

Enhancing management strategies for crop yield improvement via advanced
data-informed decisions

by

Emmanuela van Versendaal Pirez

B.S., Universidad de la República, 2023

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

2025

Abstract

Generating multi-scale information to understand crop variability and optimize crop management is crucial for increasing yields and improving input efficiency to meet global food demand. This dissertation uses available technological tools to generate actionable insights and improve our understanding of crop behavior at different scales.

This dissertation is divided into five chapters. The first chapter serves as an introduction to crop yield improvement strategies and describes technologies that facilitate the generation of knowledge. The remaining chapters present case studies focusing on yield improvement, testing different crop management practices and using data approaches. Chapter 2 uses crop modelling to explore different combinations of soybean planting dates and maturity groups at the regional level in Kansas, USA, under both current and future weather scenarios. Chapter 3 uses the Canopeo tool, crop modelling and Bayesian analysis to assess the effects of equidistant and non-equidistant soybean planting patterns at different site densities in three US states. Chapter 4 uses multiple years of geo-referenced yield monitor data from GPS-enabled combines and spatio-temporal analysis to establish a data pipeline that quantifies within-field yield variation in response to precipitation. Finally, Chapter 5 summarizes the key findings and outlines directions for future research to further improve crop management information and yield outcomes, a critical step in meeting the world's growing food demand.

Enhancing management strategies for crop yield improvement via advanced
data-informed decisions

by

Emmanuela van Versendaal Pirez

B.S., Universidad de la República, 2023

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

2025

Approved by:

Major Professor
Dr. Ignacio A. Ciampitti

Copyright

© Emmanuela van Versendaal Pirez 2025.

Abstract

Generating multi-scale information to understand crop variability and optimize crop management is crucial for increasing yields and improving input efficiency to meet global food demand. This dissertation uses available technological tools to generate actionable insights and improve our understanding of crop behavior at different scales.

This dissertation is divided into five chapters. The first chapter serves as an introduction to crop yield improvement strategies and describes technologies that facilitate the generation of knowledge. The remaining chapters present case studies focusing on yield improvement, testing different crop management practices and using data approaches. Chapter 2 uses crop modelling to explore different combinations of soybean planting dates and maturity groups at the regional level in Kansas, USA, under both current and future weather scenarios. Chapter 3 uses the Canopeo tool, crop modelling and Bayesian analysis to assess the effects of equidistant and non-equidistant soybean planting patterns at different site densities in three US states. Chapter 4 uses multiple years of geo-referenced yield monitor data from GPS-enabled combines and spatio-temporal analysis to establish a data pipeline that quantifies within-field yield variation in response to precipitation. Finally, Chapter 5 summarizes the key findings and outlines directions for future research to further improve crop management information and yield outcomes, a critical step in meeting the world's growing food demand.

Table of Contents

List of Figures	ix
List of Tables	xiv
Acknowledgements	xvi
Dedication	xvii
1 General introduction	1
1.1 Core principles and strategies for crop yield improvement	1
1.2 Technologies that facilitate the generation of management insights	2
1.3 Dissertation objectives	4
2 Integrating field data and a modeling approach to inform optimum planting date × maturity group for soybeans under current and future weather conditions in Kansas	6
2.1 Introduction	7
2.2 Materials and methods	9
2.2.1 Data collection	9
2.2.2 Spatial clustering	10
2.2.3 Model validation	11
2.2.4 Developing scenarios for yield stability for the current and future weather	12
2.3 Results	13
2.3.1 Spatial clustering	13
2.3.2 Model validation	13
2.3.3 Developing scenarios for yield stability for the current weather	14

2.3.4	Developing scenarios for yield stability for the future weather	15
2.4	Discussion	16
2.5	Conclusion	18
3	Soybean yield and seed quality in equidistant versus non-equidistant plant arrangements under different densities	24
3.1	Introduction	25
3.2	Materials and methods	27
3.2.1	Field experiment	27
3.2.2	Data collection	28
3.2.3	Data analysis	29
3.3	Results	32
3.3.1	Plant arrangement effect on soybean seed yield	32
3.3.2	Plant arrangement effect on seed quality	33
3.3.3	Canopy coverage	34
3.4	Discussion	34
3.5	Conclusion	37
4	Spatio-temporal yield variation and precipitation within a field	48
4.1	Introduction	49
4.2	Materials and methods	51
4.2.1	Data collection and processing	51
4.2.2	Statistical analysis	52
4.3	Results	54
4.3.1	Model evaluation	54
4.3.2	Spatially varying response parameters	55
4.3.3	Classify type of yield response to cumulative precipitation	56
4.4	Discussion	57

4.5 Conclusion	59
5 Final remarks	65
Bibliography	68
A Appendix A, Chapter 2	95
B Appendix B, Chapter 3	107
C Appendix C, Chapter 4	112

List of Figures

2.1	Kansas map with the three clusters (represented by triangles with distinct colors) for crop seasonal cumulative annual precipitation (mm), and the T42 GCM grid resolution (represented by solid lines) reported by Brunsell et al. (2010)	21
2.2	Observed versus simulated seed yield for the three clusters (North-central, North-east, and South-east) with soybean maturity groups III, IV, V, and VI (represented by distinct colors). The dashed line represents the 1:1 line. RMSE, root mean square error; RRMSE, relative root mean square error; PLP, percentage lack of precision; PLA, percentage lack of accuracy; KGE, Kling-Gupta model efficiency.	21
2.3	Simulated seed yield for the current weather (2011 to 2021), where darker colors represent higher yields and the size of the bubbles represent the standard deviation across years, within each cluster, for the different combinations of planting date and maturity groups. Planting dates are referred to as day-month.	22
2.4	Average difference between future and current weather yields (bars, difference between the year ranges 2042-2052 vs 2011-2021) and seed yield under future weather scenarios (points, from the year 2042 to 2052) for the different combinations of planting dates and maturity groups within the three clusters in Kansas. Vertical lines represent standard deviation of the yield difference across different sites within clusters (lines over the bars) and across years and different sites in each cluster (lines over the points). Planting dates are referred to as day-month.	23

3.1	Illustrative comparison of pictures taken with a GoPro Hero5 Black portable digital camera on each plant spatial arrangement (from top to bottom) showing the original image (left), and after processing by Canopeo app (right). . .	39
3.2	Seed yield for equidistant versus non-equidistant soybean plant spatial arrangements under different densities. Dashed lines correspond to 95% credible intervals. Black solid line corresponds to 1:1 line. Colors denote equidistant treatments and shapes represents each trial. Vertical and horizontal error bars show the 95% credible intervals from the pairwise comparisons of treatments. Equidistant treatment: 1) 13 x 13, 13 cm by 13 cm; 2) 15x15, 15 cm by 15 cm; 3) 20x20, 20 cm by 20 cm. Non-equidistant treatment: 38 cm of row space.	40
3.3	Soybean canopy coverage (CC) over thermal time for the entire growing season, presented for trials with complete seasonal data (each panel): Kansas 2021 in Irrigated and dryland conditions (KS21 Irrigated and KS21 Dryland); Kansas 2022 in an early, mid, and late-planting date (KS22 Early, KS22 Mid, and KS22 Late); and Mississippi 2022 in early-planting date. Solid lines correspond to the estimated CC from the bi-hill model. Dashed lines correspond to bi-hill parameters. Points refers to observed data. Square shapes denote moments of maximum growth rate (growth timing) and maximum decay rate (decay timing). Max CC refers to maximum CC. Light-gray and dark-gray areas correspond to the R1 to R3 and R3 to R5 phenological stages, respectively, as determined using DSSAT.	41

4.1	Model performance across case studies (cases 1, 3 for maize crop, and 2, 4 for soybean). Each panel represents a different case study, displaying yield data at specific spatial coordinates within the field for a particular year. Points are color-coded to uniquely identify each spatial coordinate. The bottom-right corner of each panel includes a field map, where colors correspond to those of the points, illustrating the spatial distribution (within field variation) of the yield data.	61
4.2	Maps of the spatially varying response parameters to optimal precipitation across fields. Columns represent each parameter of the model. Rows represent each case study. In all the parameters, darker the color higher the parameter's value. CP: cumulative precipitation. Case study 1 and 3 correspond to maize crop; Case study 2 and 4 correspond to soybean crop.	62
4.3	Curves of response of yield to cumulative precipitation, each column correspond to a group (left) that were identified in the classification and regression analyses. The response variable for this classification was the cumulative precipitation that maximizes yield. Field map demonstrating spatial distribution of each group (right) Each row corresponds to each case study. Each color represents a group. Solid black line corresponds to the mean yield. Dashed black lines correspond to the 95% confidence intervals.	63
A.1	Distribution of planting dates in the 13 sites of the dataset from 2014 to 2021.	98
A.2	Interval where setting pods occurs in the state of Kansas according to NASS-USDA (https://quickstats.nass.usda.gov/results/20E59F49-96CD-3192-AEF5-6B5B63E28C43 , visited April 2022)	99
A.3	Comparison of days of the year where anthesis occurs between the simulated data and data obtained from the National Agricultural Statistics Service (https://quickstats.nass.usda.gov/results/21B7C2DA-A180-3470-A7F4-383E92588748 , visited April 2022)	100

A.4	Simulated seed yield ($kg\ ha^{-1}$) across the 30 years of study within the three clusters in Kansas (North-central, North-east, and South-east).	102
B.1	DSSAT crop growth model performance predicting R1, R3, and R5 phenological stages in the KS22 Early, Mid, and late trials. Dashed line corresponds to 1:1 line.	108
B.2	Soybean seed yield and plant arrangement by trials. Bars represent 95% credible intervals. Points represent the average soybean seed yield. The colors of the points and bars indicate each plant arrangement. Panels represent each trial. KS21 Irrigated, MS22 Late, and SD22 Mid had differences.	109
B.3	Canopy coverage parameters distribution (from the hill function) by plant arrangement across trials (ID). BlueSky color under the curve represents the 95% confidence interval. Max Canopy (%) is the maximum canopy coverage.	110
B.4	Canopy coverage parameters distribution (from the bihill function) by plant arrangement across trials (ID). BlueSky color under the curve represents the 95% confidence interval. Max Canopy (%) is the maximum canopy coverage.	111
C.1	Elevation and Compound topographic Index (CTI) for each Farm. Left column panel represent the elevation and right column panel represent the CTI. Each row corresponds to each Farm. Farm 1 correspond to case study 1 and 2, Farm 2 is the case study 3, and Farm 3 correspond to case study 4.	113
C.2	Selection of optimum cumulative precipitation in Case study 1 and 3 (Maize). Each column of panels represents a different date of starting moment. Top panels are from case study 1 and panels in the bottom are from case study 3. Blue points represent the cumulative precipitation period with the lowest AIC.	114

C.3	Selection of optimum cumulative precipitation in Case study 2 and 4 (Soybean). Each column of panels represents a different date of starting moment. Top panels above are from case study 2 and panels in the bottom are from case study 4. Blue point represent the cumulative precipitation period with the lowest AIC.	114
C.4	Field map of case study 1 with two contrasting examples (above). Predicted yield and its derivative in response to cumulative precipitation in two contrasting examples in case study 1. Above panel represent the 1st derivative and below panel represent the predicted yield (Below). Colors represent the two contrasting examples within the field. CP: Cumulative precipitation. . .	115
C.5	Regression tree for case study 1 (A), case study 2 (B), case study 3 (C), and case study 4 (D) to classify precipitation value across a field using spatial growth and decay coefficients. Figures in the left correspond to maize fields case studies and figures in the right correspond to soybean fields case studies.	116

List of Tables

2.1	Measure of agreement between observed and simulated data, for the entire study area. RMSE, root mean square error; RRMSE, relative root mean square error; PLP, percentage lack of precision; PLA, percentage lack of accuracy; KGE, Kling-Gupta model efficiency.	20
3.1	Site-years characterization showing year, state, location, planting dates, maturity groups (MG), soybean variety, water management, number of replications (Rep.), and data collected on canopy coverage (CC).	38
3.2	Average soybean seed number, thousand seed weight, and their respective quantile 2.5% and 97.5% by plant arrangement within each trial (ID). Equidistant treatment: 1) 13x13, 13 cm by 13 cm; 2) 15x15, 15 cm by 15 cm; 3) 20x20, 20 cm by 20 cm. Non-equidistant treatment: 38Row, 38 cm of row space. . .	42
3.3	Average soybean seed protein and oil concentration (%), and their respective quantile 2.5% and 97.5% by plant spatial arrangement within each trial (ID). Equidistant treatment: 1) 13 x 13, 13 cm by 13 cm; 2) 15x15, 15 cm by 15 cm; 3) 20x20, 20 cm by 20 cm. Non-equidistant treatment: 38Row, 38 cm of row space.	45
4.1	Summary of Spatially Varying Parameters by group and case study. Includes mean and interquartile ranges ($IQR_{25-75\%}$) for Maximum yield (kg ha^{-1}); Cumulative Precipitation (CP) maximize yield (mm); Growth rate ($\text{kg ha}^{-1} \text{mm}^{-1}$); and Decay rate ($\text{kg ha}^{-1} \text{mm}^{-1}$).	64

A.1	Data sources. KSU performance test (https://www.agronomy.k-state.edu/outreach-and-services/crop-performance-tests/soybeans/ , accessed 2022/03); Ciampitti et al., 2017 (https://doi.org/10.4148/2378-5977.7434); Ciampitti, et al., 2020 (figshare: 10.6084/m9.figshare.20018015).	96
A.2	Standard planting date, soil characteristics, and weather variables used to create the clusters.	97
A.3	Parameters included in the generic genotypes used from APSIM Next generation.	97
A.4	Soybean seed yield (kg ha^{-1}) for the different planting dates $\times$ maturity groups combination in the South-east cluster. Values within different letters are significantly different at $P < 0.05$.	101
A.5	ANOVA (Type III Wald chi square tests) of the current weather (2012–2021) for the three clusters within Kansas.	101
A.6	Climate change in the future (2043–2052) compared to the current weather (2012–2021) in the three clusters in Kansas.	102
A.7	Soybean seed yield (kg ha^{-1}) for the different planting dates $\times$ maturity groups combination within the three clusters in Kansas. Values within different letters are significantly different at $P < 0.05$.	103
B.1	The mean of temperature and cumulative rainfall from May to October across trials.	107

Acknowledgments

First, I would like to thank my advisor during this process, Dr. *Ignacio A. Ciampitti*, for encouraging me to pursue a PhD and for supporting me at Kansas State University, which has helped me grow both personally and professionally. I also appreciate the valuable contributions and engaging discussions from my committee members, Drs. *P. V. Vara Prasad*, *Peter Kyveryga*, and *Trevor Helfey*.

I extend my gratitude to everyone in the Ciampitti Lab from 2021 to 2025 for your support, help, and friendship.

Finally, I want to give special thanks to my family, my parents, *Mariela* and *Henk*; my brother, *Ernesto*; my sister-in-law, *Cecilia*; and my three nieces/nephews, *Isabella*, *Olivia*, and *Francisco*. I also want to thank my friends in Uruguay for always being there for me, and my friends at MHK - Coni, Nacho, Jaz, Lucas, Caro, Fran, Martin R., Feli, and others - for making life abroad easier.

Dedication

To my parents, *Mariela* and *Henk*.

A mis padres, *Mariela* y *Henk*.

Chapter 1

General introduction

1.1 Core principles and strategies for crop yield improvement

Population growth is driving a significant increase in global food demand, from 35% in 2010 to an expected 56% increase by 2050 ([Van Dijk et al., 2021](#)). In practical terms, even a conservative estimate of a 50% increase in food demand over the next 30 years (2020-2050) implies that yield of major food crops must be accelerated by approximately 30% on the existing farmland ([Cassman & Grassini, 2020](#)). At the same time, rising land prices and production costs are putting pressure on farmers to adopt more efficient management strategies ([Tomal & Gumieniak, 2020](#)).

Efficient management practices, such as variety selection, optimizing plant density, adjusting plant arrangement, and calibrating fertilizer rates, are crucial not only for increasing yield, but also for reducing costs and mitigating risks. A clear understanding of yield variability is essential for the correct application of these practices ([Fischer, 2015](#)). Therefore, it is important to generate new management insights to improve crop management practices. These insights should be developed at three levels: the regional level, which captures variability across different production areas ([Hemathilake & Gunathilake, 2022](#)); the site level, which addresses site-specific management differences ([Abdala et al., 2025](#)); and the within-

field scale level, where fine-scale variability can be managed for optimal performance (Plant, 2001).

Crop management insights must be developed at multiple scale because each scale directly influences farm management decisions and ultimately impacts farmer’s profitability (Hein et al., 2006). At the regional level, climatic factors, such as variations in day length, seasonal precipitation, and temperature, determine, for example, the performance of different soybean genotypes. These differences mean that a soybean maturity group well-suited to one region may under-perform in another (Mourtzinis & Conley, 2017), guiding farmers in selecting the most appropriate varieties for their area. At the site level, even within a single region where climatic conditions are similar, a significant variation in soil type and fertility affects the growing environment, thereby influencing variety selection and plant density (Abdala et al., 2025; Bastos et al., 2020). Finally, within-field variability, driven by factors such as field topography and local soil properties, creates distinct management zones (Basso et al., 2009). Adapting management practices, such as modifying fertilizer rates, plant density and variety selection, to these micro-environments can maximize in-field yields while minimizing costs and risks (Cook & Bramley, 1998).

1.2 Technologies that facilitate the generation of management insights

Technologies that can be used to improve crop management (e.g., remote sensing and crop modeling) have been available for over fifty years (X. Jin et al., 2018). These tools generate valuable data that enhances our understanding of crop yield variability across different spatial scales. Without these technologies, gaining comparable insights would require numerous experiments over many years in different areas and environments (Mitra et al., 2022, 2024; Roy & Aslekar, 2022; Vitali et al., 2021; Yandun Narvaez et al., 2017).

One of the key technological tools for generating valuable data in crop management is crop growth simulation modelling. Developed in the 1960s (X. Jin et al., 2018), these models

were initially designed to identify and quantify factors that limit crop yield. The underlying hypothesis was that by measuring key variables, the system could be built numerically, allowing for a deeper understanding of crop growth and environmental interactions. Crop modelling is particularly useful for assessing the impact of management strategies and policies on regional yields, especially when observed field data are limited ([Oteng-Darko et al., 2013](#); [Sinclair & Seligman, 1996](#)). Furthermore, crop models facilitate large-scale scenario analysis, enabling researches to predict how different management decisions, such as planting dates, fertilizer rates, and irrigation strategies, affects yield outcome under varying environmental conditions ([Asseng et al., 2013](#); [Boote et al., 1998](#)). These models also help quantify yield variability by integrating climate, soil, and management data, making them a crucial tool for improving site-specific recommendations and optimizing resource use in modern agriculture ([Basso et al., 2001](#); [White et al., 2011](#)).

In the early 1980s, Geographic Information System (GIS) software emerged as a powerful tool for automating, managing, and analyzing spatial data ([Franzen & Mulla, 2015](#)). When integrated with GPS-enabled planters and combines, this technology enables the collection of detailed, geo-referenced observations across fields with varying conditions such as landscape and soil texture ([Fulton et al., 2018](#); [He et al., 2021](#)). This integration provides granular insight into within-field yield variability and generates actionable information for targeted management practices, including variable seeding rates, tailored fertilizer applications, variety selection and precise irrigation scheduling, ultimately optimizing crop performance and reducing input costs ([Lowenberg-DeBoer & Erickson, 2019](#); [Plant, 2001](#)).

In addition, a number of emerging tools are further improving our understanding of crop variability and optimizing management practices. For example, remote sensing and tool applications (e.g., Canopeo, [Patrignani & Ochsner, 2015](#)) enable monitoring of crop canopy evolution during the growing season, allowing early detection of stress and accurate identification of peak biomass or canopy closure ([Hatfield et al., 2008](#); [Patrignani & Ochsner, 2015](#)). Moreover, gridded weather data models provide high-resolution climate data in areas where traditional weather stations are scarce ([Mourtzinis et al., 2017](#)), while other tools deliver detailed information on topography and soil types without the need for extensive

field measurements ([Arrouays et al., 2014](#); [Höfle & Rutzinger, 2011](#); [Wichmann et al., 2015](#)). In addition, advanced statistical models and machine learning are now efficiently analyzing large agricultural datasets to provide accurate, actionable insights that support targeted management strategies and drive improvements in crop productivity ([Coble et al., 2018](#); [Cravero et al., 2022](#); [Tantalaki et al., 2019](#)).

1.3 Dissertation objectives

Based on the critical need to enhance yield and extract meaningful, multi-scale information for more efficient crop management, and take advantage of available technological tools, the general objective of this dissertation is to improve crop management strategies to increase yield and reduce variability at different scales by making data-driven decisions in advance. This will be achieved by using crop simulation models, GPS-enabled combine yield monitoring data and advanced statistical modeling techniques to generate actionable insights that support targeted management practices at regional, site and within-field scales.

Chapter [2](#) demonstrates the optimum combination of planting dates by maturity group on soybean at regional level in Kansas, United States, considering both the current weather and future weather scenarios. This chapter aims to i) investigate via utilization of a crop growth model the best combination for planting time x maturity group within a spatial framework, and ii) quantify the effect of future weather (30 years) on seed yield relative to the benchmark soybean production (current weather) under different scenarios in the central-eastern region of Kansas, USA.

Chapter [3](#) evaluates the effects of equidistant versus non-equidistant soybean planting arrangement under different densities at the site level, across different locations in the United States. This chapter aims to i) evaluate equidistant versus non-equidistant plant spatial arrangement to assess its effect on soybean seed yield and seed quality (protein and oil concentrations), and (ii) explore plant canopy cover development throughout the season using the Canopeo tool to understand its contributions to yield and seed quality.

Chapter [4](#) focuses on the within-field scale by assessing spatio-temporal variability of

crop yield in response to rainfall. The aim of this chapter is to establish a data pipeline to quantify spatio-temporal variation in within-field crop yield response to precipitation, using multi-year precipitation records.

Chapter 2

Integrating field data and a modeling approach to inform optimum planting date $\times$ maturity group for soybeans under current and future weather conditions in Kansas

* Published in *Sustainability*

van Versendaal, E., Carcedo, A.J.P., Adey, E., Sassenrath, G., Dooley, S., Lingenfelser, J., & Ciampitti, I.A. (2023) Integrating Field Data and a Modeling Approach to Inform Optimum Planting Date $\times$ Maturity Group for Soybeans under Current and Future Weather Conditions in Kansas. *Sustainability*, 15 (2), 1081.

<https://doi.org/10.3390/su15021081>

Abstract

Optimizing planting date by maturity group ($PD \times MG$) is critical to increase productivity and reduce production risks. Understanding the effect of management, not only under current, but also future weather conditions, is even more relevant for developing effective mitigation strategies. This paper provides an analysis of the optimum combinations of soybean $PD \times MG$ management in the central-eastern region of Kansas (United States) for both current and future weather conditions. Three geographical clusters illustrating the main environmental and management characteristics were defined within the central-eastern region of Kansas. The Agricultural Production Systems Simulator platform was employed to explore $PD \times MG$ combinations (PD from mid-April to mid-July; MG from III to VI) comparing current (2011–2021) and future (2042–2052) weather conditions. Overall, early planting dates produce greater yields, but reduce their stability over time (with a 15% increase in yield variation relative to late planting) across the clusters. Late planting dates resulted in a reduction close to 27% for soybean yields relative to those obtained by planting at early dates under current weather conditions. Furthermore, longer maturity groups (IV, V, and VI) resulted in a reduced yield penalty when planting time was delayed under the current weather conditions. However, this combination did not always represent the strategy that maximized yields.

2.1 Introduction

Soybean (*Glycine max* L. (Merr.)) is an important crop due to its contribution as a source of protein and oil (FAOSTAT, 2022). Improvements in crop management increased soybean yields (de Borja Reis et al., 2020) in concert with the increasing demand in global food production. Optimization of management technologies is still a challenge for many growers, with a high degree of uncertainty in the process of selecting the main factors influencing crop production (Carciochi et al., 2020; Correndo et al., 2022a). Less information is available for the western region of the main producing soybean area in the United States (US). Within

this region, Kansas ranked 9th as a major producer of soybeans in the US, highlighting the importance of providing critical information on management technologies for increased production.

Soybean seed yield is impacted by production factors, including selection of genotypes, planting date, and their interaction with the growing environment (soils and weather conditions). Although the environment is often the limiting factor determining yield, agronomic decisions such as modification of planting date and selection of cultivars are important components that producers can manage (Egli & Cornelius, 2009). The soybean planting window in Kansas is roughly 3 months from early April to early July, with a quantifiable reduction on maximum yields with delays in planting of 20 kg/ha/day (Ciampitti et al., 2022). Cultivars in maturity groups III and IV are the most grown throughout central and east Kansas, with longer maturity groups V and VI more commonly grown in the southeast region. Previous research has identified some of the impacts of variety selection and planting date on performance (Boyer et al., 2015; Egli & Bruening, 2000; Salmeron et al., 2014); however, the resultant recommendations were limited due to a global analysis (not site- or regional-specific) and available field data (e.g., variety trials with lack of range on planting dates and varieties). Thus, a formal new analysis is relevant to guide the process of selecting these key factors.

To provide actionable decisions based on limited field research, implementation of a crop growth model can help guide the selection of the best scenario for planting date and maturity group for soybeans under changing weather conditions. The selection of the best management practices should not only consider a few experimental years of study but consider the effects of longer trends in weather variation and their effects for optimizing management practices for increasing seed yields. This approach can be explored using the Agricultural Production Systems sIMulator – APSIM (Holzworth et al., 2014). In addition to exploring the effects of current weather to provide timely outcomes for growers, the effect of future weather conditions is a relevant research question to understand impacts on seed yield and guide future research investments. A recent study investigated the relevance of soybean growth models for climate change impact assessments, highlighting the changes in model performance

based on model utilization, assumptions, and approach to evaluate future weather. Different crop growth models highlight the impact of future weather conditions negatively impacting soybean yield, by increases in temperature and changes in precipitation patterns (Z. Jin et al., 2018; Schaubberger et al., 2017; Zabel et al., 2021). Although these predictive models may have limitations, implementation of these models can assist in targeting key plant traits from a breeding perspective and focus resources on improved management to ameliorate the potential implications on soy-bean seed yield from future weather conditions.

The introduction of modern soybean genotypes requires providing more up-to-date information on the current optimum planting dates and maturity groups across the most prominent soybean producing region in Kansas. The optimal planting times for these modern cultivars under varying soils are not well known for the current weather conditions. Following this rationale, the main aims of this study are to i) investigate via utilization of a crop growth model the best combination for planting time x maturity group within a spatial framework, and ii) quantify the effect of future weather (30 years) on seed yield relative to the benchmark soybean production (current weather) under different scenarios for planting date x maturity group in the central-east region of Kansas, US.

2.2 Materials and methods

2.2.1 Data collection

Soybean seed yield data were gathered from three sources across 13 sites in central and eastern Kansas, United States from 2014 to 2021 (Figure 2.1, Table A.1). The final dataset encompassed a planting window from April 6 to July 11, four maturity groups (III, IV, V, and VI), plant density from 25 to 38 plant m^{-2} , and harvest dates from October 8 to November 29.

For each site, the current standard planting date was defined as the most frequent planting date (statistical model) of the dataset across years (Figure A.1). Then, early planting dates were defined as those before the standard planting date, and the late dates are the plantings

carried out after this time.

Using the coordinates of each site (Table A.1), soil data were gathered from USDA-SSURGO (<https://www.nrcs.usda.gov/>, visited 10 March 2022) up to the 200 cm depth, and the soil parameters were calculated using the function within the APSIM next generation package (Miguez, 2022). Furthermore, long-term weather records (1984–2021) were obtained from Kansas Mesonet (<https://mesonet.k-state.edu/weather/historical/>, visited: 25 February 2022), using the closest weather station to each site (Table A.1, maximum distance 50 km). The variables included maximum and minimum temperatures, precipitation, and solar radiation, in daily time steps. For the Riley and Belleville sites, solar radiation data were retrieved from NASA POWER (<https://power.larc.nasa.gov>, visited: 1 March 2022). The whole dataset was employed for the spatial clustering (Section 2.2.2) and the future weather data generation (Section 2.2.3), and a subset (2011–2021) was employed to represent the current weather data (Section 2.2.4).

2.2.2 Spatial clustering

The sites (Table A.1) were grouped into regions with similar weather, soil, and management characteristics using the spatial Fuzzy c-Means (FCM) clustering algorithm (Bezdek et al., 1984) with “Geocmeans” package (Gelb & Apparicio, 2021) in R (R Core Team, 2021). The variables included (Table A.2) in this analysis were: i) management (standard planting date), ii) main soil characteristics (texture as percentage of sand, silt, and clay), and iii) weather (maximum temperature and cumulative rainfall).

Soil texture was included due to its relevance to define crop available water, and its relative invariability in time (Gijsman et al., 2002; Saxton & Rawls, 2006). The maximum temperature variable included for the clustering process was obtained as the daily maximum temperature averaged by month considering primarily the months of July, August, and September as a relevant period for yield formation (Figure A.2). This period includes the crop growth stages of pod formation (Fehr & Caviness, 1977), highly sensitive to heat stress (Kantolic & Slafer, 2001). Cumulative rainfall for the soybean growing season (May

to October) was calculated based on the observed current growing season in the database. Maximum temperature and cumulative rain were selected to be included in this analysis due to their impact on yield and relevance when defining future weather scenarios (Brunsell et al., 2010). Data sources for these variables were listed in 2.2.1.

2.2.3 Model validation

The Agricultural Production Systems Simulator (APSIM) Next-Generation is a modular modeling system that has been used in many applications, including farming systems design and the assessment of climate forecasting (Holzworth et al., 2018). APSIM generic genotypes for soybeans (Table A.3, Archontoulis et al., 2014) included in the APSIM-Soybean model (Brown et al., 2014, <https://builds.apsim.info/api/nextgen/docs/Soybean.pdf>, visited: September 2022) were used to perform simulations to compare with observed grain yield from 66 site-years, testing different planting dates and maturity groups (Table A.1). Due to the lack of capability of APSIM Next-Generation to represent the early termination of the crop cycle due to frost events, the crop growing season was terminated at the observed harvest day, when this date was available (Table A.1). Although the lack of crop phenology data provided a limitation, the simulated flowering dates agreed with those expected for the region (Figure A.3, <https://quickstats.nass.usda.gov/>, visited: 10 April 2022), denoting the satisfactory performance of the generic cultivars from APSIM (Archontoulis et al., 2014). Soil and weather data sources are listed in Table A.2.

The performance of the model was evaluated for the whole dataset as a first step, and then individually for each region defined in section 2.2.2, using root mean square error (RMSE), relative root mean square error (RRMSE), percentage lack of precision (PLP), percentage lack of accuracy (PLA), and Kling-Gupta model efficiency (KGE, Kling et al., 2012). The Metrica package was used for analysis (Correndo et al., 2022b) in R (R Core Team, 2021).

2.2.4 Developing scenarios for yield stability for the current and future weather

Future weather data (2022-2052) were generated using long-term weather records (1984-2021) for each site to define the historical seasonality, via the `decompose` function with `stats` package, in R (R Core Team, 2021). Future seasonal trends of temperature and precipitation were subtracted from the extensive study done by Brunsell et al. (2010). Briefly, this study examined the seasonal trends in air temperature and precipitation patterns in a decadal manner as the average monthly output of 21 global climate models under the Special Report on Emissions Scenarios A1B scenario used in the IPCC fourth assessment report (AR4) for six grid cells representing Kansas. These grid cells correspond to the T42 GCM grid resolution (Figure 2.1). Overall, this approach demonstrated an excellent agreement between multimodel ensemble mean output and observations for temperature ($r^2 = 0.99$, RMSE=0.48 - 1.48 °C) and a good agreement for precipitation (r^2 between 0.64 and 0.89, RMSE = 322 - 1144 mm). However, one limitation is the lack of consideration of the effect of the CO² increment. Each site was treated independently according to where the grid cell was located (Brunsell et al., 2010). One limitation of this approach is the lack of consideration of the effect of the CO² increment.

The model was set to plant soybean crop on April-15, May-1, May-15, June-1, June-15, July-1, and July-15 in each site from 1984 to 2021 with the current weather and simulated future weather from 2022 to 2052. The model included four maturity groups using the generic genotypes for soybean from APSIM (Archontoulis et al., 2014, Table A.3, these cultivars can be found in the APSIM platform as: ‘Generic_MG3’, ‘Generic_MG4’, ‘Generic_MG5’, and ‘Generic_MG6’). The range of planting dates and the selection of maturity group was based on the observed dataset. The plant density was set as 32 plants m^{-2} , and row spacing was set as 0.75 m to evaluate yield and stability across the different combinations of planting date and maturity groups in the current and future weather. Moreover, it is worth mentioning that the simulations were carried out on the same 13 sites used to validate the model.

To interpret and compare the difference between both scenarios the last decade of the

current (2011-2021) and future weather (2042-2052) were fitted with a linear mixed model using the lme4 package (Bates et al., 2015) in R (R Core Team, 2021) and then analyzed with ANOVA (R Core Team, 2021). For this model grain yield was function of maturity group, planting date, cluster, and their interactions, being the site the random component. Each region was evaluated independently, to explore the best combination and the change in every region in both decades. Pairwise comparisons were conducted using the emmeans R package (Lenth, 2021) with Tukey method at significance level $\alpha=0.05$ in R (R Core Team, 2021).

2.3 Results

2.3.1 Spatial clustering

Three clusters were defined within the region of study in Kansas: i) North-central, ii) North-east, and iii) South-east, named according to their spatial distribution (Figure 2.1). The clusters differed in their cumulative annual rainfall. These differences were driven by the spatial distribution of each cluster, as Kansas is characterized by a precipitation gradient increasing towards the South-east but moving from high to low precipitation from east to west.

The South-east cluster had the latest standard planting date and the greatest seasonal cumulative precipitation (Table A.1). In contrast, the North-central cluster had the lowest seasonal cumulative precipitation (less than 200 mm less relative to the South-east cluster). While the standard planting date and accumulated rainfall enclose notable differences between clusters, temperature and soil features differed slightly between clusters (Table A.1).

2.3.2 Model validation

The APSIM model simulated soybean seed yield reflecting an adequate performance ($KGE=0.44$) for the whole study region (Table 2.1). The difference between observed and simulated yields, expressed as RMSE was 1084 kg ha^{-1} , which represented 29% of the observed mean

(RRMSE). The model over-predicted yields relative to the observed yields (Figure 2.2). Overall, the validation model for each maturity group adequately modeled the seed yield dynamics for this study. This result highlights the capability of the generic cultivars to capture the yield variability of the maturity groups.

As a second step, the APSIM model performance was evaluated individually for each cluster (Figure 2.2). The model simulated more accurately than precisely (Figure 2.2, $PLA < PLP$) in the South-east and in the North-east clusters. In the North-east cluster, the soybean seed yield was simulated with an RRMSE of 27%. Meanwhile, in the North-central cluster, the model simulated with more precision than accuracy ($PLP = 43 < PLA = 57$), and with good efficiency ($KGE = 0.48$).

2.3.3 Developing scenarios for yield stability for the current weather

The simulated seed yield for current weather varied depending on the cluster. South-east presented the highest yield (mean = 4383 kg ha^{-1}), whereas the North-central cluster had the lowest yield (mean = 2774 kg ha^{-1}), all simulations under rainfed conditions. Conversely, the standard deviation did not change among clusters, but varied depending on the planting date. In general, early planting had higher yields but less stability, for all the maturity groups, within all the clusters (Figure 2.3).

The most drastic change in soybean yield with maturity group and planting date was observed in the South-east (Figure 2.3). In the South-east, maturity groups IV, V, and VI showed satisfactory results when they were planted from April-15 to June-1 (Figure 2.3, Table A.4). These dates maintained similar yields but increased their stability over the years at the later dates. In the latest planting date (July-15), it was notable that standard deviation changed depending on the maturity groups, with reduced values for shorter maturity groups.

The North-east presented an intermediate yield (mean = 3662 kg ha^{-1}). This cluster did not show a significant difference between the interaction of planting date and maturity group (Table A.5). In the early dates (April-15 to May-15), seed yield and standard deviation did not vary between maturity groups (Figure 2.3). Moreover, maturity group IV yield did not

change until June-15, but the standard deviation of yield for this variety decreased with late planting times. On the latest planting dates (July-1 to July-15), yield was greater when the maturity group was shorter, but stability did not change among maturity groups.

The North-central cluster, as well as North-east, did not present a significant difference between the interaction of planting date and maturity group (Table A.5). Maturity group IV showed the best results maintaining high yields and reduced variability until June-1 (Figure 2.3). Likewise for the North-central, the North-east cluster in the latest planting dates displayed better yields with the shorter maturity group, with less variation of the standard deviation among soybean varieties.

2.3.4 Developing scenarios for yield stability for the future weather

In general, the analysis predicted that in the next 30 years seed yield would decrease by 21% (Figure A.4). The extent of yield reduction varied for the different clusters. The North-central cluster was the least affected by the future weather, presenting the lowest decrease in seed yield (4%), while the South-east cluster experienced the largest yield decrease (39%). This result was associated with future weather changes in each cluster (Table A.6). The South-east cluster was the most affected by the future weather scenario, decreasing its cumulative seasonal precipitation by $\sim 26\%$. Additionally, this cluster presented the highest increase in temperature, especially for the summer season.

In the North-central, maturity groups IV, V, and VI had a higher decrease in seed yield with early planting dates (Figure 2.4). In contrast, maturity group III showed a different pattern. For the earlier dates (April-15 – May-1), the maturity group III showed an increased yield under future weather in comparison with the current weather. Conversely, late dates (July-1 – July-15) showed good yield responses with the later maturity groups (V and VI) in the future weather (2042–2052), presenting a slight increase in yield (Figure 2.4). Likewise, maturity groups III and IV had a low yield decrease at the latest planting date (July-15).

In the South-east, the most negatively impacted location under future weather scenarios (Figure 2.4), early dates had a larger decrease in yield (2235 kg ha^{-1}) than late dates (1052

$kg\ ha^{-1}$). Furthermore, the latest planting date (July-15) demonstrated a slighter yield decrease for all maturity groups. In the case of maturity groups III and IV, intermediate planting dates (May-15, and June-1) were the most affected.

In general terms, early planting dates maximized seed yield in the northern clusters, even with the impact of future weather. However, late planting dates (June-15 to July-1) increased in importance for yield maximization in the South-east cluster (Table A.7) without changing the superiority in yield of the later maturity groups.

2.4 Discussion

This study provides a quantification of the optimum planting date by soybean maturity group combinations under current and future weather in the western border of the main US soybean producing region (USDA-NASS, 2022). Even though there is a wide range of plausible planting dates in the central south of Kansas (Ciampitti et al., 2022), an extensive and formal study of management optimization has not been reported before for this region. This situation leads farmers to decide management strategies with uncertainty, not only on how to maximize yields but also to consider its stability (yield variation) over time. Lastly, although the future weather for this region has been described (Brunsell et al., 2010), the overall impact on yield and crop management adaptation practices has not been formally assessed.

Several studies in the literature suggested that early planting dates may produce higher yields in the Midwest (Bastidas et al., 2008; de Bruin & Pedersen, 2008; Egli & Cornelius, 2009), and Southern regions (Chen & Wiatrak, 2010; Egli & Bruening, 2000; Egli & Cornelius, 2009). Comparable results were shown in this research for the three clusters within Kansas. Additionally, Egli & Bruening (2000) reported a decline in yield of 13–36% on late planting dates in the Southern region of the US (United States), like the simulated yield decreases observed in this study (24–30%). Furthermore, when focusing on the maturity by planting date interaction, Salmeron et al. (2014) described a decrease from 7% (for maturity group, MG III) to 18% (for MG V), in the US Mid-South region when the planting dates were

late. Similarly, this study revealed a larger impact on yield across the explored genotypes (from 25% for MG III, to 44% for MG VI), highlighting the conclusion that maximizing yields is only possible if we tailor the selection of variety and planting dates on regional environmental factors (Chen & Wiatrak, 2010). Usually, late planting shifts the reproductive stage into a less favorable environment with shorter days (Egli & Bruening, 1992; Heatherly & Spurlock, 1999; Kessler et al., 2020), and early planting increased risk of crop failure due to late spring frosts (de Bruin & Pedersen, 2008; Egli & TeKrony, 1996). Likewise reported on this current study, early planting dates maximized yield but penalized stability, whereas late planting dates penalized yield but resulted in higher stability (Ciampitti et al., 2022; Rizzo et al., 2022).

Not surprisingly, the effect of future weather was reflected in a reduction in soybean yields (ranging from 4 to 39%), pointing out the potential variation under different environmental conditions. Furthermore, previous studies suggest that water deficit is one of the main causes of the yield gap (Battisti & Sentelhas, 2015; Zhou et al., 2021). Likewise, our study reflected comparable reductions in yield with less seasonal precipitation. Several studies have already documented reductions in soybean yield and its stability under future weather (Z. Jin et al., 2018; Mourtzinis et al., 2019, 2015; Schlenker & Roberts, 2009; Zhou et al., 2021), ranging from 19 to 82%. However, those past investigations focused on quantifying yield reductions without providing any adaptation strategies (Elli et al., 2022). In this context, crop management arises as an important mechanism to mitigate the impact of climate change on yield (Battisti & Sentelhas, 2015; Kumagai & Sameshima, 2014; Y. Liu & Dai, 2020). Among the management options, the selection of planting date (Battisti & Sentelhas, 2015; do Rio et al., 2016; Mourtzinis et al., 2019; Rurinda et al., 2015) and maturity group (Kumagai & Sameshima, 2014; Y. Liu & Dai, 2020) are crucial factors that producers can manage. According to do Rio et al. (2016) and Y. Liu & Dai (2020) delaying the planting date could mitigate the yield penalty under future predicted weather conditions as reported for Brazil and China, respectively. Similarly in central-east Kansas results showed that for future scenarios, long maturity groups (IV, V, and VI) with late planting dates provide lesser yield reductions relative to early planting dates. To obtain maximum yields under

future weather, early planting dates remain the best choice for farmers. Nonetheless, for the Southeast cluster, long maturity groups can help to express their compensation capability (do Rio et al., 2016) relative to the short duration of the growing season due to the late planting dates.

Some limitations in this study were: i) limited observed dataset on crop phenology for improved model calibration (Wallach et al., 2021; Wu et al., 2019), ii) lack of observed soil data, forcing to employ data generated via average of large scale, not related to a specific zone (Hoffmann et al., 2016), and iii) lack of more relevant geographical representation (more environments). Overall, this modeling approach captures the aim of this study correctly and provides a starting point related to improving the planting date x maturity group management in this region and a way to mitigate future weather. Future research could enhance the quality of this information by providing observed data with phenology to assemble a correct model calibration and validation (Wallach et al., 2021; Wu et al., 2019), thus improving its accuracy in predicting events. Nevertheless, it was remarkable the capability of the generic cultivar available in APSIM to capture the yield variability of the maturity groups. However, it is critical for future research studies to focus on collecting a dataset with greater spatial distribution and high-resolution soil mapping are critical to attain more suitable “site-specific” recommendations of planting date with maturity group across this region.

2.5 Conclusion

This study highlights potential benefits for farmers when selecting the optimal combinations of planting dates by maturity groups for soybean yield. The most relevant outcomes of this study under current weather conditions were: i) farmers maximize soybean yield by planting in April to early May, and ii) farmers can find more yield stability (less variation) over time when planting late but with a clear yield penalty for maximum yields. For the future weather: iii) yield reductions should be expected, but the selection of the right maturity group x planting date could be used as a mitigation practice, and iv) later maturity groups

(IV, V, and VI) can result in less yield reduction for late planting dates compared to early.

Table 2.1: Measure of agreement between observed and simulated data, for the entire study area. RMSE, root mean square error; RRMSE, relative root mean square error; PLP, percentage lack of precision; PLA, percentage lack of accuracy; KGE, Kling-Gupta model efficiency.

Metric	Value
RMSE	1084
RRMSE	0.29
PLP	68
PLA	32
KGE	0.44

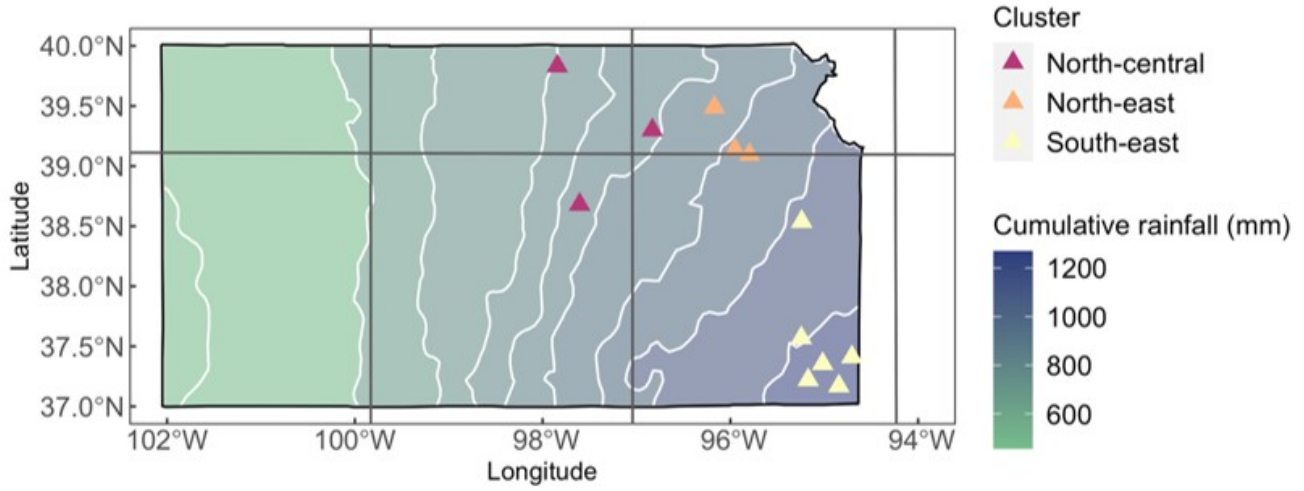


Figure 2.1: Kansas map with the three clusters (represented by triangles with distinct colors) for crop seasonal cumulative annual precipitation (mm), and the T42 GCM grid resolution (represented by solid lines) reported by [Brunsell et al. \(2010\)](#).

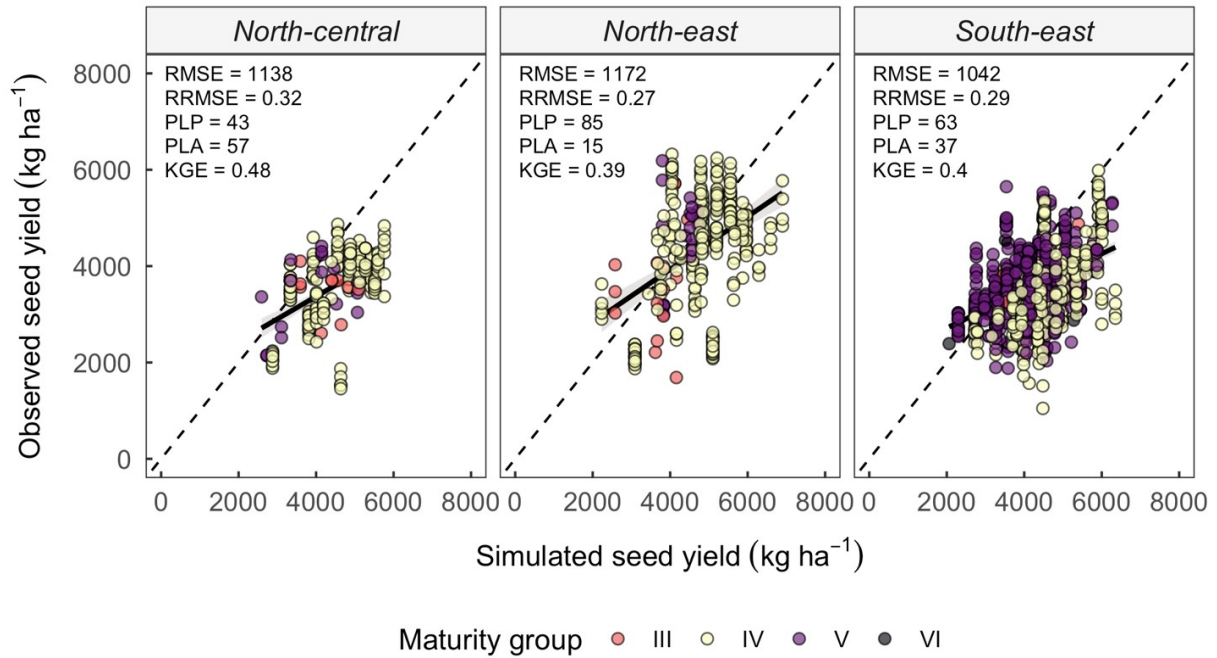


Figure 2.2: Observed versus simulated seed yield for the three clusters (North-central, North-east, and South-east) with soybean maturity groups III, IV, V, and VI (represented by distinct colors). The dashed line represents the 1:1 line. RMSE, root mean square error; RRMSE, relative root mean square error; PLP, percentage lack of precision; PLA, percentage lack of accuracy; KGE, Kling-Gupta model efficiency.

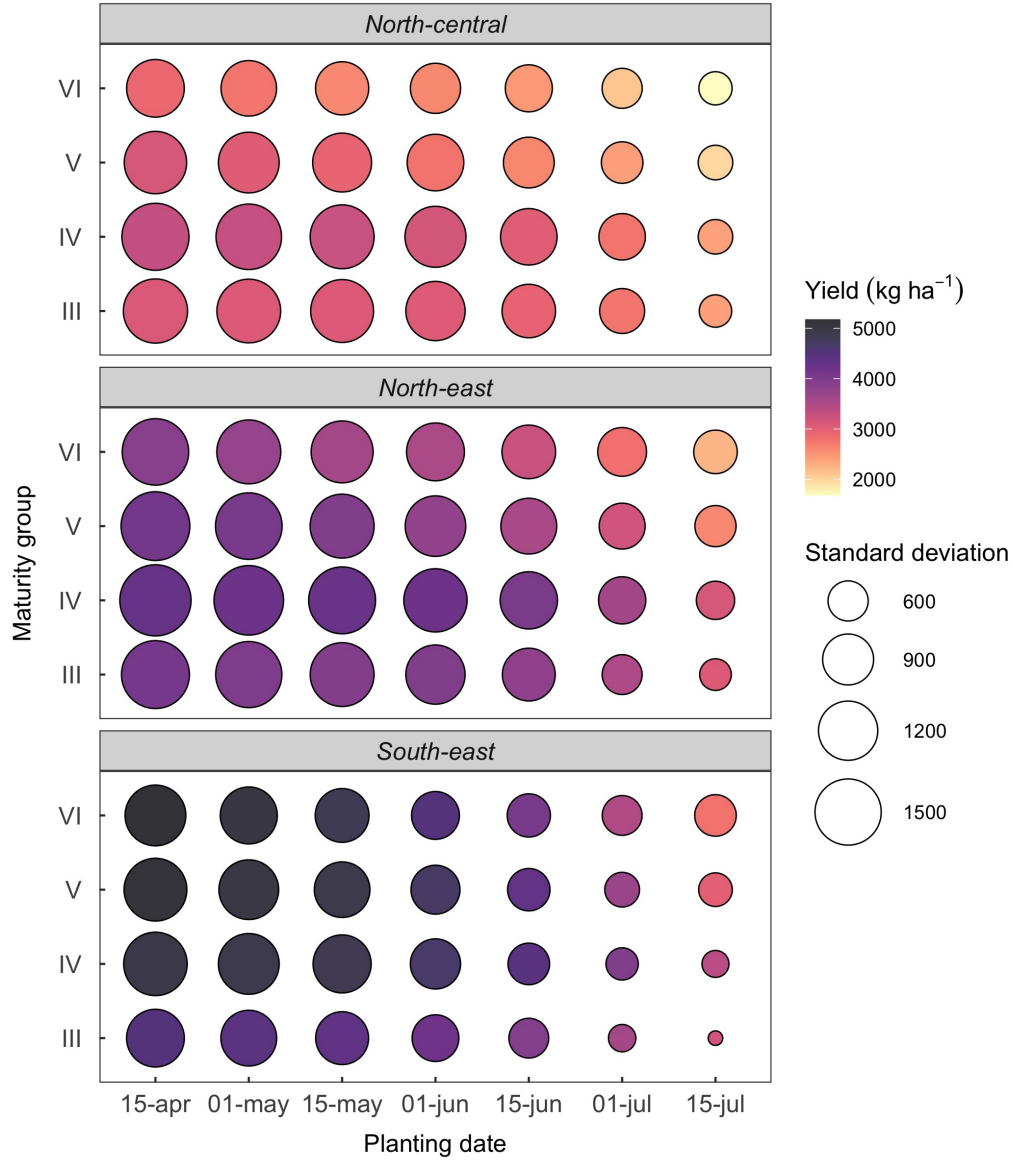


Figure 2.3: Simulated seed yield for the current weather (2011 to 2021), where darker colors represent higher yields and the size of the bubbles represent the standard deviation across years, within each cluster, for the different combinations of planting date and maturity groups. Planting dates are referred to as day-month.

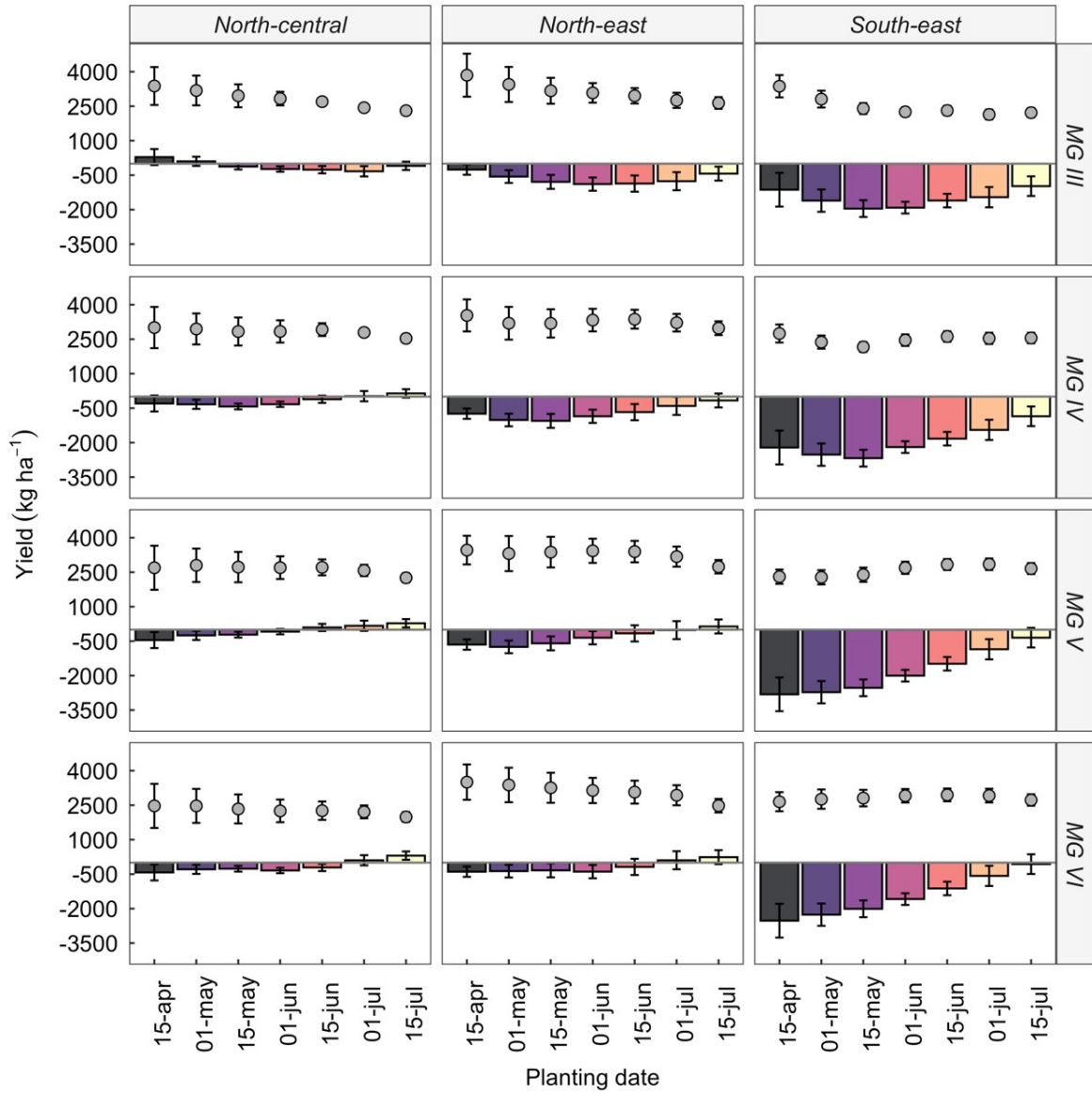


Figure 2.4: Average difference between future and current weather yields (bars, difference between the year ranges 2042-2052 vs 2011-2021) and seed yield under future weather scenarios (points, from the year 2042 to 2052) for the different combinations of planting dates and maturity groups within the three clusters in Kansas. Vertical lines represent standard deviation of the yield difference across different sites within clusters (lines over the bars) and across years and different sites in each cluster (lines over the points). Planting dates are referred to as day-month.

Chapter 3

Soybean yield and seed quality in equidistant versus non-equidistant plant arrangements under different densities

* Published in *Crop Science*

van Versendaal, E., Pereyra, V. M., Irby, T., Kovacs, P., Hefley, T., Prasad, P. V., Kyveryga, P., Van De Woestyne, B. W., & Ciampitti, I. A. (2025). Soybean yield and seed quality in equidistant versus non-equidistant plant arrangements under different densities. *Crop Science*, 65(1), e21364.

<https://doi.org/10.1002/csc2.21364>

Abstract

Equidistant plant arrangements have shown positive impacts, in the United States (US) over the last century, on soybean (*Glycine max* L.) yield and seed quality by reducing intraspecific plant competition and enhancing early canopy cover. This study aims to: i) assess the effects

of equidistant versus non-equidistant plant arrangements on soybean yield and seed quality across different regions in the US; ii) explore the effect of modified soybean plant canopy resulting from different plant arrangements on yield and seed quality. In 2021 and 2022, thirteen trials were conducted in Kansas, Mississippi, and South Dakota. The treatments consisted of three equidistant plant arrangements of 13x13, 15x15, and 20x20 cm (620,000; 430,000; 242,363 seeds ha⁻¹, respectively), and a non-equidistant with 38 cm of row space (321,236 seeds ha⁻¹). Soybean yield and yield components were collected, and canopy coverage (%) was monitored throughout the growing season. Overall, there is no evidence of any advantage of equidistant over non-equidistant arrangements in soybeans. Yield differences related to spatial arrangement were observed in only three trials: the non-equidistant increased yield compared to the 20x20 cm equidistant arrangement (lowest density) in two trials, while the 13x13 and 15x15 patterns (highest density) showed higher yield compared to the 20x20 cm arrangement in one trial. Furthermore, seed quality remained constant across spatial arrangements in the 13 trials. In contrast to prior research, our study found no correlation between canopy coverage development and yield improvement. The tested equidistant arrangement while promising did not provide substantial evidence of yield improvement relative to the non-equidistant.

3.1 Introduction

Soybean (*Glycine max* L.) is a major field crop due to its importance as a source of protein and oil (FAOSTAT, 2023). Through advancements in breeding and crop management, soybean yield has increased in response to the global food demand (Foley et al., 2011; Ray et al., 2012). The United States (US) is the world’s second-largest soybean producer, contributing 30% of the world’s production (USDA, 2023), highlighting the importance of testing modern technologies in this country to enhance on-farm production. However, uncertainties remain regarding the optimal soybean plant arrangement (Carciochi et al., 2019; Cox & Cherney, 2011; Pereyra et al., 2022) to maximize not only yield but also seed quality (e.g., protein and oil).

Soybean seed yield and quality are impacted by management practices, including plant spatial arrangement as defined by plant density and row spacing, and their interactions with the environment (J. F. Andrade et al., 2019; de Bruin & Pedersen, 2008; Lu et al., 2020). In US, soybean production has shifted to narrow row spacing, shrinking from 76 cm to 38 cm and, more recently to 19 cm, resulting in greater yields due to improved plant light interception (J. F. Andrade et al., 2019; de Bruin & Pedersen, 2008; Specht et al., 2015). Contrarily, wider row spacing might result in increased seed quality (Boydak et al., 2002). To optimize space utilization in a narrow row system, changes in seeding rates may also be beneficial, resulting in an equidistant plant pattern (Devlin et al., 1995; Schutte & Nleya, 2019). Several studies conducted in the US over the last century have explored the effects of equidistant plant arrangements on soybean yield (Duncan, 1986; Egli, 1994; Safo-Kantanka & Lawson, 1980; Shibles & Weber, 1966; Wiggans, 1939). Collectively, these studies demonstrated that planting soybeans in square patterns leads to increased yields compared to planting them in rectangular patterns. However, no studies in the last three decades have evaluated the effect of different equidistant spatial arrangements on the yield impact for modern US soybean varieties.

Equidistance arrangements may reduce intraspecific competition for nutrients and water (Shibles & Weber, 1966). Moreover, equidistant plant arrangement may promote early development canopy leaf area and greater light interception, thereby increasing crop growth rate (Roth et al., 2022; Shibles & Weber, 1966). Achieving full canopy coverage earlier in the season can increase yields (Bullock et al., 1998; Fickett et al., 2013; Moreira et al., 2019). In addition, attaining full canopy coverage mid-season may increase protein levels (Roth et al., 2022). Conversely, soybean plants have a high capacity to compensate for plant-to-plant growth differences during the seed set period (crop plasticity), making the crop less susceptible to yield variations (F. H. Andrade & Abbate, 2005; Cox & Cherney, 2011; Masino et al., 2018). Nonetheless, the ability of the plant to improve light interception is a key mechanism to improve yields (Willcott et al., 1984). Over time, there has been a gradual and consistent increase in the total amount of light interception, which can be attributed to various factors, including the reduction of lodging, the later maturation of cultivars, and the extension of

the growing season (Koester et al., 2014). This increase in light interception has been a significant contributing factor to the increased yields of soybeans and has been correlated with a higher seed protein concentration at the base of the main stem (Huber et al., 2016). Lastly, maintaining seed quality as yield continues to improve over time could be linked to a better light distribution throughout the leaf canopy (Bellaloui et al., 2012; Huber et al., 2016). Therefore, there could be an opportunity for improving both yields and seed quality with a more optimized plant arrangement.

Following this rationale, the main aims of this study were to: (i) evaluate equidistant versus non-equidistant plant spatial arrangement to assess its effect on soybean seed yield and seed quality (protein and oil concentrations), and (ii) explore plant canopy cover development throughout the season and its contributions to yield and seed quality across different US producing regions.

3.2 Materials and methods

3.2.1 Field experiment

Thirteen field trials were conducted in Kansas, Mississippi, and South Dakota, US, during the 2021 and 2022 growing seasons (Table 3.1). In 2021, a total four dryland and irrigated trials were in Kansas and Mississippi. The remaining nine trials were conducted in 2022, under dryland conditions with three planting dates in each state within Kansas, Mississippi, and South Dakota. These environments (dryland, irrigated, and the three planting dates) were selected to evaluate the impact of plant spatial arrangements under varied conditions. The inclusion of two additional planting dates for the second year was executed with the goal of exploring more seasonal variation across sites. Weather characterization is presented in Table B.1. Briefly, weather variables were retrieved from the DAYMET database (Thornton et al., 2021) and summarized from soybean growing season (May to October) at all site-years.

All studies were arranged in a randomized complete block design with three or four replicates, depending on the trial. The size of each plot was 7.62 m in length and 1.52 m

in width. Treatments were manually sown to achieve the desired arrangement and planted at a depth of 2.54 cm. All field trials presented four treatments with different soybean plant arrangements. Three treatments had an equidistant spatial arrangement, where both plant spacing and row spacing were identical (12.70 x 12.70 cm with 620,000 seeds ha^{-1} , 15.24 x 15.24 cm with 430,000 seeds ha^{-1} , and 20.32 x 20.32 cm with 242,636 seeds ha^{-1}). For simplicity, these were named 13 x 13 cm, 15 x 15 cm, and 20 x 20 cm, respectively. The last treatment was considered a regular arrangement (current farming practice), non-equidistant, with a row spacing of 38 cm and a seeding rate of 321,236 seeds ha^{-1} . In essence, our comparison involved three equidistant plant arrangements with varying densities against a single non-equidistant plant arrangement, which reflects current farming practices.

3.2.2 Data collection

Soybean yield was determined via harvesting the entire plots with a combine. Thousand seed weight was measured from a subsample and seed number (seeds m^{-2}) was estimated from the seed yield and thousand seed weight (grams). Yield data is reported with 130 g ha^{-1} of moisture. For this study, the term soybean seed quality refers specifically to both seed oil and protein concentrations. To determine seed quality on a dry basis (%), near-infrared spectroscopy (NIR) was employed using the Perten DA7200 Feed Analyzer (Perten Instruments, Stockholm, Sweden).

Canopy coverage (CC) measurements were taken using a GoPro Hero5 Black portable digital camera (GoPro, Inc., San Mateo, California) parallel to the soil at ~ 2 m height within each plot at various growth stages of the crop (Figure 3.1). In Kansas, 2021, CC was collected in 6 moments across the whole growing season. In Kansas, 2022, the CC data were obtained weekly in the three trials (with 15, 12, and 9 samples during the season in the May, June, and July planting date, respectively). In Mississippi, in both growing seasons, the photos were taken at least 3 points in time before the crop reaches the maximum CC. In South Dakota, canopy cover measurements were not available. CC (%) was obtained using Canopeo (Figure 3.1, [Patrignani & Ochsner, 2015](#)), a monitoring tool for crop growth based on colorimetry

that calculates the green canopy size and enables accurate monitoring changes in canopy size during the crop season. To accurately reflect the crop’s developmental stages, the moments of CC measurement are expressed in thermal time, using a base temperature of 10°C.

To obtain and subsequently identify the critical period of soybean development (R1-R3 and R3 – R5 phenological stages) in the CC, DSSAT model v. 4.8 was utilized. This crop growth simulation tool is extensively detailed in (Boote et al., 1998). The model was set to simulate soybean phenology for each trial using the coefficients of the default genotypes corresponding to the maturity groups. The performance of the model was evaluated based on its ability to predict phenological stages R1, R3, and R5 during the 2022 Kansas trials, which provided the phenological data. The model accurately identified these critical stages, demonstrating robust performance (Figure B.1).

3.2.3 Data analysis

The analysis of soybean seed yield (kg ha⁻¹), seed number (seeds m⁻²), and thousand seed weight (g) in relation to plant arrangement across trials was conducted using a hierarchical Bayesian framework. For this study, the term variable response will refer to the analysis of yield and its components. In this framework, the first level of the model (data layer) describes the variable response, $y_{ij(k)}$, for the i^{th} block ($i = 1, 2, 3, 4$) with the j^{th} plant arrangement ($j = 1, 2, 3, 4$), which is in the k^{th} trial ($k = 1, 2, \dots, 13$). The variable response $y_{ij(k)}$ was assumed to follow a normal distribution with expected value $\mu_{ij(k)}$ and variance σ_k^2 :

$$y_{ij(k)} \sim \text{Normal}(\mu_{ij(k)}, \sigma_k^2). \quad (3.1)$$

Additionally, the second level of the model incorporates the underlying process model. The parameter $\mu_{ij(k)}$ was defined as a linear effect:

$$\mu_{ij(k)} = \alpha_{j(k)} + b_{ij(k)}, \quad (3.2)$$

where $\alpha_{j(k)}$ is the overall value of yield or yield components associated with the j^{th} plant

arrangement within the k^{th} trial, and $b_{ij(k)}$ denotes the random effect of the i^{th} block with the j^{th} plant arrangement within the k^{th} trial. A uniform distribution was assigned to α with the minimum and maximum values depending on previous agronomy information. For yield and seed number, the prior was defined as $\alpha \sim \text{Uniform}(0, 15000)$. Similarly, for thousand seed weights, the prior distributions were defined as $\alpha \sim \text{Uniform}(0, 1000)$. Moreover, for σ_k^2 , the prior was an uninformative Half-student- t prior ($\sigma^2 \sim \text{Half-student-}t(3, 0, 2.5)$), where the parameters ν , μ , and σ represent the shape, location, and scale, respectively.

A hierarchical Bayesian framework was employed to evaluate the response of protein and oil seed concentration (seed quality) to plant arrangements across trials. The protein and oil concentration in the i^{th} block ($i = 1, 2, 3$ or 4), under the j^{th} plant arrangement ($j = 1, 2, 3, 4$), which is in the k^{th} trial ($k = 1, 2, \dots, 13$), are described at the model's top level (Eq. 3.3). The first layer of this framework was defined using a beta distribution (Ferrari & Cribari-Neto, 2004; Ospina & Ferrari, 2010) with an expected value $\mu_{ij(k)}$ and dispersion δ_k :

$$Q_{ij(k)} \sim \text{Beta}(\mu_{ij(k)}, \delta_k). \quad (3.3)$$

In the secondary level of the model, $\mu_{ij(k)}$ was defined as a linear effect:

$$\mu_{ij(k)} = \alpha_{j(k)} + b_{i(k)}, \quad (3.4)$$

where $\alpha_{j(k)}$ is the protein or oil concentration (%) obtained with the j^{th} plant arrangement within the k^{th} trial. $b_{i(k)}$ is the random effect of the i^{th} block, with the j^{th} plant arrangement within the k^{th} trial. To determine the values of α , a normally distributed prior was selected ($\alpha \sim \text{Normal}(0, 2.25)$). Moreover, for δ_k was assigned an uninformative Half-student- t ($\delta_k \sim \text{Half-Student-}t(3, 0, 2.5)$), where the parameters ν , μ , and σ represent the shape, location, and scale, respectively.

The analysis of the evolution of CC in relation to plant arrangement across trials was conducted using a hierarchical Bayesian framework. The evolution of the canopy coverage ($CC_{ij(k)}$) in the i^{th} block ($i = 1, 2, 3$ or 4) with the j^{th} plant arrangement ($j = 1, 2, 3, 4$),

within the k^{th} trial ($k = 1, 2, \dots, 13$), is described at the model's first level (data layer) (Eq. 3.5). This layer was defined by a parametrization of the beta distribution (Damgaard & Irvine, 2019) with an expected value $\mu_{ij(k)}$ and dispersion δ_k :

$$CC_{ij(k)} \sim \text{Beta}(\mu_{ij(k)}, \delta_k). \quad (3.5)$$

In the second level (underlying process model), $\mu_{ij(k)}$ was specified as a hill or bihill function, being both sigmoid functions (Archontoulis & Miguez, 2015). The hill function was employed in each trial to assess the progression of CC until it reaches its maximum level (Eq. 3.6). If data on CC was available for the entire growing season, the bihill function was used (Eq. 3.7). To obtain the values of each parameter in both functions ($\alpha, \beta, \gamma, \delta$, and θ), a normally distributed prior was selected (Eq. 3.8, Eq. 3.9, Eq. 3.10, Eq. 3.11, Eq. 3.12). Additionally, each parameter was influenced by the plant arrangement as a fixed effect and the block as a random effect within each trial:

$$\mu_{ij(k)} = \alpha_{ij(k)} \frac{\text{Thermal time}^{\beta_{ij(k)}}}{\gamma_{ij(k)}^{\beta_{ij(k)}} + \text{Thermal time}^{\beta_{ij(k)}}}, \quad (3.6)$$

$$\mu_{ij(k)} = \frac{\alpha_{ij(k)}}{\left(\frac{\gamma_{ij(k)}}{\text{Thermal time}}\right)^{\beta_{ij(k)}} \left(\frac{\text{Thermal time}}{\delta_{ij(k)}}\right)^{\theta_{ij(k)}}}, \quad (3.7)$$

$$\alpha \sim \text{Normal}(60, 40), \quad (3.8)$$

$$\beta \sim \text{Normal}(10, 10), \quad (3.9)$$

$$\gamma \sim \text{Normal}(450, 250), \quad (3.10)$$

$$\delta \sim \text{Normal}(1500, 250), \quad (3.11)$$

$$\theta \sim \text{Normal}(30, 30), \quad (3.12)$$

where $\alpha_{ij(k)}$ represents the maximum CC (%) obtained in the i^{th} block with the j^{th} plant

arrangement within the k^{th} trial. $\beta_{ij(k)}$ refers to the change in CC by unit change in thermal time ($\% \text{ } ^\circ\text{Cd}^{-1}$) until it reaches the maximum CC (Growth rate) in the i^{th} block with the j^{th} plant arrangement within the k^{th} trial. $\gamma_{ij(k)}$ is the moment in thermal time ($^\circ\text{Cd}$) where the canopy reaches the maximum growth rate ($\beta_{ij(k)}$) in the i^{th} block with the j^{th} plant arrangement within the k^{th} trial. $\delta_{ij(k)}$ is the moment in thermal time ($^\circ\text{Cd}$) where the canopy reaches the maximum decay rate ($\beta_{ij(k)}$) in the i^{th} block with the j^{th} plant arrangement within the k^{th} trial. $\theta_{ij(k)}$ refers to the decay rate, which is the change in CC by unit change in thermal time ($\% \text{ } ^\circ\text{Cd}^{-1}$) from the maximum CC until senescence in the i^{th} block with the j^{th} plant arrangement within the k^{th} trial.

The Bayesian analyses were conducted using the brms package (Bürkner, 2021) in the R software (R Core Team, 2021). In all evaluations, a warm-up phase was carried out with 5000 posterior samples, followed by 15,000 iterations. The R_{hat} values were utilized to assess the mixing and convergence of the chain, with a value of 1 indicating convergence (Vehtari et al., 2021). Furthermore, the 95% credible interval estimated from the posterior distributions of each treatment within each site was used for treatment comparisons. When these intervals are found to overlap between treatments, no effects from plant spatial arrangement are observed. The selection of a Bayesian approach was driven by the utilization of prior distributional knowledge, which accounts for the underlying data distribution, allowing for a more precise and informed analysis (O’Hagan, 2008).

3.3 Results

3.3.1 Plant arrangement effect on soybean seed yield

During the 2021 growing season, the trials in dryland conditions exhibited a higher average seed yield (ranging from 5583 to 6823 $kg \text{ } ha^{-1}$) compared to the trials conducted in 2022 (ranging from 2069 to 5742 $kg \text{ } ha^{-1}$). In 2022, it is noteworthy that Kansas and Mississippi experienced a 27% and 43% reduction in rainfall, respectively, compared to the previous year. Additionally, South Dakota received the lowest rainfall among the states during the

same period (Table B.1). Figure 3.2 demonstrates that in environments yielding higher, non-equidistant plant spatial arrangements demonstrated an advantage over equidistant arrangements. This advantage is quantified by a 95% credible interval for the regression slopes, ranging from 0.54 to 0.87, derived from the posterior distribution of a regression analysis comparing equidistant and non-equidistant yields.

Further evaluation of the posterior distribution from treatment pairwise comparisons, considering the 95% credible interval, revealed that only three of the thirteen trials (KS21 Irrigated, MS22 Late, and SD22 Mid) exhibited differences based on plant spatial arrangement. In KS21 Irrigated and MS22 Late, utilizing a non-equidistant spatial arrangement (38 cm regular) resulted in increased yields (95% credible interval: 5965-6831 $kg\ ha^{-1}$ and 4398-6191 $kg\ ha^{-1}$, respectively) compared to the 20x20 cm arrangement (95% credible interval: 5003-5878 $kg\ ha^{-1}$ and 2398-4214 $kg\ ha^{-1}$, respectively), which had the lowest plant density. Moreover, in SD22 Mid the equidistant arrangement 13x13 and 15x 15 cm, both with the highest plant density, increased yield (95% credible interval: 2507-2930 $kg\ ha^{-1}$ and 2543-2964 $kg\ ha^{-1}$, respectively) with respect to the 20x20 cm arrangement (95% credible interval: 2014-2492 $kg\ ha^{-1}$). The remaining ten trials showed non-responsive behavior across different plant arrangements (Figure B.2).

Yield components (seeds number and thousand seed weight) revealed a consistent lack of responsiveness to the different plant spatial arrangements within each trial. Seed numbers averaged from 1650 to 5474 seeds m^{-2} across trials, while thousand seed weights averaged from 163 to 108 g (Table 3.2).

3.3.2 Plant arrangement effect on seed quality

The average protein concentration of soybean seeds varied from 39% to 40% across trials. Similarly, the average oil concentration ranged from 23% to 26%. Through an analysis of treatment pairwise comparisons and the 95% credible interval of the posterior distribution, consistent results were observed in the thirteen trials. These findings indicate a consistent lack of responsiveness in terms of protein and oil concentration among different plant

arrangements (Table 3.3).

3.3.3 Canopy coverage

Similarly, to yield components and seed quality, by evaluating the posterior distribution derived from the treatment pairwise comparisons, and considering the 95% credible interval, it was found a lack of responsiveness in the canopy parameters based on the plant arrangement across trials (Figure 3.3, Figure B.3, Figure B.4). Specifically, in KS21 Irrigated, where the non-equidistant arrangement led to a higher yield range compared to the 20x20 cm pattern, no differences were observed between the two spatial arrangements for any of the canopy parameters. The 38 cm row space arrangement exhibited a range of 78-100% for max CC, 433-614 °Cd for the moment of maximum growth rate, 4-8% °Cd⁻¹ for the growth rate, 1728-1822 °Cd for the moment of maximum decay rate, and 43-71% °Cd⁻¹ for the decay rate. Conversely, the 20x20 cm spatial arrangement resulted in a range of 77-100% for max CC, 466-651 °Cd for the moment of maximum growth rate, 4-8% °Cd⁻¹ for the growth rate, 1721-1814 °Cd for the moment of maximum decay rate, and 39-65% °Cd⁻¹ for the decay rate (Figure 3.3, Figure B.4). Similar results were found in MS22 Late, where no differences were observed between the two spatial arrangements despite their varying yield ranges (Figure B.4).

Furthermore, Figure 3.3 suggests that all trials generally attained maximum CC after R1 stage. Additionally, illustrates that in most trials, all treatments typically reached maximum CC before the R3 phenological stage, except in the KS22 Late trial. In this particular trial, despite achieving maximum CC after the R3 stage, there is no differences in yield or CC parameters across treatments.

3.4 Discussion

This study quantifies the impacts of plant equidistance arrangement relative to a non-equidistance arrangements on soybean yield and seed quality across different US soybean

production regions (USDA, 2023). While diverse studies have explored optimal row spacing and plant density (Carciochi et al., 2019; Cox & Cherney, 2011; de Bruin & Pedersen, 2008; Harder et al., 2007), and older research has investigated equidistance arrangements (Duncan, 1986; Egli, 1994; Safo-Kantanka & Lawson, 1980; Shibles & Weber, 1966; Wiggans, 1939), there has been a significant lack of comprehensive and formal recent studies on the effects of different equidistant spatial arrangements on soybean in the US. This study confirms a lack of defined impact of different equidistant arrangements on both seed yield and quality under different densities, providing valuable insights to inform potential for decisions on crop arrangement.

Plant arrangement can produce positive impacts on soybean seed yields in the US (Duncan, 1986; Wilcox, 1974), highlighting a decrease in yield as rectangularity departs from the square pattern (Safo-Kantanka & Lawson, 1980; Wiggans, 1939). Furthermore, Egli (1994) found that an equidistant spatial arrangement led to yield increases of 9 to 19%. However, soybean breeding efforts during the last three decades have increased the compensatory yield per plant by increasing branches (Board, 2000; Carpenter & Board, 1997; Suhre et al., 2014), biomass, pods, and seeds per plant (Cox & Cherney, 2011; Cox et al., 2010). Recent studies showed lower variation in yield at changing row spacing (Sobko et al., 2020) and plant densities in medium-high yielding environments (Carciochi et al., 2019; Pereyra et al., 2022), as well as a 31% increase in the number of seeds plant⁻¹ when the seeding rate was 321,000 compared with 469,000 seeds ha⁻¹ (Cox & Cherney, 2011). A plausible explanation is that reduced plant densities produce more proportionally seeds from branches, while under high plant density there is a larger contribution derived from the main stem. A prior study found that wider row spacing or lower seeding rates increase seeds and pods per plant but do not affect pod and seed density per area due to compensation by more plants per unit area (Cox & Cherney, 2011), aligning with our findings. Likewise, yield produced on branches is greater on newer cultivars (Suhre et al., 2014). Regarding soybean seed quality, Boydak et al. (2002) suggest that wider row spacing may lead to improved seed quality. Contrastingly, Sobko et al. (2020) found no significant effect of plant arrangement on protein or oil content, but higher plant density can lead to higher protein content. However, the present study did

not identify any discernible differences in seed quality across different plant arrangements.

The seasonal changes in canopy throughout the growing season have been demonstrated to influence both soybean seed yield and seed quality (Roth et al., 2022; Schmitz & Kandel, 2021). A number of studies have also demonstrated that early season canopy closure (Bullock et al., 1998; Cox & Cherney, 2011; Moreira et al., 2019; Roth et al., 2022) combined with improved growth rate (Masino et al., 2018) can increase yields. Surprisingly, under this study, no similarities were observed between canopy factors and yields. The absence of a correlation between canopy evolution and yield may be attributed to the capacity of soybeans to compensate, resulting in an increased number of branches per plant in low plant densities and wider rows (F. H. Andrade & Abbate, 2005; Carpenter & Board, 1997; Cox & Cherney, 2011; Suhre et al., 2014). Additionally, it has been demonstrated that if soybean can achieve canopy closure at stage R1 or capture more than 450 MJ m^{-2} of cumulative photosynthetically active radiation, there is no need to increase plant density in order to increase yield. This lack of yield advantage can be attributed to a reduction in light interception efficiency and radiation use efficiency (de Bruin & Pedersen, 2009; Edwards et al., 2005; Purcell et al., 2002). Contrary to these findings, this study observed variations in the moment of achieving canopy closure across different trials, yet still found no differences in yield variation. In the case of seed quality, previous literature indicated that achieving full CC in the mid-season and delaying crop maturity can increase protein (Roth et al., 2022). However, this study's findings indicate that there were no differences observed in the timing of achieving full CC or the seed quality across various plant arrangements. The lack of relatedness of seed quality with plant arrangement could be also explained because seed quality depends on genotype, geographic region, and climatic parameters (Chung et al., 2003; Medic et al., 2014; Roth et al., 2022; Vollmann et al., 2000).

Potential limitations in this study could be related to the limited treatments evaluating non-equidistant patterns (i.e., a wider range of row spacings), the absence of information related to soil moisture availability, and the lack of data on biomass partitioning into vegetative and reproductive structures to assess soybean plant compensation. For the first limitation, examining a broader range of treatments with consistent plant densities could enhance

our understanding of plant spatial arrangement effects. Regarding biomass partitioning, a physiological process-based characterization of the branching plasticity may provide deeper insights into plant behavior across various spatial arrangements. Overall, this approach captures the aim of this study correctly and provides a starting point for improving soybean management. Future research should focus on comparing equidistant patterns with non-equidistant patterns under comparable plant densities and exploring different canopy architectures (with less or more plasticity on the production of branches). Improving understanding of the limits for branching and compensation for soybeans will help provide foundational knowledge to study the optimization of plant arrangement and further opportunities for yield gains.

3.5 Conclusion

The most relevant outcomes of this study are: (i) in the majority of the trials, the equidistant spatial arrangements did not demonstrate any advantages over the non-equidistant arrangements in terms of soybean yield, (ii) protein and oil seed concentration remained constant across different plant arrangements, and (iii) canopy parameters were not found to be influential in soybean yield variation and seed quality traits. Future investigations should explore further opportunities for optimizing plant arrangement considering other factors such as plant density, plant canopy architecture (varieties), and planter systems to capture potential yield gains from improved resource use efficiencies and yields.

Table 3.1: Site-years characterization showing year, state, location, planting dates, maturity groups (MG), soybean variety, water management, number of replications (Rep.), and data collected on canopy coverage (CC).

Year	State	Location	ID	Planting date	MG	Variety	Water management	Rep.	Canopy coverage
2021	Kansas	Topeka	KS21 Dryland	5/13/21	IV	P42A84E	Dryland	4	Growing season
		Topeka	KS21 Irrigated	5/13/21	IV	P42A84E	Irrigated	4	Growing season
2021	Mississippi	Starkville	MS21 Dryland	4/20/21	V	P53A67X*	Dryland	4	Until full CC
		Starkville	MS21 Irrigated	4/20/21	V	P53A67X*	Irrigated	4	Until full CC
2022	Kansas	Manhattan	KS22 Early	5/10/22	IV	P42A84E	Dryland	4	Growing season
		Manhattan	KS22 Mid	6/16/22	IV	P42A84E	Dryland	4	Growing season
		Manhattan	KS22 Late	7/22/22	III	AG30XF2	Dryland	4	Until full CC
2022	Mississippi	Starkville	MS22 Early	4/29/22	V	P53A67X*	Dryland	3	Growing season
		Stoneville	MS22 Mid	5/19/22	V	P53A67X*	Dryland	4	Until full CC
		Starkville	MS22 Late	6/7/22	V	P53A67X*	Dryland	3	Until full CC
2022	South Dakota	Beresford	SD22 Early	5/19/22	II	AG20XF1	Dryland	4	Not available
		Aurora	SD22 Mid	5/26/22	II	AG20XF1	Dryland	4	Not available
		Aurora	SD22 Late	6/08/22	II	AG20XF1	Dryland	4	Not available

* Determinate growth habit cultivars.

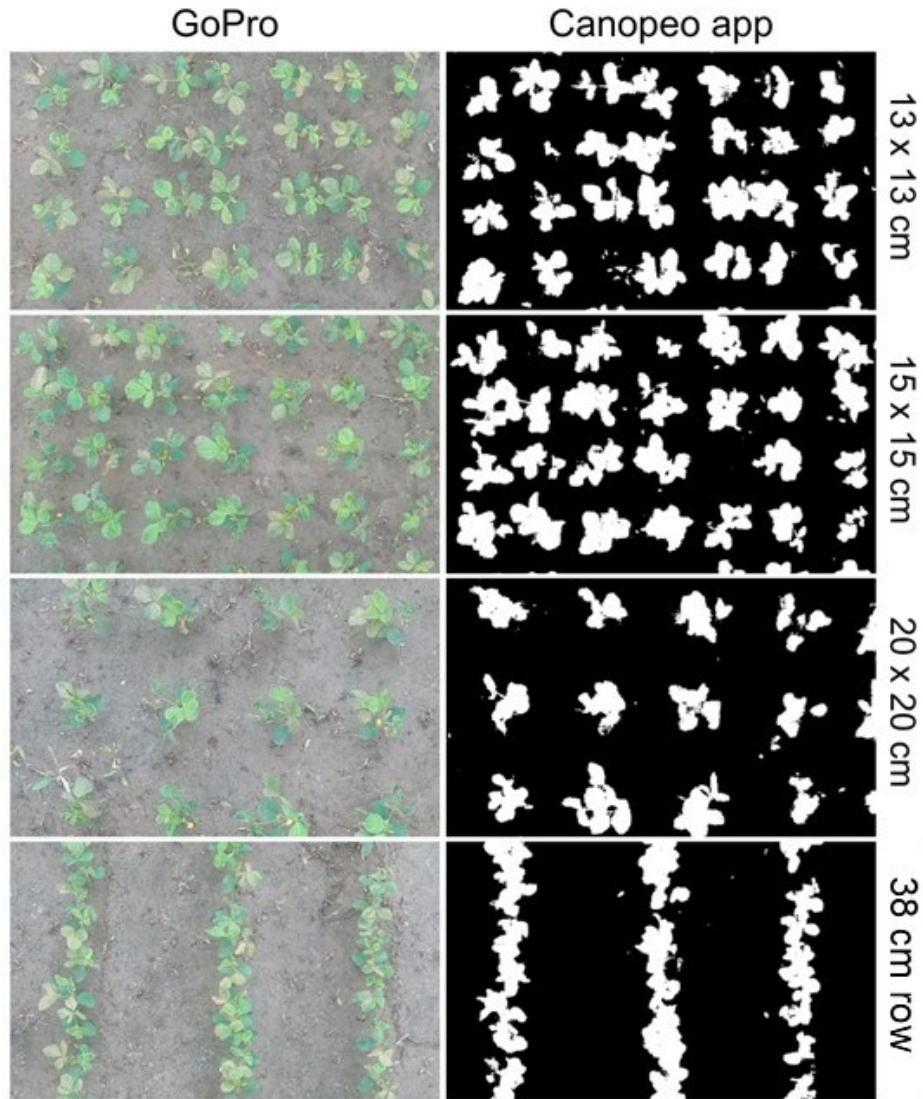


Figure 3.1: Illustrative comparison of pictures taken with a GoPro Hero5 Black portable digital camera on each plant spatial arrangement (from top to bottom) showing the original image (left), and after processing by Canopeo app (right).

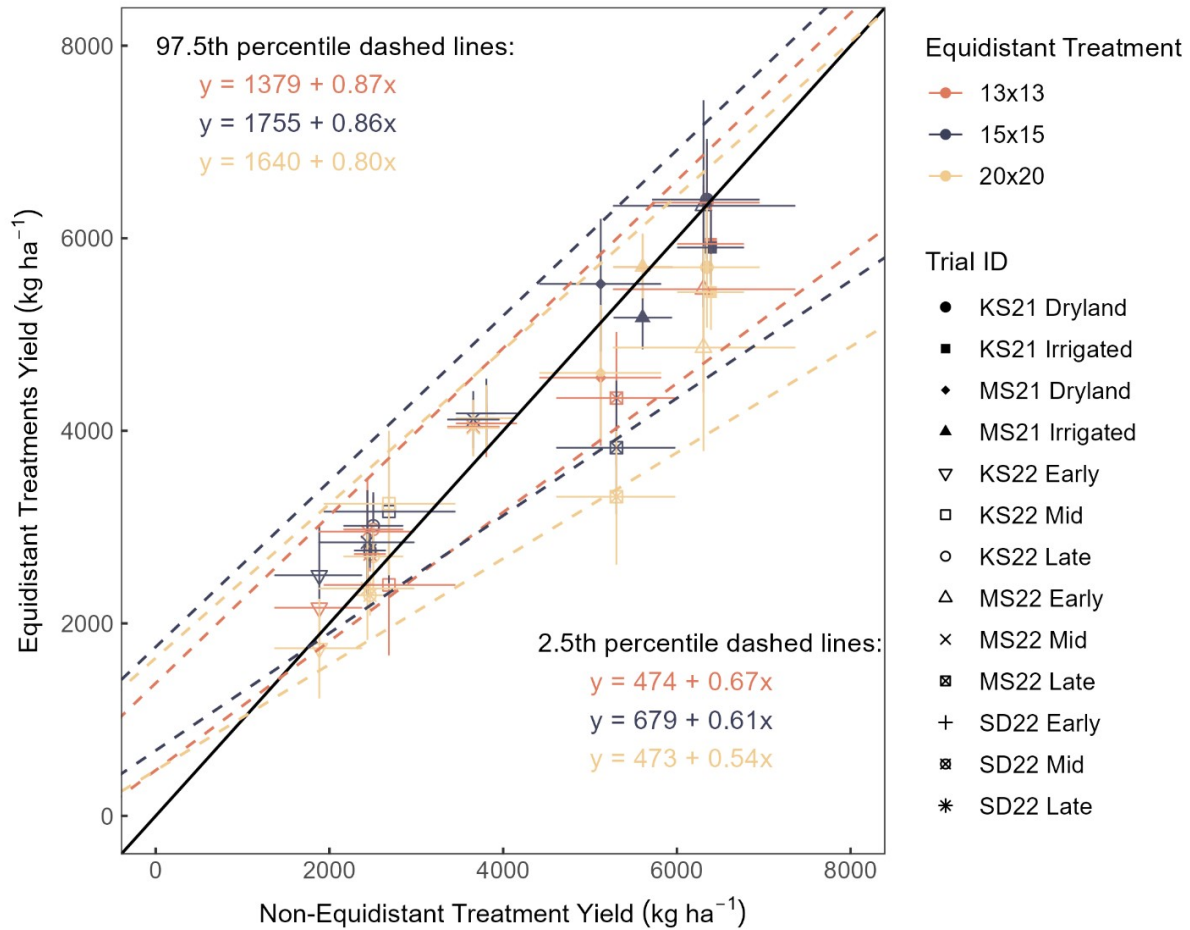


Figure 3.2: Seed yield for equidistant versus non-equidistant soybean plant spatial arrangements under different densities. Dashed lines correspond to 95% credible intervals. Black solid line corresponds to 1:1 line. Colors denote equidistant treatments and shapes represents each trial. Vertical and horizontal error bars show the 95% credible intervals from the pairwise comparisons of treatments. Equidistant treatment: 1) 13 x 13, 13 cm by 13 cm; 2) 15x15, 15 cm by 15 cm; 3) 20x20, 20 cm by 20 cm. Non-equidistant treatment: 38 cm of row space.

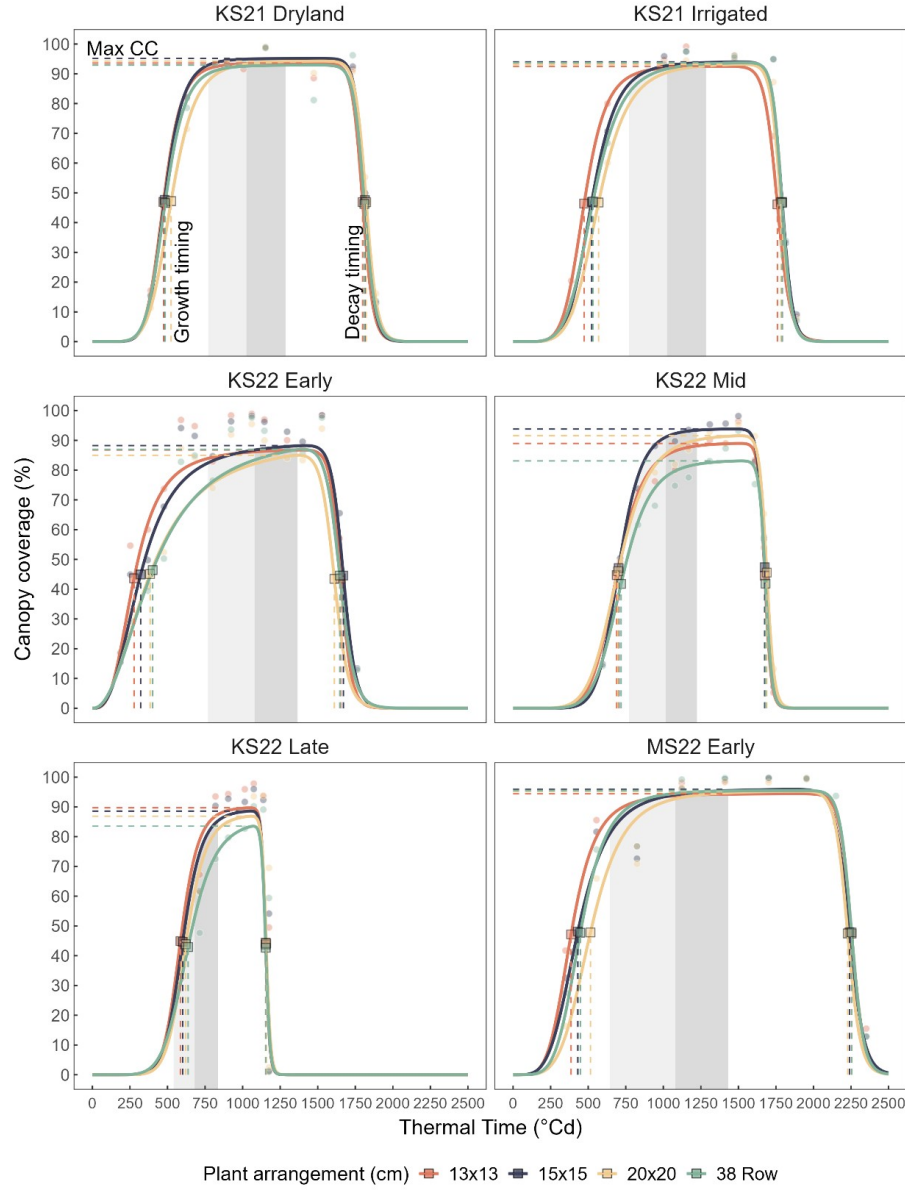


Figure 3.3: Soybean canopy coverage (CC) over thermal time for the entire growing season, presented for trials with complete seasonal data (each panel): Kansas 2021 in Irrigated and dryland conditions (KS21 Irrigated and KS21 Dryland); Kansas 2022 in an early, mid, and late-planting date (KS22 Early, KS22 Mid, and KS22 Late); and Mississippi 2022 in early-planting date. Solid lines correspond to the estimated CC from the bi-hill model. Dashed lines correspond to bi-hill parameters. Points refers to observed data. Square shapes denote moments of maximum growth rate (growth timing) and maximum decay rate (decay timing). Max CC refers to maximum CC. Light-gray and dark-gray areas correspond to the R1 to R3 and R3 to R5 phenological stages, respectively, as determined using DSSAT.

Table 3.2: Average soybean seed number, thousand seed weight, and their respective quantile 2.5% and 97.5% by plant arrangement within each trial (ID). Equidistant treatment: 1) 13x13, 13 cm by 13 cm; 2) 15x15, 15 cm by 15 cm; 3) 20x20, 20 cm by 20 cm. Non-equidistant treatment: 38Row, 38 cm of row space.

ID	Plant Arrangement	Seed Number (seeds m ⁻²)			Thousand Seed Weight (g)		
		Mean	Q: 2.5%	Q: 97.5%	Mean	Q: 2.5%	Q: 97.5%
KS21 Dryland	38Row	4087	2228	6665	155	142	170
	13x13	3714	1868	6174	156	143	171
	15x15	5347	3233	8211	162	149	178
	20x20	5175	3161	7987	146	134	160
KS21 Irrigated	38Row	4085	2916	5541	151	137	167
	13x13	4336	3177	5873	146	132	162
	15x15	4424	3219	5926	153	138	168
	20x20	3835	2709	5185	148	134	164
MS21 Dryland	38Row	3881	3233	4563	135	108	164
	13x13	3818	3176	4508	122	96	152
	15x15	3537	2921	4178	161	130	194
	20x20	3672	3052	4335	129	103	158

Continued on next page

Table 3.2: Continued from previous page

ID	Plant Arrangement	Seed Number (seeds m ⁻²)			Thousand Seed Weight (g)		
		Mean	Q: 2.5%	Q: 97.5%	Mean	Q: 2.5%	Q: 97.5%
MS21 Irrigated	38Row	3592	2998	4279	162	131	199
	13x13	3885	3250	4598	152	122	186
	15x15	3191	2654	3848	169	137	205
	20x20	3492	2905	4176	169	137	205
KS22 Early	38Row	1517	900	2160	125	110	140
	13x13	1852	1234	2498	120	105	135
	15x15	1980	1354	2621	126	111	142
	20x20	1453	840	2078	122	107	137
KS22 Mid	38Row	2082	1319	3106	137	117	157
	13x13	1799	1073	2775	145	127	167
	15x15	2603	1749	3709	127	109	147
	20x20	2322	1507	3384	150	130	172
KS22 Late	38Row	1872	1519	2289	136	124	149
	13x13	2313	1925	2780	131	120	144

Continued on next page

Table 3.2: Continued from previous page

ID	Plant Arrangement	Seed Number (seeds m ⁻²)			Thousand Seed Weight (g)		
		Mean	Q: 2.5%	Q: 97.5%	Mean	Q: 2.5%	Q: 97.5%
MS22 Early	15x15	2286	1897	2749	134	123	146
	20x20	1978	1616	2408	139	127	152
	38Row	5995	3637	9361	113	101	126
	13x13	5156	2772	8292	111	99	124
	15x15	6125	3719	9464	109	98	122
MS22 Mid	20x20	4620	2304	7555	110	98	123
	38Row	2338	1769	3038	108	99	117
	13x13	2791	2090	3685	109	100	119
	15x15	2746	2095	3529	106	98	115
	20x20	2277	1727	2982	107	98	116

Table 3.3: Average soybean seed protein and oil concentration (%), and their respective quantile 2.5% and 97.5% by plant spatial arrangement within each trial (ID). Equidistant treatment: 1) 13 x 13, 13 cm by 13 cm; 2) 15x15, 15 cm by 15 cm; 3) 20x20, 20 cm by 20 cm. Non-equidistant treatment: 38Row, 38 cm of row space.

ID	Plant Arrangement	Protein (%)			Oil (%)		
		Mean	Q: 2.5%	Q: 97.5%	Mean	Q: 2.5%	Q: 97.5%
KS21 Dryland	38Row	38	37	40	26	24	28
	13x13	39	38	41	25	23	27
	15x15	39	37	41	26	24	28
	20x20	38	36	39	25	23	27
KS21 Irrigated	38Row	38	36	39	26	24	28
	13x13	38	36	39	26	24	28
	15x15	38	36	39	26	24	28
	20x20	37	35	38	26	24	28
MS21 Dryland	38Row	42	41	43	23	22	26
	13x13	43	42	44	23	21	26
	15x15	42	41	43	24	22	26
	20x20	42	42	43	24	22	26

Continued on next page

Table 3.3: Continued from previous page

ID	Plant Arrangement	Protein (%)			Oil (%)		
		Mean	Q: 2.5%	Q: 97.5%	Mean	Q: 2.5%	Q: 97.5%
MS21 Irrigated	38Row	43	42	44	24	22	26
	13x13	43	42	44	24	22	26
	15x15	42	41	43	24	22	26
	20x20	43	42	44	24	22	26
KS22 Early	38Row	40	39	41	23	21	26
	13x13	41	40	42	23	21	25
	15x15	41	40	42	23	21	26
	20x20	41	39	42	23	21	26
KS22 Mid	38Row	38	36	40	23	21	26
	13x13	38	37	40	23	21	26
	15x15	37	36	39	24	22	26
	20x20	38	37	40	24	21	26
KS22 Late	38Row	43	42	44	24	21	27
	13x13	43	42	44	24	21	27

Continued on next page

Table 3.3: Continued from previous page

ID	Plant Arrangement	Protein (%)			Oil (%)		
		Mean	Q: 2.5%	Q: 97.5%	Mean	Q: 2.5%	Q: 97.5%
MS22 Early	15x15	43	43	44	24	21	27
	20x20	43	42	43	24	21	27
	38Row	42	41	43	23	20	27
	13x13	42	41	44	23	20	27
	15x15	42	40	43	23	20	27
MS22 Mid	20x20	42	41	44	23	20	27
	38Row	43	42	45	24	20	27
	13x13	43	42	45	24	20	27
	15x15	43	42	45	24	21	28
	20x20	43	41	44	24	21	28
MS22 Late	38Row	42	41	43	23	20	26
	13x13	42	41	43	22	20	26
	15x15	41	41	42	23	20	26
	20x20	41	40	42	23	20	27

Chapter 4

Spatio-temporal yield variation and precipitation within a field

* Published in *Computers and Electronics in Agriculture*

van Versendaal, E., Hernandez, C. M., Kyveryga, P., Hefley, T., Van De Woestyne, B. W., Prasad, P. V., & Ciampitti, I. A. (2025). Spatio-temporal yield variation and precipitation within a field. *Computers and Electronics in Agriculture*, 231, 109996.

<https://doi.org/10.1016/j.compag.2025.109996>

Abstract

In rainfed systems, precipitation significantly influence crop yield variability across years. Understanding the spatio-temporal yield variation within a field is crucial for enhancing farming production and profitability. This study aims to establish a data pipeline to quantify spatio-temporal yield variation within a field in response to precipitation, using multiple years of yield-precipitation records. Four field case studies – two with maize and two with soybean - each containing at least five years of yield monitor data of the same crop are used to evaluate the proposed data pipeline. Daily precipitation data of each case study were obtained from CHIRPS dataset and used to calculate derived variables representing

cumulative precipitation over different periods. Generalized Additive Models (GAMs) were fitted to analyze the within field variation of yield response to cumulative precipitation across all possible cumulative precipitation periods. The optimal cumulative precipitation period was selected based on the Akaike Information Criterion. Spatially varying response parameters derived from the GAM with the optimal cumulative precipitation period were then used to classify the varying yield responses to precipitation within the field using a Classification and Regression Tree methodology. The most relevant outcomes of this study include: (i) the optimal period of cumulative precipitation impacting yield change by field and crop type; (ii) within a field, some areas exhibit greater yield variability than others; and (iii) areas with equal high yield variability within the field may respond differently to precipitation. Future work could include the utilization of the spatio-temporal variation within a field for guiding input recommendations.

4.1 Introduction

Rainfed agriculture, the dominant farming practice, is crucial for worldwide food production. As global food demands grows, the pressure from climate changes, water shortages for irrigation, and land degradation are intensifying the need to increase crop yields under rainfed systems ([Rockström et al., 2010](#); [Shideed, 2017](#)). In the United States (US) Corn Belt region, which is responsible for over 85% of the country’s soybean (*Glycine max* L. Merrill)-maize (*Zea mays* L.) production, less than 37% of the area is irrigated. Typically, crop yield in rainfed systems are, on average, 50% or more below their potential due to water stress ([Lobell et al., 2009](#)). The productivity of these systems depends on field management, soil properties, environmental and weather conditions ([Wani et al., 2009](#)). Therefore, enhancing food production in rainfed agriculture could be possible by adopting existing technologies, such as site-specific input management, tailored to weather interactions ([Shideed, 2017](#); [Trevisan et al., 2021](#)).

Crop yield is often characterized by large spatial and temporal variability. Weather conditions are a key factor influencing crop yield variability across years ([Ahmed et al.,](#)

2022; Eck et al., 2020; Kukal & Irmak, 2018). Studies have shown that 36-43% of soybean yield variability and 31-42% of maize yield variability can be attributed to inter-annual climate variability (Ray et al., 2015). In particular, temperature and precipitation during the growing season significantly affect soybean and maize yields (Mourtzinis et al., 2017c; Pryor et al., 2014; Teasdale & Cavigelli, 2017). Among these factors, precipitation stress (drought or excessive precipitation) has a profound impact on crop growth development, and yield components, thereby influencing productivity (Dietz et al., 2021; Iqbal et al., 2020; Stewart & Lal, 2018). Although many studies demonstrate the varying responses of crop yields to precipitation, including how these responses differ by region and soil type (Seifert et al., 2017; Teasdale & Cavigelli, 2017; Xu et al., 2022), there is limited information on how yield responses to precipitation vary within individual farmer fields. This information can be crucial for implementing and improving site-specific management practices within the field.

The incorporation of Geographical Information Systems and yield monitor data has increased interest in site-specific variable rate input strategies, accounting for variations in the agricultural landscape to improve production (Bullock et al., 2002; Fulton et al., 2018; Goodchild, 1991; Griffin et al., 2008; Krause et al., 2020). Site-specific management involves managing crops at a smaller spatial scale than the entire field, addressing inefficiencies caused by uniform input rates that do not account for local needs (Moore & Tyndale-Biscoe, 1999; Nawar et al., 2017; Plant, 2001; Sawyer, 1994). Factors such as field topography, soil properties, and land use vary within the field, affecting crop productivity (Krause et al., 2020; Nawar et al., 2017). Additionally, these factors also influence the response of the crop to precipitation due to runoff, soil texture, waterlogged, and water storage (Kumhálová et al., 2011; Maestrini & Basso, 2018). It is well-documented that crop responses to inputs vary with seasonal precipitation conditions. Excessive precipitation can lead to nutrient runoff or leaching, especially with fertilizers, while insufficient precipitation might cause fertilizer accumulation in the soil (Basso et al., 2010; Billé & Rogné, 2022; Fu et al., 2014; R. Liu et al., 2014; Van Metre et al., 2016; Waller et al., 2021). Despite the critical role of precipitation in determining yield variation and input response, and despite advancements in site-specific technologies, a significant knowledge gap remains in understanding how yield responds to

varying precipitation within a field.

It is widely known that crop yield variability is driven by stable factors such as soil characteristics and variable factors including water (weather) and nutrient supply, pests, weeds, and disease infestation. These factors interact to affect yields across varying spatial and temporal scales, complicating the understanding of yield variability within a farm ([Kravchenko et al., 2005](#); [Simmonds et al., 2013](#); [Whelan & McBratney, 2000](#)). In this context, developing a new approach is critical for understanding yield variation within a field, reducing input costs, and enhancing production in rainfed systems. Accordingly, this study aims to establish a data pipeline to quantify spatio-temporal crop yield variation within a field in response to precipitation, using multiple year of yield-precipitation records. This, in-turn, will help to quantify and predict complex interactions of crop yield response and economic impact to different input application in variable rates within fields.

4.2 Materials and methods

4.2.1 Data collection and processing

Three US farms served as case studies to evaluate the proposed data pipeline. Farm 1, located in the state of Illinois, covers 20 hectares, and provided sixteen years of yield monitor data, comprising nine years of corn and seven years of soybean data (Case study 1 and 2, respectively). Farm 2, located in the state of Kansas, covers 4 hectares, and has five years of corn yield monitor data (Case study 3). Farm 3, located in the state of Ohio, encompasses 22 hectares, and has five years of soybean yield monitor data (Case study 4). The yield monitor data were cleaned using an outlier detection method followed by the local Moran procedure ([Córdoba, 2014](#)). Subsequently, the yield monitor data were interpolated using inverse distance weighting interpolation ([Gräler et al., 2016](#)) to create a 6x6 m grid, enabling the overlay of yield monitor data across years. Overall, Farm 2 exhibits greater variability in elevation and Compound Topographic Index compared to Farms 1 and 3 (Figure C.1).

Daily precipitation data of each farm were gathered from CHIRPS dataset ([Funk et al.,](#)

2014, 2015), and used to calculate derived variables representing cumulative precipitation over different periods. Cumulative precipitation refers to the total precipitation accumulated over a specific period of days. Since planting date was not available, these periods start from the beginning of the historical planting date (April 1st for maize and April 20th for soybean), and range from 10 to 200 days later, in 10-day increments (Baum et al., 2019; Deines et al., 2023; Mourtzinis et al., 2017c). This range was chosen to cover a broad variety of potential cumulative precipitation periods.

4.2.2 Statistical analysis

Spatially varying response model

The analysis of the on-farm variation of the yield response to cumulative precipitation was evaluated using Generalized Additive Models (GAMs, Wood, 2017), independently for each crop and farm. GAMs are flexible analytical frameworks for capturing complex and non-linear interactions within data. The form of a GAM is composed of linear sums of multi-dimensional basis functions and this is what allows complex relationships to be modelled. In summary, these models are ideal for large spatial datasets because they use smooth functions for spatial coordinates to capture trends and autocorrelation. They also efficiently handle complex interactions between spatial factors and covariables such as precipitation (Comber et al., 2024; Wood, 2003, 2017).

In this framework, the first level of the model (data layer) describes the crop yield (Y_i) for the i^{th} observation. The crop yield (Y_i) was assumed to follow a normal distribution with expected value μ_i and variance σ^2 :

$$Y_i \sim Normal(\mu_i, \sigma^2). \quad (4.1)$$

In the second level (underlying process model), μ_i was defined as:

$$\mu_i = \beta_0 + f(Lon_i, Lat_i, CP_i). \quad (4.2)$$

where β_0 is the intercept term; and $f(Lon_i, Lat_i, CP_i)$ is a tensor product smooth function of latitude (X_i), longitude (Y_i), and cumulative precipitation (CP_i) with basis functions thin plate spline for the spatial coordinates (Wood, 2003) and cubic regression spline for cumulative precipitation variable. The tensor product function allows the interactions between spatial variables and CP to be modeled, enabling the effect of precipitation to vary spatially within the field. The complexity of the model is controlled by smoothing parameters, where $k=(900,3)$ specifies the basis dimension limits, with 900 basis functions allocated for the interaction of spatial coordinates and 3 for cumulative precipitation.

The GAMs were fitted to the data using the ‘mgcv’ package (Wood, 2017) in the R software (R Core Team, 2021). Furthermore, to determine the cumulative precipitation period that best explained yield variation, we performed model selection based on the Akaike Information Criterion (AIC) (Bozdogan, 1987), selecting the model with the lowest value of AIC. This model with the corresponding period accounts for potential precipitation effects, such as waterlogging or drought stress, by incorporating interactions between the precipitation timing and duration within each field, recognizing that each field has its own optimal cumulative precipitation period. The performance of the model was evaluated using Relative Root Mean Square Error (RRMSE), Kling-Gupta model efficiency (KGE), Percentage Lack of Precision (PLP), and Percentage Lack of Accuracy (PLA), using the Metrica package (Correndo et al., 2022b).

Spatially varying response parameters were derived from the GAM. Maximum yield at each coordinate was determined based on the first derivative of yield response to cumulative precipitation: it was obtained at the point where the derivative is equal to zero, or at the point closest to zero if the maximum is not reached. For a flat yield response, the minimum extreme of cumulative precipitation was used. Additionally, the corresponding cumulative precipitation that maximizes yield was identified. Yield growth rate was calculated from the start of the increasing yield to the maximum yield, while decay rate was assessed from the maximum yield to the end of precipitation period. If the maximum yield is not reached, a rate of 0 is assigned, and it is noted in the dataset that the maximum yield response to cumulative precipitation was not achieved at that spatial coordinate.

Classify type of yield response to cumulative precipitation

To classify responses to cumulative precipitation within the field, Classification and Regression Tree (CART) methodology was used, which divides observations into increasingly homogeneous groups (Breiman et al., 2017). The response variable for this classification was the cumulative precipitation that maximizes yield. Explanatory variables included Maximum Yield, Growth Rate, Decay Rate, and a binary indicator of whether maximum yield was reached, derived from the GAM model. The CART models were implemented using the ‘rpart’ package (Therneau et al., 2015) in R (R Core Team, 2021), with a complexity parameter (cp) set at 0.1. This parameter ensured that new splits reduce the relative error by at least 10%, aiding to identify distinct response groups across the field.

4.3 Results

4.3.1 Model evaluation

As expected, the optimal cumulative precipitation period for explaining yield variation differed across case studies. For maize, the optimal period began on May 1st and lasted for 100 days in case study 1, and 70 days in case study 3. For soybean, the optimal period started on April 20th and lasted for 90 days in case study 2, and from May 1st for 90 days in case study 4 (Figure C.2, Figure C.3).

Overall, the models using the optimal cumulative precipitation period reflected a good performance, with a KGE values ranging from 0.66 to 0.92. Errors were primarily due to lack of precision (85–97 %) rather than lack of accuracy (3–15 %). The highest RRMSE were observed in the maize case studies, with values of 17 % and 13 % for case study 1 and case study 3, respectively. In contrast, the soybean case studies had lower RRMSE, with values of 11 % for case study 2 and 4 % for case study 4 (Figure 4.1).

4.3.2 Spatially varying response parameters

Figure 4.2 illustrates the ranges and spatial distribution of response parameters across field. In case study 4, 9.6 % of the area did not achieved the maximum yield within the range of cumulative precipitation observed over the years. In contrast, in the remaining case studies, less than 1 % of the area did not reach the maximum yield within the cumulative precipitation range (0.7 %, 0.1 %, and 0 % for case study 1, 2, and 3, respectively). In these areas, the maps illustrate as maximum yields the highest yield reached in the recorded cumulative precipitation range (Figure C.3, Figure 4.2). Maize case studies demonstrated the highest variability in the maximum yield reachable, with first and third interquartile ranges ($IQR_{25-75\%}$) of 12,486 to 14089 $kg\ ha^{-1}$ (total coefficient of variation, CV, equal to 10 %) for case study 1, and $IQR_{25-75\%}$ of 8960 to 11384 $kg\ ha^{-1}$ (total CV = 21 %) for case study 3. Conversely, soybean case studies exhibited the narrower ranges, with an $IQR_{25-75\%}$ of 4364 to 4804 $kg\ ha^{-1}$ (total CV = 6 %) and 5244 to 5661 $kg\ ha^{-1}$ (total CV = 7 %) for case study 2 and 4, respectively. Additionally, case studies 1 and 2, conducted in the same field, exhibited similar yield distributions (Figure 4.2).

The spatial distribution of cumulative precipitation that maximizes yield does not always correspond to the patterns of maximum yield observed across the field. This discrepancy arises because identical maximum yield peaks can be achieved under diverse precipitation conditions. For instance, in case study 1 (Maize), to attain a consistent maximum yield of approximately 12600 $kg\ ha^{-1}$, different field areas required varying levels of cumulative precipitation. One area needed 350 mm, another required more than ~ 400 mm, and a third reached this yield level with ~ 450 mm of cumulative precipitation (Figure C.4, Figure 4.2).

Growth rates for maize exhibited an $IQR_{25-75\%}$ of 25 to 34 $kg\ ha^{-1}\ mm^{-1}$ in case study 1 and 22 to 32 $kg\ ha^{-1}\ mm^{-1}$ in case study 3. For soybean, the growth rates were 13 to 20 $kg\ ha^{-1}\ mm^{-1}$ and 6 to 17 $kg\ ha^{-1}\ mm^{-1}$ in case study 2 and 4, respectively. Decay rate for maize showed an $IQR_{25-75\%}$ of 22 to 50 $kg\ ha^{-1}\ mm^{-1}$ and 13 to 21 $kg\ ha^{-1}\ mm^{-1}$ in case study 1 and 3. For soybean, decay rate ranged from 10 to 16 $kg\ ha^{-1}\ mm^{-1}$ and 18

to $25 \text{ kg ha}^{-1} \text{ mm}^{-1}$ in case study 1 and 3, respectively. Notably, maize exhibited greater variability in decay rates, with a CV values of 50 % in Case Study 1 and 39 % in Case Study 3.

4.3.3 Classify type of yield response to cumulative precipitation

The CART analysis revealed groups with varying type of responses to cumulative precipitation across each case study. Specifically, maize case studies identified four distinct groups, while soybean case studies contained three. Furthermore, the variables critical for defining distinct response groups in the CART analysis varied by case study. In maize (case study 1 and 3), groups were categorized based on the decay rate, whereas in soybean (case study 2 and 4), categorization was based on the growth rate (Figure C.5, Figure 4.3).

In case study 1, group 4 requires a greater cumulative precipitation to maximize yield, compared to groups 1, 2, and 3. Group 1 exhibits the highest decay rate, leading to a substantial yield reduction when precipitation exceeded the optimal range. Additionally, group 1 present the highest values of growth rate, while group 4 present the lowest values (Table 4.1, Figure 4.3).

In Case Study 3, Group 4 required the highest cumulative precipitation to achieve maximum yield, compared to groups 1,2, and 3. Group 1, which needed less cumulative precipitation to maximize yield, obtained a higher maximum yield, growth rate and decay rate. Conversely, Group 4 required more precipitation and achieved lower maximum yields, growth rates, and decay rate (Table 4.1, Figure 4.3).

In case study 2, all groups had similar maximum yields. However, group 1 needs less cumulative precipitation to achieve maximum yield, compared to groups 2 and 3. While groups 2 and 3 demonstrate yield increases up to the maximum yield followed by a decline, in general group 1 exhibits a negative yield response as cumulative precipitation increase. Furthermore, although Groups 1 and 2 have similar CV in yield (14 and 13%, respectively), their responses to cumulative precipitation were markedly different (Table 4.1, Figure 4.3).

Similar patterns occur in case study 4. All groups displayed a similar range of maximum

yields, and a uniform total CV in yield of 10%. However, group 1 needs a lower cumulative precipitation to maximize yield, compared to groups 2 and 3. Similar to case study 2, group 1 generally responded negatively to increases in cumulative precipitation, whereas group 2 and 3 shows yield increases up to their maximum followed by a decline (Table 4.1, Figure 4.3).

4.4 Discussion

This study develops a novel data pipeline to quantify the spatiotemporal yield variations within the field in response to precipitation. Despite extensive research over the past three decades into site-specific agricultural management zones – highlighting the importance of understanding yield variability and its underlying factors such as topography, soil texture, water storage, and historical management – there remains a gap knowledge understanding the interaction between crop, soil, and weather (Blackmore, 1994; Cook & Bramley, 1998; Hedley, 2015; Plant, 2001; Sawyer, 1994; Sudduth et al., 2015; Zhang et al., 2002). While previous studies have shown that precipitation significantly affects crop yield in dryland conditions and that the yield response to precipitation varies (Grassini et al., 2015; Li et al., 2019; Martinez-Feria & Basso, 2020; Teasdale & Cavigelli, 2017; Yu et al., 2021), there is still a lack of studies exploring the spatial yield variations within the field in response to precipitation.

Crop yield response to precipitation can substantially vary spatially within the field, as demonstrated in this study. Despite the importance of precipitation on yield variation, prescription maps for variable rate applications typically rely on soil maps and crop properties, utilizing yield monitor data and remote sensing, without integrating soil, management, and varying weather conditions (Córdoba et al., 2016; Guastaferro et al., 2010; Khosla et al., 2010; Mzuku et al., 2005; Nawar et al., 2017; Pilesjö et al., 2005). Some methodologies generate prescription maps based on whether the year is dry, intermediate, or wet, without demonstrating how the responses of each field area change with precipitation (Basso et al., 2009). On the other hand, several studies have started to evaluate the interaction between

soil, crop, and climate. For instance, areas within the field with high wetness index are significantly affected by waterlogged in wet years, resulting in lower yields, whereas they achieve higher yields in dry years (Maestrini & Basso, 2018). Another study notes that topography has less impact on yield variation in wet years on fine loamy soils (Halvorson & Doll, 1991). In this study, the spatially varying parameters not always follow the patterns of elevation and wetness index. A recent study (Riihimäki et al., 2021) revealed that the topographic wetness index was a modest proxy for soil moisture, but with severe limitations accounting for both high spatial and temporal resolution. Moreover, a previous study similarly identified more stable and unstable yield zones within field depending on precipitation, finding that areas prone to water stress are between 23–33 % below the field average in drought years and 26–33 % during flood years (Martinez-Feria & Basso, 2020). This aligns with findings from this study, which also reveals that two equally unstable zones within the field may respond differently to precipitation.

Identifying field areas with varying yield responses to precipitation and understanding how those site-specific responses can impact the effective use of inputs (Anderson et al., 2016; R. Liu et al., 2014; Ploschuk et al., 2022) can enhance prescription maps and field profitability. Studies have shown that 5–75 % of within field areas suffer economic losses due to precipitation conditions, with significant losses noted during wet springs where precipitation exceeds 300 mm. In such scenarios, each additional 10 mm of May/June precipitation can reduce median profits by \$50-\$120 per hectare in maize fields, whereas soybean fields generally show less economic impact (Fey et al., 2016). Furthermore, areas prone to flooding or drought typically exhibit a decline in productivity due to reduced nutrient uptake by plants. Overfertilizing crops under stress will not improve plant growth or may even inhibit it. Furthermore, extreme precipitation can lead to runoff or leaching of applied fertilizers, while insufficient precipitation may result in fertilizer accumulation in the soil (Billé & Rogna, 2022; R. Liu et al., 2014; Rouphael et al., 2012; Van Metre et al., 2016; Waller et al., 2021). Additionally, even minor differences in landscape position can significantly impact water movement across the field, which, in turn, affects nitrogen loss and altering optimal plant density (Assefa et al., 2016; Haarhoff & Swanepoel, 2018; Rogovska et al., 2023). By

identifying areas where precipitation positively or negatively impacts the field, it is possible to strategically plant different genotypes across the field based on their resistance to water-logging or drought, optimizing in this way the yield (Rotili et al., 2020). All previous studies emphasize the importance of precipitation for inputs management, highlighting the significance of this study. By identifying areas with variable and stable productivity in response to precipitation changes, farmers can implement targeted input management strategies. Such strategies help minimize risks, reduce costs, and maximize yields (Hatfield & Walthall, 2015; McArthur & McCord, 2017).

Some limitations of this study were: i) the lack of historical observed weather data in each specific farm location, ii) a limited number of farms with extensive (at least five years or more) yield monitor data, iii) limited number of years with high-resolution remote sensing data, and iv) lack of available data on soil physical and chemical properties, such as soil texture. Despite these challenges, the data pipeline approach adopted in this study effectively captures our objectives and serves as a starting point for understanding spatio-temporal yield variability in response to precipitation. Future research should: i) connect these spatio-temporal variation of yield responses to precipitation with input rates to further refine variable input strategies within fields and ii) use the maps obtained with this data pipeline to precisely sample soil physical and chemical properties to correctly understand variability the implication of soil factors within the field.

4.5 Conclusion

This study underscores the potential benefits for farmers in understanding spatio-temporal yield variations in response to precipitation. The most relevant outcomes of this study include: (i) the optimal period of cumulative precipitation impacting yield varies by field and crop type; (ii) within a field, some areas exhibit greater yield variability than others; and (iii) areas with similar high yield variability within the field may respond differently to precipitation. Future investigation could expand this approach by integrating high-resolution remote sensing data to develop an improved method for upscaling field characterization to

more fields with the potential for guiding environmental characterization and the decision of prescription for critical farming inputs.

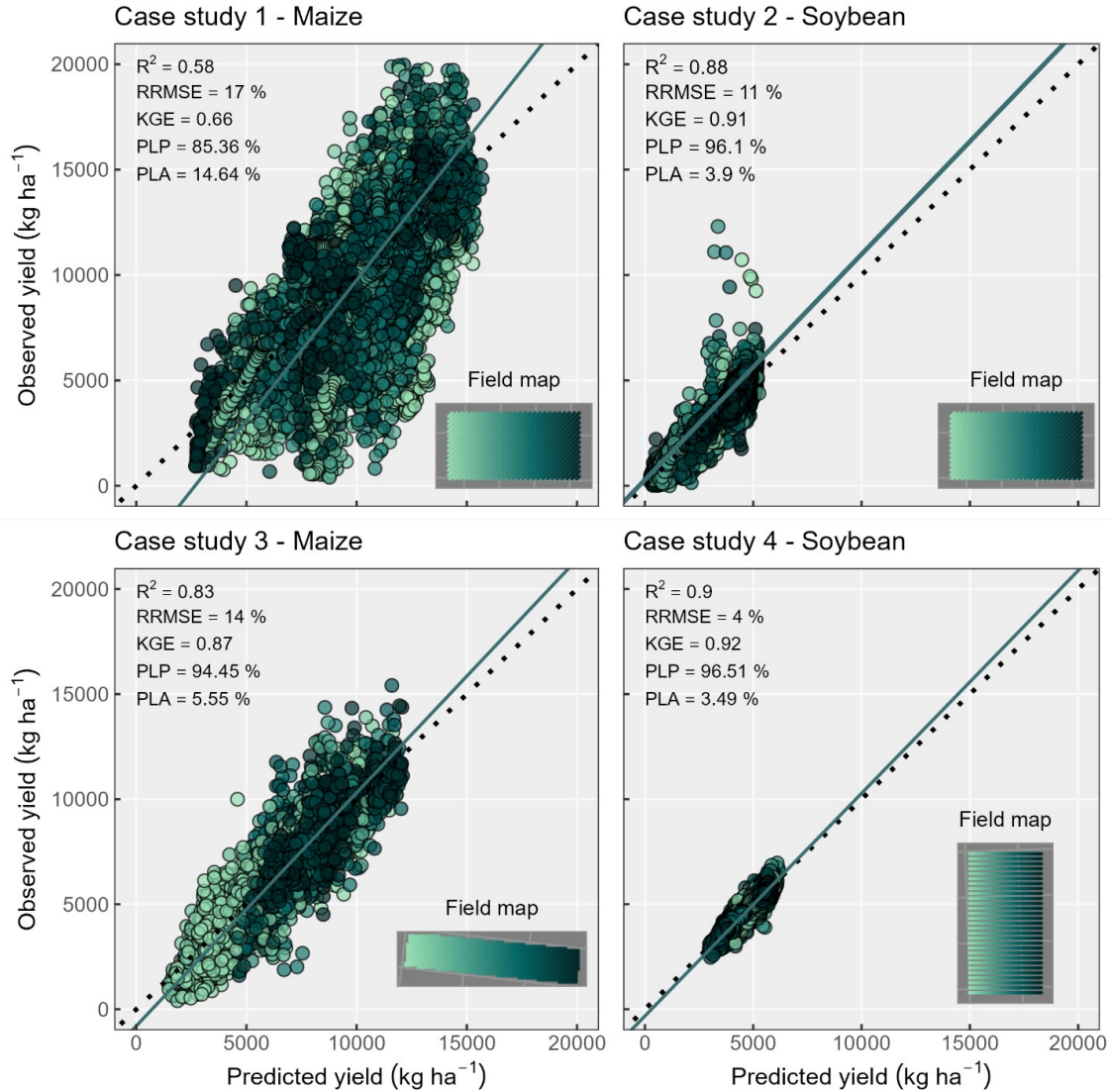


Figure 4.1: Model performance across case studies (cases 1, 3 for maize crop, and 2, 4 for soybean). Each panel represents a different case study, displaying yield data at specific spatial coordinates within the field for a particular year. Points are color-coded to uniquely identify each spatial coordinate. The bottom-right corner of each panel includes a field map, where colors correspond to those of the points, illustrating the spatial distribution (within field variation) of the yield data.

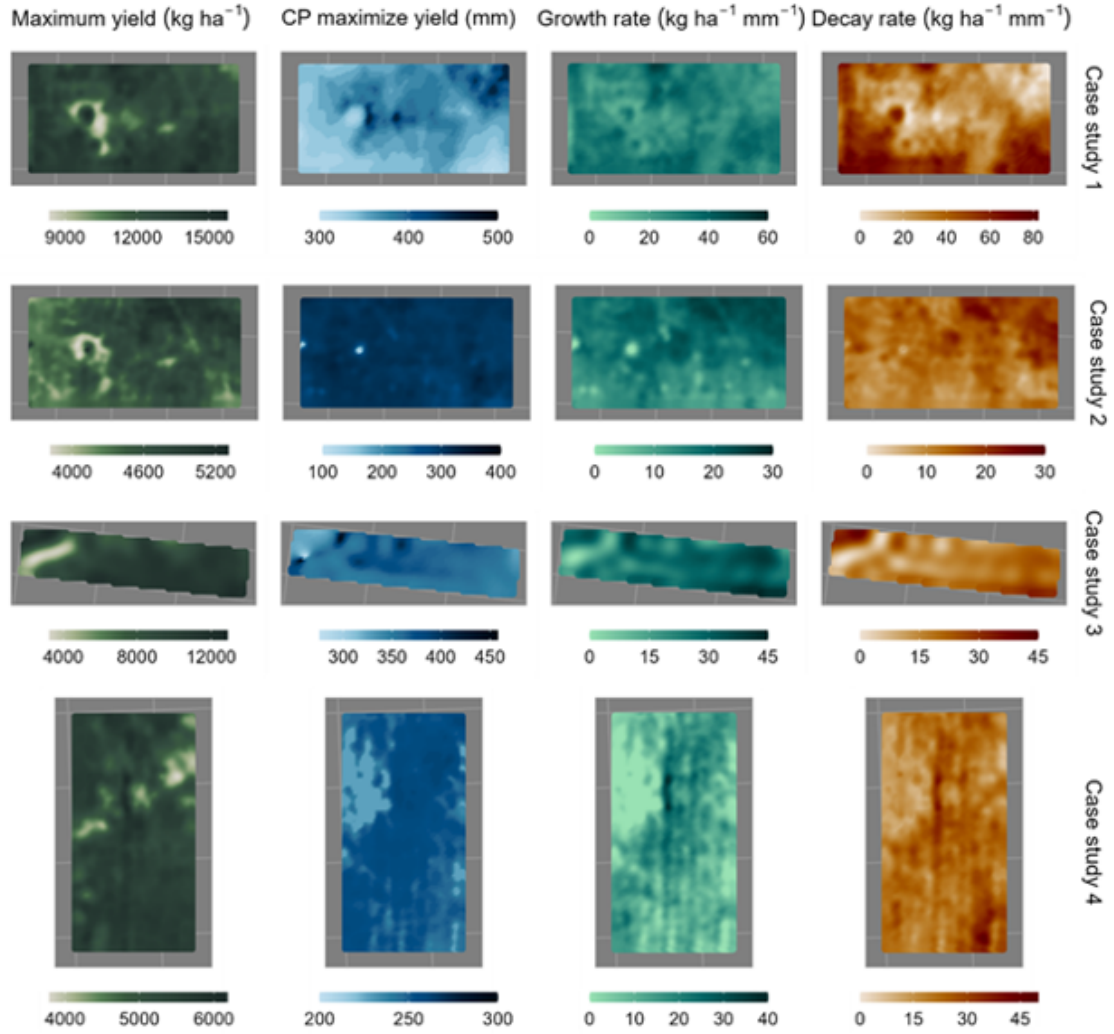


Figure 4.2: Maps of the spatially varying response parameters to optimal precipitation across fields. Columns represent each parameter of the model. Rows represent each case study. In all the parameters, darker the color higher the parameter's value. CP: cumulative precipitation. Case study 1 and 3 correspond to maize crop; Case study 2 and 4 correspond to soybean crop.

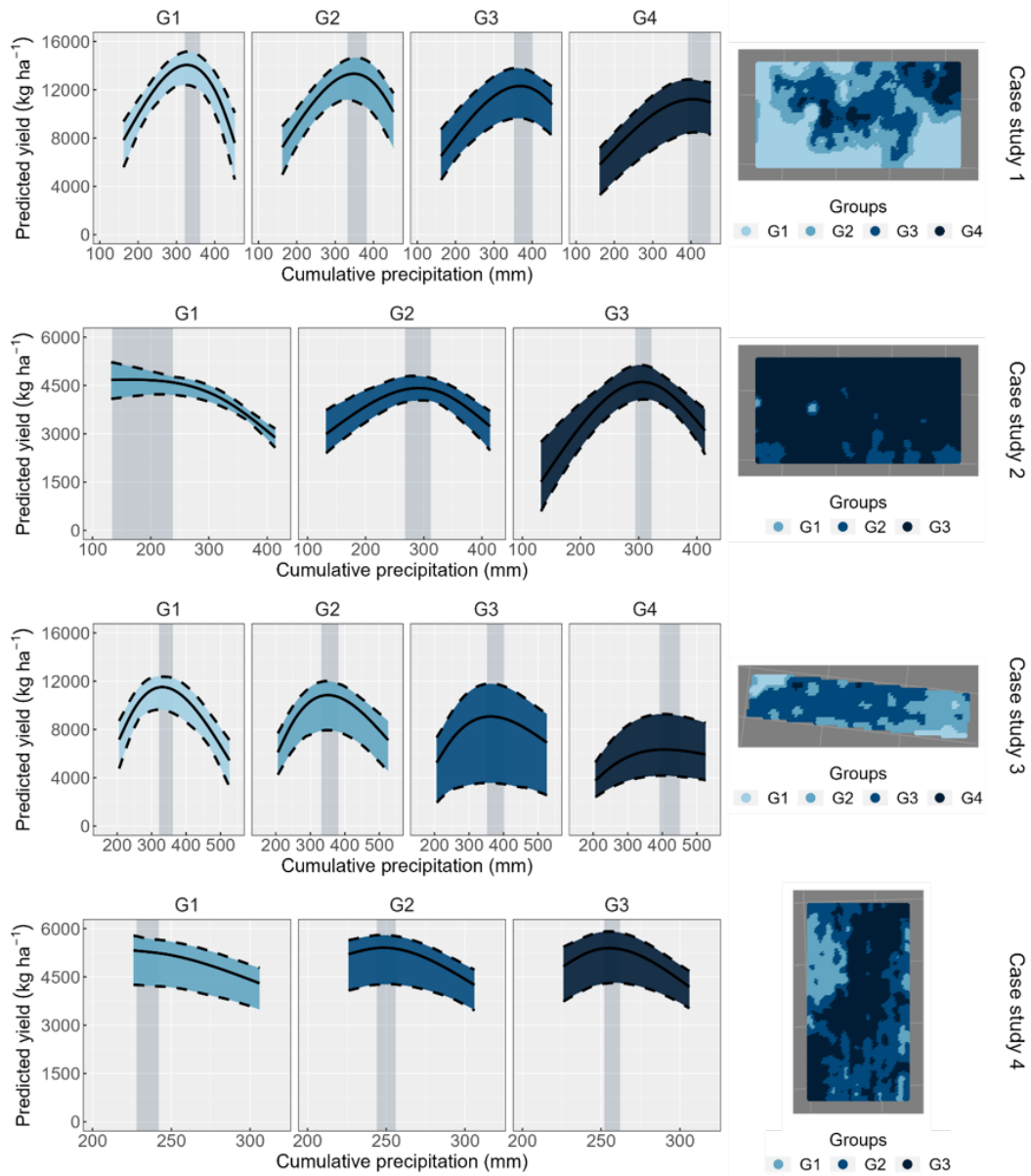


Figure 4.3: Curves of response of yield to cumulative precipitation, each column correspond to a group (left) that were identified in the classification and regression analyses. The response variable for this classification was the cumulative precipitation that maximizes yield. Field map demonstrating spatial distribution of each group (right) Each row corresponds to each case study. Each color represents a group. Solid black line corresponds to the mean yield. Dashed black lines correspond to the 95% confidence intervals.

Table 4.1: Summary of Spatially Varying Parameters by group and case study. Includes mean and interquartile ranges ($IQR_{25-75\%}$) for Maximum yield ($kg\ ha^{-1}$); Cumulative Precipitation (CP) maximize yield (mm); Growth rate ($kg\ ha^{-1}\ mm^{-1}$); and Decay rate ($kg\ ha^{-1}\ mm^{-1}$).

Case Study	Group	Maximum Yield		CP Maximize Yield		Growth Rate		Decay Rate	
		Mean	$IQR_{25-75\%}$	Mean	$IQR_{25-75\%}$	Mean	$IQR_{25-75\%}$	Mean	$IQR_{25-75\%}$
1	G1	14091	13769-14498	337	332-342	34	32-37	56	49-62
	G2	13353	12929-13967	358	352-362	30	27-33	34	30-37
	G3	12339	11869-13108	378	372-382	26	23-29	20	17-25
	G4	11259	10327-12550	417	402-422	21	19-23	7	4-10
2	G1	4752	4590-4897	179	134-224	1	0-2	8	7-9
	G2	4428	4235-4665	293	286-302	9	8-10	10	8-11
	G3	4614	4380-4817	308	304-312	18	15-20	14	11-17
3	G1	11665	11355-12091	333	322-346	34	30-39	32	29-34
	G2	10901	10552-11535	353	350-360	32	29-35	21	19-23
	G3	9119	8012-10960	365	360-372	24	20-29	13	11-16
	G4	6104	4497-7858	413	400-430	12	9-15	3	2-4
4	G1	5315	5080-5607	231	228-234	0	0-0	14	12-15
	G2	5410	5290-5653	250	248-252	8	6-10	21	17-23
	G3	5399	5229-5685	257	256-258	17	14-20	25	22-27

Chapter 5

Final remarks

Generating actionable information on crop management to enhance yields, and leveraging available technological tools at multiple scales, is essential for achieving global food security and providing viable options for farmers. This dissertation focuses on improving crop management strategies to increase yields and reduce variability using advanced data-informed decisions. To generate this information, we examined three scales and corresponding tools: (i) at the regional level, using crop modelling (APSIM) (Chapter 2); (ii) at the site level, using a Bayesian approach, the Canopeo tool and crop modelling (DSSAT) (Chapter 3); and (iii) at the field level, by analyzing multiple years of geo-referenced yield monitor data from combine harvesters and performing spatio-temporal analysis (Chapter 4).

Chapter 2 demonstrates the potential benefits for farmers in selecting the optimal combinations of planting dates by maturity group for soybean yield. Indicating that this combination strategy varies across different regions within Kansas. The main findings of this chapter were that under current weather conditions farmers can: i) maximize soybean yield by planting in April to early May, and ii) find more yield stability (less variability) over time by planting later, but with a yield penalty for maximum yield. For future weather scenarios: iii) yield losses can be expected, but selecting the right maturity group x planting date combinations could be used as a mitigation practice, and iv) later maturity groups (IV, V and VI) may result in lower yield loss for late planting dates compared to early planting

dates.

Chapter 3 examines the effects of equidistant versus non-equidistant plant spatial arrangements in soybean at the site level. The results indicate that (i) in most trials (location management), equidistant plant spatial arrangement does not offer yield advantages over non-equidistant arrangement, (ii) protein and oil seed concentration remained constant across different plant arrangements, and (iii) canopy parameters had no effect on soybean yield variation and seed quality traits.

Lastly, chapter 4 highlights the potential benefits for farmers in understanding spatio-temporal yield variations in response to precipitation at the within-field scale. The main key findings of this chapter include: (i) the optimal period of cumulative precipitation impacting yield varies by field and crop type; (ii) some areas exhibit greater yield variability than others; and (iii) areas with similar high yield variability within the field may respond differently to precipitation.

A few of the major research limitations throughout this dissertation were related to (i) lack of observed soil data, especially soil physical and chemical properties (Chapters 2, 3, and 4), (ii) limited observed data of crop phenology for model calibration (Chapter 2) and the absence of information on biomass partitioning into vegetative and reproductive structures needed to assess soybean plant compensation (Chapter 3), (iii) limited number of treatments evaluating non-equidistant patterns (Chapter 3), and (iv) insufficient geographical representation or environments (Chapter 2) and a small number of farms with at least five years of yield monitor data (Chapter 4).

Future research could enhance the information generated by incorporating observed phenological data to ensure proper model calibration and validation (Chapter 2), expanding evaluation of soybean plant spatial arrangements, particularly at consistent densities (chapter 3), and integrating the findings with remote sensing data into spatio-temporal analyses (Chapter 4). This integration would improve the understanding of regional, local, and within-field variability for each crop, addressing data gaps where field observations are limited. Additionally, incorporating economic and social analyses would provide a more comprehensive information base for farmers, enabling more effective data-driven decision-making (Chapters

2, 3, and 4).

References

- Abdala, L. J., Tamagno, S., Ruiz, A., Schwalbert, R. A., Correndo, A. A., & Martin, N. (2025). Yield environment changes the ranking of soybean genotypes. *Field Crops Research*, 321, 109661. doi: [10.1016/j.fcr.2024.109661](https://doi.org/10.1016/j.fcr.2024.109661)
- Ahmed, M., Hayat, R., Ahmad, M., ul Hassan, M., Kheir, A. M. S., ul Hassan, F., ... Ahmad, S. (2022). Impact of climate change on dryland agricultural systems: A review of current status, potentials, and further work need. *International Journal of Plant Production*, 16, 341–363. doi: [10.1007/s42106-022-00197-1](https://doi.org/10.1007/s42106-022-00197-1)
- Anderson, W., Johansen, C., & Siddique, K. H. M. (2016). Addressing the yield gap in rainfed crops: a review. *Agronomy for Sustainable Development*, 36, 18. doi: [10.1007/s13593-015-0341-y](https://doi.org/10.1007/s13593-015-0341-y)
- Andrade, F. H., & Abbate, P. E. (2005). Response of maize and soybean to variability in stand uniformity. *Agronomy Journal*, 97(4), 1263–1269. doi: [10.2134/agronj2005.0006](https://doi.org/10.2134/agronj2005.0006)
- Andrade, J. F., Rattalino Edreira, J. I., Mourtzinis, S., Conley, S. P., Ciampitti, I. A., Dunphy, J. E., ... others (2019). Assessing the influence of row spacing on soybean yield using experimental and producer survey data. *Field Crops Research*, 230, 98–106. doi: [10.1016/j.fcr.2018.10.014](https://doi.org/10.1016/j.fcr.2018.10.014)
- Archontoulis, S. V., & Miguez, F. E. (2015). Nonlinear regression models and applications in agricultural research. *Agronomy Journal*, 107(2), 786–798. doi: [10.2134/agronj2012.0506](https://doi.org/10.2134/agronj2012.0506)
- Archontoulis, S. V., Miguez, F. E., & Moore, K. J. (2014). A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM plant model: Application to soybean. *Environ. Model. Softw.*, 62, 465–477. Retrieved from

<https://www.sciencedirect.com/science/article/pii/S1364815214001133> doi:
10.1016/j.envsoft.2014.07.1133

Arrouays, D., Grundy, M. G., Hartemink, A. E., Hempel, J. W., Heuvelink, G. B. M., Hong, S. Y., ... Zhang, G.-L. (2014). Globalsoilmap. In *Advances in agronomy* (pp. 93–134). Elsevier. doi: [10.1016/b978-0-12-800137-0.00003-0](https://doi.org/10.1016/b978-0-12-800137-0.00003-0)

Assefa, Y., Vara Prasad, P. V., Carter, P., Hinds, M., Bhalla, G., Schon, R., ... Ciampitti, I. A. (2016). Yield responses to planting density for us modern corn hybrids: A synthesis-analysis. *Crop Science*, *56*, 2802–2817. doi: [10.2135/cropsci2016.04.0215](https://doi.org/10.2135/cropsci2016.04.0215)

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., ... Wolf, J. (2013). Uncertainty in simulating wheat yields under climate change. *Nature Climate Change*, *3*, 827–832. doi: [10.1038/nclimate1916](https://doi.org/10.1038/nclimate1916)

Basso, B., Cammarano, D., Chen, D., Cafiero, G., Amato, M., Bitella, G., ... Basso, F. (2009). Landscape position and precipitation effects on spatial variability of wheat yield and grain protein in southern Italy. *Journal of Agronomy and Crop Science*, *195*, 301–312. doi: [10.1111/j.1439-037x.2008.00351.x](https://doi.org/10.1111/j.1439-037x.2008.00351.x)

Basso, B., Cammarano, D., Troccoli, A., Chen, D., & Ritchie, J. T. (2010). Long-term wheat response to nitrogen in a rainfed mediterranean environment: Field data and simulation analysis. *European Journal of Agronomy*, *33*, 132–138. doi: [10.1016/j.eja.2010.04.004](https://doi.org/10.1016/j.eja.2010.04.004)

Basso, B., Ritchie, J. T., Pierce, F. J., Braga, R. P., & Jones, J. W. (2001). Spatial validation of crop models for precision agriculture. *Agricultural Systems*, *68*, 97–112. doi: [10.1016/s0308-521x\(00\)00063-9](https://doi.org/10.1016/s0308-521x(00)00063-9)

Bastidas, A. M., Setiyono, T. D., Dobermann, A., Cassman, K. G., Elmore, R. W., Graef, G. L., & Specht, J. E. (2008). Soybean sowing date: The vegetative, reproductive, and agronomic impacts. *Crop Sci.*, *48*(2), 727–740. doi: [10.2135/cropsci2006.05.0292](https://doi.org/10.2135/cropsci2006.05.0292)

- Bastos, L. M., Carciochi, W., Lollato, R. P., Jaenisch, B. R., Rezende, C. R., Schwalbert, R., ... Ciampitti, I. A. (2020). Winter wheat yield response to plant density as a function of yield environment and tillering potential: A review and field studies. *Frontiers in Plant Science*, *11*, 54. doi: [10.3389/fpls.2020.00054](https://doi.org/10.3389/fpls.2020.00054)
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, *67*(1). doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)
- Battisti, R., & Sentelhas, P. C. (2015). Drought tolerance of brazilian soybean cultivars simulated by a simple agrometeorological yield model. *Exp. Agric.*, *51*(2), 285–298. doi: [10.1017/s0014479714000283](https://doi.org/10.1017/s0014479714000283)
- Baum, M. E., Archontoulis, S. V., & Licht, M. A. (2019). Planting date, hybrid maturity, and weather effects on maize yield and crop stage. *Agronomy Journal*, *111*, 303–313. doi: [10.2134/agronj2018.04.0297](https://doi.org/10.2134/agronj2018.04.0297)
- Bellaloui, N., Smith, J. R., Gillen, A. M., Fisher, D. K., & Mengistu, A. (2012). Effect of shade on seed protein, oil, fatty acids, and minerals in soybean lines varying in seed germinability in the early soybean production system. *American Journal of Plant Sciences*, *3*(1), 84–95. doi: [10.4236/ajps.2012.31008](https://doi.org/10.4236/ajps.2012.31008)
- Bezdek, J. C., Ehrlich, R., & Full, W. (1984). Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*, 191–203.
- Billé, A. G., & Rogna, M. (2022). The effect of weather conditions on fertilizer applications: A spatial dynamic panel data analysis. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *185*, 3–36. doi: [10.1111/rssa.12709](https://doi.org/10.1111/rssa.12709)
- Blackmore, S. (1994). Precision farming: An introduction. *Outlook on Agriculture*, *23*, 275–280. doi: [10.1177/003072709402300407](https://doi.org/10.1177/003072709402300407)

- Board, J. (2000). Light interception efficiency and light quality affect yield compensation of soybean at low plant populations. *Crop Science*, 40(5), 1285–1294. doi: [10.2135/cropsci2000.4051285x](https://doi.org/10.2135/cropsci2000.4051285x)
- Boote, K. J., Jones, J. W., Hoogenboom, G., & Pickering, N. B. (1998). The cropgro model for grain legumes. In G. Y. Tsuji, G. Hoogenboom, & P. K. Thornton (Eds.), *Understanding options for agricultural production* (Vol. 7, pp. 99–128). Springer Netherlands. doi: [10.1007/978-94-017-3624-4_6](https://doi.org/10.1007/978-94-017-3624-4_6)
- Boydak, E., Alpaslan, M., Hayta, M., Gerçek, S., & Simsek, M. (2002). Seed composition of soybeans grown in the harran region of turkey as affected by row spacing and irrigation. *Journal of Agricultural and Food Chemistry*, 50(16), 4718–4720. doi: [10.1021/jf0255331](https://doi.org/10.1021/jf0255331)
- Boyer, C. N., Stefanini, M., Larson, J. A., Smith, S. A., Mengistu, A., & Bellaloui, N. (2015). Profitability and risk analysis of soybean planting date by maturity group. *Agron. J.*, 107(6), 2253–2262. doi: [10.2134/agronj15.0148](https://doi.org/10.2134/agronj15.0148)
- Bozdogan, H. (1987). Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. *Psychometrika*, 52, 345–370. doi: [10.1007/bf02294361](https://doi.org/10.1007/bf02294361)
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification and regression trees* (1st ed.). Routledge. doi: [10.1201/9781315139470](https://doi.org/10.1201/9781315139470)
- Brown, H. E., Huth, N. I., Holzworth, D. P., Teixeira, E. I., Zyskowski, R. F., Hargreaves, J. N. G., & Moot, D. J. (2014). Plant modelling framework: Software for building and running crop models on the apsim platform. *Environ. Model. Softw.*, 62, 385–398.
- Brunsell, N. A., Jones, A. R., Jackson, T. L., & Feddema, J. J. (2010). Seasonal trends in air temperature and precipitation in ipcc ar4 gcm output for kansas, usa: Evaluation and implications. *Int. J. Climatol.*, 30(8), 1178–1193. doi: [10.1002/joc.1958](https://doi.org/10.1002/joc.1958)

- Bullock, D. S., Khan, S., & Rayburn, A. (1998). Soybean yield response to narrow rows is largely due to enhanced early growth. *Crop Science*, 38(4), 1011–1016. doi: [10.2135/crop-sci1998.0011183x003800040021x](https://doi.org/10.2135/crop-sci1998.0011183x003800040021x)
- Bullock, D. S., Lowenberg-DeBoer, J., & Swinton, S. M. (2002). Adding value to spatially managed inputs by understanding site-specific yield response. *Agricultural Economics*, 27, 233–245. doi: [10.1111/j.1574-0862.2002.tb00119.x](https://doi.org/10.1111/j.1574-0862.2002.tb00119.x)
- Bürkner, P.-C. (2021). Bayesian item response modeling in r with brms and stan. *Journal of Statistical Software*, 100(5). doi: [10.18637/jss.v100.i05](https://doi.org/10.18637/jss.v100.i05)
- Carciochi, W. D., Sadras, V. O., Pagani, A., & Ciampitti, I. A. (2020). Co-limitation and stoichiometry capture the interacting effects of nitrogen and sulfur on maize yield and nutrient use efficiency. *Eur. J. Agron.*, 113, 125973. doi: [10.1016/j.eja.2019.125973](https://doi.org/10.1016/j.eja.2019.125973)
- Carciochi, W. D., Schwalbert, R., Andrade, F. H., Corassa, G. M., Carter, P., Gaspar, A. P., ... Ciampitti, I. A. (2019). Soybean seed yield response to plant density by yield environment in north america. *Agronomy Journal*, 111(4), 1923–1932. doi: [10.2134/agronj2018.10.0635](https://doi.org/10.2134/agronj2018.10.0635)
- Carpenter, A. C., & Board, J. E. (1997). Branch yield components controlling soybean yield stability across plant populations. *Crop Science*, 37(3), 885–891. doi: [10.2135/crop-sci1997.0011183x003700030031x](https://doi.org/10.2135/crop-sci1997.0011183x003700030031x)
- Cassman, K. G., & Grassini, P. (2020). A global perspective on sustainable intensification research. *Nature Sustainability*, 3, 262–268. doi: [10.1038/s41893-020-0507-8](https://doi.org/10.1038/s41893-020-0507-8)
- Chen, G., & Wiatrak, P. (2010). Soybean development and yield are influenced by planting date and environmental conditions in the southeastern coastal plain, united states. *Agron. J.*, 102(6), 1731–1737. doi: [10.2134/agronj2010.0219](https://doi.org/10.2134/agronj2010.0219)

- Chung, J., Babka, H. L., Graef, G. L., Staswick, P. E., Lee, D. J., Cregan, P. B., . . . Specht, J. E. (2003). The seed protein, oil, and yield qtl on soybean linkage group i. *Crop Science*, 43(3), 1053–1067. doi: [10.2135/cropsci2003.1053](https://doi.org/10.2135/cropsci2003.1053)
- Ciampitti, I., Correndo, A., & van Versendaal, E. (2022). *Soybean planting date and maturity group selection*. https://webapp.agron.ksu.edu/agr_social/m_eu_article.throck?article_id=3093&eu_id=490. (Last accessed May 2022)
- Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy*, 40, 79–96. doi: [10.1093/aepp/ppx056](https://doi.org/10.1093/aepp/ppx056)
- Comber, A., Harris, P., & Brunsdon, C. (2024). Multiscale spatially varying coefficient modelling using a geographical gaussian process gam. *International Journal of Geographical Information Science*, 38, 27–47. doi: [10.1080/13658816.2023.2270285](https://doi.org/10.1080/13658816.2023.2270285)
- Cook, S. E., & Bramley, R. G. V. (1998). Precision agriculture — opportunities, benefits and pitfalls of site-specific crop management in australia. *Australian Journal of Experimental Agriculture*, 38, 753. doi: [10.1071/ea97156](https://doi.org/10.1071/ea97156)
- Correndo, A., Adey, E., Rosso, L. H. M., Tremblay, N., Prasad, P. V. V., Du, J., & Ciampitti, I. A. (2022a). Footprints of corn nitrogen management on the following soybean crop. *Agron. J.*, 114(2), 1475–1488. doi: [10.1002/agj2.21023](https://doi.org/10.1002/agj2.21023)
- Correndo, A., Moro Rosso, L., Schwalbert, R., Hernandez, C., Bastos, L., Holzworth, D. P., & Ciampitti, I. (2022b). *metrica: Prediction performance metrics*. (R package version 1.2.3, <https://CRAN.R-project.org/package=metrica>)
- Cox, W. J., & Cherney, J. H. (2011). Growth and yield responses of soybean to row spacing and seeding rate. *Agronomy Journal*, 103(1), 123–128. doi: [10.2134/agronj2010.0316](https://doi.org/10.2134/agronj2010.0316)

- Cox, W. J., Cherney, J. H., & Shields, E. (2010). Soybeans compensate at low seedling rates but not at high thinning rates. *Agronomy Journal*, 102(4), 1238–1243. doi: [10.2134/agronj2010.0047](https://doi.org/10.2134/agronj2010.0047)
- Cravero, A., Pardo, S., Sepúlveda, S., & Muñoz, L. (2022). Challenges to use machine learning in agricultural big data: A systematic literature review. *Agronomy*, 12, 748. doi: [10.3390/agronomy12030748](https://doi.org/10.3390/agronomy12030748)
- Córdoba, M. A. (2014). *Herramientas estadísticas para el monitoreo y uso de la variabilidad espacial del rendimiento y propiedades de suelo intralote* (Unpublished doctoral dissertation). Universidad Nacional de Córdoba.
- Córdoba, M. A., Bruno, C. I., Costa, J. L., Peralta, N. R., & Balzarini, M. G. (2016). Protocol for multivariate homogeneous zone delineation in precision agriculture. *Biosystems Engineering*, 143, 95–107. doi: [10.1016/j.biosystemseng.2015.12.008](https://doi.org/10.1016/j.biosystemseng.2015.12.008)
- Damgaard, C. F., & Irvine, K. M. (2019). Using the beta distribution to analyse plant cover data. *Journal of Ecology*, 107(6), 2747–2759. doi: [10.1111/1365-2745.13200](https://doi.org/10.1111/1365-2745.13200)
- de Borja Reis, A. F., Tamagno, S., Moro Rosso, L. H., Ortez, O. A., Naeve, S., & Ciampitti, I. A. (2020). Historical trend on seed amino acid concentration does not follow protein changes in soybeans. *Sci. Rep.*, 10(1), 1–10. doi: [10.1038/s41598-020-74734-1](https://doi.org/10.1038/s41598-020-74734-1)
- de Bruin, J. L., & Pedersen, P. (2008). Soybean seed yield response to planting date and seedling rate in the upper midwest. *Agron. J.*, 100(3), 696–703. doi: [10.2134/agronj2007.0115](https://doi.org/10.2134/agronj2007.0115)
- de Bruin, J. L., & Pedersen, P. (2009). Growth, yield, and yield component changes among old and new soybean cultivars. *Agronomy Journal*, 101(1), 124–130. doi: [10.2134/agronj2008.0187](https://doi.org/10.2134/agronj2008.0187)

- Deines, J. M., Swatantran, A., Ye, D., Myers, B., Archontoulis, S., & Lobell, D. B. (2023). Field-scale dynamics of planting dates in the us corn belt from 2000 to 2020. *Remote Sensing of Environment*, 291, 113551. doi: [10.1016/j.rse.2023.113551](https://doi.org/10.1016/j.rse.2023.113551)
- Devlin, D. L., Fjell, D. L., Shroyer, J. P., Gordon, W. B., Marsh, B. H., Maddux, L. D., ... Duncan, S. R. (1995). Row spacing and seeding rates for soybean in low and high yielding environments. *Journal of Production Agriculture*, 8(2), 215–222. doi: [10.2134/jpa1995.0215](https://doi.org/10.2134/jpa1995.0215)
- Dietz, K., Zörb, C., & Geilfus, C. (2021). Drought and crop yield. *Plant Biology Journal*, 23, 881–893. doi: [10.1111/plb.13304](https://doi.org/10.1111/plb.13304)
- do Rio, A., Sentelhas, P. C., Farias, J. R. B., Sibaldelli, R. N. R., & Ferreira, R. C. (2016). Alternative sowing dates as a mitigation measure to reduce climate change impacts on soybean yields in southern brazil. *Int. J. Climatol.*, 36(11), 3664–3672. doi: [10.1002/joc.4583](https://doi.org/10.1002/joc.4583)
- Duncan, W. G. (1986). Planting patterns and soybean yields 1. *Crop Science*, 26(3), 584–588. doi: [10.2135/cropsci1986.0011183x002600030033x](https://doi.org/10.2135/cropsci1986.0011183x002600030033x)
- Eck, M. A., Murray, A. R., Ward, A. R., & Konrad, C. E. (2020). Influence of growing season temperature and precipitation anomalies on crop yield in the southeastern united states. *Agricultural and Forest Meteorology*, 291, 108053. doi: [10.1016/j.agrformet.2020.108053](https://doi.org/10.1016/j.agrformet.2020.108053)
- Edwards, J. T., Purcell, L. C., & Karcher, D. E. (2005). Soybean yield and biomass responses to increasing plant population among diverse maturity groups: Ii. light interception and utilization. *Crop Science*, 45(5), 1778–1785. doi: [10.2135/cropsci2004.0570](https://doi.org/10.2135/cropsci2004.0570)
- Egli, D. B. (1994). Mechanisms responsible for soybean yield response to equidistant planting patterns. *Agronomy Journal*, 86(6), 1046–1049. doi: [10.2134/agronj1994.00021962008600060021x](https://doi.org/10.2134/agronj1994.00021962008600060021x)

- Egli, D. B., & Bruening, W. (1992). Planting date and soybean yield: evaluation of environmental effects with a crop simulation model: Soygro. *Agric. For. Meteorol.*, 62(1-2), 19–29. doi: [10.1016/0168-1923\(92\)90003-m](https://doi.org/10.1016/0168-1923(92)90003-m)
- Egli, D. B., & Bruening, W. P. (2000). Potential of early-maturing soybean cultivars in late plantings. *Agron. J.*, 92(3), 532–537. doi: [10.2134/agronj2000.923532x](https://doi.org/10.2134/agronj2000.923532x)
- Egli, D. B., & Cornelius, P. L. (2009). A regional analysis of the response of soybean yield to planting date. *Agron. J.*, 101(2), 330–335. doi: [10.2134/agronj2008.0148](https://doi.org/10.2134/agronj2008.0148)
- Egli, D. B., & TeKrony, D. M. (1996). Seedbed conditions and prediction of field emergence of soybean seed. *J. Prod. Agric.*, 9(3), 365–370. doi: [10.2134/jpa1996.0365](https://doi.org/10.2134/jpa1996.0365)
- Elli, E. F., Ciampitti, I. A., Castellano, M. J., Purcell, L. C., Naeve, S., Grassini, P., ... Archontoulis, S. V. (2022). Climate change and management impacts on soybean n fixation, soil n mineralization, n2o emissions, and seed yield. *Front. Pl. Sci.*, 13, 1–15. doi: [10.3389/fpls.2022.849896](https://doi.org/10.3389/fpls.2022.849896)
- FAOSTAT. (2022). *Food and agriculture organization of the united nations. faostat statistical database*. Retrieved from <https://www.fao.org/faostat/> (Last accessed June 2022)
- FAOSTAT. (2023). *Food and agriculture organization of the united nations—crops’ data*. Retrieved from <https://www.fao.org/faostat/en/#home>
- Fehr, W. R., & Caviness, C. E. (1977). *Stages of soybean development* (Tech. Rep.). Ames, IA: Iowa State University. Retrieved from <https://core.ac.uk/download/pdf/83024475.pdf>
- Ferrari, S., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7), 799–815. doi: [10.1080/0266476042000214501](https://doi.org/10.1080/0266476042000214501)

- Fey, S., Kyveryga, P., Connor, J., Kiel, A., & Muth, D. (2016). Within-field profitability assessment: Impact of weather, field management and soils. In *13th international conference on precision agriculture*. St. Louis, Missouri, USA.
- Fickett, N. D., Boerboom, C. M., & Stoltenberg, D. E. (2013). Soybean yield loss potential associated with early-season weed competition across 64 site-years. *Weed Science*, *61*(3), 500–507. doi: [10.1614/ws-d-12-00164.1](https://doi.org/10.1614/ws-d-12-00164.1)
- Fischer, R. A. (2015). Definitions and determination of crop yield, yield gaps, and of rates of change. *Field Crops Research*, *182*, 9–18. doi: [10.1016/j.fcr.2014.12.006](https://doi.org/10.1016/j.fcr.2014.12.006)
- Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., ... Zaks, D. P. M. (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337–342. doi: [10.1038/nature10452](https://doi.org/10.1038/nature10452)
- Franzen, D., & Mulla, D. (2015). A history of precision agriculture. In *Precision agriculture technology for crop farming* (pp. 1–20). New York, NY: CRC Press.
- Fu, Q., Wang, Q., Shen, X., & Fan, J. (2014). Optimizing water and nitrogen inputs for winter wheat cropping system on the loess plateau, china. *Journal of Arid Land*, *6*, 230–242. doi: [10.1007/s40333-013-0225-4](https://doi.org/10.1007/s40333-013-0225-4)
- Fulton, J., Hawkins, E., Taylor, R., & Franzen, A. (2018). Yield monitor data: Collection, management, and usage. *Crops & Soils*, *51*, 4–51. doi: [10.2134/cs2018.51.0403](https://doi.org/10.2134/cs2018.51.0403)
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., ... Verdin, A. (2014). *A quasi-global precipitation time series for drought monitoring: U.s.* (Tech. Rep. No. 832). Geological Survey Data Series.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*, 150066. doi: [10.1038/sdata.2015.66](https://doi.org/10.1038/sdata.2015.66)

- Gelb, J., & Apparicio, P. (2021). Apport de la classification floue c-means spatiale en géographie: essai de taxinomie socio-résidentielle et environnementale à lyon. *Cybergegeo*. doi: [10.4000/cybergegeo.36414](https://doi.org/10.4000/cybergegeo.36414)
- Gijsman, A. J., Hoogenboom, G., Parton, W. J., & Kerridge, P. C. (2002). Modifying dssat crop models for low-input agricultural systems using a soil organic matter-residue module from century. *Agron. J.*, *94*(3), 462–474. doi: [10.2134/agronj2002.4620](https://doi.org/10.2134/agronj2002.4620)
- Goodchild, M. (1991). Geographic information systems. *Progress in Human Geography*, *15*, 194–200. doi: [10.1177/030913259101500205](https://doi.org/10.1177/030913259101500205)
- Grassini, P., Torrion, J., Yang, H., Rees, J., Andersen, D., Cassman, K., & Specht, J. (2015). Soybean yield gaps and water productivity in the western u.s. corn belt. *Field Crops Research*, *179*, 150–163. doi: [10.1016/j.fcr.2015.04.015](https://doi.org/10.1016/j.fcr.2015.04.015)
- Griffin, T., Dobbins, C., Vyn, T., Florax, R., & Lowenberg-DeBoer, J. (2008). Spatial analysis of yield monitor data: case studies of on-farm trials and farm management decision making. *Precision Agriculture*, *9*, 269–283. doi: [10.1007/s11119-008-9072-2](https://doi.org/10.1007/s11119-008-9072-2)
- Gräler, B., Pebesma, E., & Heuvelink, G. (2016). Spatio-temporal interpolation using gstat. *The R Journal*, *8*, 204. doi: [10.32614/rj-2016-014](https://doi.org/10.32614/rj-2016-014)
- Guastaferro, F., Castrignanò, A., De Benedetto, D., Sollitto, D., Troccoli, A., & Cafarelli, B. (2010). A comparison of different algorithms for the delineation of management zones. *Precision Agriculture*, *11*, 600–620. doi: [10.1007/s11119-010-9183-4](https://doi.org/10.1007/s11119-010-9183-4)
- Haarhoff, S., & Swanepoel, P. (2018). Plant population and maize grain yield: A global systematic review of rainfed trials. *Crop Science*, *58*, 1819–1829. doi: [10.2135/crop-sci2018.01.0003](https://doi.org/10.2135/crop-sci2018.01.0003)

- Halvorson, G., & Doll, E. (1991). Topographic effects on spring wheat yields and water use. *Soil Science Society of America Journal*, 55, 1680–1685. doi: [10.2136/sssaj1991.03615995005500060030x](https://doi.org/10.2136/sssaj1991.03615995005500060030x)
- Harder, D. B., Sprague, C. L., & Renner, K. A. (2007). Effect of soybean row width and population on weeds, crop yield, and economic return. *Weed Technology*, 21(3), 744–752. doi: [10.1614/wt-06-122.1](https://doi.org/10.1614/wt-06-122.1)
- Hatfield, J. L., Gitelson, A. A., Schepers, J. S., & Walthall, C. L. (2008). Application of spectral remote sensing for agronomic decisions. *Agronomy Journal*, 100. doi: [10.2134/agronj2006.0370c](https://doi.org/10.2134/agronj2006.0370c)
- Hatfield, J. L., & Walthall, C. L. (2015). Meeting global food needs: Realizing the potential via genetics $\times$ environment $\times$ management interactions. *Agronomy Journal*, 107, 1215–1226. doi: [10.2134/agronj15.0076](https://doi.org/10.2134/agronj15.0076)
- He, X., Zhang, D., Yang, L., Cui, T., Ding, Y., & Zhong, X. (2021). Design and experiment of a gps-based turn compensation system for improving the seeding uniformity of maize planter. *Computers and Electronics in Agriculture*, 187, 106250. doi: [10.1016/j.compag.2021.106250](https://doi.org/10.1016/j.compag.2021.106250)
- Heatherly, L. G., & Spurlock, S. R. (1999). Yield and economics of traditional and early soybean production system (esps) seedlings in the midsouthern united states. *Field Crops Res.*, 63(1), 35–45. doi: [10.1016/s0378-4290\(99\)00025-8](https://doi.org/10.1016/s0378-4290(99)00025-8)
- Hedley, C. (2015). The role of precision agriculture for improved nutrient management on farms: Precision agriculture managing farm nutrients. *Journal of Science of Food and Agriculture*, 95, 12–19. doi: [10.1002/jsfa.6734](https://doi.org/10.1002/jsfa.6734)
- Hein, L., Van Koppen, K., De Groot, R. S., & Van Ierland, E. C. (2006). Spatial scales, stakeholders and the valuation of ecosystem services. *Ecological Economics*, 57, 209–228. doi: [10.1016/j.ecolecon.2005.04.005](https://doi.org/10.1016/j.ecolecon.2005.04.005)

- Hemathilake, D. M. K. S., & Gunathilake, D. M. C. C. (2022). Agricultural productivity and food supply to meet increased demands. In *Future foods* (pp. 539–553). Elsevier. doi: [10.1016/b978-0-323-91001-9.00016-5](https://doi.org/10.1016/b978-0-323-91001-9.00016-5)
- Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., ... others (2016). Impact of spatial soil and climate input data aggregation on regional yield simulations. *PLoS ONE*, *11*, e0151782.
- Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., ... Keating, B. A. (2014). Apsim - evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.*, *62*, 327–350. doi: [10.1016/j.envsoft.2014.07.009](https://doi.org/10.1016/j.envsoft.2014.07.009)
- Holzworth, D. P., Huth, N. I., Fainges, J., Brown, H., Zurcher, E., Cichota, R., ... Snow, V. (2018). Apsim next generation: Overcoming challenges in modernising a farming systems model. *Environ. Model. Softw.*, *103*, 43–51. doi: [10.1016/j.envsoft.2018.02.002](https://doi.org/10.1016/j.envsoft.2018.02.002)
- Huber, S. C., Li, K., Nelson, R., Ulanov, A., DeMuro, C. M., & Baxter, I. (2016). Canopy position has a profound effect on soybean seed composition. *PeerJ*, *4*, e2452. doi: [10.7717/peerj.2452](https://doi.org/10.7717/peerj.2452)
- Höfle, B., & Rutzinger, M. (2011). Topographic airborne lidar in geomorphology: A technological perspective. *zfg-suppl*, *55*, 1–29. doi: [10.1127/0372-8854/2011/0055s2-0043](https://doi.org/10.1127/0372-8854/2011/0055s2-0043)
- Iqbal, M., Singh, A., & Ansari, M. (2020). Effect of drought stress on crop production. In A. Rakshit, H. Singh, A. K. Singh, U. Singh, & L. Fraceto (Eds.), *New frontiers in stress management for durable agriculture* (pp. 35–47). Springer Singapore. doi: [10.1007/978-981-15-1322-0_3](https://doi.org/10.1007/978-981-15-1322-0_3)
- Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., & Wang, J. (2018). A review of data assimilation of remote sensing and crop models. *European Journal of Agronomy*, *92*, 141–152. doi: [10.1016/j.eja.2017.11.002](https://doi.org/10.1016/j.eja.2017.11.002)

- Jin, Z., Zhuang, Q., Wang, J., Archontoulis, S. V., Zobel, Z., & Kotamarthi, V. R. (2018). The combined and separate impacts of climate extremes on the current and future us rainfed maize and soybean production under elevated co₂. *Glob. Chang. Biol.*, *23*(7), 2687–2704. doi: [10.1111/gcb.13617](https://doi.org/10.1111/gcb.13617)
- Kantolic, A. G., & Slafer, G. A. (2001). Photoperiod sensitivity after flowering and seed number determination in indeterminate soy-bean cultivars. *Field Crops Res.*, *72*(2), 109–118. doi: [10.1016/s0378-4290\(01\)00168-x](https://doi.org/10.1016/s0378-4290(01)00168-x)
- Kessler, A., Archontoulis, S. V., & Licht, M. A. (2020). Soybean yield and crop stage response to planting date and cultivar maturity in iowa, usa. *Agron. J.*, *112*(1), 382–394. doi: [10.1002/agj2.20053](https://doi.org/10.1002/agj2.20053)
- Khosla, R., Westfall, D., Reich, R., Mahal, J., & Gangloff, W. (2010). Spatial variation and site-specific management zones. In M. Oliver (Ed.), *Geostatistical applications for precision agriculture* (pp. 195–219). Dordrecht: Springer Netherlands. doi: [10.1007/978-90-481-9133-8_8](https://doi.org/10.1007/978-90-481-9133-8_8)
- Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper danube basin under an ensemble of climate change scenarios. *J. Hydrol.*, *424–425*, 264–277. doi: [10.1016/j.jhydrol.2012.01.011](https://doi.org/10.1016/j.jhydrol.2012.01.011)
- Koester, R. P., Skoneczka, J. A., Cary, T. R., Diers, B. W., & Ainsworth, E. A. (2014). Historical gains in soybean (glycine max merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *Journal of Experimental Botany*, *65*(12), 3311–3321. doi: [10.1093/jxb/eru187](https://doi.org/10.1093/jxb/eru187)
- Krause, M., Crossman, S., DuMond, T., Lott, R., Swede, J., Arliss, S., ... Gore, M. (2020). Random forest regression for optimizing variable planting rates for corn and soybean using topographical and soil data. *Agronomy Journal*, *112*, 5045–5066. doi: [10.1002/agj2.20442](https://doi.org/10.1002/agj2.20442)

- Kravchenko, A., Robertson, G., Thelen, K., & Harwood, R. (2005). Management, topographical, and weather effects on spatial variability of crop grain yields. *Agronomy Journal*, *97*, 514–523. doi: [10.2134/agronj2005.0514](https://doi.org/10.2134/agronj2005.0514)
- Kukal, M., & Irmak, S. (2018). Climate-driven crop yield and yield variability and climate change impacts on the u.s. great plains agricultural production. *Scientific Reports*, *8*, 3450. doi: [10.1038/s41598-018-21848-2](https://doi.org/10.1038/s41598-018-21848-2)
- Kumagai, E., & Sameshima, R. (2014). Genotypic differences in soybean yield responses to increasing temperature in a cool climate are related to maturity group. *Agric. For. Meteorol.*, *198*, 265–272. doi: [10.1016/j.agrformet.2014.08.016](https://doi.org/10.1016/j.agrformet.2014.08.016)
- Kumhálová, J., Kumhála, F., Kroulík, M., & Matějková, (2011). The impact of topography on soil properties and yield and the effects of weather conditions. *Precision Agriculture*, *12*, 813–830. doi: [10.1007/s11119-011-9221-x](https://doi.org/10.1007/s11119-011-9221-x)
- Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means*. (R package version 1.7.0, <https://CRAN.R-project.org/package=emmeans>)
- Li, Y., Guan, K., Schnitkey, G., DeLucia, E., & Peng, B. (2019). Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the united states. *Global Change Biology*, *25*, 2325–2337. doi: [10.1111/gcb.14628](https://doi.org/10.1111/gcb.14628)
- Liu, R., Wang, J., Shi, J., Chen, Y., Sun, C., Zhang, P., & Shen, Z. (2014). Runoff characteristics and nutrient loss mechanism from plain farmland under simulated rainfall conditions. *Science of The Total Environment*, *468–469*, 1069–1077. doi: [10.1016/j.scitotenv.2013.09.035](https://doi.org/10.1016/j.scitotenv.2013.09.035)
- Liu, Y., & Dai, L. (2020). Modelling the impacts of climate change and crop management measures on soybean phenology in china. *J. Clean. Prod.*, *262*, 121271. doi: [10.1016/j.jclepro.2020.121271](https://doi.org/10.1016/j.jclepro.2020.121271)

- Lobell, D., Cassman, K., & Field, C. (2009). Crop yield gaps: Their importance, magnitudes, and causes. *Annual Review of Environment and Resources*, *34*, 179–204. doi: [10.1146/annurev.envIRON.041008.093740](https://doi.org/10.1146/annurev.envIRON.041008.093740)
- Lowenberg-DeBoer, J., & Erickson, B. (2019). Setting the record straight on precision agriculture adoption. *Agronomy Journal*, *111*, 1552–1569. doi: [10.2134/agronj2018.12.0779](https://doi.org/10.2134/agronj2018.12.0779)
- Lu, P., Jiang, B., & Weiner, J. (2020). Crop spatial uniformity, yield and weed suppression. In *Advances in agronomy* (Vol. 161, pp. 117–178). Elsevier. doi: [10.1016/bs.agron.2019.12.003](https://doi.org/10.1016/bs.agron.2019.12.003)
- Maestrini, B., & Basso, B. (2018). Drivers of within-field spatial and temporal variability of crop yield across the us midwest. *Scientific Reports*, *8*, 14833. doi: [10.1038/s41598-018-32779-3](https://doi.org/10.1038/s41598-018-32779-3)
- Martinez-Feria, R., & Basso, B. (2020). Unstable crop yields reveal opportunities for site-specific adaptations to climate variability. *Scientific Reports*, *10*, 2885. doi: [10.1038/s41598-020-59494-2](https://doi.org/10.1038/s41598-020-59494-2)
- Masino, A., Rugeroni, P., Borrás, L., & Rotundo, J. L. (2018). Spatial and temporal plant-to-plant variability effects on soybean yield. *European Journal of Agronomy*, *98*, 14–24. doi: [10.1016/j.eja.2018.02.006](https://doi.org/10.1016/j.eja.2018.02.006)
- McArthur, J., & McCord, G. (2017). Fertilizing growth: Agricultural inputs and their effects in economic development. *Journal of Development Economics*, *127*, 133–152. doi: [10.1016/j.jdeveco.2017.02.007](https://doi.org/10.1016/j.jdeveco.2017.02.007)
- Medic, J., Atkinson, C., & Hurburgh, C. R. (2014). Current knowledge in soybean composition. *Journal of the American Oil Chemists' Society*, *91*(3), 363–384. doi: [10.1007/s11746-013-2407-9](https://doi.org/10.1007/s11746-013-2407-9)

- Miguez, F. (2022). *apsimx: Inspect, read, edit and run ‘apsim’ “next generation” and ‘apsim’ classic*. (R package version 2.3.1, <https://CRAN.R-project.org/package=apsimx>)
- Mitra, A., Vangipuram, S., Bapatla, A., Bathalapalli, V., Mohanty, S., Kougianos, E., & Ray, C. (2022). *Everything you wanted to know about smart agriculture*. (<https://doi.org/10.48550/ARXIV.2201.04754>) doi: [10.48550/arxiv.2201.04754](https://doi.org/10.48550/arxiv.2201.04754)
- Mitra, A., Vangipuram, S., Bapatla, A., Bathalapalli, V., Mohanty, S., Kougianos, E., & Ray, C. (2024). Smart agriculture: A comprehensive overview. *SN Computer Science*, 5, 969. doi: [10.1007/s42979-024-03319-w](https://doi.org/10.1007/s42979-024-03319-w)
- Moore, G., & Tyndale-Biscoe, J. (1999). Estimation of the importance of spatially variable nitrogen application and soil moisture holding capacity to wheat production. *Precision Agriculture*, 1, 27–38. doi: [10.1023/a:1009973802295](https://doi.org/10.1023/a:1009973802295)
- Moreira, F. F., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2019). Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. *Plant Methods*, 15(1), 139. doi: [10.1186/s13007-019-0519-4](https://doi.org/10.1186/s13007-019-0519-4)
- Mourtzinis, S., & Conley, S. P. (2017). Delineating soybean maturity groups across the united states. *Agronomy Journal*, 109, 1397–1403. doi: [10.2134/agronj2016.10.0581](https://doi.org/10.2134/agronj2016.10.0581)
- Mourtzinis, S., Gaspar, A., Naeve, S., & Conley, S. (2017c). Planting date, maturity, and temperature effects on soybean seed yield and composition. *Agronomy Journal*, 109, 2040–2049. doi: [10.2134/agronj2017.05.0247](https://doi.org/10.2134/agronj2017.05.0247)
- Mourtzinis, S., Rattalino Edreira, J. I., Conley, S. P., & Grassini, P. (2017). From grid to field: Assessing quality of gridded weather data for agricultural applications. *European Journal of Agronomy*, 82, 163–172. doi: [10.1016/j.eja.2016.10.013](https://doi.org/10.1016/j.eja.2016.10.013)
- Mourtzinis, S., Specht, J. E., & Conley, S. P. (2019). Defining optimal soybean sowing dates across the us. *Sci. Rep.*, 9(1), 1–7. doi: [10.1038/s41598-019-38971-3](https://doi.org/10.1038/s41598-019-38971-3)

- Mourtzinis, S., Specht, J. E., Lindsey, L. E., Wiebold, W. J., Ross, J., Nafziger, E. D., ... Conley, S. P. (2015). Climate-induced reduction in us-wide soybean yields underpinned by region-and in-season-specific responses. *Nat. Plants*, *1*, 8–11. doi: [10.1038/nplants.2014.26](https://doi.org/10.1038/nplants.2014.26)
- Mzuku, M., Khosla, R., Reich, R., Inman, D., Smith, F., & MacDonald, L. (2005). Spatial variability of measured soil properties across site-specific management zones. *Soil Science Society of America Journal*, *69*, 1572–1579. doi: [10.2136/sssaj2005.0062](https://doi.org/10.2136/sssaj2005.0062)
- Nawar, S., Corstanje, R., Halcro, G., Mulla, D., & Mouazen, A. (2017). Delineation of soil management zones for variable-rate fertilization. In *Advances in agronomy* (pp. 175–245). Elsevier. doi: [10.1016/bs.agron.2017.01.003](https://doi.org/10.1016/bs.agron.2017.01.003)
- O’Hagan, A. (2008). The bayesian approach to statistics. In T. Rudas (Ed.), *Handbook of probability: Theory and applications* (pp. 85–100). SAGE Publications.
- Ospina, R., & Ferrari, S. L. P. (2010). Inflated beta distributions. *Statistical Papers*, *51*(1), 111–126. doi: [10.1007/s00362-008-0125-4](https://doi.org/10.1007/s00362-008-0125-4)
- Oteng-Darko, P., Yeboah, S., Addy, S. N. T., Amponsah, S., & Owusu Danquah, E. (2013). Crop modeling: A tool for agricultural research – a review. *E3 Journal of Agricultural Research and Development*, *2*, 001–006.
- Patrignani, A., & Ochsner, T. E. (2015). Canopeo: A powerful new tool for measuring fractional green canopy cover. *Agronomy Journal*, *107*(6), 2312–2320. doi: [10.2134/agronj15.0150](https://doi.org/10.2134/agronj15.0150)
- Pereyra, V. M., Bastos, L. M., de Borja Reis, A. F., Melchiori, R. J. M., Maltese, N. E., Appelans, S. C., ... Ciampitti, I. A. (2022). Early-season plant-to-plant spatial uniformity can affect soybean yields. *Scientific Reports*, *12*(1), 17128. doi: [10.1038/s41598-022-21385-z](https://doi.org/10.1038/s41598-022-21385-z)

- Pilesjö, P., Thylén, L., & Persson, A. (2005). Topographical data for delineation of agricultural management zones. In *Precision agriculture '05* (pp. 819–826). Wageningen, Netherlands: Wageningen Academic Publishers.
- Plant, R. E. (2001). Site-specific management: the application of information technology to crop production. *Computers and Electronics in Agriculture*, *30*, 9–29. doi: [10.1016/S0168-1699\(00\)00152-6](https://doi.org/10.1016/S0168-1699(00)00152-6)
- Ploschuk, R., Miralles, D., & Striker, G. (2022). A quantitative review of soybean responses to waterlogging: agronomical, morpho-physiological and anatomical traits of tolerance. *Plant Soil*, *475*, 237–252. doi: [10.1007/s11104-022-05364-x](https://doi.org/10.1007/s11104-022-05364-x)
- Pryor, S., Scavia, D., Downer, C., Gaden, M., Iverson, L., Nordstrom, R., ... Robertson, G. (2014). *Midwest. climate change impacts in the united states: The third national climate assessment*. U.S. Global Change Research Program. doi: [10.7930/j0j1012n](https://doi.org/10.7930/j0j1012n)
- Purcell, L. C., Ball, R. A., Reaper, J. D., & Vories, E. D. (2002). Radiation use efficiency and biomass production in soybean at different plant population densities. *Crop Science*, *42*(1), 172–177. doi: [10.2135/cropsci2002.1720](https://doi.org/10.2135/cropsci2002.1720)
- R Core Team. (2021). *R: A language and environment for statistical computing*. <https://www.r-project.org/>. (R Foundation for Statistical Computing)
- Ray, D. K., Gerber, J. S., MacDonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, *6*, 5989. doi: [10.1038/ncomms6989](https://doi.org/10.1038/ncomms6989)
- Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., & Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications*, *3*(1), 1293. doi: [10.1038/ncomms2296](https://doi.org/10.1038/ncomms2296)

- Riihimäki, H., Kemppinen, J., Kopecký, M., & Luoto, M. (2021). Topographic wetness index as a proxy for soil moisture: The importance of flow-routing algorithm and grid resolution. *Water Resources Research*, 57, e2021WR029871. doi: [10.1029/2021wr029871](https://doi.org/10.1029/2021wr029871)
- Rizzo, G., Mazzilli, S. R., Ernst, O., Baethgen, W. E., & Berger, A. G. (2022). Season-specific management strategies for rainfed soybean in the south american pampas based on a seasonal precipitation forecast. *Ag. Syst.*, 196, 103331. doi: [10.1016/j.agry.2021.103331](https://doi.org/10.1016/j.agry.2021.103331)
- Rockström, J., Karlberg, L., Wani, S., Barron, J., Hatibu, N., Oweis, T., ... Qiang, Z. (2010). Managing water in rainfed agriculture—the need for a paradigm shift. *Agricultural Water Management*, 97, 543–550. doi: [10.1016/j.agwat.2009.09.009](https://doi.org/10.1016/j.agwat.2009.09.009)
- Rogovska, N., O'Brien, P. L., Malone, R., Emmett, B., Kovar, J. L., Jaynes, D., ... Kyveryga, P. (2023). Long-term conservation practices reduce nitrate leaching while maintaining yields in tile-drained midwestern soils. *Agricultural Water Management*, 288, 108481. doi: [10.1016/j.agwat.2023.108481](https://doi.org/10.1016/j.agwat.2023.108481)
- Roth, L., Barendregt, C., Bétrix, C.-A., Hund, A., & Walter, A. (2022). High-throughput field phenotyping of soybean: Spotting an ideotype. *Remote Sensing of Environment*, 269, 112797. doi: [10.1016/j.rse.2021.112797](https://doi.org/10.1016/j.rse.2021.112797)
- Rotili, D., De Voil, P., Eyre, J., Serafin, L., Aisthorpe, D., Maddonni, G., & Rodríguez, D. (2020). Untangling genotype x management interactions in multi-environment on-farm experimentation. *Field Crops Research*, 255, 107900. doi: [10.1016/j.fcr.2020.107900](https://doi.org/10.1016/j.fcr.2020.107900)
- Rouphael, Y., Cardarelli, M., Schwarz, D., Franken, P., & Colla, G. (2012). Effects of drought on nutrient uptake and assimilation in vegetable crops. In R. Aroca (Ed.), *Plant responses to drought stress* (pp. 171–195). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: [10.1007/978-3-642-32653-0_7](https://doi.org/10.1007/978-3-642-32653-0_7)

- Roy, R., & Aslekar, A. (2022). Iot in farm productivity enhancement. In *Proceedings of the 2022 international conference on decision aid sciences and applications (dasa)* (pp. 1034–1039). Chiangrai, Thailand: IEEE. doi: [10.1109/dasa54658.2022.9765273](https://doi.org/10.1109/dasa54658.2022.9765273)
- Rurinda, J., van Wijk, M. T., Mapfumo, P., Descheemaeker, K., Supit, I., & Giller, K. E. (2015). Climate change and maize yield in southern africa: What can farm management do? *Glob. Change Biol.*, *21*(12), 4588–4601. doi: [10.1111/gcb.13061](https://doi.org/10.1111/gcb.13061)
- Safo-Kantanka, O., & Lawson, N. C. (1980). The effect of different row spacings and plant arrangements on soybeans. *Canadian Journal of Plant Science*, *60*(1), 227–231. doi: [10.4141/cjps80-031](https://doi.org/10.4141/cjps80-031)
- Salmeron, M., Gbur, E. E., Bourland, F. M., Buehring, N. W., Earnest, L., Fritschi, F. B., ... Purcell, L. C. (2014). Soybean maturity group choices for early and late plantings in the midsouth. *Agron. J.*, *106*(5), 1893–1901. doi: [10.2134/agronj14.0222](https://doi.org/10.2134/agronj14.0222)
- Sawyer, J. (1994). Concepts of variable rate technology with considerations for fertilizer application. *Journal of Production Agriculture*, *7*, 195–201. doi: [10.2134/jpa1994.0195](https://doi.org/10.2134/jpa1994.0195)
- Saxton, K. E., & Rawls, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.*, *70*(5), 1569–1578. doi: [10.2136/sssaj2005.0117](https://doi.org/10.2136/sssaj2005.0117)
- Schauberger, B., Archontoulis, S., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., ... Frieler, K. (2017). Consistent negative response of us crops to high temperatures in observations and crop models. *Nat. Commun.*, *8*. doi: [10.1038/ncomms13931](https://doi.org/10.1038/ncomms13931)
- Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to u.s. crop yields under climate change. *Proc. Natl. Acad. Sci.*, *106*(37), 15594–15598. doi: [10.1073/pnas.0906865106](https://doi.org/10.1073/pnas.0906865106)

- Schmitz, P. K., & Kandel, H. J. (2021). Using canopy measurements to predict soybean seed yield. *Remote Sensing*, 13(16), 3260. doi: [10.3390/rs13163260](https://doi.org/10.3390/rs13163260)
- Schutte, M., & Nleya, T. (2019). Row spacing and seeding rate effects on soybean seed yield. In M. Kasai (Ed.), *Soybean—biomass, yield and productivity*. IntechOpen. doi: [10.5772/intechopen.80748](https://doi.org/10.5772/intechopen.80748)
- Seifert, C., Roberts, M., & Lobell, D. (2017). Continuous corn and soybean yield penalties across hundreds of thousands of fields. *Agronomy Journal*, 109, 541–548. doi: [10.2134/agronj2016.03.0134](https://doi.org/10.2134/agronj2016.03.0134)
- Shibles, R. M., & Weber, C. R. (1966). Interception of solar radiation and dry matter production by various soybean planting patterns. *Crop Science*, 6(1), 55–59. doi: [10.2135/crop-sci1966.0011183x000600010017x](https://doi.org/10.2135/crop-sci1966.0011183x000600010017x)
- Shideed, K. (2017). Rainfed agriculture and food security in dry areas. In S. Murad, E. Baydoun, & N. Daghir (Eds.), *Water, energy food sustainability in the middle east* (pp. 299–340). Cham: Springer International Publishing. doi: [10.1007/978-3-319-48920-9_14](https://doi.org/10.1007/978-3-319-48920-9_14)
- Simmonds, M., Plant, R., Peña-Barragán, J., Van Kessel, C., Hill, J., & Linqvist, B. (2013). Underlying causes of yield spatial variability and potential for precision management in rice systems. *Precision Agriculture*, 14, 512–540. doi: [10.1007/s11119-013-9313-x](https://doi.org/10.1007/s11119-013-9313-x)
- Sinclair, T. R., & Seligman, N. G. (1996). Crop modeling: From infancy to maturity. *Agronomy Journal*, 88, 698–704. doi: [10.2134/agronj1996.00021962008800050004x](https://doi.org/10.2134/agronj1996.00021962008800050004x)
- Sobko, O., Zikeli, S., Claupein, W., & Gruber, S. (2020). Seed yield, seed protein, oil content, and agronomic characteristics of soybean (glycine max l. merrill) depending on different seeding systems and cultivars in germany. *Agronomy*, 10(7), 1020. doi: [10.3390/agronomy10071020](https://doi.org/10.3390/agronomy10071020)

- Specht, J. E., Diers, B. W., Nelson, R. L., de Toledo, J. F. F., Torrion, J. A., & Grassini, P. (2015). Soybean. In S. Smith, B. Diers, J. Specht, & B. Carver (Eds.), *Yield gains in major u.s. field crops* (pp. 311–355). CSSA, ASA, SSSA. doi: [10.2135/cssaspecpub33.c12](https://doi.org/10.2135/cssaspecpub33.c12)
- Stewart, B., & Lal, R. (2018). Increasing world average yields of cereal crops. In *Advances in agronomy* (pp. 1–44). Elsevier. doi: [10.1016/bs.agron.2018.05.001](https://doi.org/10.1016/bs.agron.2018.05.001)
- Sudduth, K., Drummond, S., Birrell, S., & Kitchen, N. (2015). Analysis of spatial factors influencing crop yield. In P. Robert, R. Rust, & W. Larson (Eds.), *Asa, cssa, and sssa books* (pp. 129–139). Madison, WI, USA: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. doi: [10.2134/1996.precisionagproc3.c13](https://doi.org/10.2134/1996.precisionagproc3.c13)
- Suhre, J. J., Weidenbenner, N. H., Rowntree, S. C., Wilson, E. W., Naeve, S. L., Conley, S. P., ... Davis, V. M. (2014). Soybean yield partitioning changes revealed by genetic gain and seeding rate interactions. *Agronomy Journal*, 106(5), 1631–1642. doi: [10.2134/agronj14.0003](https://doi.org/10.2134/agronj14.0003)
- Tantalaki, N., Souravlas, S., & Roumeliotis, M. (2019). Data-driven decision making in precision agriculture: The rise of big data in agricultural systems. *Journal of Agricultural & Food Information*, 20, 344–380. doi: [10.1080/10496505.2019.1638264](https://doi.org/10.1080/10496505.2019.1638264)
- Teasdale, J., & Cavigelli, M. (2017). Meteorological fluctuations define long-term crop yield patterns in conventional and organic production systems. *Scientific Reports*, 7, 688. doi: [10.1038/s41598-017-00775-8](https://doi.org/10.1038/s41598-017-00775-8)
- Therneau, T., Atkinson, B., & Ripley, B. (2015). Package "rpart" [Computer software manual].
- Thornton, P. E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., & Wilson, B. E. (2021). Gridded daily weather data for north america with comprehensive uncertainty quantification. *Scientific Data*, 8(1), 190. doi: [10.1038/s41597-021-00973-0](https://doi.org/10.1038/s41597-021-00973-0)

- Tomal, M., & Gumieniak, A. (2020). Agricultural land price convergence: Evidence from polish provinces. *Agriculture*, 10, 183. doi: [10.3390/agriculture10050183](https://doi.org/10.3390/agriculture10050183)
- Trevisan, R., Bullock, D., & Martin, N. (2021). Spatial variability of crop responses to agronomic inputs in on-farm precision experimentation. *Precision Agriculture*, 22, 342–363. doi: [10.1007/s11119-020-09720-8](https://doi.org/10.1007/s11119-020-09720-8)
- USDA. (2023). *Soybean 2023 world production*. <https://ipad.fas.usda.gov/cropeexplorer/cropview/commodityView.aspx?cropid=222200>. (Accessed: 2023)
- USDA-NASS. (2022). *United states department of agriculture–national agricultural statistics service*. <https://www.nass.usda.gov>. (Accessed on 15 April 2022)
- Van Dijk, M., Morley, T., Rau, M. L., & Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food*, 2, 494–501. doi: [10.1038/s43016-021-00322-9](https://doi.org/10.1038/s43016-021-00322-9)
- Van Metre, P., Frey, J., Musgrove, M., Nakagaki, N., Qi, S., Mahler, B., ... Button, D. (2016). High nitrate concentrations in some midwest united states streams in 2013 after the 2012 drought. *Journal of Environmental Quality*, 45, 1696–1704. doi: [10.2134/jeq2015.12.0591](https://doi.org/10.2134/jeq2015.12.0591)
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of mcmc (with discussion). *Bayesian Analysis*, 16(2). doi: [10.1214/20-ba1221](https://doi.org/10.1214/20-ba1221)
- Vitali, G., Francia, M., Golfarelli, M., & Canavari, M. (2021). Crop management with the iot: An interdisciplinary survey. *Agronomy*, 11, 181. doi: [10.3390/agronomy11010181](https://doi.org/10.3390/agronomy11010181)
- Vollmann, J., Fritz, C. N., Wagentristl, H., & Ruckebauer, P. (2000). Environmental and genetic variation of soybean seed protein content under central european growing conditions.

- Journal of the Science of Food and Agriculture*, 80(9), 1300–1306. doi: [10.1002/1097-0010\(200007\)80:9<1300::aid-jsfa640>3.0.co;2-i](https://doi.org/10.1002/1097-0010(200007)80:9<1300::aid-jsfa640>3.0.co;2-i)
- Wallach, D., Palosuo, T., Thorburn, P., Gourdain, E., Asseng, S., Basso, B., . . . Seidel, S. J. (2021). How well do crop modeling groups predict wheat phenology, given calibration data from the target population? *Eur. J. Agron.*, 124, 126195. doi: [10.1016/j.eja.2020.126195](https://doi.org/10.1016/j.eja.2020.126195)
- Waller, D., Meyer, A., Raff, Z., & Apfelbaum, S. (2021). Shifts in precipitation and agricultural intensity increase phosphorus concentrations and loads in an agricultural watershed. *Journal of Environmental Management*, 284, 112019. doi: [10.1016/j.jenvman.2021.112019](https://doi.org/10.1016/j.jenvman.2021.112019)
- Wani, S., Rockström, J., & Oweis, T. (2009). *Rainfed agriculture: Unlocking the potential*. Wallingford: CABI.
- Whelan, B., & McBratney, A. (2000). The "null hypothesis" of precision agriculture management. *Precision Agriculture*, 2, 265–279. doi: [10.1023/a:1011838806489](https://doi.org/10.1023/a:1011838806489)
- White, J. W., Hoogenboom, G., Kimball, B. A., & Wall, G. W. (2011). Methodologies for simulating impacts of climate change on crop production. *Field Crops Research*, 124, 357–368. doi: [10.1016/j.fcr.2011.07.001](https://doi.org/10.1016/j.fcr.2011.07.001)
- Wichmann, V., Bremer, M., Lindenberger, J., Rutzinger, M., Georges, C., & Petrini-Monteferrì, F. (2015). Evaluating the potential of multispectral airborne lidar for topographic mapping and land cover classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3/W5, 113–119. doi: [10.5194/isprsannals-ii-3-w5-113-2015](https://doi.org/10.5194/isprsannals-ii-3-w5-113-2015)
- Wiggans, R. G. (1939). The influence of space and arrangement on the production of soybean plants. *Agronomy Journal*, 31(4), 314–321. doi: [10.2134/agronj1939.00021962003100040005x](https://doi.org/10.2134/agronj1939.00021962003100040005x)

- Wilcox, J. R. (1974). Response of three soybean strains to equidistant spacings. *Agronomy Journal*, 66(3), 409–412. doi: [10.2134/agronj1974.00021962006600030022x](https://doi.org/10.2134/agronj1974.00021962006600030022x)
- Willcott, J., Herbert, S. J., & Zhi-Yi, L. (1984). Leaf area display and light interception in short-season soybeans. *Field Crops Research*, 9, 173–182. doi: [10.1016/0378-4290\(84\)90023-6](https://doi.org/10.1016/0378-4290(84)90023-6)
- Wood, S. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65, 95–114. doi: [10.1111/1467-9868.00374](https://doi.org/10.1111/1467-9868.00374)
- Wood, S. (2017). *Generalized additive models: An introduction with r* (2nd ed.). Chapman and Hall/CRC. doi: [10.1201/9781315370279](https://doi.org/10.1201/9781315370279)
- Wu, Y., Wang, E., He, D., Liu, X., Archontoulis, S. V., Huth, N. I., ... Yang, W. (2019). Combine observational data and modelling to quantify cultivar differences of soybean. *Eur. J. Agron.*, 111, 125940. doi: [10.1016/j.eja.2019.125940](https://doi.org/10.1016/j.eja.2019.125940)
- Xu, R., Li, Y., Guan, K., Zhao, L., Peng, B., Miao, C., & Fu, B. (2022). Divergent responses of maize yield to precipitation in the united states. *Environmental Research Letters*, 17, 014016. doi: [10.1088/1748-9326/ac3cee](https://doi.org/10.1088/1748-9326/ac3cee)
- Yandun Narvaez, F., Reina, G., Torres-Torriti, M., Kantor, G., & Cheein, F. A. (2017). A survey of ranging and imaging techniques for precision agriculture phenotyping. *IEEE/ASME Transactions on Mechatronics*, 22, 2428–2439. doi: [10.1109/tmech.2017.2760866](https://doi.org/10.1109/tmech.2017.2760866)
- Yu, S., Khan, S., Mo, F., Ren, A., Lin, W., Feng, Y., ... Gao, Z. (2021). Determining optimal nitrogen input rate on the base of fallow season precipitation to achieve higher crop water productivity and yield. *Agricultural Water Management*, 246, 106689. doi: [10.1016/j.agwat.2020.106689](https://doi.org/10.1016/j.agwat.2020.106689)

- Zabel, F., Müller, C., Elliott, J., Minoli, S., Jägermeyr, J., Schneider, J. M., ... Asseng, S. (2021). Large potential for crop production adaptation depends on available future varieties. *Glob. Chang. Biol.*, 27(16), 3870–3882. doi: [10.1111/gcb.15649](https://doi.org/10.1111/gcb.15649)
- Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture—a worldwide overview. *Computers and Electronics in Agriculture*, 36, 113–132. doi: [10.1016/s0168-1699\(02\)00096-0](https://doi.org/10.1016/s0168-1699(02)00096-0)
- Zhou, W., Guan, K., Peng, B., Wang, Z., Fu, R., Li, B., ... Chen, Z. (2021). A generic risk assessment framework to evaluate historical and future climate-induced risk for rainfed corn and soybean yield in the u.s. midwest. *Weather Clim. Extrem.*, 33. doi: [10.1016/j.wace.2021.100369](https://doi.org/10.1016/j.wace.2021.100369)

Appendix A

Appendix A, Chapter 2

Table A.1: Data sources. KSU performance test (<https://www.agronomy.k-state.edu/outreach-and-services/crop-performance-tests/soybeans/>, accessed 2022/03); Ciampitti et al., 2017 (<https://doi.org/10.4148/2378-5977.7434>); Ciampitti, et al., 2020 (figshare: [10.6084/m9.figshare.20018015](https://figshare.com/figures/articles/10.6084/m9.figshare.20018015)).

Site	Lon	Lat	Year	Planting dates	Maturity group	Plant density (pl/m ²)	Harvest dates	Source
Assaria	-97.60	38.68	2014–2021	May 15 to June 14	IV and V	34 and 38	October 16 to November 1	K-State soybean performance trials
Belleville	-97.84	39.83	2014–2020	May 15 to June 17	III, IV and V	34 and 35	October 15 to October 27	K-State soybean performance trials
Columbus	-94.84	37.16	2016; 2017; 2019	June 27 to June 29	IV, V and VI	25 and 30	November 16 to November 21	K-State soybean performance trials
Erie	-95.24	37.56	2014; 2016–2021	May 19 to July 11	IV and V	25, 30 and 39	November 11 to December 10	K-State soybean performance trials
Kiro	-95.79	39.09	2017–2021	April 06 to June 10	III, IV and V	25 and 35	October 12 to October 22	K-State soybean performance trials
McCune	-95.01	37.35	2014–2016; 2018; 2019	June 06 to July 01	III, IV, V and VI	25, 35, 38	November 02 to November 29	K-State soybean performance trials
Onaga	-96.16	39.48	2014–2018	May 06 to June 22	IV	34 and 38	October 08 to November 19	K-State soybean performance trials
Ottawa	-95.24	38.53	2014–2021	May 16 to July 6	III, IV and V	34 and 35	October 14 to October 22	K-State soybean performance trials
			2014–2016	May 04 to July 15	IV and V	32	Not available	Ciampitti et al. (2017)
			2018–2019	May 10 to June 29	III and IV	32	October 18 to October 25	Field research studies
Parsons	-95.17	37.22	2014; 2015; 2018; 2020	June 06 to July 07	IV, V and VI	25 and 30	November 05 to November 25	K-State soybean performance trials
			2014–2015	May 02 to July 01	IV and V	32	Not available	Ciampitti et al. (2017)
Pittsburg	-94.70	37.41	2014–2018; 2020–2021	June 10 to July 08	IV and V	30 and 38	November 03 to November 16	K-State soybean performance trials
Riley	-96.83	39.29	2019–2021	May 13 to June 04	IV	37 and 38	October 22 to October 24	K-State soybean performance trials
Rossville	-95.95	39.13	2016	May 30 to June 02	IV	25	October 19	K-State soybean performance trials
Topeka	-95.67	39.04	2014–2017; 2021	May 07 to June 02	IV	34	October 14 to October 16	K-State soybean performance trials
			2014–2016	April 30 to June 09	III, V and VI	32	Not available	Ciampitti et al. (2017)
			2018–2019	May 10 to June 29	III and IV	32	October 08 to October 17	Ciampitti et al. (2020)

Table A.2: Standard planting date, soil characteristics, and weather variables used to create the clusters.

Site	Standard planting date (DOY)	Weather characteristics			Soil characteristics			
		Cumulative rainfall (mm)			Temp. max (°C)	Sand (%)	Silt (%)	Clay (%)
		Jul	Aug	Sep				
Assaria	155–163	595	34	33	30	11	67	22
Belleville	150–160	601	33	31	29	10	59	31
Riley	150–160	691	32	31	29	4	61	35
Kiro	120–150	734	32	31	29	6	70	24
Onaga	127–140	713	32	31	29	10	65	25
Rossville	150–153	735	32	31	29	15	66	19
Topeka	120–160	717	32	31	29	7	60	33
Columbus	178–180	809	32	32	30	42	37	21
Erie	150–180	836	33	32	30	13	67	20
McCune	165–180	836	33	32	30	4	54	42
Ottawa	145–170	768	32	32	29	7	70	23
Parsons	150–180	819	32	32	30	13	67	20
Pittsburg	165–175	868	33	32	30	13	67	20

Table A.3: Parameters included in the generic genotypes used from APSIM Next generation.

Parameter	MG3	MG4	MG5	MG6
Vegetative.Target.FixedValue (°Cd)	380	388	396	404
EarlyFlowering.Target.FixedValue (°Cd)	120	140	160	180
EarlyGrainFilling.Target.FractionofGrainfillPeriod.FixedValue	0.361	0.324	0.072	0.056
LateGrainFilling.Target.EntireGrainfillPeriod.FixedValue (°Cd)	664	664	696	712
VegetativePhotoperiodModifier.XYPairs.X	13.4	13.1	12.83	12.58
	16.91	16.49	16.13	15.8
ReproductivePhotoperiodModifier.XYPairs.X	13.4	13.1	12.83	12.58
	16.91	16.49	16.13	15.8

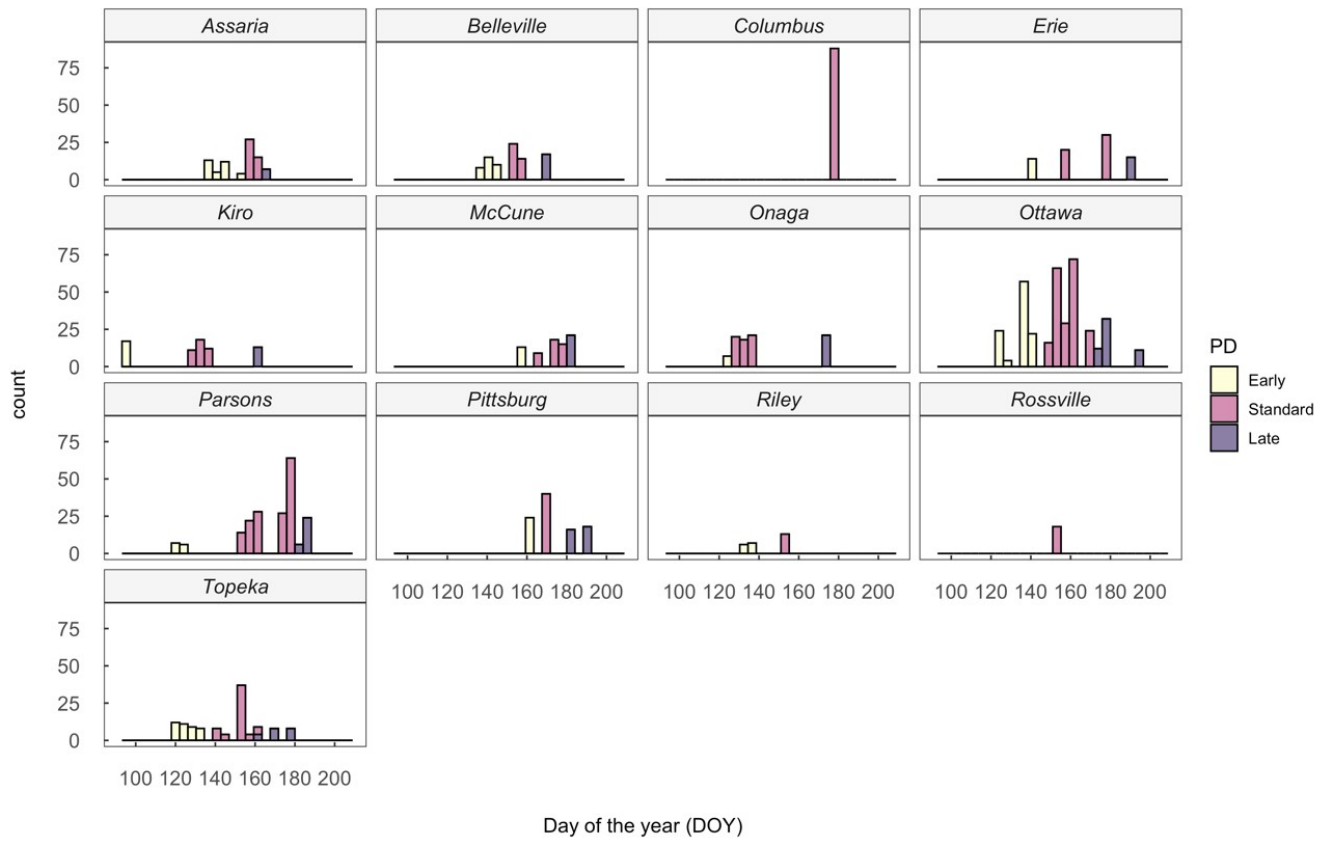


Figure A.1: Distribution of planting dates in the 13 sites of the dataset from 2014 to 2021.

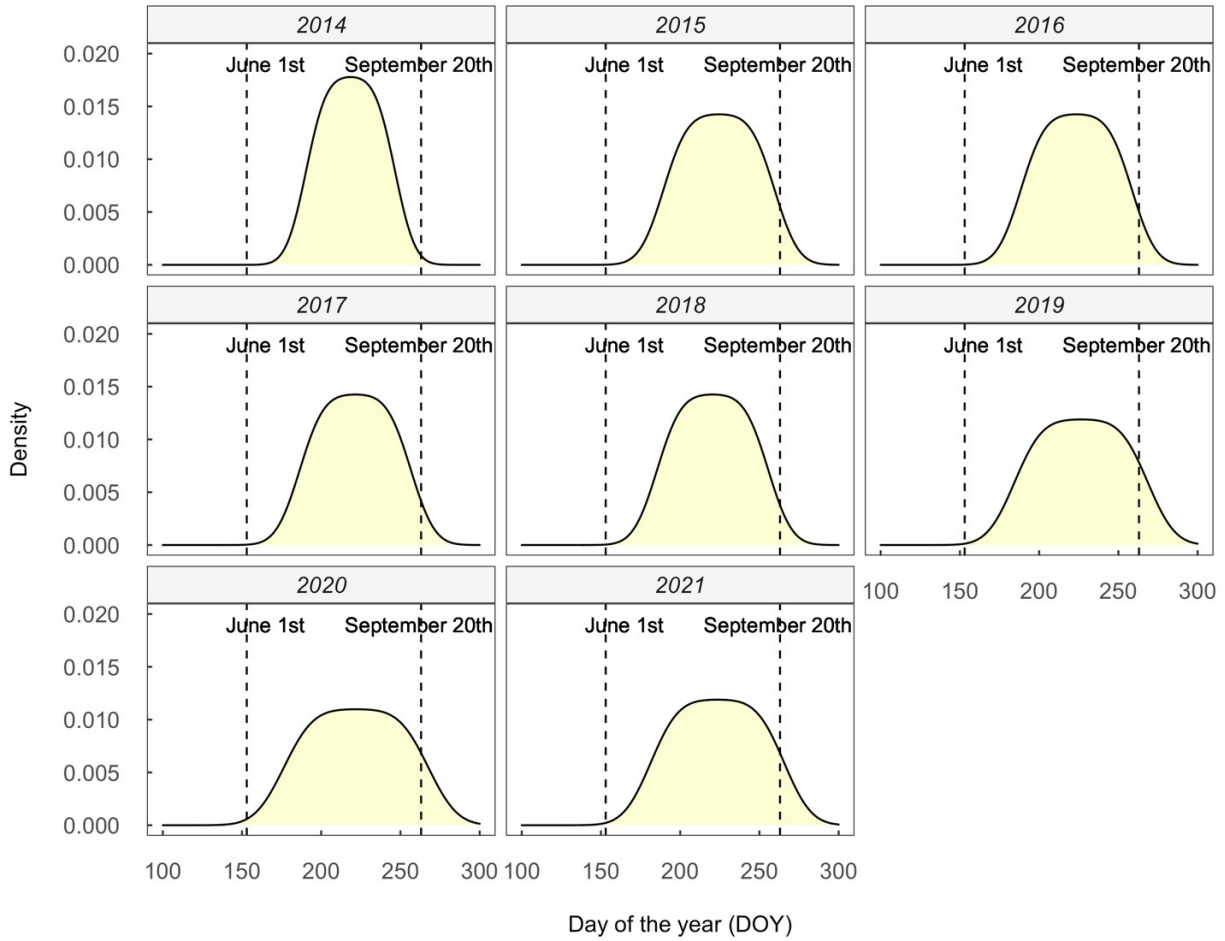


Figure A.2: Interval where setting pods occurs in the state of Kansas according to NASS-USDA (<https://quickstats.nass.usda.gov/results/20E59F49-96CD-3192-AEF5-6B5B63E28C43>, visited April 2022)

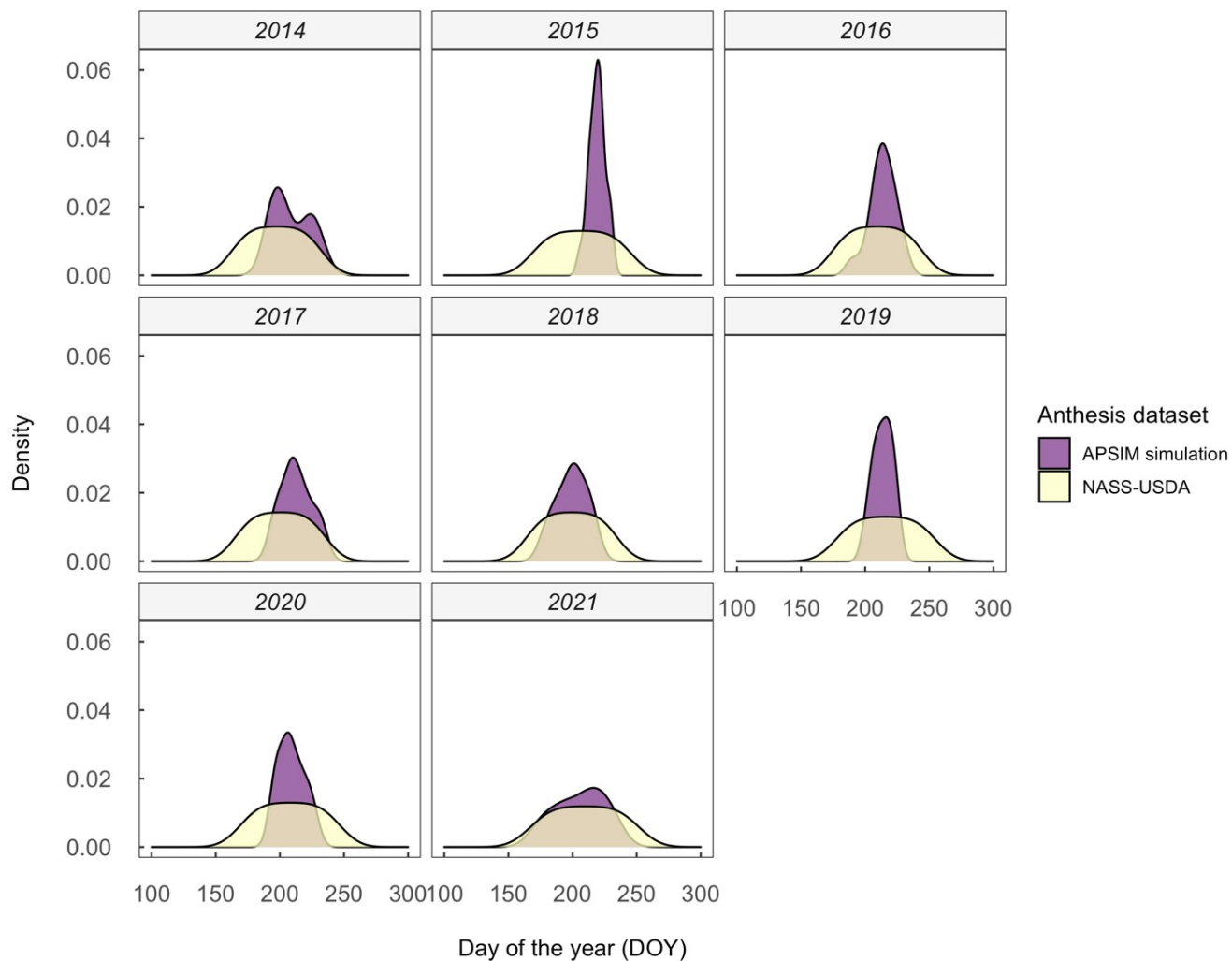


Figure A.3: Comparison of days of the year where anthesis occurs between the simulated data and data obtained from the National Agricultural Statistics Service (<https://quickstats.nass.usda.gov/results/21B7C2DA-A180-3470-A7F4-383E92588748>, visited April 2022)

Table A.4: Soybean seed yield (kg ha⁻¹) for the different planting dates $\times$ maturity groups combination in the South-east cluster. Values within different letters are significantly different at $P < 0.05$.

Cluster	Planting date	Maturity group	Mean yield (kg ha ⁻¹)
South-east	15-Apr	VI	4958 a
	15-Apr	V	4931 ab
	1-May	VI	4821 abc
	1-May	V	4818 abc
	15-Apr	IV	4808 abc
	1-May	IV	4750 abcd
	15-May	V	4742 abcd
	15-May	IV	4701 abcd
	15-May	VI	4624 bcde
	1-Jun	V	4530 cdef
	1-Jun	V	4519 cdef
	15-Apr	III	4429 defg
	15-Jun	IV	4319 efgh
	1-Jun	VI	4318 efgh
	1-May	III	4311 efgh
	15-May	III	4256 fghi
	15-Jun	V	4173 ghij
	1-Jun	III	4078 hijk
	15-Jun	VI	3949 ijk
	1-Jul	IV	3891 jkl
	15-Jun	III	3832 klm
	1-Jul	V	3618 lmn
	1-Jul	III	3527 mn
	1-Jul	VI	3520 mn
	15-Jul	IV	3357 no
	15-Jul	III	3145 op
	15-Jul	VI	3025 pq
	15-Jul	VI	2803 q

Table A.5: ANOVA (Type III Wald chi square tests) of the current weather (2012–2021) for the three clusters within Kansas.

Cluster	Component	P-value
North-central	Planting date	$< 2.2e-16$ ***
	Maturity group	$6.54e-10$ ***
	Planting date : Maturity group	0.9991
North-east	Planting date	$< 2.2e-16$ ***
	Maturity group	$2.15e-09$ ***
	Planting date : Maturity group	0.9989
South-east	Planting date	$< 2.2e-16$ ***
	Maturity group	$3.39e-10$ ***
	Planting date : Maturity group	0.0284 *
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1		

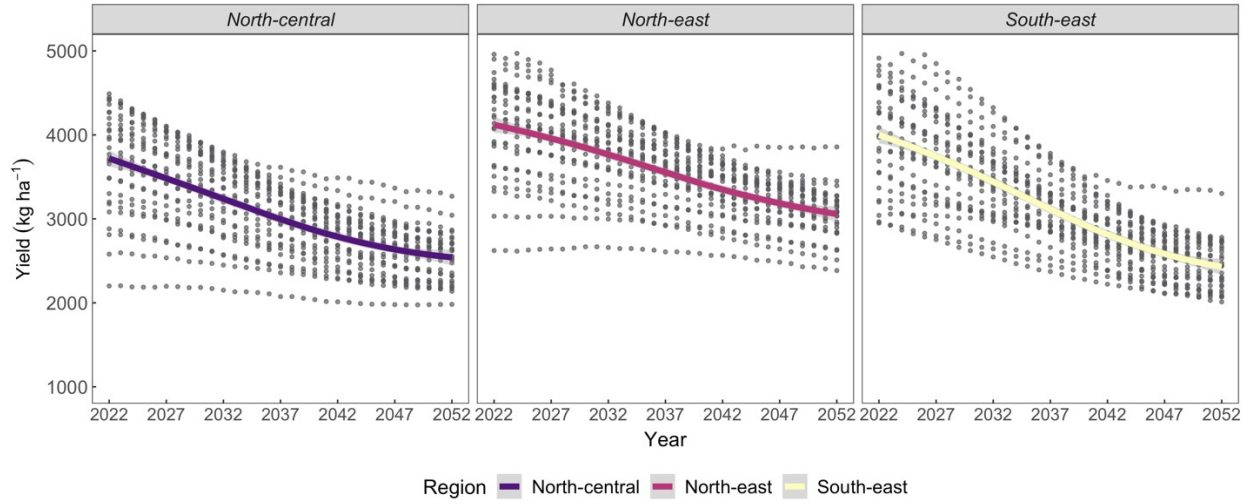


Figure A.4: Simulated seed yield ($kg\ ha^{-1}$) across the 30 years of study within the three clusters in Kansas (North-central, North-east, and South-east).

Table A.6: Climate change in the future (2043–2052) compared to the current weather (2012–2021) in the three clusters in Kansas.

Cluster	Weather variable	Winter	Spring	Summer	Fall	Year
North-central	Maximum temperature ($^{\circ}C$)	-0.33	-0.63	1.34	1.03	
	Minimum temperature ($^{\circ}C$)	0.04	-0.17	1.72	1.63	
	Cumulative rainfall (mm)	34	65	-171	-77	-148
North-east	Maximum temperature ($^{\circ}C$)	-0.22	-0.27	1.38	1.65	
	Minimum temperature ($^{\circ}C$)	0.19	-0.22	1.39	1.81	
	Cumulative rainfall (mm)	30	121	-162	-61	-73
South-east	Maximum temperature ($^{\circ}C$)	0.16	0.19	1.70	1.65	
	Minimum temperature ($^{\circ}C$)	0.46	0.15	1.55	1.81	
	Cumulative rainfall (mm)	-21	24	-223	-87	-308

Table A.7: Soybean seed yield ($kg ha^{-1}$) for the different planting dates $\times$ maturity groups combination within the three clusters in Kansas. Values within different letters are significantly different at $P < 0.05$.

Cluster	Planting date	Maturity group	Mean yield ($kg ha^{-1}$)
North-central	15-Apr	III	3380 a
	1-May	III	3185 b
	15-Apr	IV	3008 c
	15-May	III	2954 cd
	1-May	IV	2947 cd
	15-Jun	IV	2919 cd
	1-Jun	IV	2838 de
	15-May	IV	2836 de
	1-Jun	III	2832 de
	1-May	V	2799 de
	1-Jul	IV	2795 de
	15-May	V	2721 ef
	15-Jun	V	2710 ef
	15-Jun	III	2699 efg
	1-Jun	V	2697 efg
	15-Apr	V	2693 efg
	1-Jul	V	2575 fgh
	15-Jul	IV	2534 gh
	15-Apr	VI	2469 hi
	1-May	VI	2468 hi
	1-Jul	III	2436 hi

Continued on next page

Table A.7 – continued from previous page

Cluster	Planting date	Maturity group	Mean yield (kg ha ⁻¹)
North-east	15-May	VI	2336 ij
	15-Jul	III	2305 ij
	15-Jun	VI	2260 j
	15-Jul	V	2256 j
	1-Jun	VI	2250 j
	1-Jul	VI	2210 j
	15-Jul	VI	1986 k
	15-Apr	III	3851 a
	15-Apr	IV	3536 b
	15-Apr	VI	3504 bc
	15-Apr	V	3461 bcd
	1-May	III	3448 bcd
	1-Jun	V	3430 cd
	15-Jun	V	3396 de
	1-May	VI	3379 de
	15-May	V	3373 de
	15-Jun	IV	3370 de
	1-Jun	IV	3332 ef
	1-May	V	3310 efg
	15-May	VI	3256 fgh
	1-Jul	IV	3217 ghi
	1-May	IV	3194 hi
	15-May	IV	3188 hi

Continued on next page

Table A.7 – continued from previous page

Cluster	Planting date	Maturity group	Mean yield (kg ha ⁻¹)
	1-Jul	V	3177 hi
	15-May	III	3172 hij
	1-Jun	VI	3140 ijk
	1-Jun	III	3080 jk
	15-Jun	VI	3068 kl
	15-Jul	IV	2981 lm
	15-Jun	III	2954 m
	1-Jul	VI	2932 m
	1-Jul	III	2754 n
	15-Jul	V	2737 n
	15-Jul	III	2641 o
	15-Jul	VI	2478 p
	15-Apr	III	3369 a
	15-Jun	VI	2947 b
	1-Jul	VI	2917 b
	1-Jun	VI	2909 bc
	1-Jul	V	2847 cd
	15-Jun	V	2834 d
	1-May	III	2809 de
	15-May	VI	2807 de
	1-May	VI	2762 ef
	15-Apr	IV	2748 efg
	15-Jul	VI	2725 fg

Continued on next page

Table A.7 – continued from previous page

Cluster	Planting date	Maturity group	Mean yield (kg ha ⁻¹)
	1-Jun	V	2691 gh
	15-Jul	V	2657 hi
	15-Apr	VI	2650 hi
	15-Jun	IV	2617 i
	15-Jul	IV	2552 j
	1-Jul	IV	2537 j
	1-Jun	IV	2459 k
	15-May	III	2399 kl
	15-May	V	2390 l
	1-May	IV	2375 lm
	15-Jun	III	2313 mn
	15-Apr	V	2304 n
	1-May	V	2278 no
	1-Jun	III	2255 no
	15-Jul	III	2219 op
	15-May	IV	2163 pq
	1-Jul	III	2139 q

Appendix B

Appendix B, Chapter 3

Table B.1: The mean of temperature and cumulative rainfall from May to October across trials.

Year	State	Site	Temperature (°C)	Cumulative rainfall (mm)
2021	Kansas	Topeka	20	930
	Mississippi	Starkville	23	1211
2022	Kansas	Manhattan	20	668
	Mississippi	Starkville	23	761
		Stoneville	24	609
	South Dakota	Aurora	16	431
		Beresford	17	298

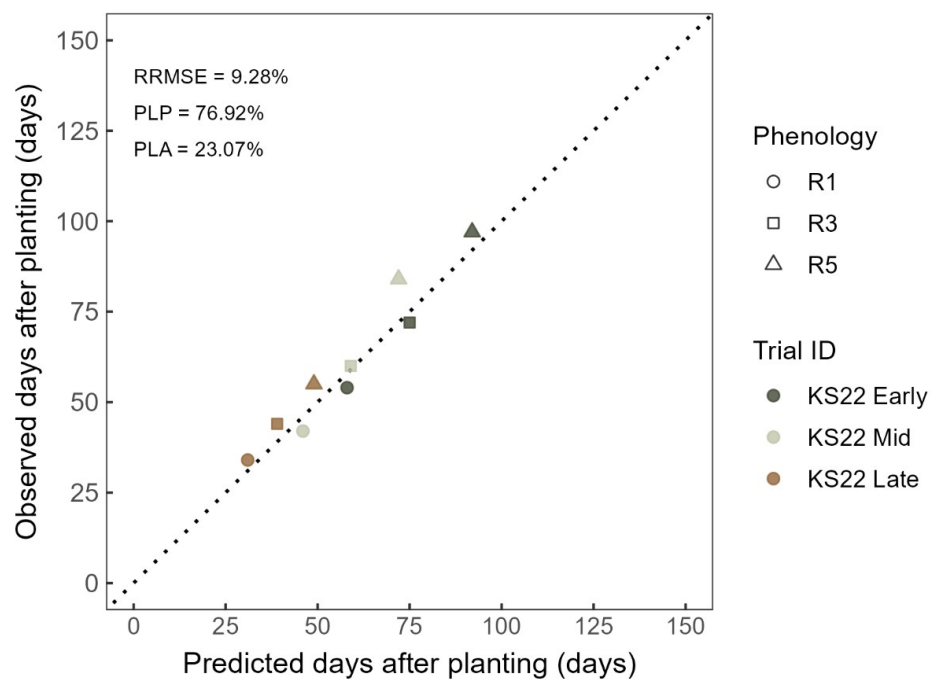


Figure B.1: DSSAT crop growth model performance predicting R1, R3, and R5 phenological stages in the KS22 Early, Mid, and late trials. Dashed line corresponds to 1:1 line.

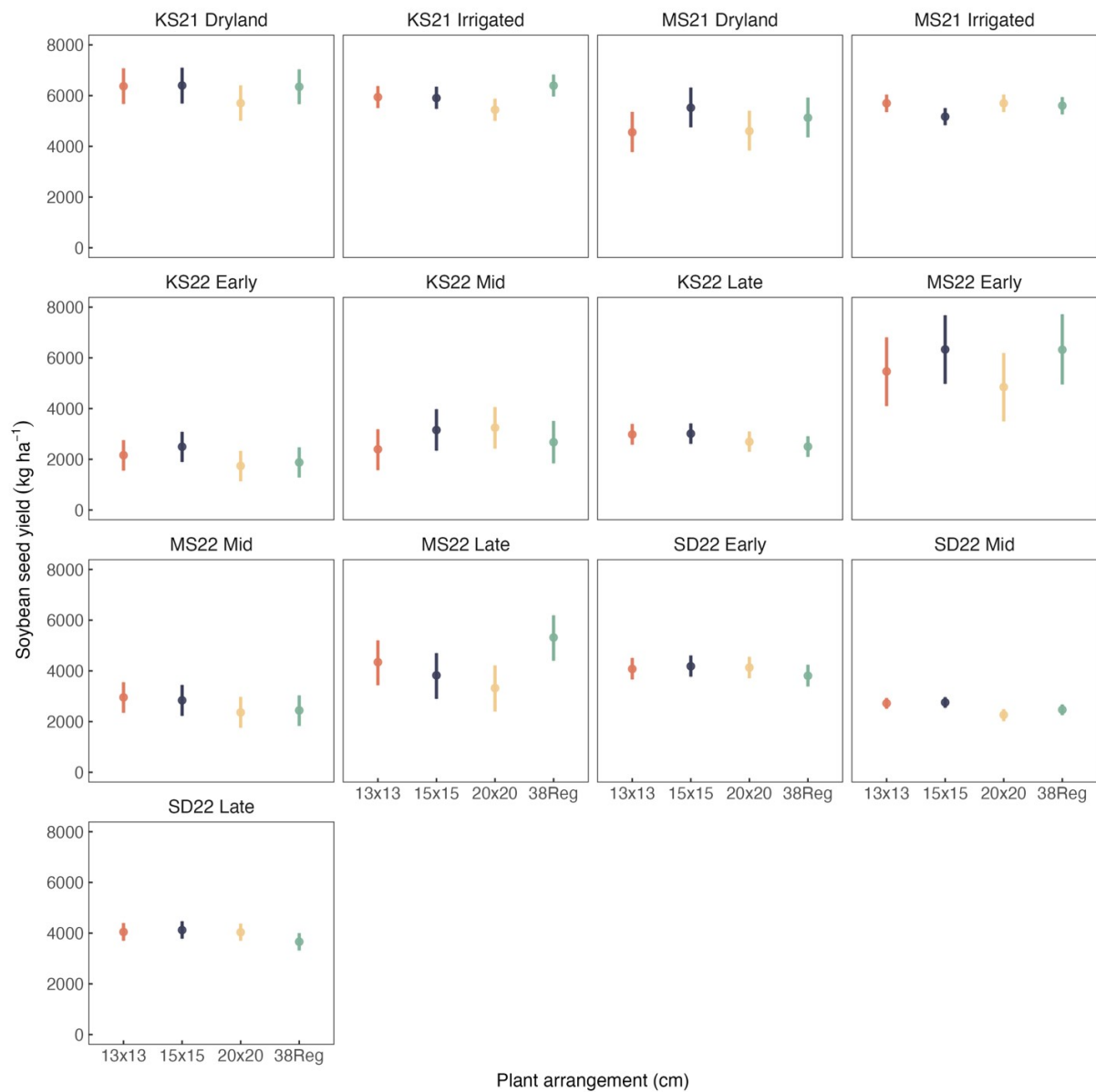


Figure B.2: Soybean seed yield and plant arrangement by trials. Bars represent 95% credible intervals. Points represent the average soybean seed yield. The colors of the points and bars indicate each plant arrangement. Panels represent each trial. KS21 Irrigated, MS22 Late, and SD22 Mid had differences.

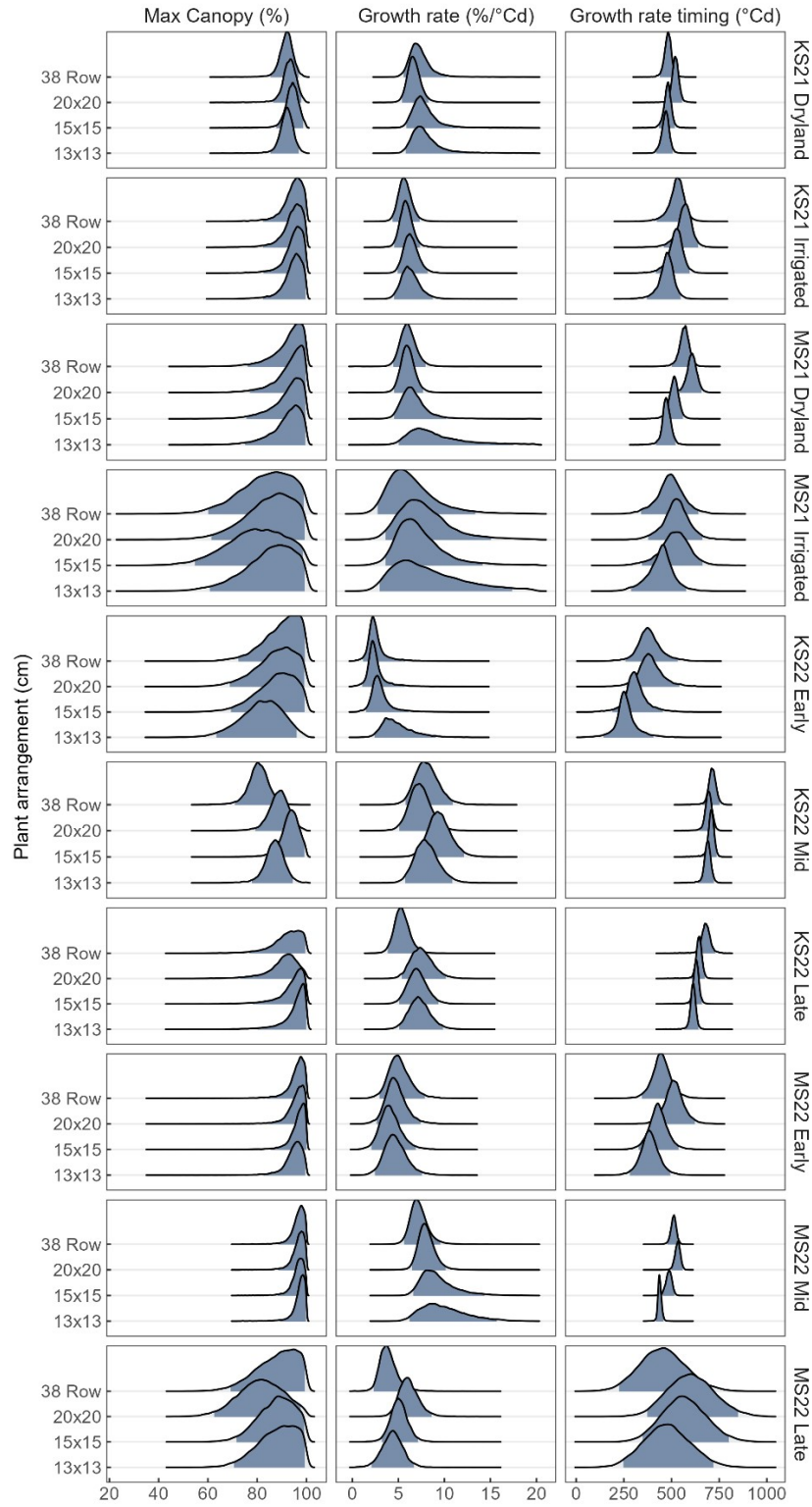


Figure B.3: Canopy coverage parameters distribution (from the hill function) by plant arrangement across trials (ID). BlueSky color under the curve represents the 95% confidence interval. Max Canopy (%) is the maximum canopy coverage.

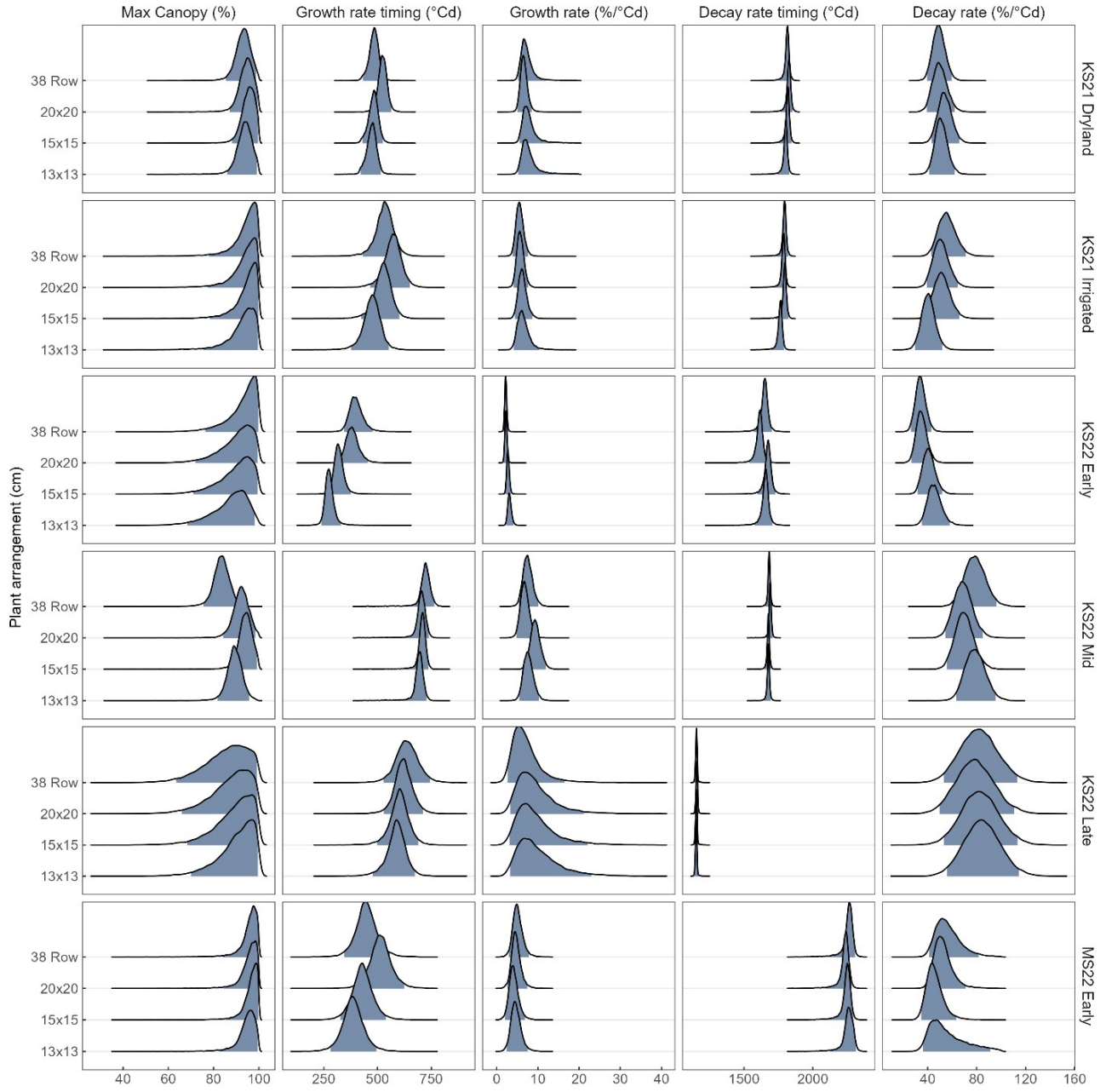


Figure B.4: Canopy coverage parameters distribution (from the bihill function) by plant arrangement across trials (ID). BlueSky color under the curve represents the 95% confidence interval. Max Canopy (%) is the maximum canopy coverage.

Appendix C

Appendix C, Chapter 4

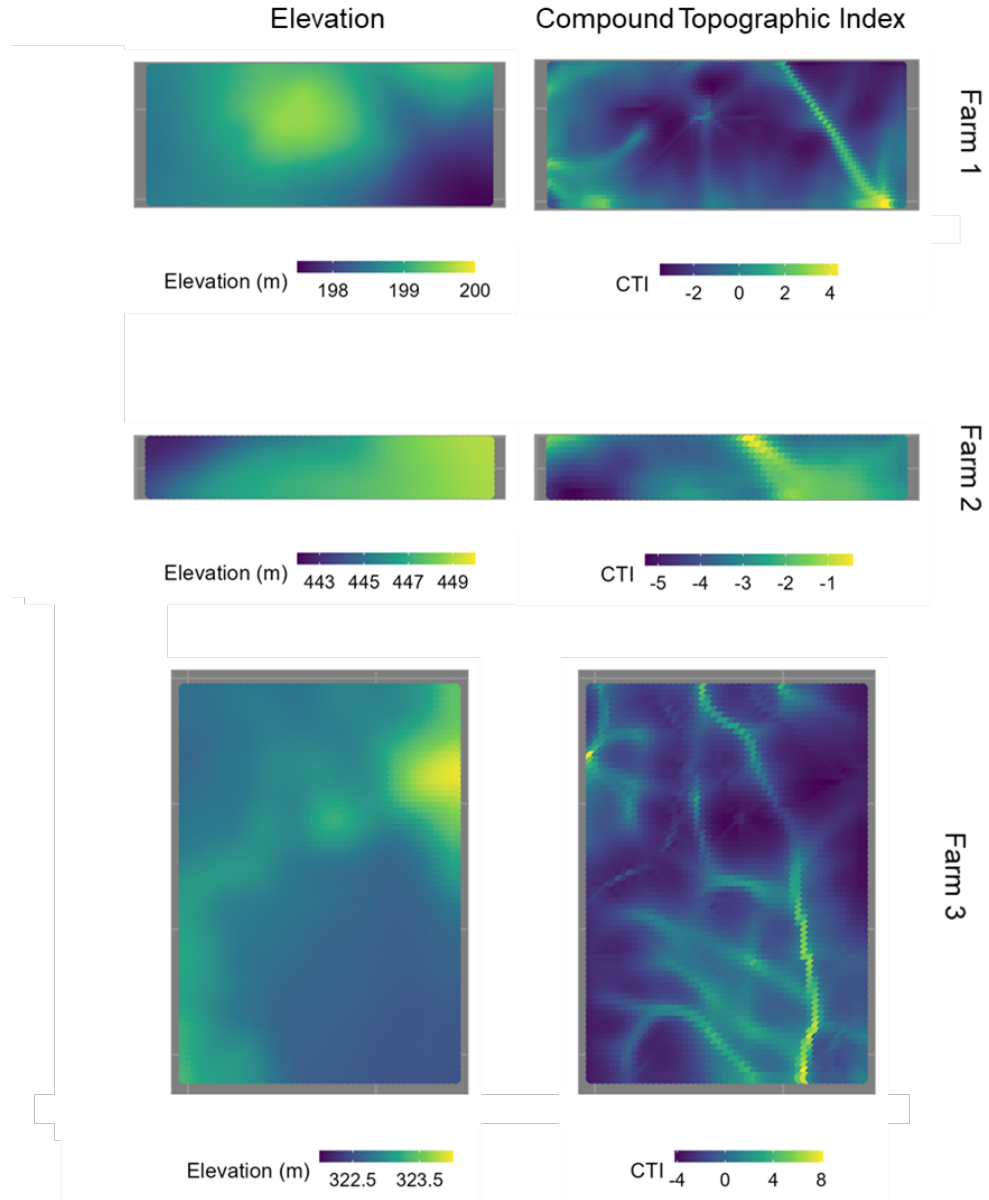


Figure C.1: Elevation and Compound topographic Index (CTI) for each Farm. Left column panel represent the elevation and right column panel represent the CTI. Each row corresponds to each Farm. Farm 1 correspond to case study 1 and 2, Farm 2 is the case study 3, and Farm 3 correspond to case study 4.

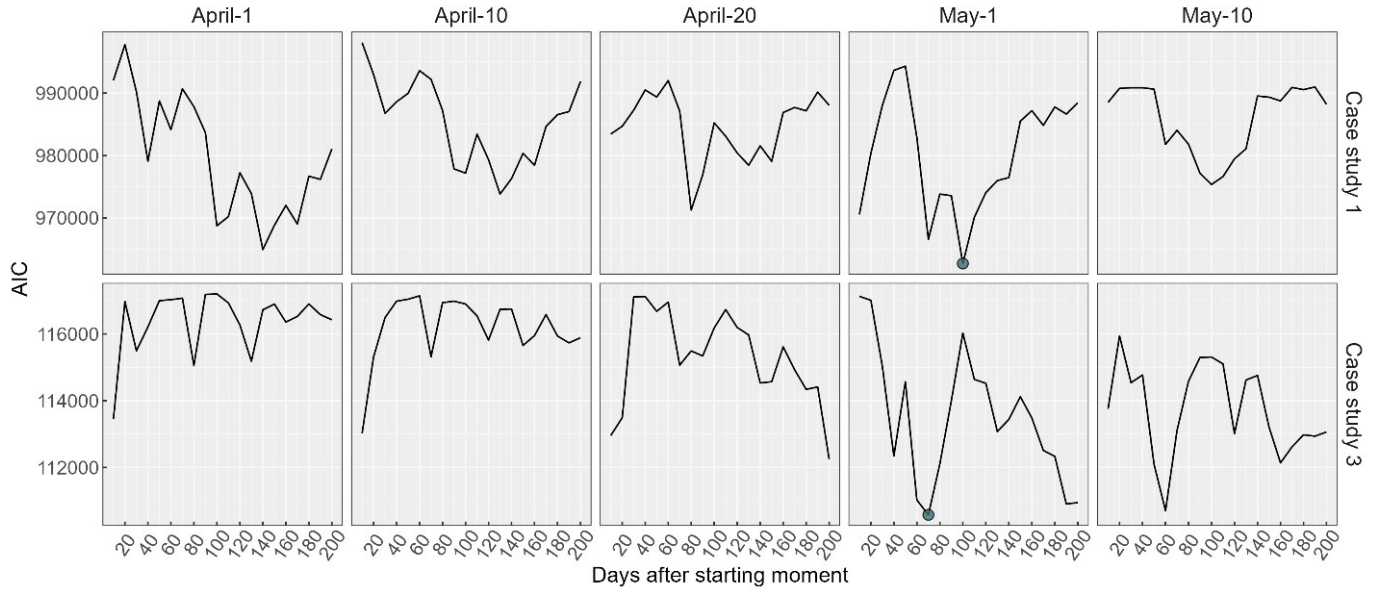


Figure C.2: Selection of optimum cumulative precipitation in Case study 1 and 3 (Maize). Each column of panels represents a different date of starting moment. Top panels are from case study 1 and panels in the bottom are from case study 3. Blue points represent the cumulative precipitation period with the lowest AIC.

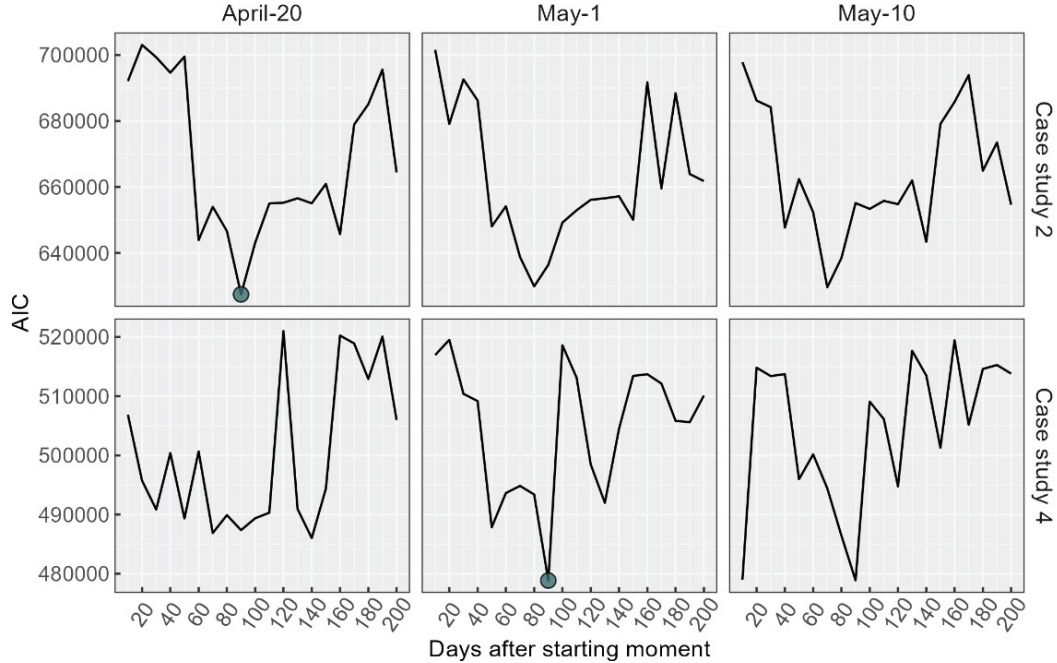


Figure C.3: Selection of optimum cumulative precipitation in Case study 2 and 4 (Soybean). Each column of panels represents a different date of starting moment. Top panels above are from case study 2 and panels in the bottom are from case study 4. Blue point represent the cumulative precipitation period with the lowest AIC.

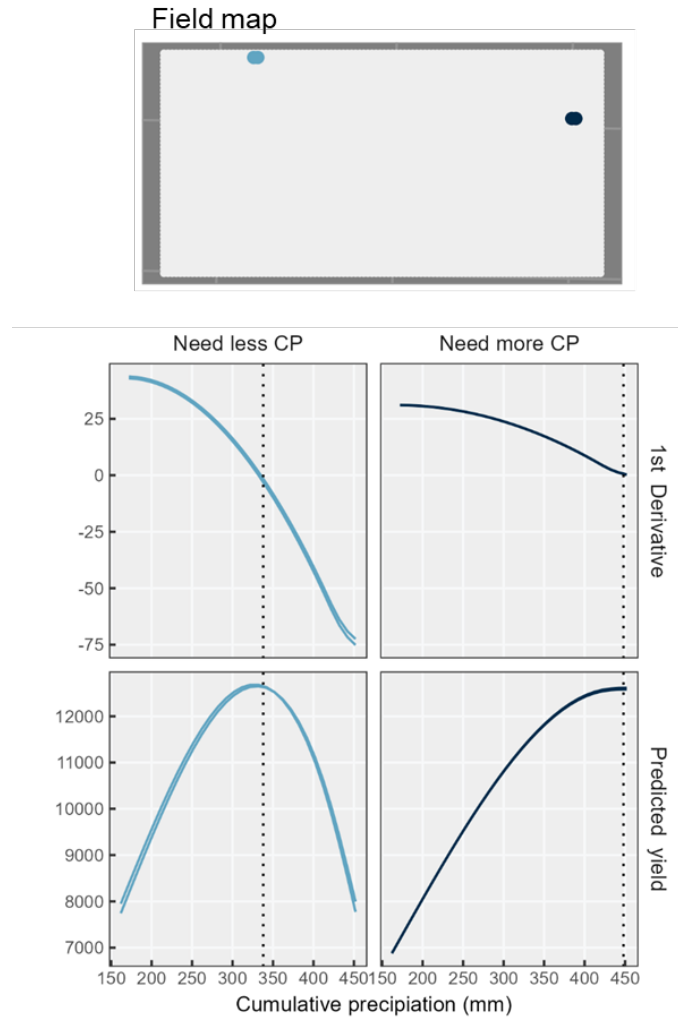


Figure C.4: Field map of case study 1 with two contrasting examples (above). Predicted yield and its derivative in response to cumulative precipitation in two contrasting examples in case study 1. Above panel represent the 1st derivative and below panel represent the predicted yield (Below). Colors represent the two contrasting examples within the field. CP: Cumulative precipitation.

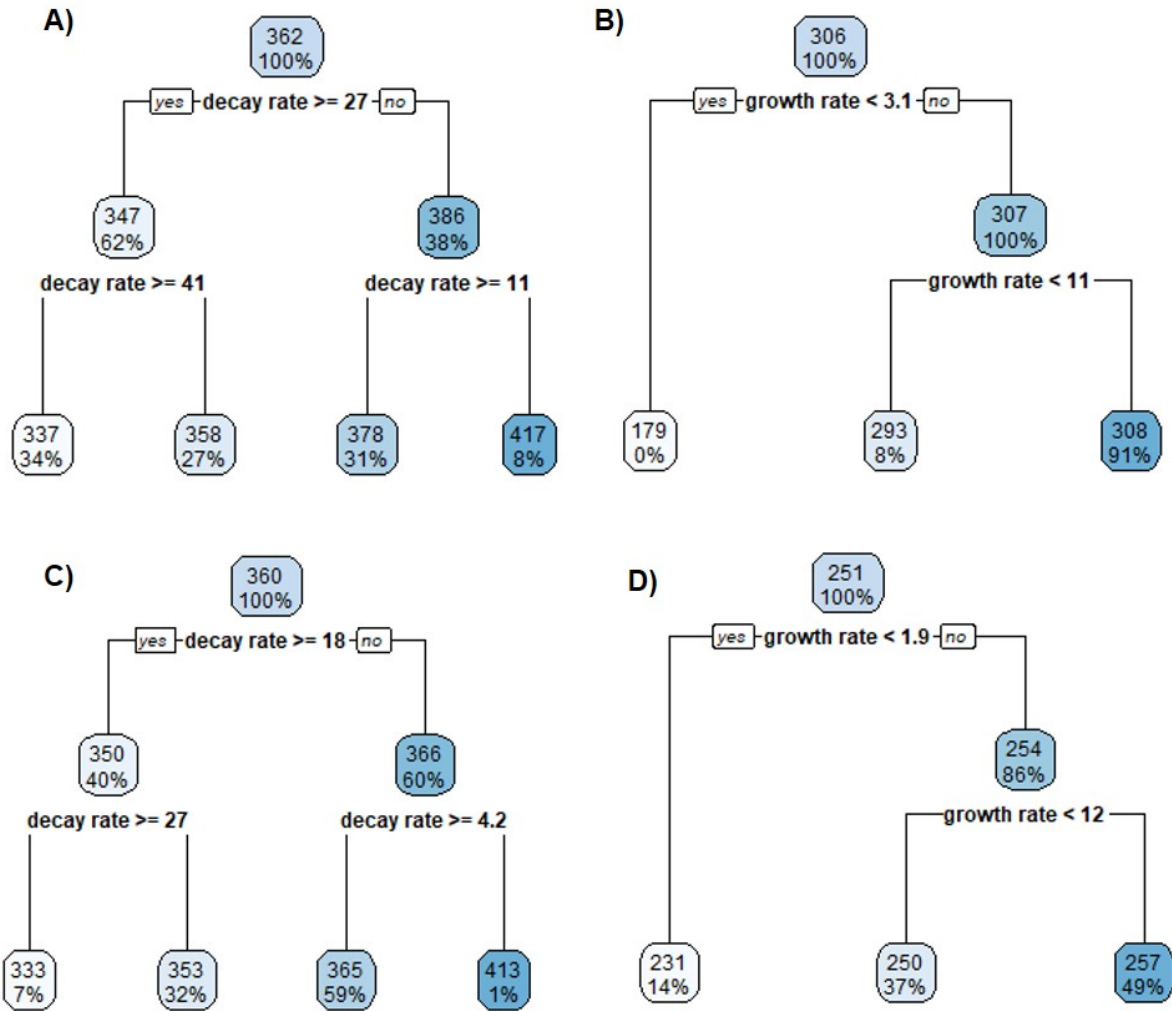


Figure C.5: Regression tree for case study 1 (A), case study 2 (B), case study 3 (C), and case study 4 (D) to classify precipitation value across a field using spatial growth and decay coefficients. Figures in the left correspond to maize fields case studies and figures in the right correspond to soybean fields case studies.