A non-parametric approach to estimating nonlinear impacts of early season soil moisture on dryland corn yields

by

Parker Vulgamore

B.S., Kansas State University, 2023

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Approved by:

Major Professor
Jennifer Ifft, Ph.D.

Copyright

Abstract

Early season soil moisture is one of the few predictive factors available to agricultural decision-makers before planting begins. Farmers, merchandisers, insurers, and futures market participants often incorporate early season moisture conditions into their production forecasts and risk assessments, yet it remains unclear whether these assessments are supported by measurable impacts on dryland corn yield outcomes. While previous studies have explored the relationship between growing season moisture and yield, the role of early season soil moisture in isolation has been largely overlooked.

This study makes three key contributions. First, it provides a model for early season yield forecasting. Results indicate that shifting just one of the 30 early-season days from average to the driest soil moisture conditions is associated with a 3% national yield decline—roughly 446 million fewer bushels relative to April 2025 WASDE projections. Second, it demonstrates that the effects of early season soil moisture are highly nonlinear, underscoring the importance of using a flexible, exposure-based approach rather than assuming a parametric functional form. Third, it highlights significant regional differences in soil moisture-yield relationships, emphasizing the need for spatially specific modeling approaches. We found that exposure to low soil moisture is associated with increased yield in the central and northern corn belt, whereas it is associated with decreases in yield in other regions and the national model. This study offers a framework for improving the assessment of early-season production expectations, particularly for farmers, merchandisers, insurers, and policymakers whose decisions rely on production assumptions informed by limited information available at that time.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This work was supported by effort from the Rural and Farm Finance Policy Analysis Center (RaFF) at the University of Missouri. RaFF aims to provide objective analysis and inform decisionmakers on issues affecting farm and rural finances.

There are several people I am obligated—by social convention and moral duty—to thank for their invaluable contributions to this research. First, profound thanks to ChatGPT, whose assistance in debugging code was greatly appreciated. Admittedly, its true magnum opus was helping craft parts of this very acknowledgments section, though exactly which parts shall remain undisclosed.

On a marginally more genuine note, heartfelt gratitude goes out to my family for their unwavering support throughout graduate school. Thanks to my dad, whose repeated declarations of learning "more in one year of grad school than in four years of undergrad combined" were both motivational and concerning. Special appreciation to my mom, my original human ChatGPT, whose English-teacher expertise served as my personal writing assistant long before artificial intelligence. I'd also be remiss not to acknowledge my brother and sister, whose exciting adventures consistently reminded me why staying close to home for grad school was the only reasonable choice.

Huge thanks to my fellow grad students, friends, and especially my roommates, whose unwavering commitment to Nancys at Taco Lucha and impromptu pickleball matches kept us all sane—or at least distracted. Special appreciation goes to my fiancée, Marissa, who managed the remarkable feat of working part-time, attending night classes, and planning our wedding, all while patiently tolerating the whims and frustrations of a stressed-out grad student. Together, we accumulated not only knowledge but enough empty wine bottles to sustain the economy of a small developing nation for several years.

Profound thanks also to the professors who invested heavily in my personal and professional growth. Dr. Britton, my undergraduate advisor turned lifelong mentor and friend, deserves a special shoutout. From countless hours of advice to exactly 1,314 letters of recommendation (yes, I counted), I genuinely wouldn't be where I am without him. Deep appreciation also goes to my major professor, Dr. Jenny Ifft, who enthusiastically supported every weird research rabbit hole I insisted on chasing—this study included—despite them being far beyond her typical domain of expertise.

The lion's share of credit for this research, however, belongs unequivocally to Dr. Micah Harp, who is responsible for ideation, coding wizardry, data preparation, and significant portions of the writing. His patience is matched only by his genius, and he's precisely the type of advisor every confused grad student hopes they'll stumble upon. Additionally, Dr. Jesse Tack deserves sincere thanks for his invaluable contributions and insights throughout the development of this study.

Thank you all, and sincere apologies that I've betrayed you by selling out and deciding to attend law school instead.

Chapter 1 - Introduction

The complex interplay between soil moisture and production can lead farmers to make decisions regarding crop selection, input application, marketing strategies, and even insurance coverage before their seeds are ever in the ground (Jagnani, et al. 2020, Bille and Rogna 2021). Understanding the significance of early season soil moisture is crucial, as it is one of the few predictive factors available to decision-makers before planting begins. Furthermore, this perception of early season soil conditions impacts stakeholders beyond just producers.

Soil moisture is one of the few predictive indicators available to decision-makers, shaping not only farmers' strategies but also those of elevator managers, grain merchandisers, feedlot operators, and futures market participants who rely on early season field conditions for logistical and economic planning. Feedlot managers and other local end-users responsible for sourcing inputs and forward contracting also have a stake in future production. Speculators and futures market participants similarly monitor precipitation and soil conditions in the weeks leading up to planting, seeking to predict how supply will be impacted.

However, there are many factors throughout the growing season contributing substantially to eventual crop yields (Bailey-Serres, et al. 2019, Schlenker and Roberts 2009, Chen, et al. 2023, Ortiz-Bobea, Wang, et al. 2019). Given this dynamic nature, it is difficult to ascertain the effect of any one individual contributor such as soil moisture (Denmead and Shaw 1960). These widespread discussions raise an important question: Are decisions made based on early season soil moisture information supported by measurable impacts on dryland corn yields? While several studies explore the dynamic relationship between soil moisture or precipitation and crop yields throughout the growing season (Steduto, et al. 2012, Cakir 2004, Ortiz-Bobea, Wang, et al. 2019), early season conditions in isolation are often overlooked.

In this paper, we explore whether early season soil moisture data offers meaningful predictive value for dryland corn yields in the United States, separate from the growing season conditions that remain uncertain at planting time. We further examine whether the effects of soil moisture are nonlinear, recognizing that the impact of moisture on yield may not be constant across different exposure levels (Y. Li, et al. 2019). Rather than assuming a linear or parametric functional form, we estimate the change associated with soil moisture exposure at different levels, allowing the data to reveal how yields respond across the full distribution of early season moisture conditions. This exposure-based approach is inspired by Schlenker and Roberts 2009 and Ortiz-Bobea 2019. It provides a more flexible and interpretable estimation, identifying whether specific thresholds or inflection points exist where moisture levels become particularly beneficial or harmful.

Additionally, we evaluate the spatial heterogeneity of this relationship across growing regions. Agricultural productivity is shaped by regional differences in soil type, climate, and cropping systems, meaning that soil moisture conditions do not influence all regions equally (Dayanto, Wang and Jacinthe 2016, Wu, et al. 2007). A national model may mask important localized dynamics, leading to misleading conclusions about the role of early season moisture in yield determination. Understanding these regional variations is critical for ensuring that yield predictions reflect real-world agronomic conditions and provide actionable insights for producers, insurers, and policymakers.

The results indicate that exposure to the lowest extreme levels of soil moisture have the strongest effect on yield. Across nearly every region and specification, there was a nonlinear effect of exposure across soil moisture levels. However, the strength of the correlation and magnitude of the effect varies substantially across different regions, emphasizing the importance of testing spatial heterogeneity.

Lastly, we extend the primary yield-based analysis by estimating a model using crop insurance loss ratio data as the dependent variable. Loss ratios offer a complementary perspective, capturing the extent to which realized revenue or yield falls short of guaranteed coverage and thus triggering indemnity payments. This provides a more direct economic lens through which to assess the impact of early season moisture stress. The loss ratio models yield similar patterns to the yield-based results, reinforcing the evidence of threshold effects and regional heterogeneity. By examining both yield and loss ratio outcomes, this paper aims to broaden our understanding of how early season soil moisture influences production, risk exposure, and decision-making across the agricultural sector.

Our study makes three primary contributions:

1.1 Predictive Ability

The first major contribution of this study is the development of a model that evaluates the predictive power of early season soil moisture data. By relying solely on soil moisture information available at planting and incorporating a linear time trend, the model enables early season yield forecasting and projection, providing a valuable tool for decision-making before the growing season unfolds.

To ensure the robustness of our approach, we tested multiple model specifications, varying both the time window length and the placement of soil moisture breakpoints. The early season period represents an arbitrary choice in most agronomic and economic studies, yet it has profound implications for predictive performance. We systematically evaluated models using different early season window lengths, ranging from 30 to 90 days, to determine whether expanding the exposure period improved or weakened the model's ability to explain yield variability. While longer time windows provide more data points for estimation, they also risk diluting the effects of early season moisture exposure, as other environmental conditions begin influencing yields.

Our results demonstrate that a 30-day early season window provides the best balance—ensuring enough observations within exposure bins to maintain statistical power while preserving the assumption that early season soil moisture effects are cumulative and additive over time. This specification captures substantial variation in yield outcomes while preventing the inclusion of unrelated seasonal effects that would weaken the model's predictive ability.

Across countless time window specifications, national and subnational models, and different geographic aggregations, we consistently find soil moisture exposure to be a statistically significant determinant of yield. Model fit indicators, such as the adjusted R^2, range from 0.2 to over 0.5, suggesting that in some cases, half of the variation in yield outcomes can be explained using only information available before planting even begins and a county fixed effect. While model fit alone is not a definitive measure of true predictive ability, these findings underscore the potential value of early season soil moisture as a forecasting tool.

1.2 Non-linearity

The second major contribution of this study is how the change associated with soil moisture exposure varies across different moisture levels. Rather than assuming a parametric functional form—such as a linear, quadratic, or logistic relationship, we employ an exposure-based approach that allows for flexible estimation of coefficients at different moisture thresholds. This method follows the approach of Schlenker and Roberts 2009, who estimated nonlinear temperature effects on yield using a similar framework (Schlenker and Roberts 2009). Additionally, Ortiz-Bobea 2019 followed a similar exposure approach to breakdown how different factors contribute to yield outcomes throughout the growing season (Ortiz-Bobea, Wang, et al. 2019).

A key challenge in estimating nonlinear effects is determining where to place breakpoints to best capture threshold effects. If breakpoints are spaced too narrowly, the model risks overfitting, particularly when the number of observations per bin is small. Conversely, if breakpoints are too widely spaced, the model may miss important inflection points where yield responses change direction. To address this, we systematically tested multiple binning strategies to find a data-driven balance—ensuring that each bin contained enough observations to provide stable estimates while still preserving flexibility in capturing nonlinear effects.

This approach offers two key advantages. First, it allows us to estimate how the yield impact of soil moisture exposure varies across its full distribution, rather than assuming that each additional day of moisture exposure has a constant effect. This is particularly important for detecting critical moisture thresholds, beyond which additional wet or dry days may have diminishing or compounding effects on yield.

Second, it provides a more precise and interpretable estimation of associated coefficients compared to traditional models that rely on average soil moisture values or assume a predefined nonlinear structure. By allowing soil moisture exposure to enter flexibly into the model, this study uncovers insights into the timing and intensity of moisture stress or sufficiency, helping to better inform decision-makers.

1.3 Spatial Heterogeneity

The third major contribution of this study is the identification of significant spatial heterogeneity in the relationship between soil moisture exposure and yield across regions and states. To assess how well the national model generalizes, we estimated separate models for ERS Farm Resource Regions, agroecological zones, and an individual state. While some regions—consistent with the national model—showed that low soil moisture levels were detrimental to yield, others exhibited the opposite relationship, where drier conditions were positively correlated with yield.

Regions receiving above average annual rainfall tended to benefit from drier early season conditions, aligning with the intuitive notion that excessive rainfall can delay planting, leach nutrients, and lead to waterlogging. Conversely, in drought-prone regions, lower soil moisture levels were strongly negatively correlated with yield, reinforcing the importance of sufficient early season moisture in areas with lower baseline precipitation.

These findings underscore a key insight: the relationship between early season soil moisture and yield is highly region-dependent, and even the functional form of this relationship varies across spatial scales. The national model effectively averages over these distinct relationships, which suggests that it may perform poorly out of sample compared to more localized regional models. This highlights the importance of spatially tailored modeling approaches when estimating the effects of soil moisture on yield, particularly for applications in yield forecasting, risk management, and policy design.

Chapter 2 - Background

2.1 Soil Moisture and Yield Response Discussion

Many factors prior to and throughout the growing season influence eventual crop yields. The interactions between crops, climate, water, and soil are complex, and much of the existing literature has explored these dynamics from various angles.

The specific factors discussed in the literature include drought (Li, et al. 2009), excessive rainfall (Y. Li, et al. 2019), extreme planting conditions (Urban, et al. 2015), temperature exposure (Tack, Barkley and Nalley 2015, Schlenker and Roberts 2009), technological advancements (Bailey-Serres, et al. 2019), and agronomic change (Pingali 2012). Additionally, research has focused on the adaptive decisions farmers make in response to these conditions (Jagnani, et al. 2020, Bille and Rogna 2021) .

Soil moisture is a key variable in hydrology, meteorology, and agriculture, with surface soil moisture being particularly variable across space and time (Korres, Reichenau and Schneider 2013). In agricultural landscapes, soil moisture patterns are influenced by both natural factors such as precipitation, soil composition, and topography, and agro-economic factors including soil management and fertilization practices (Korres, Reichenau and Schneider 2013). This variability makes it challenging to establish direct cause-and-effect relationships between yield and its influencing variables.

Short periods of very high spring soil moisture can significantly affect crop yields (Urban, et al. 2015). Large precipitation events can reduce the number of workable days in the field, delaying planting and increasing the likelihood of maize losing pollen viability. Flooding can also cause cellular damage through oxygen-related pathways, as seen in both maize (Yan, et al. 1996) and soybeans (VanToai and Bolles 1991). In some cases, severely affected areas may need to be replanted or left unplanted.

As precipitation deviates toward either extreme dryness or excessive rainfall, yield losses increase substantially (Y. Li, et al. 2019). On average, extreme drought conditions ($<-2\sigma$) reduce maize yield by −32.2%, while extreme rainfall ($>2.5\sigma$) leads to a −16.6% reduction. The most severe rainfall events ($>3.5\sigma$) cause yield losses (−34.1%) that are comparable to the most extreme droughts ($<-2.5\sigma$, −36.6%), indicating that excessive rainfall can be just as detrimental as drought under certain conditions.

Other research confirms that both drought and excessive rainfall have different effects depending on the crop growth stage at the time. (Dayanto, Wang and Jacinthe 2016, Kanwar and Mukhtar 1988, Evans and Fausey 1999). Extreme weather events throughout different stages of the growing season have been studied extensively, though most work has not isolated pre-planting conditions (Chen, et al. 2023). Additional studies have shown that the effects of soil moisture stress differ across different stages of corn growth, with evidence from both irrigated and dryland field studies (Denmead and Shaw 1960, Cakir 2004).

Beneficial dryland crop management practices—such as crop rotation, reduced or no-tillage, and residue management systems—are essential to conserving and efficiently utilizing limited water resources, particularly in regions where water availability at critical growth stages largely determines crop productivity (Unger, et al. 1991, Farahani, Peterson and Westfall 1998, Schlegel, et al. 2019). Several studies show that practices enhancing soil water conservation directly contribute to higher yields (Lafond, Loeppky and Derksen 1992, Peterson, et al. 1996).

Our study contributes to this body of literature by explicitly examining how early season soil moisture levels, as indicators of available soil water at planting, influence dryland corn yields

2.2 Mechanisms Linking Early Season Soil Moisture to Yield Outcomes

This study measures the net impact of soil moisture exposure on yield outcomes driven by: (1) the direct, agronomic relationship, (2) correlation with seasonal precipitation, and (3) producer-induced impacts. These mechanisms are broken down in the following sections.

2.2.1 Agronomic Relationship Between Early Season Soil Moisture and Yield

We hypothesize that early season soil moisture promotes strong root development, setting a foundation for robust plant growth and higher yields. Pre-growing-season soil moisture conditions affect soil inorganic nitrogen (SIN) content, with implications for corn yield (Z. Li, et al. 2022). In wetter early season conditions, increased leaching reduces SIN content, leading to yield losses of 0.54–0.86 Mg/ha (5–14%) under no fertilizer and 0.21–0.33 Mg/ha (1–3%) with normal nitrogen fertilization (167 kg N/ha).

Other research has examined the role of planting conditions in influencing crop outcomes. One study reported that short periods of very high spring soil moisture significantly reduce corn and soybean yields, and incorporating these extreme conditions into yield models improves predictive performance (Urban, et al. 2015). Additional work has focused on the use of irrigation to manage planting conditions. For example, research has explored how pre-plant irrigation can be strategically applied to improve planting conditions and ultimately support crop growth (Kisekka, et al. 2017). Similarly, studies have examined the effects of managing off-season soil moisture through irrigation, finding that while in-season water availability remains critical for grain yields, optimal pre-planting moisture can also lead to positive yield outcomes (Stone, et al. 1987, Stone, et al. 2008). These findings further support the hypothesis that soil moisture conditions prior to planting play a meaningful role in determining yield potential.

2.2.2 Correlation with Seasonal Precipitation

Another avenue of exploration considers whether early season soil moisture correlates with growing season precipitation through El Nino-Southern Oscillation (“ENSO”) effects. The hypothesis suggests that wet conditions before planting could signal a wetter-than-average growing season, potentially enhancing yields. Conversely, dry early season conditions could indicate sustained drought throughout the growing season.

One study examined precipitation patterns in eastern China, finding that higher spring precipitation in the Yangtze River valley correlates with increased summer precipitation in northeastern China but decreased rainfall in southeastern China (Liu, Zhang and Zuo 2017). Although geographically specific, these findings highlight the potential for inter-seasonal moisture relationships.

Studies in the United States have proposed that reduced soil moisture during late winter or spring in mid-continental regions (e.g., central U.S.) could contribute to warm and dry summers through decreased evaporation and altered atmospheric circulation (Namias 1991). Additionally, extremely dry months tend to cluster more frequently than wet months, suggesting asymmetry in precipitation patterns (Gilman and Riedel 1951).

Another study used singular value decomposition (SVD) to analyze spring soil moisture and summer precipitation across the continental U.S (Wu, et al. 2007). Positive and negative correlations were observed in different U.S. regions. For instance, the Southern Great Plains exhibited strong negative correlations, meaning high spring soil moisture sometimes corresponded with reduced summer precipitation whereas the opposite was true in other regions. Geographic specific, land-atmosphere feedbacks were identified as primary mechanism to this phenomenon. While there is evidence that spring soil moisture correlates with summer precipitation, the relationship is spatially and temporally variable. Certain regions, like the Southern Great Plains, exhibit strong interactions, while others show mixed or weak patterns. To consider this effect, the current study incorporates region-specific models.

2.2.3 Producer-Induced Impacts

Farmers may adapt their practices in response to early season moisture conditions, such as adjusting fertilization rates or planting schedules (Bille and Rogna 2021, Jagnani, et al. 2020). These adaptations, while critical to understanding real-world yield outcomes, introduce additional variability into the analysis.

For instance, producers may react positively to favorable conditions (e.g., increasing inputs) or negatively to unfavorable conditions (e.g., reducing inputs), potentially influencing yields independently of soil moisture. To mitigate the influence of these adaptations on the results, this study focuses exclusively on dryland corn yields, excluding irrigated fields. This study captures the net effect of producer-induced impacts on yield.

Chapter 3 - Data

3.1 Yield Data

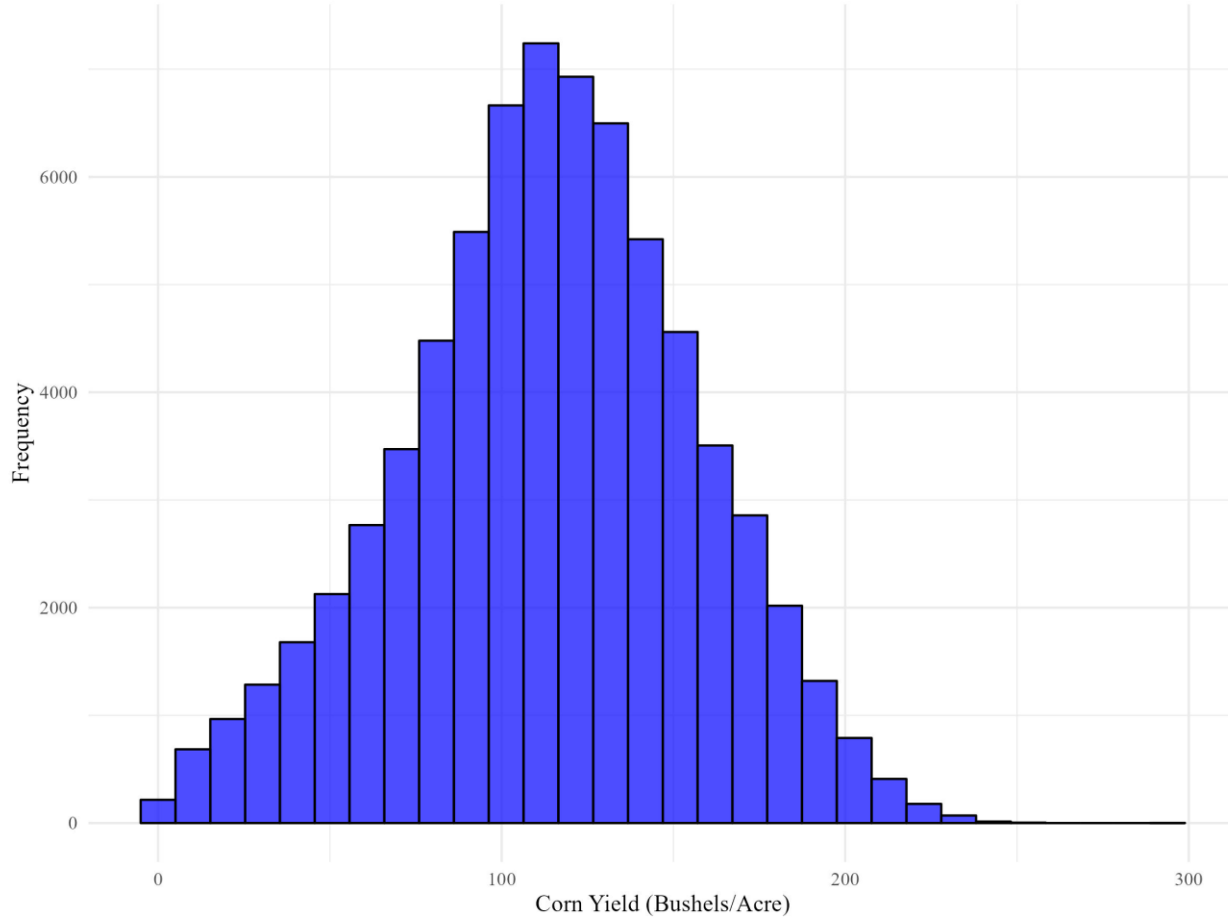
This study utilizes an unbalanced panel of county-level crop yields, compiled from the United States Department of Agriculture (USDA) Risk Management Agency (RMA) (USDA Risk Management Agency 2022). The RMA dataset was chosen over the commonly used NASS county-level yield datasets due to its ability to provide more representative data when controlling for irrigation status, which has a substantial effect on the variability and outcomes of yields. The current study focuses exclusively on dryland/non-irrigated corn yields.

When matched with counties for which soil moisture data is available, the panel includes yield data for 2,318 U.S. counties from 1990 to 2022. Due to the unbalanced nature of the panel, not all counties have a complete series of 31 years, with the minimum number of observations from any one county being 22. We also estimate the model with a balanced panel as a robustness check with the results reported in the appendix. In total, the dataset includes 71,644 unique county-year observations. Descriptive statistics for the yield data are provided in **Table 3.1. Figure 3.1** is a histogram of the yield observations used in the model.

Table 3.1 Descriptive Statistics: Crop Yields

Crop	Mean	Minimum	Maximum	Std. Dev.	Count
Corn (Bushels/Acre)	113.77	1	294.8	42.59	71,644

Figure 3.1 Histogram of Dryland Corn Yield Observations



We apply a logarithmic transformation to the yield data for the regression analysis. To accommodate this, eight observations with a yield of 0 and 44 observations with yields below 1 bu/ac were dropped, facilitating an easy logarithmic transformation. Additionally, four extreme outliers where the reported yield exceeded 300 bu/ac were excluded from the analysis. Including these observations in the model did not significantly alter the results.

3.2 Soil Moisture Data

Soil moisture data are currently available through a variety of in situ/surface measurements, climate and land-surface models, and satellite remote sensing at both point and spatially distributed scales. However, surface-based point measurements of soil moisture are often sparse and unevenly distributed, limiting their ability to provide estimates of soil moisture dynamics over larger spatial scales like the current study employs (Tavakol, et al. 2019). This study integrates relative soil moisture measurements derived from the Short-term Prediction Research and Transition-Land Information System (SPoRT-LIS) provided by NASA (NASA n.d.). Additionally, the SPoRT-LIS data employed in the current study has a longer time series than common alternative sources like Soil Moisture Active Passive (SMAP) (NASA n.d.).

Generally, satellite- and model-based soil moisture data offer broader geographical coverage, but model accuracy is highly dependent on meteorological forcing data and the parameterization schemes of land surface models (LSMs). Errors in surface forcing data—including precipitation, temperature, radiation, humidity, pressure, and wind—can significantly affect soil moisture simulations (Mitchell and et. al. 2004).

The North American Land Data Assimilation System (NLDAS) project was designed to integrate the best available observation- and model-based data, creating a quality-controlled dataset with consistent temporal and spatial resolution. While NLDAS provides a strong foundation for soil moisture modeling, its coarse resolution presents limitations, particularly for its application in agriculture and this study (Tavakol, et al. 2019).

To address these limitations, the NASA SPoRT Center runs the Noah LSM within the Land Information System (LIS) framework, producing model-based soil moisture estimates over the United States at a finer resolution than NLDAS (Tavakol, et al. 2019). Although SPoRT-LIS uses high-quality boundary conditions from NLDAS, it provides improved resolution (~1 km) and incorporates land-atmosphere states and fluxes with greater accuracy (Case and White 2014, Case, White and Guyer, et al. 2016).

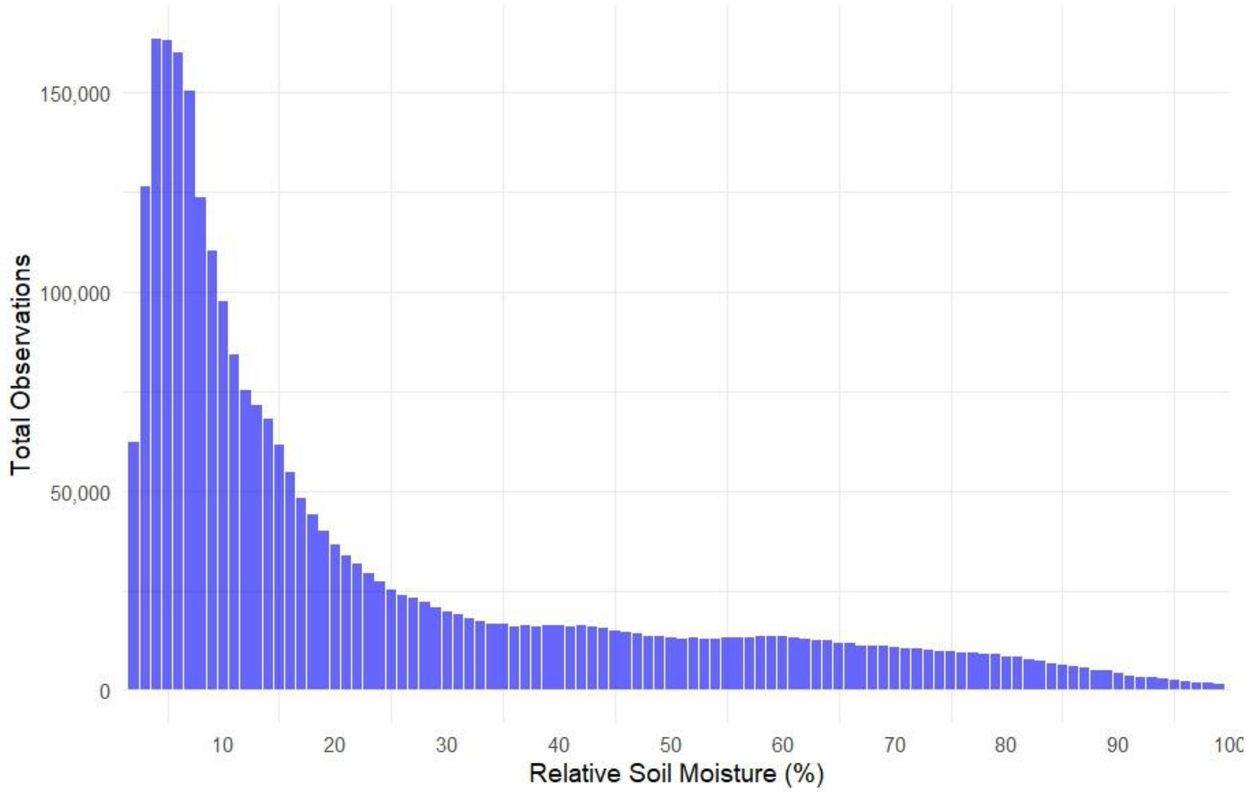
SPoRT-LIS outputs daily volumetric soil moisture for four soil layers (0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm) at a spatial resolution of approximately 1 km^2, which has been shown to be effective in water resource management applications (Kumar, et al. 2006, McDonough, et al. 2018, Peters-Lidard, et al. 2007).

In a 2018 study, surface soil moisture estimates from the SPoRT-LIS 0–10 cm layer were compared against in situ soil moisture measurements from the International Soil Moisture Network, focusing on the Missouri and Arkansas-Red-White River Basins (McDonough, et al. 2018). Validation was conducted at 5-cm and 10-cm depths, with performance evaluated across different soil types, land cover, depths, slopes, aspects, and pixel heterogeneity to assess the conditions under which SPoRT-LIS provides accurate estimates.

Results indicated that 53% of measurements at 5-cm depth and 51% at 10-cm depth were significantly correlated with in situ data, with Spearman's ρ values exceeding 0.5 on a daily basis. These validation results support the use of SPoRT-LIS surface soil moisture estimates for this research.

Figure 3.2 displays a distribution of soil moisture for each day used in the national model analysis.

Figure 3.2 Distribution of Soil Moisture Data



Chapter 4 - Methods

4.1 Soil Moisture Data Aggregation to County-Level

The soil moisture dataset used in this study includes daily values at roughly 1 km^2 resolution from 1981 to 2020. Because yield data is available at the county level, we needed to scale the soil moisture data up to match that same geographic unit. But rather than averaging across entire counties, we first limited our focus to the areas within each county that are consistently used for row crop production.

To do this, we used the Cropland Data Layer from 2002 to 2021 to identify which gridded cells are regularly planted corn, wheat, or soybeans (USDA NASS 2024). For each 1 km^2 pixel, we counted how many of those 20 years it was planted with one of those crops. If a pixel planted corn, wheat, or soybeans in more than half of the years, we kept it; otherwise, we dropped it. That gave us a binary mask showing which parts of each state were consistently in crop production. These masks were created in Google Earth Engine and exported for use in R.

Once we had those masks, we applied them to the soil moisture data. We cropped the soil moisture rasters to each state, converted the values from volumetric to relative soil moisture, and then multiplied them by the mask so that only cropland areas remained. That allowed us to zero in on the soil conditions that are most relevant for crop development. We used the 0–20 cm depth layer since it's most influential during germination and early growth (Archontoulis and Licht 2021).

To aggregate the filtered soil moisture data to the county level, we used the *terra* package in R to take the spatially weighted median value for each day. County boundaries don't line up perfectly with the soil moisture grid, so some grid cells fall partly in one county and partly in another. To deal with that, we applied weights based on how much of each cell overlaps with a county. For example, if half a grid cell is inside a county, it gets half the weight of a fully contained cell when calculating the median.

The end result is a daily, county-level dataset of relative soil moisture that reflects conditions specifically on land that has a long-term history of corn, wheat, or soybean production.

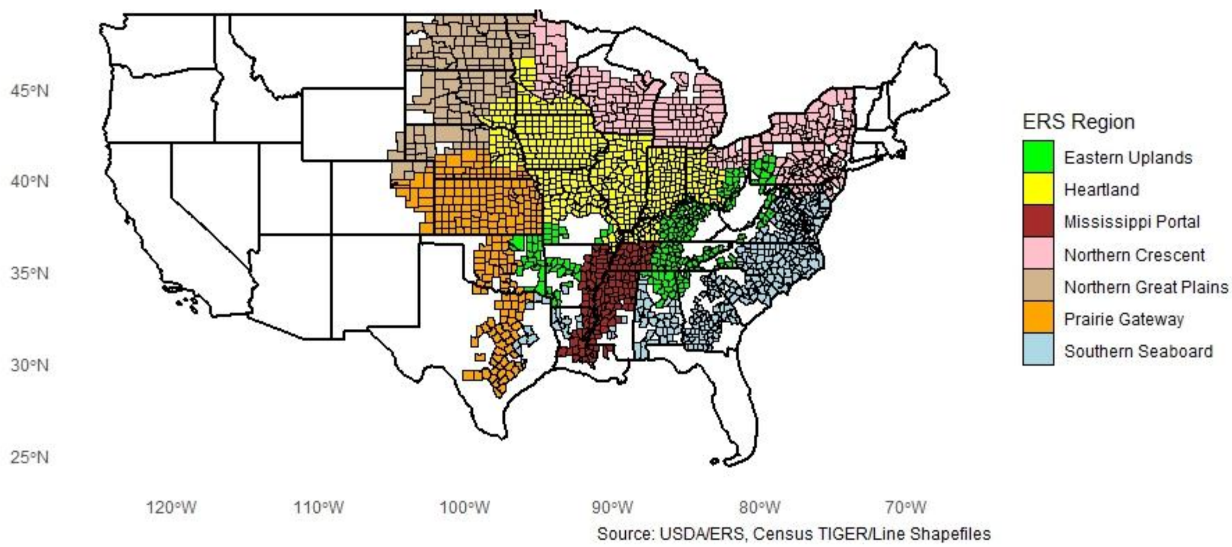
4.2 Spatial Heterogeneity Analysis

Excessive rainfall on crop yields is significantly influenced by regional conditions (Y. Li, et al. 2019). They find that extreme drought consistently led to significant yield reductions in most regions, with 94% of county-year samples showing negative yield anomalies. The 12 states, responsible for 88% of U.S. maize production over the past decade, were particularly affected. Yield losses due to extreme drought ranged from −45.7% in Illinois to −12.2% in Nebraska, with irrigated states such as Nebraska, Kansas, Oklahoma, and Texas generally experiencing less severe impacts than non-irrigated states like Illinois, Wisconsin, and Missouri.

The effects of excessive rainfall, however, were even more spatially variable than those of drought, with both positive and negative yield responses observed depending on the region. Negative yield anomalies were reported in 53% of county-year samples. When only counties with negative yield anomalies were considered, the average impact of excessive rainfall was much greater (−30%) compared to when all counties were included (−16.6%). This suggests that positive yield impacts in some counties offset the negative impacts in others, reducing the overall average effect at regional and national scales. Similar patterns of mixed positive and negative responses to excessive rainfall have been observed in other regions, such as China, underscoring the complexity of its impact on crop yields (Chen, et al. 2019).

To better understand these dynamics, our study divided the data into seven* subregions based on the USDA Economic Research Service (ERS) Farm Resource Regions. The Farm Resource Regions are derived from a combination of factors, including U.S. farm characteristics, historical Farm Production Regions, USDA Land Resource Regions, and NASS Crop Reporting Districts. By grouping counties based on dominant commodities, these regions reflect the localized climate, soil, water, and topographic conditions that shape agricultural productivity. **Figure 4.1** provides a visual representation of these regions.

Figure 4.1 Farm Resource Regions Used in Analysis (USDA ERS)

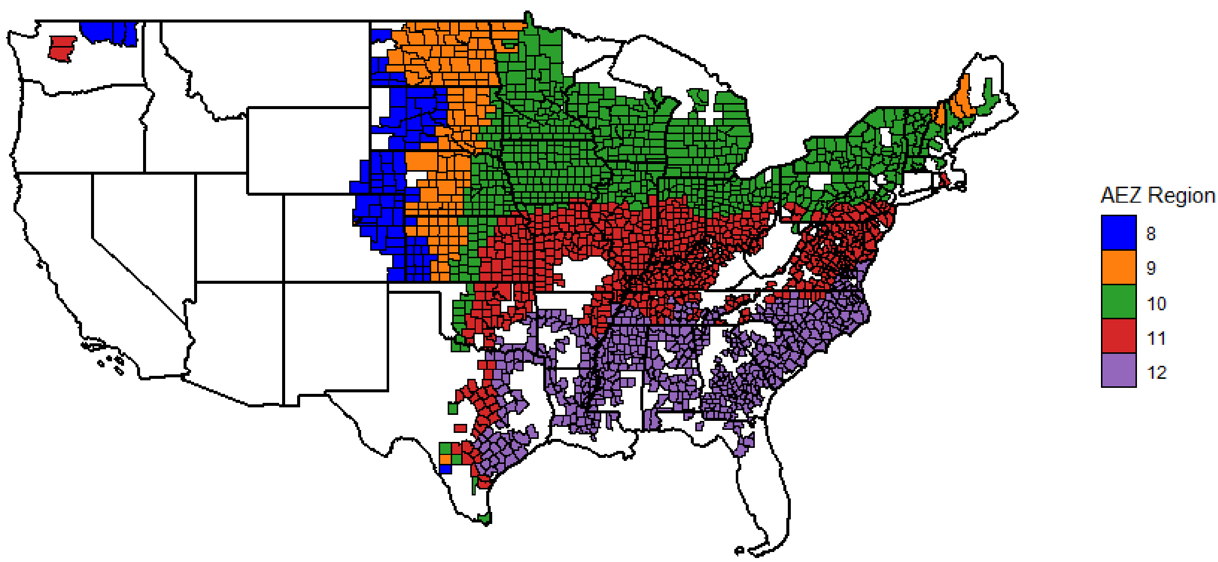


* The "Fruitful Rim" and "Basin and Range" farm resource regions were excluded from the analysis due to having limited observations of corn grown in those regions.

However, these regions still pose potential issues. For example, Eastern Colorado, which often receives less than 20 inches of precipitation annually, is grouped into the same zone as northeast Kansas where it is common for counties to receive twice that amount.

To test the robustness of ERS results, this study also splits observations into agroecological zones. These zones are land resource mapping units that are defined by climate, landform, soils, and/or land cover (FAO Land and Water Development Division 1996, Fischer, et al. 2021). With an emphasis on climactic and agronomic conditions, these zones offer further insight into regional dynamics influenced by external conditions. **Figure 4.2** is a map of the counties used in the analysis by agroecological zone ("AEZ").

Figure 4.2 Agroecological Zones Used in Analysis (Fischer, et al. 2021)

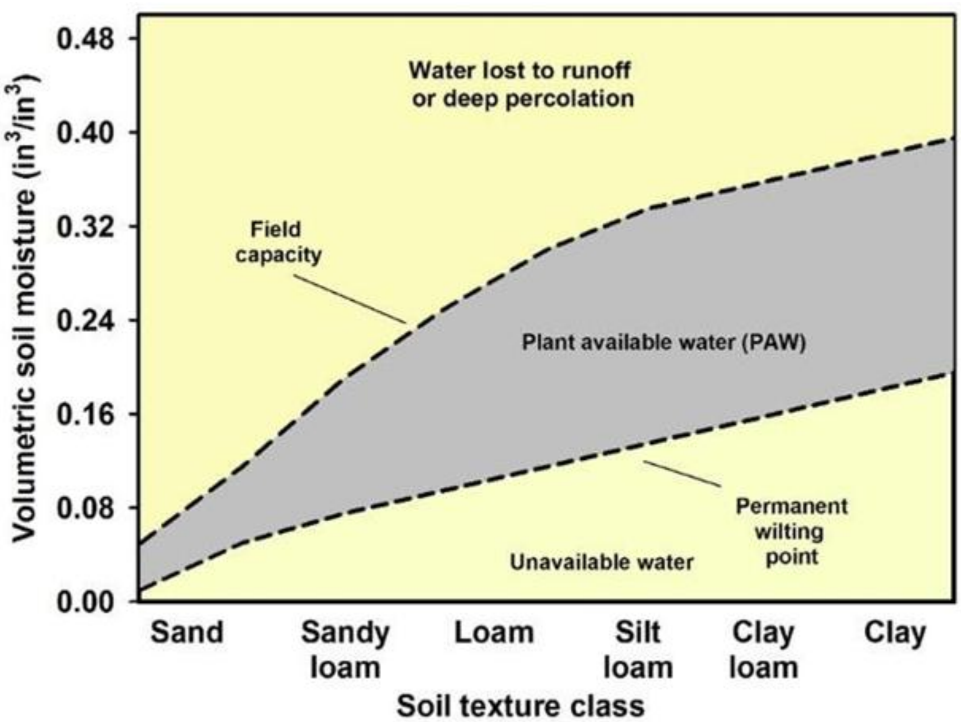


Finally, we analyzed the state of Iowa individually. Iowa is not broken up in either the AEZ regions or the ERS regions, so it provides a good baseline for comparison to see which region's results appear most representative of the smaller sample. Iowa has a long corn yield history providing over 6,000 county-year observations for analysis.

4.3 Conversion to Relative Soil Moisture Values

The influence of soil moisture on agricultural yields is significantly impacted by the hydrological properties of the soil. These properties vary according to soil types, ranging from sandy soils to clay (Archontoulis and Licht 2021). These properties include the wilting point (the moisture level below which plants can no longer extract water), field capacity (the moisture level representing near-optimal conditions for plant growth), and saturation (where soil pores are completely filled with water). **Figure 4.3** illustrates the variation in hydrological properties across different soil types.

Figure 4.3 Soil Hydrological Properties (Zotarelli, Dukes and Morgan 2019)



Deviations from optimal soil moisture levels can negatively affect root development, plant transpiration, and nitrogen mineralization in the soil (Kisekka, et al. 2017). Excessive moisture can also lead to nitrogen loss through denitrification and leaching, as well as increased evaporation rates (Urban, et al. 2015). Therefore, accurately assessing soil moisture, while accounting for variations in soil type, is critical to understanding its effect on yields.

To address the variation in soil properties across different counties, we use a relative measure of soil moisture, indexing moisture values based on key soil characteristics such as field capacity and wilting point. The relative soil moisture value used in the regression is calculated as follows:

$$Relative\ soil\ moisture = \frac{Volumetric\ soil\ moisture - wilting\ point}{Saturation - wilting\ point} \tag{3.1}$$

This formula effectively represents soil moisture as a ratio between the wilting point and saturation level for any given soil type. This allows for cross-geographic comparisons, as soil moisture is measured on a scale from 0 to 1, representing the percentage of saturation. A value of 0 indicates that soil moisture is at the wilting point (very dry), while a value of 1 indicates full saturation (very wet). A value of 0.54 suggests near-optimal moisture for plant growth, while values below 0.2 and above 0.8 indicate drought and excess moisture stress, respectively (Archontoulis and Licht 2021).

4.4 Soil Moisture Exposure

An alternative approach would be to use the value of soil moisture on a specified day or over a range of days as a continuous variable in the model. However, this approach faces several challenges. First, the model could be highly sensitive to the somewhat arbitrary choice of date, and it raises questions about which depth of soil moisture should be used as the volatile nature of the shallow 0-20cm depth made the model especially sensitive to the selected date. This method also requires imposing a functional form on the continuous soil moisture variable that is not clearly supported by existing literature.

There is also agronomic justification for this method. Consider two scenarios that lead to identical relative soil moisture readings on the chosen date, March 15th. In the first case, moisture could have gradually accumulated over multiple light precipitation events. In the second, a single substantial precipitation event could have occurred on the night of March 14th. While the soil moisture readings on March 15th that would be used in the first model are identical, the underlying agronomic dynamics are significantly different. The exposure approach helps capture this nuance.

To flexibly estimate the nonlinear relationship between soil moisture and yield, and to account for large precipitation events occurring near the chosen date, this study adopts a non-parametric approach. Instead of focusing on a single day's moisture value, the number of days during which soil moisture exposure falls within defined thresholds over a specific early season interval is calculated. This method draws on the work of previous literature that used a similar approach to analyze the nonlinear effects of temperature and other climatic drives of crop yields during the growing season (Schlenker and Roberts 2009, Ortiz-Bobea, Wang, et al. 2019).

In the current study, soil moisture exposure is measured by first setting an exposure time window. Various start and end date specifications are considered for each model and region. The testing and selection of the exposure window is described further in **4.5 Early Season Time Window Specification.** Then, relative soil moisture values are rounded to the nearest 0.01 and grouped into 100 levels, ranging from 0 to 100. A value of 0 indicates that the soil moisture is at or below its permanent wilting point, while a value of 100 represents full saturation.

To estimate the change associated with different soil moisture exposure levels, we began by combining the 100 distinct soil moisture bins (each representing 1% increments of relative soil moisture, or RSM) into broader groupings. These new groupings serve as regressors for the analysis. The aggregated bins are represented by $\phi_{k,it}$ which denotes the total number of days county *i* in year *t* was exposed to relative soil moisture levels within bin *k*. Mathematically, $\phi_{k,it}$ is defined as:

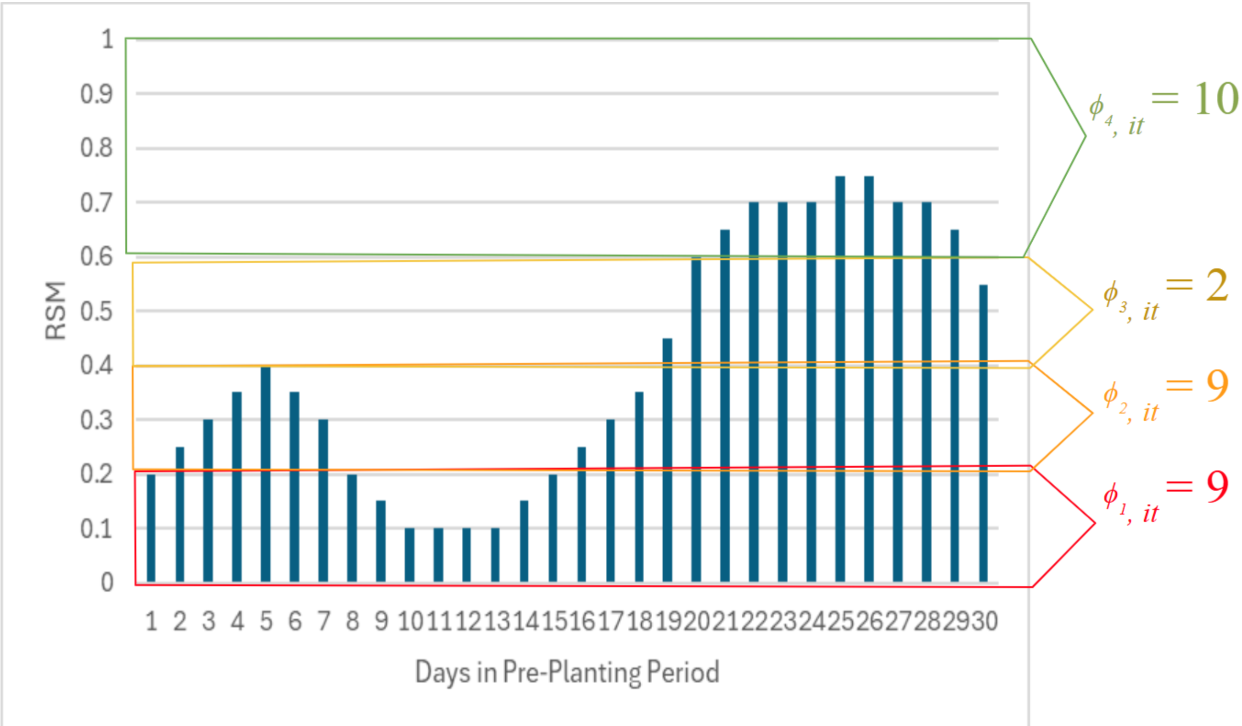
$$\phi_{k,it} = \sum_{d=1}^{T} 1(RSM_{it,d} \in \mathrm{Bin_k}) \tag{3.2}$$

where T represents the total days in the early season period. 1(·) is an indicator function that equals 1 if the daily relative soil moisture value ($RSM_{it,\,d}$) for county *i* on day *d* falls within the range defined by bin *k*, and 0 otherwise.

To illustrate, consider **Figure 4.4,** with an early season period of 31 days and a bin configuration comprising four ranges: 0%−20%, 21%−40%, 41%−60%, and 61%−100%. If a county experiences 9 days of 0%−20% relative soil moisture, 9 days of 21%−40%, 2 days of 41%-60% and 10 days of 61%−100% during the specified early season period, the corresponding bin counts are:

$$\phi_{0\%-20\%\,,it} = 9, \qquad \phi_{21\%-40\%\,,it} = 9, \qquad \phi_{41\%-60\%\,,it} = 2, \qquad \phi_{61\%-100\%\,,it} = 10$$

Figure 4.4 Bin Methodology Example



This binning approach avoids assumptions about the functional form of the relationship between RSM and crop yields, allowing the model to flexibly capture threshold effects and nonlinearities.

4.5 Early Season Time Window Specification

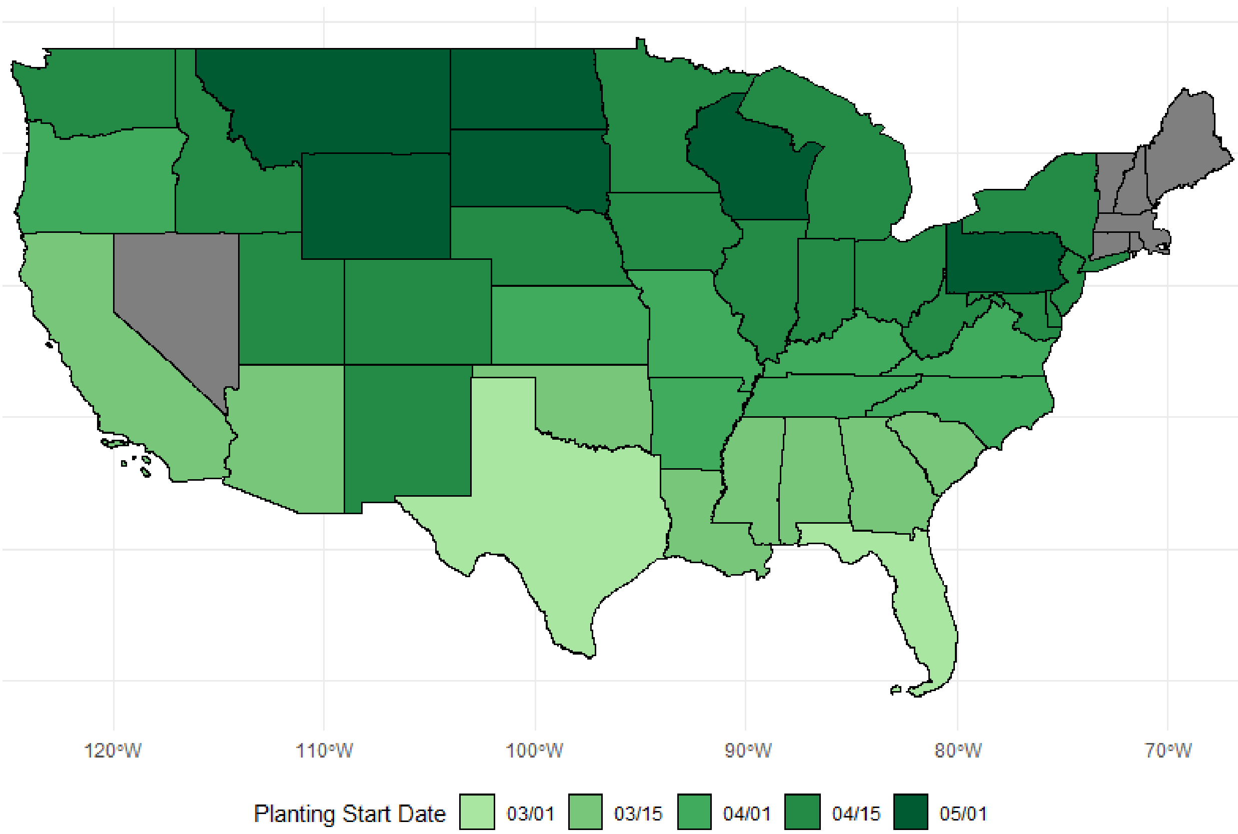
An important implicit assumption of the exposure approach is that soil moisture effects accumulate over time and that yield is proportional to total exposure (Ortiz-Bobea 2013). In other words, soil moisture exposure on any given day within the specified early season window has the same effect as any other day in the same soil moisture bin. For example, an observation on day 2 is assumed to have the same impact as an observation on day 19 within that bin.

This assumption is particularly important when defining the length of the early season window. We test this assumption and account for it by specifying a variety of different models all the way out to 90-day periods reported in the appendix. However, for the primary study, we fix the planting period to be 30 days across all national and regional models. This window is long enough that there are sufficient observations spread across exposure bins (30 to be exact), yet short enough that this assumption remains plausible.

With the length of the time window fixed to 30 days, we have to define a start and end date for the 30-day period. Since planting dates vary across the country, we rely on the most recent study (USDA NASS 2010), which used 20 years of historical crop progress estimates and expert input to determine planting windows by state. Specifically, we use the state-specific planting start date, defined as the date when 5% of planting is completed (USDA NASS 2010).

To standardize regional comparisons, we round each county's planting start date to the nearest beginning or midpoint of the month and define a fixed 30-day early season window. This means that for each county, the 30 days leading up to its adjusted planting start date define its early season exposure period. As a result, regions such as the Heartland may contain multiple counties with different time windows, but they are grouped together for analysis. **Figure 4.5** displays the planting date for each state used in the analysis.

Figure 4.5 Planting Start Date by State



4.6 Soil Moisture Bin Breakpoint Specification

The size and number of exposure bins is an important consideration for model specification as it is possible to increase the model fit by adding parameters, but doing so may result in overfitting (Schwarz 1978). The selection of the number of bins and their corresponding breakpoints used in the regressions was determined through an iterative, data-driven process.

For each bin possible configuration, we estimated the regression model and calculated the Bayesian Information Criterion (BIC) and adjusted R^2 (Schwarz 1978). These metrics were used to assess the trade-off between model complexity and explanatory power. We used the national model for testing and ultimately selected the specification that maximized the adjusted R^2. We adjusted this specification slightly to ensure the bins reflected the full nature of the nonlinear relationship. This adjustment split the lowest relative soil moisture bin into two as this allowed us to better capture the yield response at the lowest moisture levels. We held this specification constant for each of the subsequent regional models.

4.7 Regression Model Specification

The models are estimated using OLS regression. The regression specification can be expressed as:

$$\ln y_{it} = \sum_k \beta_k \phi_{k,it} + z_{it}\delta + c_i + \varepsilon_{it} \qquad (3.3)$$

where $\ln y_{it}$ represents the natural logarithm of yield for county *i* in year *t*. The variable $\phi_{k,it}$ represents the number of days county *i* was exposed to relative soil moisture within bin *k* summed across all *k* (bins). The coefficients (β_k) quantify how these exposures contribute to crop yield. β_k represents the change associated with one additional day of exposure to bin *k*. The term z_{it} includes additional control variables, such as a quadratic time trend (t and t^2) to account for technological change. County-specific fixed effects (c_i) are included to control for unobserved heterogeneity, such as soil quality and other factors that do not vary over time. Finally, we expect moisture conditions to be correlated in adjacent counties, so we allow the error terms (ε_{it}) to be spatially correlated using Conley standard errors (Conley 1999).

4.8 Economic Significance Testing

As an extension to the primary yield-based analysis, we re-estimate the model using county-year loss ratio data from the USDA Risk Management Agency (RMA). The loss ratio captures the extent to which actual revenue or yield falls below expected revenue or yield and thus triggers an indemnity payment. We use the aggregate loss ratio for all types of corn crop insurance policies, but revenue protection (RP) policies are the most common form of crop insurance. While yield provides a direct agronomic outcome, the loss ratio reflects economic losses and offers a complementary perspective on how early season soil moisture influences producer risk. The model specification remains unchanged, with the dependent variable replaced by the county-level loss ratio and the same fixed effects, time trends, and Conley standard errors applied. This analysis is conducted both at the national level and across four major corn-producing ERS Farm Resource Regions.

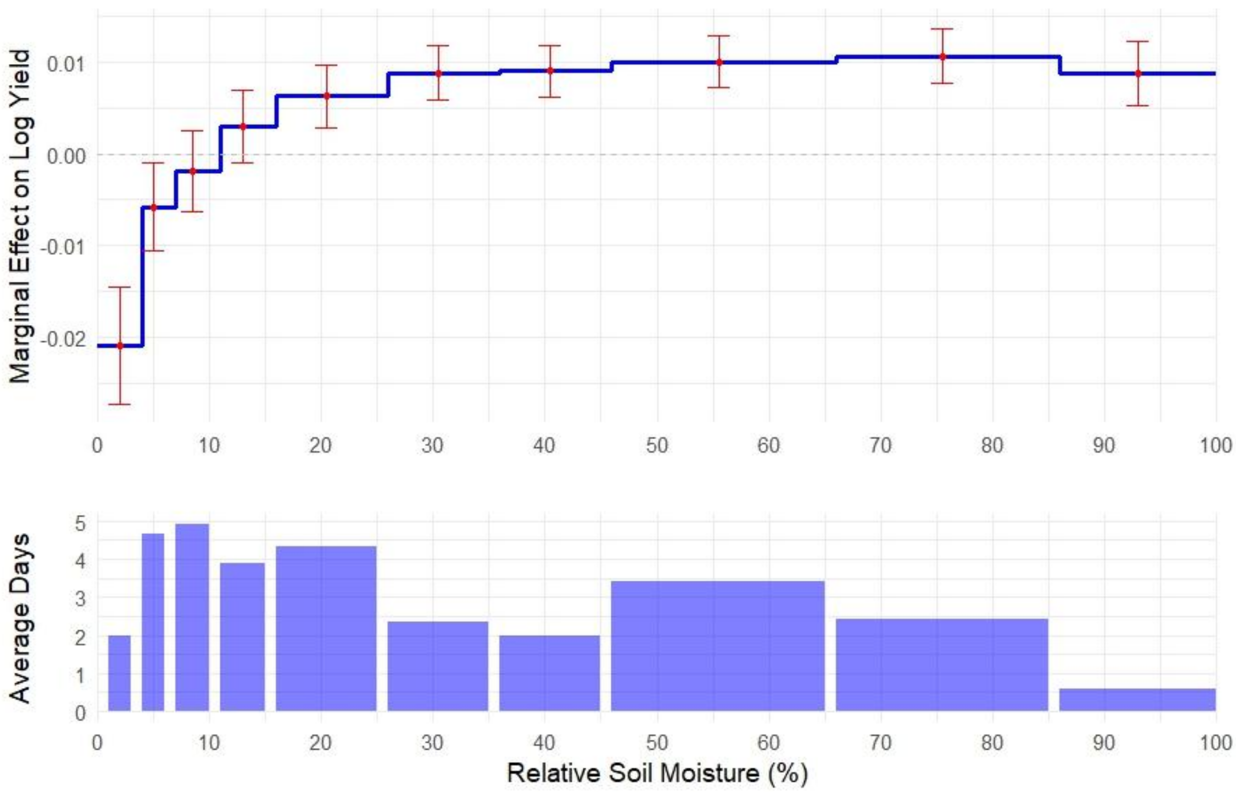
Chapter 5 - Results

5.1 National Model

First, we specified a nationwide model, summarized in **Figure 5.1**, which examines the relationship between relative soil moisture exposure during the 30-day, state-specific early season period and final crop yields. **Figure 5.1** includes error bars representing one standard deviation, while the lower panel shows the distribution of average days spent in each bin across the sample.

The coefficients from the regression model represent the estimated association between one additional day of exposure to a given soil moisture bin and the natural logarithm of dryland corn yield, holding all other variables constant. Importantly, the model includes the full 30-day early season exposure distribution across soil moisture bins, meaning the total number of days across all bins is fixed. As such, increasing exposure in one bin implicitly reduces exposure in another. This introduces a compositional constraint to interpretation—coefficient estimates reflect relative changes in exposure across bins, not absolute changes. Therefore, the estimated effect of a single bin should be understood in comparison to another or some combination of other bins.

Figure 5.1 National Model Coefficients and Distribution



To illustrate the interpretation, consider shifting one of the 30 early season days from the average bin (45–65% RSM, coefficient ≈ 0.01) to the driest bin (1–3% RSM, coefficient ≈ -0.02). This change is thus associated with a yield reduction of approximately 3%.

For reference, the April 2025 WASDE report projects 14,867 million bushels of corn production for the 2024/2025 marketing year, with an average yield of 179.3 bushels per harvested acre (USDA March 2025). Based on our model results, if a widespread drought were to cause this illustrated shift to the 1–3% RSM bin, the projected dryland corn yield would decline to approximately 173.9 bu/acre, leading to a back-of-the-envelope production loss of 446 million bushels relative to the current estimate.

5.2 ERS Farm Resource Region Models

Table 5.1 presents a comparative look between the national and regional model results.

Table 5.1 National and ERS Model Estimates

	National	Heartland	Northern Crescent	Northern Great Plains	Prairie Gateway	Eastern Uplands	Southern Seaboard	Mississippi Portal
1 - 3%	-0.0209 ***	0.0807 ***	0.0207 ***	-0.0597 ***	0.0588 *	0.0381	0.0128	0.0297 ***
	(0.0033)	(0.0036)	(0.0040)	(0.0079)	(0.0230)	(0.0319)	(0.0136)	(0.0083)
4 - 6%	-0.0058 *	0.0251 ***	0.0163 ***	-0.0394 ***	0.0832 ***	0.0389	0.0176	0.0301 ***
	(0.0025)	(0.0036)	(0.0030)	(0.0033)	(0.0231)	(0.0310)	(0.0137)	(0.0077)
7 - 10%	-0.0018	0.0188 ***	0.0129 ***	-0.0259 ***	0.0911 ***	0.0413	0.0188	0.0327 ***
	(0.0022)	(0.0034)	(0.0026)	(0.0034)	(0.0232)	(0.0312)	(0.0135)	(0.0080)
11 - 15%	0.0030	0.0128 ***	0.0117 ***	-0.0152 ***	0.0969 ***	0.0406	0.0224	0.0324 ***
	(0.0020)	(0.0028)	(0.0024)	(0.0034)	(0.0229)	(0.0307)	(0.0136)	(0.0073)
16 - 25%	0.0063 ***	0.0098 ***	0.0078 ***	-0.0051	0.1048 ***	0.0395	0.0215	0.0332 ***
	(0.0018)	(0.0025)	(0.0022)	(0.0031)	(0.0230)	(0.0303)	(0.0137)	(0.0074)
26 - 35%	0.0088 ***	0.0089 ***	0.0055 **	0.0003	0.1099 ***	0.0305	0.0236 .	0.0380 ***
	(0.0015)	(0.0018)	(0.0017)	(0.0027)	(0.0229)	(0.0303)	(0.0137)	(0.0077)
36 - 45%	0.0090 ***	0.0093 ***	0.0047 **	-0.0023	0.1112 ***	0.0389	0.0248 .	0.0395 ***
	(0.0014)	(0.0014)	(0.0016)	(0.0026)	(0.0230)	(0.0300)	(0.0137)	(0.0078)
46 - 65%	0.0100 ***	0.0095 ***	0.0032 **	0.0020	0.1135 ***	0.0327	0.0281 *	0.0415 ***
	(0.0014)	(0.0013)	(0.0012)	(0.0032)	(0.0234)	(0.0301)	(0.0137)	(0.0079)
66 - 85%	0.0106 ***	0.0095 ***	0.0024 .	-0.0020	0.1254 ***	-0.0096	0.1027 ***	0.0428 ***
	(0.0015)	(0.0015)	(0.0014)	(0.0027)	(0.0235)	(0.0316)	(0.0139)	(0.0082)
86 - 100%	0.0087 ***	0.0069 ***	-0.0015	0.0023	0.1247 ***			0.0496 ***
	(0.0018)	(0.0016)	(0.0025)	(0.0051)	(0.0253)			(0.0089)
Time	0.0013	0.0110 ***	0.0091 ***	0.0267 *	-0.0523 ***	0.0088 *	-0.0012	0.0061 *
	(0.0018)	(0.0017)	(0.0019)	(0.0105)	(0.0086)	(0.0039)	(0.0030)	(0.0028)
Time^2	0.0005 ***	0.0001 *	0.0002 **	0.0003	0.0020 ***	0.0002	0.0005 ***	0.0003 ***
	(0.0001)	(0.0000)	(0.0001)	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
N	59846	16554	9081	4073	7324	7116	9827	4267
Adj. R-squared	0.5092	0.4300	0.4244	0.4647	0.2539	0.3189	0.3286	0.4228
RMSE	0.3798	0.2247	0.2465	0.5502	0.6438	0.3105	0.3593	0.2093

In high-productivity regions like **Figure 5.2 The Heartland** and **Figure 5.3 Northern Crescent**, lower early season soil moisture levels are associated with higher yields. This aligns with the relatively stable growing conditions in these areas, where excessive early season moisture—rather than drought—is often the primary concern. High soil moisture can delay planting, leach nutrients, and cause waterlogging, all of which are detrimental to yield (Bille and Rogna 2021, Ziyi Li 2022)

In the Heartland model, exposure to the lowest soil moisture bin (1–3% RSM) is associated with a coefficient of 0.0807, statistically significant at the 99% confidence level. Relative to the average day in this region's distribution (coefficient = 0.0095), an additional day of exposure to extremely dry conditions corresponds to a net increase in dryland corn yield of over 7%. This stands in contrast to the national model, which finds a 3% yield reduction associated with the same level of dryness.

Figure 5.2 The Heartland

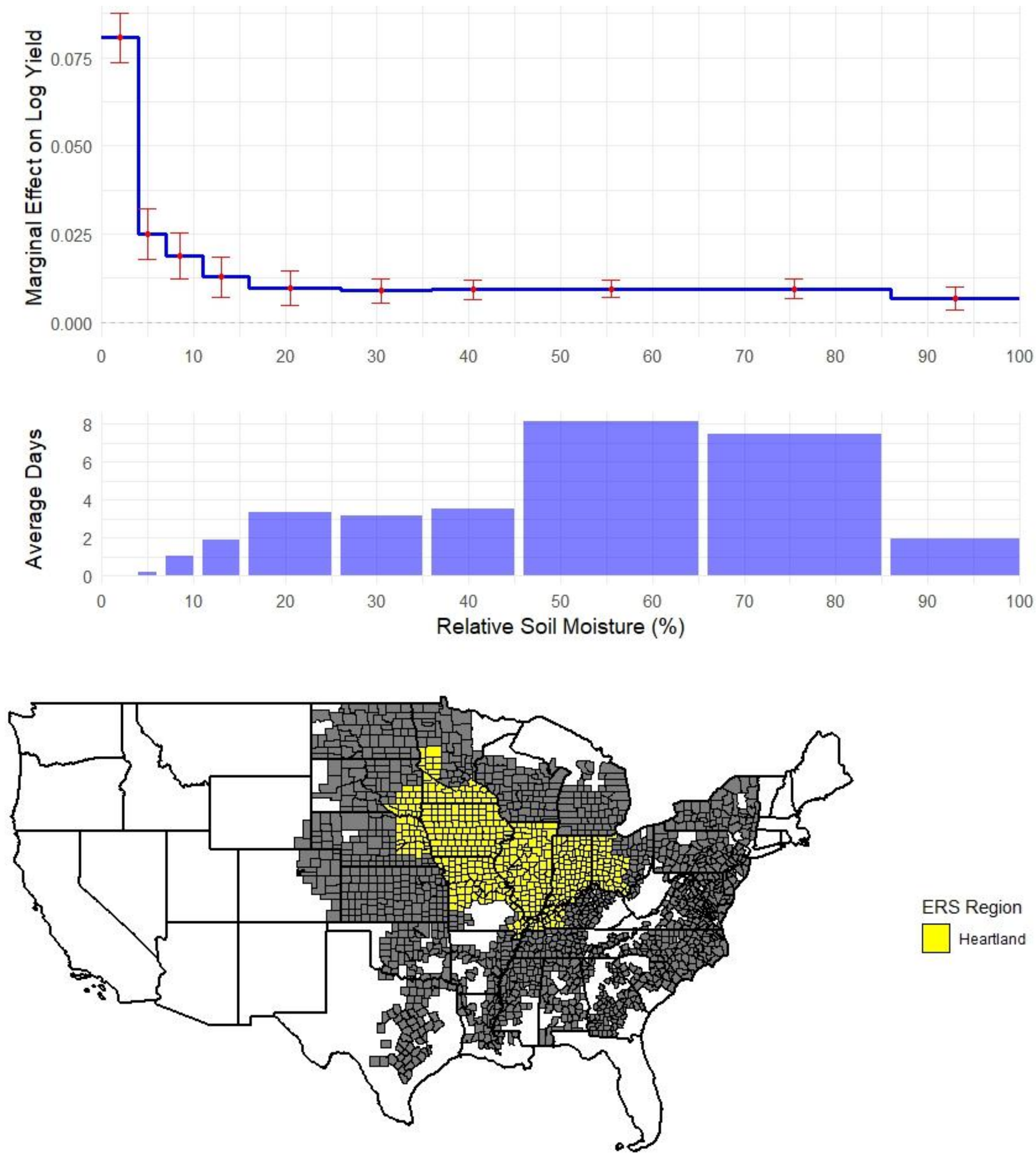
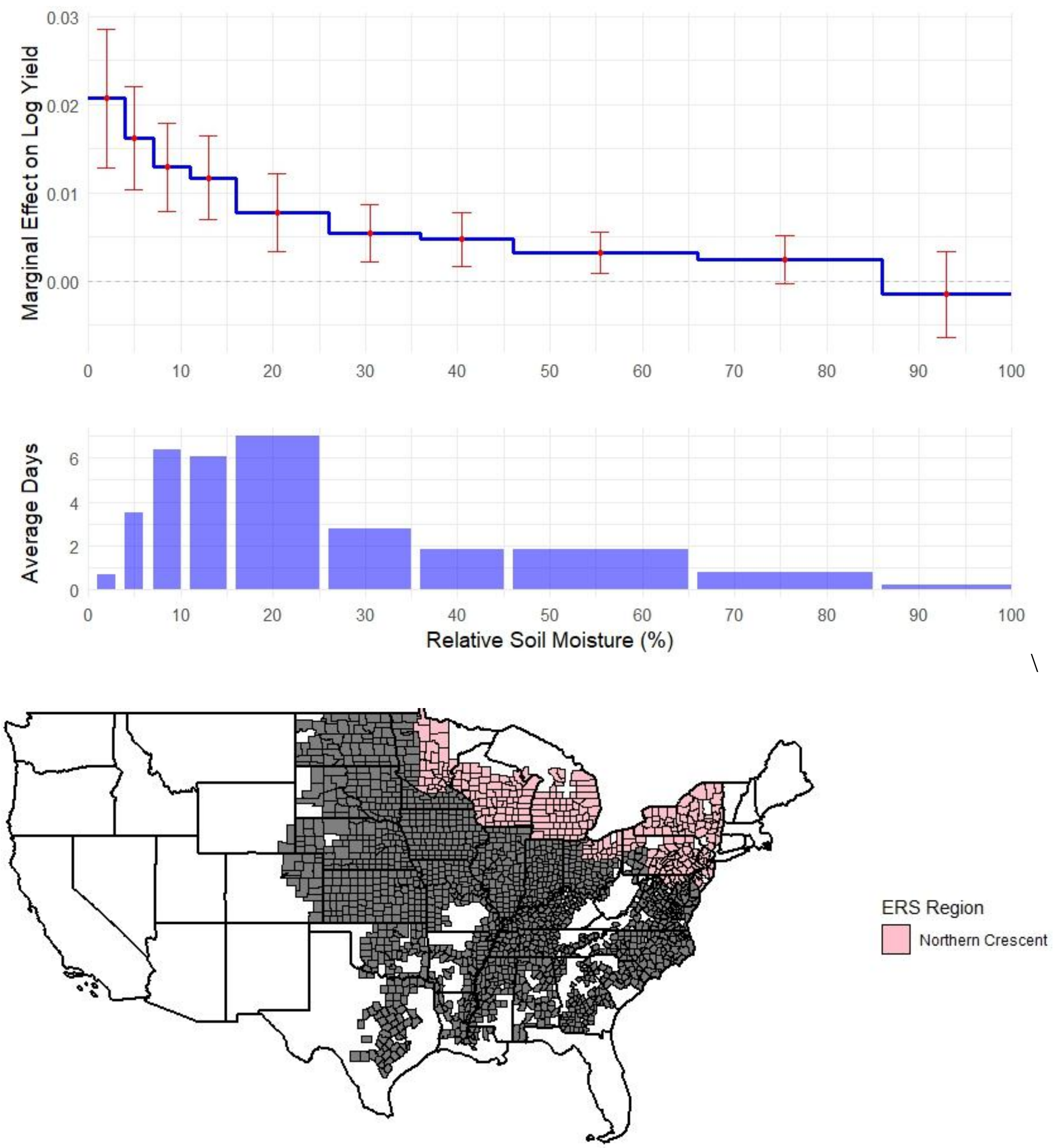


Figure 5.3 Northern Crescent



Several other regions exhibit the opposite relationship. For example, in the **Figure 5.4 Northern Great Plains** and the **Figure 5.5 Prairie Gateway**, low early season soil moisture is strongly negatively associated with yield—consistent with agronomic and climatic principles in those regions. In these areas with deep, high-capacity soils, shallow surface moisture may be less important early on, since crops can draw from deeper reserves (Evans and Fausey 1999, Liu, Zhang and Zuo 2017). However, once those reserves are depleted, crops become entirely dependent on rainfall, and early season dryness can be a critical warning sign of insufficient subsurface moisture.

In the Northern Great Plains model, an additional day of exposure to the lowest moisture bin, relative to the average day (26–35% RSM bin, coefficient = 0.0003), is associated with a yield reduction of approximately 6% (–0.0597 + 0.0003). This further illustrates how the national model's results are attenuated toward zero, likely due to unmodeled spatial heterogeneity that biases the estimates toward the average, masking important region-specific relationships.

Figure 5.4 Northern Great Plains

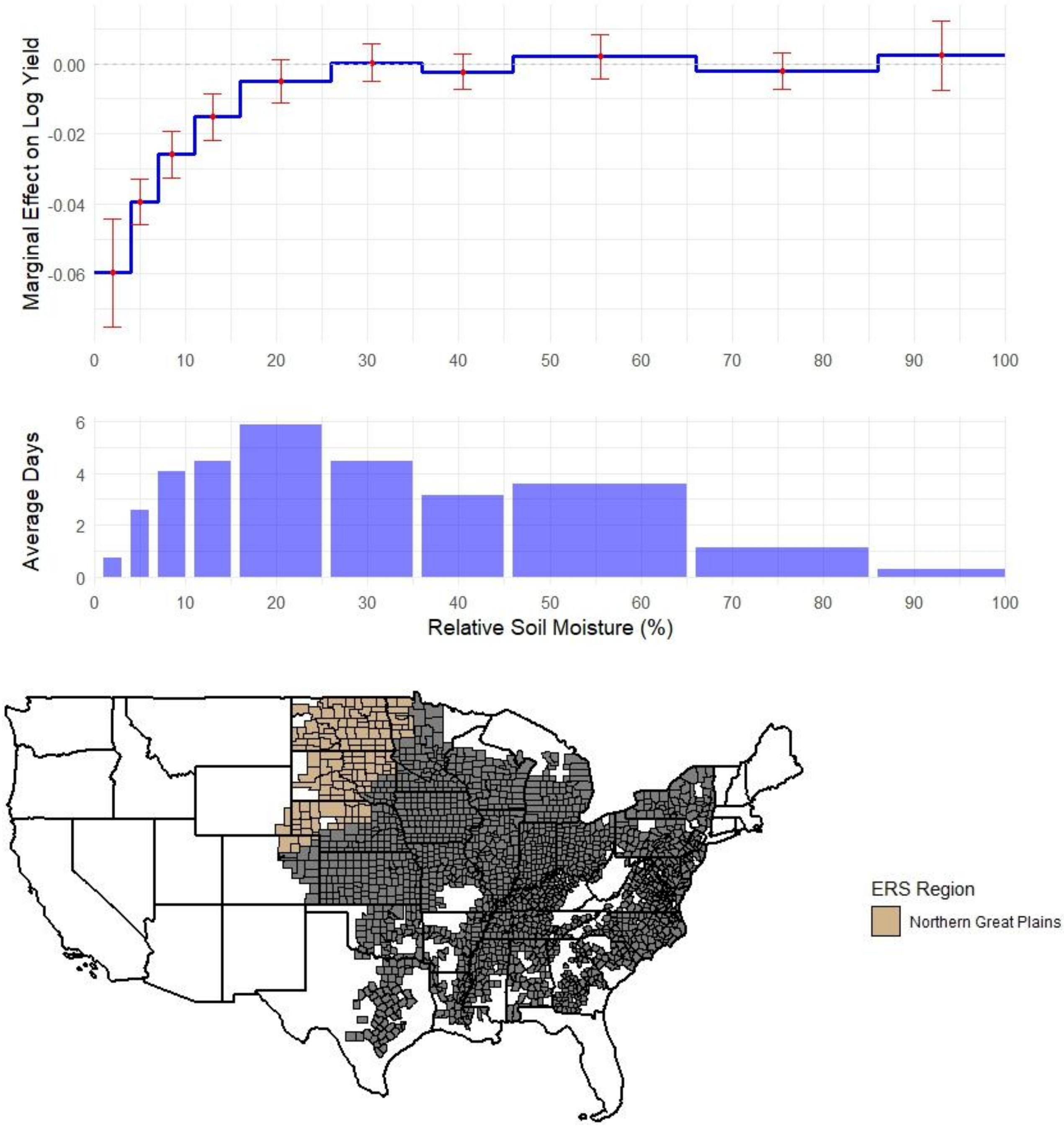
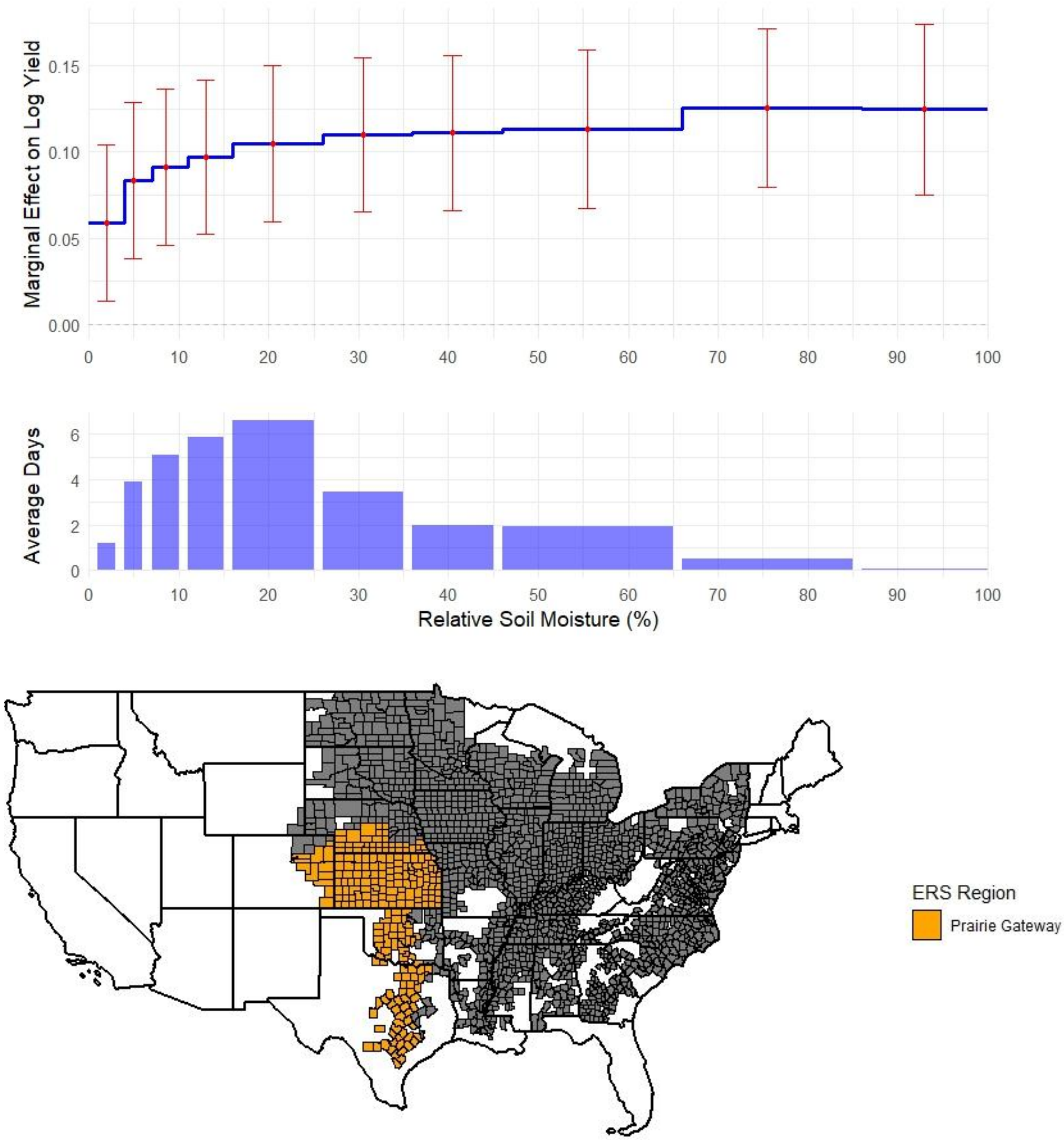
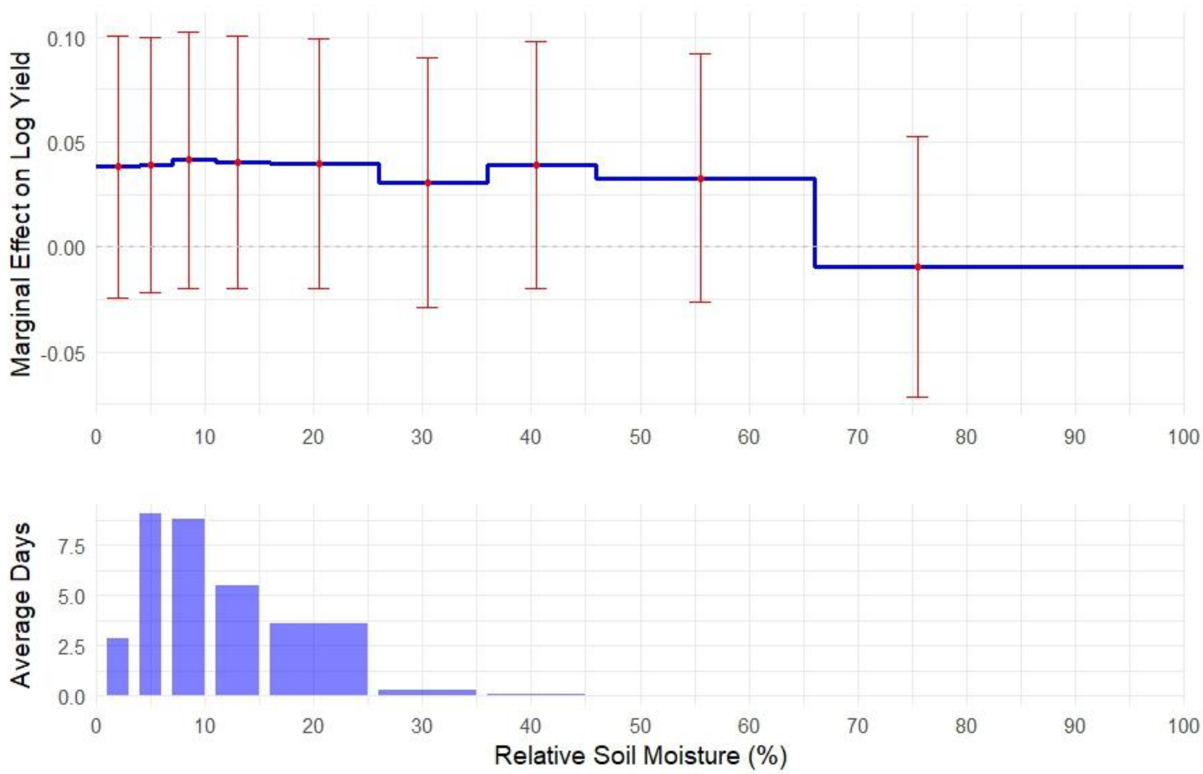


Figure 5.5 Prairie Gateway



Another interesting finding in the regional analysis is the apparent lack of correlation between early season conditions and dryland corn yields in some regions. Results from **Figure 5.6 Eastern Uplands**, **Figure 5.7 Southern Seaboard**, and **Figure 5.8 Mississippi Portal** exhibit substantially larger standard errors as well as lower model fit indicators. These areas also have relatively low levels of corn production, resulting in fewer observations available for analysis. This is an important factor that may contribute to the reduced statistical precision. Additionally, it is possible that meaningful heterogeneity within these broad regions remains unaccounted for, suggesting that further spatial disaggregation may be necessary to detect localized effects. These results underscore the challenge of balancing geographic specificity with the statistical power needed to produce reliable estimates.

Figure 5.6 Eastern Uplands



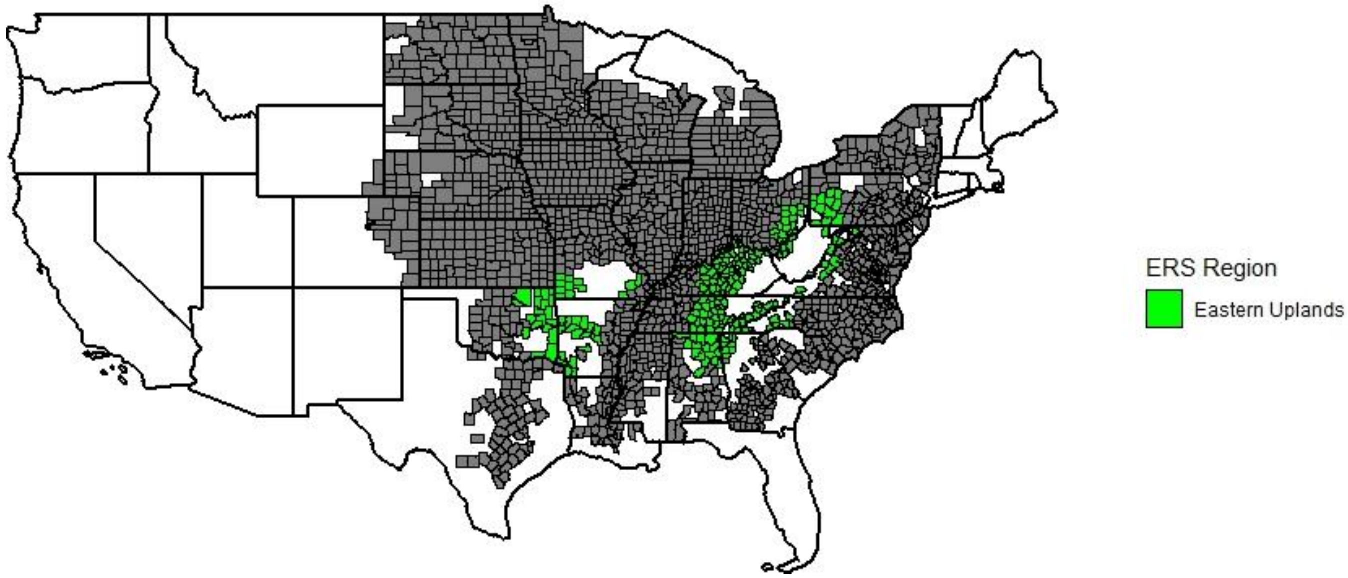
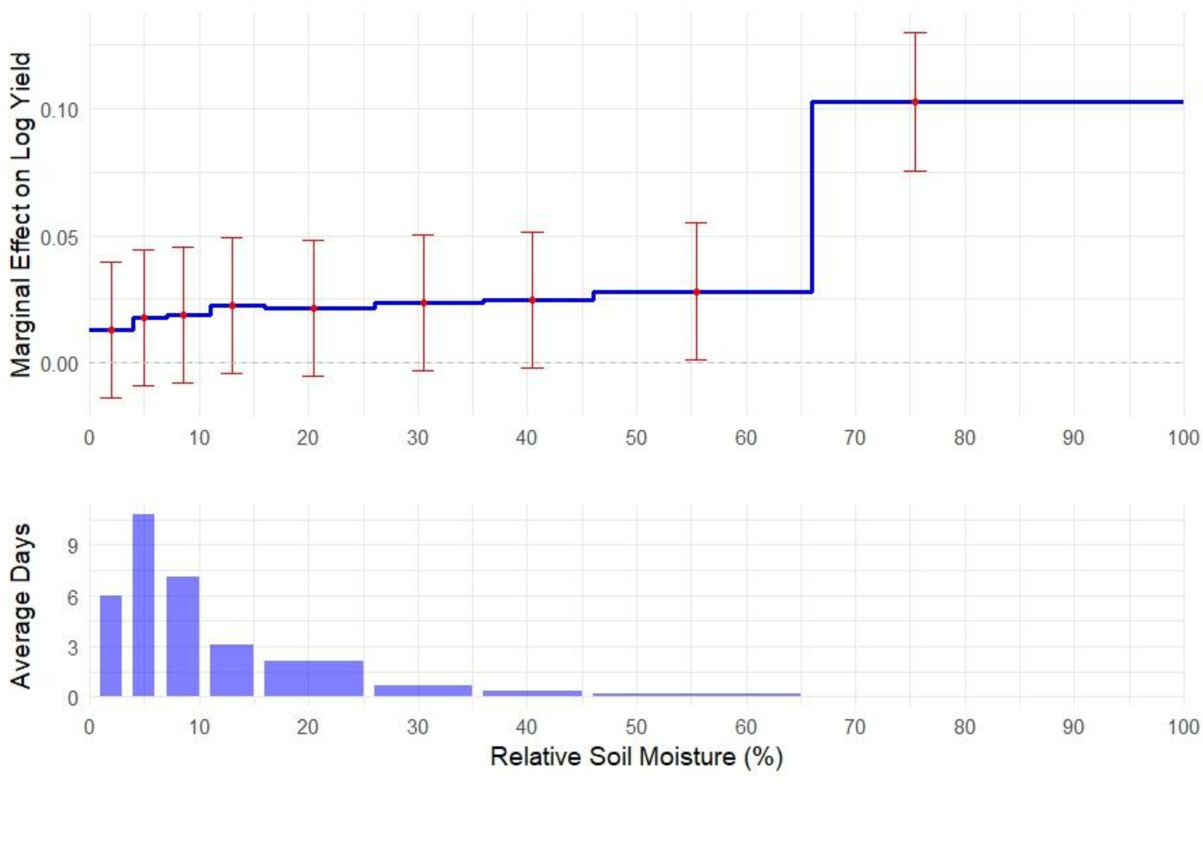


Figure 5.7 Southern Seaboard



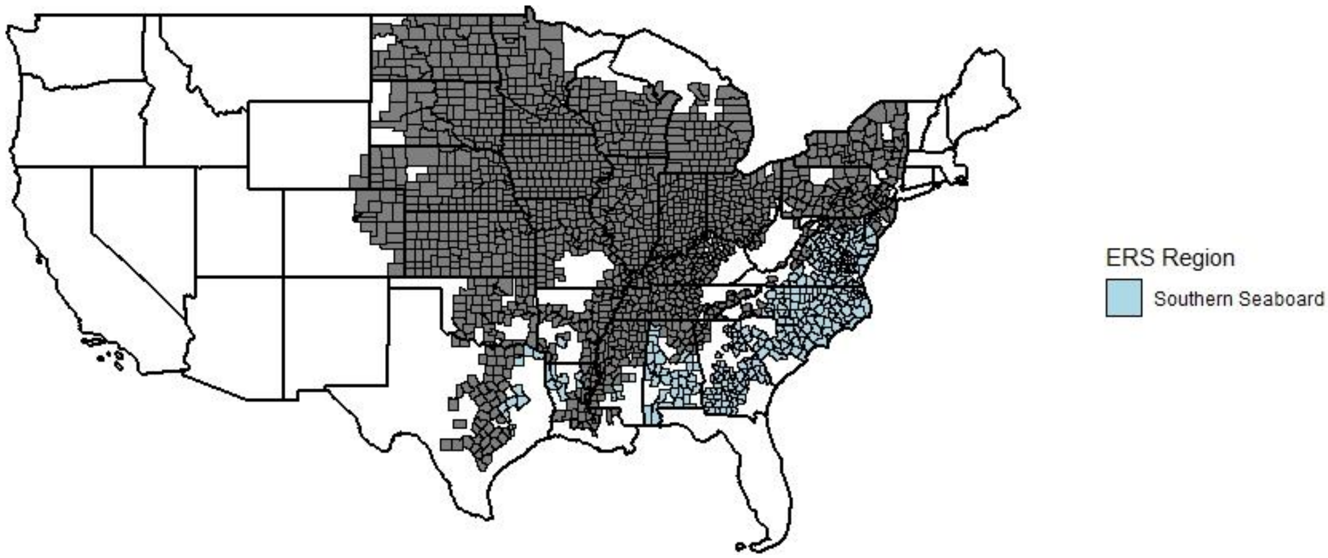
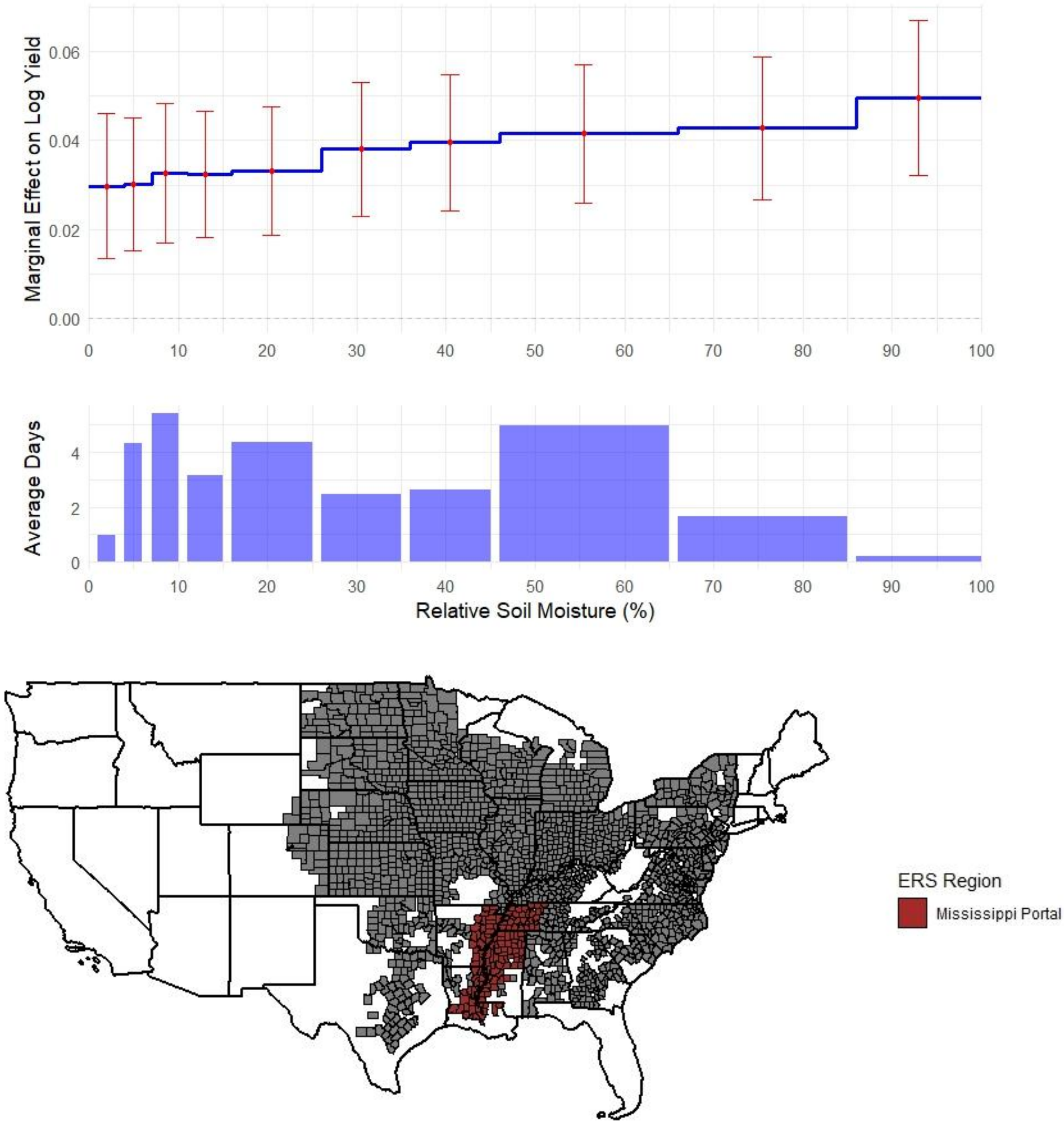

ERS Region
Southern Seaboard

Figure 5.8 Mississippi Portal



5.3 Economic Significance

As an extension of the primary yield analysis, we re-estimate the national model using loss ratios as the dependent variable to better capture potential economic losses. Loss ratios reflect the degree to which actual revenue or yield falls short of the insured revenue or yield guarantee, thereby triggering indemnity payments. The consistent statistical significance across exposure bins suggests that early season soil moisture is a meaningful predictor of insurance outcomes and risk exposure.

To explore spatial variation in how early season soil moisture affects insurance losses, we re-estimate the model across four major corn-producing ERS regions: the Heartland, Northern Crescent, Northern Great Plains, and Prairie Gateway. In the Heartland and Northern Crescent, drier early season conditions are associated with lower loss ratios, consistent with the yield models and agronomic concerns about excess moisture in high-productivity regions.

In contrast, the Northern Great Plains and Prairie Gateway show the opposite pattern—dry conditions are associated with higher loss ratios, underscoring the financial risks of insufficient early season moisture in drier climates. Notably, the Prairie Gateway model shows the steepest gradient, with coefficients becoming increasingly negative across the soil moisture distribution.

We conducted Wald tests for equality of coefficients across soil moisture bins within each loss ratio model. The null hypothesis, that all soil moisture bin coefficients are equal, was strongly rejected in every region, indicating that the impact of early season moisture on loss ratios varies significantly across the distribution. This provides statistical justification for interpreting differences in coefficients across bins and confirms that not all levels of early season exposure carry the same financial risk.

These regional findings reinforce the spatial heterogeneity observed in the yield models and demonstrate that early season soil moisture is not only agronomically important but also economically consequential for producers’ financial outcomes.

Table 5.2 RMA Loss Ratio Model Results

	(1) National	(2) Heartland	(3) Northern Crescent	(4) Northern Great Plains	(5) Prairie Gateway
1 - 3%	-0.0500**		-0.0832***	0.0426***	-0.1565**
	(0.0167)		(0.0224)	(0.0126)	(0.0592)
4 - 6%	-0.0700***	-0.1747***	-0.0760***	0.0458***	-0.1988***
	(0.0162)	(0.0246)	(0.0197)	(0.0098)	(0.0595)
7 - 10%	-0.0764***	-0.1428***	-0.0651***	0.0312***	-0.2150***
	(0.0160)	(0.0223)	(0.0190)	(0.0086)	(0.0593)
11 - 15%	-0.0816***	-0.1212***	-0.0586**	0.0268**	-0.2242***
	(0.0159)	(0.0215)	(0.0187)	(0.0092)	(0.0593)
16 - 25%	-0.0785***	-0.1072***	-0.0382*	0.0111	-0.2324***
	(0.0157)	(0.0205)	(0.0184)	(0.0083)	(0.0593)
26 - 35%	-0.0794***	-0.0988***	-0.0276.	0.0142*	-0.2422***
	(0.0156)	(0.0192)	(0.0167)	(0.0068)	(0.0597)
36 - 45%	-0.0783***	-0.0991***	-0.0175	0.0185*	-0.2474***
	(0.0154)	(0.0181)	(0.0162)	(0.0078)	(0.0604)
46 - 65%	-0.0774***	-0.0980***	-0.0121	0.0142.	-0.2481***
	(0.0154)	(0.0176)	(0.0144)	(0.0077)	(0.0587)
66 - 85%	-0.0830***	-0.1038***	-0.0100	0.0360***	-0.2612***
	(0.0167)	(0.0190)	(0.0182)	(0.0098)	(0.0610)
86 - 100%	-0.0716***	-0.0902***	0.0265*	0.0064	-0.2960***
	(0.0186)	(0.0220)	(0.0130)	(0.0144)	(0.0710)
Time	-0.0679***	-0.0168.	-0.0913***	-0.1551***	-0.0163
	(0.0064)	(0.0091)	(0.0115)	(0.0161)	(0.0151)
Time^2	0.0014***	0.0005.	0.0019***	0.0033***	-0.0003
	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0004)
Observations	55344	16827	8390	4635	7536
Adj. ^2	0.1061	0.0219	0.0835	0.1827	0.1329
RMSE	1.4667	1.3769	1.3581	1.277	1.6565

Chapter 6 - Discussion

This study provides critical insights into the relationship between early season soil moisture exposure and dryland corn yield outcomes, highlighting both the significance of extreme moisture conditions and the necessity of regional specificity in agricultural modeling. While the national model offers a broad understanding of how soil moisture extremes influence yields, its breakdown at the regional level underscores the need for localized approaches. Differences in soil types, cropping systems, and climatic conditions mean that models optimized for one region may fail to capture key dynamics in another.

When the analysis shifts from the national scale to examining regional variations, the estimates, model fit, and optimal specifications diverge significantly. This breakdown highlights the importance of spatial heterogeneity in soil moisture-yield relationships, as factors such as climate and agricultural practices vary widely across regions. The ERS Farm Resource Region (FRR) models provide valuable insight that aligns with agronomic intuition.

These findings offer a valuable contribution by confirming that early season soil moisture conditions are indeed related to yield outcomes. Our results provide empirical evidence that this information carries predictive power. However, the degree to which early season conditions influence final yields varies by region and exposure level, suggesting that current decisions may not always be fully informed or optimal. This opens the door for future research not only to refine estimates of these effects but also to explore how soil moisture data can be most effectively harnessed in practice. Our approach offers a useful proof of concept: practitioners could track moisture exposure in a defined window, apply estimated coefficients, and generate expected yield deviations relative to a baseline. This type of information could improve the targeting of crop insurance products, guide drought response efforts, and support more strategic investments in irrigation and soil management.

Despite its utility, this study raises important questions for future research. One key consideration is the appropriate level of spatial and temporal granularity. The stark regional differences observed suggest that finer disaggregation—by county, soil type, or cropping system—could yield more precise insights. However, greater granularity risks overfitting, particularly in regions with limited historical data. It also limits the number of observations thus decreasing the explanatory power of the model in some regions. Balancing model complexity and practical applicability remains essential for developing robust and scalable yield prediction models.

Another crucial consideration is rigorous out-of-sample testing. While the models effectively capture historical patterns, their predictive validity must be assessed against out-of-sample data to ensure continued relevance under evolving climatic and agricultural conditions. This is particularly critical in the context of climate change, which is likely to alter the frequency and intensity of extreme soil moisture events.

Finally, while the exposure approach effectively captures nonlinear soil moisture effects, future research could explore more flexible methods, such as machine learning, to further refine these relationships. However, any advancements in modeling must remain interpretable and accessible to practitioners to ensure widespread adoption in real-world agricultural decision-making.

In conclusion, this research underscores the dual importance of extreme soil moisture levels and regional specificity in yield prediction. By addressing these considerations and refining the methodological framework, future studies can enhance these models to better support agricultural stakeholders. In an increasingly uncertain agricultural landscape, such tools will be vital for adapting to new challenges and optimizing outcomes.

Chapter 7 - References

Archontoulis, S., and M. Licht. 2021. "FACTS soil moisture benchmarking tool." *Integrated Crop Management.* Iowa State University Extension and Outreach. https://crops.extension.iastate.edu/blog/mark-licht-mike-castellano-sotirios-archontoulis/facts-soil-moisture-benchmarking-tool.

Bailey-Serres, Julia, Jane E Parker, Elizabeth A Ainsworth, Giles E.D Oldroyd, and Julian I Schroeder. 2019. "Genetic strategies for improving crop yields." *Nature.* doi:10.1038/s41586-019-1679-0.

Bille, Anna G, and Marco Rogna. 2021. "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.* doi:10.1111/rssa.12709.

Cakir, Recep. 2004. "Effect of water stress at different development stages on vegetative and reproductive growth of corn." *Field Crops Research.* doi:10.1016/j.fcr.2004.01.005.

Case, Jonathan L., and Kristopher D. White. 2014. "Assessment of the 3-km SPoRT-Land Information System for Drought Monitoring and Hydrological Forecasting." *NASA Marshall Space Flight Center.*

Case, Jonathan L., Kristopher D. White, Brian Guyer, Jim Meyer, Jayanthi Srkishen, Clay Blankenship, and Bradley T. Zavodsky. 2016. "Real-time Land Information System over the continental US for situational awareness and local numerical weather predication applications." *American Meteorological Society (AMS) Annual Meeting.*

Chen, Dan, Ying Guo, Rui Wang, Yunmeng Zhao, Kaiwei Li, Jiquan Zhang, Xingpeng Liu, Zhijun Tong, and Chunli Zhao. 2023. "Quantifying multi-hazards and impacts over different growth periods of maize: A study based on index construction." *International Journal of Disaster Risk Science.* doi:10.1007/s13753-023-00516-8.

Chen, Huili, Quihua Liang, Zhongyao Liang, Yong Liu, and Shuguang Xie. 2019. "Remote-sensing disturbance detection index to identify spatio-temporal varying flood impact on crop production." *Agricultural and Forest Meteorology.* doi:https://doi.org/10.1016/j.agrformet.2019.02.002.

Conley, TG. 1999. "GMM estimation with cross sectional dependence." *J. Econometrics.*

Dayanto, S., L. Wang, and P.A. Jacinthe. 2016. "Global synthesis of drought effects on maize and wheat production." *PLoS ONE.* https://doi.org/10.1371/journal.pone.0156362.

Denmead, O.T, and R.H Shaw. 1960. "The effects of soil moisture stress at different stages of growth on the development and yield of corn." *Agronomy Journal.* doi:10.2134/agronj1960.00021962005200050010x.

Evans, R.O., and N.R. Fausey. 1999. "Effects of inadequate drainage on crop growth and yield." *Agricultural Drainage.*

FAO Land and Water Development Division. 1996. *Agro-ecological zoning guidlines.* Rome: FAO Soils Bulletin 73, Soil Resources, Management and Conservation Service.

Farahani, H.J., G.A. Peterson, and D.G. Westfall. 1998. "Dryland cropping intensification: A fundamental solution to efficient use of precipitation." *Adv. Agron.* https://doi.org/10.1016/S0065-2113(08)60505-2.

Fischer, G., F.O. Nachtergaele, H.T. van Velthuizen, F. Chiozza, G. Franceschini, M. Henry, D. Muchoney, and S. Tramberend. 2021. *Global Agro-Ecological Zones v4 – Model documentation.* Rome: FAO. doi:https://doi.org/10.4060/cb4744en.

Gilman, C.S., and J.T. Riedel. 1951. "Persistence of Extremely Wet and Extremely Dry Months in the United States." *Monthly Weather Review.* doi:https://doi.org/10.1175/1520-0493(1951)079%3C0045:POEWAE%3E2.0.CO;2.

Jagnani, Maulik, Christopher B Barrett, Yanyan Liu, and Liangzhi You. 2020. "Within-season producer response to warmer temperatures: Defensive investments by Kenyan farmers." *The Economic Journal.* doi:10.1093/ej/ueaa063.

Kanwar, R.S., and S. Mukhtar. 1988. "Excessive soil water effects at various stages of development on the growth and yield of corn." *Transactions of the ASAE.* https://doi.org/10.13031/2013.30678.

Kisekka, I, A Schlegel, L Ma, P.H Gowda, and P.V.V Prasad. 2017. "Optimizing preplant irrigation for maize under limited water in the High Plains." *Agricultural Water Management.* doi:10.1016/j.agwat.2017.03.023.

Korres, W., T.G. Reichenau, and K. Schneider. 2013. "Patterns and scaling prperties of surface soil moisture in an agricultural landscape: An ecohydrological modeling study." *Journal of Hydrology.* doi:https://doi.org/10.1016/j.jhydrol.2013.05.050.

Kumar, S.V., DC Peters-Lidard, Y Tian, PR Houser, J Geiger, S Olden, L Lighty, JL Eastman, B Doty, and P. Adams, J. Dirmeyer. 2006. "Land information system: An interoperable framework for high resolution land surface modeling." *Environmental modeling and software.* doi:https://doi.org/10.1016/j.envsoft.2005.07.004.

Lafond, G.P., H. Loeppky, and D.A. Derksen. 1992. "The effects of tillage systems and crop rotations on soil water conservation, seeding establishment and crop yield." *Can. J. Plant Sci.* https://doi.org/10.4141/cjps92-011.

Li, Yan, Kaiyu Guan, Gary D Schnitkey, Evan DeLucia, and Bin Peng. 2019. "Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States." *Global Change Biology.* doi: 10.1111/gcb.14628.

Li, Yinpeng, Wei Ye, Meng Wang, and Xiaodong Yan. 2009. "Climate change and drought: a risk assessment of crop-yield impacts." *Climate Research.* doi:10.3354/cr00797.

Li, Ziyi, Kaiyu Guan, Wang Zhou, Bin Peng, Zhenong Jin, Jinyun Tang, Robert F. Grant, et al. 2022. "Assessing the impacts of pre-growing-season weather conditions on soil nitrogen dynamics and corn productivity in the U.S. Midwest." *Field Crops Research.* doi:https://doi.org/10.1016/j.fcr.2022.108563.

Liu, Li, Renhe Zhang, and Zhiyan Zuo. 2017. "Effect of Spring Precipitation on Summer Precipitation in Eastern China: Role of Soil Moisture." *American Meteorological Society.* doi:https://doi.org/10.1175/JCLI-D-17-0028.1.

McDonough, Kelsey R., Stacy L. Hutchinson, JM S. Hutchinson, Jonathan L. Case, and Vahid Rahmani. 2018. "Validation and assessment of SPoRT-LIS surface soil moisture estimates for water resources management applications." *Journal of Hydrology.* doi:https://doi.org/10.1016/j.jhydrol.2018.09.007.

Mitchell, Kenneth E., and et. al. 2004. "The multi-instiution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrologoical modeling system." *Journal of Geophysical Research Atmospheres, An AGU Journal.* doi:https://doi.org/10.1029/2003JD003823.

Namias, Jerome. 1991. "Spring and Summer 1988 Drought over the Contiguous United States—Causes and Prediction." *Journal of Climate.* doi:https://doi.org/10.1175/1520-0442(1991)004%3C0054:SASDOT%3E2.0.CO;2.

NASA. n.d. " Short-term Prediction Research and Transition-Land Information System."

NASA . n.d. *Soil Moisture Active Passive SMAP.* https://smap.jpl.nasa.gov/.

Ortiz-Bobea, Ariel. 2013. "Is weather really additive in agricultural production? Implications for climate change impacts." *Discussion Paper - Resources for the Future (RFF).*

Ortiz-Bobea, Ariel, Haoying Wang, Carlos M Carrillo, and Toby Ault. 2019. "Unpacking the climatic drivers of US agricultural yields." *Environmental Research Letters.* https://doi.org/10.1088/1748-9326/ab1e75.

Peters-Lidard, C.D., P.R. Houser, Y. Tian, S.V. Kumar, J. Geiger, S. Olden, L. Lighty, et al. 2007. "High-performance Earch system modeling with NASA/GSFC's Land Information System." *Innovations in Systems and Software Engineering.* doi:https://doi.org/10.1007/s11334-007-0028-x.

Peterson, G.A., A.J. Schlegel, D.L. Tanaka, and O.R. Jones. 1996. "Precipitation use efficiency as affected by cropping and tillage systems." *J. Prod. Agric.* https://doi.org/10.2134/jpa1996.0180.

Pingali, Prahbu L. 2012. "Green Revolution: Impacts, limits, and the path ahead." *(PNAS) Proceedings of the National Academy of Sciences of the United States of America.* doi:10.1073/pnas.0912953109.

Schlegel, Alan J., Yared Assefa, Lucas A. Haag, Curtis Thompson, and Loyd R. Stone. 2019. "Soil Water and Water Use in Long-Term Dryland Crop Rotations." *Agronomy Journal.* https://doi.org/10.2134/agronj2018.09.0623.

Schlenker, W., and M.J. Roberts. 2009. "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change." *Proceedings of the National Academy of Sciences.* https://doi.org/10.1073/pnas.0906865106.

Schoonover, J.E., and J.F. Crim. 2015. "An introduction to soil concepts and the role of soils in watershed management." *Journal of Contemporary Water Research and Education.* doi:https://doi.org/10.1111/j.1936-704X.2015.03186.x.

Schwarz, Gideon. 1978. "Estimating the dimension of a model." *The Annals of Statistics.* doi:https://doi.org/10.1214%2Faos%2F1176344136.

Steduto, Pasquale, Theodore C Hsiao, Elias Fereres, and Dirk Raes. 2012. "Crop yield responses to water." *FAO Rome, Irrigation and Drainage Paper* .

Stone, L. R, R. E Gwin, P. J Gallagher, and M. J. Hattendorf. 1987. "Dormant-season irrigation grain yield, water use, and water loss." *Agronomy Journal.*

Stone, Loyd R, Freddie R Lamm, Alan J Schlegel, and Norman L Klocke. 2008. "Storage efficiency of off-season irrigation." *Agronomy Journal.* doi:10.2134/agronj2007.0242.

Tack, Jesse, A. Barkley, and L. Nalley. 2015. "Effect of warming temperatures on US wheat yields." *Proceedings of the National Academy of Sciences.* https://doi.org/10.1073/pnas.1415181112.

Tavakol, Ameneh, Vahid Rahmani, Steven Quiring, and Sujay Kumar. 2019. "Evaluation analysis of NASA SMAP L3 and L4 and SPoRT-LIS soil moisture data in the United States." *Remote Sensing of Environment.* doi:https://doi.org/10.1016/j.jhydrol.2018.09.007.

Unger, P.W., B.A. Stewart, J.F. Parr, and R.P. Singh. 1991. "Crop management and tillage methods for conserving soil and water in semi-arid regions." *Soil Tillage Res.* https://doi.org/10.1016/0167-1987(91)90041-U.

Urban, Daniel W, Michael J Roberts, Wolfram Schlenker, and David B Lobell. 2015. "The effects of extremely wet planting conditions on maize and soybean yields." *Climatic Change.* doi:10.1007/s10584-015-1362-x.

USDA NASS. 2024. *CroplandCROS and Cropland Data Layer.*

USDA NASS. 2010. "Field Crops Usual Planting and Harvesting Dates."

USDA Risk Management Agency. 2022. "Crop Yield by County."

USDA. March 2025. "World Agricultural Supply and Demand Estimates (WASDE)."

VanToai, Tara, and Christopher Bolles. 1991. "Postanoxic injury in soybean (glycine max) seedlings." *Plant Physiology.* doi:10.1104/pp.97.2.588.

Wu, Wanru, Robert E. Dickinson, Hui Wang, Yongqiang Liu, and Muhammad Shaikh. 2007. "Covariabilities of spring soil moisture and summertime." *International Journal of Climatology.* doi:10.1002/joc.1419.

Yan, Bin, Quijie Dai, Xiaozhong, Liu Liu, Shaobai Huang, and Zixia Wang. 1996. "Flooding-induced membrane damage, lipid oxidation and activated oxygen generation in corn leaves." *Plant and Soil.* doi:10.1007/BF00009336.

Ziyi Li, Kaiyu Guan, Wang Zhou, Bin Peng, Zhenong Jin, Jinyun Tang, Robert F. Grant, Emerson D. Nafziger, Andrew J. Margenot, Lowell E. Gentry, Evan H. DeLucia, Wendy H. Yang, Yaping Cai, Ziqi Qin, Sotirios V. Archontoulis, Fabián G. Fernández, et. al. 2022. "Assessing the impacts of pre-growing-season weather conditions on soil nitrogen dynamics and corn productivity in the U.S. Midwest." *Field Crops Research.* doi:https://doi.org/10.1016/j.fcr.2022.108563.

Zotarelli, L., M. Dukes, and K. Morgan. 2019. "Interpretation of soil moisture content to determine soil field capacity and avoid over-irrigating sandy soils using soil moisture sensors." *University of Florida Institute of Food and Agricultural Sciences* (UF IFAS Extension). https://edis.ifas.ufl.edu/publication/AE460.

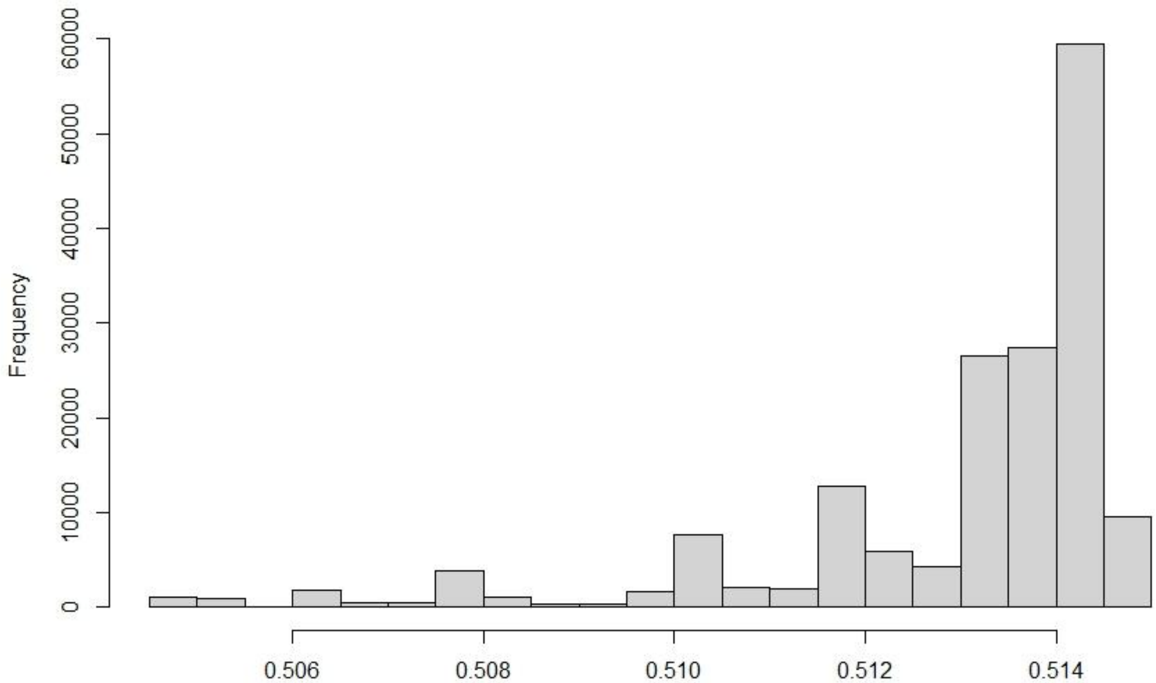
Appendix A - National Model Specification and Testing

Appendix A.1 Soil Moisture Bin Breakpoint Specification

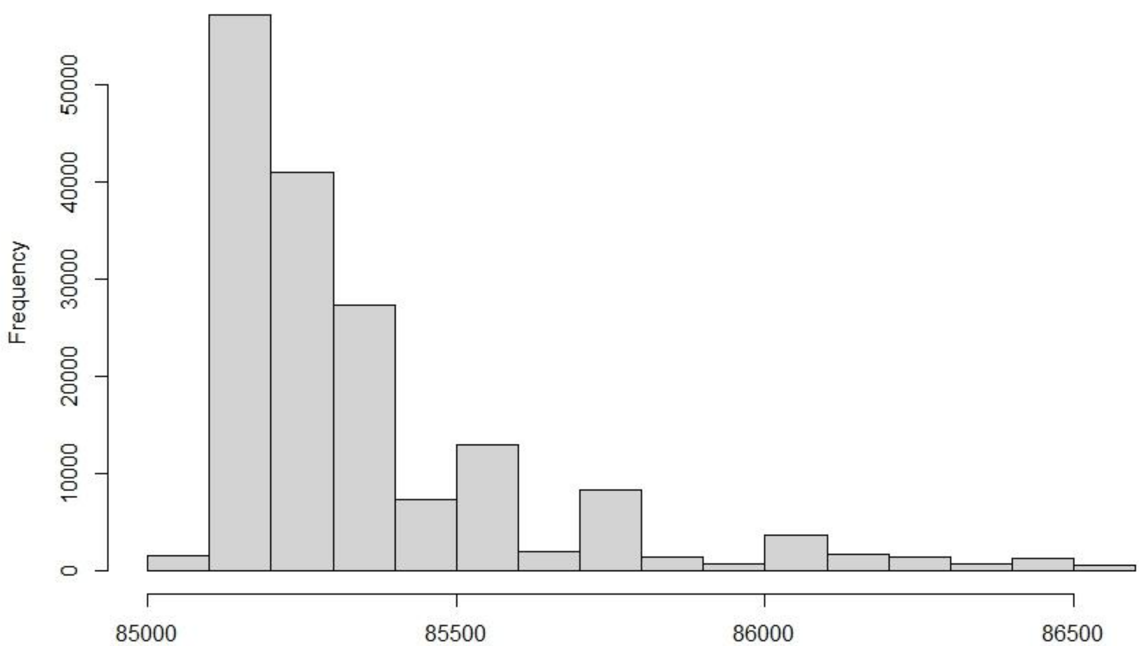
To determine the optimal number and size of soil moisture exposure bins, we employed an iterative, data-driven testing process. Recognizing the trade-off between model complexity and explanatory power, we evaluated millions of different bin configurations by estimating regression models for each scenario. We tested for the optimal number, placement, and size of the bins. For every specification, we calculated two metrics: the adjusted R^2 and the Bayesian Information Criterion (BIC).

Appendix Figure A.1 displays a histogram of the adjusted r2 values for all bin specifications. Similarly, **Appendix Figure A.2** shows the Bayesian information coefficients. Ultimately, we selected the bin configuration that maximized Adjusted R2, slightly adjusting it by dividing the lowest moisture bin to better capture critical nonlinear moisture effects. This specification provided slightly more bins than that which minimized BIC but its ability to capture additional nonlinearity is worth the additional complexity. This final specification was consistently applied across all subsequent regional models.

Appendix Figure A.1 Distribution of Adjusted R2 Values Over All Breakpoint Specifications



Appendix Figure A.2 Distribution of BIC Values Over All Breakpoint Specifications



Appendix A.2 Balanced Panel Robustness Check

Appendix Table A.1

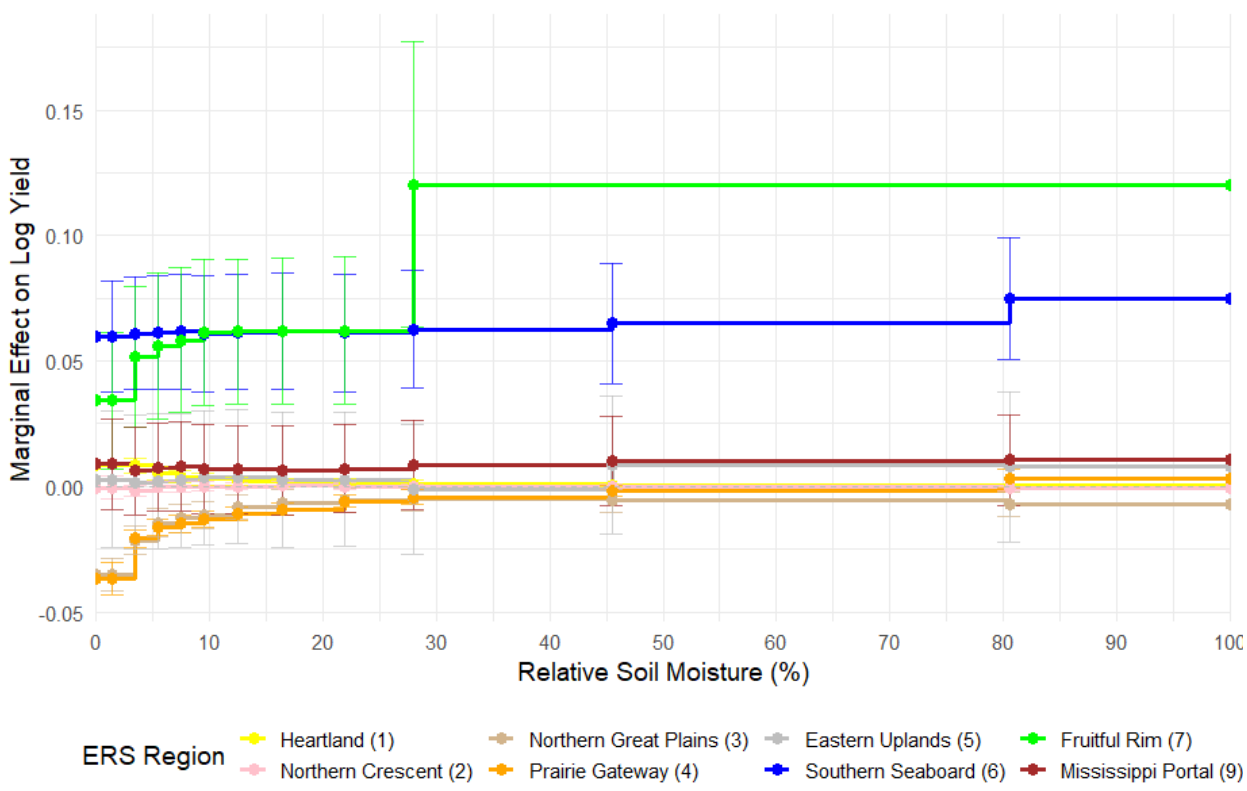
	(1) Unbalanced Panel	(2) Balanced Panel
1 - 3%	-0.0209***	-0.0164***
	(0.0033)	(0.0033)
4 - 6%	-0.0058*	-0.0046
	(0.0025)	(0.0024)
7 - 10%	-0.0018	-0.0012
	(0.0022)	(0.0022)
11 - 15%	0.0030	0.0030
	(0.0020)	(0.0020)
16 - 25%	0.0063***	0.0052**
	(0.0018)	(0.0016)
26 - 35%	0.0088***	0.0081***
	(0.0015)	(0.0014)
36 - 45%	0.0090***	0.0084***
	(0.0014)	(0.0013)
46 - 65%	0.0100***	0.0090 ***
	(0.0014)	(0.0013)
66 - 85%	0.0106***	0.0098***
	(0.0015)	(0.0014)
86 - 100%	0.0087***	0.0076***
	(0.0018)	(0.0016)
Time	0.0013	0.0030
	(0.0018)	(0.0016)
Time^2	0.0005***	0.0004***
	(0.0001)	(0.0000)
Observations	59846	57598
Adj. ^2	0.5092	0.4966
RMSE	0.3789	0.3484

Appendix B - ERS Farm Resource Region Models

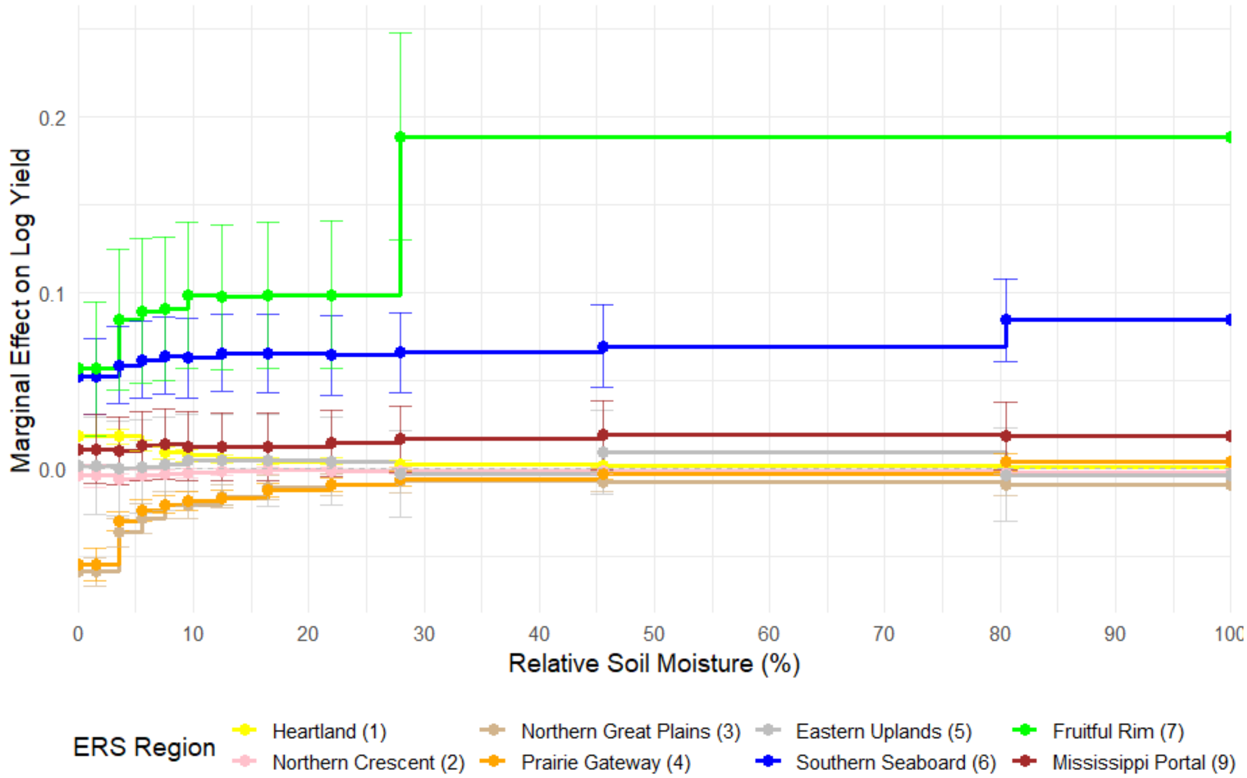
Appendix B.1 Time Separability Testing

The following figures test the relationship between regions over different time window specifications. It provides some support to the implicit assumption of time separability, indicating that regional results remain relatively consistent regardless of the time window specification.

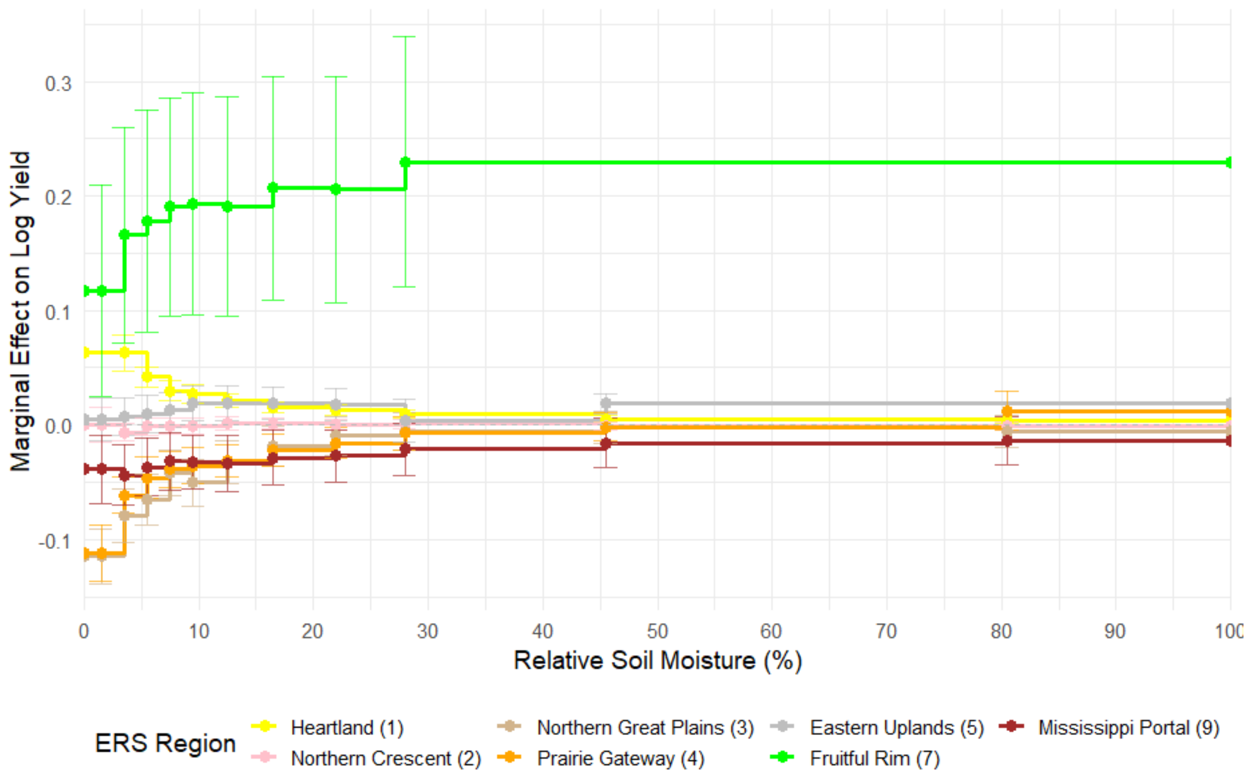
Appendix Figure B.1 Exposure-response Estimates by Region (January 1st - March 15th)



Appendix Figure B.2 Exposure-response Estimates by Region (February 1st -March 15th)



Appendix Figure B.3 Exposure-response Estimates by Region (March 1st - March 15th)



Appendix B.2 ERS Model Results

The following tables display detailed model results for the ERS regional models reported in the study.

Appendix Table B.1 ERS Heartland

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0807	0.0036	22.7013	<0.001 ***
4 - 6%	0.0251	0.0036	6.9421	<0.001 ***
7 - 10%	0.0188	0.0034	5.5908	<0.001 ***
11 - 15%	0.0128	0.0028	4.5091	<0.001 ***
16 - 25%	0.0098	0.0025	3.8951	<0.001 ***
26 - 35%	0.0089	0.0018	5.0231	<0.001 ***

36 - 45%	0.0093	0.0014	6.7391	<0.001 ***
46 - 65%	0.0095	0.0013	7.3156	<0.001 ***
66 - 85%	0.0095	0.0015	6.4964	<0.001 ***
86 - 100%	0.0069	0.0016	4.1810	<0.001 ***
time	0.0110	0.0017	6.4071	<0.001 ***
time^2	0.0001	0.0000	2.2512	0.024 *
Observations:	16554			
Adj. R-squared:	0.4300			
Standard-errors:	Conley (50km)			
RMSE:	0.22469			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table B.2 ERS Northern Cresent

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0207	0.0040	5.1549	<0.001 ***
4 - 6%	0.0163	0.0030	5.4585	<0.001 ***
7 - 10%	0.0129	0.0026	5.0140	<0.001 ***
11 - 15%	0.0117	0.0024	4.7881	<0.001 ***
16 - 25%	0.0078	0.0022	3.4850	<0.001 ***
26 - 35%	0.0055	0.0017	3.2833	0.001 **
36 - 45%	0.0047	0.0016	3.0416	0.002 **
46 - 65%	0.0032	0.0012	2.6678	0.008 **
66 - 85%	0.0024	0.0014	1.7707	0.077 .
86 - 100%	-0.0015	0.0025	-0.6035	0.546
time	0.0091	0.0019	4.8270	<0.001 ***
time^2	0.0002	0.0001	3.1372	0.002 **
Observations:	9081			
Adj. R-squared:	0.4244			
Standard-errors:	Conley (50km)			
RMSE:	0.24648			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table B.3 ERS Northern Great Plains

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	-0.0597	0.0079	-7.5634	<0.001 ***
4 - 6%	-0.0394	0.0033	-11.9287	<0.001 ***
7 - 10%	-0.0259	0.0034	-7.7227	<0.001 ***
11 - 15%	-0.0152	0.0034	-4.4769	<0.001 ***
16 - 25%	-0.0051	0.0031	-1.6206	0.105
26 - 35%	0.0003	0.0027	0.1146	0.909
36 - 45%	-0.0023	0.0026	-0.9026	0.367
46 - 65%	0.0020	0.0032	0.6341	0.526
66 - 85%	-0.0020	0.0027	-0.7559	0.450
86 - 100%	0.0023	0.0051	0.4620	0.644
time	0.0267	0.0105	2.5395	0.011 *
time^2	0.0003	0.0003	1.1038	0.270
Observations:	4073			
Adj. R-squared:	0.4647			
Standard-errors:	Conley (50km)			
RMSE:	0.55020			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table B.4 ERS Prairie Gateway

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0588	0.0230	2.5591	0.011 *
4 - 6%	0.0832	0.0231	3.6022	<0.001 ***
7 - 10%	0.0911	0.0232	3.9266	<0.001 ***
11 - 15%	0.0969	0.0229	4.2365	<0.001 ***
16 - 25%	0.1048	0.0230	4.5525	<0.001 ***
26 - 35%	0.1099	0.0229	4.8026	<0.001 ***
36 - 45%	0.1112	0.0230	4.8416	<0.001 ***

46 - 65%	0.1135	0.0234	4.8456	<0.001 ***
66 - 85%	0.1254	0.0235	5.3343	<0.001 ***
86 - 100%	0.1247	0.0253	4.9363	<0.001 ***
time	-0.0523	0.0086	-6.0735	<0.001 ***
time^2	0.0020	0.0002	8.4874	<0.001 ***
Observations:	7324			
Adj. R-squared:	0.2539			
Standard-errors:	Conley (50km)			
RMSE:	0.64378			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table B.5 ERS Eastern Uplands

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0381	0.0319	1.1944	0.232
4 - 6%	0.0389	0.0310	1.2532	0.210
7 - 10%	0.0413	0.0312	1.3225	0.186
11 - 15%	0.0406	0.0307	1.3221	0.186
16 - 25%	0.0395	0.0303	1.3028	0.193
26 - 35%	0.0305	0.0303	1.0071	0.314
36 - 45%	0.0389	0.0300	1.2951	0.195
46 - 65%	0.0327	0.0301	1.0840	0.278
66 - 85%	-0.0096	0.0316	-0.3021	0.763
86 - 100%	*NA*	*NA*	*NA*	*NA*
time	0.0088	0.0039	2.2861	0.022 *
time^2	0.0002	0.0001	1.3331	0.183
Observations:	7116			
Adj. R-squared:	0.3189			
Standard-errors:	Conley (50km)			
RMSE:	0.31046			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table B.6 ERS Southern Seaboard

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0128	0.0136	0.9461	0.344
4 - 6%	0.0176	0.0137	1.2895	0.197
7 - 10%	0.0188	0.0135	1.3874	0.165
11 - 15%	0.0224	0.0136	1.6408	0.101
16 - 25%	0.0215	0.0137	1.5731	0.116
26 - 35%	0.0236	0.0137	1.7193	0.086 .
36 - 45%	0.0248	0.0137	1.8080	0.071 .
46 - 65%	0.0281	0.0137	2.0501	0.040 *
66 - 85%	0.1027	0.0139	7.3665	<0.001 ***
86 - 100%	*NA*	*NA*	*NA*	*NA*
time	-0.0012	0.0030	-0.3928	0.694
time^2	0.0005	0.0001	5.5522	<0.001 ***
Observations:	9827			
Adj. R-squared:	0.3286			
Standard-errors:	Conley (50km)			
RMSE:	0.35929			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table B.7 ERS Mississippi Portal

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0297	0.0083	3.5633	<0.001 ***
4 - 6%	0.0301	0.0077	3.9189	<0.001 ***
7 - 10%	0.0327	0.0080	4.0879	<0.001 ***
11 - 15%	0.0324	0.0073	4.4621	<0.001 ***
16 - 25%	0.0332	0.0074	4.5009	<0.001 ***
26 - 35%	0.0380	0.0077	4.9651	<0.001 ***
36 - 45%	0.0395	0.0078	5.0751	<0.001 ***
46 - 65%	0.0415	0.0079	5.2441	<0.001 ***

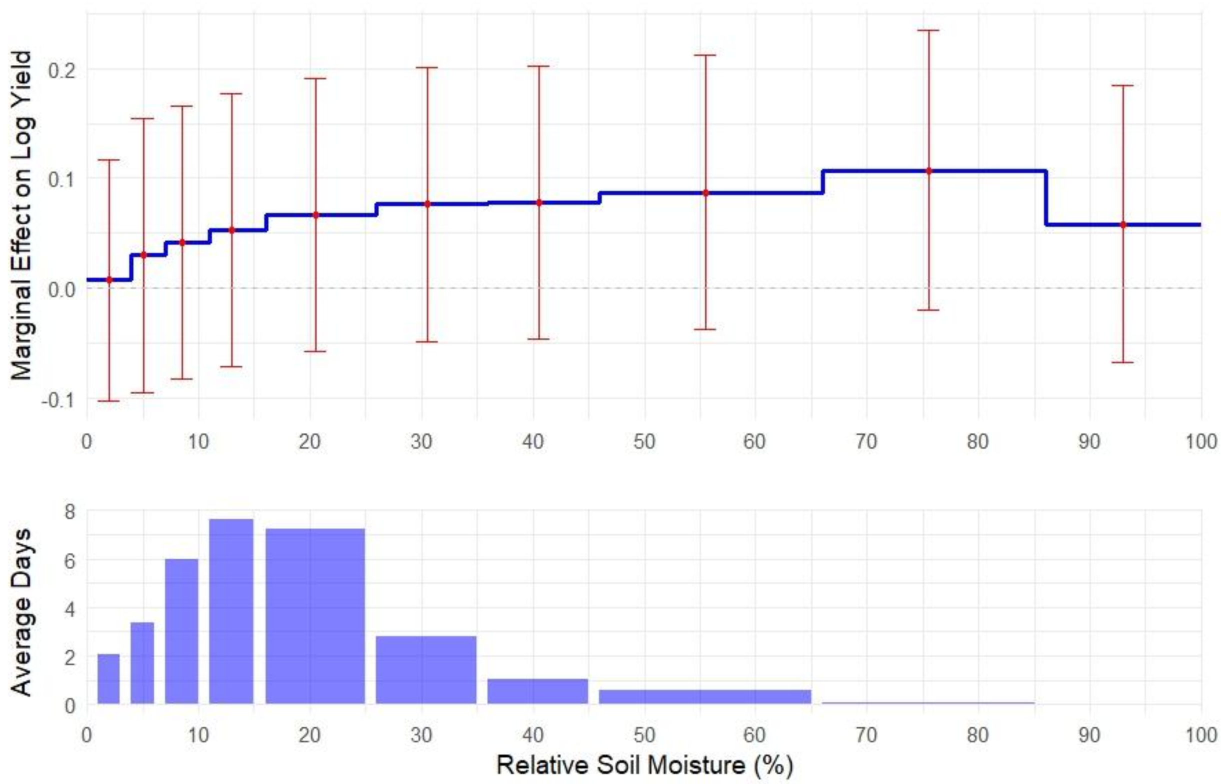
66 - 85%	0.0428	0.0082	5.2129	<0.001 ***
86 - 100%	0.0496	0.0089	5.5668	<0.001 ***
time	0.0061	0.0028	2.1689	0.030 *
time^2	0.0003	0.0001	3.7042	<0.001 ***
Observations:	4267			
Adj. R-squared:	0.4228			
Standard-errors:	Conley (50km)			
RMSE:	0.20928			

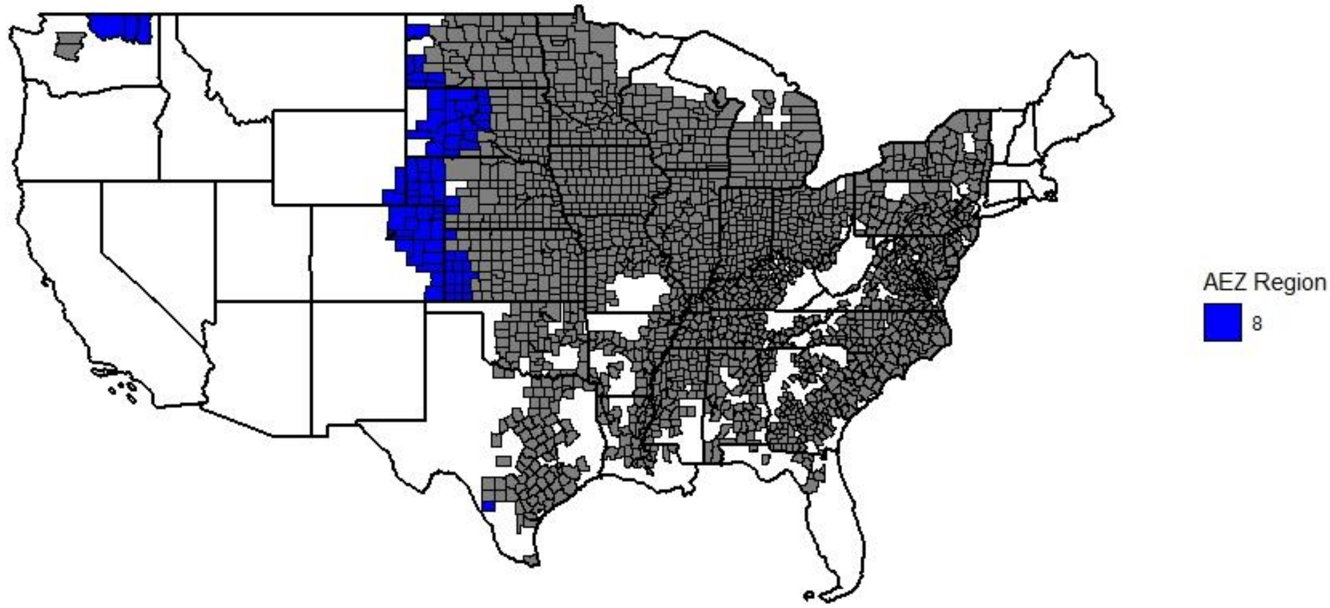
Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix C - Spatial Heterogeneity Robustness Tests

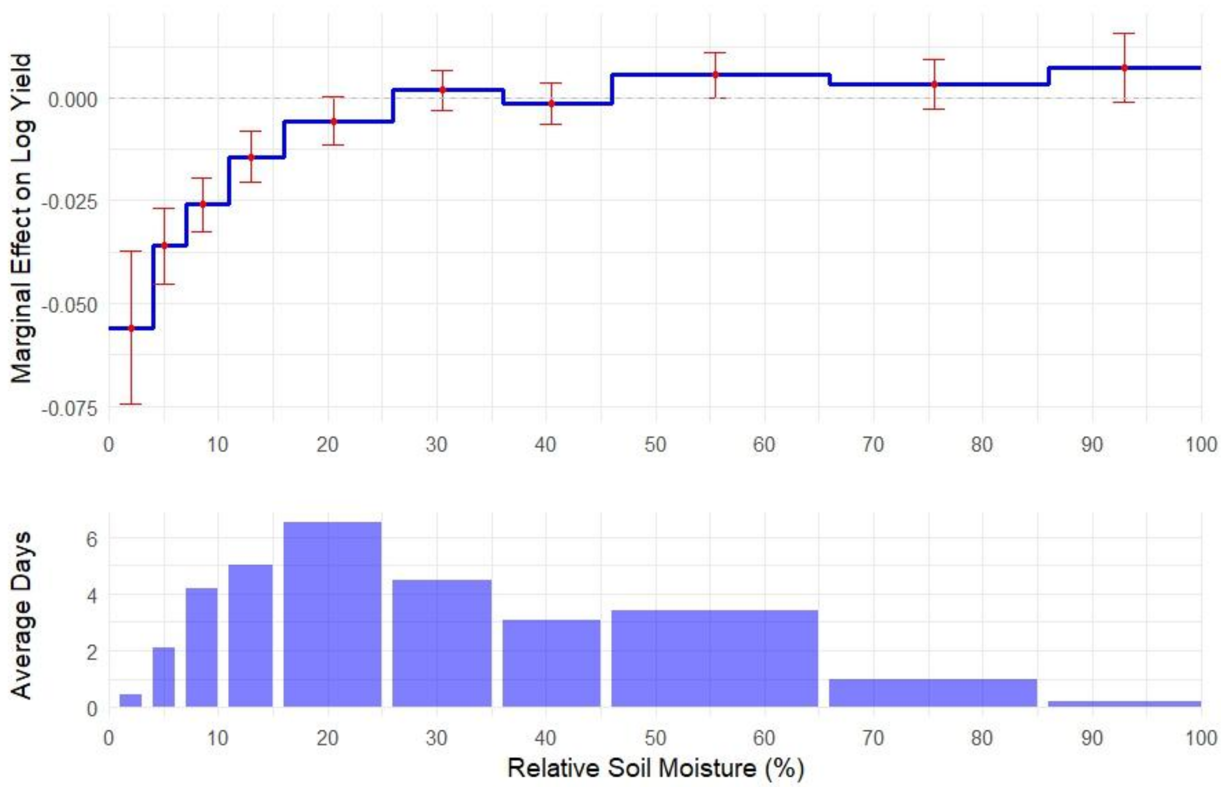
To test the robustness of the regional analysis, the following figures display the results of the agroecological zones and Iowa models.

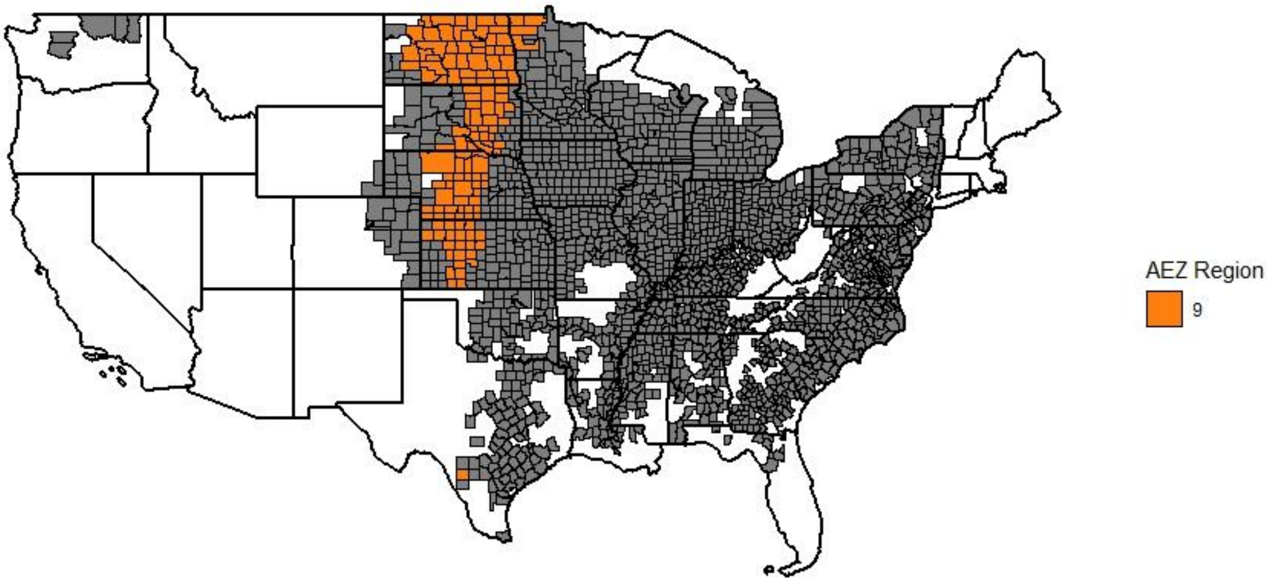
Appendix Figure C.1 AEZ 8



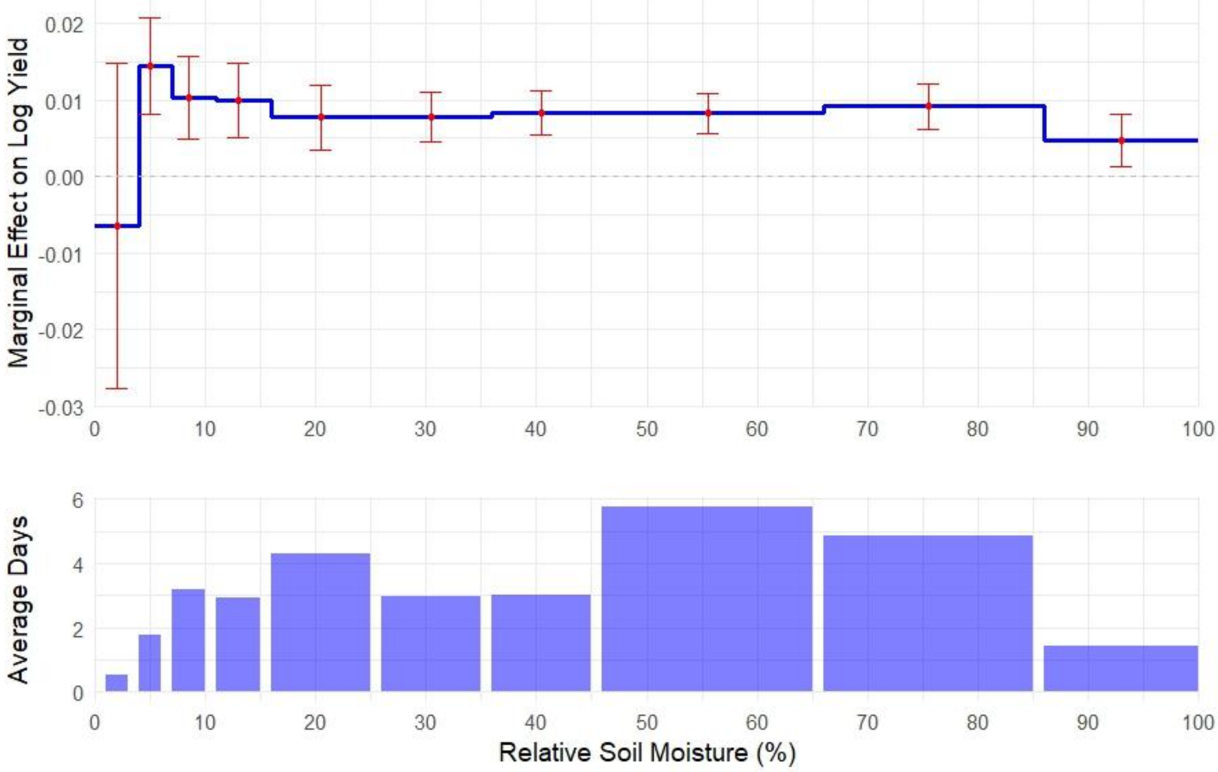


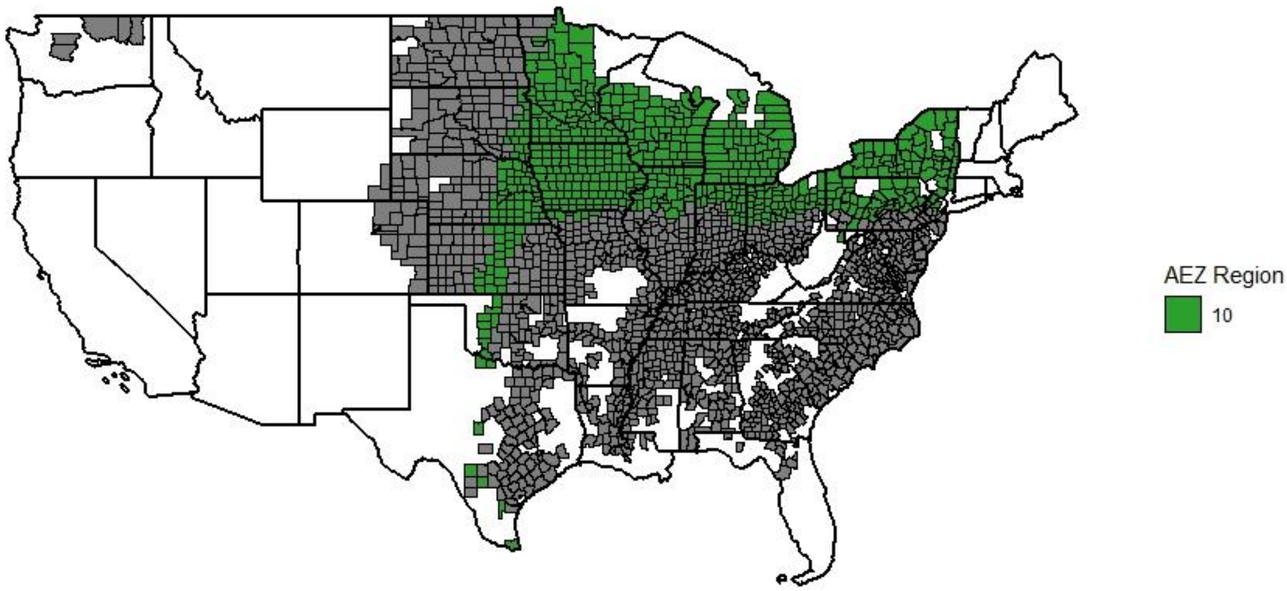
Appendix Figure C.2 AEZ 9



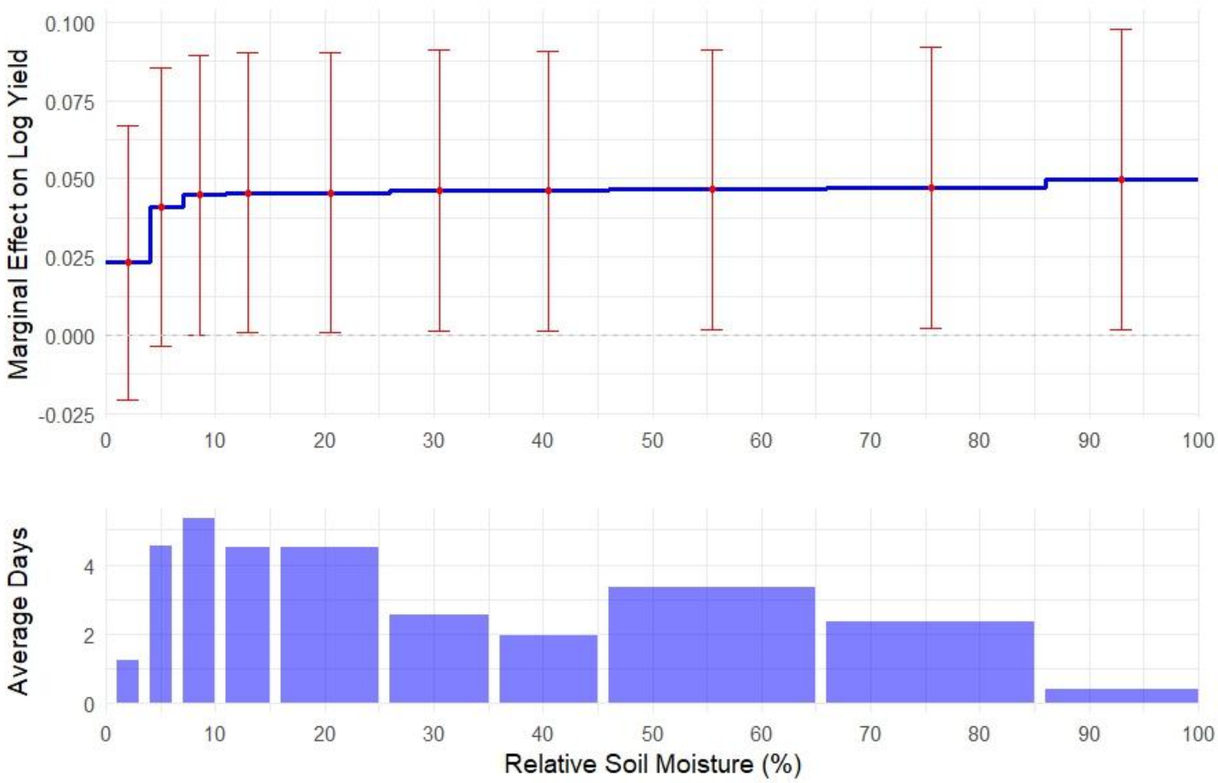


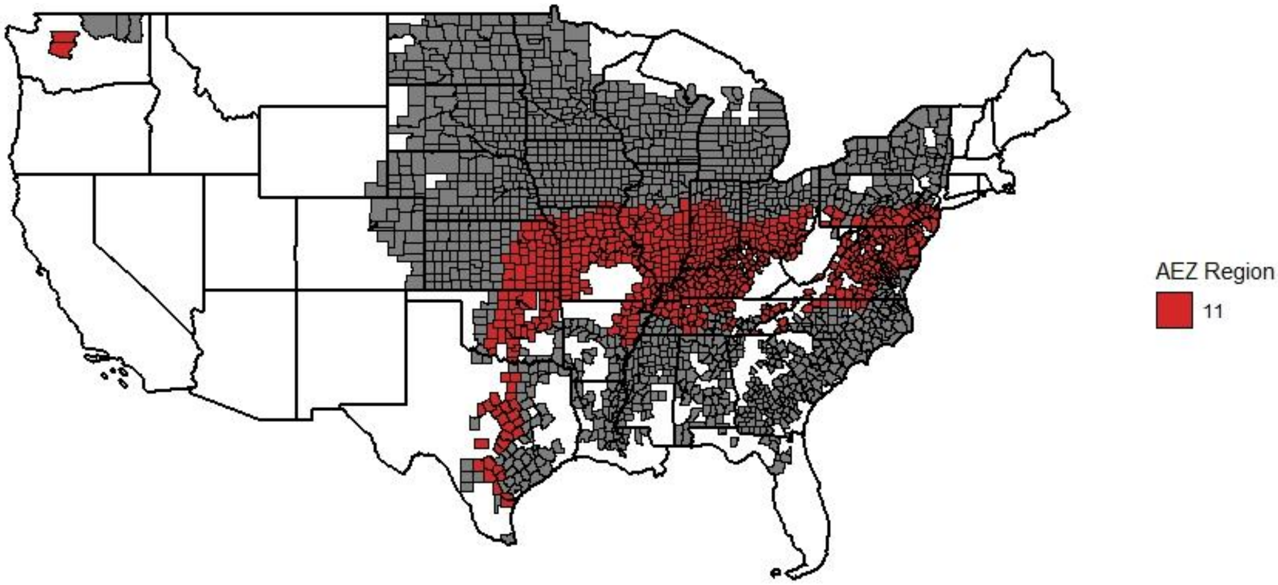
Appendix Figure C.3 AEZ 10



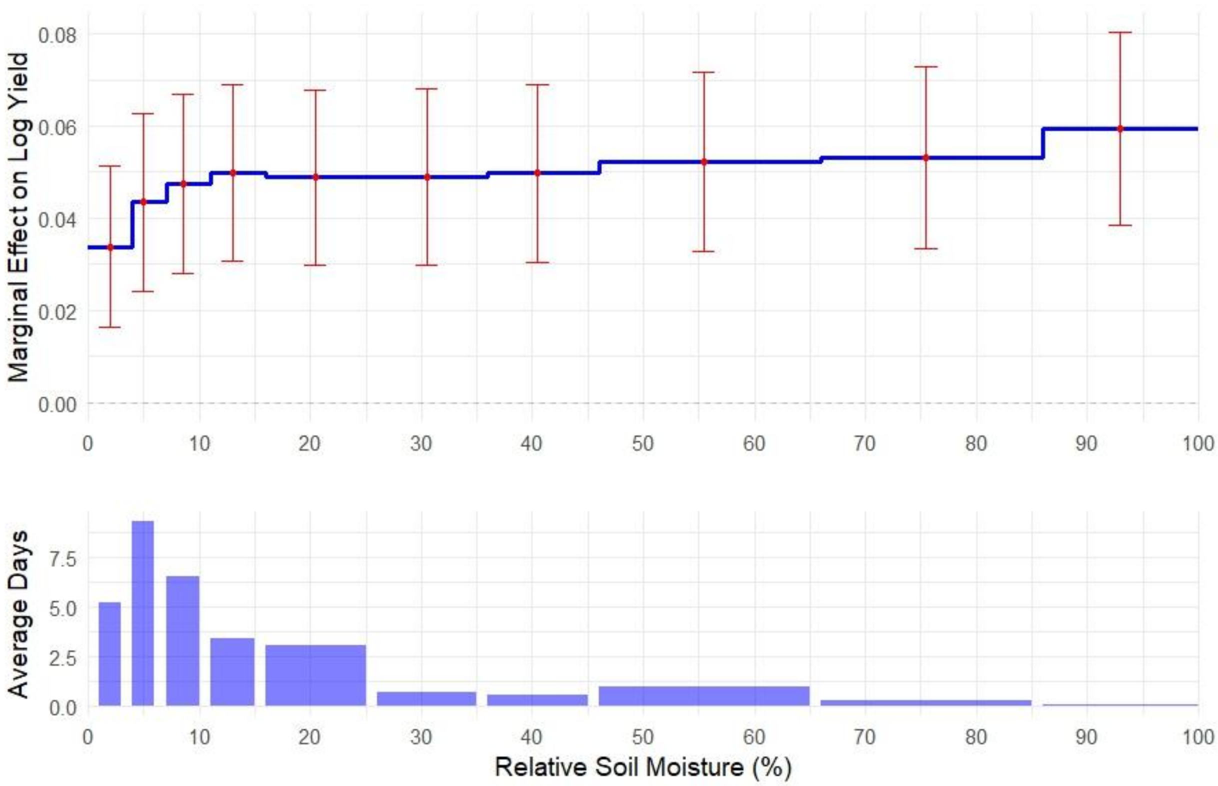


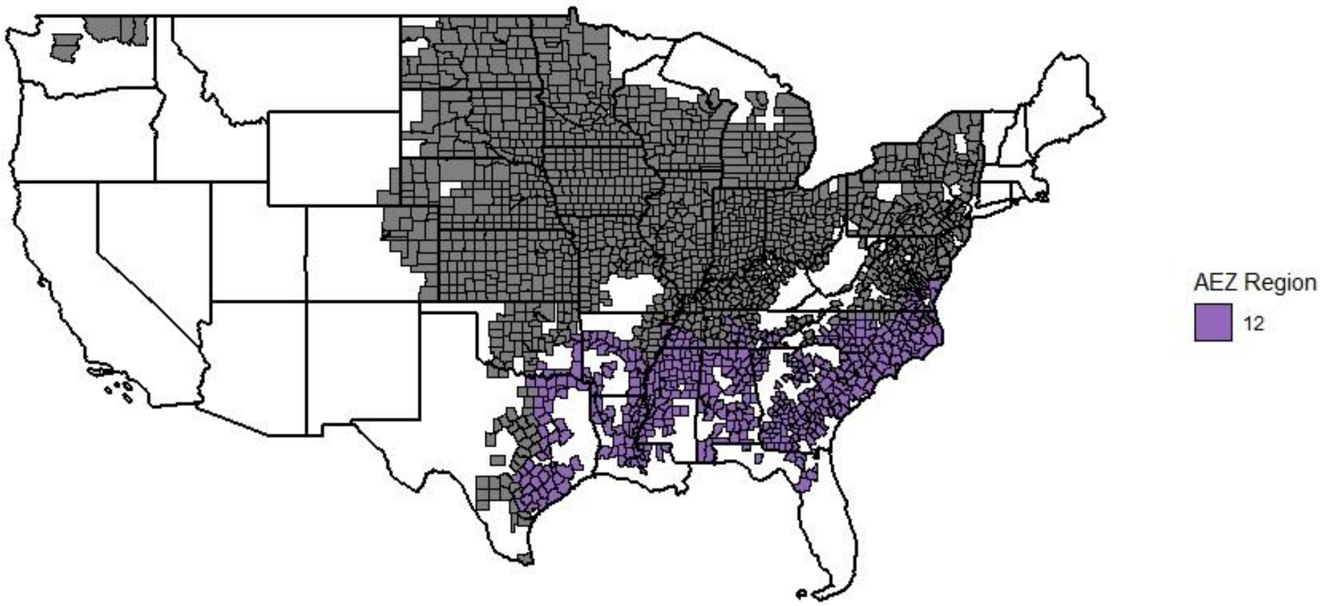
Appendix Figure C.4 AEZ 11



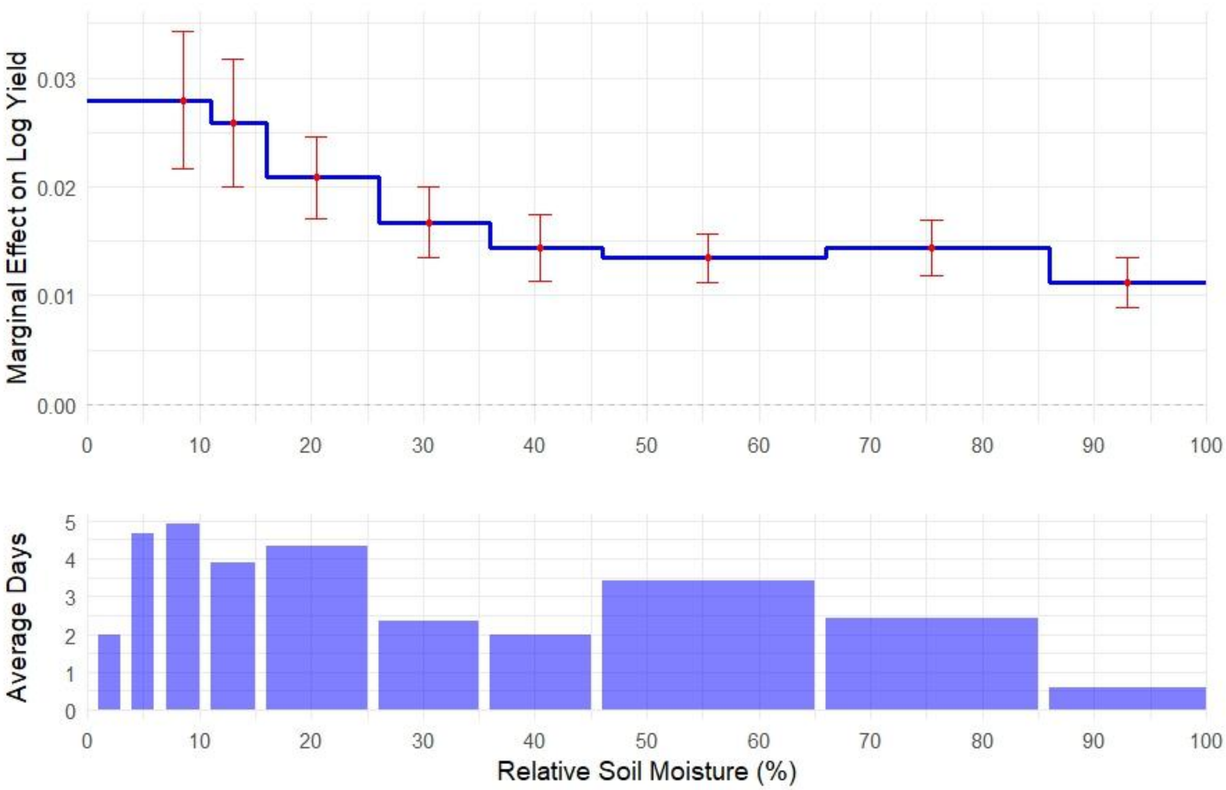


Appendix Figure C.5 AEZ 12





Appendix Figure C.6 Iowa Model



Appendix Table C.1 AEZ 8

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0074	0.0560	0.1316	0.895
4 - 6%	0.0299	0.0637	0.4697	0.639
7 - 10%	0.0421	0.0633	0.6644	0.507
11 - 15%	0.0529	0.0634	0.8337	0.405
16 - 25%	0.0669	0.0634	1.0556	0.291
26 - 35%	0.0761	0.0638	1.1927	0.233
36 - 45%	0.0783	0.0633	1.2372	0.216
46 - 65%	0.0872	0.0637	1.3679	0.171
66 - 85%	0.1074	0.0650	1.6511	0.099 .
86 - 100%	0.0584	0.0644	0.9063	0.365
time	-0.0545	0.0201	-2.7060	0.007 **
time^2	0.0023	0.0005	4.6023	<0.001 ***
Observations:	2272			
Adj. R-squared:	0.3037			
Standard-errors:	Conley (50km)			
RMSE:	0.7026			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table C.2 AEZ 9

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	-0.0559	0.0095	-5.9075	<0.001 ***
4 - 6%	-0.0360	0.0047	-7.6920	<0.001 ***
7 - 10%	-0.0260	0.0033	-7.8808	<0.001 ***
11 - 15%	-0.0143	0.0031	-4.6005	<0.001 ***
16 - 25%	-0.0056	0.0029	-1.8924	0.059 .
26 - 35%	0.0018	0.0025	0.7503	0.453
36 - 45%	-0.0013	0.0025	-0.4935	0.622
46 - 65%	0.0057	0.0028	1.9902	0.047 *
66 - 85%	0.0033	0.0031	1.0786	0.281

86 - 100%	0.0074	0.0043	1.7167	0.086 .
time	-0.0109	0.0131	-0.8346	0.404
time^2	0.0013	0.0003	3.7329	<0.001 ***
Observations:	4042			
Adj. R-squared:	0.3086			
Standard-errors:	Conley (50km)			
RMSE:	0.61386			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table C.3 AEZ 10

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	-0.0064	0.0108	-0.5960	0.551
4 - 6%	0.0144	0.0032	4.4436	<0.001 ***
7 - 10%	0.0103	0.0027	3.7704	<0.001 ***
11 - 15%	0.0100	0.0025	4.0548	<0.001 ***
16 - 25%	0.0077	0.0021	3.6305	<0.001 ***
26 - 35%	0.0078	0.0016	4.7558	<0.001 ***
36 - 45%	0.0083	0.0015	5.6059	<0.001 ***
46 - 65%	0.0083	0.0014	6.1132	<0.001 ***
66 - 85%	0.0092	0.0015	6.0433	<0.001 ***
86 - 100%	0.0047	0.0018	2.6714	0.008 **
time	0.0109	0.0021	5.2077	<0.001 ***
time^2	0.0001	0.0001	2.4554	0.014 *
Observations:	18537			
Adj. R-squared:	0.5514			
Standard-errors:	Conley (50km)			
RMSE:	0.29091			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table C.4 AEZ 11

Variable	Estimate	Std. Error	t Value	p Value

1 - 3%	0.0231	0.0223	1.0369	0.300
4 - 6%	0.0410	0.0227	1.8063	0.071 .
7 - 10%	0.0447	0.0228	1.9613	0.050 *
11 - 15%	0.0453	0.0228	1.9899	0.047 *
16 - 25%	0.0455	0.0228	1.9967	0.046 *
26 - 35%	0.0462	0.0228	2.0244	0.043 *
36 - 45%	0.0461	0.0227	2.0275	0.043 *
46 - 65%	0.0465	0.0227	2.0461	0.041 *
66 - 85%	0.0469	0.0229	2.0492	0.040 *
86 - 100%	0.0499	0.0244	2.0408	0.041 *
time	-0.0010	0.0019	-0.5098	0.610
time^2	0.0005	0.0001	8.2889	<0.001 ***
Observations:	20126			
Adj. R-squared:	0.4570			
Standard-errors:	Conley (50km)			
RMSE:	0.33810			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1

Appendix Table C.5 AEZ 12

Variable	Estimate	Std. Error	t Value	p Value
1 - 3%	0.0337	0.0089	3.7862	<0.001 ***
4 - 6%	0.0435	0.0099	4.4155	<0.001 ***
7 - 10%	0.0476	0.0099	4.8025	<0.001 ***
11 - 15%	0.0498	0.0097	5.1277	<0.001 ***
16 - 25%	0.0489	0.0097	5.0593	<0.001 ***
26 - 35%	0.0489	0.0098	5.0089	<0.001 ***
36 - 45%	0.0498	0.0098	5.0802	<0.001 ***
46 - 65%	0.0522	0.0099	5.2658	<0.001 ***
66 - 85%	0.0531	0.0100	5.2845	<0.001 ***
86 - 100%	0.0595	0.0106	5.5887	<0.001 ***
time	0.0030	0.0026	1.1529	0.249

time^2	0.0004	0.0001	4.6901	<0.001 ***
Observations:	14869			
Adj. R-squared:	0.3895			
Standard-errors:	Conley (50km)			
RMSE:	0.35770			

Note: Standard errors in parentheses. ***p<0.001, **p<0.01, *p<0.05, .p<0.1