLIANCE	460	380	400	4.0	3.9
ARLIN	340	520	540	4.0	2.8
ARMOUR	300	620	900	4.0	2.9
Art	420	360	640	4.0	3.9
BOB DOLE	260	680	820	4.0	3.9
CLARA-CL	500	300	460	4.0	2.8
DANBY	440	360	300	4.0	4.0
EVEREST	260	620	640	4.0	3.7
FULLER	440	340	300	4.0	3.3
HEYNE	400	480	600	3.9	4.0
IKE	400	480	600	3.9	4.0
JAGALENE	340	560	380	4.0	3.8
JAGGER	380	400	340	4.0	4.0
KARL-92	200	800	720	4.0	3.6
LAKIN	440	360	300	4.0	4.0
LARNED	460	420	320	4.0	3.7
LARRY	320	580	420	3.9	4.0
LCS-MINT	480	320	720	4.0	3.6
LONGHORN	420	380	700	4.0	4.0
OVERLEY	320	480	900	4.0	3.7
POSTROCK	400	480	600	3.9	4.0
RONL	220	780	500	4.0	3.9
SANTA-FE	420	360	640	4.0	3.9
STANTON	500	300	460	4.0	2.8
SY-GOLD	380	460	720	4.0	4.0
SY-WOLF	500	300	460	4.0	2.8
TATANKA	260	680	820	4.0	3.9
TREGO	300	620	460	4.0	3.2
WB-CEDAR	200	760	340	4.0	4.0
WB-GRAINFIELD	420	420	680	4.0	3.7
WB4458	420	360	640	4.0	3.9
WINTERHAWK	360	520	340	4.0	2.8
ZENDA	260	680	820	4.0	3.9

**Table 2.9-** APSIM thermal time accumulation, photo period sensitivity and vernalization sensitivity coefficients by genotype for the Kansas Wheat Breeding Program

<b>Genotype</b>	<b>End of Juvenile TT</b>	<b>Floral Initiation TT</b>	<b>Start of Grain Fill TT</b>	<b>Photoperiod Sensitivity</b>	<b>Vernalization Sensitivity</b>
BOB DOLE	420	440	900	3.8	4.0
EVEREST	420	320	860	4.0	4.0
GALLAGHER	420	400	700	3.9	3.9
JOE	380	440	300	4.0	4.0
KS090049K-8	440	300	620	4.0	3.9
KS090387K-20	360	460	560	3.9	4.0
KS090413K-4	340	460	400	3.9	4.0
KS090438K-9	360	400	460	4.0	4.0
KS090616K-1	400	480	600	3.9	4.0
KS100028K-10	420	440	900	3.8	4.0
KS100028K-11	420	440	900	3.8	4.0
KS100060K-19	300	560	680	4.0	3.8
KS100509K-2	340	460	400	3.9	4.0
KS120081M~5	420	400	700	3.9	3.9
KS120125M~9	380	500	360	3.9	3.6
KS120129M~4	280	620	400	4.0	4.0
KS120252M~14	400	440	900	4.0	3.9
KS120506M~7	500	320	400	4.0	3.7
KS120513M~5	280	720	400	3.8	4.0
KS120559M~5	420	440	900	3.8	4.0
KS120648M~5	400	420	560	3.9	4.0
KS12DH0090-172	400	420	560	3.9	4.0
KS12DH0156-88	320	600	740	4.0	3.8
KS13DH0008-23	400	440	900	4.0	3.9
KS13DH0030-28	380	460	840	4.0	3.8
KS13DH0035-66	300	540	580	4.0	4.0
KS13DH0041-35	420	440	900	3.8	4.0
KS14HW106-6-6	420	400	700	3.9	3.9
LARRY	360	460	560	3.9	4.0
NUSAKA15-3	420	380	700	4.0	4.0
SY-MONUMENT	420	440	900	3.8	4.0
WB4269	440	340	420	3.9	4.0
WB4458	340	440	680	4.0	4.0
ZENDA	420	440	900	3.8	4.0

## References

- Aslam, M. A., Ahmed, M., Stockle, C. O., Higgins, S. S., ul Hassan, F., & Hayat, R. (2017). Can Growing Degree Days and Photoperiod Predict Spring Wheat Phenology? *Frontiers in Environmental Science*, 5. <https://doi.org/ARTN> 57  
10.3389/fenvs.2017.00057
- Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R package version*, 2(1), 74.
- Bergkamp, B., Impa, S., Asebedo, A., Fritz, A., & Jagadish, S. K. (2018). Prominent winter wheat varieties response to post-flowering heat stress under controlled chambers and field based heat tents. *Field Crops Research*, 222, 143-152.
- Ceglar, A., van der Wijngaart, R., de Wit, A., Lecerf, R., Boogaard, H., Seguini, L., van den Berg, M., Toreti, A., Zampieri, M., Fumagalli, D., & Baruth, B. (2019). Improving WOFOST model to simulate winter wheat phenology in Europe: Evaluation and effects on yield. *Agricultural Systems*, 168, 168-180. <https://doi.org/10.1016/j.agsy.2018.05.002>
- Crain, J., Reynolds, M., & Poland, J. (2017). Utilizing High-Throughput Phenotypic Data for Improved Phenotypic Selection of Stress-Adaptive Traits in Wheat. *Crop Science*, 57(2), 648-659.
- Eversmeyer, M., & Kramer, C. (2000). Epidemiology of wheat leaf and stem rust in the central great plains of the USA. *Annual review of Phytopathology*, 38(1), 491-513.
- Fritz, A., Martin, J., Shroyer, J. (2007). *Fuller Hard Red Wheat*. <https://bookstore.ksre.ksu.edu/pubs/L927.pdf>
- Giri, A. (2019). *Wheat improvement for heat and drought stress tolerance*. <https://krex.k-state.edu/dspace/handle/2097/39551>
- Grogan, S. M., Anderson, J., Baenziger, P. S., Frels, K., Guttieri, M. J., Haley, S. D., Kim, K. S., Liu, S. Y., McMaster, G. S., Newell, M., Prasad, P. V. V., Reid, S. D., Shroyer, K. J., Zhang, G. R., Akhunov, E., & Byrne, P. F. (2016). Phenotypic Plasticity of Winter Wheat Heading Date and Grain Yield across the US Great Plains. *Crop Science*, 56(5), 2223-2236. <https://doi.org/10.2135/cropsci2015.06.0357>
- Halse, N., & Weir, R. (1970). Effects of vernalization, photoperiod, and temperature on phenological development and spikelet number of Australian wheat. *Australian Journal of Agricultural Research*, 21(3), 383-393.
- Hassan, M. A., Yang, M. J., Rasheed, A., Yang, G. J., Reynolds, M., Xia, X. C., Xiao, Y. G., & He, Z. H. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, 282, 95-103. <https://doi.org/10.1016/j.plantsci.2018.10.022>

- Herndl, M., White, J. W., Hunt, L., Graeff, S., & Claupein, W. (2008). Field-based evaluation of vernalization requirement, photoperiod response and earliness per se in bread wheat (*Triticum aestivum* L.). *Field Crops Research*, *105*(3), 193-201.
- Hodges, T., & Ritchie, J. (1991). The CERES-Wheat phenology model. *Predicting crop phenology*, *133*, 141.
- Hossain, I., Epplin, F. M., & Krenzer, E. G. (2003). Planting date influence on dual-purpose winter wheat forage yield, grain yield, and test weight. *Agronomy Journal*, *95*(5), 1179-1188.
- Hu, Q., Weiss, A., Feng, S., & Baenziger, P. S. (2005). Earlier winter wheat heading dates and warmer spring in the US Great Plains. *Agricultural and Forest Meteorology*, *135*(1-4), 284-290. <https://doi.org/10.1016/j.agrformet.2006.01.001>
- Kipp, S., Mistele, B., Baresel, P., & Schmidhalter, U. (2014). High-throughput phenotyping early plant vigour of winter wheat. *European Journal of Agronomy*, *52*, 271-278.
- Kiss, T., Balla, K., Cseh, A., Berki, Z., Horváth, Á., Vida, G., Veisz, O., & Karsai, I. (2021). Assessment of the genetic diversity, population structure and allele distribution of major plant development genes in bread wheat cultivars using DArT and gene-specific markers. *Cereal Research Communications*, 1-9.
- Lollato, R. P., Bavia, G. P., Perin, V., Knapp, M., Santos, E. A., Patrignani, A., & DeWolf, E. D. (2020). Climate-risk assessment for winter wheat using long-term weather data. *Agronomy Journal*, *112*(3), 2132-2151. <https://doi.org/10.1002/agj2.20168>
- Lollato, R. P., & Edwards, J. T. (2015). Maximum attainable wheat yield and resource-use efficiency in the southern Great Plains. *Crop Science*, *55*(6), 2863-2876.
- Maeoka, R. E., Sadras, V. O., Ciampitti, I. A., Diaz, D. R., Fritz, A. K., & Lollato, R. P. (2020). Changes in the phenotype of winter wheat varieties released between 1920 and 2016 in response to in-furrow fertilizer: Biomass allocation, yield, and grain protein concentration. *Frontiers in Plant Science*, *10*, 1786.
- Mazerolle, M. J., & Mazerolle, M. M. J. (2020). Package ‘AICcmodavg’. *CRAN.[Google Scholar]*.
- McMaster, G. S., White, J. W., Hunt, L. A., Jamieson, P. D., Dhillon, S. S., & Ortiz-Monasterio, J. I. (2008). Simulating the Influence of Vernalization, Photoperiod and Optimum Temperature on Wheat Developmental Rates. *Annals of Botany*, *102*(4), 561-569. <https://doi.org/10.1093/aob/mcn115>
- McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, *87*(4), 291-300. [https://doi.org/10.1016/S0168-1923\(97\)00027-0](https://doi.org/10.1016/S0168-1923(97)00027-0)

- Munaro, L., Hefley, T., DeWolf, E., Haley, S., Fritz, A., Zhang, G., Haag, L., Schlegel, A., Edwards, J., & Marburger, D. (2020). Exploring long-term variety performance trials to improve environment-specific genotype $\times$  management recommendations: A case-study for winter wheat. *Field Crops Research*, 255, 107848.
- Parida, B. R., & Ranjan, A. K. (2019). Wheat acreage mapping and yield prediction using landsat-8 OLI satellite data: A case study in Sahibganj Province, Jharkhand (India). *Remote Sensing in Earth Systems Sciences*, 2(2), 96-107.
- Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. *Journal of Atmospheric and Oceanic Technology*, 37(12), 2167-2183.  
<https://doi.org/10.1175/Jtech-D-19-0214.1>
- Porter, J. R., & Gawith, M. (1999). Temperatures and the growth and development of wheat: a review. *European Journal of Agronomy*, 10(1), 23-36.
- Rezaei, E. E., Siebert, S., Hugging, H., & Ewert, F. (2018). Climate change effect on wheat phenology depends on cultivar change. *Scientific Reports*, 8. <https://doi.org/ARTN 489110.1038/s41598-018-23101-2>
- Rife, T. W., & Poland, J. A. (2014). Field Book: An Open-Source Application for Field Data Collection on Android. *Crop Science*, 54(4), 1624-1627.  
<https://doi.org/10.2135/cropsci2013.08.0579>
- Robertson, G. (1968). A biometeorological time scale for a cereal crop involving day and night temperatures and photoperiod. *International Journal of Biometeorology*, 12, 191-223.
- Sadras, V. O., & Monzon, J. P. (2006). Modelled wheat phenology captures rising temperature trends: Shortened time to flowering and maturity in Australia and Argentina. *Field Crops Research*, 99(2-3), 136-146.
- Saiyed, I. M., Bullock, P. R., Sapirstein, H. D., Finlay, G. J., & Jarvis, C. K. (2009). Thermal time models for estimating wheat phenological development and weather-based relationships to wheat quality. *Canadian Journal of Plant Science*, 89(3), 429-439.  
<https://doi.org/Doi 10.4141/Cjps07114>
- Salazar-Gutierrez, M., Johnson, J., Chaves-Cordoba, B., & Hoogenboom, G. (2013). Relationship of base temperature to development of winter wheat. *International Journal of Plant Production*, 7(4), 741-762.
- Sciarresi, C., Patrignani, A., Soltani, A., Sinclair, T., & Lollato, R. P. (2019). Plant traits to increase winter wheat yield in semiarid and subhumid environments. *Agronomy Journal*, 111(4), 1728-1740.
- Silva, P. (2021). *Breeding strategies for improving pest and disease resistance in wheat*.  
<https://krex.k-state.edu/dspace/handle/2097/41279>

- Streck, N. A., Weiss, A., Xue, Q., & Baenziger, P. S. (2003). Improving predictions of developmental stages in winter wheat: a modified Wang and Engel model. *Agricultural and Forest Meteorology*, 115(3-4), 139-150.
- Wang, B., Li Liu, D., Asseng, S., Macadam, I., & Yu, Q. (2015). Impact of climate change on wheat flowering time in eastern Australia. *Agricultural and Forest Meteorology*, 209, 11-21.
- Wang, E. L., & Engel, T. (1998). Simulation of phenological development of wheat crops. *Agricultural Systems*, 58(1), 1-24. <https://doi.org/Doi> 10.1016/S0308-521x(98)00028-6
- Wang, X., Xuan, H., Evers, B., Shrestha, S., Pless, R., & Poland, J. (2019). High-throughput phenotyping with deep learning gives insight into the genetic architecture of flowering time in wheat. *Gigascience*, 8(11). <https://doi.org/ARTN> giz120  
10.1093/gigascience/giz120
- Watson, S. (2002). *Wheat Varieties for Kansas and the Great Plains*. Lone Tree Publ. Co., Topeka, KS.
- WestBred. (2018). WB4269 Hard Red Winter. <https://www.westbred.com/content/dam/westbred/us/en/products/spring-wheat/WB4269.pdf>
- Yan, W., & Wallace, D. H. (1998). Simulation and prediction of plant phenology for five crops based on photoperiod× temperature interaction. *Annals of Botany*, 81(6), 705-716.
- Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). Decimal Code for Growth Stages of Cereals. *Weed Research*, 14(6), 415-421. <https://doi.org/DOI> 10.1111/j.1365-3180.1974.tb01084.x
- Zheng, B., Chenu, K., & Doherty, A. The APSIM-wheat module (7.5 R3008).

# **Chapter 3 - Assessment of Spatial Variability within the KS Wheat Breeding Program**

## **Introduction**

In plant breeding the impacts of genotype by environment interactions and the challenges to quantify these interactions have long been recognized (Allard & Bradshaw, 1964). Both macro and micro environment variations in precipitation, temperature, and soil nutrient availability have shown to impact breeder selections (Reynolds, 2002). Traditionally, breeders mitigate these interactions by evaluating genotype performance across varying environments over multiple years. However, limitations in labor, equipment and seed availability can restrict the number of testing locations that a breeding program can reasonably maintain throughout a growing season. Furthermore, in some instances, the annual weather variations can exceed long-term weather extremes for a location (Atlin et al., 2017). This makes breeding genotypes for both regional and global adaptation challenging. However, global food security depends on breeders ability to account for environmental variations in the presence of climate change (Braun et al., 2010).

Environmental impacts on wheat yield are well documented and mostly modulated by heat and drought stress (Lollato et al., 2017). These impacts have been demonstrated to affect with-in season and across year production of global crops (Morgounov et al., 2018). In addition to impacting grain yield, spatial variability, particularly with field available nitrogen, can also impact grain quality (Stafford, 1999). Micronutrient spatial variability and field topography can also cause quality variability in micronutrient, protein, and oil concentrations of key staple grain crops (Kravchenko & Bullock, 2002; Lin et al., 2009). Similarly, wheat diseases can cause genotype by environment interactions both on plant health and mycotoxin concentrations (Miedaner et al., 2001). Genetic resistance effectiveness to diseases can also be impacted by

environmental parameters such as low or high temperatures (Rodriguez-Algaba et al., 2020).

Accounting for these spatial effects on yield, quality and disease response are critical for breeders as they develop new genetics

Like weather fluctuations, variations in soil physical and chemical properties can greatly affect genotype performance both at a micro and macro levels (Blum, 2018). Soil compaction can significantly impact genotype performance and can particularly be a problem in small fields where high foot traffic and small plot research equipment is used. In combination with weather variation, soil compaction can have variable impact on wheat root development and overall grain yield (Correa et al., 2019; Liu et al., 2021). In the US Central Great Plains, acid soils and the consequent increased aluminum levels, have significant implications for wheat breeding and production (Lollato et al., 2019). This has driven the development of acid-tolerant varieties that are have the ability to produce higher yields in these environments (Kariuki, 2007; Munaro et al., 2020). In a similar pairing, alkaline and sodic soils have also shown an evolutionary connection in plant breeding (Bui, 2013). Soil salinity can have impacts on crop production as it can have inverse effects on soil available water. Within Kansas, salinity issues are typically isolated within fields (Lamond & Whitney, 1992). However, saline inclusion in fields have significant implications on breeding of cereal crops (Isla et al., 2003). Having the ability to quantify these environmental parameters and test germplasm for tolerance in extreme conditions could lead to the release of improved varieties that can thrive in conditions that are traditionally associated with low production

Breeding programs commonly use a variety of experimental designs and statistical analyses to account for spatial variability, as improved experimental designs in combination with proper spatial analysis can impact selection accuracies (Qiao et al., 2000). In early generation

studies where field space and seed availability are limited, single rep experiments are usually needed. In these instances, breeders either use augmented or partial replicated experimental designs (Cullis et al., 2006). These experiments allow for spatial adjustments without genotype replication by adjusting to checks placed throughout the experiment. These partial and un-replicated trials can increase genetic gain by increasing genetic variance (Smith et al., 2006). Despite the efficiencies of these trials, their lack replication limits the ability to delineate the genetic variance and error, although computational solutions have been proposed (You et al., 2016). Typically in later generation yield testing, breeding programs utilize replicated trials such as randomized complete block and alpha-lattice designs (Kumar et al., 2020). In addition to experimental design, there are multiple spatial statistical methods that can also account for variability within a trial; for instance, using mixed effect methods for genotype by environment analysis can quantify variance across environments (Elias et al., 2016).

In-season remote sensing sensor measurements can account for yield spatial variability by analyzing plant growth throughout the season (Zarco-Tejada et al., 2005). Similarly, this in-season collection method can geostatistically characterize grain quality variability based on spatially delineated zones (Diacono et al., 2012). Precision agriculture technologies can improve nitrogen management through delimitating management zones in crop production situations (Khosla et al., 2002; Schwalbert et al., 2019). However, using in-season sensor measurements to quantify spatial variability within breeding programs poses a unique challenge, compared to production or management trials, with an added genetic variation component. In-season vegetation indices can quantify this genetic variation (Babar et al., 2006) and although this genetic variation can impact spatial estimates, in-season vegetative indices using UAV's can be used to estimate spatial effects within large breeding programs (Haghighattalab et al., 2017).

A twofold limitation of cost and time typically restricts the application of sampling for spatial variability within a breeding program. Although the total area of a test location for breeding programs is relatively small compared to production agriculture, these small plots require added spatial resolution to effectively quantify variability. Additionally, statistical spatial analysis requires robust and large sample sets. The novelty of collecting sensor-based soil parameters is that the sample size can be significantly increased without added time in the field or the added cost and time of laboratory soil analysis. Therefore, many on-the-go sensor platforms have been commercially developed for precision spatial quantification (Lund et al., 1999; Sudduth et al., 2001). Primarily early generation on-the-go platforms were equipped with electromagnetic induction sensors to measure the apparent soil electrical conductivity ( $EC_a$ ) to quantify spatial changes (Brevik et al., 2006). Although spectrometers in soil lab analysis have widely been used to relate to soil properties (Ben-Dor et al., 2009; Nocita et al., 2015) the technology has not always been mobile or cost effective for field applications. Recently, this technology has lowered in cost and has become portable, allowing for quick and cost-effectively soil property analysis and fertilizer recommendations (Ng et al., 2020). Additionally, this optical sensor technology has recently been integrated into mobile sensors allowing for real time assessment of soil organic matter (Kweon et al., 2013). Modern soil sensor platforms have also implemented methods to quantify soil compaction through load cells and sensing cones (Cho & Sudduth, 2015; Zeng et al., 2008) The continued development of these sensors makes the possibility of application within a breeding program feasible.

## **Materials and Methods**

### **Experiment locations**

Soil samples were collected at two locations, Reno (RN) and Thomas (TH) Counties (Table 3.1). The TH site was sampled prior to planting for two growing seasons (2020 and 2021), whereas the RN site was only sampled preplant for the 2021 growing season. Due to crop rotation unique fields were sampled at each location between years. In addition to unique fields between years, there were two distinctly managed fields, one irrigated (THI) and one dryland (THD), sampled at the TH location both years. Likewise at Reno County, samples were collected at distinct north (RNN) and south (RNS) fields. Furthermore, although the RNN field was sampled as a continuous field there are four distinctly managed blocks within the whole field. Two of the blocks were planted in soybeans during the summer of 2020 and were followed by winter wheat plots. Whereas the other two blocks were summer fallow and planted to winter canola in the fall of 2020. In total seven site years were sampled for this experiment. The locations and fields were selected on their relative importance within the breeding program and their diverse geographic location which separated them in distinct wheat growing regions (Roozeboom et al., 2008).

Soil types varied by location and year. The principal soil type for the 20THI, 20THD, and 21THI fields was a Keith silt loam (Fine-silty, mixed, superactive, mesic Aridic Argiustolls). However, the 20THD field had a three-acre inclusion of Richfield silty clay loam (Fine, smectitic, mesic Aridic Argiustolls) which accounted for approximately 25% of the total sampling area. The 21THD field was comprised of three soil types, with approximately 60% of the field was Richfield silty clay loam, though there was a two-hectare inclusion of Keith silt loam and a one eighth-hectare inclusion of Pleasant silty clay loam (Fine, smectitic, mesic Aquic

Argiustolls). The primary soil type at both RNN and RNS was an Ost loam (Fine-loamy, mixed, superactive, mesic Udic Argiustolls) (Soil Survey Staff, 2021).

## **Plant Material**

To evaluate spatial variability within the Kansas State Wheat Breeding program only populations grown in a modified augmented type 2 (MAD-2) were selected. These un-replicated, rectangular plots use check varieties to allow for row column corrections. For these experiments the primary check was a, three-way blend of commercially adapted lines for the region. A Western Blend was used for all fields at the TH location and the Central Blend was used as the primary check at the RN location. In addition to the primary check each experiment had two unique secondary checks that were selected based on agronomic potential and producer adoption within the testing region.

The population at the 20THD location consisted of 351 wheat *Ae. tauschii* introgressions (AeT<sub>a</sub>). The lines were developed by crossing six hexaploidy wheat lines with 21 diploid AeT<sub>a</sub> lines to capture the genetic diversity of the D-genome. Full details of the population establishment is detailed in Nyine et al. (2020). The 20THI population consisted of elite winter durum lines (*Triticum turgidum* subsp. *durum*) being tested in the Durum Preliminary Yield Nursery (DPYN). The DPYN lines are fix lines that are being evaluated in a yield nursery for the first time after limited selection intensity. Likewise, the DPYN population was evaluated at the 21THI location. However, due to evolving nature of breeding population, the entries at the 20THI-DPYN and 21THI-DPYN were not equivalent. The population evaluated at the 21THD location was the winter wheat PYN1A experiment. At RNN a population of F<sub>4</sub> segregating winter wheat populations planted in 3 row plots were evaluated. Two separate experiments were evaluated on the RNS location; the winter wheat PYN experiment which was

comprised into 4 distinct PYN populations identified as PYN 4-8 and a yield trial (YT) experiment which contained PYN populations 1-3 and a population of F<sub>5</sub> segregating populations.

Due to seed availability and evaluation stage, plot dimensions varied by location and experiment. The plot dimensions for the 20THD-Aeta, 20RN-PYN and 20THD-PYN1A were 2.4 m x 1.4 m (3.36 m<sup>2</sup>). Whereas, the 20THI-DPYNA, 21THI-DPYNA and 21RNS-YT experiments were 4.5 m x 1.4 m (6.75 m<sup>2</sup>). Each plot consisted of 6 rows spaced at 20cm with a seeding rate of 15g m<sup>-2</sup>. The 20RNN-F4 plots were planted as 3 row plots and were 2.3 m x 0.6 m. The plots were planted in September and October of each year and harvested in late June or early July.

### **Soil Sensors**

Two soil field sensor platforms were used to obtain indirect measurements of soil properties. The Veris MSP-3 mobile sensor cart platform (Veris Technologies, USA) was used to collect apparent electric conductivity (EC<sub>a</sub>) and spectral reflectance. This sensor has two pairs of coulter disks equipped with rolling electrodes. While one disk electrode emits a small electrical current, its paired disk receives the current after passing through the soil medium. The recorded received electric pulse is an indirect measurement of soil resistivity. These soil resistivity measurements correlate with changes in soil texture, soil moisture, cation exchange capacity and salinity (Kweon et al., 2013; Naderi-Boldaji et al., 2014). The two pairs of disks allow for resistivity measurements at the 0-30 and 0-90 cm depths. Additionally, the cart contains a LED optical sensor mounted within a planter row that captures soil reflectance in the red (660nm) and near-infrared (940nm) spectral regions at the 5 cm depth. The reflectance values of the optic sensor are related to soil organic matter (Lund & Maxton, 2011). The cart was pulled at a speed

of  $1.5 \text{ m s}^{-1}$  on a 10-meter grid, soil readings were obtained every 1 to 2 seconds along the transect. However, in 2020 only 10 meter transects were obtained for the whole field and only portions of the field were cross hatched to obtain the 10-meter true grid pattern (Table 3.1). The cart was also equipped with a GPS unit that allowed for all samples to be geo-referenced. However, to improve transect accuracy the cart was pulled with a John Deere 5055E tractor equipped with a Trimble RTK autosteer navigation system.

The second sensor used was the Veris P4000 DW-EC-Force Probe. This platform was used on all locations at all years. Like the MSP-3 the P4000 probe is equipped with an optical and electromagnetic sensor. Additionally, the P4000 is equipped with load cell force penetrometer that allows for an indirect measurement of soil compaction (Cho & Sudduth, 2015). The main advantage of the P4000 is that it allows for point sampling at controlled depths. The probe is pushed through the soil medium using through the tractor hydraulic system at a rate not to exceed  $30 \text{ mm s}^{-1}$ . The probe also has a GPS system for geo-reference. However, this system did not have the accuracy needed for sampling within small plots (Fig. 3.1) Therefore, all geo-referenced points were collected with the RTK Trimble system detailed above. Probe readings were collected at two different intervals. On the whole field probe readings were taken on 30m grids, whereas smaller areas within the large field had a higher density of samples and probe measurements were taken on 5m grids.

### **Soil Analysis**

In addition to soil sensor measurement, physical soil cores were also obtained with the Veris P4000 using the core attachment. These cores were 5.1cm in diameter and taken to the one-meter depth. Within the whole field, cores were collected on 60m grids. All soil cores were paired with P4000 probe locations. Samples were preserved in plastic sleeves and immediately

frozen for future analysis. The cores were split into three depths 0-20, 20-60 and 60-100 cm. A sub sample of the soil was oven dried to obtain volumetric water content (VWC) and bulk density (BD) (Blake & Hartge, 1986). The remaining samples were air dried and sieved to pass through a 0.25mm sieve. The ground samples were then used to determine other soil properties.

All sample depths were analyzed for electrical conductivity using a 1:5 soil to water (He et al., 2012). All other soil parameters were only measured for the 0-20 and 20-60 cm depths. Soil pH was measured in a 1:5 soil to water solution (Brown, 1998) and phosphorus (P) concentrations were obtained through a Mehlich-3 extract (Mehlich, 1984). Cation exchange capacity (CEC) was measured using the displacement method (Chapman, 1965). Soil potassium (K), calcium (Ca), magnesium (Mg), and sodium (Na) were obtained with ammonium acetate extraction (Warncke & Brown, 1998). Total nitrogen and organic carbon were taken using the dry combustion method (Nelson & Sommers, 1996), and soil texture was recorded through the hydrometer method (Bouyoucos, 1962).

## **Spatial Analysis**

To assess spatial patterns ordinary kriging (OK) was applied to the observed soil properties from all platforms. The first step in this process was to assess the spatial covariance structure of the data by constructing a variogram. A variogram model was then fitted using the variogram function in the ‘gstat’ package (Pebesma & Graeler, 2015). This function allows for variogram model selection that minimizes the root mean squared error (RMSE). The variogram assess the degree of spatial autocorrelation of points exhibit and allows for optimal performance of the kriging models. Five model structures were evaluated for each soil parameter variogram, including exponential, linear, Gaussian, and spherical distribution. In this study all variograms

were fit with the assumption of isotropic distribution. The theoretical variogram model is shown by equation 1.

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \quad [\text{Eq. 1}]$$

Where;

$\gamma(h)$  is the variogram value at a distance of  $h$

$Z(x_i)$  is the observed value of a given parameter at location  $x_i$

$Z(x_i + h)$  is the observed value location with distance  $h$

$N(h)$  are the pairs at distance  $h$

After a variogram model was selected for each combination of parameters and locations, a spatial grid point was created for spatial predictions. The OK method was selected to make spatial predictions and was executed using the ‘gstat’ package in R (Pebesma & Graeler, 2015). The OK algorithm estimates a property  $Z$  at any point in the spatial grid by weighting the average of observed points near the estimated grid point.

$$Z(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad [\text{Eq. 2}]$$

Where,  $Z(x_0)$  is the predicted value at each location and  $\lambda_i$  is the weight assigned to each neighbor.

In addition to individually kriged grid points the interpolated kriged data was partitioned into zones based on the K-means algorithm (Eq. 3). This hard clustering method minimizes the sum of squares within clusters where  $\mu^{(i)}$  is centroid of the  $i^{th}$  cluster and is calculated by the mean of points in matrix  $X_i$

$$\text{agr min } \sum_{i=1}^k \sum_{x \in X_i} \|x - \mu^{(i)}\|^2 \quad [\text{Eq. 3}]$$

The silhouette method was used to quantitatively determine the optimum number of clusters for each parameter (Rousseeuw, 1987). This method compares the fit of each point to its own cluster compared to other clusters. The maximum obtainable silhouette value is 1 and indicates the point fully belongs to the cluster whereas lower values indicate a poor fit with a value of zero meaning the point does not fit to one particular cluster. The `fviz_nbclust` function within the ‘factoextra’ package in R was used to evaluate 2 to 10 clusters for each field and soil parameter where the cluster with the highest silhouette value was determined to be the optimum number of clusters.

The produced spatial grid and delineated zones from the kriging and k-means process was uploaded to Quantum GIS (*QGIS: A Free and Open Source Geographic Information System.* ), along with a point file obtained from the GPS planting file that represented the middle of the plot. Each plot was then assigned a kriged value for each soil parameter along with a delineated zone through the join attributes by nearest distance function in QGIS.

### **Statistical Analysis**

Yield adjustments were executed by incorporating row and column adjustments using method 3 consistent with the MAD-2 design with Agrobases Generation II software (Agrobases Generation II 2014, Agronomix, Winnipeg, MB, Canada; <https://www.agronomix.com/> ). Method 3 adjusts parameters based on both the secondary check and the primary check subplot within each whole plot represented in equation 4. Where  $b$  is the regression coefficient of the mean control subplots.

$$Y_{ij(k)} = Y_{ij(k)} - b(X_{ij(A)} - \bar{X}_A) \quad [\text{Eq. 4}]$$

To test spatial zone effects on yield, a Tukey test was performed on each individual parameter zone at the 0.95 significance level. In addition to evaluating zonal effects on yield the

kriged value of each soil parameter was evaluated in a linear model to assess each trait's correlation to yield.

In order to make to make spatial corrections with plot level kriged values, a mixed multivariate model [Eq. 5] was used within the 'ASReml-R' package (Butler et al., 2009)

$$y = X\beta + Z\tau + \varepsilon \quad [\text{Eq. 5}]$$

where  $y$  is a vector of observed yields,  $\beta$  is a vector of known fixed effects representing the measured soil traits and  $\tau$  is the genotype random effects vector.  $X$  and  $Z$  are the incidence matrices for  $\beta$  and  $\tau$  and  $\varepsilon$  is a vector of random effects residual effects, normally distributed with zero mean. For each model the fixed effect covariates were the soil properties of interest and the random effects were genotype, range and row.

The experimental design and spatial correction models for yield were compared by calculating the coefficient of variation (CV)

$$CV = \frac{\sigma}{\mu} \quad [\text{Eq. 6}]$$

Where  $\sigma$  is the population standard deviation and  $\mu$  is the population mean.

## Results and Discussions

The soil parameter distribution varied by location and year. While most properties had normal distribution, there were several locations that exhibited bi-modal distribution (Fig. 3.2, 3.3 and 3.4). This variability in distribution may be expected due to the heterogeneous nature of soil (Adamchuk et al., 2010), particularly in fields with NRCS-mapped soil type inclusions or fields where multiple management experiments have occurred over time. This variability was also commonly observed by depth, particularly with the P4000. While these depth variations would be expected due to soil profile horizon variations, the location and measured soil property affected the magnitude of the shift and distribution. The 21RNN and 21RNS fields had lower

EC<sub>a</sub> values than all THI and THD fields for both years (Figure 3.2). This likely was due to a combination of low pre-sampling precipitation and a summer soybean crop that was harvested just prior to the sampling period. In general, the Red and IR reflectance trends by depth paralleled each other. This would be expected since both the Red and IR bands correlate to organic matter content (Kweon et al., 2013) and large scale organic matter variations would not be expected for the soil types tested.

Comparable to the P4000 data, trends by location and by depth were observed in the MSP3 data (Fig. 3.3). This observation was particularly evident for EC<sub>a</sub> at 21RNN field (Fig. 3.3C) where there was a strong bi-modal trend. As discussed in the field description, although this field was sampled as one field, it was managed as four blocks. The two summer fallow, fall sown canola fields likely had a higher VWC resulting in a higher EC reading as the Veris EC measurement is directly correlated to VWC (Hezarjaribi & Sourell, 2007). For this reason, the 21RNN spatial extrapolations were conducted as four sub-fields within the 21RNN whole field. In comparison to the P4000, the MSP3 red and IR measurements did not show as much variation or range at all locations. This is likely due to the 5 cm measurement depth of the MSP3 in comparison to the 20cm sections captured by the P4000.

Lab analysis of the physical soil cores revealed location and depth patterns of multiple soils parameters. At all locations the VWC for the 0-20cm depth were lower than the values at the 20-60cm depth (Fig 3.4). This would be expected based on precipitation events and previous crop management. Even though the irrigated site at TH receives a pre-plant application to aid in germination and stand establishment, the application timing in respects to soil sample timing would favor higher VWC values at the 20-60cm depth. It is also notable that the VWC values at the RN location were lower than TH values for both years. The combination of limited summer

precipitation and the summer soybean crop are likely the cause of this discrepancy. In general, organic matter content was higher at the surface than the subsurface. This trend is also consistent with relative measurements obtained and previously discussed for both soil sensor platforms. Bulk density samples varied by depth and location both on magnitude and distribution. There were instances where the bulk density exceeded normal expected limits which was likely caused by compression within the sample tube from the hydraulic probe (Rogers & Carter, 1987). However, the relative values were still used to account for compaction variability within the field. Soil macronutrients were also within expected ranges. However, it is worth noting that phosphorus content was both higher and had a greater distribution range at the 20-60cm depth. While this is not abnormal, it is notable that the distribution would affect plant growth and could lead to spatial zones with higher phosphorous content.

Like the soil physical properties there were also trends in the soil chemical properties and micronutrient distribution. Primarily, there was an observed location difference between TH and RN in soil pH, particularly at the surface (Fig 3.5). The TH surface pH values tended to be more neutral and ranged from 7.02 to 8.76. Whereas the RN pH surface values were more acidic and ranged from 4.8 to 8.33. However, the 20-60cm pH values were similar at both locations with TH samples and ranging from 7.86 to 8.9 and RN samples ranging from 7.64 to 8.46. This wider range of pH values at RN surface could significantly impact plant performance as it has been shown that acidic soils can impact genotypic performance of wheat yield (Lollato et al., 2019). The increased pH for the second depth at RN could be associated with the increased concentration of Ca also observed at the 20-60cm depth (Fig 3.5C and D). While processing the RN cores inclusions of calcium carbonate were visible at the 40cm depth (Fig. 3.6). The presence of these extra cations also caused the CEC to increase at the 20-60cm depth although the CEC at

both RN and TH were comparable across all years. However, the distribution of the Ca inclusion could affect soil property distribution and genotype performance at the location.

Soil textural classification differences occurred at all locations for all years (Fig. 3.7). The textural changes were expected at TH as the USDA NRCS soil survey included soil series with both silt loam and silty clay loam textures. However, the soil survey at RN only included one soil texture while multiple textures were observed. The discrepancy in observed and mapped soil textures was expected as the scale of soil survey maps are inadequate for precision agriculture applications as the accuracies and scale are often inconsistent (Brevik et al., 2003). Overall, all locations and years had at least two textures present, with the 20THD, 21THD and 21THI having four textural classes represented. The spatial distribution of textures can affect genetic performance as texture has been associated with water holding capacity and wheat yield (Wu et al., 2011). Since soil texture is a soil property that does not rapidly change having the ability to construct an accurate soil texture map may be helpful to breeders particularly for experimental design avoidance where experiments can be placed in areas of uniform texture.

The fitted variogram model varied for each soil parameter and sensor platform. The exponential model resulted in the best fit a majority of the time for the MSP3 data (Table 3.2) as it accounted for 19 of the 40 (47.5%) soil parameter depth combinations. Additionally, the spherical model accounted for 27.5% (11 of 40) parameter-depth combinations while the linear model accounted for 20% (8 of 40) and the gaussian model accounted for 3.1% (2 of 64). The P4000 distribution of model fits was not as distinct (Table 3.3) with the spherical (39.1% 25 of 64), linear (31.3% 20 of 64) and exponential (26.6% 17 of 64) models all being significantly represented. However, like the MSP3, the gaussian model was selected the least, 3.1% (2 of 64). The variogram range for both the MSP3 (Table 3.2) and the P4000 (Table 3.3) sensors have

similar distributions. However, the average range for the P4000 (32.75m) was higher than the MSP3 (25.62m) indicating more autocorrelation with the MSP3 platform. However, this maybe expected with the increased number of observations provided by the MSP3 (Griffith, 1987).

The k-means cluster groupings for the MSP3 platform ranged from two to ten clusters with two clusters being selected at the highest frequency (70% 28 of 40). This trend was also observed for P4000, however at a lower magnitude, where two clusters were selected 59.4% (38 of 64) of the time. The wider distribution of clusters with the P4000 could be attributed to sample size. Since k-means is a hard clustering technique, when datasets have lower number of samples, each point has a lower chance to distinctly belong to a group (Hot & Popović-Bugarin, 2015). In this study it is important to recognize that the increased sample density advantage the MSP3 has over the P4000 due to time and labor requirements. This sample density can lead to better geospatial corrections by limiting the effect that one sample point can have on the geo-spatial corrections. Although, it has also been documented that geostatistical analysis can be performed with a limited number of observations over large sample areas (Schloeder et al., 2001). Several studies show that the optimum number of samples for OK is 100-150 (Voltz & Webster, 1990), which limits the implementation of the P4000 within a breeding program.

In most cases, field experiments covered multiple k-means cluster groups as shown in Figure 3.7. However, in some cases not all cluster groups were represented (Fig. 3.8 and 3.9) which would be expected as the whole field was tested but the individual experiments only spanned a small area of the whole field. Cluster avoidance could be one strategy breeders can utilize if spatial samples are collected before experiments are planned and mapped (Heil & Schmidhalter, 2017). However, this would require samples be taken prior to plot design or the establishment of long-term soil trends at a sight. At most locations the MSP3 soil parameters

were associated with yield effects, however, the cluster trends were inconsistent depending on location and germplasm being tested. For example, the lower valued cluster for ECSH exhibited a yield increase at 20THD, 21THD and 21RNN. However, the low ECSH cluster showed a significant yield decrease at 20THI, 21RNN and 21RNS and no significant effect at 21THD. A similar trend was observed for the MSP3, ECDP measurement and all EC measurement depths with the P4000. Although  $EC_a$  is associated with texture and moisture (Kalopesas et al., 2015) the actual groupings might not quantify the exact differences of soil property rather than the relative effect of that specific location to population being tested. Like the  $EC_a$  values, P4000 force data also correlated with yield, however, the trends were again inconsistent. In general, lower force, indicating less compaction, had a positive effect on yield. However, in some instances (21RNN- Fig 3.7) a higher force at the surface layer had a yield benefit. While this may not be expected other factors such as soil moisture and soil texture could have an impact on soil compaction and yield potential through stand establishment (Tesar & Jackobs, 1972). This example shows that one soil parameter alone does not always indicate increased yields, particularly in a stressed environment like 21RNN where the increased force could have been caused by soil textural changes that were conducive to higher yield potential.

Similar to the k-means groups, the individual kriged values also showed trends to grain yield. However, the correlations were not strong as the  $R^2$  values ranges from <0.01 to 0.15 for the MSP3, from <0.01 to 0.23 for the P4000 and <0.01 to 0.28 for the lab analyzed soil cores. Of the most significant MSP3 trends was ESCH where a general weak negative correlation to yield was observed. Inversely, the P4000 EC depth 1 had a general positive correlation to yield. The difference in effective depth and sensor specific sphere of influence all could affect the correlation of the values. It is also worth noting that since these are kriged extrapolated

datapoints that the spatial corrections and sample density could also have a significant impact on the correlation to grain yield.

For all experimental years and locations, the spatial and statistical corrections improved the CV over the raw data (Fig 3.16, Table 3.5). It is notable that in general the raw CV for the 2020 fields was higher than the 2021 fields except for the 21THI-DPYNA. This likely could be explained by a late spring freeze in 2020 that caused significant damage to the plants. Both a spatial and genetic effect of freeze damage potentially could account for variation in performance. However, the disparity in effectiveness may also be related to overall number of plots and area of each experiment. For example, the 21THI-DPYNA had the lowest number of samples and the covered the smallest area (Table 3.3) which makes it susceptible to higher CV's if the genetic or spatial variation is high. This is also supported by literature as larger experiments can be impacted by spatial corrections more than smaller experiments (Hoeffler et al., 2020). However, this experiment shows that spatial corrections alone can be as effective or more effective than experimental design corrections.

The soil platform geo-spatial corrections improved CV in all instances except for the AeTa experiment at THD in 2020. However, as previously mentioned, this experiment had potential freeze damage and the blocking effect of the MAD-2 may have quantified that event better than the soil corrections. In all other experiments the soil geo-spatial corrections were at least 2% better than the experimental design corrections. Additionally, the MSP3 sensor had the best CV in 4 of the 7 experiments (57.1%), while the soil cores were best in 2 of the 7 experiments (28.6%) and the experimental design was best once. However, the differences between the soil platforms were never more than 1.71% meaning the soil platform corrections were similar which was expected as they quantify similar soil properties.

Although experimental CV values were similar for spatial and statistical methods there was disparity in correction magnitude and its effect on rank order. For example, several spatial adjustments lowered the raw yield whereas the statistical adjusted yield increased the raw yield for the same genotype (Table B.1). In extreme cases this change was as much as 0.67 tons ha<sup>-1</sup> which changed the rank order by as many as 259 positions in an experiment of 485 entries. Although, in general the top tiers were mostly unchanged, these large changes in rank order can have significant impacts on breeder selections and warrants the further investigation on yield correction impacts on breeder selections.

Overall, the sensor platform did not significantly affect the CV correction or rank order as much as the spatial correction vs the statistical correction did (Table 3.5). In most experimental locations and years, the MSP3 had the lowest CV of the single correction methods, however, improved performance was not significant. This likely was due to the increased sampling density that limited the extrapolation needs during OK. Even though there was rank order change between platforms the degree of change was not as significant as observed between the statistical and spatial corrections. Since all platforms were quantifying the variability of the soil, it would be expected that the differences were not as great. However, each platform used different techniques, sample depths and quantity of observations, so some variation would be expected.

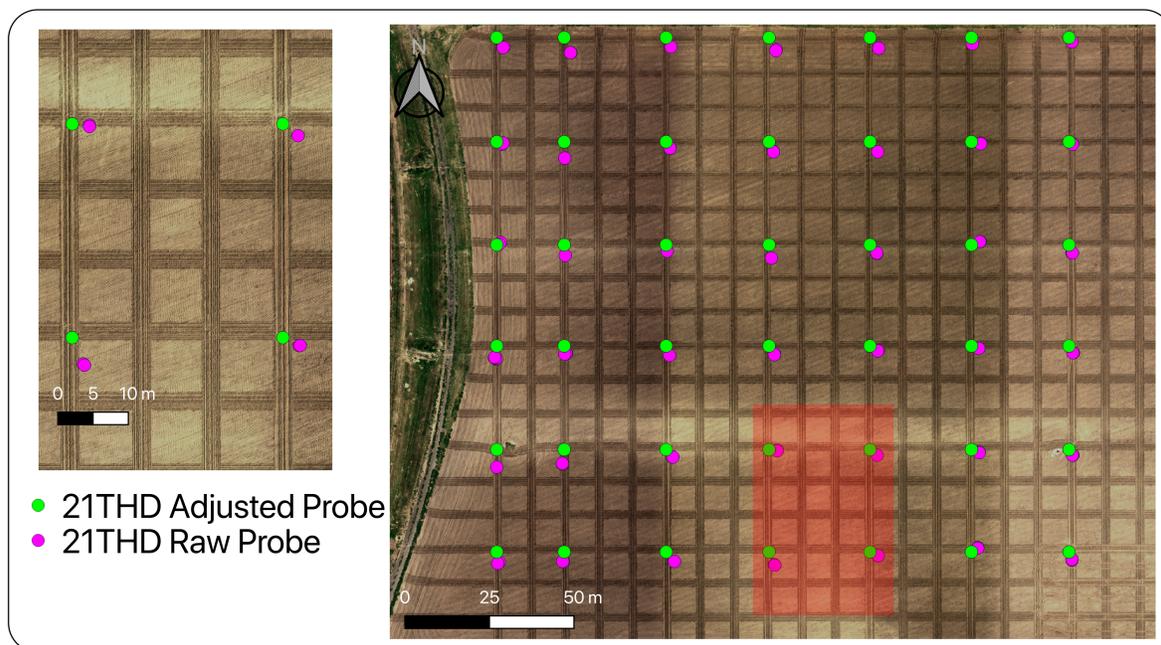
The advantage that soil property corrections have over the experimental statistical methods is that blocking strategies may not always capture the variation. As shown in both Fig. 3.7 and B.3, the soil clusters and geo-spatial corrections do not occur in discrete squares or rectangles and span across several blocks, which can negate the value of blocking. Additionally, the quality and placement of checks in a MAD-2 are susceptible to spatial and management

variability (Lin & Poushinsky, 1983). For example, poor stand establishment, or isolated stressor events can diminish the effectiveness of the check corrections. These factors support the implementation of geo-spatial corrections from soil parameters over experimental design.

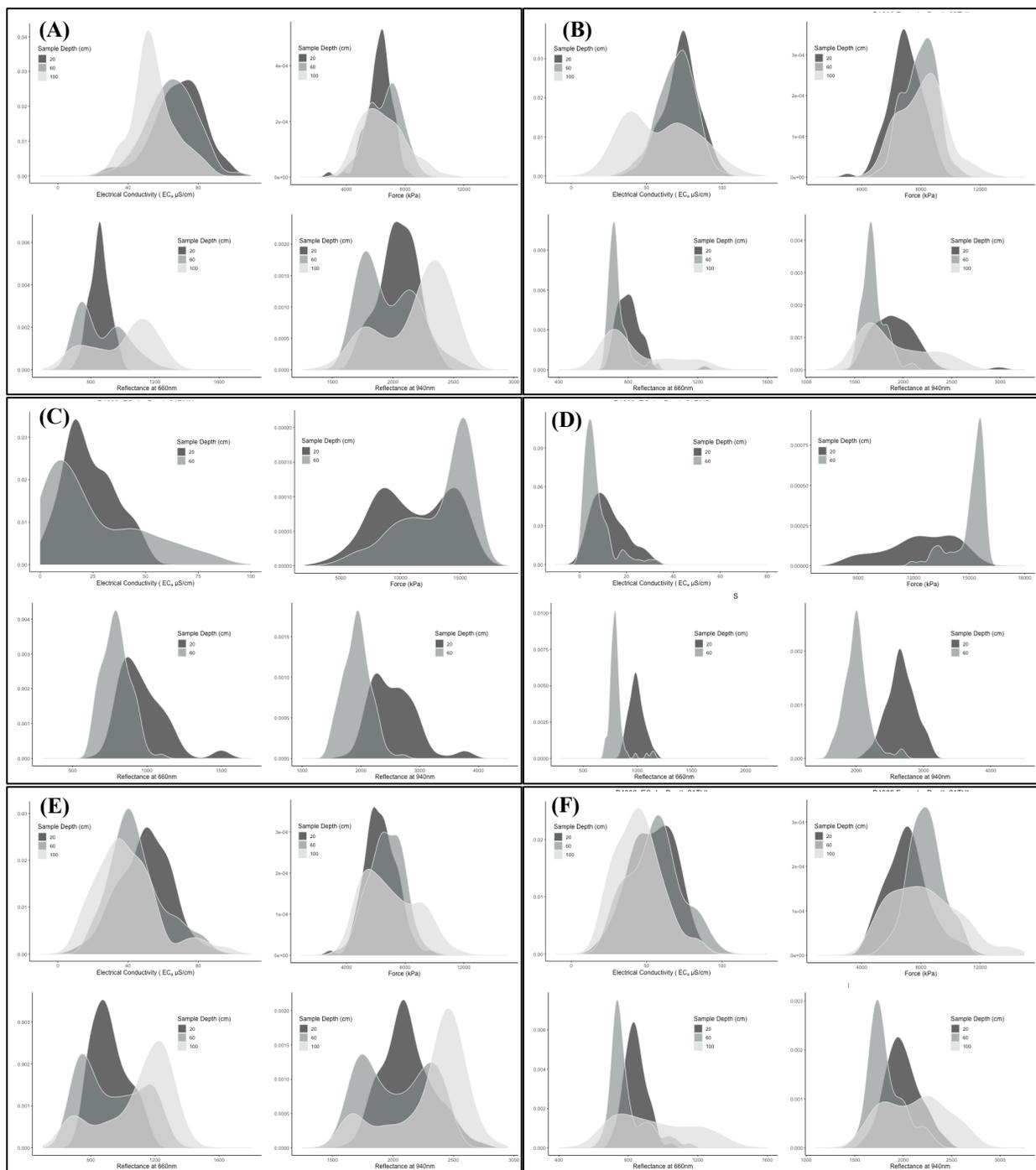
## **Conclusions**

This experiment confirms that soil spatial variation within breeding experiments exist and that it has significant impacts on genotype performance and ranking estimation. Additionally, there is support that on-the-go precision soil sensors have the capability to capture this variation. However, it is unknown if these platforms would need to be used on every field for every growing season, or if more stable soil properties such as soil texture can be quantified once and used multiple years. Furthermore, the cost of equipment and operation need to be explored prior to implementation of this technology into a breeding program. From this study a high-density sampling system such as the MSP3 was favored as it both improved the CV and was easier to operate, making it more likely to be implemented in a program.

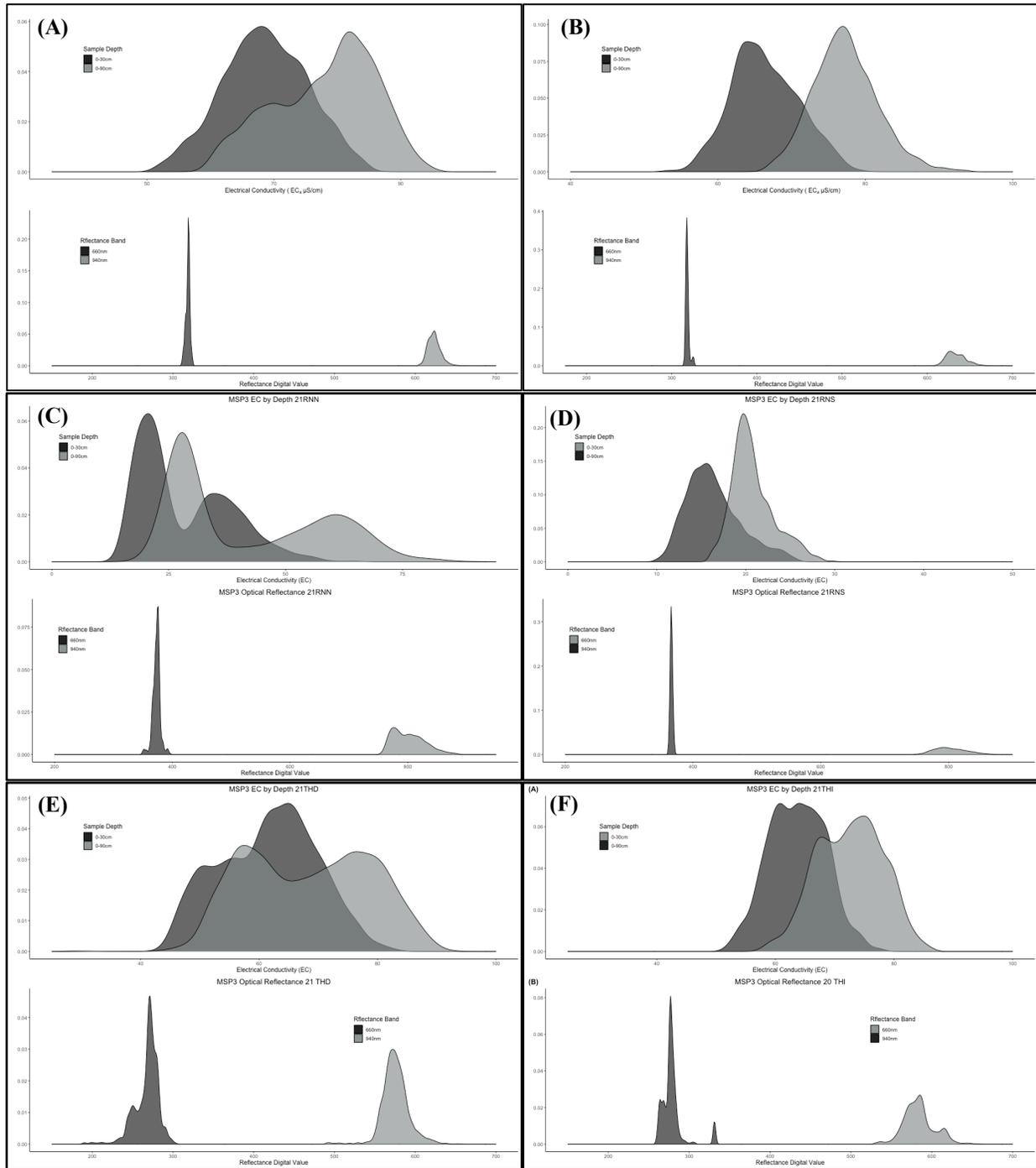
Additionally, underlying questions of soil parameter effects on the germplasm being tested still exist. Continued research, to quantify blocking affects these spatial corrections is needed. These techniques should be tested on germplasm that has been through several rounds of selection and are now in replicated field trials. At this stage, this technology could not only improve breeder selections but could also push towards agronomic recommendations upon selection for release. Furthermore, these soil sensors should be compared to remote sensing techniques where both pre-season and in-season monitoring with UAV's may have the potential to capture soil surface variability with bare ground sensing and full profile variability with genotype reflectance data.



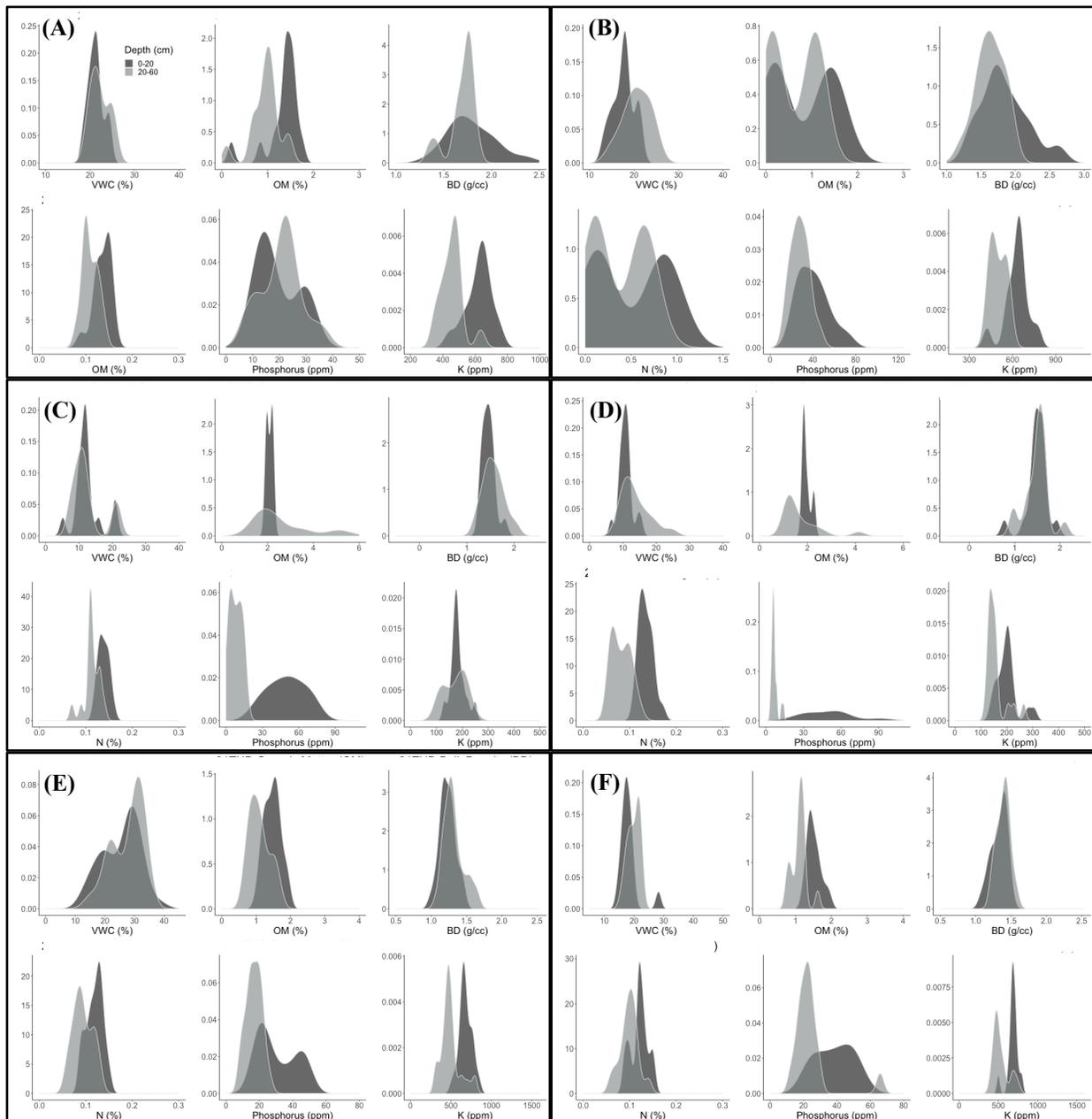
**Figure 3.1.** 21THD Veris P4000 sample points. Points in pink represent the raw GPS positions obtained from the P4000 and the green points represent RTK corrected points. Ariel image was obtained after sampling and grid marks are sampling paths created by the MSP3 sensor.



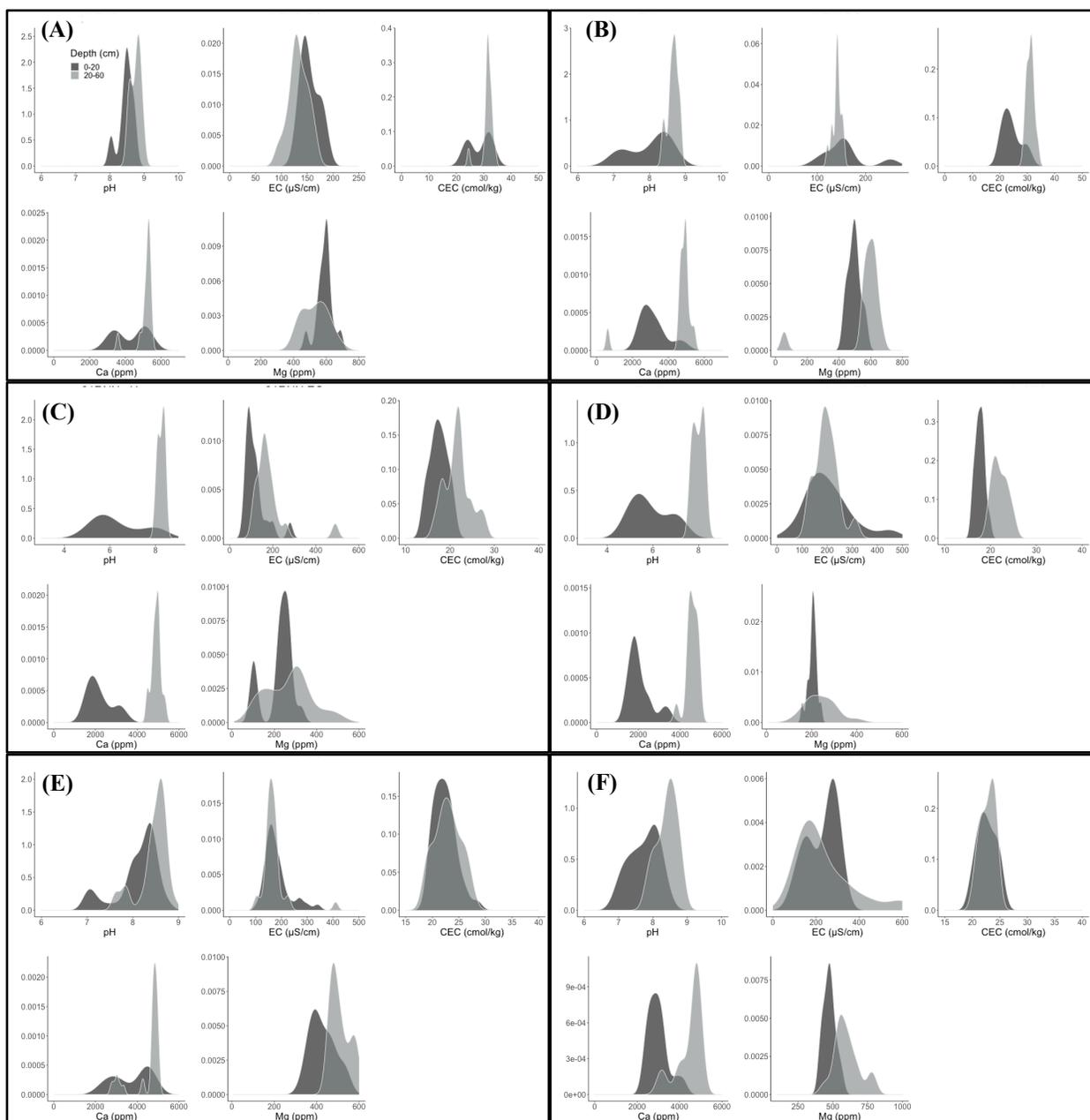
**Figure 3.2.** Distribution of soil values obtained by the Veris P4000 by depth for EC, Force, Red and IR parameters for 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F). The x-axis is the measured value of the parameter, force is in KPa, apparent electrical conductivity ( $EC_a$ ) is in  $\mu\text{S cm}^{-1}$ , and reflectance are 16-bit digital values. Each sensor sample depth is represented by color.



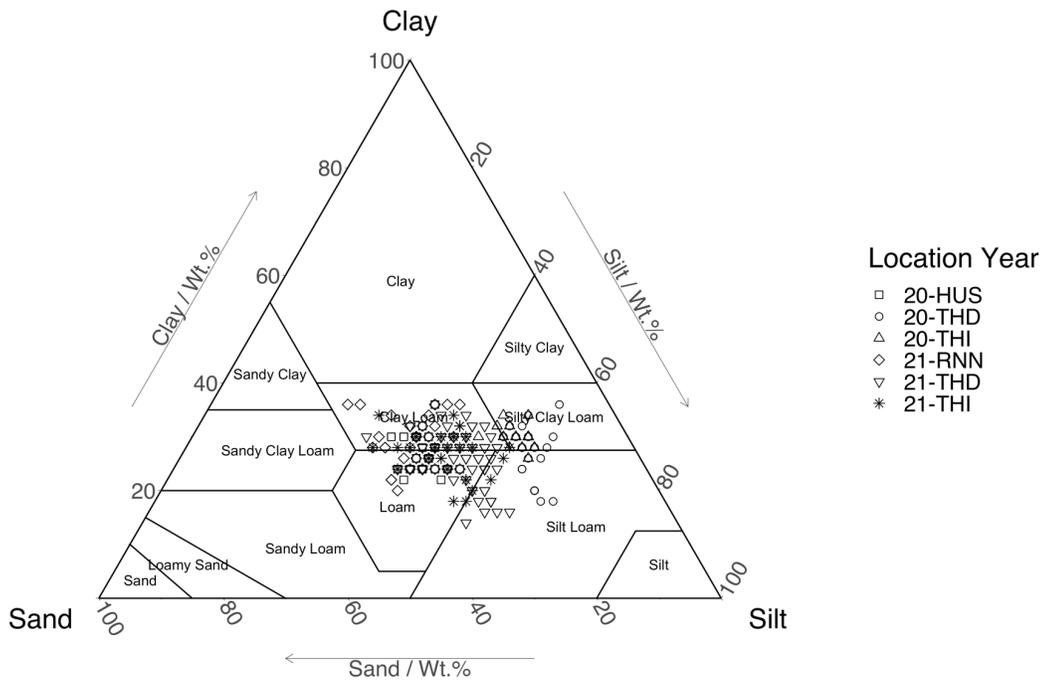
**Figure 3.3.** Distribution of soil values obtained by the Veris MSP3 for shallow EC (30cm), deep EC (90cm), Red and IR parameters at 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F). The x-axis is the measured value of the parameter, force is in KPa, apparent electrical conductivity ( $EC_a$ ) is in  $\mu\text{S cm}^{-1}$ , and reflectance are 16-bit digital values. The colors in the  $EC_a$  graphs represent sample depth and represent the spectral band for the reflectance graphs.



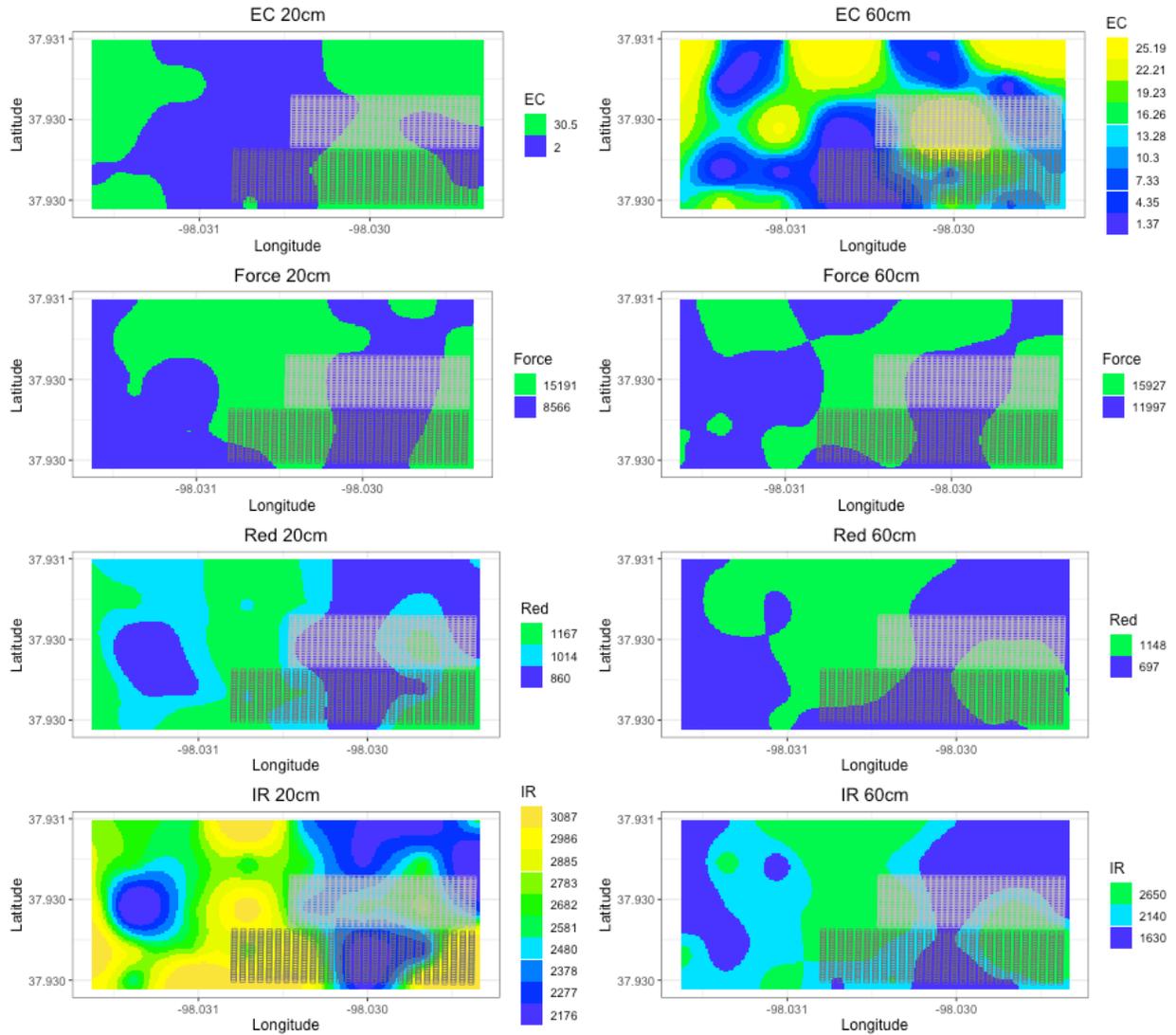
**Figure 3.4.** Distribution of soil values obtained from lab analysis of physical soil cores by depth for VWC, pH, EC, N, P, K, CEC, OM, and BD at 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F). The x-axis is the measured value of each soil parameter, volumetric water content (VWC), organic matter (OM) and nitrogen (N) are percents. Phosphorous (P) and potassium (K) are ppm and bulk density (BD) is  $\text{g cc}^{-1}$ . Each sample depth is represented by color.



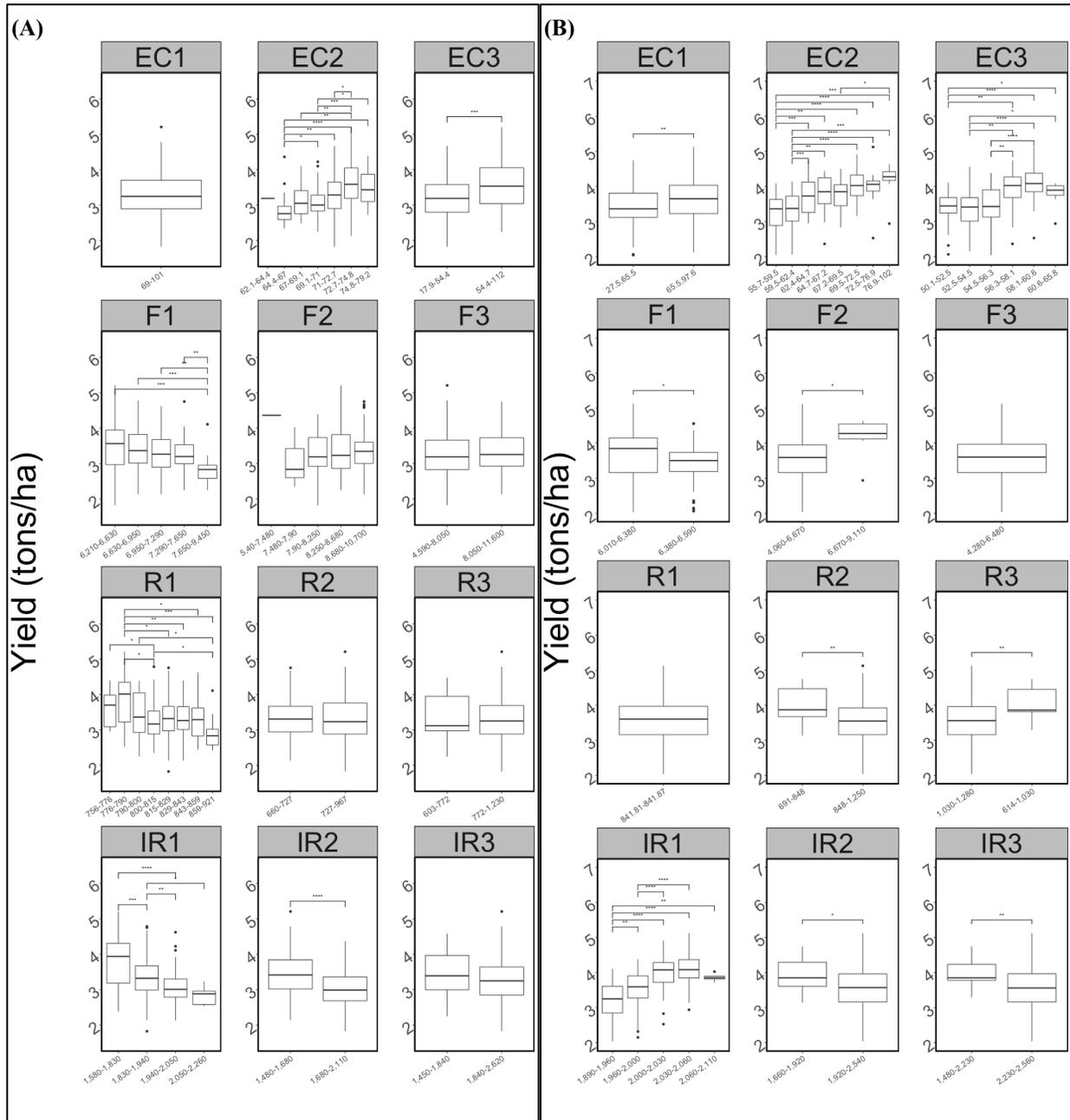
**Figure 3.5.** Distribution of soil values obtained from lab analysis of physical soil cores by depth for pH, EC, CEC, Ca, and Mg at 20THD (A), 20THI (B), 21RNN (C), 21RNS(D), 21THD (E) and 21THI (F) The x-axis is the measured value of each soil parameter, electrical conductivity (EC) is in  $\mu\text{S cm}^{-1}$ , and cation exchange capacity (CEC) is  $\text{cmol kg}^{-1}$ . Calcium (Ca), and magnesium (Mg) are ppm and pH is unitless. Each sample depth is represented by color.

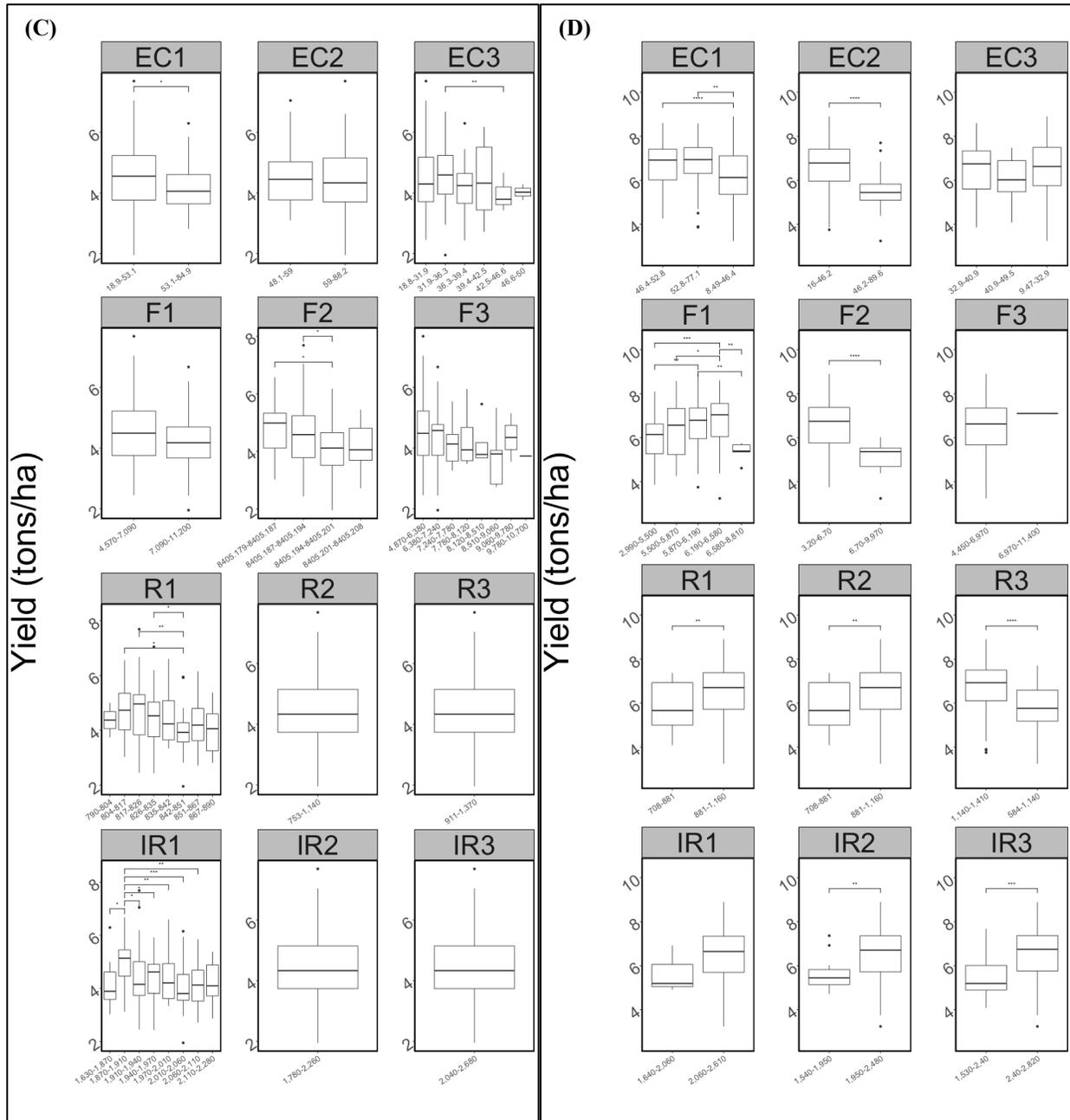


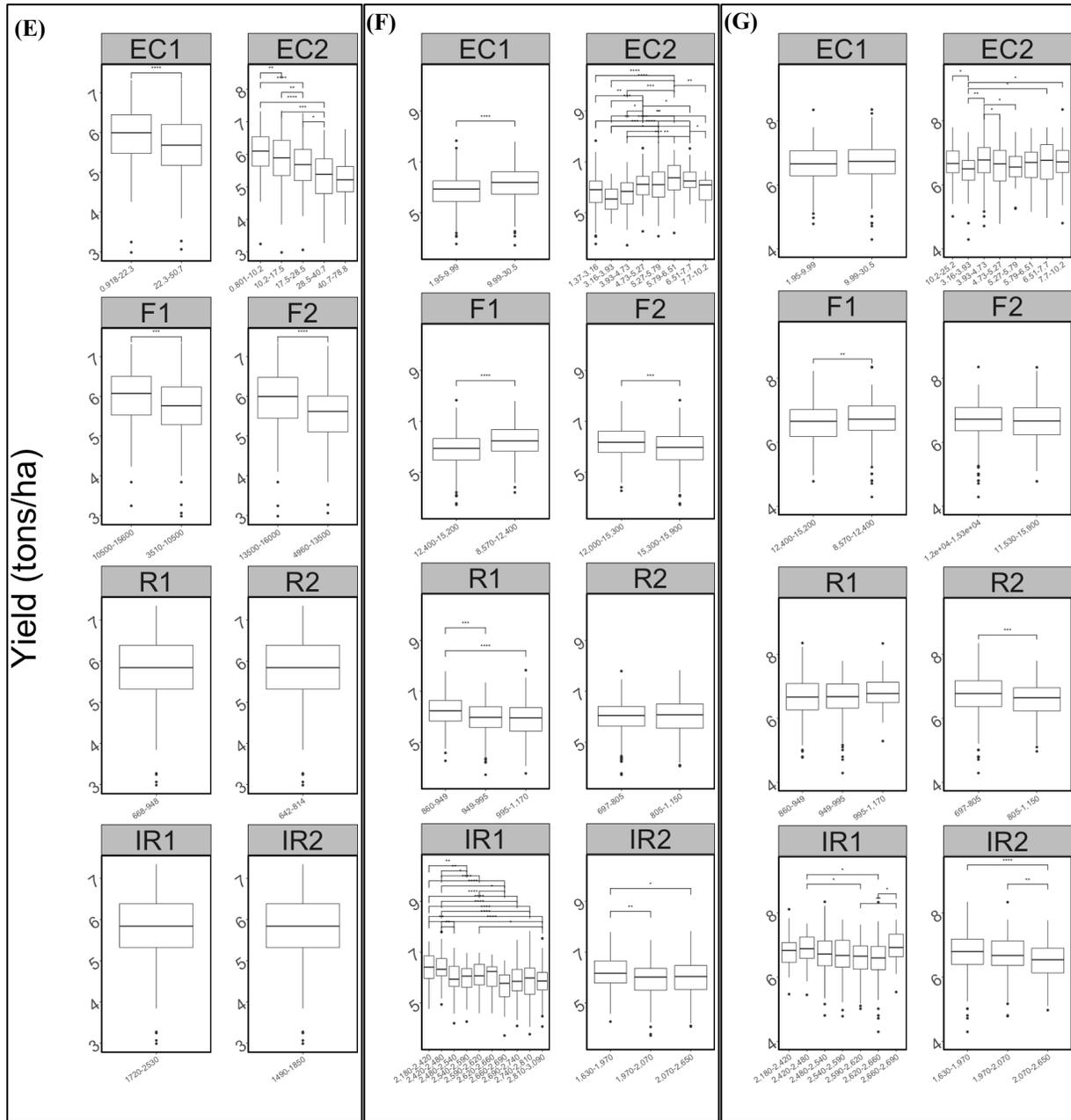
**Figure 3.6.** Soil texture distribution by location and year. Soil textural categories are in accordance to the USDA textural soil classification system (García-Gaines & Frankenstein, 2015).



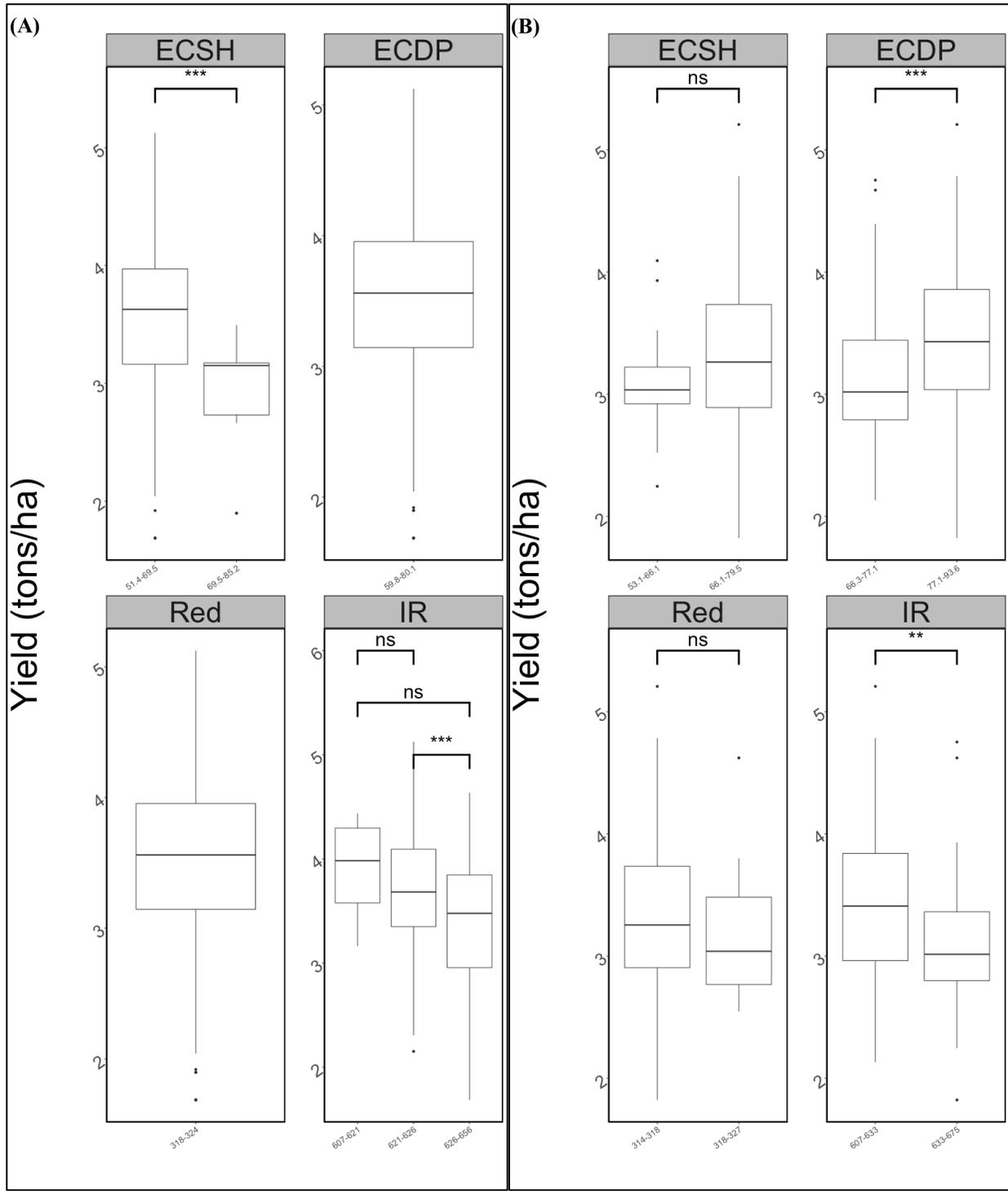
**Figure 3.7.** K-means cluster contour map for EC, Force, Red and IR measurements from the P4000 at 21RNS with PYN and YT plot map overlays. Number of color clusters were determined by k-means clusters and the values for the cluster color represent the median value of the cluster range.

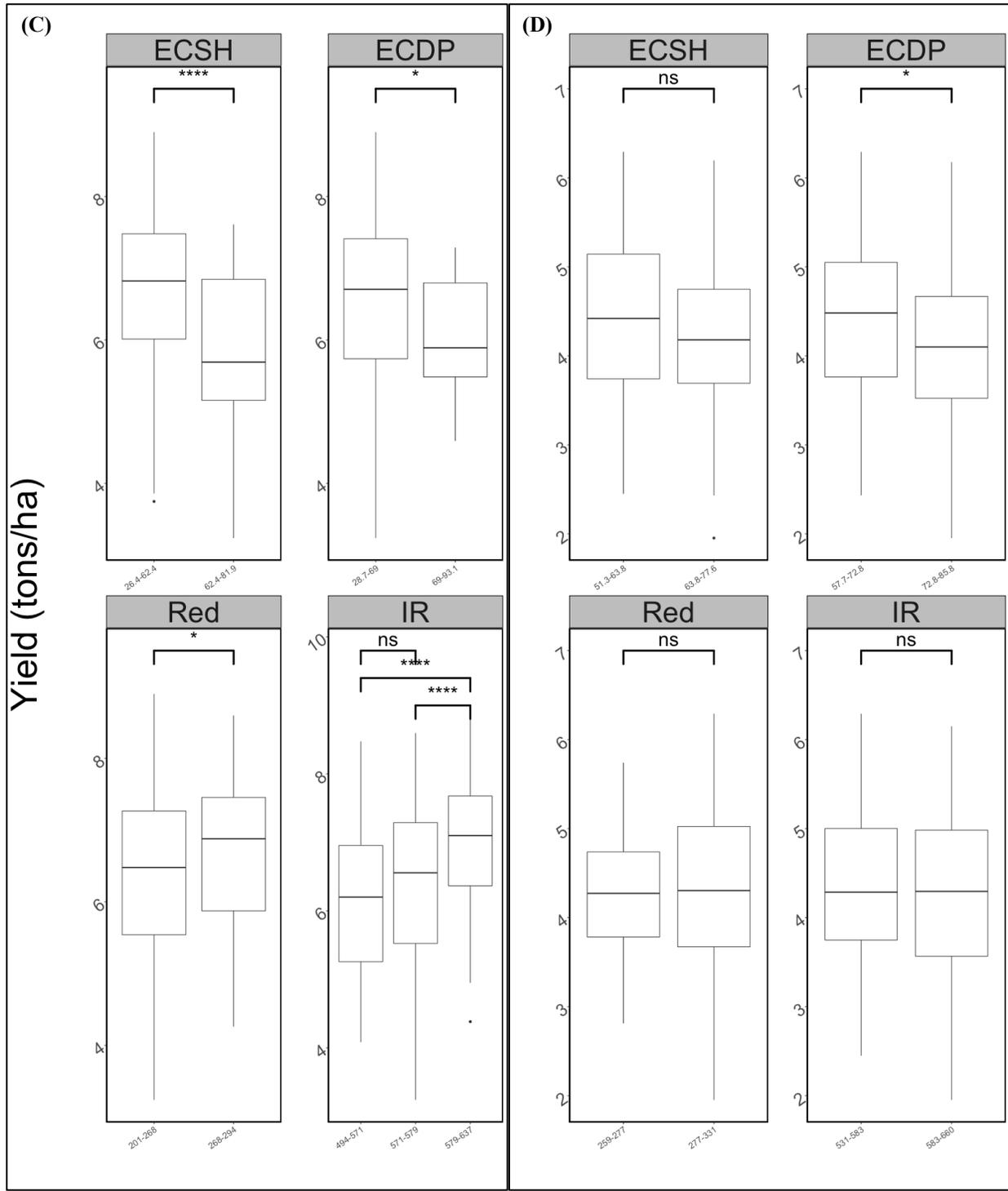


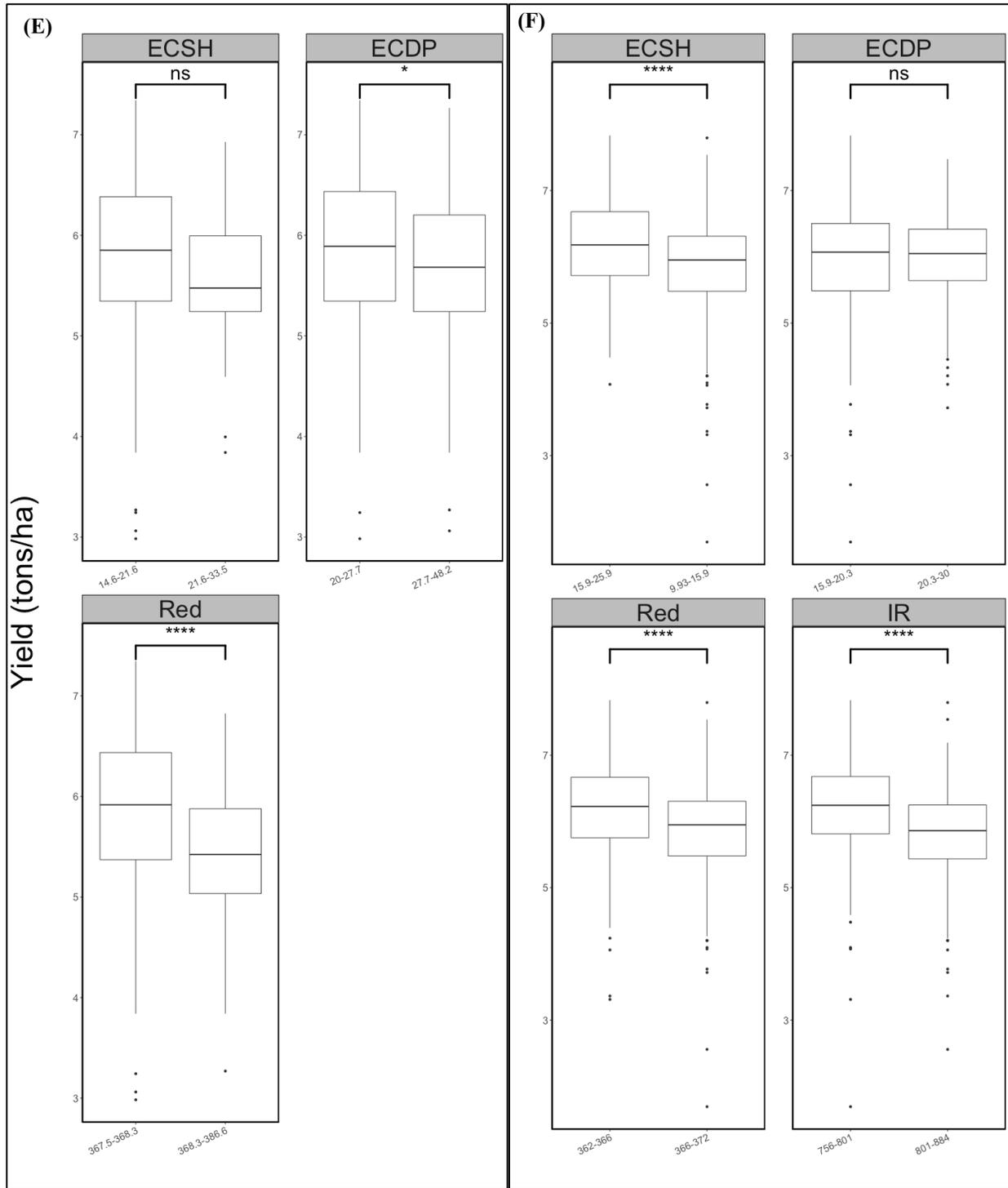


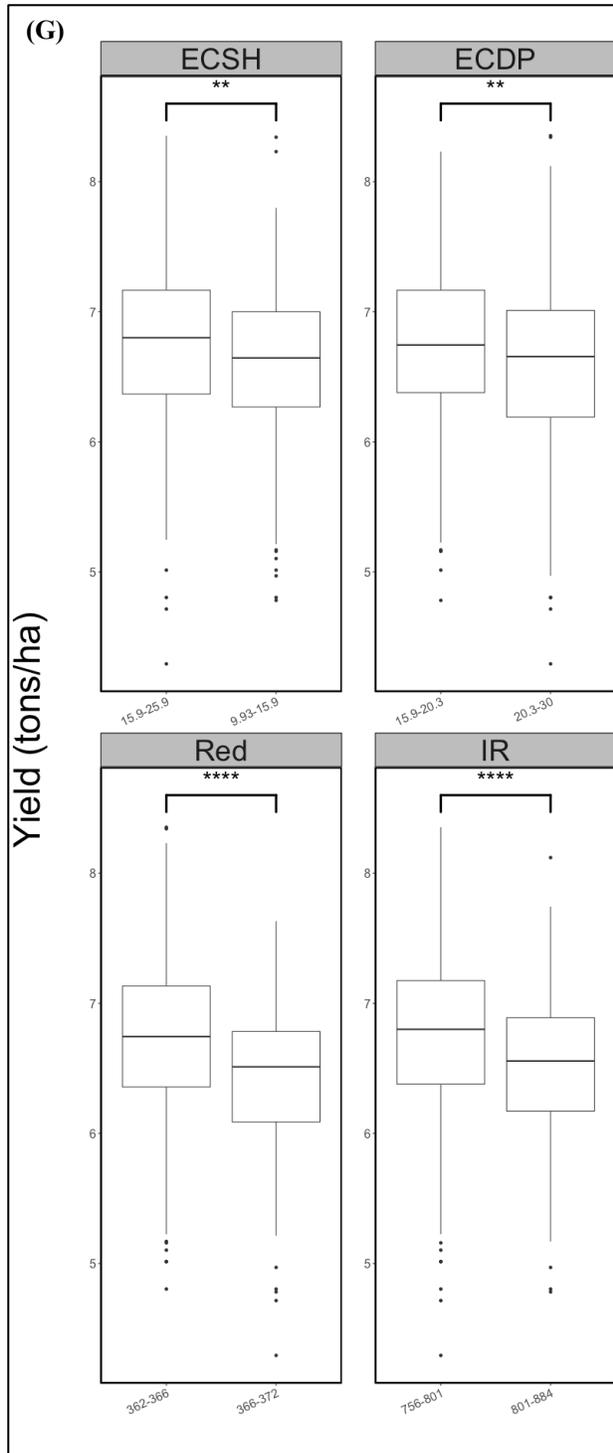


**Figure 3.8.** P4000 k-means cluster contour effects of EC, Force (F), Red (R) and IR on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-YT (F) and 21RNS-PYN (G). The title in the gray bar represents the soil measurement and depth. Depth-1 is 0-40 cm Depth 2 is 40-60cm and depth 3 is 60-100cm. The trait abbreviations are electrical conductivity (EC), force (F), red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S}/\text{cm}$ , force is kPa and both red and near infrared are digital reflectance values. Symbols denote Wilcoxon t-test significance where "\*\*\*\*\*"=0.001, "\*\*\*\*"=0.001, "\*\*\*"=0.01, "\*\*"=0.05 and ns= not significant.

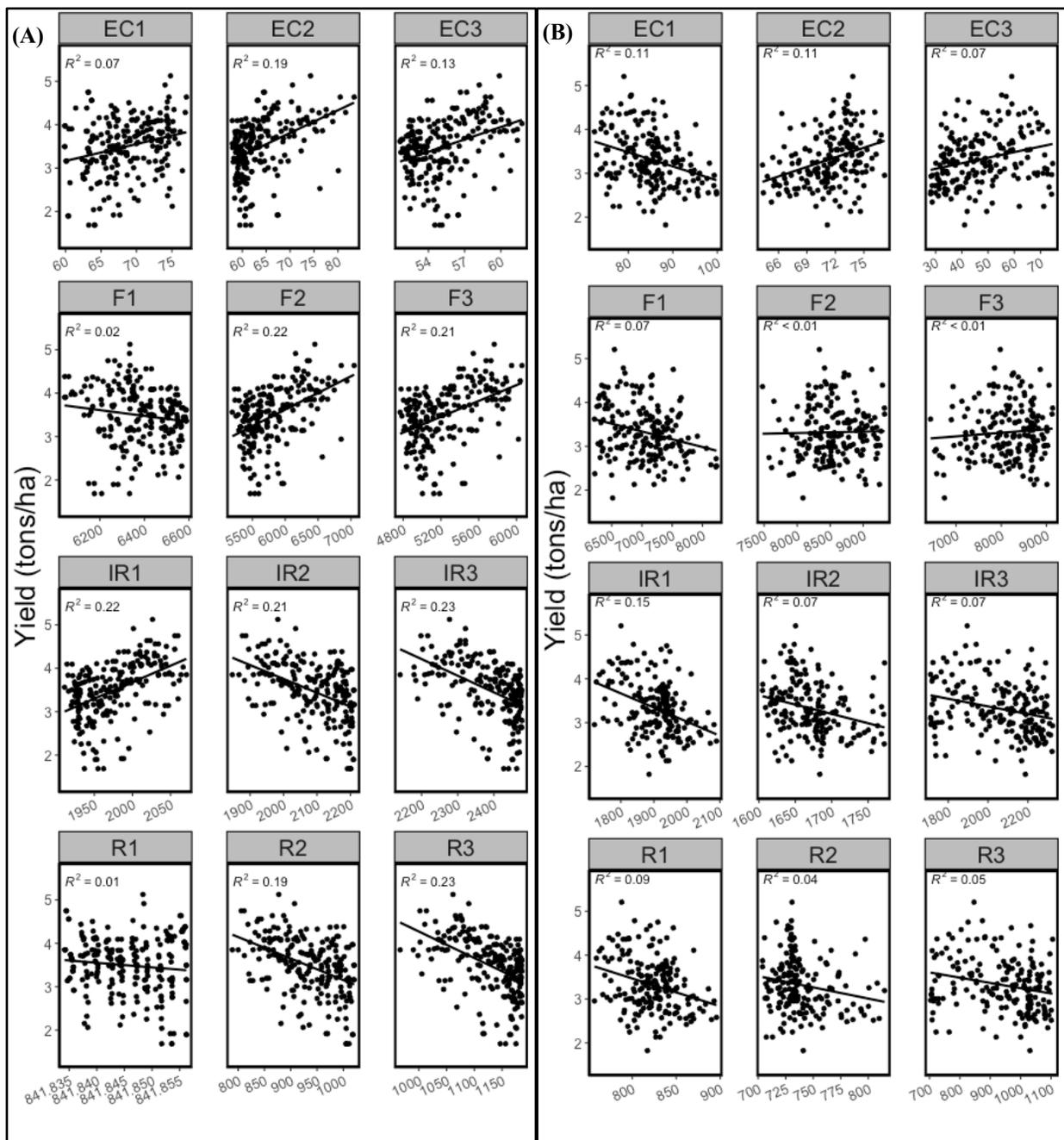


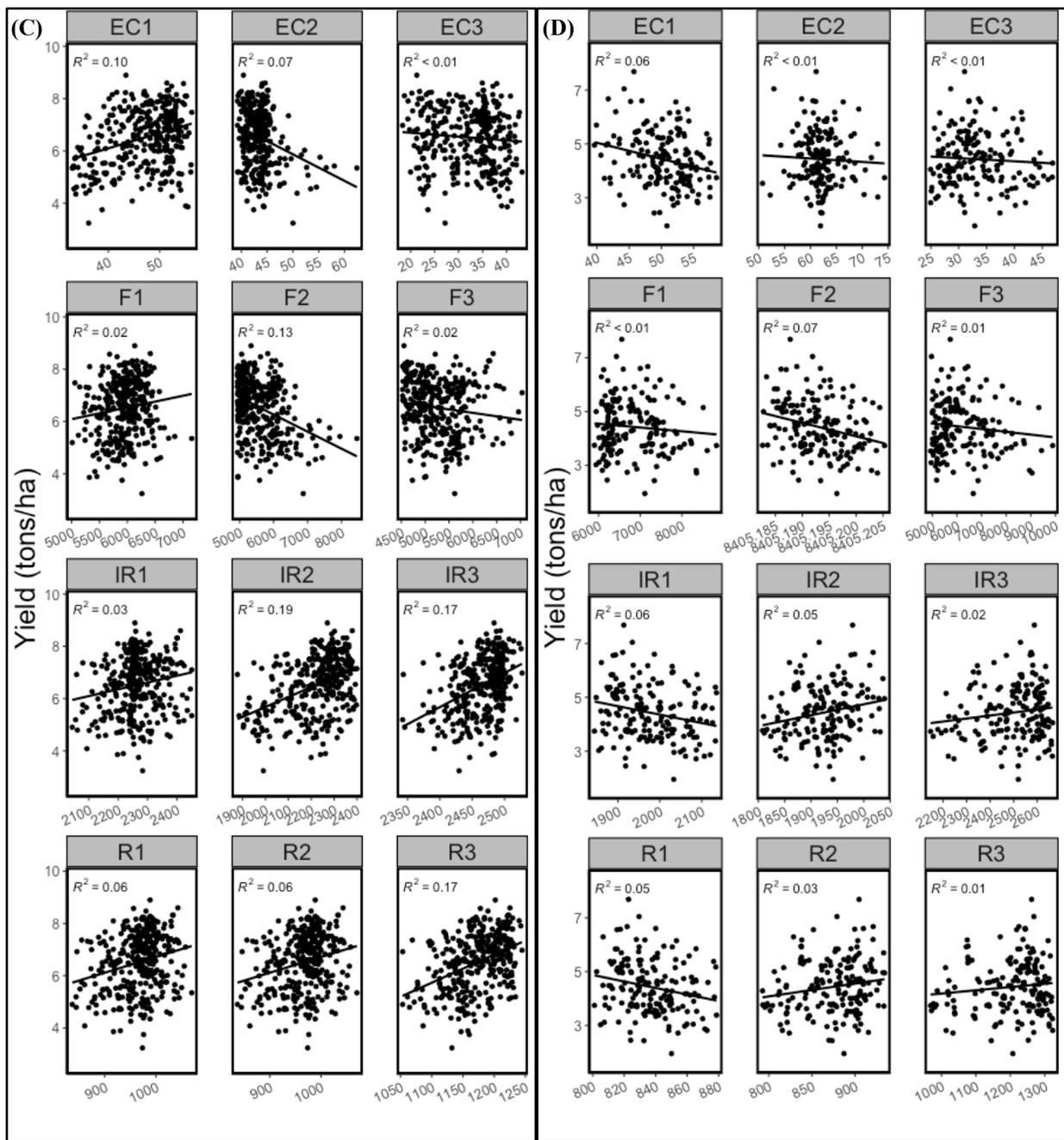


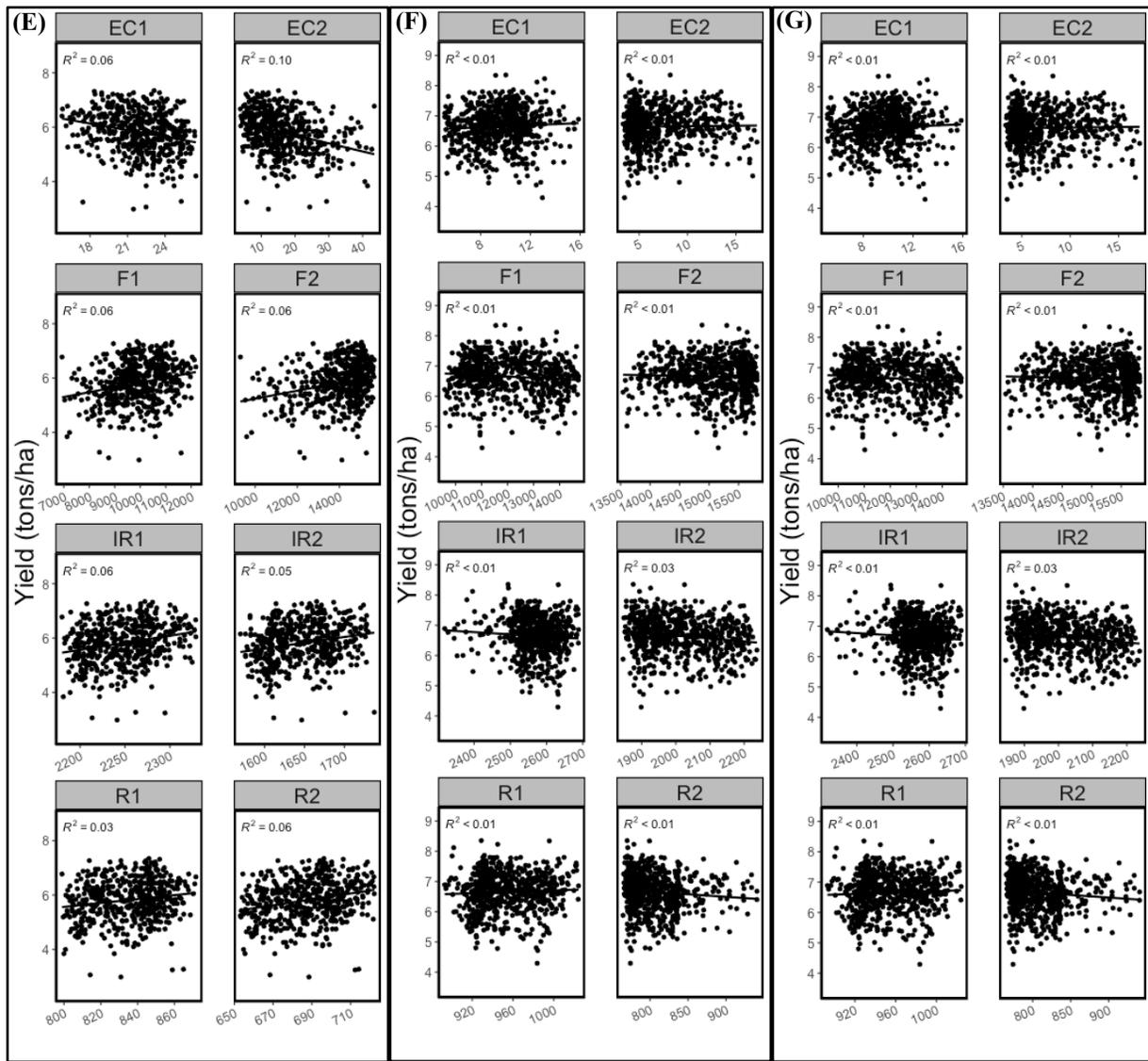




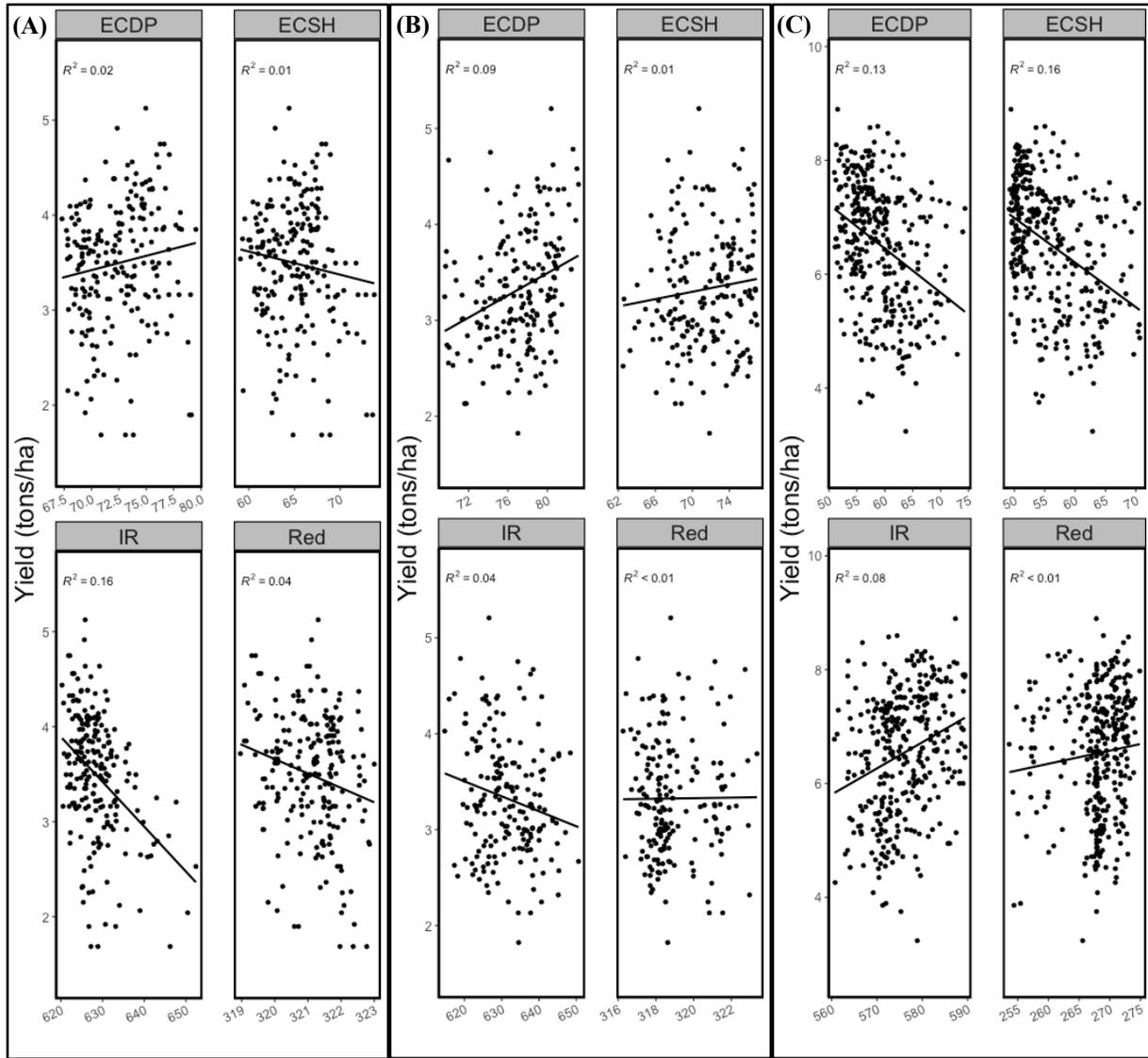
**Figure 3.9.** MSP3 k-means cluster contour effects of EC, Force (F), Red (R) and IR on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-YT (F) and 21RNS-PYN (G). The title in the gray bar represents the soil measurement and depth. ESH is 0-30 cm and ECDP is 0-90cm. The trait abbreviations are electrical conductivity (EC) , red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S}/\text{cm}$  and both red and near infrared are digital reflectance values. Symbols denote Wilcoxon t-test significance where “\*\*\*\*”=0.001, “\*\*\*”=0.001, “\*\*”=0.01, “\*”=0.05 and ns= not significant.

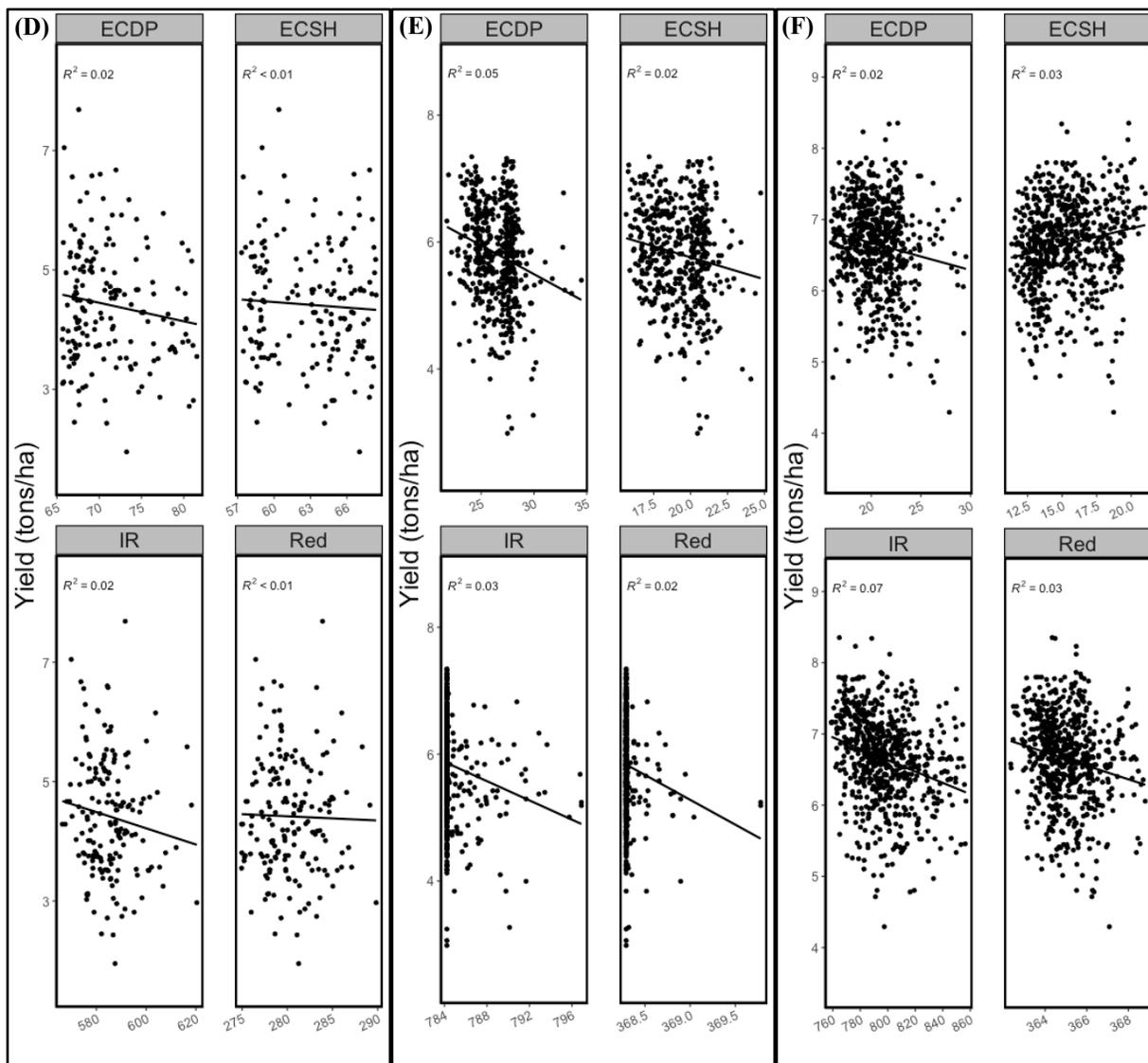


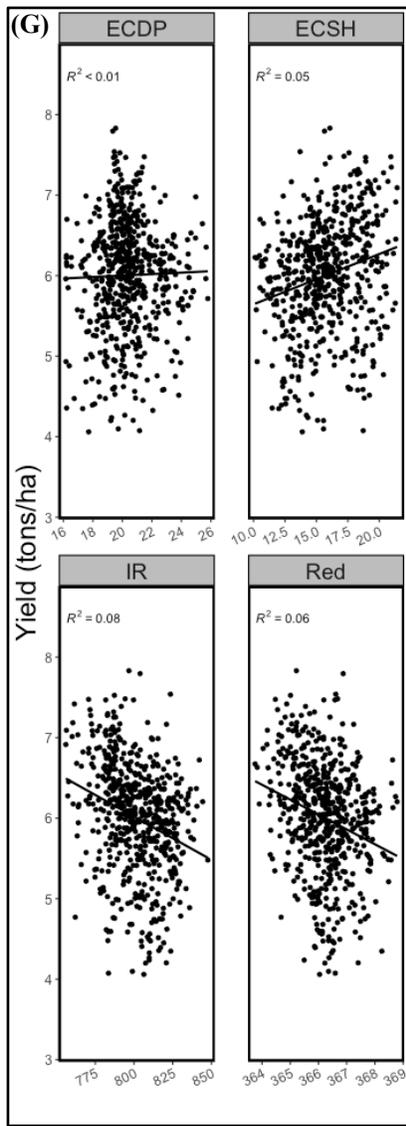




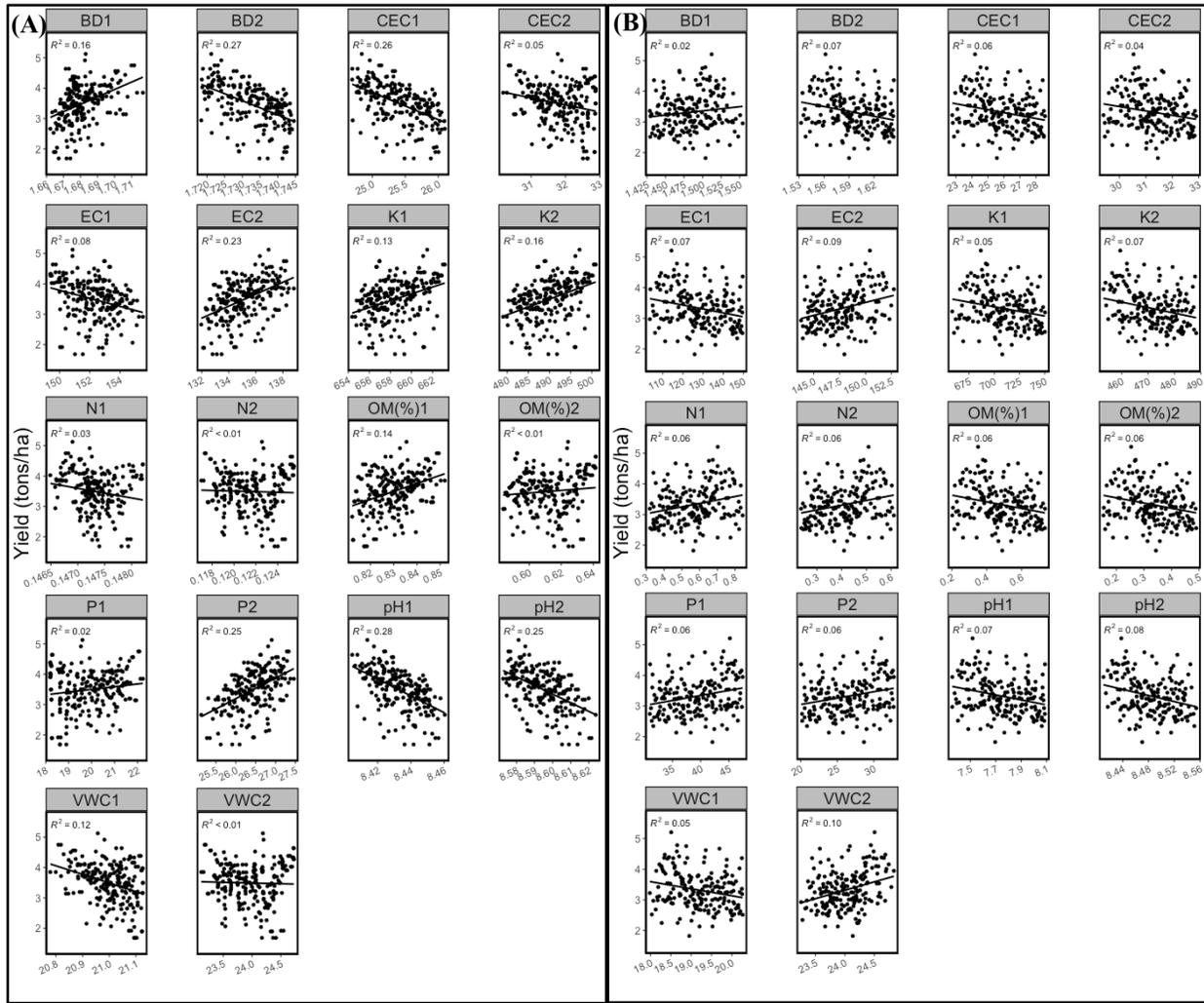
**Figure 3.10.** P4000 Kriged soil property correlation with grain yield of EC, Force (F), Red (R) and IR at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-PYN (F) and 21RNS-YT (G). The title in the gray bar represents the soil measurement and depth. Depth-1 is 0-40 cm Depth 2 is 40-60cm and depth 3 is 60-100cm. The trait abbreviations are electrical conductivity (EC), force (F), red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S}/\text{cm}$ , force is kPa and both red and near infrared are digital reflectance values.

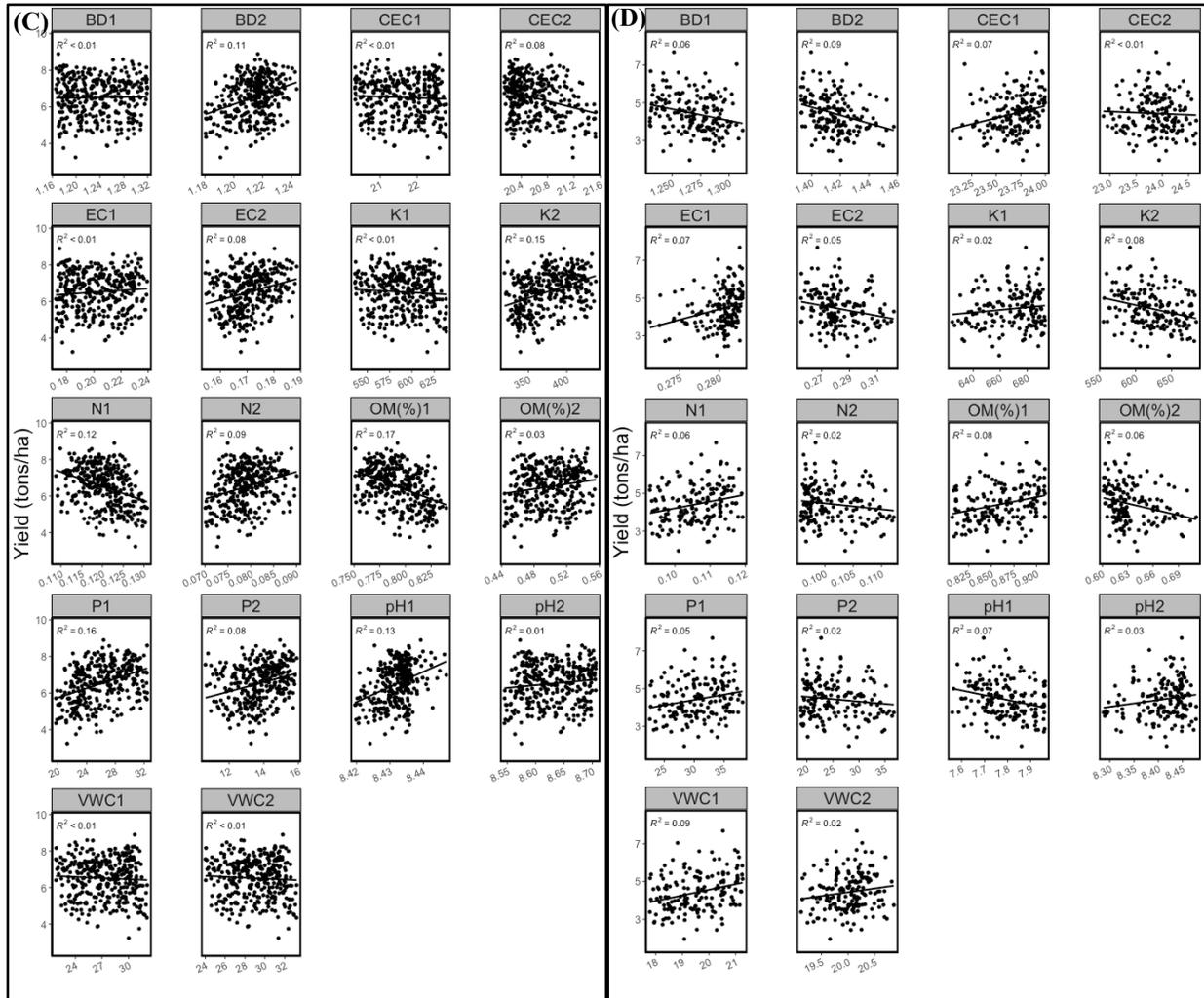




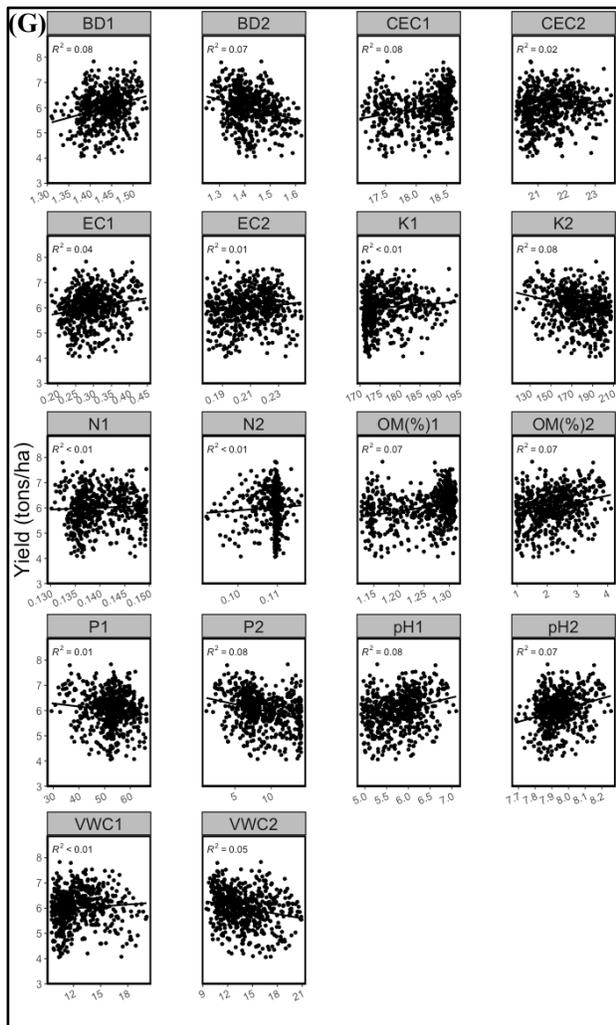


**Figure 3.11.** MSP3 Kriged soil property correlation with grain yield of EC, Force (F), Red (R) and IR on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-PYN (F) and 21RNS-YT (G). The title in the gray bar represents the soil measurement and depth. ECSH is 0-30 cm and ECDP is 0-90cm. The trait abbreviations are electrical conductivity (EC) , red (R), and near infrared (IR) The x-axis is the k-means cluster group for each individual parameter, EC is  $\mu\text{S/cm}$  and both red and near infrared are digital reflectance values.

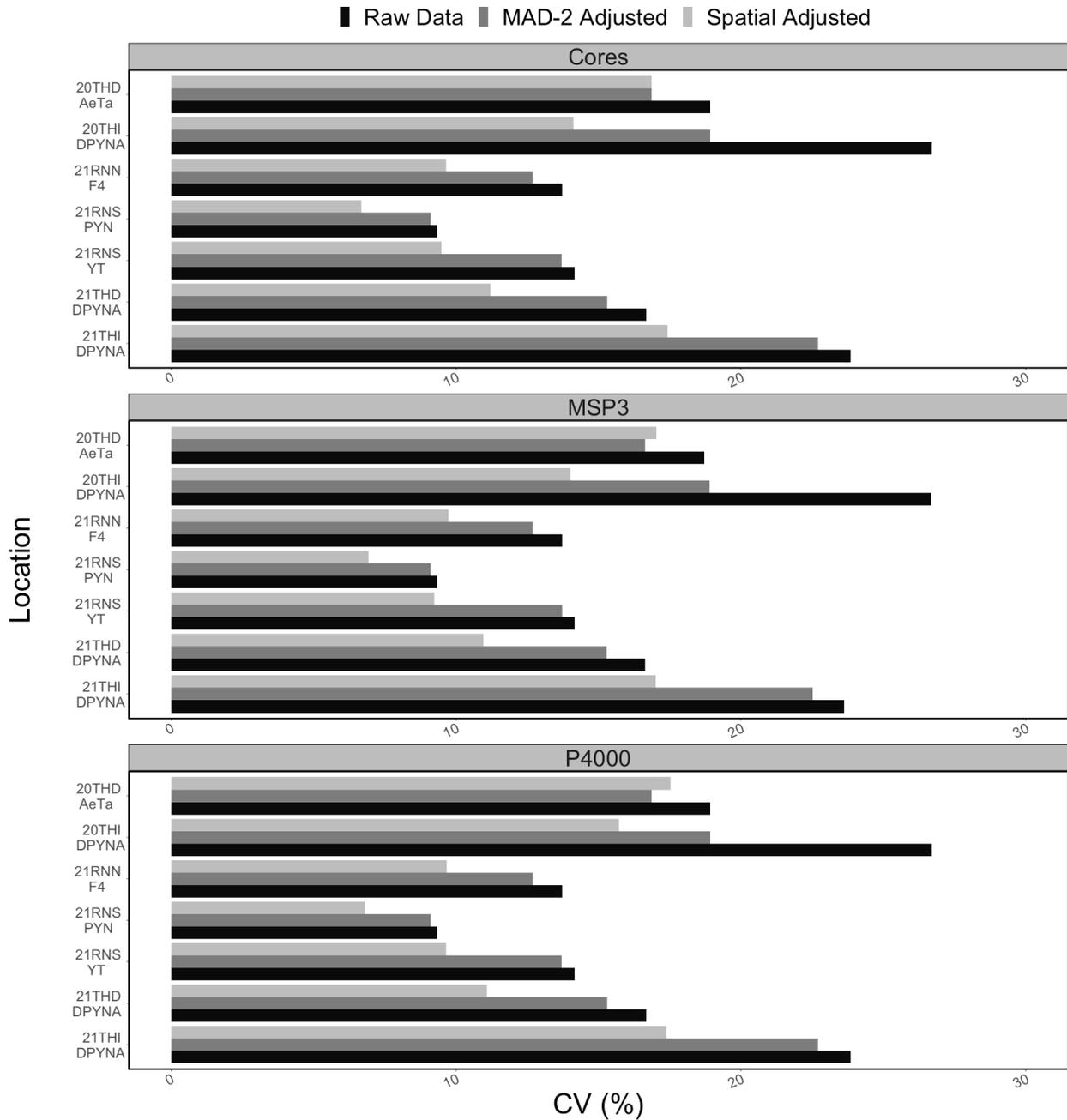








**Figure 3.12.** Soil analysis kriged soil property correlation with grain yield of BD, EC, CEC, K, N, OM, P, pH and VWC on yield at 20THD (A), 20THI (B), 21THD (C), 21THI (D), 21RNN (E), 21RNS-PYN (F) and 21RNS-YT (G). The title in the gray bar represents the soil measurement and depth. Depth-1 is 0-40 cm and Depth 2 is 40-60cm. The x-axis is the measured value of each soil parameter, volumetric water content (VWC), organic matter (OM) and nitrogen (N) are percents. Phosphorous (P) and potassium (K) are ppm and bulk density (BD) is  $\text{g cc}^{-1}$ . Electrical conductivity (EC) is in  $\mu\text{S cm}^{-1}$ , and cation exchange capacity (CEC) is  $\text{cmol kg}^{-1}$ . Calcium (Ca), and magnesium (Mg) are ppm and pH is unitless. Each sample depth is represented by color.



**Figure 3.13.** Coefficient of variation (CV%) by trial for raw data, data adjusted by the MAD-2 method 3, spatially adjusted yields and spatio-statistical adjusted yield. The x-axis represents CV% and the y-axis represents the location, year and experiment. The colored bars represent the type of spatial correction.

**Table 3.1-** Field experimental locations sample pattern, sensor platform and sample density for spatial soil sampling

Year	Location ID	Field Area (ha)	Sampling Grid Size (m)	Sensor Platform	Number of Observations
2020	THD	2.19	10*	Veris MSP3	3,379
			30 x 30	Veris P4000	120
			60 x 60	Soil Cores	17
2020	THI	2.74	10*	Veris MSP3	2,129
			30 x 30	Veris P4000	86
			60 x 60	Soil Cores	19
2021	RNN	3.15	10 x 10	Veris MSP3	4,651
			30 x 30	Veris P4000	120
			60 x 60	Soil Cores	20
2021	RNS	2.36	10 x 10	Veris MSP3	3,432
			30 x 30	Veris P4000	80
			60 x 60	Soil Cores	16
2021	THD	4.83	10 x 10	Veris MSP3	5,908
			30 x 30	Veris P4000	208
			60 x 60	Soil Cores	36
2021	THI	4.27	10*	Veris MSP3	3,432
			30 x 30	Veris P4000	70
			60 x 60	Soil Cores	20

\* Data was collected on single spaced transects opposed to grids

**Table 3.2-** Field experimental locations and weather station details for the Association Mapping Panel experiments.

Year	Location ID	Experiment ID	Experiment Area (ha)	Number of plots	Primary Check	Secondary Checks
2020	THD	Aeta	0.11	210	Western Blend	Joe WB 4792
2020	THI	DPYNA	0.14	196	Western Blend	AF03-1 KS12D0096-1
2021	RNN	F4	0.13	504	Central Blend	Zenda WB 4699
2021	RNS	PYN	0.33	692	Central Blend	Zenda WB 4699
2021	RNS	YT	0.43	588	Central Blend	Zenda WB 4699
2021	THD	PYN1A	0.16	336	Western Blend	Joe WB 4792
2021	THI	DPYNA	0.07	169	Western Blend	AF03-1 KD 1133

**Table 3.3-** MSP3 K-means clustering, variogram model type, sill and range for all locations and all years.

Location		K—means	Variogram Model		
Year	Measurement	clusters	Type	Sill	Range
20THD	Shallow EC	2	Linear	46.18	44.82
	Deep EC	2	Spherical	68.53	77.36
	Red	3	Exponential	9.31	7.18
	IR	2	Linear	70.76	6.57
20THI	Shallow EC	2	Exponential	19.38	25.60
	Deep EC	10	Exponential	19.12	20.49
	Red	2	Linear	5.18	5.50
	IR	2	Spherical	108.39	16.02
21HUS	Shallow EC	2	Exponential	9.69	27.01
	Deep EC	2	Spherical	4.88	29.42
	Red	7	Exponential	3.12	8.57
	IR	2	Exponential	645.51	15.63
21RNN F1	Shallow EC	2	Exponential	7.13	19.74
	Deep EC	2	Exponential	4.65	11.04
	Red	2	Exponential	9.41	3.34
	IR	3	Spherical	537.62	7.38
21RNN F2	Shallow EC	2	Exponential	24.59	13.88
	Deep EC	2	Exponential	136.24	24.24
	Red	3	Exponential	51.55	3.56
	IR	8	Spherical	296.50	7.71
21RNN F3	Shallow EC	3	Spherical	8.89	26.48
	Deep EC	2	Linear	24.85	15.95
	Red	2	Linear	17.16	4.01
	IR	4	Spherical	328.90	13.96
21RNN F4	Shallow EC	3	Exponential	101.22	29.38
	Deep EC	2	Exponential	201.06	27.91
	Red	2	Spherical	15.36	5.05
	IR	4	Exponential	910.44	26.49
21RNN Full	Shallow EC	2	Linear	118.41	65.09
	Deep EC	2	Linear	356.33	64.42
	Red	4	Exponential	34.04	8.15
	IR	2	Exponential	724.47	13.89
21THI	Shallow EC	2	Spherical	22.47	37.93
	Deep EC	2	Spherical	30.44	28.57
	Red	2	Gaussian	146.28	15.38
	IR	2	Gaussian	312.30	10.97
21THD	Shallow EC	2	Exponential	104.69	89.34
	Deep EC	2	Spherical	138.40	155.77
	Red	2	Linear	189.21	5.73
	IR	4	Exponential	213.02	5.27

**Table 3.4-** P4000 K-means clustering, variogram model type, sill and range for all locations and all years.

Location Year	Measurement	Sample Depth	K-means clusters	Variogram Model Type	Sill	Range	
20THD	EC	0-20	2	Spherical	203.57	6.46	
		20-60	5	Spherical	208.83	16.12	
		60-100	10	Linear	146.41	13.28	
	Force	0-20	10	Spherical	655331.50	18.38	
		20-60	2	Spherical	1162210.00	24.82	
		60-100	2	Spherical	2162789.00	28.33	
	Red	0-20	9	Gaussian	3435.08	25.18	
		20-60	2	Spherical	22922.76	77.97	
		60-100	2	Linear	45527.98	72.41	
	IR	0-20	6	Exponential	24319.55	25.72	
		20-60	2	Spherical	54928.93	72.70	
		60-100	2	Linear	87780.92	57.89	
20THI	EC	0-20	2	Spherical	127.68	36.64	
		20-60	3	Linear	136.09	23.16	
		60-100	9	Spherical	479.80	49.61	
	Force	0-20	2	Linear	1246620.00	8.13	
		20-60	10	Linear	1144286.00	14.59	
		60-100	9	Exponential	2419179.00	14.50	
	Red	0-20	10	Linear	7106.53	10.71	
		20-60	2	Spherical	3062.37	39.38	
		60-100	2	Exponential	51949.75	66.97	
	IR	0-20	8	Linear	55405.83	11.97	
		20-60	2	Exponential	12733.54	14.52	
		60-100	2	Exponential	144501.30	51.06	
21RNS	EC	0-20	2	Exponential	50.81	16.87	
		20-60	9	Exponential	36.56	11.24	
	Force	0-20	2	Spherical	3359784.00	24.12	
		20-60	2	Linear	857948.20	20.14	
	Red	0-20	3	Linear	4523.56	10.05	
		20-60	2	Spherical	5391.08	5.66	
	IR	0-20	10	Spherical	34758.32	8.51	
		20-60	2	Spherical	32611.45	18.98	
	21RNN	EC	0-20	2	Exponential	161.67	29.14
			20-60	5	Spherical	508.82	46.01
Force		0-20	2	Exponential	11091563.00	24.11	
		20-60	2	Spherical	9533657.00	57.18	
Red		0-20	2	Spherical	21632.19	59.39	
		20-60	2	Spherical	8593.81	48.01	
IR		0-20	2	Spherical	140358.50	63.64	
		20-60	3	Exponential	52648.88	17.54	
21THI	EC	0-20	2	Linear	226.86	8.97	
		20-60	3	Linear	289.77	8.59	
		60-100	9	Spherical	199.05	19.21	
	Force	0-20	2	Linear	2135037.00	12.82	
		20-60	10	Gaussian	1327583.00	18.93	
		60-100	9	Linear	5099018.00	6.38	
	Red	0-20	10	Linear	4307.55	13.97	
		20-60	2	Spherical	11974.96	47.09	
		60-100	2	Linear	54128.69	57.75	
		0-20	8	Linear	32719.71	12.26	
IR	20-60	2	Linear	30848.67	14.38		

		60-100	2	Spherical	112107.30	61.69
		0-20	2	Spherical	222.09	14.92
	EC	20-60	2	Spherical	160.70	9.82
		60-100	5	Spherical	213.28	21.36
		0-20	3	Linear	1206741.00	4.76
	Force	20-60	2	Linear	1441630.00	16.18
		60-100	4	Exponential	3051430.00	7.33
21THD		0-20	5	Exponential	9544.43	14.69
	Red	20-60	2	Exponential	74352.82	149.48
		60-100	2	Exponential	84451.22	143.01
		0-20	2	Exponential	31431.17	5.07
	IR	20-60	2	Exponential	108677.90	84.13
		60-100	2	Exponential	150455.80	102.29

**Table 3.5-** Coefficient of Variation (CV%) of yield data from seven trial across two years. Values are obtained from the raw yields, the experimental design spatial corrections, and spatial corrections from soil core, MSP3 and P4000 data.

<b>Year</b>	<b>Loc</b>	<b>Exp.</b>	<b>Raw Yield</b>	<b>MAD-2 Corrections</b>	<b>Soil Core Corrections</b>	<b>MSP3 Corrections</b>	<b>P4000 Corrections</b>
2020	THD	AeTa	18.9%	16.6%	16.9%	17.0%	17.5%
2020	THI	DPYNA	26.7%	18.9%	14.1%	14.0%	15.7%
2021	RNN	F4	13.7%	12.7%	9.7%	9.7%	9.7%
2021	RNS	PYN	9.3%	9.1%	6.7 %	6.9%	6.8%
2021	RNS	YT	14.12%	13.7%	9.5%	9.2%	9.7%
2021	THD	PYNA	16.7%	15.3%	11.2%	11.0%	11.1%
2021	THI	DPYNA	23.9%	22.7%	17.4%	17.0%	17.4%

## References

- Adamchuk, V. I., Ferguson, R. B., & Hergert, G. W. (2010). Soil heterogeneity and crop growth. In *Precision Crop Protection-the Challenge and Use of Heterogeneity* (pp. 3-16). Springer.
- Allard, R. W., & Bradshaw, A. D. (1964). Implications of genotype-environmental interactions in applied plant breeding 1. *Crop Science*, 4(5), 503-508.
- Atlin, G. N., Cairns, J. E., & Das, B. (2017). Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Global food security*, 12, 31-37.
- Babar, M., Reynolds, M., Van Ginkel, M., Klatt, A., Raun, W., & Stone, M. (2006). Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. *Crop Science*, 46(3), 1046-1057.
- Ben-Dor, E., Chabrillat, S., Demattê, J. A. M., Taylor, G., Hill, J., Whiting, M., & Sommer, S. (2009). Using imaging spectroscopy to study soil properties. *Remote sensing of environment*, 113, S38-S55.
- Blake, G. R., & Hartge, K. (1986). Bulk density. *Methods of soil analysis: Part 1 Physical and mineralogical methods*, 5, 363-375.
- Blum, A. (2018). *Plant breeding for stress environments*. CRC press.
- Bouyoucos, G. J. (1962). Hydrometer method improved for making particle size analyses of soils 1. *Agronomy Journal*, 54(5), 464-465.
- Braun, H.-J., Atlin, G., & Payne, T. (2010). Multi-location testing as a tool to identify plant response to global climate change. *Climate change and crop production*, 1, 115-138.
- Brevik, E. C., Fenton, T. E., & Jaynes, D. B. (2003). Evaluation of the accuracy of a central Iowa soil survey and implications for precision soil management. *Precision Agriculture*, 4(3), 331-342.
- Brevik, E. C., Fenton, T. E., & Lazari, A. (2006). Soil electrical conductivity as a function of soil water content and implications for soil mapping. *Precision Agriculture*, 7(6), 393-404.
- Brown, J. R. (1998). *Recommended chemical soil test procedures for the North Central Region*. Missouri Agricultural Experiment Station, University of Missouri--Columbia.
- Bui, E. (2013). Possible role of soil alkalinity in plant breeding for salt-tolerance. *Biology letters*, 9(5), 20130566.
- Butler, D., Cullis, B. R., Gilmour, A., & Gogel, B. (2009). ASReml-R reference manual. *The State of Queensland, Department of Primary Industries and Fisheries, Brisbane*.

- Chapman, H. (1965). Cation-exchange capacity. *Methods of Soil Analysis: Part 2 Chemical and Microbiological Properties*, 9, 891-901.
- Cho, Y., & Sudduth, K. A. (2015). Estimation of soil profile physical and chemical properties using a VIS-NIR-EC-force probe. 2015 ASABE Annual International Meeting,
- Correa, J., Postma, J. A., Watt, M., & Wojciechowski, T. (2019). Soil compaction and the architectural plasticity of root systems. *Journal of Experimental Botany*, 70(21), 6019-6034. <https://doi.org/10.1093/jxb/erz383>
- Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of agricultural, biological, and environmental statistics*, 11(4), 381-393.
- Diacono, M., Castrignanò, A., Troccoli, A., De Benedetto, D., Basso, B., & Rubino, P. (2012). Spatial and temporal variability of wheat grain yield and quality in a Mediterranean environment: A multivariate geostatistical approach. *Field Crops Research*, 131, 49-62.
- Elias, A. A., Robbins, K. R., Doerge, R., & Tuinstra, M. R. (2016). Half a century of studying genotype× environment interactions in plant breeding experiments. *Crop Science*, 56(5), 2090-2105.
- García-Gaines, R. A., & Frankenstein, S. (2015). USCS and the USDA soil classification system: Development of a mapping scheme.
- Griffith, D. A. (1987). Spatial autocorrelation. *A Primer (Washington, DC, Association of American Geographers)*.
- Haghighattalab, A., Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2017). Application of geographically weighted regression to improve grain yield prediction from unmanned aerial system imagery. *Crop Science*, 57(5), 2478-2489.
- He, Y., DeSutter, T., Prunty, L., Hopkins, D., Jia, X., & Wysocki, D. A. (2012). Evaluation of 1: 5 soil to water extract electrical conductivity methods. *Geoderma*, 185, 12-17.
- Heil, K., & Schmidhalter, U. (2017). Improved evaluation of field experiments by accounting for inherent soil variability. *European Journal of Agronomy*, 89, 1-15.
- Hezarjaribi, A., & Sourell, H. (2007). Feasibility study of monitoring the total available water content using non-invasive electromagnetic induction-based and electrode-based soil electrical conductivity measurements. *Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage*, 56(1), 53-65.
- Hoefler, R., González-Barrios, P., Bhatta, M., Nunes, J. A., Berro, I., Nalin, R. S., Borges, A., Covarrubias, E., Diaz-Garcia, L., & Quincke, M. (2020). Do spatial designs outperform classic experimental designs? *Journal of Agricultural, Biological and Environmental Statistics*, 25(4), 523-552.

- Hot, E., & Popović-Bugarin, V. (2015). Soil data clustering by using K-means and fuzzy K-means algorithm. 2015 23rd Telecommunications Forum Telfor (TELFOR),
- Isla, R., Aragüés, R., & Royo, A. (2003). Spatial variability of salt-affected soils in the middle Ebro Valley (Spain) and implications in plant breeding for increased productivity. *Euphytica*, 134(3), 325-334.
- Kalopesas, C., Galanis, G., Kalopesa, E., Katsogiannos, F., Kalafatis, P., Bilas, G., Patakas, A., & Zalidis, G. (2015). Spatial distribution of soil moisture in precision farming using integrated soil scanning and field telemetry data. EGU General Assembly Conference Abstracts,
- Kariuki, S. K. (2007). *Hard red winter wheat cultivar responses to a pH and aluminum concentration gradient*. Oklahoma State University.
- Khosla, R., Fleming, K., Delgado, J., Shaver, T., & Westfall, D. (2002). Use of site-specific management zones to improve nitrogen management for precision agriculture. *Journal of Soil and Water Conservation*, 57(6), 513-518.
- Kravchenko, A., & Bullock, D. (2002). Spatial variability of soybean quality data as a function of field topography: I. Spatial data analysis. *Crop Science*, 42(3), 804-815.
- Kumar, A., Bharti, B., Kumar, J., Bhatia, D., Singh, G., Jaiswal, J., & Prasad, R. (2020). Improving the efficiency of wheat breeding experiments using alpha lattice design over randomised complete block design. *Cereal Research Communications*, 48(1), 95-101.
- Kweon, G., Lund, E., & Maxton, C. (2013). Soil organic matter and cation-exchange capacity sensing with on-the-go electrical conductivity and optical sensors. *Geoderma*, 199, 80-89.
- Lamond, R., & Whitney, D. (1992). Management of Saline and Sodic Soils. MF-1022. In: Kansas State University, Cooperative Extension Service Manhattan, Kansas.
- Lin, C.-S., & Poushinsky, G. (1983). A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. *Biometrics*, 553-561.
- Lin, W., Jia-Ping, W., Yan-Xuan, L., HUANG, H.-Q., & Qian-Fang, F. (2009). Spatial variability of micronutrients in rice grain and paddy soil. *Pedosphere*, 19(6), 748-755.
- Liu, H., Colombi, T., Jäck, O., Keller, T., & Weih, M. (2021). Effects of soil compaction on grain yield of wheat depend on weather conditions. *Science of The Total Environment*, 150763.
- Lollato, R. P., Edwards, J. T., & Ochsner, T. E. (2017). Meteorological limits to winter wheat productivity in the US southern Great Plains. *Field Crops Research*, 203, 212-226.
- Lollato, R. P., Ochsner, T. E., Arnall, D. B., Griffin, T. W., & Edwards, J. T. (2019). From field experiments to regional forecasts: Upscaling wheat grain and forage yield response to acidic soils. *Agronomy Journal*, 111(1), 287-302.

- Lund, E., & Maxton, C. (2011). Proximal sensing of soil organic matter using the Veris® OpticMapper™. 2nd Global Workshop on Proximal Soil Sensing, Montreal, Quebec, Canada,
- Lund, E. D., Colin, P., Christy, D., & Drummond, P. E. (1999). Applying soil electrical conductivity technology to precision agriculture. Proceedings of the fourth international conference on precision agriculture,
- Mehlich, A. (1984). Mehlich 3 soil test extractant: A modification of Mehlich 2 extractant. *Communications in soil science and plant analysis*, *15*(12), 1409-1416.
- Miedaner, T., Reinbrecht, C., Lauber, U., Schollenberger, M., & Geiger, H. (2001). Effects of genotype and genotype—environment interaction on deoxynivalenol accumulation and resistance to Fusarium head blight in rye, triticale, and wheat. *Plant Breeding*, *120*(2), 97-105.
- Morgounov, A., Sonder, K., Abugalieva, A., Bhadauria, V., Cuthbert, R. D., Shamanin, V., Zelenskiy, Y., & DePauw, R. M. (2018). Effect of climate change on spring wheat yields in North America and Eurasia in 1981-2015 and implications for breeding. *Plos One*, *13*(10), e0204932.
- Munaro, L., Hefley, T., DeWolf, E., Haley, S., Fritz, A., Zhang, G., Haag, L., Schlegel, A., Edwards, J., & Marburger, D. (2020). Exploring long-term variety performance trials to improve environment-specific genotype× management recommendations: A case-study for winter wheat. *Field Crops Research*, *255*, 107848.
- Naderi-Boldaji, M., Sharifi, A., Hemmat, A., Alimardani, R., & Keller, T. (2014). Feasibility study on the potential of electrical conductivity sensor Veris® 3100 for field mapping of topsoil strength. *biosystems engineering*, *126*, 1-11.
- Nelson, D. W., & Sommers, L. E. (1996). Total carbon, organic carbon, and organic matter. *Methods of soil analysis: Part 3 Chemical methods*, *5*, 961-1010.
- Ng, W., Husnain, Anggria, L., Siregar, A. F., Hartatik, W., Sulaeman, Y., Jones, E., & Minasny, B. (2020). Developing a soil spectral library using a low-cost NIR spectrometer for precision fertilization in Indonesia. *Geoderma Regional*, *22*. <https://doi.org/ARTN e00319>  
10.1016/j.geodrs.2020.e00319
- Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Dor, E. B., Brown, D. J., Clairotte, M., & Csorba, A. (2015). Soil spectroscopy: An alternative to wet chemistry for soil monitoring. *Advances in agronomy*, *132*, 139-159.
- Nyine, M., Adhikari, E., Clinesmith, M., Jordan, K. W., Fritz, A. K., & Akhunov, E. (2020). Genomic patterns of introgression in interspecific populations created by crossing wheat with its wild relative. *G3: Genes, Genomes, Genetics*, *10*(10), 3651-3661.

- Pebesma, E., & Graeler, B. (2015). gstat: Spatial and spatio-temporal geostatistical modelling, prediction and simulation. *R package version*, 1.1-5.
- QGIS: A Free and Open Source Geographic Information System*. . [www.qgis.org](http://www.qgis.org)
- Qiao, C., Basford, K., DeLacy, I., & Cooper, M. (2000). Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theoretical and Applied Genetics*, 100(1), 9-16.
- Reynolds, M. (2002). Physiological approaches to wheat breeding. *FAO Plant Production and Protection Series*, 567.
- Rodriguez-Algaba, J., Sørensen, C. K., Labouriau, R., Justesen, A. F., & Hovmøller, M. S. (2020). Susceptibility of winter wheat and triticale to yellow rust influenced by complex interactions between vernalisation, temperature, plant growth stage and pathogen race. *Agronomy*, 10(1), 13.
- Rogers, J., & Carter, C. (1987). Soil core sampling for hydraulic conductivity and bulk density. *Soil Science Society of America Journal*, 51(5), 1393-1394.
- Roozeboom, K. L., Schapaugh, W. T., Tuinstra, M. R., Vanderlip, R. L., & Milliken, G. A. (2008). Testing wheat in variable environments: genotype, environment, interaction effects, and grouping test locations. *Crop Science*, 48(1), 317-330.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20, 53-65.
- Schloeder, C., Zimmerman, N., & Jacobs, M. (2001). Comparison of methods for interpolating soil properties using limited data. *Soil Science Society of America Journal*, 65(2), 470-479.
- Schwalbert, R., Amado, T., Reimche, G., & Gebert, F. (2019). Fine-tuning of wheat (*Triticum aestivum*, L.) variable nitrogen rate by combining crop sensing and management zones approaches in southern Brazil. *Precision Agriculture*, 20(1), 56-77.
- Smith, A., Lim, P., & Cullis, B. R. (2006). The design and analysis of multi-phase plant breeding experiments. *The Journal of Agricultural Science*, 144(5), 393-409.
- Soil Survey Staff, N. R. C. S., United States Department of Agriculture. (2021). *Web Soil Survey*. Retrieved 09/17/2021 from online at the following link: <http://websoilsurvey.sc.egov.usda.gov/>
- Stafford, J. (1999). An investigation into the within-field spatial variability of grain quality. Precision agriculture'99, Part 1 and Part 2. Papers presented at the 2nd European Conference on Precision Agriculture, Odense, Denmark, 11-15 July 1999,
- Sudduth, K., Drummond, S., & Kitchen, N. (2001). Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. *Computers and Electronics in Agriculture*, 31(3), 239-264.

- Tesar, M., & Jackobs, J. (1972). Establishing the stand. *Alfalfa science and technology*, 15, 415-435.
- Voltz, M., & Webster, R. (1990). A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *Journal of soil Science*, 41(3), 473-490.
- Warncke, D., & Brown, J. (1998). Potassium and other basic cations. *Recommended chemical soil test procedures for the North Central Region, 1001*, 31.
- Wu, Y., Huang, M., & Gallichand, J. (2011). Transpirational response to water availability for winter wheat as affected by soil textures. *Agricultural Water Management*, 98(4), 569-576.
- You, F. M., Song, Q. J., Jia, G. F., Cheng, Y. Z., Duguid, S., Booker, H., & Cloutier, S. (2016). Estimation of genetic parameters and their sampling variances for quantitative traits in the type 2 modified augmented design. *Crop Journal*, 4(2), 107-118. <https://doi.org/10.1016/j.cj.2016.01.003>
- Zarco-Tejada, P. J., Ustin, S., & Whiting, M. (2005). Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. *Agronomy Journal*, 97(3), 641-653.
- Zeng, Q., Sun, Y., Lammers, P. S., Ma, D., Lin, J., & Hueging, H. (2008). Improvement of a dual-sensor horizontal penetrometer by incorporating an EC sensor. *Computers and Electronics in Agriculture*, 64(2), 333-337. <https://doi.org/10.1016/j.compag.2008.05.012>

# **Chapter 4 - Rank order Phenotypic Selection for Wheat Breeding using UAVs**

## **Introduction**

Major technological advancements within plant breeding have the potential to allow breeders to evaluate larger populations and with higher confidence. However, despite these advancements most breeder selections are based on grain yield and visual selection, particularly for complex traits (Bentley & Mackay, 2017). Efforts continue to make both genotyping and phenotyping advancement to increase genetic gain and secure food stability for an increasing global population (Velu & Singh, 2013). Furthermore, both genomic and phenomic computational advancements need to be explored to maximize the potential of the digital age and overcome current bottle necks.

Sequencing technology has greatly improved in plant breeding, allowing for cheaper sequencing with higher coverage. Additionally, whole genome sequencing has allowed for multiple wheat reference genomes to be created (Walkowiak et al., 2020). This technology combined with advancements in genomic selection have shown potential to supplement traditional breeding selection and are currently being implemented in breeding programs for both grain yield and grain quality (Battenfield et al., 2016; Sehgal et al., 2020). However, it has been shown even with improvements, prediction accuracies of these models vary and are still dependent on high-throughput and high-precision phenotype data to improve the models. Furthermore, genomic selection accuracies will need to be higher in breeding programs with faster selection cycles (Poland et al., 2012).

Similar to sequencing, advancements in sensor technology and data management strategies have enabled substantial advancements in high-throughput phenotyping. Low cost, accessible options have provided opportunities for non-destructive plant phenotyping. Currently a number of phenotyping carts, tractors and UAV platforms have been effectively deployed in wheat breeding systems (Barker III et al., 2016; Beauchêne et al., 2019; Busemeyer et al., 2013). When equipped with multi-spectral, thermal and hyperspectral sensors these platforms are capable of capturing large amounts high spectral, radiometric and spatial resolution data with increased temporal resolution. However, implementation and sensor bottlenecks still exist, limiting the widespread adoption of these technologies in breeding programs (Song et al., 2021). Continued advancements in data management and computational pipelines, have the potential to overcome these bottlenecks and further the adoption of these technologies (Crain et al., 2021).

It is well demonstrated that HTP data can predict phenotypic traits such as height, lodging and even disease response (Singh et al., 2019; Su et al., 2018; Wang et al., 2018). Furthermore, combining HTP data with machine learning computational advancements, such as convolutional neural networks, has the potential to aid with genotype selection within a breeding program for heading date and spike detection (Fernandez-Gallego et al., 2020; Wang et al., 2019). Additionally, using vegetative indices (VI's) such as green normalized difference vegetation index (GNDVI), normalized difference vegetation index (NDVI) and normalized difference red edge (NDRE), with machine learning techniques have been used to quantify field variability for grain yield and quality prediction (Zhou et al., 2021). However, NDVI has been shown to have saturated values as crops develop and leaf area index (LAI) exceeds values of two (Haboudane et al., 2004). This saturation phenomenon has been shown to limit the ability to detect genotype specific classification when compared to NDRE (Bonfil, 2017). Likewise, it has

been reported that GNDVI can outperform NDVI correlations for grain yield and biomass, particularly at early reproductive stages (Gordillo-Salinas et al., 2021; Kyratzis et al., 2017). However, recent research shows that combined VI's have greater potential to correlate to physiological traits such as flag leaf nitrogen values in wheat (Eitel et al., 2008). Furthermore, it has shown that using a combination of GNDVI, NDVI and NDRE in machine learning applications can outperform single VI predictions for grain yield in wheat (Ramos et al., 2020).

In addition to incorporating multiple VI's it has been explored to incorporate multi-temporal points from key growth stages to improve trait prediction. In multiple crops, single date multi-spectral data collected around anthesis has shown to have the highest correlation to grain yield regardless of the VI index (García-Martínez et al., 2020; Kyratzis et al., 2017; Potgieter et al., 2017). However, it has been shown that using multi-temporal VI's to quantify the VI growth rate between two dates of key physiological dates such as vegetative to heading and head to grain fill can improve yield prediction (Bonfil, 2017). Similarly, (Zhou et al., 2017) accounted for multiple collection dates through VI summation and an accumulative VI index. Furthermore, it has been shown that incorporating multi-temporal collection points through multi-linear regression can improve yield prediction. Further progression has been shown by using machine learning tools such as LASSO, support vector regression and random forest regression in combination with multi-temporal VI data (Fu et al., 2020; Shafiee et al., 2021).

Ultimately a combination of genotype and phenotype technologies will be needed to maximize the technological advancements in plant breeding. Furthermore, these technologies need to be combined with data management and computational advancements. Developing solid pipelines that allow for easy implementation and adoptions within breeding programs will be pivotal moving forward (Kim, 2020).

## **Materials and Methods**

### **Experimental Locations**

Field experiments were conducted across seven locations over 5 years throughout the central corridor of Kansas (Fig. 4.1). However, not all seven locations were represented each year due to either limitations in UAV data collection or crop loss due to extreme environmental conditions such as late spring freeze and hail damage. Those site years were omitted since they would not represent normal production conditions and making selections on those sites would not be beneficial to the program. Despite the loss of some locations this experiment still covers eighteen site years with diverse temperature and moisture conditions (Fig 4.3). Furthermore, these locations represent prominent and distinct wheat growing regions making them priority target locations for the breeding program (Roozeboom et al., 2008).

### **Plant Material**

Diverse germplasm developed by the KSWB program was used to evaluate VI predication accuracies. Like all breeding programs, the KSWB program has a continuous cycle of germplasm evaluation and advancement demonstrated in Fig. 4.2. Since prediction at any advancement stage can save time and increase genetic gain, genotypes at all stages in the selection schematic were evaluated.

Early generation lines were grown in the Preliminary Yield Nursery (PYN) under a modified augmented design-type2. These smaller rectangular plots were 2.4 m x 1.4 m (3.36 m<sup>2</sup>) and contain between 100 and 200 unique genotypes. Advanced lines were grown in either the Advanced Yield Nursery (AYN), or the Kansas Interstate Nursery (KIN) within the scope of the Kansas State Wheat Breeding program. These larger plots were 4.5 m x 1.4 m (6.75 m<sup>2</sup>) and

grown in an alpha-lattice experimental design and typically contain 40 genotypes per experiment. Additionally, a segregating F<sub>5</sub> population was also grown in the larger plot size with an alpha-lattice experimental design. It is notable that there are multiple AYN and PYN grouping within each year.

The genotypes within a PYN, AYN or KIN experiment are fluid year to year. For example, in early generations the lines are either selected or dropped meaning the selected PYNs become AYN's while the unselected are removed from the testing cycle. This narrows the genetic diversity to more desirable and typically higher yielding genotypes. By the time a genotype reaches the KIN it has gone through several selection cycles. Furthermore, genotypes can be selected and reentered into the KIN experiment for several years prior to release.

## **UAS and Sensor Specifications**

All UAV phenotyping was completed using a DJI Matrice 100 (DJI, Shenzhen, China) quadcopter, equipped with a Micasense (MicaSense Inc. Seattle, WA) RedEdge-M multispectral sensor. The five-band multispectral sensor that captures images in the blue (455-495 nm), green (540-560 nm), red (658-678 nm), red-edge (RE; 707-727 nm), and near-infrared (NIR; 690-730 nm) regions of the light spectrum. To increase radiometric resolution all images were captured and stored as 16-bit tiff. GPS planting coordinates were used with the CSIRO mission planner application to create UAV flight mission plans. All flight plans were conducted using the Litchi mobile application. To ensure proper spatial resolution and image overlapping for image analysis a 20 m above ground level flight height and 2 m s<sup>-1</sup>, flight speed were maintained for all flights. Furthermore, the RedEdge-M capture rate is one second, when paired with the other flight parameters the resulting images contain an 80% overlap both the x and y directions of the flight pattern.

The individual images describe above were then used to create a composite orthomosaic using Agisoft Photoscan (Version 1.4.0, Agisoft LLC, Russia) as described in Wang et al. (2020). The stitching process blends pixels values of adjacent photos together resulting in one stitched image. To maintain true geospatial position, ground control points (GCPs) were used to make corrections to the orthomosaic image. White square tiles (90cm<sup>2</sup>) were, distributed to the outer corners of each trial and surveyed using an Emlid Reach RS (Emlid Ltd. Hong Kong), Real-Time Kinematic (RTK) Global Navigation Satellite System (GNSS) (Haghighattalab et al., 2016).

Individual five band reflectance values were extracted from the orthomosaic image using multi-step process in Quantum Geographic Information System (QGIS, [www.qgis.org](http://www.qgis.org)). Step one consisted of creating a polygon shape file with the plot-id for the targeted plot of interest. Next, the HTP Geoprocessor plugin described in (Wang et al., 2016) was used to extract the plot level digital and then were assigned to the plot-id. The extracted digital values were used to calculate, GNDVI, NDRE, and NDVI. All VI calculations are shown in Table 4.2.

### **Thermal Time Corrections**

Between spring green up and plant senescence, the target flight temporal resolution for all experiments was 10 to 14 days. However, travel time and equipment failure affected the revisit time. Furthermore, the RedEdge-M is a passive sensor meaning it needs the sun as an external light source and ideal flight time was within 1.5 hours before or after solar noon. This made environmental conditions such as wind and cloud cover limit flight dates.

Due to the diverse growing conditions cause by both location and year effects all flight dates were normalized based on growing degree days. Weather data for thermal time corrections was obtained from weather stations that are maintained by the Kansas Mesonet (Patrignani et al.,

2020). Based on results discussed in chapter 2 the growing degree day method (GDD1) thermal index with accumulation starting on January 1<sup>st</sup>, to avoid vernalization effects, was used. Four thermal time grouping targets were chosen for multi-temporal analysis 400, 700, 1200 and 1500 °Cd. Flight dates within 100°Cd of the target thermal time were assigned to that thermal grouping. These dates were chosen as 400°Cd is estimated to represent tillering while 700°Cd represents jointing, 1200°Cd represents heading or flowering and 1500°Cd represents mature kernels.

To capture temporal impacts with multiple VI's two different summarization methods were implemented. First the sum VI method described by Zhou et al. (2017) was used to quantify the cumulative VI between jointing and anthesis (400-1200 °Cd). Secondly, a normalized temporal VI between two thermal time was used to quantify growth rate between two dates (Bonfil, 2017). The summation method was only implemented for the 400, 700 and 1200 °Cd as these are key growing stages. Whereas 1500 °Cd is a maturity time point and higher values at this stage could have a negative impact. However, all thermal times were used for the two-point normalized temporal indices. Calculations for all the temporal indices are included in Table 4.2.

### **Spatial Corrections**

A 2D spatial p-spline method was implemented to evaluate the spatial trends of the raw phenotypic VI's. All spatial correction were completed with the R package 'SpATS' (Rodriguez-Alvarez et al., 2018). The SpATS model makes spatial correction through smoothing p-splines ANOVA in a two-dimensional interaction in one model. Additionally, the SpATS package has the capability to construct the variogram from the row and column parameters through the variogram function. The variogram produced from this function is based on the deviance of the column ( $cd_{ij}$ ), row ( $rd_{ij}$ ) and is the residual ( $v_{ij}$ ) displacement. In this

experiment a three dimensional variogram was construction from the row column deviance to visualize the variation (Fig. 4.7). The SpATS model also allows for the addition of environmental and genotypic effects resulting in equation 1.

$$y = f(u, v) + X_d b_d + Z_d h_d + \varepsilon \quad [\text{Eq. 1}]$$

Where is the  $y$  is the VI phenotypic response,  $f(u, v)$  is the smoothing p-spline function that accounts for row ( $u$ ) and column ( $v$ ) effects;  $X_d$  and  $Z_d$  are fixed and random blocking factor design matrices,  $b_d$  and  $h_d$  are fixed and random effect design matrix;  $\varepsilon$  is the residual errors assumed in a normal distribution. From equation 1 both Best Linear Unbiased Estimators (BLUEs) and Best Linear Unbiased Predictors (BLUP) values were calculated and used in the ensuing analysis.

Broad-sense heritability per line was calculated for each phenotypic collection date and VI combination using equation 2

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_e^2}{r}} \quad [\text{Eq. 2}]$$

Where  $\sigma_G^2$  represent the genotypic variance,  $\sigma_e^2$  is the residual error variance and  $r$  is the number of replications.

To account for spatial variability individual yield trials were analyzed with Agrobase Generation II software (Agrobase Generation II 2014, Agronomix, Winnipeg, MB, Canada; <https://www.agronomix.com/> ) according to the chosen experimental designs. The PYN experiments were analyzed as a modified augmented design type 2, using method 3 where adjustment are based on the slope form regressing the average of all primary and secondary check plots (Lin & Poushinsky, 1985). The KIN and AYN experiments were conducted as an alpha-lattice with 2 or three replicates(Yau, 1997). For both experimental designs the resulting BLUPs for lines tested accounted for site X trial X year interactions. Additionally, the

phenotypic VI values were also corrected based on the respective experimental design. To further assess rank correlation between VI and yield data, both raw and fitted data were placed in rank order by year, location, and experiment.

### **Statistical Analyses**

Linear models were used to assess both VI values and VI rank order effects with grain yield and grain yield rank. using the ‘lme4’ package in R (Bates et al., 2007).

$$y_i = \mu + \beta_i x + \varepsilon \quad [\text{Eq. 3}]$$

Where,  $y_i$  is the grain yield or grain yield rank,  $\mu$  is the overall mean,  $\beta_i$  is the coefficient estimator for the VI or VI rank of interest and  $\varepsilon$  is the residual errors assumed in a normal distribution. The accuracy of the models was evaluated using coefficient of determination ( $R^2$ ). Additionally a Pearson’s correlation ( $r$ ) using ‘GGally’ package in R (Schloerke et al., 2018) was conducted. Additionally, a Spearman’s correlation was conducted to assess the rank order correlation of VI’s to grain yield at all locations.

### **Random Forest**

To assess top yield group prediction ability a random forest classification analysis was performed using the ‘randomForest’ package in R (Liaw & Wiener, 2002). The classification method is used to predict non-numeric groupings of a data set. In this experiment a binary classification for yield groups was used where 1 indicated the genotype was in the top yield group and a value of 0 indicated the genotype was not part of the yield group. Yield groupings of the top 10%, 20% and 40% were evaluated for selection.

The dataset was randomly subset into a training and testing set at a 60%/40% split. The training set is used to create the model through a boot strapping method. To stabilize the error rate 500 trees were used for each random forest algorithm.

The performance of the classifications was evaluated through a confusion matrix which contains multiple metrics for model performance. The Cohen's Kappa statistic was calculated and used as a criteria for model fit. The measurement quantifies the diagonal parentages of the confusion matrix and compares those values to the expected percentages of random chance. The kappa values range from 0 to 1 whereas a value close to 0 indicates a poor or no fit and a value of 1 represents complete agreement.

$$K = \frac{\text{observed probability} - \text{chance probability}}{1 - \text{chance probability}} \quad [\text{Eq. 4}]$$

Model accuracy was calculated based on the number of correct predictions made in comparison to the total number of elements in the table.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP + FP + FN + TN)} \quad [\text{Eq. 5}]$$

Where TP is the true positives, TN are true negatives, FP are false positives and FN are false negatives. However, to fully understand the accuracy metric, the no information rate must also be considered. The no information rate metric indicates the value of majority classification and the probability of selecting that classification with no prior information. A hypothesis test was also conducted to determine if the model accuracy rate is greater than the no information rate. As a result of the large, unbalanced selection data in general the no information rate for these data sets are very high.

Other metrics calculated from the confusion matrix include the sensitivity and specificity of the model. The sensitivity measures the model's ability to predict TP in respect to TP and FN. A sensitivity value of 1 indicates that there are no false negative.

$$\text{Sensitivity} = \frac{(TP)}{(TP + FN)} \quad [\text{Eq. 6}]$$

The specificity or true negative rate metric assess the model's ability to quantify the TN in relationship to TN and FP. Due to the unbalanced nature of this data set both the sensitivity and specificity metrics need to be considered.

$$\textit{Specificity} = \frac{(TN)}{(FP + TN)} \quad [\text{Eq. 7}]$$

To evaluate the importance of each predictor VI a Mean Decrease Accuracy plot was constructed. The plot indicates the loss of accuracy that would occur if the variable was dropped from the model. The higher the loss of accuracy score the more important that variable is to the model. Additionally, the contain the Gini coefficient which measures the variables contributions to the homogeneity of the tree nodes. Like the accuracy score the higher Gini score indicates higher importance of the variable to the model. Both metrics are obtained from the out-of-bag portion of the data where the MSE and residual sum of squares is minimized.

## **Results and Discussions**

Despite all locations being grown through the central corridor of Kansas both location and year temperature and moisture variation impacted the yield environment. The 2018 growing season observed lower temperature throughout most of the growing season until May when the temperatures were elevated compared to the other four growing season (Fig. 4.3A). Additionally, the 2018 growing season had lower monthly precipitation compared to the other locations (Fig. 4.3B). These conditions resulted in the lowest average yield performance of all years. In total none of the years were statistically comparable and only 2020 Washington, was similar to 2017 McPherson and 2021 Saline (Fig 4.4). This potentially could impact prediction and correlation of VI's to yield performance as it has been shown using dissimilar predictor and training target environments reduces prediction accuracy (Cooper et al., 1997). However, developing a VI model that works across yield groups would be highly impactful for a breeding program.

Temporal trends in VI distribution were observed at all locations. With all indices the trend was to increase VI from 400 to 700 and decrease from 700 to 1200 and 1200 to 1500 °Cd (Fig 4.5). Although the decrease in VI values from 1200 to 1500 °Cd was greater than any other thermal time interval, which is consistent with other UAV and remote sensing research (Hassan et al., 2019). Furthermore, the range VI's at 1200 and 1500 °Cd in general are wider than at 400 and 700 °Cd. Heading (1200 °Cd) is a critical growth stage for overall yield potential of winter wheat, particularly in Kansas where fertility can be impacted by heat (Prasad & Djanaguiraman, 2014). With a greater genotype response at this stage it could lead to significant yield detection. This is also consistent with literature that has found VI at or near anthesis have the greatest correlation to grain yield (Wang et al., 2014). It is also notable that the GNDVI and NDVI indices approach saturation early in the growing season which may narrow the potential for genotype specific detection (Haboudane et al., 2004). However, both indices have been shown to correlate to grain yield and have been widely used as tools in breeding programs (Haghighattalab et al., 2016).

In general heritability values were high for all, VI, year and flight date. However, differences were observed in both VI heritability between flight dates within year and across years (Fig 4.6). Overall, 2018 (0.529) and 2021 (0.545) had the lowest mean heritability with 2017 (0.784), 2019 (0.721) and 2020 (0.707) all having similar full season heritability. It is interesting to note that the two lowest heritable VI years were also the highest and lowest yielding years during this study. Whereas the three middle yielding groups all had similar heritability. In contrast, heritability by VI differences were not observed as GNDVI (0.651), NDRE (0.649) and NDVI (0.671) had similar mean heritability across all years. This indicates

that the environmental condition of the location year has a greater impact than the VI on heritability.

Previous studies have shown that large scale breeding nurseries can be impacted by spatial differences in soil, micro-climates and past management (Miller et al., 1988). The variogram (Fig. 4.7) shows that significant VI spatial trends were observed for the NDRE, VI at Reno County. This trend was observed for all traits at all location years. The positive and negative spatial deviation varies with the location in the field, which impacts raw VI values of some genotypes more than others, particularly in partial or single rep experiments. The advantage the p-splines method describe by Rodriguez-Alvarez et al. (2018) provides is that spatially adjusted values, in addition to BLUP and BLUE values, are calculated in a single step. This single step correction is visualized in Fig. 4.8 where the global and local trends are corrected to provide distributed genotypic BLUP values. Overall, the combination of observed spatial trends and previous reports of spatial adjustments for plant breeding applications, supports the use of automated spatial correction tools for VI data.

A Pearson's correlation was conducted on whole locations to assess the correlation of thermal time flight date VI's to both corrected and raw grain yields. Additionally, a Spearman correlation was conducted on whole locations to assess rank order values. The rank order is of interest since breeders are more interested in relative yield performance rather than the absolute value when making selections. This was done for both spatially fitted and BLUP values. Although, significant correlation was observed for all absolute values within a location for both fitted (Fig. 4.9) and BLUP VI (Fig C.1) in general the correlation for the fitted data was better. Overall, although significant, correlation in general was low. Similar observations were observed with the Spearman rank order correlation. In some instances, the rank correlation was improved,

however, the magnitude and directionality of the change was inconsistent. This indicates the rank order correlation follows a similar trend to VI/grain yield correlation.

From the correlation matrix (Fig. 4.9) and Table 4.3 the 1200<sup>0</sup>Cd flight date had the highest correlations 61.5% of the time while the 700<sup>0</sup>Cd flight date accounted for the other 31.5% of the time. However, when assessing the rank order correlation, the highest correlation by flight time was flipped with the 700<sup>0</sup>Cd flight date accounting for the highest correlation 61.5% and the 1200<sup>0</sup>Cd flight or the other 31.5% of the time. Although the flight times were categorized by thermal time the variance in both the thermal time stage prediction and actual match of thermal time to the group likely contributed to this inconsistency. However, from these correlations it is evident that both the 700 and 1200<sup>0</sup>Cd are highly correlated to genotype performance compared to 400 and 1500<sup>0</sup>Cd thermal times. This is consistent with previous reports as the 700 and 1200<sup>0</sup>Cd are closely related to reproductive growth stages of wheat.

In addition to VI correlations the  $\Delta$ VI and  $\Sigma$ VI values also showed correlation to grain yield (Fig 4.11). However, like the individual VI's the correlation improved with the rank order compared to the actual grain yield. In all cases the  $\Sigma$ VI had higher correlation than any of the  $\Delta$ VI period indices. Of the  $\Delta$ VI indices the 700 to 1200<sup>0</sup>Cd period had the highest correlation which is also consistent with the individual correlations. In contrast the 1200 to 1500<sup>0</sup>Cd period had the lowest correlation and had a negative correlation to the other time periods. As this period is associated with grain fill and maturity it is likely that a combination of stay green and later maturing genotypes affected the VI's correlation to grain yield, although the rank order was strong (Christopher et al., 2008; Cox et al., 1988).

Due to the genetic diversity and difference in prior selection intensity of each experiment linear models were used to assess rank order by experiment at each location for 2021. Overall R<sup>2</sup>

values were low ranging from  $<0.01$  to 0.40 (Table 4.5, Fig. 4.11). Additionally, to evaluate the rank as a potential selection tool, the percent of lines needed to be kept to capture the top 10% of the highest yielding genotypes was evaluated. Similar to the linear model performance the rank order numbers were poor and a range between 16% and 100% of the lines needed to be kept to capture the top 10% (Table 4.6). Although 16% would be an acceptable keep rate, only 12 of the potential 216 temporal, VI combinations resulted in keep rates lower than 50%.

Overall random forest prediction accuracies depended on yield grouping and VI's used (Table 4.7). The top 10% yield groups had the highest predictions accuracies (0.890-0.899), followed by the top 20% (0.775-0.797) and top 40% (0.599-0.629) yield groupings. While these prediction accuracies are high, that metric alone cannot quantify the performance of the random forest model. Further comparison of the accuracies to the no information rate shows that in most cases the random forest actually performs worse than if the largest proportioned class was chosen every time. Furthermore, the poor performance of these models is reflected in the hypothesis testing where the large p-values conclude that the none of the models are significantly better than the no information rate. Finally, the near zero values of the Cohen's Kappa values for all models indicates that models are not good fits compared to random probability.

Despite the overall poor performance of the models, some models still show potential and warrant further study, as the sensitivity of the model had overall good performance. Recalling from equation 6, sensitivity measures a model's ability to detect true positives. In this study a true positive occurs when an unselected genotype in the test data set was also unselected in the predicted data set. A false negative occurred when an unselected genotype in the test data set was selected in the predicted dataset. Although, the split nature of all these datasets favors high sensitivity values, due to the majority of the genotypes being unselected, all the test populations

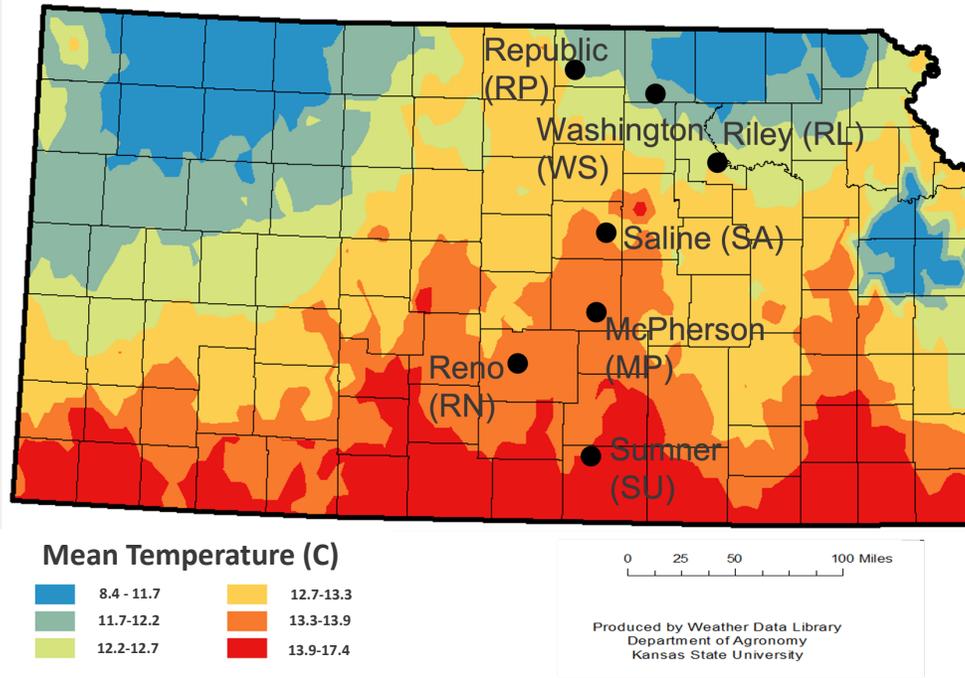
had higher sensitivity than the actual construction of the dataset. Meaning the observed false negative rate was low. Inversely, the specificity numbers were poor, indicating that the false positive rate was high in respects to the true negative rate. A false positive is considered a genotype that was selected in the real dataset but unselected in the predicted set. While a true negative was a genotype that was selected in the real dataset and the predicted dataset. While the high sensitivity numbers indicate only a small portion of the unselected genotypes from the data set were selected in the predicted data set the specificity numbers indicate that a large portion of the selected genotypes in the data set were unselected in the predicted data set.

The random forest function also provides an importance score for each variable used. This score relates to the predictive power and the order of importance the variable has on the model's performance. The prediction accuracy is based on the out-of-bag sampling during node construction of each tree. A rank pattern can be assembled by incorporating all VI and temporal combinations into the random forest model for each categorical yield group. As seen in Fig 4.12 the NDRE value at the 700<sup>0</sup>Cd time point had the highest importance for the top 10% and 20% yield groups but had lower importance for the top 40% group. Overall, importance of both VI's and temporal points were inconsistent across all three models, indicating no one set of VI or temporal points is collectively more important. With the generally poor and inconsistent model, the lack of well-defined input performance maybe expected.

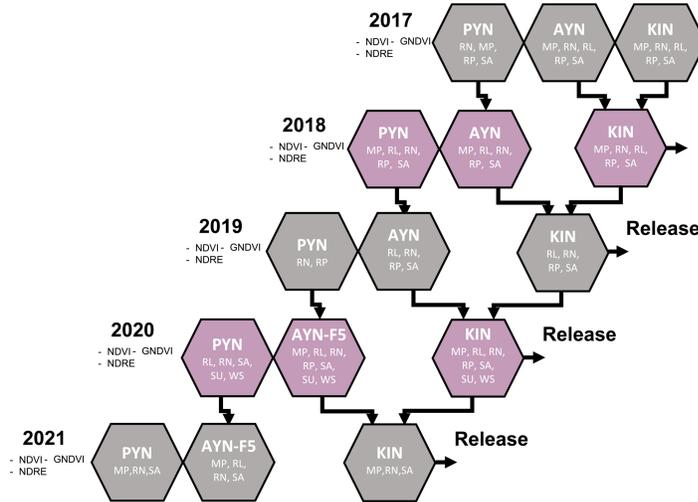
## **Conclusions**

This experiment confirms that UAV spectral VI's are correlated to grain yield of genotypes within KSWB program. Furthermore, this study highlights the potential of using HTP systems with targeted temporal resolution for potential future machine learning applications. However, this study still shows that fluid nature and the genetic diversity of a breeding program

poses challenges to implementing these techniques on a large scale. Further research is needed to incorporate other phenotypic traits with VI's to find ideal trait combinations for selection prediction. Furthermore, continued studies on G x E impacts on HTP data collection and model incorporation are needed.



**Figure 4.1.** Experiment locations for the Kansas Wheat Breeding experiments over a five-year period. Temperature contours are derived from the 2020 observed mean monthly temperatures. The maps shows the north to south temperature gradient for the locations that impacts the cumulative growing degree units used to normalize UAV flight dates.



**Figure 4.2.** Schematic diagram of proposed experimental selection using UAV VI's within the KS Wheat Breeding program across 5 years.

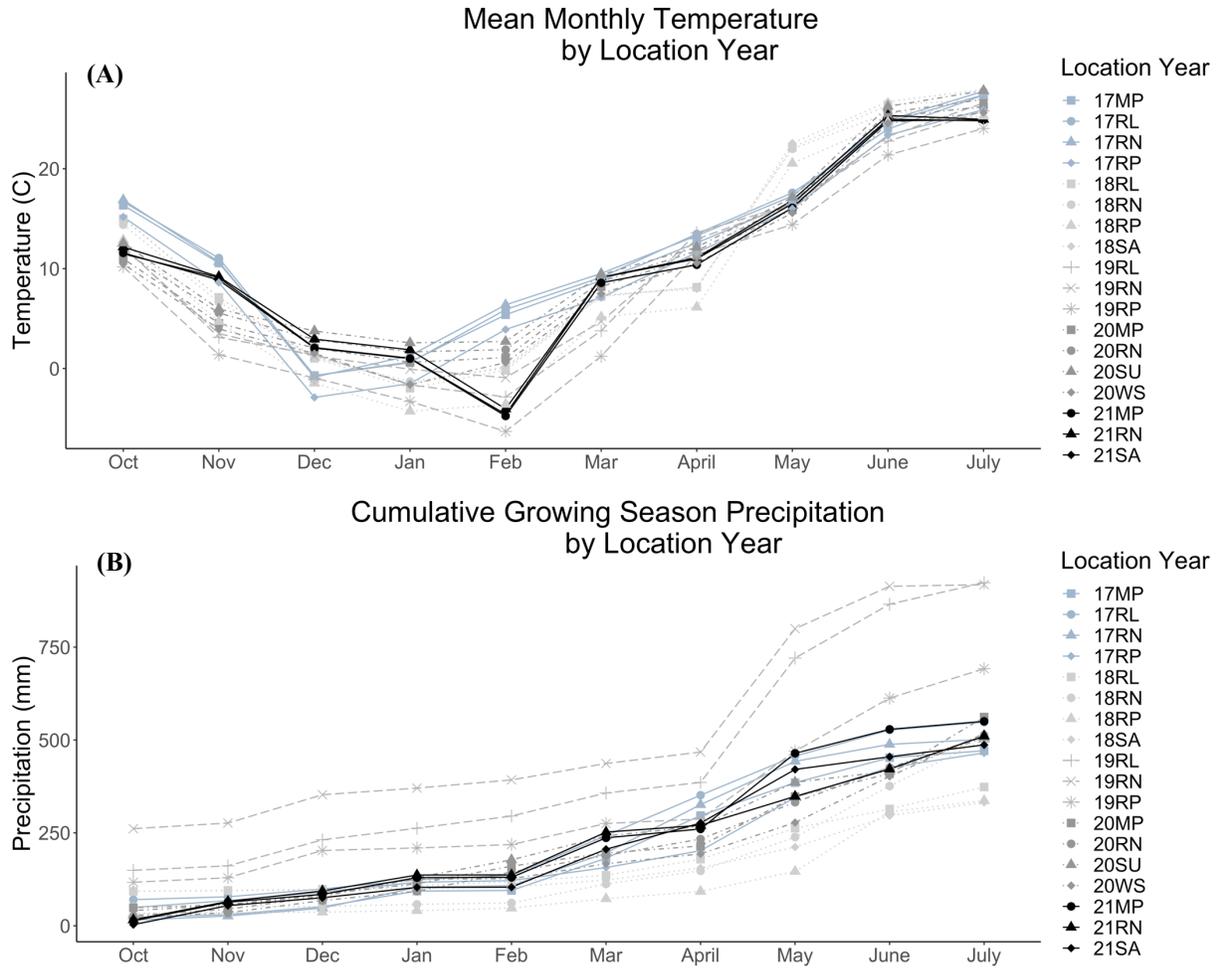
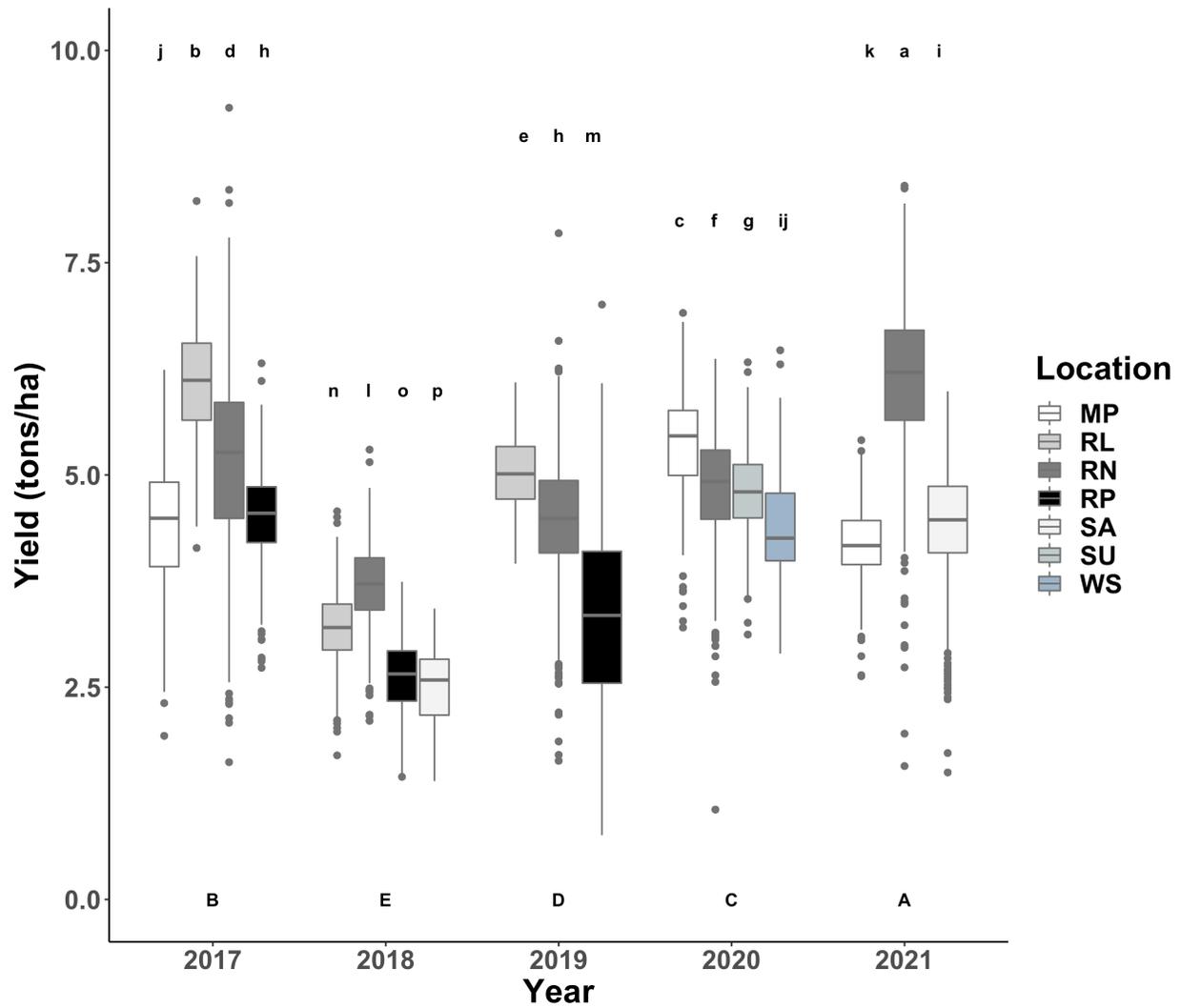
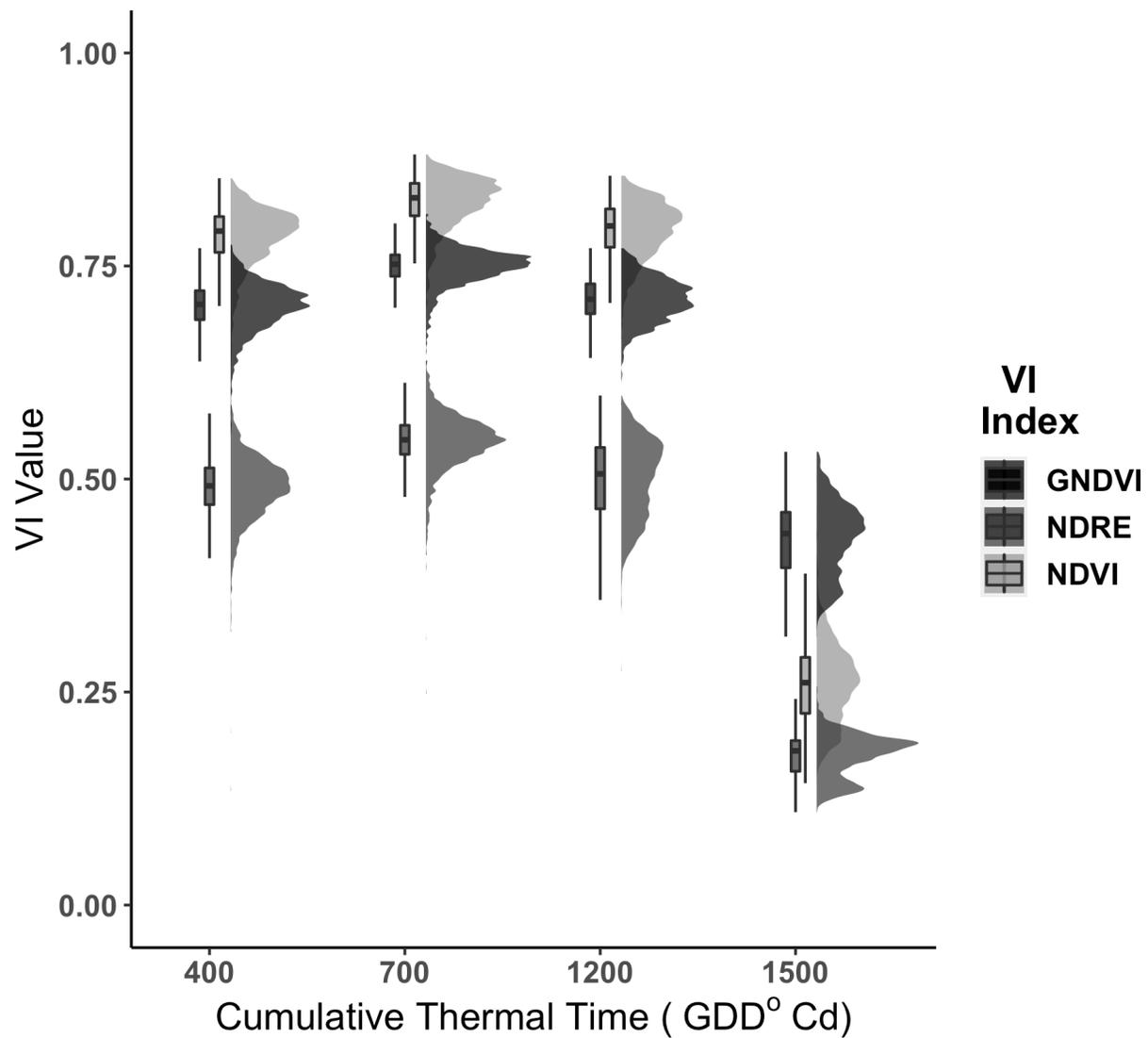


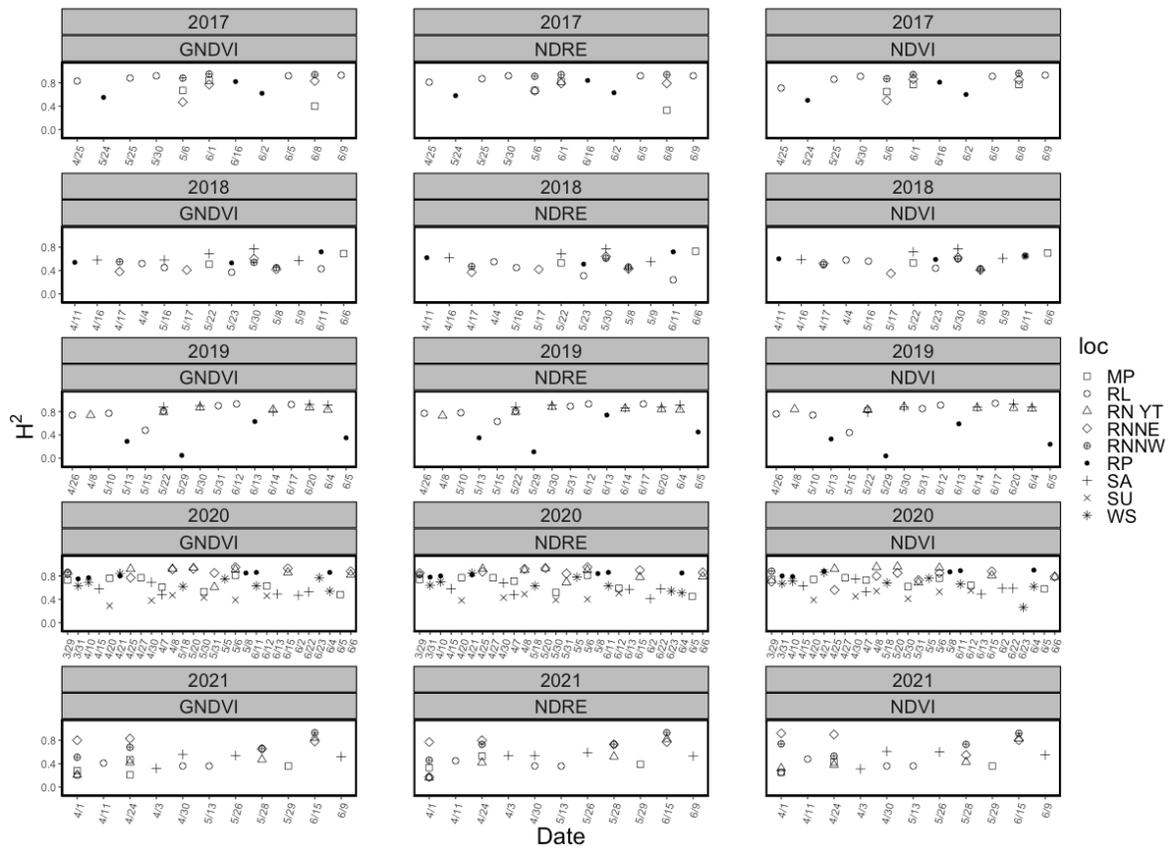
Figure 4.3. Mean monthly temperature (A) and mean monthly precipitation (B) by location year across 18 site years. The x-axis is the month during the growing season and the y-axis is the recorded mean monthly temperature (A) and mean monthly precipitation (B) from the Kansas Mesonet weather stations. Symbol and line combinations represent the location and year the data was recorded.



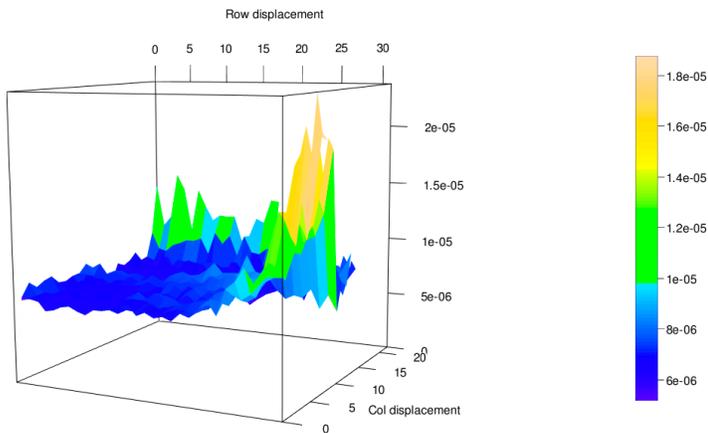
**Figure 4.4.** Grain yield by year and location for the KSWB experiments over 5 years. The x-axis is the harvest year growing season and the y-axis is adjusted grain yield in tons ha<sup>-1</sup>. Colors represent location that data was collected for during each season.



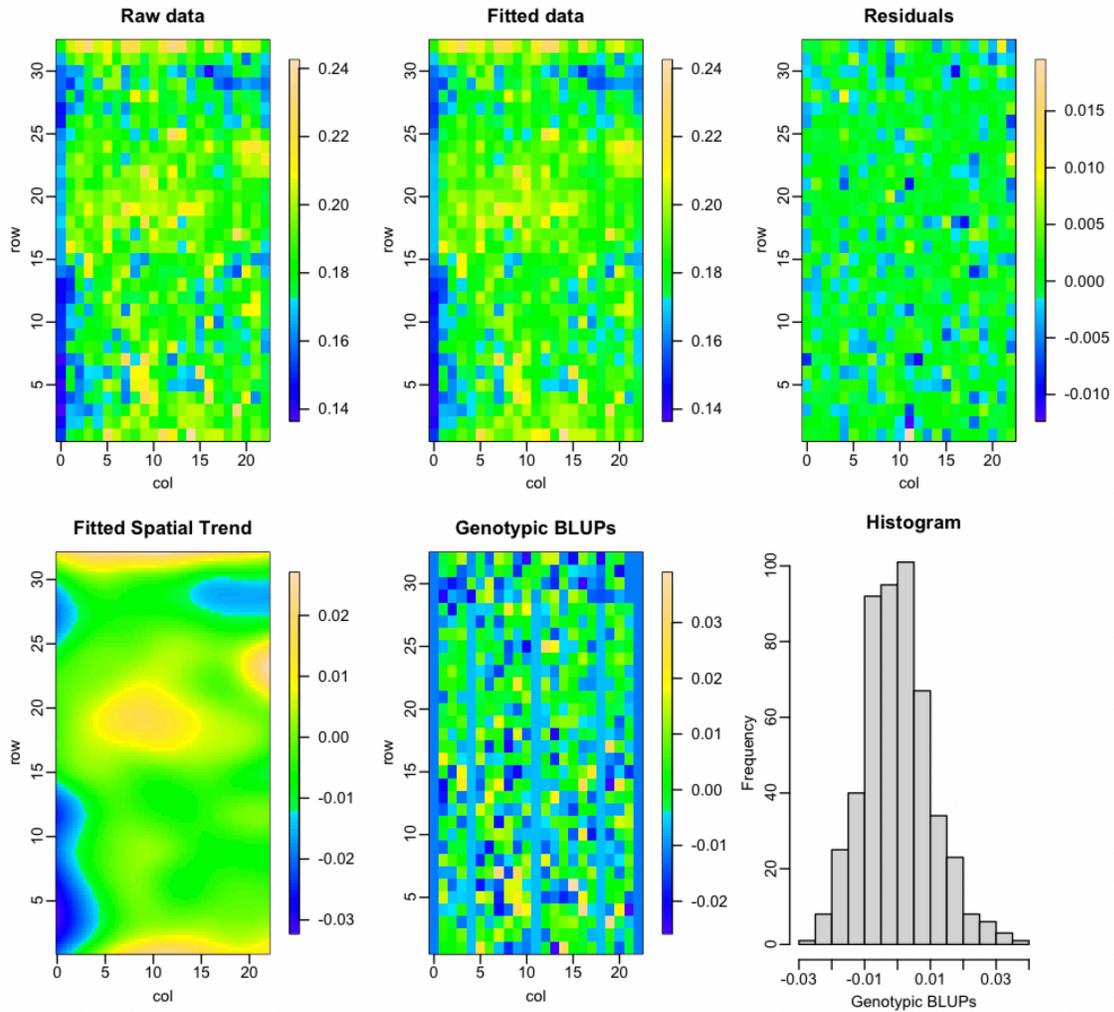
**Figure 4.5.** VI distribution by thermal time flight date for GNDVI, NDRE and NDVI at the 2021 Reno County location. The x-axis is the thermal time at the flight date, the y-axis is the calculated VI value and the colors indicate the VI.



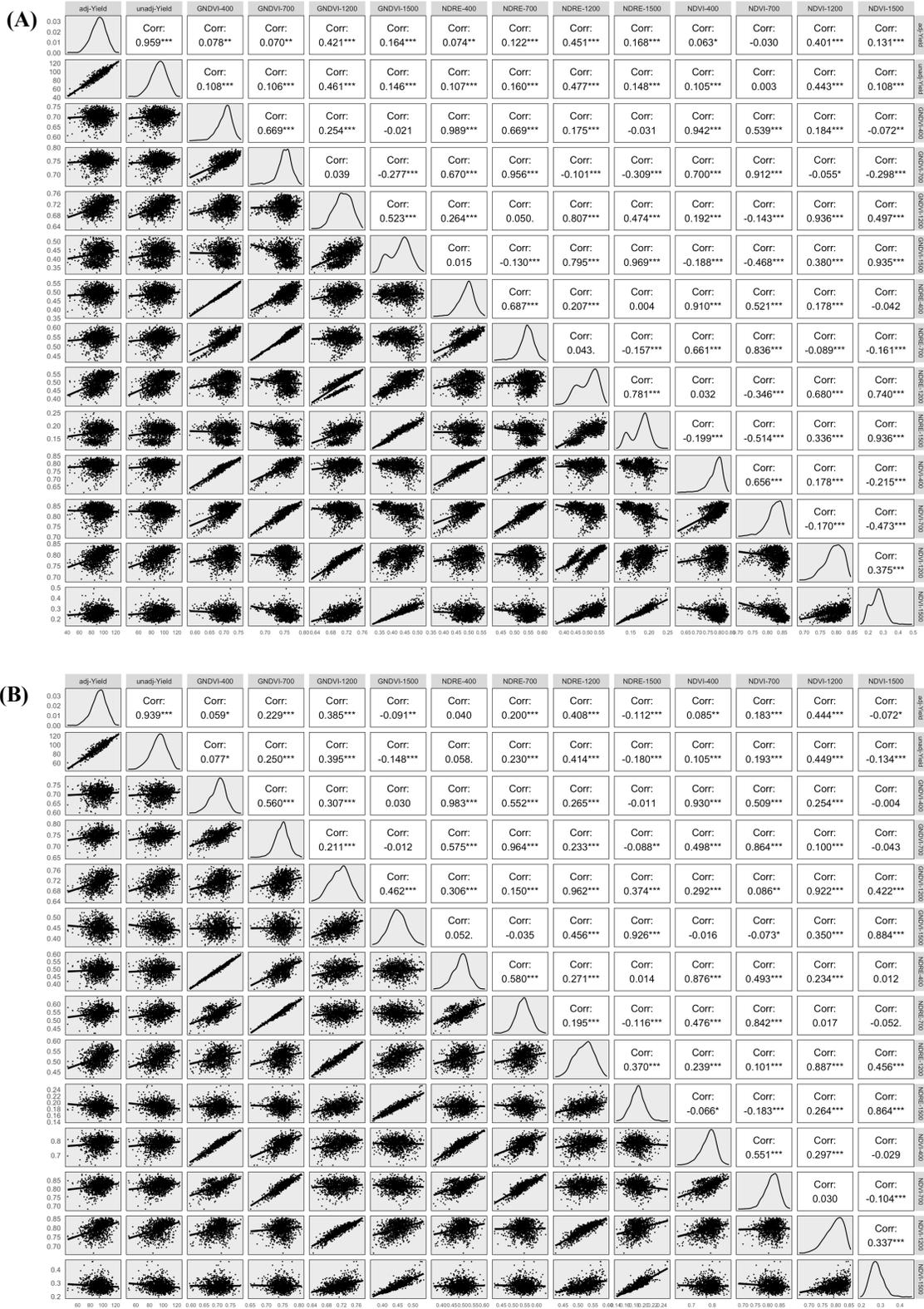
**Figure 4.6.** Broad sense heritability for GNDVI, NDRE and NDVI by flight date. Graphs include 5 years and 9 locations. However, not all locations are represented every year.



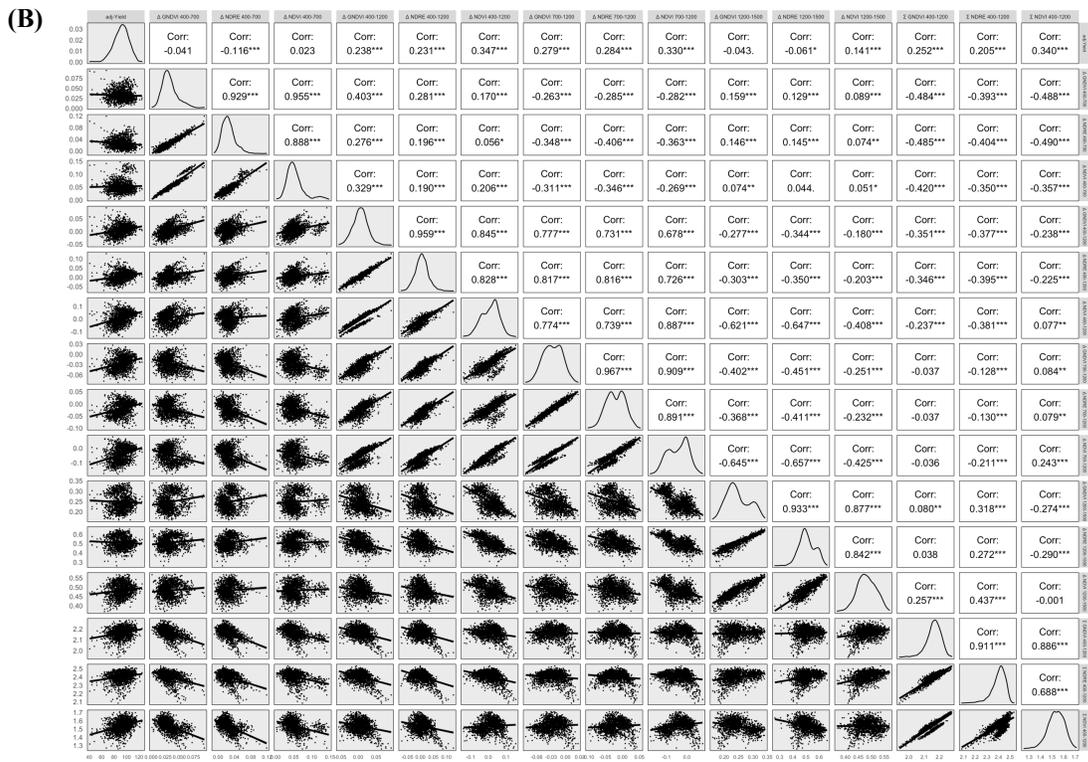
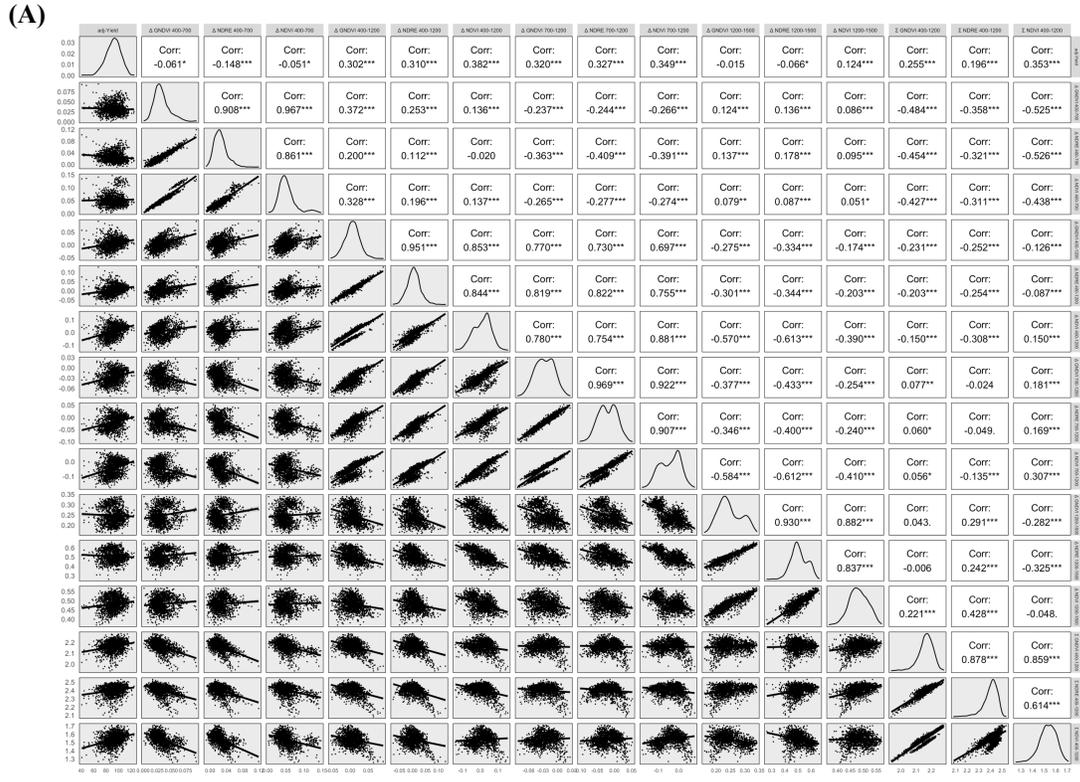
**Figure 4.7.** 3D Variogram created from 2D splines for NDRE collected 6/15/21 for the 2021 Reno Yield trial. The x-axis is the row effect the y-axis is the column displacement, and the z-axis is the spatial trend.



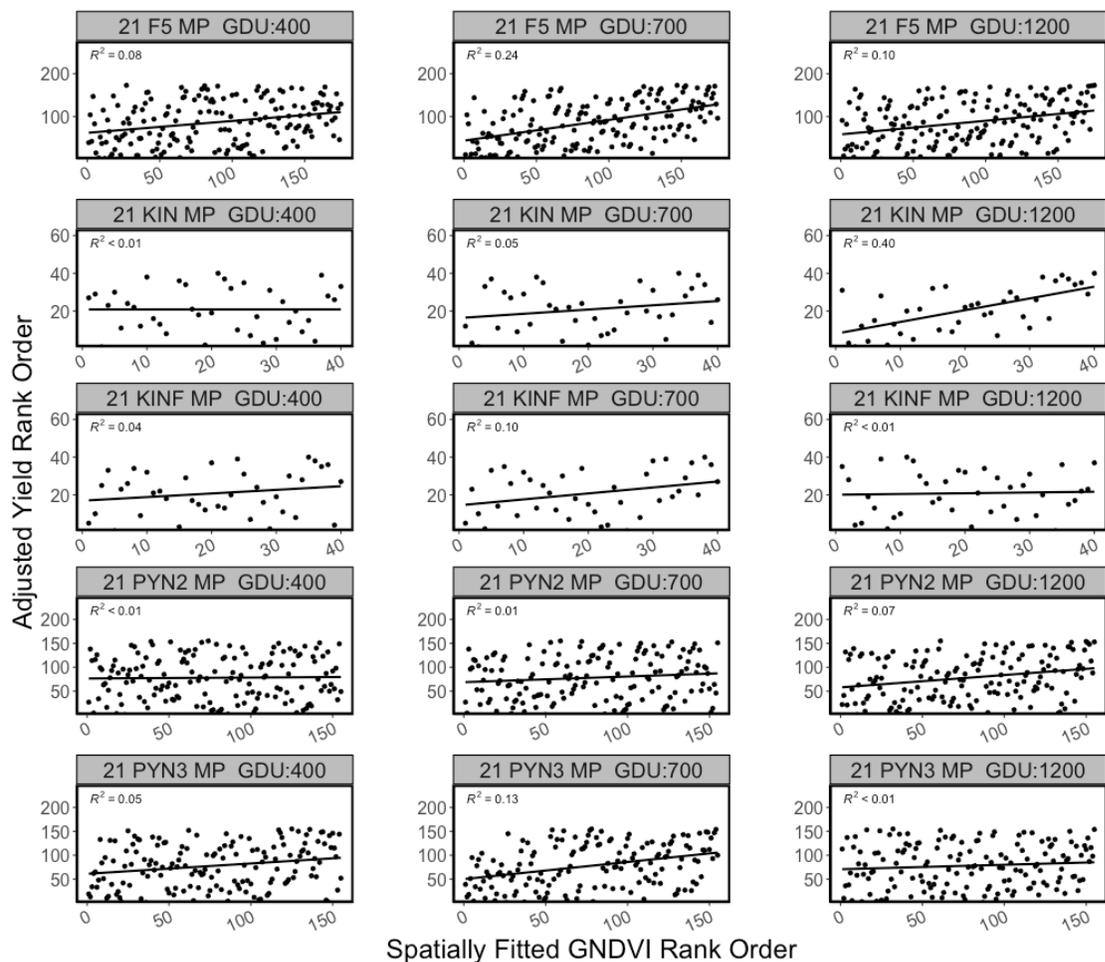
**Figure 4.8** 2D splines spatial correction workflow example. Data is the NDRE, VI collected 6/15/21 from the 2021 Reno Yield trial. The x-axis on the maps are columns and the y-axis are the rows.



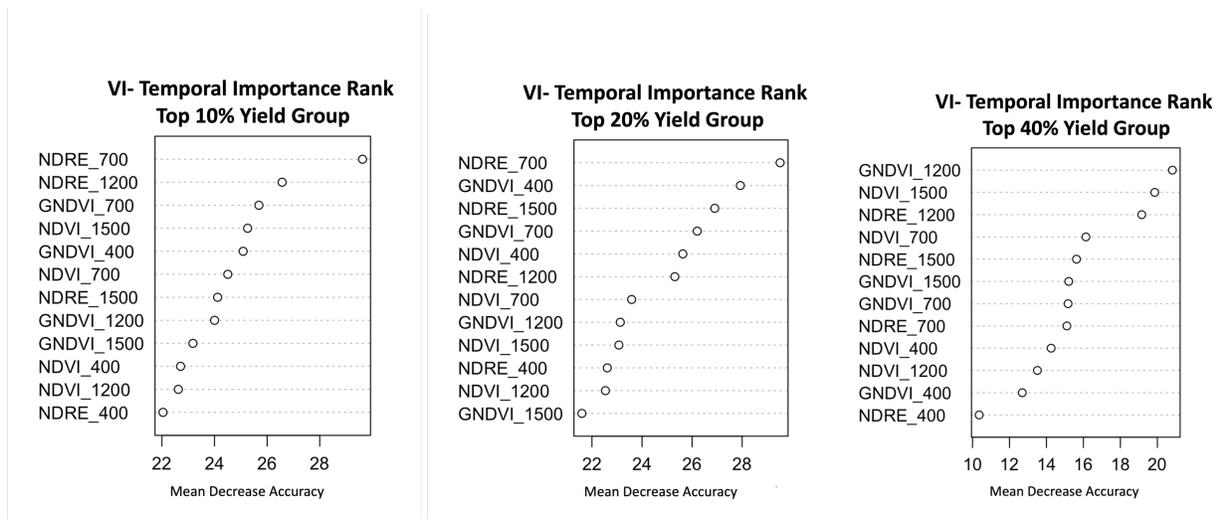
**Figure 4.9.** Pearson correlation (A) and Spearman rank order correlation (B) values for adjusted grain yield, raw grain yield, and spatially fitted GNDVI, NDRE and NDVI at 2021 Reno County.



**Figure 4.10.** Pearson correlation (A) and Spearman rank order correlation (B) values for adjusted grain yield, raw grain yield, and spatially fitted  $\Delta VI$  at 2021 Reno County.



**Figure 4.11.** Linear model performance for grain yield rank order and fitted rank order by experiment at the 2021 McPherson location. The x-axis is the rank order of the fitted GNDVI values, and the y-axis is the adjusted grain yield rank order. The gray title box indicates the year location, experiment, and thermal time flight date.



**Figure 4.12.** Rank order importance for the random forest model incorporating all VI and temporal combinations for 3 distinct yield groupings. The x-axis is the decreased mean accuracy and the y-axis the VI, thermal time combination. Higher decrease accuracy numbers indicate higher importance.

**Table 4.1-** Growing degree accumulation at flight date for all locations. Thermal time categories are designated at the top of the table. These categories were used for all analysis.

		400			700			1200			1500		
Growing season	LOC	Date	GDU	LOC	Date	GDU	LOC	Date	GDU	LOC	Date	GDU	
2021	MP	4/1/21	385	MP	4/24/21	592	MP	5/29/21	1222	MP			
2021	RN	4/1/21	427	RN	4/24/21	649	RN	5/28/21	1235	RN	6/15/21	1628	
2021	SA	4/3/21	410	SA	4/30/21	735	SA	5/26/21	1166	SA	6/9/21	1439	
2020	RN	3/29/20	434	RN	5/6/20	737	RN	5/31/20	1126	RN	6/15/20	1688	
2020	MP	4/7/20	456	MP	4/27/20	681	MP	5/30/20	1201	MP	6/12/20	1523	
2020	WS	4/10/20	442	WS	5/5/20	736	WS	6/4/20	1230	WS	6/23/20	1674	
2020	SU	4/8/20	579	SU	4/30/20	854	SU	5/30/20	1364	SU	6/12/20	1691	
2019	RL	---	---	RL	4/26/19	610	RL	5/31/19	1171	RL	6/17/19	1538	
2019	RN	4/8/19	427	RN	---	---	RN	5/30/19	1238	RN	6/14/19	1569	
2019	RP	---	---	RP	5/13/19	659	RP	6/13/19	1228	RP	---	---	
2018	RP	4/11/18	323	RP	5/23/18	894	RP	6/11/18	1355	RP	---	---	
2018	RN	4/17/18	486	RN	5/8/18	792	RN	5/30/18	1179	RN	---	---	
2018	SA	4/16/18	480	SA	5/9/18	819	SA	5/22/18	1125	SA	---	---	
2017	RP	---	---	RP	---	---	RP	5/24/17	1137	RP	6/16/17	1642	
2017	MP	---	---	MP	---	---	MP	5/6/17	1185	MP	6/8/17	1607	
2017	RL	---	---	RL	---	---	RL	5/25/17	1259	RL	6/5/17	1609	
2017	RN	---	---	RN	---	---	RN	5/6/17	1272	RN	6/1/17	1545	

**Table 4.2-** Equations for vegetation indices used to evaluate rank order predictions.

VI	Index	Equation
GNDVI	Green Normalized Difference Vegetation Index	$GNDVI = \frac{NIR - Green}{NIR + Green}$
NDRE	Normalized Difference Red Edge	$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$
NDVI	Normalized Difference Vegetation Index	$NDVI = \frac{NIR - Red}{NIR + Red}$
$\Delta GNDVI_{400-700}$	Change in Green Normalized Difference Vegetation Index 400 to 700 °Cd	$\Delta GNDVI_{400-700} = \frac{GNDVI_{700} - GNDVI_{400}}{GNDVI_{700} + GNDVI_{400}}$
$\Delta NDRE_{400-700}$	Change in Normalized Difference Red Edge 400 to 700 °Cd	$\Delta NDRE_{400-700} = \frac{NDRE_{700} - NDRE_{400}}{NDRE_{700} + NDRE_{400}}$
$\Delta NDVI_{400-700}$	Change in Normalized Difference Vegetation Index 400 to 700 °Cd	$\Delta NDVI_{400-700} = \frac{NDVI_{700} - NDVI_{400}}{NDVI_{700} + NDVI_{400}}$
$\Sigma GNDVI_{4-7-12}$	Accumulation of Green Normalized Difference Vegetation Index 400 , 700 and 1200 °Cd	$\Sigma GNDVI_{4-7-12} = GNDVI_{400} + GNDVI_{700} + GNDVI_{1200}$
$\Sigma NDRE_{4-7-12}$	Accumulation of Normalized Difference Red Edge 400 , 700 and 1200 °Cd	$\Sigma NDRE_{4-7-12} = NDRE_{400} + NDRE_{700} + NDRE_{1200}$
$\Sigma NDVI_{4-7-12}$	Accumulation of Normalized Difference Vegetation Index 400 , 700 and 1200 °Cd	$\Sigma NDVI_{4-7-12} = NDVI_{400} + NDVI_{700} + NDVI_{1200}$

**Table 4.3-** Pearson’s correlation of spatially fitted VI temporal points to adjusted grain yield by location.

Spatially Fitted VI to GRYLD- Pearson Correlation													
Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.266	0.305	0.392	---	0.268	0.293	0.393	---	0.252	0.339	0.383	---
RN	2021	0.078	0.070	0.421	0.164	0.074	0.122	0.451	0.168	0.063	-0.030	0.401	0.131
SA	2021	0.161	0.200	0.202	0.117	0.149	0.172	0.179	0.103	0.165	0.219	0.221	0.140
MP	2020	0.276	0.292	0.754	0.554	0.312	0.210	0.764	0.410	0.229	0.225	0.743	0.550
RN	2020	0.115	0.483	0.367	0.040	0.076	0.445	0.388	0.170	0.117	0.341	0.439	0.088
SU	2020	0.297	0.353	0.308	-0.065	0.313	0.465	0.318	-0.182	0.283	0.345	0.351	-0.093
WS	2020	0.292	0.388	-0.008	-0.080	0.320	0.416	-0.045	-0.083	0.304	0.332	0.022	-0.022
RL	2019	---	0.659	0.521	0.520	---	0.688	0.569	0.463	---	---	0.666	0.570
RN	2019	0.198	---	0.525	0.333	0.142	---	0.520	0.309	0.245	---	0.527	0.380
RP	2019	---	0.376	0.488	---	---	0.360	0.532	---	---	0.370	0.499	---
RN	2018	0.301	0.318	0.513	---	0.326	0.323	0.521	---	0.368	0.364	0.524	---
RP	2018	0.305	0.504	0.431	---	0.335	0.536	0.388	---	0.347	0.490	0.429	---
SA	2018	0.450	0.412	0.617	---	0.521	0.456	0.681	---	0.379	0.252	0.567	---
MP	2017	---	---	-0.065	-0.411	---	---	-0.047	-0.352	---	---	-0.169	-0.032
RL	2017	---	---	0.004	-0.084	---	---	0.015	-0.149	---	---	-0.040	-0.137
RN	2017	---	---	0.202	0.233	---	---	0.307	0.220	---	---	0.167	0.243
RP	2017	---	---	0.106	0.016	---	---	0.185	0.024	---	---	0.131	0.269

**Table 4.4-** Spearman’s correlation of spatially fitted VI temporal points to adjusted grain yield by location.

Spatially Fitted VI to GRYLD- Spearman Correlation													
Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.300	0.340	0.335	---	0.296	0.326	0.341	---	0.286	0.377	0.335	---
RN	2021	0.059	0.229	0.385	-0.091	0.040	0.200	0.408	-0.112	0.085	0.183	0.444	-0.072
SA	2021	0.145	0.187	0.186	0.105	0.127	0.173	0.174	0.093	0.146	0.190	0.185	0.114
MP	2020	0.263	0.311	0.706	0.477	0.285	0.234	0.725	0.308	0.224	0.202	0.655	0.477
RN	2020	0.321	0.494	0.725	0.483	0.353	0.511	0.720	0.497	0.160	0.588	0.638	0.432
SU	2020	0.281	0.331	0.296	-0.074	0.299	0.443	0.333	-0.167	0.256	0.322	0.327	-0.049
WS	2020	0.235	0.245	-0.077	0.066	0.251	0.281	-0.108	0.013	0.241	0.178	-0.065	0.141
RL	2019	---	0.714	0.585	0.580	---	---	0.615	0.541	---	---	0.706	0.611
RN	2019	0.164	---	0.522	0.325	0.141	---	0.509	0.310	0.227	---	0.537	0.383
RP	2019	---	0.376	0.478	---	---	0.354	0.524	---	---	0.363	0.496	---
RN	2018	0.296	0.275	0.459	---	0.320	0.291	0.483	---	0.364	0.297	0.483	---
RP	2018	0.301	0.525	0.437	---	0.338	0.565	0.400	---	0.330	0.501	0.436	---
SA	2018	0.439	0.383	0.621	---	0.523	0.445	0.691	---	0.397	0.231	0.549	---
MP	2017	---	---	-0.215	-0.416	---	---	-0.191	-0.370	---	---	-0.308	-0.094
RL	2017	---	---	-0.037	-0.124	---	---	-0.049	-0.173	---	---	-0.144	-0.161
RN	2017	---	---	0.208	0.199	---	---	0.323	0.187	---	---	0.145	0.233
RP	2017	---	---	0.153	0.050	---	---	0.231	0.049	---	---	0.166	0.272

**Table 4.5-** P-values for linear model VI rank order predictions to observed grain yield rank order

Exp.	Loc.	GNDVI- 400	GNDVI- 700	GNDVI- 1200	NDRE- 400	NDRE- 700	NDRE- 1200	NDVI- 400	NDVI- 700	NDVI- 1200
F5	MP	0.08	0.24	0.10	0.09	0.26	0.14	0.06	0.20	0.12
KIN	MP	<0.01	0.05	0.40	<0.01	<0.01	0.41	<0.01	0.06	0.40
KINF	MP	0.04	0.10	<0.01	0.05	0.09	<0.01	<0.01	0.11	<0.01
PYN2	MP	<0.01	<0.01	0.07	<0.01	<0.01	0.08	<0.01	<0.01	0.06
PYN3	MP	0.05	0.13	<0.01	0.05	0.12	<0.01	<0.01	0.17	0.03
AYN1	RN	<0.01	0.07	<0.01	<0.01	0.08	0.02	<0.01	0.04	0.03
AYN2	RN	0.02	0.10	<0.01	<0.01	0.07	<0.01	<0.01	0.09	0.04
F4	RN	<0.01	<0.01	0.15	<0.01	<0.01	0.15	<0.01	<0.01	0.12
F5	RN	<0.01	<0.01	0.14	<0.01	<0.01	0.12	<0.01	<0.01	0.17
KINF	RN	<0.01	0.02	0.09	<0.01	<0.01	0.04	0.07	<0.01	0.04
PYN2	RN	<0.01	0.05	0.17	<0.01	0.03	0.17	<0.01	0.03	0.22
PYN3	RN	<0.01	0.02	0.10	<0.01	<0.01	0.06	<0.01	0.02	0.15
PYN4	RN	0.06	0.17	0.06	0.06	0.14	0.07	0.06	0.13	0.12
PYN5	RN	0.03	0.10	0.15	0.03	0.09	0.19	<0.01	0.06	0.18
PYN6	RN	0.21	0.23	0.13	0.20	0.22	0.10	0.16	0.19	0.14
PYN7	RN	0.07	0.14	0.02	0.06	0.10	0.02	0.05	0.16	0.02
AYN1	SA	0.15	0.12	0.06	0.18	0.08	0.09	0.16	0.08	0.14
AYN2	SA	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
PYN2	SA	0.04	0.09	0.04	0.04	0.08	0.03	0.04	0.09	0.03
PYN3	SA	<0.01	0.02	0.03	<0.01	<0.01	<0.01	<0.01	0.02	0.02
PYN4	SA	0.13	0.19	0.11	0.13	0.15	0.07	0.08	0.12	0.10
PYN5	SA	0.25	0.35	0.30	0.24	0.34	0.30	0.20	0.30	0.28
PYN6	SA	0.13	0.17	0.16	0.11	0.19	0.16	0.10	0.13	0.14
PYN7	SA	0.04	0.06	0.07	0.03	0.05	0.07	0.03	0.07	0.07

**Table 4.6-** Table of percentages of the total population needed to capture the top 10% yield group using single temporal VI's as a selection criteria.

Exp.	Loc.	GNDVI-400	GNDVI-700	GNDVI-1200	NDRE-400	NDRE-700	NDRE-1200	NDVI-400	NDVI-700	NDVI-1200
F5	MP	63.7%	68.4%	77.4%	65.8%	63.2%	73.7%	63.2%	83.2%	85.8%
KIN	MP	72.0%	40.0%	16.0%	72.0%	56.0%	16.0%	68.0%	30.0%	18.0%
KINF	MP	78.0%	52.0%	68.0%	78.0%	58.0%	60.0%	78.0%	58.0%	74.0%
PYN2	MP	83.8%	95.0%	85.6%	91.3%	96.3%	86.3%	83.8%	85.6%	92.5%
PYN3	MP	91.9%	68.8%	65.6%	92.5%	66.3%	65.6%	90.0%	70.6%	70.0%
AYN1	RN	47.5%	62.5%	62.5%	52.5%	65.0%	57.5%	50.0%	85.0%	87.5%
AYN2	RN	90.0%	55.0%	87.5%	92.5%	60.0%	87.5%	100.0%	52.5%	72.5%
F4	RN	95.8%	97.5%	98.0%	95.8%	97.5%	98.0%	95.5%	97.5%	98.0%
F5	RN	78.9%	91.1%	94.4%	77.2%	91.7%	93.9%	69.4%	88.3%	94.4%
KINF	RN	67.5%	77.5%	77.5%	65.0%	77.5%	80.0%	72.5%	95.0%	75.0%
PYN2	RN	79.4%	94.4%	91.9%	70.0%	94.4%	90.6%	74.4%	90.0%	88.8%
PYN3	RN	86.9%	95.6%	92.5%	91.9%	95.6%	95.0%	91.9%	95.0%	92.5%
PYN4	RN	88.8%	72.5%	70.0%	87.5%	70.0%	70.0%	80.0%	58.8%	61.3%
PYN5	RN	72.5%	69.2%	59.2%	75.0%	77.5%	68.3%	61.7%	76.7%	80.8%
PYN6	RN	77.6%	67.1%	91.2%	75.3%	74.1%	88.8%	73.5%	71.8%	83.5%
PYN7	RN	93.5%	94.7%	90.0%	93.5%	96.5%	91.2%	93.5%	86.5%	81.8%
AYN1	SA	90.0%	77.5%	42.5%	87.5%	87.5%	70.0%	87.5%	65.0%	47.5%
AYN2	SA	62.5%	95.0%	65.0%	55.0%	92.5%	65.0%	60.0%	100.0%	75.0%
PYN2	SA	88.1%	86.9%	76.3%	90.0%	87.5%	87.5%	93.1%	81.3%	65.6%
PYN3	SA	90.6%	94.4%	93.1%	91.9%	94.4%	91.3%	91.9%	93.8%	91.3%
PYN4	SA	80.0%	71.3%	73.8%	56.3%	82.5%	76.3%	87.5%	81.3%	71.3%
PYN5	SA	64.2%	53.3%	42.5%	40.8%	53.3%	45.0%	64.2%	71.7%	53.3%
PYN6	SA	60.6%	91.2%	94.1%	65.9%	81.2%	90.0%	79.4%	95.3%	90.6%
PYN7	SA	80.0%	74.7%	70.6%	94.7%	65.9%	81.8%	90.6%	92.4%	55.9%

**Table 4.7-** Random forest model performance metrics. All models used a 60% training population and a 40% testing population. Additionally, all models use the GNDVI, NDRE, and NDVI indices. The thermal time column indicates the temporal thermal times used to create and test the model.

Thermal Times	Yield Group	Prediction Accuracy	95 % CI	No Info. Rate	P-Value [Acc > NIR]	Kappa Coefficient (K)	Sensitivity	Specificity
400-700-1200-1500	Top 10%	0.899	(0.8823, 0.9148)	0.898	0.451	0.025	1	0.014
400-700-1200-1500	Top 20%	0.797	(0.7751, 0.8184)	0.799	0.569	0.069	0.982	0.065
400-700-1200-1500	Top 40%	0.604	(0.5778, 0.6302)	0.598	0.34	0.144	0.747	0.392
400-700-1200	Top 10%	0.896	(0.8822, 0.9086)	0.898	0.657	0.003	0.997	0.005
400-700-1200	Top 20%	0.791	(0.7735, 0.8085)	0.802	0.899	0.046	0.972	0.06
400-700-1200	Top 40%	0.629	(0.6081, 0.6496)	0.608	0.026	0.194	0.757	0.43
400-700	Top 10%	0.896	(0.8822, 0.9086)	0.898	0.657	0.003	0.997	0.005
400-700	Top 20%	0.791	(0.7735, 0.8085)	0.802	0.899	0.046	0.972	0.06
400-700	Top 40%	0.629	(0.6081, 0.6496)	0.608	0.026	0.194	0.757	0.43
400-1200	Top 10%	0.894	(0.8818, 0.9052)	0.898	0.786	0.013	0.994	0.014
400-1200	Top 20%	0.789	(0.7731, 0.804)	0.802	0.954	0.061	0.965	0.079
400-1200	Top 40%	0.599	(0.5808, 0.6177)	0.606	0.765	0.132	0.732	0.395
700-1200	Top 10%	0.894	(0.8805, 0.9069)	0.898	0.753	-0.008	0.995	0
700-1200	Top 20%	0.794	(0.7667, 0.8187)	0.798	0.643	0.044	0.983	0.046
700-1200	Top 40%	0.602	(0.581, 0.6228)	0.608	0.725	0.122	0.762	0.354
1200-1500	Top 10%	0.89	(0.8772, 0.9014)	0.887	0.325	0.088	0.996	0.056
1200-1500	Top 20%	0.775	(0.7583, 0.7904)	0.789	0.962	0.057	0.961	0.08
1200-1500	Top 40%	0.611	(0.5916, 0.6291)	0.59	0.017	0.16	0.765	0.388

## References

- Barker III, J., Zhang, N., Sharon, J., Steeves, R., Wang, X., Wei, Y., & Poland, J. (2016). Development of a field-based high-throughput mobile phenotyping platform. *Computers and Electronics in Agriculture*, *122*, 74-85.
- Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R package version*, *2*(1), 74.
- Battenfield, S. D., Guzmán, C., Gaynor, R. C., Singh, R. P., Peña, R. J., Dreisigacker, S., Fritz, A. K., & Poland, J. A. (2016). Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. *The plant genome*, *9*(2), plantgenome2016.2001.0005.
- Beauchêne, K., Leroy, F., Fournier, A., Huet, C., Bonnefoy, M., Lorgeou, J., De Solan, B., Piquemal, B., Thomas, S., & Cohan, J.-P. (2019). Management and characterization of abiotic stress via PhénoField®, a high-throughput field phenotyping platform. *Frontiers in Plant Science*, *10*, 904.
- Bentley, A. R., & Mackay, I. (2017). Advances in wheat breeding techniques. In *Achieving sustainable cultivation of wheat Volume 1* (pp. 109-111). Burleigh Dodds Science Publishing.
- Bonfil, D. J. (2017). Wheat phenomics in the field by RapidScan: NDVI vs. NDRE. *Israel Journal of Plant Sciences*, *64*(3-4), 41-54.  
<https://doi.org/10.1080/07929978.2016.1249135>
- Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H. P., Reif, J. C., Würschum, T., & Müller, J. (2013). BreedVision—A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors*, *13*(3), 2830-2847.
- Christopher, J., Manschadi, A., Hammer, G., & Borrell, A. (2008). Developmental and physiological traits associated with high yield and stay-green phenotype in wheat. *Australian Journal of Agricultural Research*, *59*(4), 354-364.
- Cooper, M., Stucker, R., DeLacy, I., & Harch, B. (1997). Wheat breeding nurseries, target environments, and indirect selection for grain yield. *Crop Science*, *37*(4), 1168-1176.
- Cox, T., Shroyer, J., Ben-Hui, L., Sears, R., & Martin, T. (1988). Genetic improvement in agronomic traits of hard red winter wheat cultivars 1919 to 1987. *Crop Science*, *28*(5), 756-760.
- Crain, J., Wang, X., Lucas, M., & Poland, J. (2021). Experiences of Applying Field-Based High-Throughput Phenotyping for Wheat Breeding. In *High-Throughput Crop Phenotyping* (pp. 71-99). Springer.

- Eitel, J. U. H., Long, D. S., Gessler, P. E., & Hunt, E. R. (2008). Combined Spectral Index to Improve Ground-Based Estimates of Nitrogen Status in Dryland Wheat. *Agronomy Journal*, 100(6), 1694-1702. <https://doi.org/10.2134/agronj2007.0362>
- Fernandez-Gallego, J. A., Lootens, P., Borra-Serrano, I., Derycke, V., Haesaert, G., Roldán-Ruiz, I., Araus, J. L., & Kefauver, S. C. (2020). Automatic wheat ear counting using machine learning based on RGB UAV imagery. *The Plant Journal*, 103(4), 1603-1613.
- Fu, Z. P., Jiang, J., Gao, Y., Krienke, B., Wang, M., Zhong, K. T., Cao, Q., Tian, Y. C., Zhu, Y., Cao, W. X., & Liu, X. J. (2020). Wheat Growth Monitoring and Yield Estimation based on Multi-Rotor Unmanned Aerial Vehicle. *Remote Sensing*, 12(3). <https://doi.org/ARTN508>  
10.3390/rs12030508
- García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O. R., & Vázquez-Peña, M. A. (2020). Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and RGB images acquired with unmanned aerial vehicles. *Agriculture*, 10(7), 277.
- Gordillo-Salinas, V. M., Flores-Magdaleno, H., Ortiz-Solorio, C. A., & Arteaga-Ramirez, R. (2021). Evaluation of nitrogen status in a wheat crop using unmanned aerial vehicle images. *Chilean Journal of Agricultural Research*, 81(3), 408-419. <https://doi.org/10.4067/S0718-58392021000300408>
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote sensing of environment*, 90(3), 337-352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Haghighattalab, A., Perez, 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. <https://doi.org/ARTN35>  
10.1186/s13007-016-0134-6
- Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., Xiao, Y., & He, Z. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, 282, 95-103.
- Kim, J. Y. (2020). Roadmap to high throughput phenotyping for plant breeding. *Journal of Biosystems Engineering*, 45(1), 43-55.
- Kyratzis, A. C., Skarlatos, D. P., Menexes, G. C., Vamvakousis, V. F., & Katsiotis, A. (2017). Assessment of Vegetation Indices Derived by UAV Imagery for Durum Wheat Phenotyping under a Water Limited and Heat Stressed Mediterranean Environment. *Frontiers in Plant Science*, 8. <https://doi.org/ARTN1114>  
10.3389/fpls.2017.01114

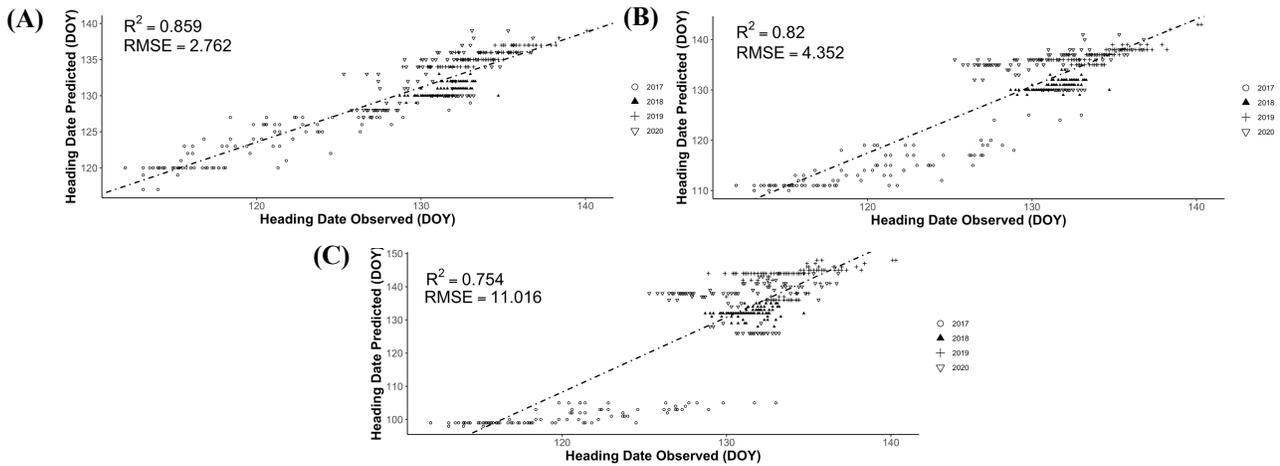
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.
- Lin, C.-S., & Poushinsky, G. (1985). A modified augmented design (type 2) for rectangular plots. *Canadian Journal of Plant Science*, 65(3), 743-749.
- Miller, M. P., Singer, M. J., & Nielsen, D. R. (1988). Spatial variability of wheat yield and soil properties on complex hills. *Soil Science Society of America Journal*, 52(4), 1133-1141.
- Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. *Journal of Atmospheric and Oceanic Technology*, 37(12), 2167-2183.  
<https://doi.org/10.1175/Jtech-D-19-0214.1>
- Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S. Y., Manes, Y., Dreisigacker, S., Crossa, J., Sanchez-Villeda, H., Sorrells, M., & Jannink, J. L. (2012). Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome*, 5(3), 103-113.  
<https://doi.org/10.3835/plantgenome2012.06.0006>
- Potgieter, A. B., George-Jaeggli, B., Chapman, S. C., Laws, K., Suárez Cadavid, L. A., Wixted, J., Watson, J., Eldridge, M., Jordan, D. R., & Hammer, G. L. (2017). Multi-spectral imaging from an unmanned aerial vehicle enables the assessment of seasonal leaf area dynamics of sorghum breeding lines. *Frontiers in Plant Science*, 8, 1532.
- Prasad, P. V., & Djanaguiraman, M. (2014). Response of floret fertility and individual grain weight of wheat to high temperature stress: sensitive stages and thresholds for temperature and duration. *Functional Plant Biology*, 41(12), 1261-1269.
- Ramos, A. P. M., Osco, L. P., Furuya, D. E. G., Goncalves, W. N., Santana, D. C., Teodoro, L. P. R., da Silva, C. A., Capristo-Silva, G. F., Li, J., Baio, F. H. R., Marcato, J., Teodoro, P. E., & Pistori, H. (2020). A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Computers and Electronics in Agriculture*, 178.  
<https://doi.org/ARTN 105791>  
10.1016/j.compag.2020.105791
- Rodriguez-Alvarez, M. X., Boer, M. P., van Eeuwijk, F. A., & Eilers, P. H. C. (2018). Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spatial Statistics*, 23, 52-71. <https://doi.org/10.1016/j.spasta.2017.10.003>
- Roozeboom, K. L., Schapaugh, W. T., Tuinstra, M. R., Vanderlip, R. L., & Milliken, G. A. (2008). Testing wheat in variable environments: genotype, environment, interaction effects, and grouping test locations. *Crop Science*, 48(1), 317-330.
- Schloerke, B., Crowley, J., & Cook, D. (2018). Package ‘GGally’. In: Extension to ‘ggplot2.’ See <https://cran.r-project.org/web/packages/GGally> ....
- Sehgal, D., Rosyara, U., Mondal, S., Singh, R., Poland, J., & Dreisigacker, S. (2020). Incorporating genome-wide association mapping results into genomic prediction models

for grain yield and yield stability in CIMMYT spring bread wheat. *Frontiers in Plant Science*, 11, 197.

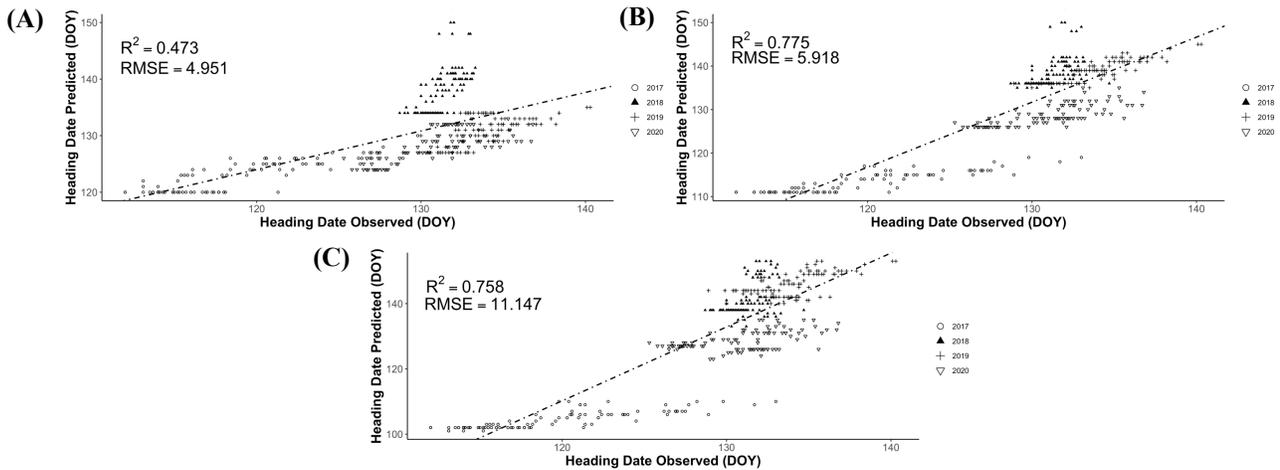
- Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., & Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Computers and Electronics in Agriculture*, 183. <https://doi.org/ARTN106036>  
10.1016/j.compag.2021.106036
- Singh, D., Wang, X., Kumar, U., Gao, L., Noor, M., Imtiaz, M., Singh, R. P., & Poland, J. (2019). High-Throughput Phenotyping Enabled Genetic Dissection of Crop Lodging in Wheat [Original Research]. *Frontiers in Plant Science*, 10(394).  
<https://doi.org/10.3389/fpls.2019.00394>
- Song, P., Wang, J. L., Guo, X. Y., Yang, W. N., & Zhao, C. J. (2021). High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop Journal*, 9(3), 633-645. <https://doi.org/10.1016/j.cj.2021.03.015>
- Su, J., Liu, C., Coombes, M., Hu, X., Wang, C., Xu, X., Li, Q., Guo, L., & Chen, W.-H. (2018). Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Computers and Electronics in Agriculture*, 155, 157-166.
- Velu, G., & Singh, R. P. (2013). Phenotyping in wheat breeding. In *Phenotyping for Plant Breeding* (pp. 41-71). Springer.
- Walkowiak, S., Gao, L. L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., Ramirez-Gonzalez, R. H., Kolodziej, M. C., Delorean, E., Thambugala, D., Klymiuk, V., Byrns, B., Gundlach, H., Bandi, V., Siri, J. N., Nilsen, K., Aquino, C., Himmelbach, A., Copetti, D., Ban, T., Venturini, L., Bevan, M., Clavijo, B., Koo, D. H., Ens, J., Wiebe, K., N'Diaye, A., Fritz, A. K., Gutwin, C., Fiebig, A., Fosker, C., Fu, B. X., Accinelli, G. G., Gardner, K. A., Fradgley, N., Gutierrez-Gonzalez, J., Halstead-Nussloch, G., Hatakeyama, M., Koh, C. S., Deek, J., Costamagna, A. C., Fobert, P., Heavens, D., Kanamori, H., Kawaura, K., Kobayashi, F., Krasileva, K., Kuo, T., McKenzie, N., Murata, K., Nabeka, Y., Paape, T., Padmarasu, S., Percival-Alwyn, L., Kagale, S., Scholz, U., Sese, J., Juliana, P., Singh, R., Shimizu-Inatsugi, R., Swarbreck, D., Cockram, J., Budak, H., Tameshige, T., Tanaka, T., Tsuji, H., Wright, J., Wu, J. Z., Steuernagel, B., Small, I., Cloutier, S., Keeble-Gagnere, G., Muehlbauer, G., Tibbets, J., Nasuda, S., Melonek, J., Hucl, P. J., Sharpe, A. G., Clark, M., Legg, E., Bharti, A., Langridge, P., Hall, A., Uauy, C., Mascher, M., Krattinger, S. G., Handa, H., Shimizu, K. K., Distelfeld, A., Chalmers, K., Keller, B., Mayer, K. F. X., Poland, J., Stein, N., McCartney, C. A., Spannagl, M., Wicker, T., & Pozniak, C. J. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, 588(7837). <https://doi.org/10.1038/s41586-020-2961-x>
- Wang, L., Tian, Y., Yao, X., Zhu, Y., & Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *Field Crops Research*, 164, 178-188.

- Wang, X., R. Thorp, K., W. White, J., N. French, A., & A. Poland, J. (2016). Approaches for Geospatial Processing of Field-Based High-Throughput Plant Phenomics Data from Ground Vehicle Platforms. *Transactions of the ASABE*, 59(5), 1053-1067. <https://doi.org/https://doi.org/10.13031/trans.59.11502>
- Wang, X., Silva, P., Bello, N. M., Singh, D., Evers, B., Mondal, S., Espinosa, F. P., Singh, R. P., & Poland, J. (2020). Improved Accuracy of High-Throughput Phenotyping From Unmanned Aerial Systems by Extracting Traits Directly From Orthorectified Images [Original Research]. *Frontiers in Plant Science*, 11(1616). <https://doi.org/10.3389/fpls.2020.587093>
- Wang, X., Singh, D., Marla, S., Morris, G., & Poland, J. (2018). Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods*, 14. <https://doi.org/ARTN 53> 10.1186/s13007-018-0324-5
- Wang, X., Xuan, H., Evers, B., Shrestha, S., Pless, R., & Poland, J. (2019). High-throughput phenotyping with deep learning gives insight into the genetic architecture of flowering time in wheat. *Gigascience*, 8(11). <https://doi.org/ARTN giz120> 10.1093/gigascience/giz120
- Yau, S. (1997). Efficiency of alpha-lattice designs in international variety yield trials of barley and wheat. *The Journal of Agricultural Science*, 128(1), 5-9.
- Zhou, X., Kono, Y., Win, A., Matsui, T., & Tanaka, T. S. (2021). Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. *Plant Production Science*, 24(2), 137-151.
- Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., Cheng, T., Zhu, Y., Cao, W. X., & Tian, Y. C. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *Isprs Journal of Photogrammetry and Remote Sensing*, 130, 246-255. <https://doi.org/10.1016/j.isprsjprs.2017.05.003>

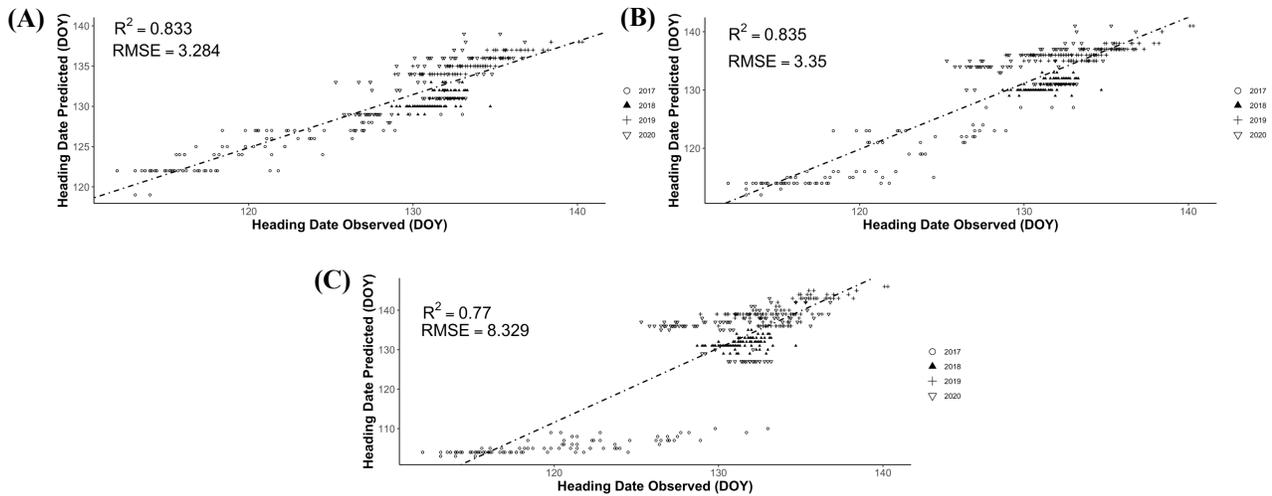
## Appendix A - Supplementary Material Chapter 2



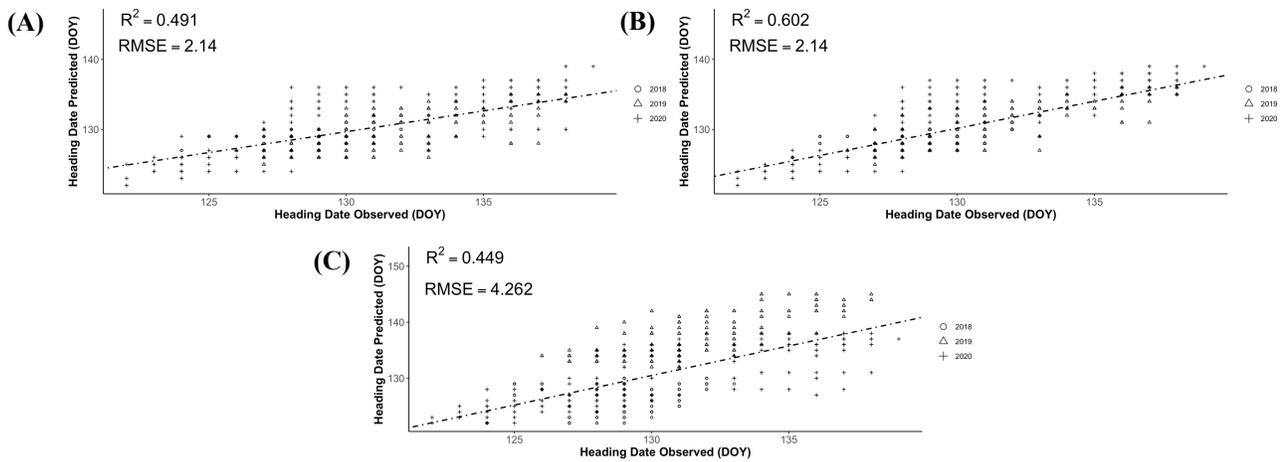
**Figure A.1.** Wang and Engel growing degree days ( $GDD_{WE}$ ) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods for the AM Panel. The x-axis is the observed heading date, calendar day of year (DOY)



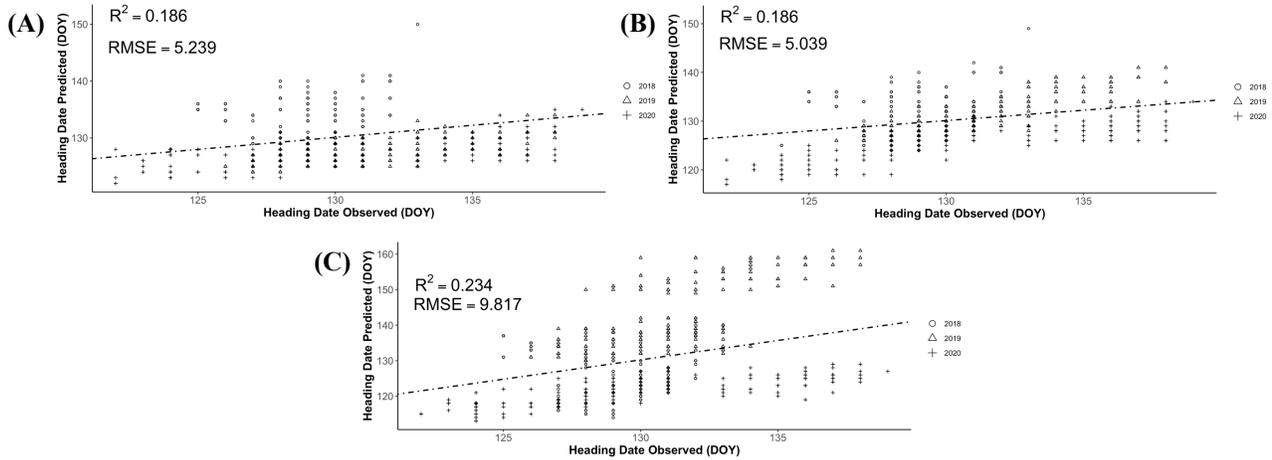
**Figure A.2** Physiological days (Pdays) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods for the AM Panel. The x-axis is the observed heading date, calendar day of year (DOY)



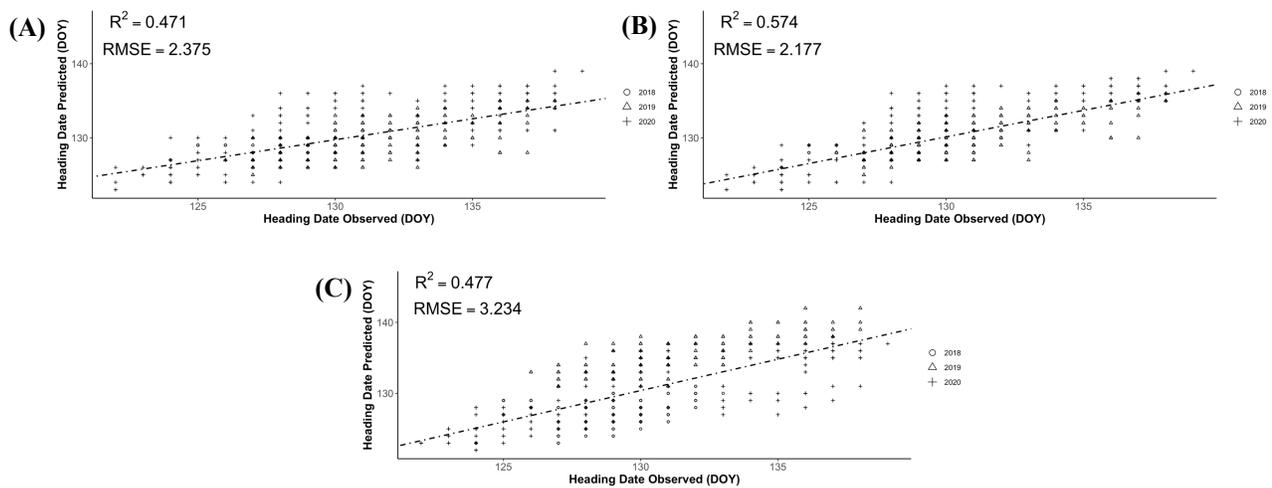
**Figure A.3** Photo Growing Degree Days (PGDD) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods for the AM Panel. The x-axis is the observed heading date, calendar day of year (DOY)



**Figure A.4** Wang and Engel growing degree days (GDD<sub>WE</sub>) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods for the Kansas State University Wheat Breeding program. The x-axis is the observed heading date, calendar day of year (DOY)

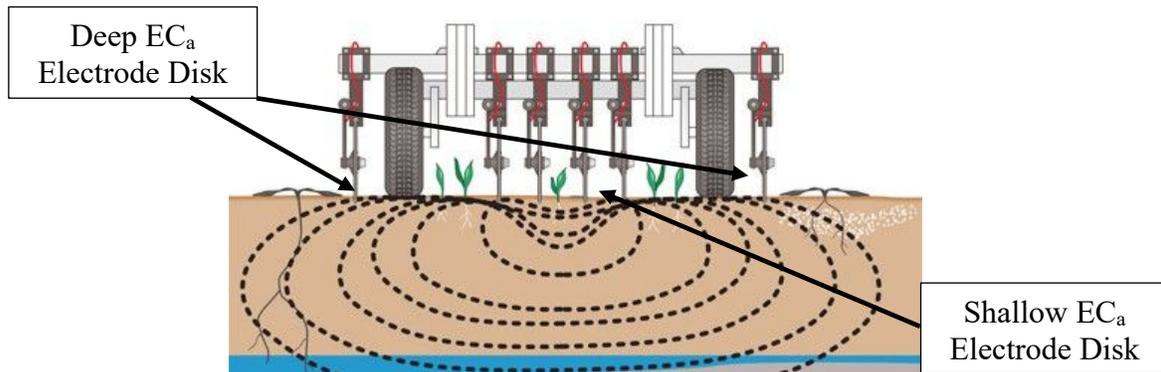


**Figure A.5** Physiological days (Pdays) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods for the Kansas State University Wheat Breeding program. The x-axis is the observed heading date, calendar day of year (DOY).

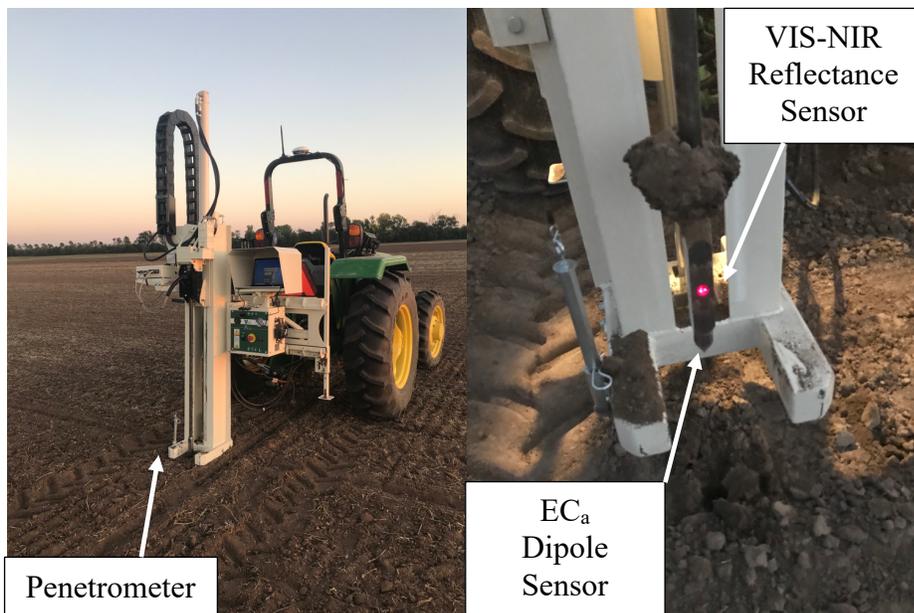


**Figure A.6** Photo Growing Degree Days (PGDD) Predicted vs Observed Heading Dates for full season (A), January 1<sup>st</sup> (B) and March 1<sup>st</sup> (C) accumulation periods for the Kansas State University Wheat Breeding program. The x-axis is the observed heading date, calendar day of year (DOY).

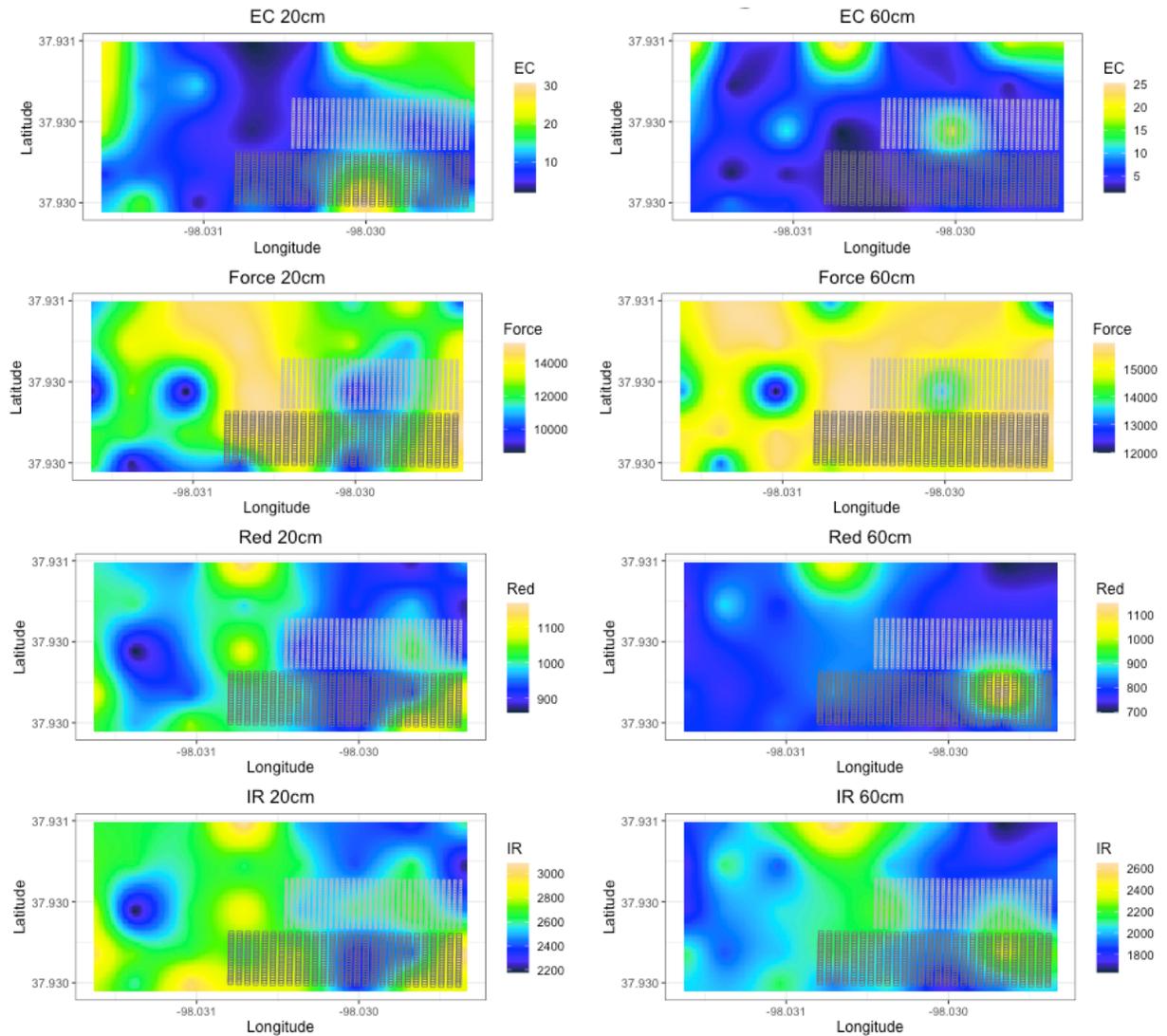
## Appendix B - Supplementary Material Chapter 3



**Figure B.1.** Diagram of the Veris MSP-3 mobile sensor cart platform used to collect apparent electric conductivity (EC<sub>a</sub>). This sensor has two pairs of coulter disks equipped with rolling electrodes. The disk pairs emit and receive small electrical current, giving EC<sub>a</sub> measurements at the 0-30 and -90 cm depths.



**Figure B.2** The P4000 DW-EC-Force Probe 3-point mounted on John Deere 5055E tractor equipped with Trimble RTK GNSS. The Veris P4000 sensor probe is equipped with penetrometer, VIS-NIR optical and electromagnetic dipole sensor.



**Figure B.3** Kriged map for EC, Force, Red and IR measurements from the P4000 at 21RNS with PYN and YT plot map overlays. Number of color clusters were determined by k-means clusters and the values for the cluster color represent the median value of the cluster range.

**Table B.1** 21RNS-PYN yield rank order of raw yields experimental design adjusted yields, spatially adjusted yields based on soil core kriged values, spatially adjusted yields based on soil MSP3 kriged values, and spatially adjusted yields based on soil P4000 kriged values.

Variety Name	Raw Yield	Raw Rank Order	Adj. Yield	Adj. Rank Order	Core Spatial Adj. Yield	Core Rank Order	MSP3 Spatial Adj. Yield	MSP3 Rank Order	P4000 Spatial Adj. Yield	P4000 Rank Order
KS170087D-10	124.25	1	125.05	1	117.34	1	118.10	1	118.43	1
KS160998S-2	124.08	2	124.57	2	115.47	2	116.21	2	116.08	2
KS160563S-2	120.78	3	120.77	3	114.38	3	114.91	3	115.13	3
KS160563S-3	116.66	5	116.65	4	111.38	6	112.35	6	112.33	4
KS170025D-11	116.99	4	116.55	5	111.95	4	112.72	4	112.05	5
KS160897S-4	114.51	15	116.29	6	108.07	30	109.19	22	108.58	27
KS170001D-19	116	7	116.19	7	111.12	8	112.07	7	110.79	11
KS160911S-3	115.83	10	115.64	8	110.91	9	111.79	8	111.22	8
KS170275D-1	112.7	26	115.58	9	106.40	54	108.03	35	107.16	44
KS170013D-7	116	8	115.4	10	110.51	11	109.95	16	111.05	10
KS170087D-8	116	9	115.23	11	111.35	7	112.39	5	111.93	6
KS20DH20385	115.67	11	115.05	12	111.47	5	111.24	10	110.61	12
KS20DH201071	108.57	69	114.67	13	104.92	83	106.18	68	105.58	79
KS160765S-1	116	6	114.57	14	110.67	10	111.35	9	111.53	7
KS170013D-27	114.84	14	114.4	15	110.23	12	111.01	12	110.34	13
KS160419S-4	115.5	12	114.36	16	110.02	14	111.19	11	109.95	17
KS160897S-6	115.17	13	114.3	17	108.87	22	109.87	18	110.15	16
KS170013D-5	113.36	20	114.16	18	109.54	17	110.00	15	109.62	19
KS160786S-4	113.69	17	114.05	19	109.44	18	109.93	17	109.21	20
KS170087D-3	113.03	24	113.83	20	109.77	16	110.09	14	111.20	9
KS170025D-18	114.18	16	113.74	21	110.11	13	110.70	13	110.16	15
KS160786S-3	113.19	21	113.55	22	109.26	19	109.60	20	109.02	22
KS160393S-5	112.04	29	113.55	23	107.68	37	108.82	26	108.09	31
KS160853S-1	113.52	18	113.47	24	107.99	31	107.65	42	107.70	34
KS170225D-1	110.55	44	113.43	25	104.52	90	106.29	66	105.40	82
KS160786S-1	113.36	19	113.35	26	109.90	15	109.75	19	110.30	14
KS170087D-9	111.87	33	112.87	27	107.03	44	107.91	36	108.92	23
KS160720S-1	110.88	37	112.82	28	107.28	40	108.31	28	108.23	30
KS170025D-6	113.19	22	112.8	29	108.30	26	108.83	24	109.80	18
KS170210D-5	113.19	23	112.76	30	108.24	27	107.74	40	108.70	25
KS1720024A-1	109.23	57	112.64	31	105.27	77	106.36	63	105.67	73
KS160382S-6	111.87	31	112.58	32	107.79	34	107.77	39	107.60	36
KS160519S-1	112.7	25	111.67	33	108.88	21	108.83	25	108.65	26
KS160524S-9	112.53	27	111.5	34	107.77	35	108.14	33	107.63	35
KS170025D-19	112.04	30	111.44	35	107.69	36	107.72	41	109.07	21
KS170013D-3	110.39	48	111.3	36	106.54	49	107.56	43	106.60	49
KS170225D-2	111.71	34	111.28	37	107.26	42	106.62	56	107.47	40
KS160472S-2	111.87	32	111.13	38	107.11	43	108.22	30	107.38	41
KS160786S-6	110.55	43	110.91	39	106.60	48	106.88	51	106.41	55
KS160786S-2	110.88	38	110.87	40	107.98	32	108.06	34	107.75	33
KS20DH21171	105.44	112	110.87	41	101.90	145	102.50	135	102.68	128
KS20DH21164	105.27	114	110.7	42	101.74	150	102.55	133	102.70	126

KS160786S-5	110.22	49	110.58	43	106.84	45	107.24	46	106.57	52
KS170013D-9	110.22	50	109.78	44	107.28	41	107.87	37	107.30	43
KS160458S-1	110.88	36	109.74	45	106.64	46	107.84	38	106.24	58
KS170013D-25	109.23	56	109.74	46	104.36	93	105.51	82	104.81	90
KS160382S-3	110.39	47	109.65	47	106.60	47	106.99	48	106.05	62
KS160390S-3	108.08	74	109.59	48	105.00	80	106.03	71	105.46	81
KS160365S-4	109.89	51	109.56	49	105.83	61	106.45	58	106.15	60
KS160366S-3	109.89	52	109.56	50	105.81	62	106.54	57	105.78	70
KS160524S-10	110.55	42	109.52	51	107.38	38	107.07	47	107.09	45
KS20DH20893	109.23	58	109.15	52	105.45	74	105.71	79	105.67	74
KS170134D-4	109.56	55	109.13	53	105.62	68	105.72	78	107.38	42
KS170275D-3	108.24	72	109.1	54	104.88	85	104.23	99	104.38	95
KS20DH21193	103.62	142	109.05	55	100.75	173	101.09	168	101.02	169
KS170087D-7	107.91	80	108.91	56	104.37	92	105.22	87	105.83	68
KS160911S-1	109.07	59	108.77	57	106.14	56	106.94	49	106.27	57
KS170013D-23	108.08	76	108.59	58	103.28	112	104.73	93	103.90	102
KS20DH20780	101.81	177	108.58	59	98.79	234	99.01	226	98.42	245
KS170025D-1	109.07	60	108.47	60	105.78	63	105.31	84	106.91	47
KS160906S-9	108.74	64	108.44	61	105.76	64	106.73	53	105.86	67
KS170013D-2	106.26	98	108.34	62	103.10	115	102.81	122	103.10	119
KS160332S-9	108.57	66	108.24	63	104.68	88	105.74	77	105.70	72
KS160786S-12	108.24	71	108.23	64	106.49	51	106.21	67	106.40	56
KS20DH20879	107.91	82	108.23	65	104.53	89	104.57	96	104.64	92
KS160469S-3	108.9	63	108.16	66	105.55	72	106.10	69	105.58	77
KS170013D-30	107.25	89	108.16	67	104.91	84	105.52	81	104.87	89
KS160332S-3	108.57	65	108.15	68	105.60	69	106.34	65	105.40	83
KS20CFB-21	103.79	139	108.14	69	102.66	129	102.83	121	102.09	142
KS160366S-1	108.41	70	108.08	70	104.85	86	105.30	85	104.91	88
KS160674S-1	109.56	54	107.95	71	105.86	60	106.62	54	106.57	51
KS20DH200727	105.6	108	107.79	72	102.95	122	103.57	109	103.23	115
KS160855S-2	107.75	84	107.77	73	105.59	70	105.96	73	105.95	63
KS160352S-1	108.08	73	107.75	74	104.40	91	105.30	86	105.35	84
KS160765S-6	107.75	83	107.74	75	104.96	81	105.67	80	106.59	50
KS170134D-3	108.08	77	107.73	76	105.89	59	106.35	64	105.93	64
KS160352S-3	107.09	90	107.65	77	103.04	119	103.41	112	103.60	103
KS20CFB-24	103.29	148	107.64	78	102.60	133	102.64	132	101.90	145
KS160366S-2	106.59	93	107.15	79	103.10	116	104.21	100	103.99	100
KS160855S-4	107.09	92	107.11	80	105.00	79	105.35	83	105.29	85
KS160911S-4	107.25	88	107.06	81	105.36	76	105.80	76	105.59	75
KS160473S-3	108.08	75	107.05	82	105.56	71	105.82	75	105.58	78
KS160332S-11	107.42	85	107	83	105.37	75	105.13	90	104.08	99
KS170275D-12	106.1	101	106.96	84	103.54	107	102.67	130	103.43	109
KS161024S-1	107.91	78	106.94	85	105.64	67	106.43	60	107.79	32
KS170001D-14	107.91	79	106.94	86	105.69	65	106.62	55	106.54	53
KS170134D-5	105.11	119	106.81	87	103.27	113	103.93	104	103.59	104
KS20DH20851	106.43	97	106.75	88	102.97	121	103.09	117	102.78	123
KS170313D-10	101.81	175	106.7	89	99.61	207	100.37	189	100.02	199
KS20DH200649	104.45	128	106.64	90	101.66	153	101.18	164	101.70	154
KS160383S-3	105.11	115	106.62	91	102.68	127	103.75	106	103.46	107
KS20DH20867	106.26	99	106.58	92	103.29	111	103.29	114	103.26	114
KS160856S-2	107.42	86	106.55	93	103.82	101	103.53	110	103.47	106
KS170313D-8	101.64	182	106.53	94	101.27	161	101.32	162	100.96	172
KS20DH20807	99.66	217	106.43	95	97.79	263	98.21	248	97.43	266

KS160352S-2	105.6	106	106.16	96	101.92	144	101.85	150	102.14	140
KS160473S-6	107.25	87	106.11	97	103.92	100	105.21	88	103.46	108
KS160572S-2	105.93	103	106.11	98	104.05	97	103.75	107	103.49	105
KS160998S-1	105.6	107	106.09	99	102.99	120	103.44	111	104.18	97
KS20DH20806	99.17	230	105.94	100	97.32	271	97.66	263	96.94	283
KS160617S-1	105.93	104	105.92	101	104.14	96	104.83	92	105.05	86
KS160524S-13	105.44	109	105.89	102	103.58	106	103.21	115	103.39	110
KS170013D-18	103.79	138	105.87	103	100.64	178	100.94	174	99.75	208
KS170313D-4	100.98	191	105.87	104	99.54	211	100.13	196	99.73	210
KS170013D-26	104.94	122	105.85	105	101.68	152	102.53	134	101.98	143
KS160390S-2	105.11	116	105.82	106	102.93	123	102.68	128	103.17	117
KS160332S-8	106.1	100	105.77	107	103.96	99	103.83	105	103.95	101
KS20DH20746	99	236	105.77	108	96.48	291	96.38	293	95.50	321
KS160393S-4	104.78	124	105.75	109	102.62	131	103.03	119	102.71	125
KS160677S-1	107.09	91	105.66	110	104.85	87	104.94	91	105.04	87
KS160570S-1	105.44	110	105.62	111	103.63	105	103.35	113	102.63	134
KS170013D-13	104.94	121	105.53	112	101.75	148	101.89	148	100.55	182
KS1720021A-2	110.55	45	105.37	113	108.13	29	108.21	31	108.33	28
KS160897S-1	104.28	130	105.31	114	102.31	137	101.65	153	101.82	150
KS170001D-20	105.11	117	105.3	115	103.50	108	104.19	101	103.00	120
KS170013D-29	104.28	131	105.19	116	101.83	147	102.67	129	101.92	144
KS170210D-4	108.57	67	105.16	117	104.93	82	106.43	59	105.46	80
KS170087D-12	104.12	136	105.12	118	102.19	139	102.25	140	101.79	152
KS160971S-1	105.27	113	105.08	119	104.24	94	104.44	97	104.62	93
KS160906S-1	103.29	147	105.07	120	100.54	181	101.15	165	100.86	174
KS160897S-3	105.93	105	105.06	121	102.71	126	102.66	131	103.23	116
KS20DH20161	110.88	40	104.75	122	108.21	28	108.29	29	107.56	38
KS1720020A-1	112.2	28	104.6	123	109.13	20	109.59	21	108.80	24
KS20DH20189	110.72	41	104.59	124	107.31	39	107.44	45	106.70	48
KS170087D-4	105.11	118	104.34	125	104.19	95	104.68	94	105.59	76
KS20DH20350	104.94	123	104.32	126	103.69	103	102.89	120	102.68	127
KS160786S-7	104.12	134	104.07	127	101.93	143	102.18	143	102.40	137
KS160897S-5	102.14	168	103.92	128	99.80	199	100.42	188	100.17	194
KS20DH20052	101.81	176	103.88	129	98.30	245	99.38	210	98.88	230
KS170013D-31	102.96	155	103.87	130	102.28	138	102.49	136	102.16	139
KS160332S-5	104.28	129	103.86	131	102.73	125	103.17	116	102.56	135
KS20DH20838	103.46	144	103.78	132	100.74	174	101.05	169	100.47	187
KS20CFB-42	104.78	126	103.75	133	102.83	124	103.97	103	103.35	112
KS170087D-1	104.12	135	103.68	134	103.46	109	103.59	108	103.35	113
KS20DH20609	98.34	247	103.63	135	95.70	318	96.73	287	96.23	299
KS170087D-5	104.28	132	103.51	136	103.79	102	104.14	102	104.71	91
KS20CFB-40	103.62	141	103.3	137	101.44	158	100.93	175	101.50	158
KS1720020A-2	110.88	39	103.28	138	108.56	23	108.97	23	108.23	29
KS160971S-3	103.46	143	103.27	139	103.09	117	103.04	118	104.13	98
KS170001D-8	101.97	171	103.23	140	100.93	166	101.40	159	100.86	173
KS160383S-2	102.47	162	103.18	141	101.54	155	101.12	167	101.28	165
KS20DH20415	103.79	140	103.17	142	103.42	110	102.76	125	102.80	122
KS160410S-5	104.78	125	103.14	143	102.08	140	102.45	137	102.67	130
KS160519S-3	104.12	133	103.09	144	103.17	114	102.69	127	102.77	124
KS20DH20040	104.12	137	102.98	145	102.62	130	101.97	145	103.13	118
KS170013D-16	102.96	154	102.97	146	100.61	179	101.38	160	100.55	181
KS20DH20184	109.07	62	102.94	147	106.43	52	106.38	62	105.70	71
KS160650S-2	104.45	127	102.84	148	103.06	118	102.80	123	102.64	132

KS20DH200791	100.65	199	102.84	149	99.95	197	100.37	190	100.16	195
KS20DH20244	98.84	242	102.79	150	95.83	312	95.68	314	95.70	313
KS20DH21159	97.35	270	102.78	151	96.37	295	96.99	283	97.60	262
KS1720021A-3	107.91	81	102.73	152	106.40	55	106.41	61	106.42	54
KS160671S-4	102.96	152	102.68	153	101.75	149	101.94	147	101.89	146
KS160856S-3	101.64	179	102.67	154	100.10	193	99.75	203	99.99	202
KS160430S-4	103.29	146	102.55	155	101.85	146	101.85	149	101.87	147
KS170013D-4	103.13	151	102.53	156	102.06	141	100.64	182	101.68	155
KS160671S-3	102.8	159	102.52	157	101.40	159	101.59	154	101.60	156
KS160618S-1	102.47	163	102.46	158	101.36	160	102.44	138	102.68	129
KS160650S-1	102.47	164	102.46	159	100.76	172	102.21	142	102.67	131
KS20DH20826	102.14	169	102.46	160	99.72	202	100.18	195	99.42	216
KS160806S-1	102.47	165	102.42	161	100.69	177	100.24	192	100.68	178
KS160524S-12	101.97	170	102.42	162	101.48	156	100.93	176	101.14	168
KS160419S-5	103.13	150	102.39	163	101.47	157	101.48	157	101.81	151
KS160377S-1	101.81	174	102.37	164	100.17	191	101.04	170	101.00	171
KS160388S-2	101.64	178	102.35	165	100.88	168	100.51	186	100.79	175
KS170013D-17	100.16	205	102.24	166	97.99	257	98.27	247	96.98	281
KS20DH20064	103.29	149	102.15	167	101.96	142	101.84	151	102.64	133
KS20DH20089	99.66	216	102.13	168	100.08	194	99.54	207	99.06	223
KS160856S-4	102.96	153	102.09	169	100.80	171	100.22	193	100.71	177
KS160634S-2	101.15	186	102.03	170	99.57	210	99.43	208	100.00	200
KS20DH20080	99.5	219	101.97	171	99.69	203	99.21	215	98.81	233
KS160975S-3	101.48	183	101.89	172	101.00	165	100.97	173	101.16	167
KS160388S-3	101.15	185	101.86	173	100.37	187	99.95	199	100.50	186
KS20DH20493	102.96	157	101.81	174	101.55	154	101.22	163	101.34	163
KS160524S-6	102.8	158	101.77	175	101.22	162	100.77	180	101.27	166
KS160897S-2	99.99	208	101.77	176	98.75	235	99.17	218	98.80	235
KS20CFB-37	109.07	61	101.74	177	106.41	53	106.92	50	105.92	66
KS170013D-20	101.64	180	101.65	178	100.00	195	100.74	181	99.71	211
KS20DH20694	96.2	289	101.49	179	93.36	364	94.34	344	93.46	364
KS170313D-7	96.53	285	101.42	180	97.37	270	97.33	273	97.17	274
KS160382S-2	102.14	167	101.4	181	100.83	169	101.04	171	100.24	192
KS170025D-2	101.97	172	101.37	182	100.92	167	100.81	178	102.33	138
KS160752S-1	99.33	221	101.27	183	99.82	198	100.01	198	100.43	189
KS170961D-2	108.57	68	100.97	184	106.53	50	106.84	52	106.17	59
KS20DH20765	94.05	335	100.82	185	93.33	365	92.83	372	92.67	377
KS160765S-7	100.82	196	100.81	186	100.48	184	100.82	177	101.78	153
KS170225D-5	97.85	259	100.73	187	96.39	294	97.34	271	96.98	280
KS160393S-1	99.99	207	100.7	188	99.46	215	98.94	230	99.66	213
KS160634S-1	100.65	198	100.64	189	99.77	200	101.13	166	101.44	160
KS20DH21182	99.99	211	100.61	190	99.59	209	99.04	224	99.82	207
KS170134D-2	101.31	184	100.54	191	101.17	163	101.56	156	101.40	162
KS160671S-5	100.82	194	100.54	192	100.45	185	100.47	187	100.53	184
KS160383S-1	99	231	100.51	193	98.42	241	99.38	211	99.20	220
KS170210D-9	97.52	267	100.4	194	95.73	316	97.03	282	96.32	296
KS170087D-6	101.15	187	100.38	195	101.73	151	101.95	146	102.41	136
KS170013D-11	100.82	197	100.38	196	100.70	176	101.03	172	100.60	179
KS170087D-11	99.33	222	100.33	197	99.17	220	99.00	228	98.28	248
KS170087D-13	99.33	223	100.33	198	98.82	233	99.00	227	99.06	224
KS160786S-13	100.32	202	100.31	199	101.10	164	100.60	184	100.74	176
KS170025D-5	99.5	218	100.3	200	100.54	180	100.29	191	100.43	188
KS170001D-4	99	234	100.26	201	98.13	252	98.57	236	97.92	253

KS20DH20967	100.32	203	100.24	202	99.61	205	99.61	205	99.99	201
KS160393S-3	99.17	227	100.14	203	98.88	230	98.91	231	98.97	229
KS160410S-7	100.82	193	100.08	204	99.53	212	99.23	214	100.12	198
KS170210D-6	97.19	275	100.07	205	95.72	317	97.14	279	95.88	306
KS170210D-7	97.19	276	100.07	206	95.54	321	96.92	285	96.02	303
KS20CFB-18	95.7	294	100.05	207	95.31	328	96.03	308	95.59	316
KS20DH20673	94.71	322	100	208	92.50	379	93.21	362	92.98	373
KS20DH20125	110.55	46	99.99	209	108.37	24	108.20	32	107.07	46
KS20CFB-5	100.98	192	99.84	210	99.72	201	99.13	219	100.14	196
KS170013D-15	99.83	215	99.84	211	98.47	239	99.08	222	98.40	246
KS170100D-3	100.16	206	99.81	212	100.72	175	100.79	179	100.60	180
KS20CFB-47	99.33	225	99.79	213	97.27	273	97.49	266	97.39	269
KS160897S-7	98.01	251	99.79	214	97.01	282	97.47	268	97.30	271
KS170013D-10	98.51	243	99.77	215	99.01	223	99.09	220	98.88	231
KS170313D-9	94.88	315	99.77	216	97.04	277	96.70	289	96.55	290
KS20CFB-25	111.21	35	99.71	217	108.33	25	108.59	27	107.56	39
KS20CFB-28	99.99	209	99.67	218	98.99	225	98.51	239	98.80	234
KS20DH200258	93.56	341	99.66	219	94.61	345	92.94	369	93.78	356
KS160671S-11	97.68	262	99.62	220	97.51	268	98.85	233	99.13	222
KS160410S-9	100.32	201	99.58	221	99.31	219	99.04	225	99.84	205
KS20DH20131	100.65	200	99.48	222	100.50	182	100.10	197	99.31	217
KS160524S-11	99	233	99.45	223	99.60	208	98.87	232	99.18	221
KS160765S-4	100.82	195	99.39	224	100.50	183	100.64	183	101.29	164
KS170001D-11	98.84	240	99.25	225	98.21	250	97.84	256	97.23	273
KS20CFB-20	94.88	316	99.23	226	96.19	303	96.12	304	95.76	310
KS170134D-8	97.52	265	99.22	227	98.03	256	98.46	241	98.26	249
KS20CFB-39	101.15	190	99.19	228	100.39	186	100.60	185	100.14	197
KS170013D-19	97.02	279	99.1	229	96.56	289	96.17	301	95.76	311
KS1720020A-3	106.59	94	98.99	230	105.94	57	106.10	70	105.79	69
KS1720020A-4	106.59	95	98.99	231	105.93	58	105.99	72	105.92	65
KS160911S-5	99.83	214	98.96	232	98.89	229	99.42	209	100.20	193
KS1720024A-2	95.54	299	98.95	233	96.33	299	96.23	299	95.91	305
KS160906S-8	99.17	229	98.87	234	99.07	222	99.76	202	99.01	227
KS170013D-24	98.34	245	98.85	235	97.03	279	97.78	258	97.37	270
KS160524S-1	98.01	250	98.85	236	99.96	196	98.56	237	99.43	215
KS1720021A-1	106.43	96	98.83	237	105.66	66	105.91	74	106.06	61
KS20DH21010	99.99	210	98.82	238	99.35	217	99.84	200	99.90	203
KS160332S-2	99.17	226	98.75	239	99.61	206	99.66	204	98.73	237
KS20CFB-27	99	235	98.68	240	98.26	247	97.88	255	97.89	254
KS170013D-22	98.01	254	98.52	241	96.57	288	97.70	261	97.04	277
KS20DH21167	97.85	260	98.47	242	98.05	254	97.76	259	98.61	238
KS160720S-3	99.83	213	98.4	243	100.13	192	99.83	201	100.29	190
KS20CFB-23	99.33	224	98.3	244	99.34	218	100.19	194	99.73	209
KS160419S-6	99	232	98.26	245	98.88	231	98.72	235	99.01	228
KS20CFB-32	109.73	53	98.23	246	107.94	33	107.55	44	107.60	37
KS170275D-2	101.64	181	98.23	247	100.82	170	101.72	152	101.01	170
KS20DH20539	97.35	269	98.21	248	96.42	292	96.25	298	96.51	291
KS170001D-22	98.01	253	98.2	249	98.18	251	98.48	240	97.58	263
KS160410S-4	99.83	212	98.19	250	98.82	232	99.07	223	99.27	218
KS20DH20082	98.84	241	98.18	251	99.00	224	98.79	234	99.57	214
KS160419S-2	98.84	238	98.1	252	98.40	242	98.09	251	98.86	232
KS170013D-28	97.19	274	98.1	253	96.93	283	97.47	269	96.97	282
KS20DH21018	98.01	256	97.93	254	97.31	272	98.34	245	98.80	236

KS20DH20068	95.37	307	97.84	255	95.74	315	95.30	322	95.43	325
KS160971S-2	98.01	252	97.82	256	99.53	213	99.27	213	100.25	191
KS20CFB-36	105.11	120	97.78	257	104.01	98	104.30	98	103.37	111
KS160352S-5	97.19	271	97.75	258	96.77	285	97.32	274	97.46	264
KS20DH21142	103.46	145	97.74	259	100.33	189	101.47	158	101.87	148
KS170225D-7	101.15	188	97.74	260	100.21	190	101.34	161	100.52	185
KS20DH21195	92.24	354	97.67	261	93.20	367	93.07	364	93.02	372
KS170001D-3	96.36	287	97.62	262	95.97	310	96.20	300	95.59	315
KS160410S-3	99.17	228	97.53	263	99.66	204	98.52	238	98.57	239
KS160365S-2	97.85	257	97.52	264	97.85	261	98.01	252	98.10	251
KS160430S-2	98.18	248	97.44	265	98.39	243	98.18	249	98.44	242
KS20DH21011	97.52	268	97.44	266	97.06	276	97.95	253	98.37	247
KS160752S-4	98.84	239	97.41	267	99.40	216	99.21	216	99.84	206
KS170013D-6	98.01	255	97.41	268	98.67	237	97.12	280	98.44	243
KS20DH20348	93.06	345	97.41	269	95.15	332	93.91	353	94.22	348
KS20DH20010	98.51	244	97.37	270	98.04	255	97.69	262	99.03	225
KS20DH20074	94.88	317	97.35	271	96.04	309	95.43	319	95.39	327
KS170210D-3	97.52	266	97.17	272	98.11	253	98.46	242	98.47	240
KS170076D-5	97.52	264	97.13	273	97.91	258	97.31	275	97.76	257
KS160410S-1	100.16	204	97.09	274	100.33	188	99.08	221	99.69	212
KS170001D-27	97.02	278	97.03	275	96.85	284	96.94	284	96.41	295
KS170001D-10	97.85	258	96.88	276	99.14	221	99.36	212	100.55	183
KS160352S-8	97.02	277	96.69	277	97.22	274	97.47	267	97.71	260
KS20CFB-16	95.54	301	96.67	278	94.49	348	95.38	320	95.07	334
KS160524S-8	97.68	261	96.65	279	97.78	264	97.51	265	97.67	261
KS20DH20341	92.24	353	96.59	280	94.35	351	92.90	371	93.52	361
KS160906S-7	96.53	284	96.55	281	97.89	259	97.80	257	97.98	252
KS160475S-2	95.7	292	96.54	282	98.33	244	96.82	286	97.71	259
KS20DH20151	102.63	161	96.5	283	102.57	134	102.37	139	102.12	141
KS160473S-2	97.52	263	96.49	284	98.43	240	98.37	244	98.42	244
KS160430S-5	97.19	272	96.45	285	97.83	262	97.74	260	97.71	258
KS20DH20526	95.54	303	96.4	286	95.90	311	95.15	326	95.75	312
KS20DH20124	97.02	280	96.36	287	98.25	248	97.25	277	97.02	278
KS160383S-8	99.33	220	96.26	288	98.90	228	98.34	246	98.15	250
KS170001D-12	97.19	273	96.22	289	98.73	236	99.00	229	99.89	204
KS20DH200513	90.09	381	96.19	290	92.17	386	90.52	404	91.17	395
KS160524S-5	95.7	293	96.15	291	96.58	287	95.62	315	96.50	292
KS160332S-6	96.53	282	96.11	292	97.60	267	97.34	272	97.42	267
KS170134D-1	96.69	281	95.92	293	98.29	246	98.44	243	98.46	241
KS160856S-1	94.88	313	95.91	294	95.75	314	95.04	327	95.46	323
KS170087D-14	94.88	314	95.88	295	95.76	313	96.02	309	96.78	287
KS170001D-2	94.55	324	95.81	296	94.57	346	94.83	331	94.09	352
KS160147S-2	98.84	237	95.77	297	97.74	265	98.14	250	96.79	286
KS20DH20306	87.12	414	95.77	298	88.57	427	88.71	424	88.62	427
KS160469S-5	96.36	286	95.62	299	97.12	275	97.28	276	96.99	279
KS160528S-7	94.71	319	95.59	300	95.22	331	94.92	329	95.56	317
KS160856S-7	94.55	323	95.58	301	95.57	320	94.58	337	95.20	331
KS160524S-7	96.53	283	95.5	302	97.02	280	96.56	291	96.93	284
KS170001D-21	94.71	320	95.2	303	96.06	307	95.21	325	96.00	304
KS20CFB-35	102.47	166	95.14	304	102.61	132	102.71	126	101.83	149
KS20DH20389	95.7	297	95.08	305	98.24	249	97.16	278	97.16	275
KS160856S-6	94.05	333	95.08	306	95.14	333	94.28	346	94.84	339
KS160524S-2	94.88	310	95.06	307	94.84	339	95.55	317	94.37	344

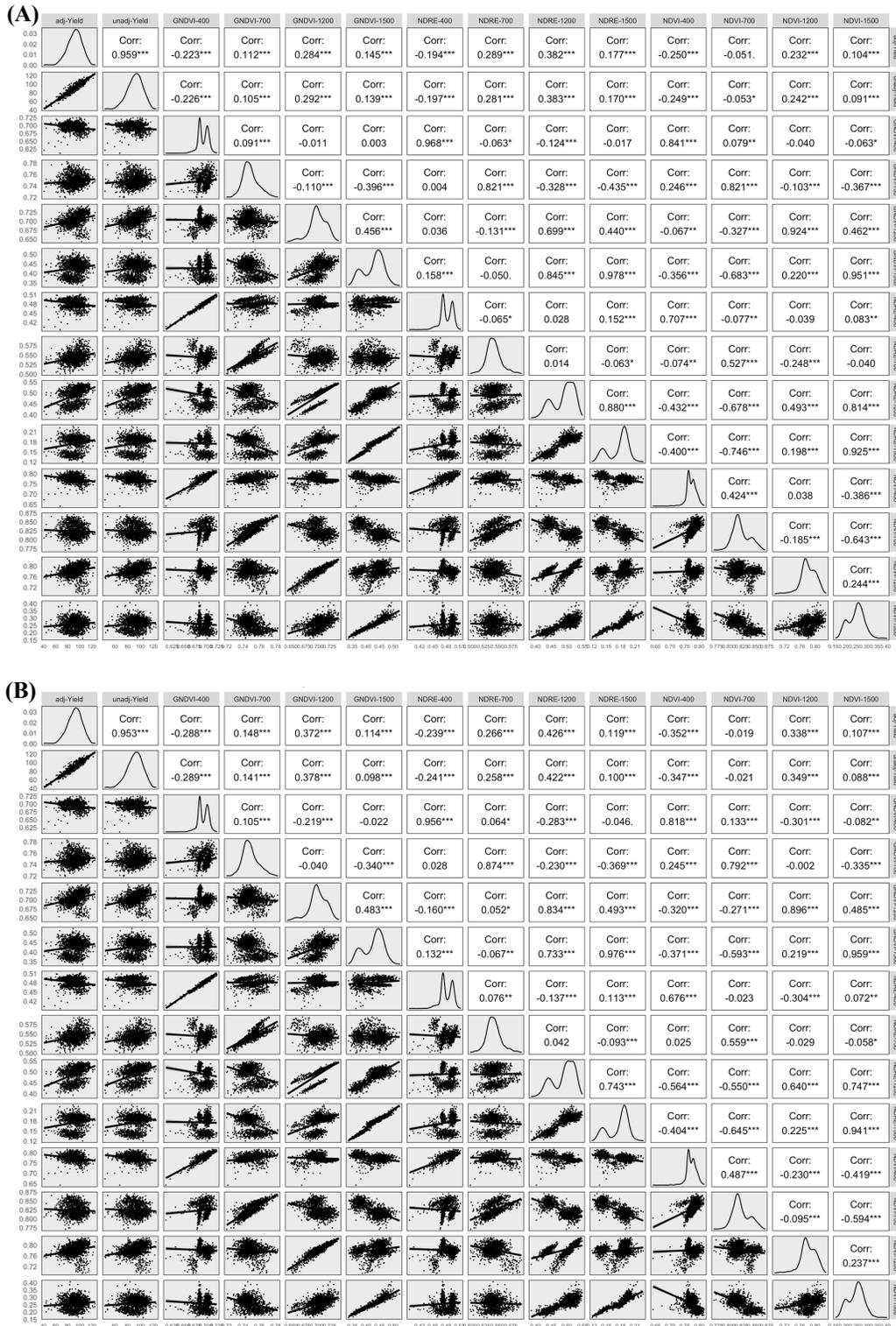
KS170100D-2	95.37	305	95.02	308	97.65	266	97.36	270	97.40	268
KS170100D-1	95.37	304	94.94	309	96.15	305	96.09	306	97.86	255
KS170313D-2	94.05	334	94.91	310	94.64	344	94.31	345	93.05	370
KS20DH20138	105.44	111	94.88	311	105.05	78	104.61	95	104.18	96
KS20DH20087	95.54	302	94.88	312	97.02	281	96.33	295	96.44	294
KS20DH20865	94.55	327	94.87	313	95.45	322	94.76	332	95.18	332
KS160855S-3	94.88	312	94.83	314	95.68	319	94.56	339	94.82	340
KS170961D-1	98.18	249	94.77	315	98.92	226	99.18	217	99.02	226
KS20CFB-15	93.56	340	94.69	316	92.82	374	93.83	355	93.46	363
KS20CFB-29	106.1	102	94.6	317	105.50	73	105.20	89	104.50	94
KS20DH20682	89.27	386	94.56	318	88.82	422	89.23	417	89.12	419
KS20DH20247	90.59	376	94.54	319	90.38	406	90.02	410	89.76	413
KS160786S-9	93.72	338	94.08	320	95.32	327	94.76	333	94.89	337
KS160816S-1	94.05	332	94.07	321	96.55	290	96.35	294	96.82	285
KS160572S-1	93.89	336	94.07	322	96.04	308	95.23	324	94.99	336
KS160671S-8	92.07	358	94.01	323	94.73	342	95.01	328	95.46	322
KS20DH20941	93.89	337	93.81	324	95.34	325	94.89	330	95.40	326
KS20CFB-4	95.7	295	93.74	325	96.19	304	96.16	303	96.16	300
KS20DH20227	95.7	296	93.74	326	95.41	323	95.59	316	95.51	318
KS160430S-3	94.88	309	93.74	327	96.34	297	96.60	290	95.84	307
KS160528S-2	94.71	318	93.68	328	96.11	306	94.56	338	95.51	319
KS160855S-1	93.23	343	93.59	329	94.75	341	93.96	351	94.40	342
KS20DH20434	94.71	321	93.56	330	96.33	298	96.09	307	96.65	289
KS20DH20243	89.6	383	93.55	331	89.78	412	89.26	416	89.61	415
KS20DH21054	92.9	348	93.52	332	94.67	343	94.51	341	95.28	329
KS20CFB-41	95.37	306	93.41	333	95.37	324	95.25	323	95.64	314
KS160472S-3	94.05	331	93.31	334	95.14	334	95.49	318	95.23	330
KS160671S-2	94.88	311	93.27	335	96.41	293	96.32	296	96.69	288
KS20DH20436	94.38	329	93.23	336	96.31	300	95.94	310	96.27	297
KS160975S-2	92.73	349	93.22	337	94.40	350	94.43	342	95.50	320
KS1720022A-1	98.34	246	93.16	338	99.52	214	99.59	206	99.23	219
KS160475S-1	96.2	288	93.13	339	98.49	238	96.71	288	97.84	256
KS20DH20013	90.92	370	92.99	340	90.31	408	90.73	400	90.24	409
KS20DH201114	96.03	290	92.79	341	97.03	278	97.55	264	97.14	276
KS160671S-7	94.22	330	92.79	342	96.27	302	95.73	313	96.23	298
KS160650S-3	94.38	328	92.77	343	96.34	296	95.87	311	96.08	301
KS20CFB-38	94.55	326	92.59	344	96.30	301	96.16	302	95.81	308
KS20DH20573	87.29	412	92.58	345	88.61	425	89.09	419	88.88	424
KS20DH20302	83.82	439	92.47	346	85.96	455	85.87	450	86.09	446
KS20DH20982	93.56	342	92.39	347	95.12	335	95.33	321	95.43	324
KS20CFB-7	91.91	360	92.37	348	91.60	393	91.88	384	91.47	391
KS160906S-3	92.57	351	92.27	349	93.70	361	93.74	356	93.64	359
KS20DH20638	86.96	415	92.25	350	87.79	433	88.23	427	88.21	430
KS20DH201081	102.8	160	92.24	351	102.67	128	102.22	141	101.46	159
KS1720023A-1	88.77	392	92.18	352	92.52	378	92.39	376	92.28	381
KS160765S-5	92.07	359	92.02	353	94.03	354	93.84	354	94.39	343
KS1720023A-2	88.61	396	92.02	354	92.16	387	92.13	379	92.01	384
KS170313D-3	87.12	413	92.01	355	90.13	410	90.25	408	90.19	410
KS170001D-6	90.59	374	91.85	356	93.21	366	93.00	366	92.94	374
KS170013D-12	91.25	366	91.84	357	92.94	371	92.17	378	91.54	390
KS20DH20460	92.9	347	91.75	358	95.25	330	94.41	343	94.88	338
KS160375S-1	92.07	355	91.74	359	93.97	356	93.93	352	93.56	360
KS160906S-5	92.57	352	91.7	360	93.97	357	94.03	350	95.01	335

KS20DH21058	91.08	367	91.7	361	93.53	363	93.07	365	94.05	353
KS170001D-18	91.41	363	91.6	362	95.05	337	94.60	336	94.33	346
KS20CFB-31	102.96	156	91.46	363	103.64	104	102.79	124	102.83	121
KS170001D-25	91.25	365	91.44	364	93.16	368	92.95	368	92.54	378
KS160524S-4	90.92	368	91.37	365	93.58	362	92.48	375	93.35	365
KS160393S-2	90.26	377	91.23	366	92.99	370	92.21	377	93.03	371
KS160671S-14	89.27	385	91.21	367	92.47	380	93.09	363	93.50	362
KS20CFB-19	86.79	417	91.14	368	89.68	414	89.84	412	89.69	414
KS160671S-13	89.1	387	91.04	369	92.03	388	92.96	367	93.34	366
KS160147S-4	91.41	361	90.99	370	94.98	338	94.08	349	93.10	369
KS160671S-6	92.57	350	90.96	371	94.79	340	94.63	335	95.13	333
KS170313D-11	90.09	380	90.95	372	92.43	381	91.47	389	90.97	399
KS20DH20236	92.9	346	90.94	373	92.66	376	93.54	359	93.15	368
KS160410S-8	92.07	356	90.93	374	94.55	347	94.52	340	94.31	347
KS20DH21040	90.92	371	90.84	375	92.38	383	93.31	361	94.16	349
KS20DH20251	82.17	449	90.82	376	84.34	470	84.34	463	83.95	469
KS20CFB-49	90.26	379	90.72	377	90.87	403	91.00	396	90.95	400
KS20CFB-26	101.97	173	90.47	378	102.43	136	102.14	144	101.43	161
KS20DH20242	86.46	420	90.41	379	87.17	443	87.78	435	88.01	431
KS170076D-3	89.6	382	90.4	380	94.25	352	93.50	360	94.14	350
KS20CFB-17	89.1	388	90.23	381	90.44	405	91.21	393	90.84	402
KS20CFB-2	85.64	426	89.99	382	89.50	416	89.09	420	89.20	418
KS20DH21051	95.7	298	89.98	383	94.45	349	96.40	292	97.29	272
KS170134D-6	90.26	378	89.91	384	93.79	358	93.64	357	93.71	357
KS1720024A-3	95.54	300	89.82	385	95.34	326	95.81	312	95.37	328
KS20DH20307	85.31	430	89.66	386	88.13	430	87.88	433	87.85	435
KS20CFB-30	101.15	189	89.65	387	102.46	135	101.58	155	101.50	157
KS20DH20048	90.75	372	89.61	388	93.74	360	92.67	374	94.13	351
KS20DH20147	90.75	373	89.58	389	93.76	359	92.94	370	92.78	375
KS20CFB-50	90.59	375	89.45	390	92.85	373	91.97	382	93.30	367
KS160906S-4	89.43	384	89.13	391	91.91	389	92.03	380	91.73	387
KS20CFB-9	88.61	397	89.07	392	89.50	415	89.85	411	88.93	423
KS160474S-2	92.07	357	89	393	95.26	329	93.61	358	94.50	341
KS20DH20019	86.79	418	88.86	394	87.52	436	87.55	437	87.77	438
KS20DH20287	80.19	463	88.84	395	83.13	475	82.76	474	83.14	474
KS1720025A-2	94.55	325	88.83	396	94.16	353	94.73	334	94.33	345
KS160534S-2	88.61	394	88.79	397	91.27	399	91.19	394	90.73	403
KS160534S-1	88.28	399	88.46	398	90.70	404	90.84	397	90.10	411
KS20DH20085	89.1	389	88.44	399	92.55	377	91.97	383	92.76	376
KS160327S-2	91.41	362	88.34	400	93.10	369	92.77	373	92.10	382
KS160906S-6	88.61	395	88.31	401	91.54	394	91.73	386	91.33	393
KS20DH20553	87.45	407	88.31	402	89.40	418	89.31	415	89.47	416
KS160752S-3	88.28	400	88.23	403	91.69	391	91.49	388	92.07	383
KS160671S-1	88.44	398	88.16	404	91.42	395	90.79	398	91.56	389
KS160524S-14	87.29	409	88.13	405	92.89	372	91.26	391	92.35	380
KS160524S-3	87.45	404	87.9	406	91.37	397	90.13	409	91.05	397
KS20DH20449	88.94	391	87.79	407	92.75	375	92.03	381	92.49	379
KS20DH20756	81.02	456	87.79	408	84.64	466	83.49	471	83.67	472
KS20DH20346	88.94	390	87.77	409	92.28	385	91.84	385	91.88	385
KS170076D-7	88.11	402	87.72	410	91.83	390	90.35	407	90.85	401
KS160528S-3	88.61	393	87.58	411	92.42	382	91.28	390	91.31	394
KS170275D-6	90.92	369	87.51	412	94.01	355	94.13	348	93.95	355
KS20DH20056	85.31	429	87.38	413	87.63	434	87.96	430	87.85	433

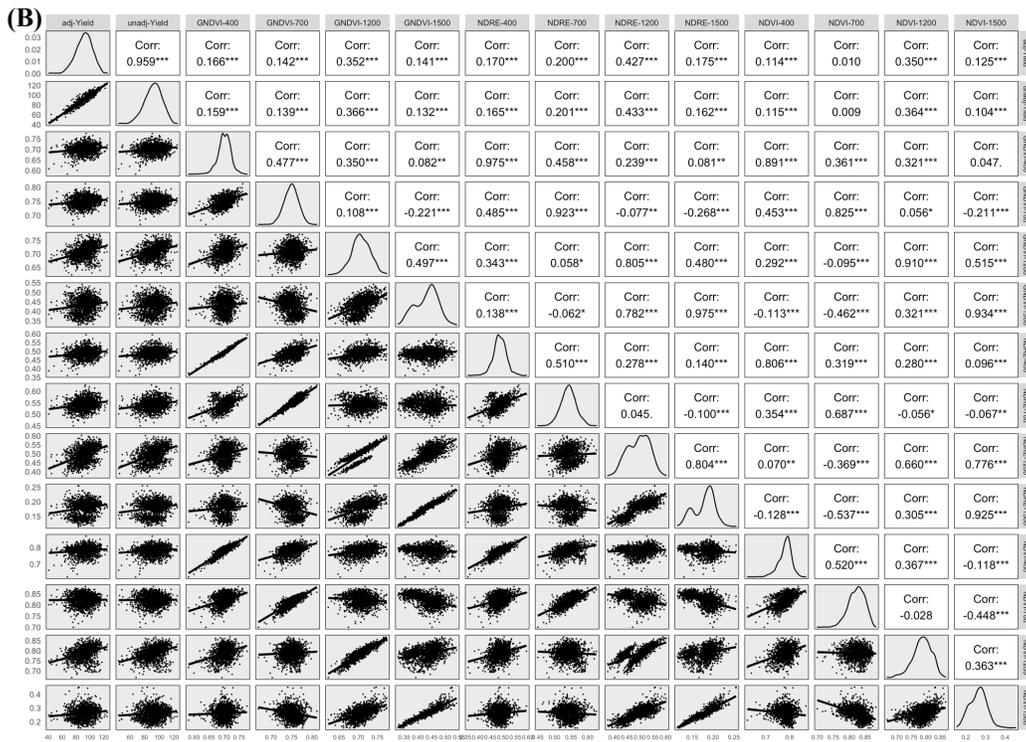
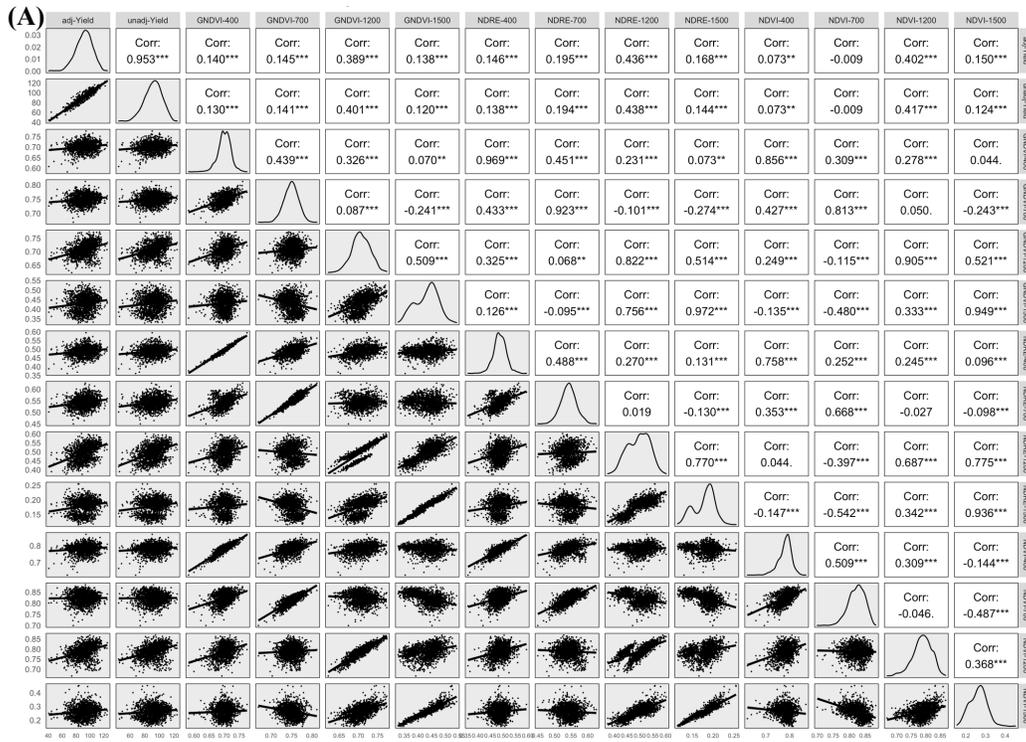
KS20DH20065	84.81	432	87.28	414	88.15	429	87.81	434	87.97	432
KS160975S-1	87.45	405	87.26	415	92.32	384	91.64	387	93.65	358
KS170076D-10	87.29	410	86.9	416	91.09	400	90.43	405	91.59	388
KS20DH20511	85.97	421	86.83	417	89.77	413	88.74	422	89.42	417
KS160382S-4	87.45	403	86.71	418	91.60	392	90.75	399	90.69	404
KS20CFB-14	85.31	428	86.44	419	87.09	445	87.59	436	87.48	440
KS20DH20057	83.82	438	86.29	420	87.20	442	87.12	443	87.15	442
KS160528S-4	85.14	431	86.02	421	89.13	420	88.10	429	88.96	421
KS172019A-1	86.79	416	85.76	422	91.33	398	91.25	392	91.36	392
KS20CFB-33	93.06	344	85.73	423	96.63	286	96.09	305	95.79	309
KS20DH20193	87.45	406	85.49	424	91.42	396	90.58	402	90.66	405
KS172019A-2	86.46	419	85.43	425	91.07	401	91.08	395	91.13	396
KS20CFB-43	87.29	411	85.33	426	89.43	417	89.46	414	89.87	412
KS20DH20134	95.87	291	85.31	427	98.91	227	97.89	254	97.44	265
KS160375S-5	85.64	424	85.31	428	89.05	421	89.12	418	88.38	428
KS20DH20180	91.41	364	85.28	429	95.09	336	94.26	347	93.97	354
KS20DH20119	95.21	308	84.65	430	97.87	260	97.04	281	96.44	293
KS20CFB-10	83.49	441	84.62	431	86.00	452	86.39	446	85.87	451
KS20DH20044	82.17	447	84.24	432	84.89	465	84.99	456	85.27	459
KS170001D-7	83.49	440	84.08	433	88.58	426	87.12	442	87.77	437
KS170076D-8	83.16	443	83.96	434	90.17	409	89.00	421	90.38	407
KS160382S-5	84.65	433	83.91	435	89.82	411	88.71	423	88.96	422
KS160671S-12	84.15	435	83.87	436	88.70	424	87.91	432	88.72	425
KS20DH20336	79.37	467	83.72	437	85.02	463	83.56	470	84.33	467
KS160634S-3	83.82	437	83.54	438	88.18	428	87.19	439	88.30	429
KS20DH200578	81.18	453	83.37	439	85.97	454	84.10	466	85.23	460
KS20CFB-11	82.17	446	83.3	440	85.00	464	85.20	454	85.23	461
KS170013D-14	83.16	442	83.17	441	87.44	440	87.16	440	87.04	443
KS20DH20136	93.72	339	83.16	442	97.48	269	96.28	297	96.07	302
KS20DH200205	76.89	475	82.99	443	83.63	474	83.46	472	83.78	470
KS170001D-5	82.83	444	82.84	444	87.35	441	86.85	445	86.79	445
KS160906S-2	81.02	454	82.8	445	85.95	456	85.46	453	85.77	452
KS170001D-16	82.34	445	82.75	446	87.06	446	86.18	448	85.74	454
KS20DH20215	84.65	434	82.69	447	89.29	419	88.41	425	88.62	426
KS1720025A-1	88.28	401	82.56	448	90.33	407	90.39	406	90.32	408
KS20DH21050	87.45	408	81.73	449	88.82	423	90.56	403	91.80	386
KS20CFB-46	81.18	452	81.64	450	84.58	467	85.00	455	85.29	458
KS160528S-6	80.52	458	81.4	451	85.83	458	84.94	457	85.91	450
KS160856S-5	80.36	461	81.39	452	85.99	453	84.72	458	85.55	457
KS160410S-2	80.36	460	81.33	453	86.69	448	85.72	451	85.94	449
KS170313D-1	80.36	462	81.22	454	86.37	450	84.53	462	85.22	462
KS170001D-9	80.52	459	81.11	455	86.22	451	84.69	459	84.80	465
KS170134D-7	81.35	451	81	456	87.60	435	87.26	438	87.57	439
KS1720022A-2	85.8	423	80.62	457	91.02	402	90.62	401	90.58	406
KS20DH21047	85.97	422	80.25	458	87.44	439	89.53	413	90.99	398
KS20DH20235	82.17	448	80.21	459	85.85	457	86.04	449	86.03	447
KS160489S-3	81.35	450	80.21	460	86.58	449	86.85	444	85.75	453
KS160327S-1	79.37	466	79.93	461	84.53	468	83.36	473	84.12	468
KS172010A-1	85.64	425	79.92	462	88.04	431	88.29	426	87.85	434
KS172009A-1	85.47	427	79.75	463	87.92	432	88.12	428	87.84	436
KS160489S-1	80.52	457	79.49	464	87.44	438	86.29	447	86.90	444
KS160410S-6	80.03	464	78.39	465	85.74	459	84.63	460	85.99	448
KS20DH21057	83.99	436	78.27	466	87.13	444	87.91	431	89.03	420

KS20CFB-6	77.55	474	78.01	467	82.13	477	81.80	477	81.89	477
KS20CFB-45	79.86	465	77.9	468	84.23	471	84.22	464	84.69	466
KS160375S-4	78.21	471	77.88	469	84.50	469	84.01	468	83.69	471
KS160382S-1	78.05	472	77.72	470	83.79	472	83.57	469	83.06	476
KS170313D-5	81.02	455	77.61	471	87.51	437	87.13	441	87.30	441
KS20DH21031	78.71	469	77.54	472	85.27	461	84.59	461	85.63	456
KS20DH20486	78.54	470	77.39	473	85.54	460	84.03	467	84.91	463
KS160671S-15	77.72	473	76.11	474	85.27	462	84.17	465	84.89	464
KS20DH20021	73.92	480	75.99	475	79.05	481	78.59	482	79.21	481
KS20DH20224	76.73	476	74.77	476	83.76	473	82.41	475	83.11	475
KS170001D-23	74.58	478	74.77	477	82.48	476	81.64	478	81.51	478
KS20CFB-13	74.58	479	74.26	478	81.74	479	80.83	479	80.57	479
KS20DH20170	78.71	468	72.58	479	86.94	447	85.51	452	85.67	455
KS170001D-1	71.12	483	71.53	480	80.50	480	79.34	480	80.15	480
KS172009A-4	71.45	481	71.13	481	79.00	482	78.60	481	77.59	483
KS20DH21098	75.9	477	70.18	482	82.00	478	82.01	476	83.32	473
KS172010A-2	70.13	484	69.81	483	78.39	484	77.77	484	76.99	484
KS172014A-3	71.45	482	65.73	484	78.81	483	78.45	483	78.02	482
KS170313D-6	63.86	485	64.72	485	74.41	485	72.79	485	72.27	485

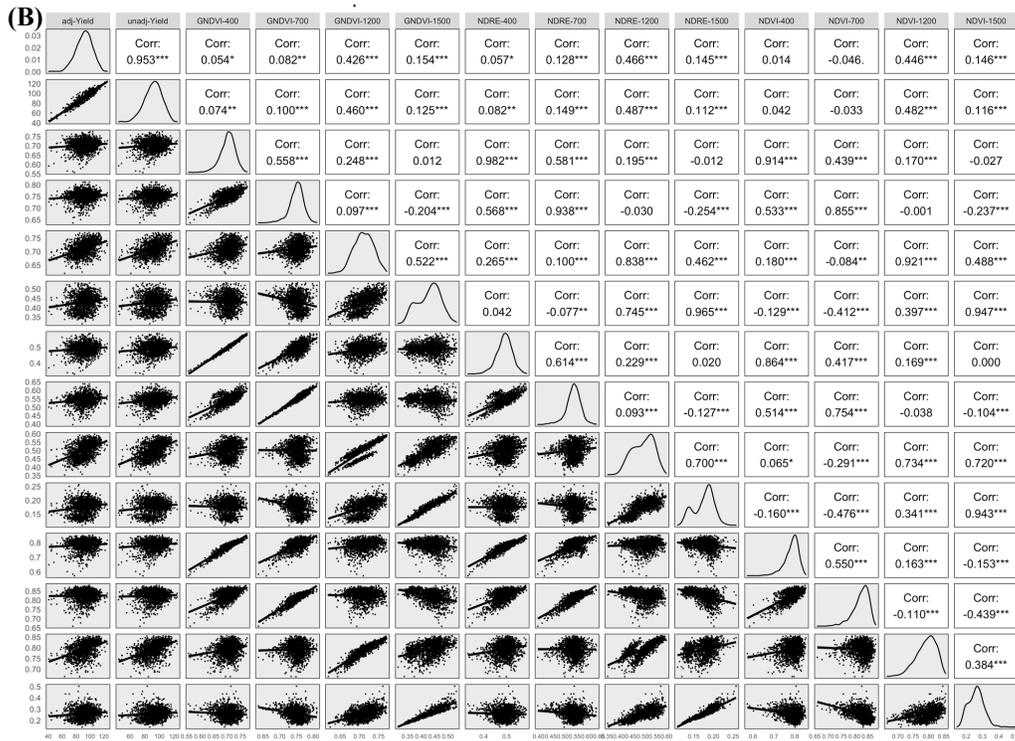
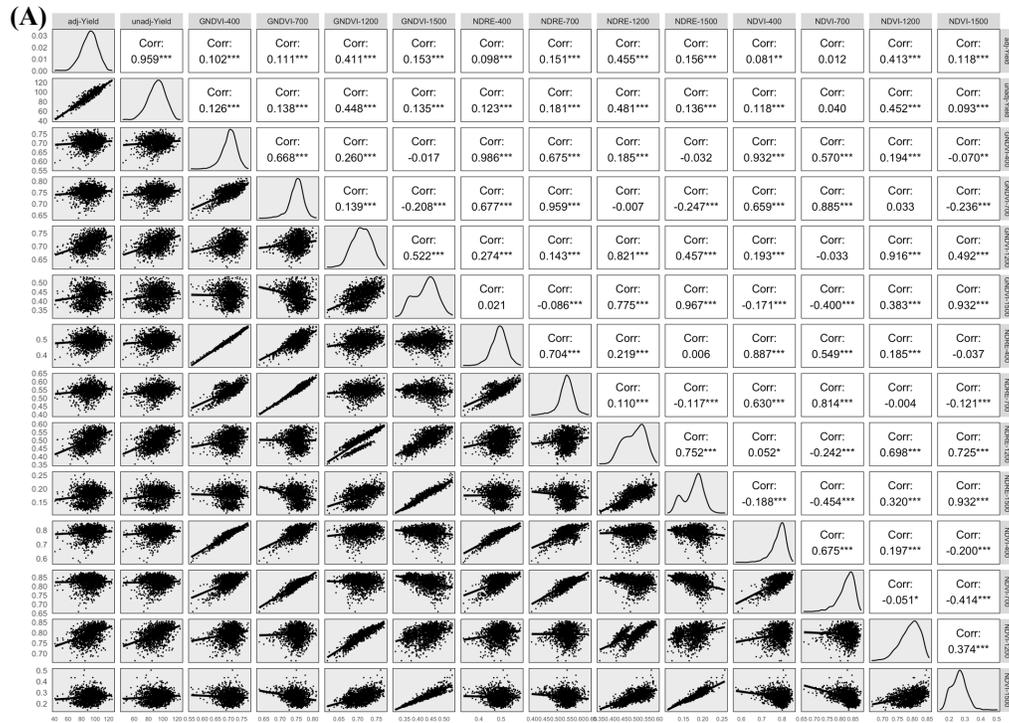
# Appendix C - Supplementary Material Chapter 4



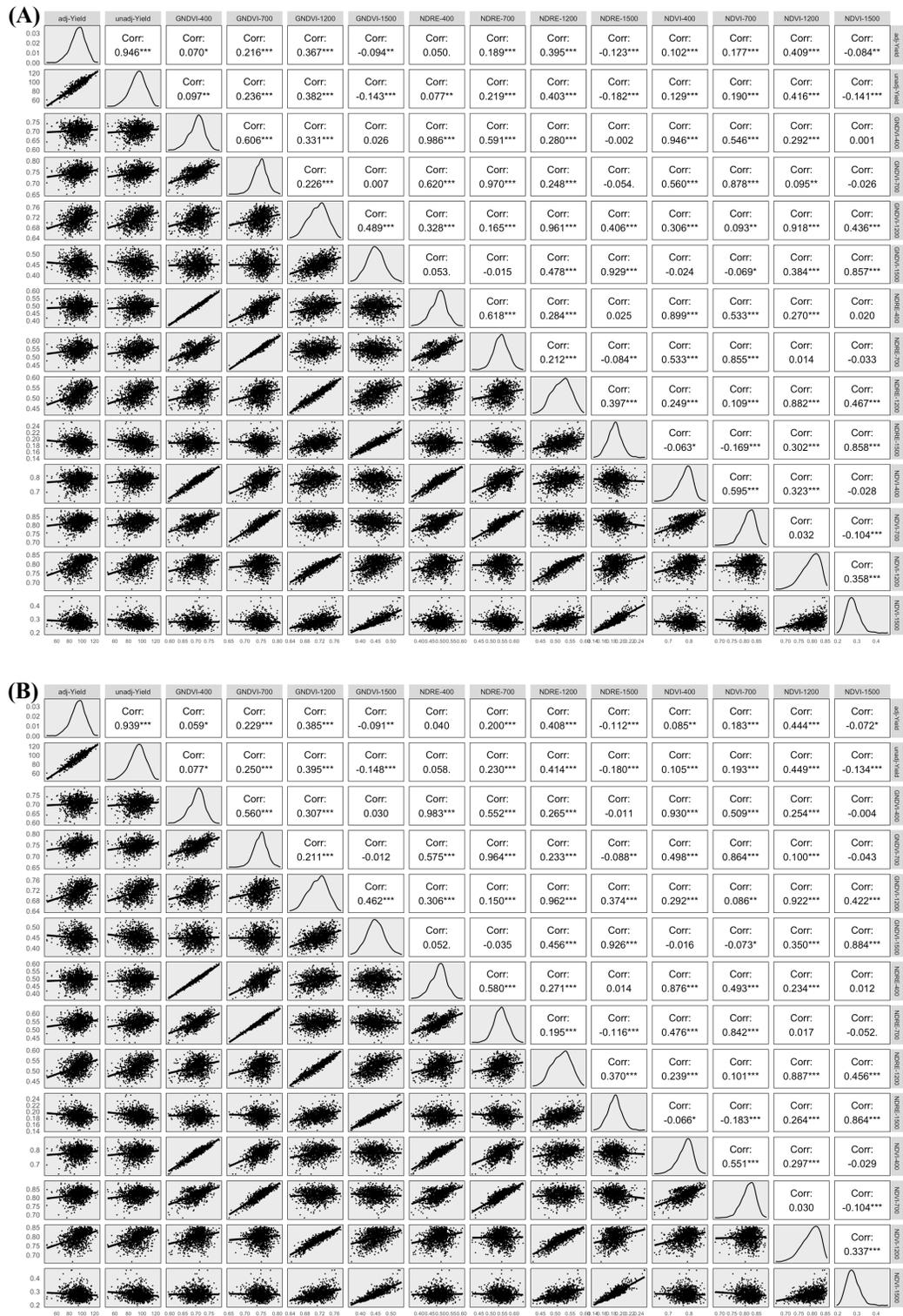
**Figure C.1** 2021 RN BLUP VI Pearson (A) and Spearman (B) rank correlation to GRYLD correlation.



**Figure C.2** 2021 RN BLUE VI Pearson (A) and Spearman (B) rank correlation to GRYLD correlation.



**Figure C.3** 2021 RN Raw VI Pearson (A) and Spearman (B) rank correlation to GRYLD correlation



**Figure C.4** 2021 RN Experimental Corrected VI Pearson (A) and Spearman (B) rank correlation to GRYLD correlation

**Table C.1** - Pearson's correlation of BLUP VI temporal points to grain yield by location

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.149	0.267	0.211	---	0.167	0.258	0.217	---	0.115	0.261	0.227	---
RN	2021	-0.223	0.112	0.284	0.145	-0.194	0.289	0.382	0.177	-0.250	-0.051	0.232	0.104
SA	2021	0.154	0.130	0.129	0.071	0.142	0.111	0.123	0.065	0.193	0.166	0.161	0.086
MP	2020	0.186	0.288	0.379	0.216	0.206	0.295	0.391	0.163	0.094	0.207	0.312	0.155
RN	2020	0.451	0.557	0.185	-0.124	0.459	0.573	0.272	-0.098	0.371	0.454	0.188	-0.225
SU	2020	0.405	0.402	0.260	0.039	0.409	0.402	0.289	-0.007	0.374	0.374	0.258	-0.035
WS	2020	0.349	0.489	0.238	-0.215	0.387	0.506	0.244	-0.176	0.364	0.421	0.223	-0.158
RL	2019	---	0.283	0.169	-0.003	---	0.289	0.155	-0.049	---	0.278	0.251	0.057
RN	2019	0.136	---	0.337	0.156	0.108	---	0.315	0.147	0.196	---	0.375	0.189
RP	2019	---	0.313	0.395	---	---	0.289	0.452	---	---	0.313	0.397	---
RN	2018	0.324	0.372	0.325	---	0.359	0.376	0.303	---	0.354	0.380	0.336	---
RP	2018	0.317	0.369	0.173	---	0.305	0.390	0.069	---	0.321	0.328	0.119	---
SA	2018	0.256	0.324	0.386	---	0.302	0.387	0.461	---	0.239	0.239	0.399	---
MP	2017	---	---	0.261	-0.012	---	---	0.288	0.015	---	---	0.130	0.078
RL	2017	---	---	0.050	-0.115	---	---	0.072	-0.151	---	---	0.081	-0.120
RN	2017	---	---	0.320	0.263	---	---	0.397	0.231	---	---	0.275	0.270
RP	2017	---	---	0.439	0.099	---	---	0.402	0.196	---	---	0.468	0.370

**Table C.2** Spearman's correlation of BLUP VI temporal points to grain yield by location.

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.149	0.260	0.161	---	0.166	0.252	0.161	---	0.122	0.262	0.199	---
RN	2021	-0.288	0.148	0.372	0.114	-0.239	0.266	0.426	0.119	-0.352	-0.019	0.338	0.107
SA	2021	0.135	0.160	0.146	0.080	0.129	0.159	0.146	0.060	0.186	0.162	0.141	0.084
MP	2020	0.211	0.314	0.327	0.155	0.222	0.314	0.331	0.070	0.101	0.233	0.262	0.080
RN	2020	0.415	0.533	0.221	-0.134	0.425	0.544	0.263	-0.087	0.254	0.447	0.216	-0.231
SU	2020	0.387	0.349	0.290	0.055	0.392	0.357	0.315	0.009	0.323	0.335	0.275	-0.005
WS	2020	0.316	0.409	0.196	-0.210	0.343	0.416	0.200	-0.176	0.324	0.349	0.180	-0.177
RL	2019	---	0.283	0.169	-0.003	---	0.289	0.155	-0.049	---	0.278	0.251	0.057
RN	2019	0.147	---	0.304	0.148	0.138	---	0.288	0.139	0.182	---	0.330	0.198
RP	2019	---	0.308	0.401	---	---	0.295	0.451	---	---	0.308	0.396	---
RN	2018	0.308	0.367	0.297	---	0.358	0.373	0.289	---	0.351	0.364	0.309	---
RP	2018	0.298	0.366	0.173	---	0.284	0.388	0.067	---	0.293	0.318	0.100	---
SA	2018	0.258	0.364	0.365	---	0.316	0.431	0.435	---	0.263	0.284	0.349	---
MP	2017	---	---	0.131	-0.040	---	---	0.176	0.005	---	---	0.010	-0.008
RL	2017	---	---	-0.082	-0.146	---	---	-0.063	-0.149	---	---	-0.075	-0.153
RN	2017	---	---	0.285	0.213	---	---	0.364	0.186	---	---	0.222	0.248
RP	2017	---	---	0.442	0.119	---	---	0.425	0.024	---	---	0.532	0.265

**Table C.3 - Pearson's correlation of BLUE VI temporal points to grain yield by location**

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.186	0.293	0.186	---	0.196	0.288	0.195	---	0.169	0.287	0.200	---
RN	2021	0.166	0.142	0.352	0.141	0.170	0.200	0.427	0.175	0.114	0.010	0.350	0.125
SA	2021	0.137	0.192	0.191	0.089	0.152	0.173	0.178	0.074	0.158	0.208	0.211	0.107
MP	2020	0.210	0.282	0.368	0.182	0.223	0.282	0.363	0.118	0.121	0.211	0.313	0.128
RN	2020	0.391	0.535	0.139	-0.129	0.394	0.558	0.250	-0.097	0.314	0.417	0.152	-0.222
SU	2020	0.362	0.401	0.321	0.102	0.362	0.442	0.344	0.009	0.358	0.398	0.334	-0.017
WS	2020	0.410	0.523	0.285	-0.234	0.446	0.540	0.235	-0.220	0.424	0.458	0.265	-0.188
RL	2019	---	0.269	0.163	-0.025	---	0.281	0.144	-0.049	---	0.270	0.265	0.045
RN	2019	0.219	---	0.402	0.196	0.188	---	0.383	0.179	0.259	---	0.417	0.241
RP	2019	---	0.294	0.448	---	---	0.289	0.494	---	---	0.301	0.449	---
RN	2018	0.361	0.387	0.369	---	0.374	0.374	0.342	---	0.387	0.418	0.370	---
RP	2018	0.378	0.494	0.207	---	0.362	0.524	0.085	---	0.368	0.463	0.099	---
SA	2018	0.295	0.387	0.432	0.338	0.338	0.463	0.515	---	0.281	0.300	0.446	---
MP	2017	---	---	0.233	-0.017	---	---	0.278	0.103	---	---	0.096	0.036
RL	2017	---	---	0.058	-0.108	---	---	0.081	-0.145	---	---	0.081	-0.114
RN	2017	---	---	0.261	0.253	---	---	0.360	0.229	---	---	0.214	0.270
RP	2017	---	---	0.417	0.098	---	---	0.396	0.196	---	---	0.438	0.357

**Table C.4 Spearman's correlation of BLUE VI temporal points to grain yield by location**

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.187	0.293	0.130	---	0.194	0.282	0.140	---	0.176	0.291	0.161	---
RN	2021	0.140	0.145	0.389	0.138	0.146	0.195	0.436	0.168	0.073	-0.009	0.402	0.150
SA	2021	0.142	0.177	0.169	0.086	0.176	0.174	0.167	0.068	0.165	0.180	0.168	0.090
MP	2020	0.239	0.315	0.332	0.143	0.238	0.312	0.331	0.063	0.134	0.242	0.262	0.069
RN	2020	0.358	0.514	0.178	-0.130	0.369	0.529	0.253	-0.085	0.189	0.386	0.166	-0.234
SU	2020	0.292	0.338	0.307	0.099	0.306	0.391	0.332	0.079	0.278	0.340	0.305	0.040
WS	2020	0.363	0.461	0.235	-0.237	0.390	0.467	0.201	-0.193	0.371	0.388	0.209	-0.175
RL	2019	---	0.265	0.181	0.044	---	0.298	0.150	0.026	---	0.211	0.265	0.095
RN	2019	0.198	---	0.360	0.175	0.178	---	0.340	0.158	0.239	---	0.372	0.237
RP	2019	---	0.299	0.450	---	---	0.298	0.490	---	---	0.304	0.451	---
RN	2018	0.318	0.356	0.349	---	0.351	0.350	0.332	---	0.365	0.373	0.360	---
RP	2018	0.361	0.505	0.203	---	0.346	0.537	0.083	---	0.348	0.459	0.086	---
SA	2018	0.311	0.442	0.392	---	0.352	0.505	0.488	---	0.313	0.347	0.390	---
MP	2017	---	---	0.103	-0.043	---	---	0.168	0.091	---	---	-0.050	-0.054
RL	2017	---	---	-0.076	-0.133	---	---	-0.060	-0.141	---	---	-0.073	-0.155
RN	2017	---	---	0.210	0.213	---	---	0.311	0.197	---	---	0.155	0.254
RP	2017	---	---	0.415	0.136	---	---	0.411	0.050	---	---	0.498	0.271

**Table C.5 - Pearson's correlation of Raw VI temporal points to grain yield by location**

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.243	0.326	0.426	---	0.253	0.317	0.437	---	0.217	0.330	0.442	---
RN	2021	0.102	0.111	0.411	0.153	0.098	0.151	0.455	0.156	0.081	0.012	0.413	0.118
SA	2021	0.189	0.236	0.228	0.125	0.182	0.198	0.206	0.105	0.195	0.244	0.239	0.149
MP	2020	0.267	0.294	0.794	0.528	0.307	0.241	0.806	0.370	0.228	0.223	0.775	0.520
RN	2020	0.083	0.503	0.302	0.007	0.054	0.478	0.357	0.146	0.075	0.264	0.364	0.059
SU	2020	0.368	0.432	0.400	-0.013	0.377	0.518	0.416	-0.157	0.350	0.418	0.409	-0.083
WS	2020	0.324	0.435	0.099	-0.102	0.356	0.459	0.056	-0.099	0.337	0.379	0.113	-0.042
RL	2019	---	0.639	0.512	0.510	---	0.671	0.557	0.446	---	0.599	0.654	0.563
RN	2019	0.234	---	0.504	0.356	0.177	---	0.493	0.331	0.272	---	0.516	0.406
RP	2019	---	0.364	0.482	---	---	0.367	0.520	---	---	0.359	0.468	---
RN	2018	0.356	0.372	0.511	---	0.385	0.379	0.515	---	0.411	0.418	0.519	---
RP	2018	0.322	0.521	0.414	---	0.339	0.559	0.362	---	0.341	0.488	0.408	---
SA	2018	0.431	0.385	0.594	---	0.499	0.435	0.661	---	0.358	0.235	0.548	---
MP	2017	---	---	0.013	-0.291	---	---	0.031	-0.270	---	---	-0.092	-0.001
RL	2017	---	---	0.011	-0.082	---	---	0.023	-0.144	---	---	-0.029	-0.133
RN	2017	---	---	0.291	0.240	---	---	0.397	0.216	---	---	0.214	0.259
RP	2017	---	---	0.164	0.015	---	---	0.225	0.019	---	---	0.203	0.262

**Table C.6 Spearman's correlation of Raw VI temporal points to grain yield by location**

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.246	0.327	0.320	---	0.252	0.316	0.339	---	0.217	0.345	0.339	---
RN	2021	0.054	0.082	0.426	0.154	0.057	0.128	0.466	0.145	0.014	-0.046	0.446	0.146
SA	2021	0.192	0.209	0.197	0.109	0.191	0.192	0.187	0.094	0.196	0.205	0.193	0.115
MP	2020	0.269	0.330	0.743	0.483	0.299	0.273	0.751	0.293	0.210	0.238	0.702	0.468
RN	2020	0.022	0.495	0.307	-0.022	0.028	0.469	0.301	0.076	-0.055	0.221	0.336	-0.020
SU	2020	0.344	0.401	0.047	-0.018	0.362	0.489	0.109	-0.126	0.306	0.383	0.108	-0.027
WS	2020	0.261	0.295	0.009	0.042	0.278	0.328	-0.027	-0.019	0.265	0.235	0.018	0.099
RL	2019	---	0.642	0.568	0.574	---	0.683	0.601	0.526	---	0.577	0.699	0.605
RN	2019	0.194	---	0.467	0.336	0.160	---	0.454	0.321	0.249	---	0.489	0.400
RP	2019	---	0.370	0.480	---	---	0.370	0.520	---	---	0.361	0.471	---
RN	2018	0.345	0.321	0.459	---	0.382	0.340	0.484	---	0.410	0.346	0.486	---
RP	2018	0.301	0.534	0.424	---	0.327	0.577	0.377	---	0.318	0.492	0.418	---
SA	2018	0.423	0.378	0.598	---	0.502	0.434	0.669	---	0.356	0.235	0.530	---
MP	2017	---	---	-0.153	-0.316	---	---	-0.118	-0.286	---	---	-0.262	-0.072
RL	2017	---	---	-0.043	-0.126	---	---	-0.041	-0.165	---	---	-0.153	-0.166
RN	2017	---	---	0.252	0.198	---	---	0.358	0.181	---	---	0.155	0.232
RP	2017	---	---	0.201	0.060	---	---	0.253	0.051	---	---	0.233	0.263

**Table C.7** - Pearson's correlation of experimental corrected VI temporal points to grain yield by location

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.237	0.305	0.421	---	0.247	0.300	0.431	---	0.223	0.317	0.443	---
RN	2021	0.070	0.216	0.367	-0.094	0.050	0.189	0.395	-0.123	0.102	0.177	0.409	-0.084
SA	2021	0.221	0.175	0.235	0.113	0.222	0.146	0.217	0.099	0.230	0.195	0.239	0.140
MP	2020	0.261	0.292	0.799	0.548	0.301	0.236	0.812	0.397	0.222	0.220	0.777	0.551
RN	2020	0.051	0.486	0.268	0.018	0.022	0.461	0.339	0.169	0.040	0.221	0.324	0.080
SU	2020	0.105	0.076	0.079	0.054	0.130	0.091	0.093	0.041	0.128	0.085	0.089	0.032
WS	2020	-0.310	-0.298	-0.307	-0.313	-0.298	-0.280	-0.298	-0.309	-0.297	-0.289	-0.304	-0.269
RL	2019	---	0.433	0.004	0.018	---	0.388	0.041	-0.035	---	0.475	0.323	0.208
RN	2019	0.273	---	0.533	0.438	0.202	---	0.530	0.429	0.304	---	0.544	0.493
RP	2019	---	0.378	0.501	---	---	0.389	0.540	---	---	0.371	0.485	---
RN	2018	-0.116	-0.140	-0.128	---	-0.049	-0.114	-0.080	---	0.015	-0.126	-0.081	---
RP	2018	-0.431	-0.393	-0.415	---	-0.419	-0.386	-0.367	---	-0.412	-0.393	-0.357	---
SA	2018	0.435	0.396	0.594	---	0.503	0.452	0.662	---	0.360	0.236	0.550	---
MP	2017	---	---	-0.113	-0.232	---	---	-0.087	-0.278	---	---	-0.225	0.040
RL	2017	---	---	0.136	0.042	---	---	0.124	-0.021	---	---	0.109	-0.012
RN	2017	---	---	0.266	0.235	---	---	0.386	0.208	---	---	0.182	0.259
RP	2017	---	---	0.002	-0.169	---	---	0.065	-0.182	---	---	0.076	0.131

**Table C.8** Spearman's correlation of experimental corrected VI temporal points to grain yield by location

Loc	Year	GNDVI-400	GNDVI-700	GNDVI-1200	GNDVI-1500	NDRE-400	NDRE-700	NDRE-1200	NDRE-1500	NDVI-400	NDVI-700	NDVI-1200	NDVI-1500
MP	2021	0.247	0.302	0.342	---	0.258	0.297	0.348	---	0.231	0.330	0.351	---
RN	2021	0.059	0.232	0.388	-0.087	0.041	0.202	0.411	-0.108	0.084	0.183	0.447	-0.068
SA	2021	0.222	0.162	0.206	0.106	0.228	0.140	0.193	0.091	0.227	0.176	0.201	0.119
MP	2020	0.260	0.332	0.760	0.500	0.286	0.267	0.777	0.328	0.195	0.234	0.710	0.508
RN	2020	0.001	0.476	0.273	-0.008	0.005	0.450	0.284	0.099	-0.087	0.166	0.277	0.003
SU	2020	0.347	0.426	0.131	0.000	0.367	0.515	0.214	-0.113	0.315	0.421	0.177	-0.014
WS	2020	0.285	0.318	0.026	-0.015	0.311	0.350	-0.019	0.023	0.301	0.268	0.035	0.112
RL	2019	---	0.382	-0.076	0.004	---	0.356	-0.003	-0.040	---	0.472	0.252	0.153
RN	2019	0.264	---	0.512	0.425	0.197	---	0.507	0.425	0.297	---	0.532	0.491
RP	2019	---	0.379	0.494	---	---	0.387	0.532	---	---	0.369	0.484	---
RN	2018	0.165	0.152	0.213	---	0.192	0.166	0.221	---	0.217	0.168	0.222	---
RP	2018	-0.243	-0.064	-0.132	---	-0.225	-0.059	-0.184	---	-0.231	-0.088	-0.154	---
SA	2018	0.434	0.403	0.596	---	0.508	0.452	0.671	---	0.371	0.238	0.534	---
MP	2017	---	---	-0.149	-0.259	---	---	-0.130	-0.305	---	---	-0.267	-0.025
RL	2017	---	---	0.043	-0.049	---	---	0.034	-0.118	---	---	0.029	-0.095
RN	2017	---	---	0.225	0.201	---	---	0.346	0.179	---	---	0.119	0.249
RP	2017	---	---	0.002	-0.191	---	---	0.070	-0.176	---	---	0.079	0.110