Understanding Missing National Agricultural Statistics Survey Yields

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

William Hence Duncan
May 2024

ACKNOWLEDGEMENTS

I would like to say thank you to everyone in the department of Agricultural and Resource Economics at The University of Tennessee for my time in the department. I have enjoyed being a part of the department for my undergraduate and graduate studies. A special thanks to Dr. Boyer and Dr. Smith for investing in me during my undergraduate career as well as continued investment throughout my masters. I would also like to thank Mr. Jon Walton for his support and mentorship throughout my time in Knoxville.

ABSTRACT

This thesis consists of two chapters that explore factors that impact county corn and soybean yields that are not reported by United States Department of Agriculture National Agricultural Statistics Service (NASS). Yields are important to several stakeholders for decision making and constructing policy. The objective of the first chapter is to determine factors that impact the likelihood of NASS corn or soybean yields being reported at the county-level. Factors, such as county land coverage and average farm size are used to explain the likelihood of a yield being reported. The second chapter focuses on utilizing other publicly available yield data to replace or fill in gaps for NASS yields when they are not reported. Specifically, this thesis compares NASS yields to other yield data across Tennessee. A spatial analysis was conducted to determine hot and cold spots of similar yield values for the various sources. The results from chapter two may be helpful to determine a county level yield when NASS yields are not reported.

TABLE OF CONTENTS

Introduction 1
CHAPTER I UNDERSTANDING MISSING NASS YIELD CHARACTERISTICS 4
Abstract 5
Introduction 6
Data 9
Methods 11
Results 12
Conclusion 13
CHAPTER II NASS YIELD ALTERTATIVES 15
Abstract 16
Introduction 17
Data 19
Methods 21
Results 22
Corn 22
Soybeans 25
Comparison of Corn and Soybeans 27
NASS Yield Substitution 27
Next Steps 28
Conclusion 29
Conclusion 31
References 32
Appendix 36
Vita 74

LIST OF TABLES

Table 1. NASS Corn Yield Reporting by County Landscape Logit Model Variable Summary Stats 42
Table 2. NASS Soybean Yield Reporting by County Landscape Logit Model Variable Summary Stats 43
Table 3. Logit Model Results for NASS Yield Reporting by County Landscape 44
Table 4. Average Marginal Effects of NASS Yield Reporting 45
Table 5. Tennessee Corn Olympic Average Yields by County, 2018-2022 47
Table 6. Tennessee Soybean Olympic Average Yields by County, 2018-2022 48
Table 7. Tennessee Corn Olympic Average Yield Summary Stats, 2018-2022 49
Table 8. Tennessee Soybean Yield Olympic Average Summary Stats, 2018-2022 50
Table 9. University of Tennessee Corn Variety Trial Average Yields, 2018 to 2022 69
Table 10. University of Tennessee Soybean Variety Trial Average Yields, 2018 to 2022 70
Table 11. University of Tennessee Variety Trial Summary Stats, 2018-2022 71

LIST OF FIGURES

Figure 1. Percentage of Missing NASS County Corn and Soybean Yields, 2011-2022 .. 37
Figure 2. Average Planted Corn Acres 2011-2022 USDA CroplandCROS 38
Figure 3. Average Planted Soybean Acres 2011-2022 USDA CroplandCROS 39
Figure 4. Percentage of NASS Corn Yields Reported 2011-2022 40
Figure 5. Percentage of NASS Soybean Yields Reported 2011-2022 41
Figure 6. Predicted Probability of a NASS Yield Being Reported by Farm Size 46
Figure 7. Hot Spot Analysis of NASS Corn Olympic Yield Averages 51
Figure 8. Hot Spot Analysis of FSA Corn Olympic Yield Averages 52
Figure 9. Hot Spot Analysis of APH Corn Yield Olympic Averages 53
Figure 10. Hot Spot Comparison of the Olympic Averages of NASS and FSA Corn Yields from 2018 to 2022 54
Figure 11. Hot Spot Comparison of Olympic Averages of NASS and APH Corn Yields from 2018 to 2022 55
Figure 12. Hot Spot Comparison of Olympic Averages of FSA and APH Corn Yields from 2018 to 2022 56
Figure 13. Multivariate Cluster Analysis of the Olympic Averages of Corn NASS Yields and FSA Yields from 2018 to 2022 57
Figure 14. Multivariate Cluster Analysis of the Olympic Averages of Corn NASS Yields and APH Yields from 2018 to 2022 58
Figure 15. Multivariate Cluster Analysis of the Olympic Averages of Corn FSA Yields and APH Yields from 2018 to 2022 59
Figure 16. Hot Spot Analysis of NASS Soybean Olympic Yield Averages 60
Figure 17. Hot Spot Analysis of FSA Soybean Olympic Yield Averages 61
Figure 18. Hot Spot Analysis of APH Soybean Olympic Yield Averages 62
Figure 19. Hot Spot Comparison of Olympic Averages of NASS and FSA Soybean Yields from 2018 to 2022 63
Figure 20. Hot Spot Comparison of Olympic Averages of NASS and APH Soybean Yields from 2018 to 2022 64
Figure 21. Hot Spot Comparison of Olympic Averages of FSA and APH Soybean Yields from 2018 to 2022 65
Figure 22. Multivariate Cluster Analysis of the Olympic Averages of Soybean NASS Yields and FSA Yields from 2018 to 2022 66
Figure 23. Multivariate Cluster Analysis of the Olympic Averages of Soybean NASS Yields and APH Yields from 2018 to 2022 67
Figure 24. Multivariate Cluster Analysis of the Olympic Averages of Soybean FSA Yields and APH Yields from 2018 to 2022 68
Figure 25. Corn Variety Trial Yields as a Percentage of Corn NASS Olympic Yield Averages, 2018-2022 72
Figure 26. Soybean Variety Trial Yields as a Percentage of NASS Soybean Olympic Yield Averages, 2018-2022 73

INTRODUCTION

The United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) is a government agency primarily focused on collecting and publishing numerous agricultural statistics. One of the most utilized statistics collected by USDA NASS are annual county crop yield averages. These county yield estimates are often used to analyze crop payments, evaluate USDA and State programs , and for research (Li et al. 2020). For example, prior to the Agricultural Improvement Act of 2018, NASS yields were the primary yield data used when calculating payments for the Agricultural Risk Coverage-County (ARC-CO) payments, which are payments when the actual county crop revenue is less than the ARC-CO guarantees. These are foundational payments that farms depend on for managing risk and remaining profitable.

This yield data is collected as a part of the USDA NASS annual Agricultural Yield survey, which is administered for all states excluding Alaska and Hawaii. The county yields are based on numerous data sources including producer reported yields from proportional samples of various crops in each county, which starts in the fall and continues through the end of January. Respondents, or producers, report their yields based on their average field-based estimates. The NASS yield survey is collected through mail, phone interviews, in person interviews, and electronically via email with most responses collected through phone calls. Statisticians check the results to analyze for outliers before being inserted into a computer system to analyze for outliers again. Then after correcting the outliers and extreme data points it is summarized by county and then released.

There has always been a concern that NASS yield estimates are incomplete or potentially inaccurate. The primary reason is potential errors in estimated yields due to collect self-reported, survey-based data, which can be subject to unreliable responses (Schbnepf 2017). Specifically, producers can intentionally or unintentionally report a yield that does not truly reflect their actual production. Another growing concern is the recent decline in survey responses to NASS yield surveys. Response rates for the NASS survey collecting yield data have declined between the years 1992-2016 with the trend accelerating beginning in 2012 (Johansson et al. 2017; Schbnepf 2017). Figure 1 displays the percentage of missing NASS county corn and soybean yields from 2011-2022. All tables and figures are located in an appendix at the end of the document.

The declining response rates creates a challenge for USDA because they will not report yield data if the identity of the respondent can be revealed. Thus, NASS will not publish data for counties where low numbers of producers or acres in production may disclose the business of individual producers. NASS requires at least 30 producers or 25% of harvested acres to be reported to release a yield for a county. While the sampling frame at a national level is still accurate, the declining response rate and their report criteria of yield can result in magnifying inaccurate reflections of county-level yield estimates (Johansson et al. 2017; Schbnepf 2017). The growing decline in non-reported yields could be a function of increased farm consolidation over time as well as the declining response rate. There were 1.9 million farms in the United States in 2002 (USDA NASS Ag Census 2002), in 2022 there were 2.2 million farms (USDA NASS Ag Census 2022). However, there is little understanding if farm consolidation has impacted

yield reporting by USDA NASS. The first chapter of this thesis will specifically address this gap in literature.

The limitations of NASS yields resulted in the Agricultural Improvement Act of 2018 switching from using NASS yields as the preferred yield for the ARC-CO payment calculations to USDA Risk Management Agency (RMA) yields as the primary yield used to calculate payments. USDA RMA datasets have estimated to be missing over 20% of counties for irrigated corn, 26% for non-irrigated corn, 36% for irrigated soybeans, and almost 40% for non-irrigated soybeans (Ishee 2020).

Therefore, there is a need for novel ways to estimate missing county yields. Researchers have looked to fill the data gaps using novel methods such as Kriging (Ishee 2020; Park, Harri, and Coble 2022). Another approach is to utilize other data sources like USDA RMA Summary of Business statistics to estimate yields. The RMA compiles their yearly Summary of Business report from several sources including, but not limited to, policyholder data, participating insurance companies, and other USDA agencies. A recent report by the National Academy of Sciences, recommended that NASS should collaborate with RMA to be able to analyze and compare RMA data to NASS data to improve accuracy (National Academies of Sciences, Engineering, and Medicine 2017). Other data sources recommended in the report include data from precision agricultural databases and the USDA Farm Service Agency (FSA). The second chapter of this thesis specifically addresses the need to find alternative yield sources when USDA NASS yields are not reported.

CHAPTER I
UNDERSTANDING MISSING NASS YIELD CHARACTERISTICS

Abstract

The United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) yield data are valuable to many groups such as government agencies, researchers, market analysts, and producers. Having reliable data on United States food and fiber production are required to set policies, make decisions, and model the impact of weather and other factors on production. However, over the course of time the survey's that contribute to USDA NASS County yield estimates have experienced a decline in response rates. The trend of declining response rates and USDA NASS rules for reporting yields can result in NASS yields being biased which has potential implications for research and payment calculations. The objective in this chapter is to determine if the quantity of acres reported for corn and soybeans within individual counties impacts the likelihood of corn and soybean yields being reported by USDA NASS. This research specifically contributes to the literature by testing if farm consolidation has resulted in yield not being reported for a county. This could generate important discussion about USDA NASS relaxing their criteria for reporting yields.

Introduction

The United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) is a government agency primarily focused on collecting and publishing agricultural statistics. One of the most utilized statistics collected by USDA NASS are annual county crop yield averages. These yield estimates are frequently used by numerous groups such as government agencies, researchers, market analysts, and producers. Thus, they provide much needed information about United States food and fiber production that is frequently used for policy development, farm payment calculations, decision making, and model impact of several factors on production.

USDA NASS yields are collected as a part of an annual survey administered by NASS, which includes all states except Alaska and Hawaii. The county yield estimates are partially determined by self-reported average field-based estimates from producers. The survey, which starts in November and ends mid-January, is collected through mail, phone interviews, in person interviews, and electronically via email with most responses collected through phone calls. Statisticians check the results to analyze for outliers before being inserted into a computer system to analyze for outliers again. Then after correcting the outliers, it is summarized by county and then released.

However, over the last 15 years, there has been a decline in survey responses and the number of county yields being reported, which has resulted in growing concerns about yield accuracy. Johansson et al. (2017) found that response rates for the NASS yield survey in 1992 were around 80 to 85 percent and in 2016 response rates declined to 55 to 70 percent response rates. Also, USDA NASS cannot release results from the Yield

Survey if they do not have enough responses. NASS can only publish a county yield estimate if at least 30 producer yield surveys are returned for that county or at least three responses are received from producers representing a minimum of 25% of the total county acreage (Schnepf 2017).

Researchers have noted that the decline in survey responses can introduce bias or error to United States crop yields, which has major implications for policy analysis and decision making. Seventy percent of empirical policy related articles published in 2011 and 2012 in the *American Journal of Agricultural Economics* relied on publicly available data, most of which is collected by the USDA (Council on Food, Agricultural and Resource Economics 2016), demonstrating how important these data are to published research. Extension services across the United States utilize NASS yields as a base to help calculate expected farm revenue when building budgets and educating producers on various programs their states.

Additionally, multiple government agencies rely on NASS surveys for various payment calculations and policy recommendations. For example, the USDA utilized NASS yields in calculating ARC-CO payments. The average ARC-CO program payment between 2014-2017 was $54.18/acre for corn and $32.18 for soybeans (Ishee 2020). However, when there are lower response rates to NASS yield surveys, the county-level yield used in these calculations may not accurately reflect the actual crop productivity within that county. Consequently, the resulting Agriculture Risk Coverage-CO (ARC-CO) payments can either be higher or lower than they would be with more precise yield data that aligns closely with the actual production in each county.

Rejesus, Coble, and Knight (2010) found that when reference yields for RMA were based on NASS data, it was not a true representation of yield information for the producers enrolled in the Federal Crop Insurance Program. As these yields were not updated the problem increased due to technology in the agricultural sector improving (Rejesus, Coble, and Knight 2010). As a result of the study, it was recommended to use a reference yield calculation based on RMA yield (Rejesus, Coble, and Knight 2010). Similarly, Li et al. (2020) examined the variability and reliability of a yield estimator based on NASS yields to RMA data for corn, soybeans, and wheat. The paper used NASS and RMA yields from the years 1991-2015 to examine the feasibility of using RMA yields, rather than NASS yields to calculate ARC-CO payments. They found slight but no major difference between the two yield values. RMA yields were typically higher than NASS yields, as well as that the result of using RMA yield for ARC-CO payments results in less variability between counties within an area.

In a similar study, Ishee (2020), analyzed three different methods of recovering missing corn and soybean yield data. The three different methods observed were: two different predictive linear models with variations in weather variables, as well as the Ordinary Kriging model. Ordinary Kriging serves is utilized to estimate values using data in the neighboring areas of the estimation (Wackernagel). After determining that Ordinary Kriging was the most effective method, ARC-CO payments were calculated using the yields from the Ordinary Kriging method. The estimated ARC-CO payments were then compared to actual ARC-CO payments for corn and soybeans during 2014 and 2015, to determine the difference in payments based on the yield data source. Park, Harri, and

Coble (2022) investigated the problem of missing corn and winter wheat yield values in NASS and RMA data and proposed an innovative approach of estimating county-level densities where data was missing. They utilized the Bayesian hierarchical approach and found that their estimates proved to be more accurate and less effected even with data omission issues than the benchmark approach.

The objective in this chapter is to determine if a county landscape and average size farm impacts the likelihood of a NASS yield being reported. NASS survey yield data plays a significant role in agricultural markets as it serves as an indicator of crop production in the United States. This information allows agricultural market prices to adjust in response to current yield values. The increasing trend of missing yield data poses a challenge and can contribute to increased price volatility in agricultural markets. Consequently, market participants may face a higher level of financial uncertainty, potentially leading to less informed decision-making in the agricultural industry.

Data

Two datasets were developed and used for the study: one to assess the reporting likelihood of corn NASS yields and the other for soybean NASS yields, both covering the years 2011-2022. Data on county yield values for corn and soybeans across the United States were sourced from USDA NASS. Additionally, land cover information was obtained from USDA CroplandCROS for all states (USDA CDL). This land cover data from USDA CroplandCROS was then integrated with the USDA NASS yield data to form separate datasets for corn and soybeans. These datasets were refined by removing records of the observed crops that occupied less than a thousand acres in a county, as per

USDA CroplandCROS data. Subsequently, data on the average farm size for each county, crop, and year was incorporated, derived from USDA RMA Summary of Business Statistics data (USDA RMA).

From USDA CroplanCROS, percentages of a county landscape by different land cover classification were calculated for the classifications: soybeans, corn, cotton, pasture/hay (combined variable of the two classifications), developed (i.e., residential, commercial, or industrial uses), and forest. These percentages were calculated by taking the classifications and dividing them by the sum of all the classifications for a county. This calculation was done for each county by year. USDA RMA Summary of Business Statistics data were used to calculate a proxy of average farm size by dividing the total insured acres in a county for a given year by the number of insurance policies issued in that county during the same year, serving as an indirect measure of average farm size. A county was marked having a missing NASS yield value if the county had a thousand acres or more of the observed crop in a year according to USDA CroplandCROS and did not report a NASS yield that year.

Figure 2 displays the average corn acres planted by county from 2011-2022 for the United States from USDA CroplandCROS data. This figure shows that counties with greater corn acres are concentrated in the upper Midwest and Northern Plains. Figure 3 displays the average soybean acres planted by county from 2011-2022 for the United States. Soybean acres are concentrated along the Mississippi River, Northern Plains, and Midwest. Figure 4 shows the percentage of years where a NASS corn yield was reported by county from 2011-2022 for the United States. A visual inspection between Figure 2

and Figure 4 suggests there is a relationship in yield report rate and corn acres planted. Figure 5 shows the percentage of years where a NASS soybean yield was reported by county from 2011-2022 for the United States. The areas where counties have higher report rates are along the Mississippi River, Northern Plains, and Midwest is also where soybean planted acres are more concentrated.

Table 1 displays the summary statistics of the variables observed for corn and used in the model. The variables are the percent of the county that is in each crop and average farm size scaled. The summary stats indicate forest and pasture/hay provide the most land cover in counties used for the corn analysis. On average, corn acres cover about 12% of the land within a county and soybeans cover 11% of the land within a county. The average corn farm size was 134 acres. The summary statistics of the variables observed for soybeans are displayed in Table 2. The percentage of land cover in a county was similar for soybeans. Both corn and soybeans covered 13% of the land area for the soybean data. The average soybean farm size was 148 acres.

Methods

A logit model was utilized to determine how a county's landscape and the average size of corn and soybean farms influence the probability of a NASS yield being reported. A logit model was used because of the dataset's binary nature, which held two outcomes: either a NASS yield being reported, or a NASS yield not being reported. The findings from the logit model are expressed in terms of odds ratios, enabling an interpretation that quantifies how changes in a variable's value affect the odds of a NASS yield being reported, either increasing or decreasing these odds by a specified percentage

The equation is written below.

(3) Pr ($NASS_{ic} = 1$)

$$= F(\beta_0 + \beta_1 PB_{ic} + \beta_2 PC_{ic} + \beta_3 PCT_{ic} + \beta_4 PPH_{ic} + \beta_5 PD_{ic} + \beta_6 PF_{ic}$$

$$+ \beta_7 SAVGSIZE_{ic} + \beta_8 SAVGSIZE^2{}_{ic} + \sum_{y=1}^{S} \beta_i state_s) \sum_{y=1}^{12} \beta_i year_y)$$

where Pr ($NASS_{ic}$=1) if a NASS yield were reported for crop *c* in *i*th county, Pr ($NASS_{ic} = 0$) if a NASS yield were not reported for crop *c* in *i*th county, where PB_{ic} is the percentage of soybean acres in *i*th county and crop *c*; PC_{ic} is percentage of corn acres; PCT_{ic} is percentage of cotton acres; PPH_{ic} is percentage of pasture and hay acres in *i*th county; PD_{ic} is percentage of developed acres; PF_{ic} is percentage of forest acres in *i*th county; $SAVGSIZE_{ic}$ is average farm size proxy scaled by a thousand, $SAVGSIZE^2{}_{ic}$ is average farm size proxy squared scaled by a thousand, state and year fixed effects were included. Average marginal effects were calculated in R using the margins package.

Results

The results from the logistic regression analysis aimed to explore the impact of county landscape and average farm size on the likelihood of a NASS yield being reported are summarized in Table 3.

In the case of corn, the proportion of county landscape used for corn production exhibited a strong positive relationship ($p < .001$) with the likelihood of a NASS corn yield being reported. Additionally, the proportion of county landscape dedicated to corn production, along with percentages of landscape in cotton and pasture/hay, average farm size, and average farm size squared, all showed significant relationships ($p < .001$).

Similarly, in the soybean model, the proportion of county landscape allocated to soybean production had a significant relationship ($p < .001$). In both models, the percentage of the county that was developed was not significant for a NASS yield being reported. In both models, state and year fixed effects were incorporated into the model to control variations across different states and years. The average marginal effects are shown in Table 4. For example, a 1% increase in the percentage of the landscape in corn results in a 0.82% increase in the likelihood of a NASS corn yield. For the soybean model a 1% increase in the percentage of the landscape in soybeans results in a 0.86% increase in the likelihood of reporting a NASS soybean yield.

In both models, the proxy of average farm size and the proxy of average farm size squared were found to be significant, indicating a positive influence on the likelihood of reporting a NASS yield until a certain threshold. Figure 6 illustrates the predicted probability curve, showing that for corn, the likelihood starts decreasing after 228.33 acres, and for soybeans, after 253.13 acres. This could indicate farm consolidation could negatively impact the likelihood of NASS yields being reported. This would suggest that as farms continue to consolidate USDA NASS revisit their criteria of reporting yields in a county to avoid not reporting counties with a reportable level of acres but not enough farms.

Conclusion

This research shows the critical issue of declining response rates to NASS yield surveys, reduced reporting of NASS county yields, and its implications for agricultural research, policymaking, and market dynamics. By investigating the impact of county landscape

and average farm size on the likelihood of reporting NASS yields for corn and soybeans, this study contributes valuable insights to the existing literature.

The results of the logit model, indicate that a county's landscape impacts the likelihood of a NASS yield being reported. Counties that have a higher percentage of their landscape in agricultural production are more likely to have a NASS yield reported. It also indicates that average farm size plays a role in the likelihood of a NASS yield being reported as well. Counties that have a high number of acres concentrated in a few farms may not have a NASS yield reported due to large farm size resulting in the county having fewer than 30 producers or 25% of harvested acres responding to the survey. This will continue to be an issue if we continue to see farm consolidation across the United States. Having a county that produces many acres for either crop or not reporting a yield could result in reporting bias. Reporting bias could occur because gaps in a county NASS yield history impacting historical county average. One county not receiving enough survey responses could also impact state and national averages as well.

Addressing the challenges posed by the increase in non-reported yields is crucial for policymakers, researchers, and market participants to make informed decisions and foster a more resilient and sustainable agricultural sector. Continued efforts to improve data collection methods and enhance the accuracy of yield reporting are imperative to ensure the reliability of NASS data.

CHAPTER II
NASS YIELD ALTERTATIVES

Abstract

Having comprehensive and dependable county crop yield averages is crucial for stakeholders, including government agencies, researchers, market analysts, and producers. Reliable data on the United States food and fiber production is essential for policymaking, decision-making, and accurately modeling the effects of weather and other variables on production. However, over the course of the last 30 years NASS county yields have experienced a decline in response rates. The trend of declining response rates to the NASS survey can bias yield estimates used in research or cause lack of data to calculate payments. The objective of this research is to determine if alternative yield values can be used in substitution of USDA NASS yields or to fill in when there is no NASS yield available. The results of this research will benefit the various groups that utilize USDA NASS yields.

Introduction

Prior to the Agricultural Improvement Act of 2018, NASS yields were the primary yield data used to calculate payments for the Agricultural Risk Coverage-County (ARC-CO) program. These payments are made when the actual county crop revenue is less than the ARC-CO guarantees (USDA Farm Service Agency [FSA]). The average ARC-CO program payment between 2014-2017 was $54.18/acre for corn and $32.18/acre for soybeans (Ishee 2020). As discussed in the previous chapter, response rates for the NASS Acreage and Production Surveys declined between the years 1992-2016 with the trend accelerating beginning in 2012 (Johansson et al. 2017; Schbnepf 2017). The decline in responses, trend of larger farms, and the criteria for reporting a county-level yield has resulted in an increase in non-reported county yields. When there are yields not being reported, the county-level yield used in these calculations may not accurately reflect the actual crop productiion within that county. Consequently, the resulting ARC-CO payments can either be higher or lower than they would be with more precise yield data that aligns closely with the actual production in each county.

Therefore, there is a need for novel ways to estimate county level yields when USDA NASS yields are not reported. Researchers have looked to fill the data gaps using other data sources like USDA RMA Summary of Business statistics to estimate yields. The RMA compiles their yearly Summary of Business report from several sources including, but not limited to, policyholder data, participating insurance companies, and other USDA agencies. In a recent report done by the National Academy of Sciences they recommend that NASS should collaborate with RMA to be able to analyze and compare

RMA data to NASS (National Academies of Sciences, Engineering, and Medicine 2017). Other data sources recommended in the report include data from precision agricultural databases and the USDA FSA.

Researchers have tried different methods to fill in the gaps of missing yield data. Rejesus, Coble, and Knight (2010) found that when reference yields for FSA payments were based using the NASS yields, it was not a true representation of yield information for corn, cotton, soybean, and wheat producers enrolled in the federal crop insurance programs. Another study found no major differences between NASS yields and RMA yields corn, soybeans, and wheat. The study stated RMA yields were typically higher than NASS and resulted in less variability in ARC-CO payments between counties in an area (Li e al. 2020).

Ishee (2020), analyzed three different methods of recovering missing corn and soybean yield data. The three different methods observed were: two different predictive linear models with variations in weather variables, as well as the Ordinary Kriging model. After determining that Ordinary Kriging was the most effective method, ARC-CO payments were calculated using the yields from the Ordinary Kriging method. The estimated ARC-CO payments were then compared to actual ARC-CO payments for corn and soybeans during 2014 and 2015, to determine the difference in payments based on the yield data source. Park, Harri, and Coble (2022) extended this work by using a Bayesian hierarchical approach to find missing corn and winter wheat yields and reported that their estimates proved to be more accurate and less effected even with data omission issues than the benchmark approaches.

The goal of this research is to observe how closely aligned alternative corn and soybean yield data sets match with reported NASS yields to see if they can be used to fill in the gaps of the Agricultural Yield Survey. We use county level data in Tennessee from USDA NASS, USDA FSA, and USDA Summary of Business for the years 2018-2022. The methods we used to evaluate if they align are hot spot analysis, hot spot comparison, and multivariate clustering in ArcGIS Pro. A hot spot analysis identifies spatially where hot (high values) and cold spot (low values) of a data set are. A hot spot comparison is a comparison of the location of the hot and cold spots between two different hot spot analysis. Identifying areas that are higher or lower in similarity value depending on the overlay of the hot and cold spots of the two hot spot analysis. Multivariate clusters cluster areas from two data set based on similar features with the aim of minimizing variance withing a cluster and maximizing the difference between the clusters.

These analyses will determine how the different yield values compare, not only in value but also spatially. Having spatial components to the analyses will strengthen the results to ensure that yield values are similar on individual county levels as well as a state level geographically. These results will help Tennessee producers and Extension educators find alternative ways to estimate county yields and could serve as a model to other states on finding alternative yield estimates when NASS yields are not reported.

Data

Tennessee county level corn and soybean yield data was collected from USDA NASS and FSA program payment data for the years 2018-2022 for the state of Tennessee. A novel approach of using the USDA RMA Summary of Business Statics and price

discovery data to estimate county level corn and soybean yields for the years 2018-2022 in Tennessee. APH yields are not made available to the public however, the USDA RMA Summary of Business database provides total liability, total acres insured, and coverage level, which were used to estimate a county-level average APH for all producers and unit structures. Total liability is equal to the guaranteed amount from each policy. Therefore, total liability can be expressed as

(1) $TL_{ikcm} = (p^{G}_{ism} y^{APH}_{ikcm} \delta_k) \times TA_{ikcm}$

where TL_{ikcm} is the total liability for the policy for county c ($c = 1,\ldots,C$); p^{G}_{ic} is the insured price and δ_k is the coverage level; and TA_{ikcm} is the total acres insured by the policy. APH yields (y^{APH}_{ikcm}) were found by rearranging (1) to

(2) $y^{APH}_{ikcm} = (TL_{ikcm}/TA_{ikcm} p^{G}_{ism} \delta_k)$.

A five-year Olympic average was calculated for the three different yield values for the corn and soybeans. Olympic averages were calculated by taking the sum of the values from 2018-2022 then subtracting the highest and lowest values from the period. After subtracting the minimum and maximum values the sum is then divided by three. Olympic averages for the observed counties can be found in Table 5 and Table 6. Olympic averages were chosen to be observed as they are the value that is used in ARC-CO payments.

Summary statistics by yield source for the corn Olympic averages can be found in Table 7. On average for corn FSA had the highest yield at 180 bu/acre and reported the highest value for the data set at 207 bu/acre. Calculated county APH yield derived from Summary of Business statistics reported the lowest value at 114 bu/acre and had the

lowest average of 146 bu/acre. NASS Olympic average yields were180 bu/acre and had a minimum value of 119 bu/acre and a maximum value of 187 bu/acre. NASS had the least amount of variance of the three yield values with a standard deviation of 8.56.

Table 8 displays the summary statistics by yield source for the soybean Olympic averages. NASS on average had the highest yield with an average value of 63.62 bu/acre with the minimum value of 55 bu/acre and maximum value of 70 bu/acre. Soybean APH yields derived from Summary of Business statistics averaged the lowest at 41 bu/acre. APH yields had the smallest standard deviation of 3.53 reporting minimum value of 37 bu/acre and maximum value of 49 bu/acre. FSA yields had the largest amount of variance with a standard deviation of 5.49. FSA yields averaged 49 bu/acre with a low value of 33 bu/acre and a high value of 60 bu/acre.

Methods

ArcGIS Pro was utilized to conduct a series of different analyses of observing the Olympic averages of the three sources. The first step was to conduct a hot spot analysis to observe the hot and cold spots of the three yield values for crops corn and soybeans. A hot spot analysis is a spatial analysis that is used to identify statistically significant cluster patterns of high (hot spots) or low (cold spots) yield values within the data. After the hot spot analysis for the yields, we conducted a pairwise comparison hot spot analysis where the NASS yields were compared to the FSA yields, the NASS yields were compared to APH yields, and the FSA yields were compared to the APH yields for both corn and soybeans.

Then, a multivariate cluster analysis was conducted to cluster the counties in the datasets into four groups based on similarities between two yield values at a time. A pairwise comparison was also done between the yield sources for this analysis. A multivariate cluster is completed in ArcGIS to analyze and group spatial patterns between variables. Clusters are assigned based on similar features with the aim of minimizing variance within a cluster and maximizing the difference between the clusters. This allows natural patterns to be recognized.

Results

Corn

Figure 7 shows the hot spot analysis for NASS corn Olympic average yields. The results of the analysis classify two counties as having a cold spot while nine counties are a hot spot. Dyer County is indicated as being a cold spot with 95% confidence and Obion County was indicated as being a cold spot with 90% confidence. These two counties indicate that at the 95% and 90% confidence level that the NASS corn Olympic averages are significantly lower than the expected values based on the spatial distribution. With the nine counties that are reported as hot spots, their NASS Olympic yield averages are significantly higher than expected based on the spatial distribution of the yields. Figure 8 shows the hot spot analysis for FSA corn Olympic yield averages. The figure indicates Dyer County being a cold spot at the 90% confidence level in the western region of the state. There are three counties touching one another in the southern part of Middle Tennessee that indicates a hot spot at the 90%, 95%, and 95% confidence levels.

Results of the APH corn Olympic yield averages hot spot analysis, shown in Figure 9, indicate two different clustering of hot spots. The larger cluster consisting of twelve counties in the southern part of middle Tennessee and the smaller cluster consisting of four counties in the northern part of middle Tennessee. Eight of the sixteen counties indicated being a hot spot with 99% confidence level. There are five counties on the western edge of the state that indicate being a cold spot, four out of the five at the 90% confidence level, one at the 95% confidence level. There is also one county located in East Tennessee that indicates being a cold spot at the 95% confidence level.

Between the three hot spot comparison for corn yields, FSA yield had the least amount of hot and cold spots across the state. Having the least amount of hot or cold spots across the states indicates that it has the least amount of spatially significant outliers in the data set. While APH yields had the highest amount of hot and cold spots, indicating that spatially, and in value, there were more spots that were significantly higher or lower than the values spatially around it.

Figure 10 displays the results of the hot spot comparison of corn yields between NASS yields and FSA yields Olympic averages. The counties displayed in dark orange have a lower similarity value. Having minimal spatial agreement indicates that the distribution of hot and cold spots for these two yields do not align well with one another spatially. When examining the hot spot analyses of corn NASS yields and FSA yields hot spot analyses individually, we notice that FSA yields do not have the concentration of counties indicating being a hot spot in the southern middle Tennessee area that NASS

indicates. The difference in the spot analyses for these two years results in these counties having a lower similarity value.

Results of the corn NASS yields and APH yields Olympic average hot spot comparison is displayed in Figure 11. The figure indicate that the spatial agreement between NASS yields and APH yields in 10 out of the 31 counties is below 0.5, indicating low agreement. The counties that report being between less than 0.2 in similar value are Maury, Marshall, Bedford, and Rutheford. There are 14 out of the 31 counties that have similar values 0.9 or above with a large concentration of them being in the northern part of West Tennessee.

Out of the three yield hot spot comparisons for corn FSA yields and APH yields have the largest number of counties that have similar values below 0.2, displayed in Figure 12. The counties that we observed as being a hot spot in APH corn yield hot spot (Figure 9) are the same counties that are indicating as having low similarity values. The western part of Tennessee is the area of the state where most counties with similarity value above 0.5 are located.

Results from the clustering of corn NASS yields and FSA yields displayed in Figure 13 indicate that within one cluster there are 14 counties with a large portion of these counties being in West Tennessee. Another cluster has nine counties all in Middle Tennessee. One cluster has one county in it, being in Middle Tennessee. The fourth cluster contains seven counties in all three regions of the state. Figure 14 shows the clusters from NASS and APH corn yields. This analysis resulted in clusters with

fourteen, eight, six, and three counties. Corn APH and FSA clusters shown in Figure 15 contain clusters with seven, eleven, eleven, and two counties.

Soybeans

NASS soybean Olympic yield averages hot spot indicates a six-county cluster in West Tennessee that is a cold spot with three counties being a hot spot but not adjacent to each other (Figure 16). Five of the six counties in the cold spot cluster are at the 95% confidence level and one being at the 90% confidence level. The three counties being a hot spot are at the 90%, 95%, and 99% confidence level. Figure 17 shows three counties are a cold spot, and two counties are a hot spot for FSA soybean Olympic yield averages in the state of Tennessee. The cold spot counties are in West Tennessee at the 90%, 90%, and 95% confidence level. The hot spots located in Middle Tennessee are at the 90% confidence level, with the other at the 95% confidence level in East Tennessee.

As a result of FSA yields being similar in value spatially the hot spot analysis for FSA yields displayed the lowest number of hot or cold spots among the three yield values. This indicates that FSA yields represent the area well spatially, not having high and low values right beside one another.

Results for the APH soybean Olympic yield averages hot spot are displayed in Figure 18. The results of the hot spot indicate a four-county cluster in the western part of the state being a cold spot. Three of the four counties are a cold spot at the 95% confidence level with one being at the 99% confidence level. There are four counties that are hot spots at the 90%, 95%, 95%, and 99% confidence level located in Middle and East Tennessee.

Figure 19 shows the similarity values between the hot spot analysis between NASS and FSA soybean yields. Many counties are classified as being 90% or greater in similar value. However, in West Tennessee there is a cluster of counties that are in the range of 50% to 80% in similarity value. Overall, the NASS yield and FSA yield soybean hot spots do not indicate any strong differences. NASS yield and APH soybean Olympic yield averages hot spot comparison can be found in Figure 20. In this figure, we observe all counties having similarity values of 0.5 and above. There is one county in the 0.5 to 0.6 similarity value range and in West Tennessee. While four counties are in the 0.7 to 0.8 similarity value range and are in West and Middle Tennessee. The remaining 17 counties are 0.9 or greater in similarity values. Figure 21 shows that in the soybean FSA yield and APH yield Olympic averages hot spot comparison there are only six counties that are not in the 0.9 or greater similarity value range. There are five counties that are in the 0.7 to 0.8 similarity value range, three being in West Tennessee, one in Middle Tennessee, and one in East Tennessee. In Middle Tennessee there is one county that is in 0.5 to 0.6 similarity value range.

The results from the multivariate cluster of NASS and FSA soybean yields are displayed in Figure 22. One cluster consists of 14 counties while another one consists of six counties and two clusters containing one county a piece. Figure 23 displays NASS and APH clusters. The clusters results have three clusters that contain five counties and a cluster containing five counties. NASS and FSA soybean clusters are shown in Figure 24. Two clusters contain just one county a piece while the two clusters contain fourteen, and six counties a piece.

Comparison of Corn and Soybeans

NASS soybean yields had fewer hot spots and more cold spots than NASS corn yields. When looking at the hot spot comparison results of NASS yields and FSA yields versus NASS yields and APH yields, soybeans had higher similarity values compared to corn. NASS corn Olympic yields had the lowest amount of deviation between the three yield values while in soybeans it was in the middle. For corn yields, the NASS Olympic average was the middle value while in soybeans it was the highest.

FSA yields for both corn and soybeans had very few hot or cold spots. Comparing the hot spot comparison results of NASS yields and FSA yields as well as FSA yields and APH yields between crops, corn had more counties with lower similarity values than soybeans. FSA Olympic averages have the highest amount of deviation of the three yields for both corn and soybeans.

The hot spot analysis ran on soybean APH yields had fewer hot spots than the corn APH yield hot spot analysis. The corn hot spots had more cold spots than soybeans, however soybeans were at a higher confidence level than corn. For both corn and soybeans at the state level APH Olympic averages were lower than NASS Olympic averages.

NASS Yield Substitution

With NASS yield data set continuing to have an increase in yield not being reported over time, it is important to determine other county yield values to substitute in for NASS county yields. Based on the analysis, FSA appears to be the best alternative for NASS county corn and soybean yields. FSA yields hot and cold spots aligned closer to NASS

yields hot spots for both corn and soybeans. This is displayed in the NASS yield FSA yield Olympic average hot spot comparison for both corn and soybeans. Most counties in these two comparisons have similar values 0.9 or above. FSA corn yields are on average around 10 bu/acre higher than NASS yields. For soybeans, FSA yields are on average 14 bu/acre lower than NASS yields. These values align closer with NASS yield than APH yield averages for both corn and soybeans. The combination of FSA yields and NASS yields have stronger spatial alignment as well as having their value on averages align closer indicates that FSA is the better yield alternative for NASS yields than APH yields.

Next Steps

Another county level yield that is publicly available are University of Tennessee Corn and Soybean variety trials. These variety trails occur annual at various Research and Education Center (REC) across the state as well as county standard test (CST) locations. The counties for REC locations include Roberston, Greene, Knox, Madison, and Gibson County for both corn and soybeans. For the years 2018-2022 there were 10 counties that had a CST location as well as a NASS Olympic average for the period. There are 23 counties for soybeans that meet the same requirement. Not every CST location has a variety trial done every year because they are done in partnership between a University of Tennessee Extension and local producers.

Due to not every CST having a variety trial yield every year simple averages of the variety trial data were used to compare against NASS corn and soybean Olympic Averages. Some counties had more than one observation each year. The entries were grouped by county by year and averaged to calculate a yearly average for each county.

The simple averages of all the years a county reported a yield were taken to calculate a county average. These averages are being compared to the NASS Olympic Averages.

The corn and soybean variety trial yields by county can be found in Tables 9 and 10. Tables 11 display the summary statistics of the variety trial yields for the state of Tennessee from the years 2018-2022. Tennessee corn variety trial yields range from 121 bu/acre and 247 bu/acre. Compared to the average of the corn Olympic averages for the state of Tennessee of 169 bu/acre Variety Trial yields were higher with an average of 184 bu/acre. Tennessee soybean variety trial yields range from 26 bu/acre and 62 bu/acre. Tennessee NASS soybean Olympic averages are 1.4 times greater than Tennessee variety trial data.

Variety trial yields as a percentage of NASS yields were mapped for the counties that hosted a variety trial as well as reported enough NASS yield data to calculate an Olympic average for the years 2018-2022. These results are displayed in Figure 25 for corn and Figure 26 for soybeans.

Conclusion

This study uses different spatial analysis to identify viable alternatives to county-level yield data in response to the increasing number of missing USDA NASS County yields. Observing data from FSA, RMA Summary of Business, and the University of Tennessee Corn and Soybean Variety Trials to evaluate their potential as replacements or supplements to traditional NASS yield figures. Using a variety of spatial analysis, including hot spot analysis, hot spot comparisons, and multivariate clustering, has allowed a further understanding between the alternative data sets and the yields reported

by NASS. Among these alternatives, FSA appears to be the best alternative yield value for NASS county corn and soybean yields due to its geographical distribution and average yields for corn and soybeans. Yields from the University of Tennessee Variety Trials were also observed as a NASS yield alternative. Despite the inconsistency in annual data availability across all counties, this data could be used as a NASS yield alternative.

CONCLUSION

With the issue of missing NASS county corn and soybean yields becoming more problematic over time this research answers relevant questions and provides insight into potential solutions. In the first chapter, the county's landscape and average farm size were examined to determine factors that impact the likelihood of a NASS yield being reported. The results of the logit model indicated that a county's landscape and average farm size can impact the likelihood of a NASS yield being reported.

Counties that have a larger percentage of their land in agricultural production are more likely to report a NASS yield. Average farm size was also found to impact the likelihood of a NASS yield being reported. In counties that have an average farm size greater than 250 acres the likelihood of a NASS yield being reported starts to decrease due to NASS reporting requirements. NASS requires 30 producers or at least 25% of harvested acres to be reported to release a county yield to the public. The trend of farm size continuing to grow and response rates to NASS survey dropping poses a risk of reporting bias increasing.

Finding a yield alternative to help assist in the problem of NASS yields is crucial to provide accurate county level yield information for policymakers, researchers, and market shareholders to make well-informed decisions. A further understanding of how USDA FSA and APH yields derived from USDA RMA Summary of Business report compare to NASS yield data came from using spatial analysis in the second chapter. From observing three different yield values at the state and county level it appears FSA yields are the most suitable alternative to NASS yields.

REFERENCES

Agricultural Act of 2014. Pub. L. No. 113-79, 128 Stat. 649 (2014).

Agriculture Improvement Act of 2018. Pub. L. No. 115-334, 132 Stat. 4490 (2018).

Council on Food, Agricultural and Resource Economics (C-FARE). (2016). From Farm Income to Food Consumption: Valuing USDA Data Products. Updated July, 2016.

Ishee, Z. S. (2020). Recovering missing yield values: Ramifications for the USDA's ARC-CO program. (Master's thesis). Mississippi State University. Retrieved from https://thesis.msstate.edu/handle/11668/16817.

Johansson, R., Effland, A., & Coble, K. (2017). Falling response rates to USDA crop surveys: Why it matters. American Journal of Agricultural Economics, 99(2), 432-438. https://doi.org/10.1093/ajae/aaw081

Li, X., Guo, Z., Huang, Y., & Zheng, X. (2020). Comparing survey-based and program-based yield data: Implications for the U.S. Agricultural Risk Coverage-County program. Geneva Papers on Risk and Insurance Issues and Practice, 45(1), 184-202. https://doi.org/10.1057/s41288-019-00146-4

National Academies of Sciences, Engineering, and Medicine. 2017. Improving Crop Estimates by Integrating Multiple Data Sources. Washington, DC: The National Academies Press. https://doi.org/10.17226/24892.

National Agricultural Statistics Service. (2023). Yield Forecasting Program of NASS: Methodology and Data Quality. Retrieved from https://www.nass.usda.gov/Publications/Methodology_and_Data_Quality/Advanced_Topics/Yield%20Forecasting%20Program%20of%20NASS_2023.pdf

Park, E., Harri, A., & Coble, K. H. (2022). Estimating crop yield densities for counties with missing data. Journal of Agricultural and Resource Economics, 47(3), 634-S10. https://doi.org/10.35423/jare2021-042

Rejesus, R., Goodwin, B., Coble, K., & Knight, T. (2010). Evaluation of the reference yield calculation method in crop insurance. Agricultural Finance Review, 70, 427-445. https://doi.org/10.1108/00021461011088530.

Schnepf, R. (2017, April 13). NASS and U.S. Crop Production Forecasts: Methods and Issues. Specialist in Agricultural Policy, Congressional Research Service.

US Department of Agriculture Economic Research Service. (n.d.). Crop Insurance at a Glance: Cost of Federal Crop Insurance. Retrieved April 5, 2023, from https://www.ers.usda.gov/topics/farm-practices-management/risk-management/crop-insurance-at-a-glance/#fcipcost

USDA, Farm Service Agency. ARC/PLC Program Data. https://www.fsa.usda.gov/programs-and-services/arcplc_program/arcplc-program-data/index

United States Department of Agriculture. Cropland Data Layer. Agricultural Research Service. https://nassgeodata.gmu.edu/CropScape/

US Department of Agriculture Risk Management Agency. Summary of Business. https://www.rma.usda.gov/

USDA, NASS. Quick Stats. https://quickstats.nass.usda.gov/

USDA National Agricultural Statistics Service, 2002 Census of Agriculture. Complete data available at www.nass.usda.gov/AgCensus

USDA National Agricultural Statistics Service, 2022 Census of Agriculture. Complete data available at www.nass.usda.gov/AgCensus

Wackernagel, H. (2003). Ordinary Kriging. In: Multivariate Geostatistics. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-05294-5_11

Zulauf, C., Schnitkey, G., Paulson, N., & Coppess, J. (2017). Comparing NASS and RMA county yields for corn. Farmdoc Daily, (7):202. Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

APPENDIX

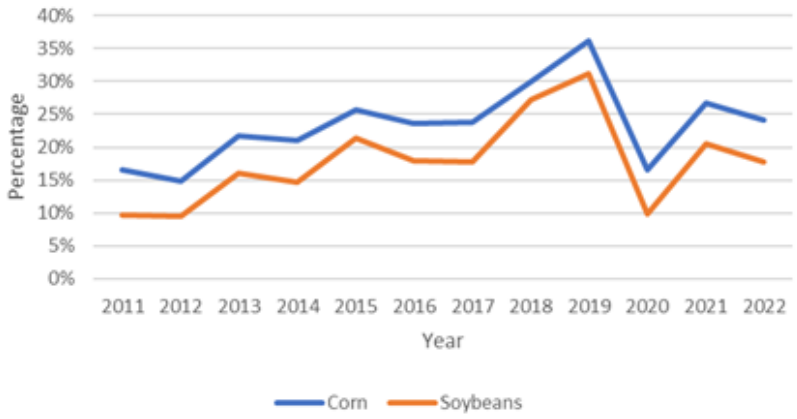


Figure 1. Percentage of Missing NASS County Corn and Soybean Yields, 2011-2022

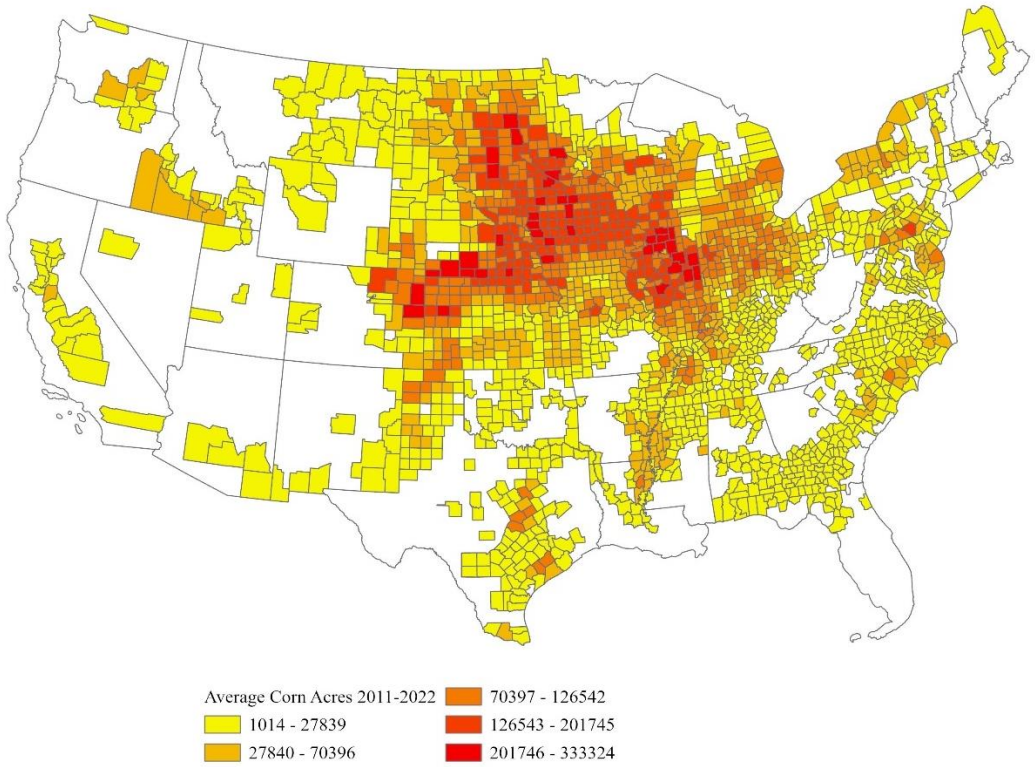


Figure 2. Average Planted Corn Acres, 2011-2022 USDA CroplandCROS

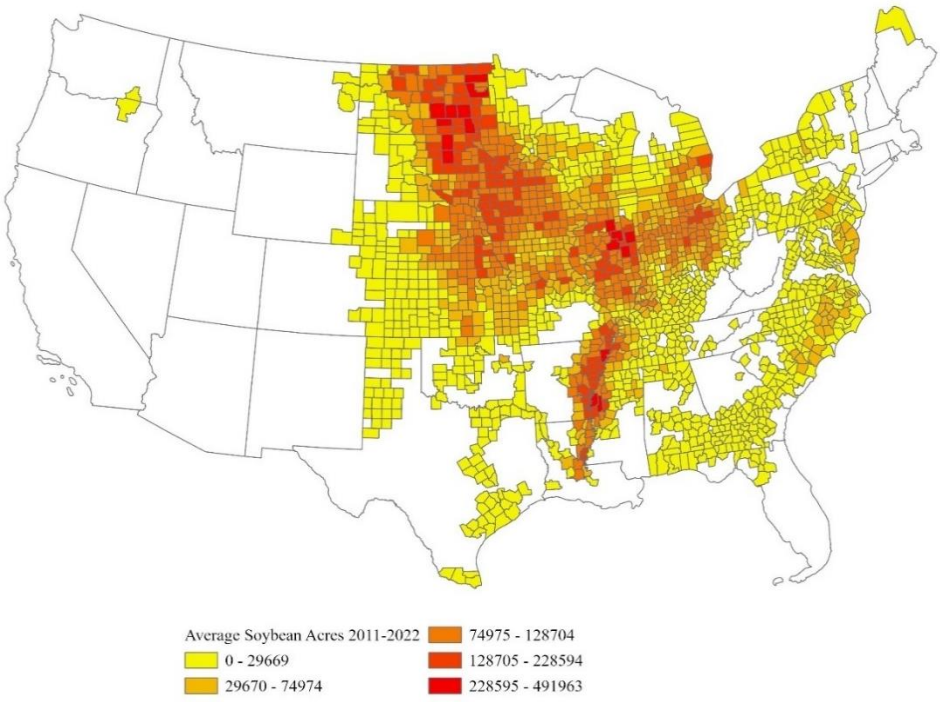


Figure 3. Average Planted Soybean Acres, 2011-2022 USDA CroplandCROS

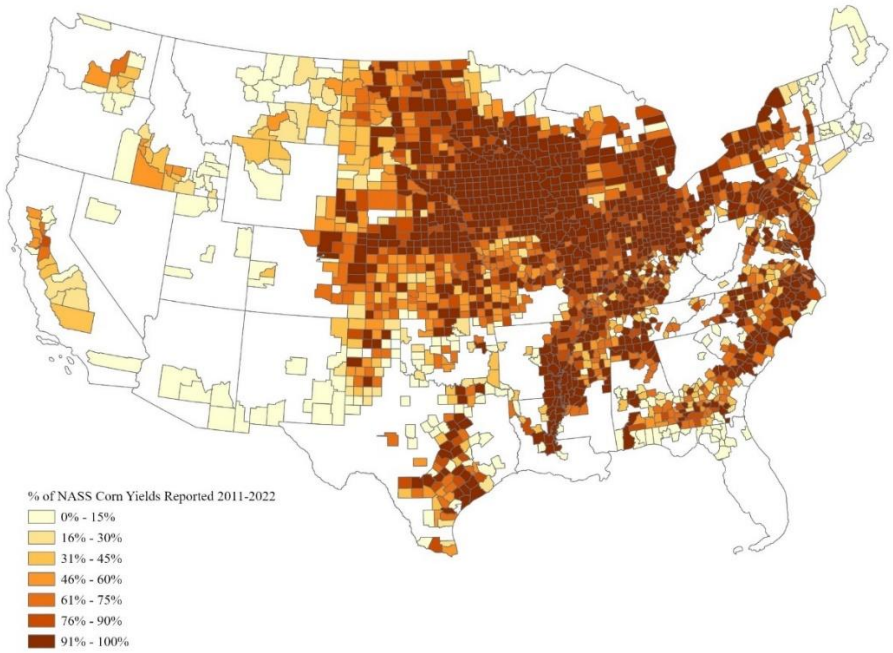


Figure 4. Percentage of NASS Corn Yields Reported, 2011-2022

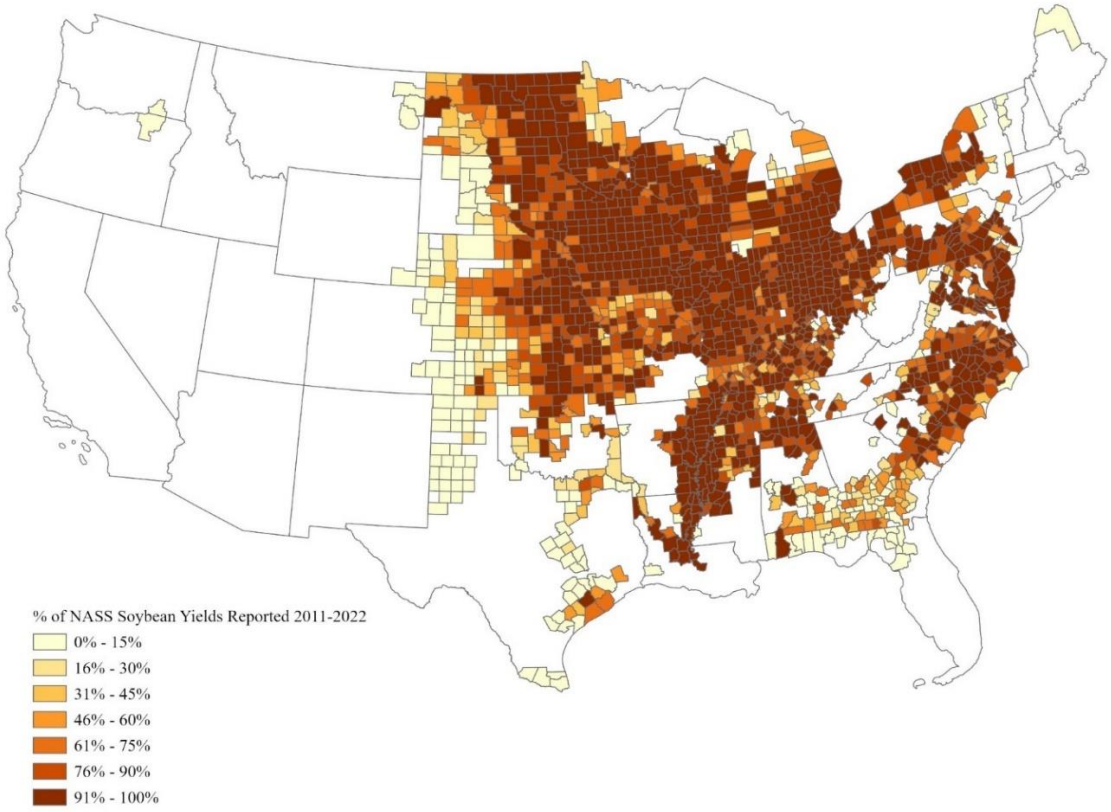


Figure 5. Percentage of NASS Soybean Yields Reported, 2011-2022

Table 1. NASS Corn Yield Reporting by County Landscape Logit Model Variable Summary Stats

	Avg	St. Dev	Min	Max
%Beans	0.11	0.13	0.00	0.68
%Corn	0.12	0.14	0.00	0.63
%Cotton	0.01	0.05	0.00	0.71
%PastureHay	0.22	0.19	0.00	0.97
%Developed	0.08	0.08	0.00	0.85
%Forest	0.23	0.22	0.00	0.83
SAvgSize	0.13	0.08	0.00	1.17
SAvgSize^2	23.45	36.99	0.00	1370.24

%Variable is equal to the percentage of a county's total landscape that is covered by that variable.

Table 2. NASS Soybean Yield Reporting by County Landscape Logit Model Variable Summary Stats

	Avg	St. Dev	Min	Max
%Beans	0.13	0.13	0.00	0.68
%Corn	0.13	0.14	0.00	0.63
%Cotton	0.01	0.04	0.00	0.70
%PastureHay	0.20	0.17	0.00	0.97
%Developed	0.08	0.08	0.00	0.85
%Forest	0.24	0.22	0.00	0.83
SAvgSize	0.14	0.09	0.00	1.33
SAvgSize^2	27.09	44.83	0.00	1780.22

%Variable is equal to the percentage of a county's total landscape that is covered by that variable.

Table 3. Logit Model Results for NASS Yield Reporting by County Landscape

	Corn	Soybeans
Variable	Estimate	Estimate
(Intercept)	1.013***	1.31***
%Beans	2.189***	7.486***
%Corn	6.144***	1.341**
%Cotton	1.878***	3.886**
%PastureHay	0.918***	0.5*
%Developed	-0.013	0.511
%Forest	-0.065	-0.079
SAvgSize	5.48***	5.569***
SAvgSize^2	-0.012***	-0.011***

*,**, *** represent significance at the 10%, 5%, and 1% levels, respectively. %Variable is equal to the percentage of a county's total landscape that is covered by that variable.

Table 4. Average Marginal Effects of NASS Yield Reporting

	Corn	Soybeans
Variable	Estimate	Estimate
%Beans	0.2904***	0.8602***
%Corn	0.8149***	0.1613**
%Cotton	0.2492***	0.5548**
%PastureHay	0.1217***	0.0629*
%Developed	-0.017	0.030
%Forest	-0.0086	-0.015
SAvgSize	0.7269***	0.00007***
SAvgSize^2	-0.016***	-0.0013***

*,**, *** represent significance at the 10%, 5%, and 1% levels, respectively. %Variable is equal to the percentage of a county's total landscape that is covered by that variable.

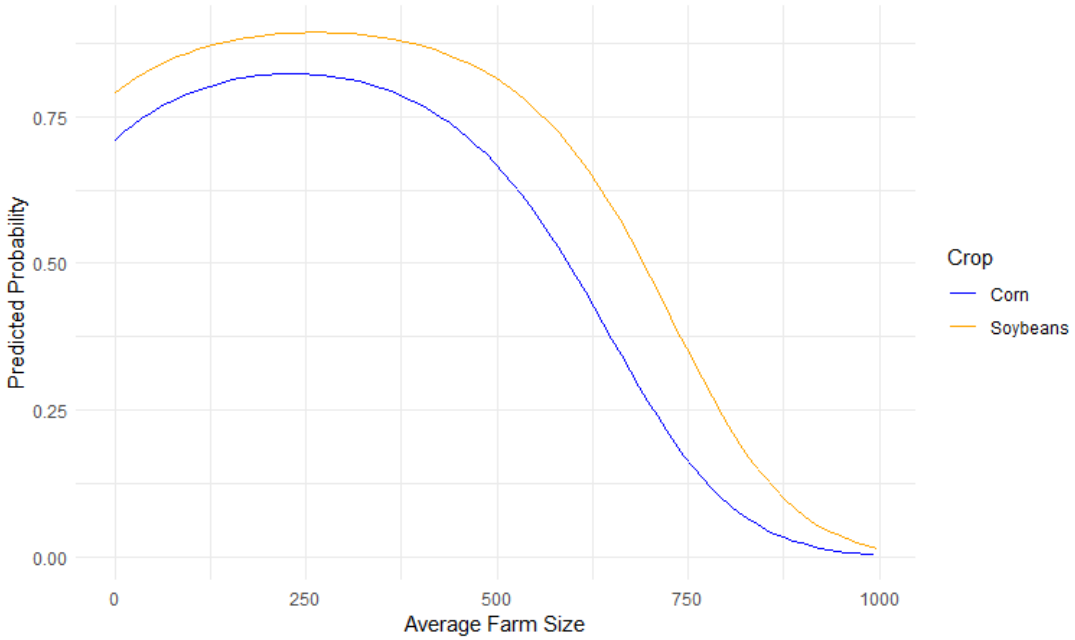


Figure 6. Predicted Probability of a NASS Yield Being Reported by Farm Size

Table 5. Tennessee Corn Olympic Average Yields by County, 2018-2022

County	NASS	FSA	APH
Bedford	156.90	174.80	150.87
Bledsoe	161.10	153.68	137.28
Carroll	163.70	170.61	134.12
Coffee	184.90	201.28	170.44
Crockett	169.40	183.99	134.31
Decatur	162.70	145.52	126.08
Dyer	170.93	180.18	139.99
Fayette	163.30	170.18	134.01
Franklin	178.37	195.46	158.72
Gibson	167.83	175.24	141.68
Giles	187.87	207.17	169.48
Greene	155.37	161.70	114.41
Henry	170.87	190.04	161.45
Lauderdale	166.80	174.79	132.79
Lawrence	173.03	202.56	169.63
Lincoln	181.40	194.14	151.47
Macon	163.00	156.58	137.11
Madison	173.33	193.38	144.02
Marshall	166.03	188.00	142.83
Maury	173.70	190.67	158.62
McMinn	167.13	172.32	128.82
Montgomery	177.43	187.45	164.79
Obion	170.40	190.24	160.16
Robertson	181.97	193.33	161.27
Rutherford	166.73	184.93	161.49
Stewart	150.30	119.33	122.94
Sumner	171.20	181.93	152.32
Tipton	164.03	181.76	133.86
Warren	180.03	193.10	144.90
Weakley	164.07	180.60	142.98
White	176.03	186.37	163.38

Table 6. Tennessee Soybean Olympic Average Yields by County, 2018-2022

County	NASS	FSA	APH
Bledsoe*	69.40	59.50	48.82
Cannon	70.53	60.32	49.84
Carroll	61.00	46.80	37.03
Chester	61.60	47.91	39.97
Decatur	55.93	33.70	37.21
Dyer	61.13	50.65	40.63
Fayette	62.67	44.25	37.12
Gibson	60.90	47.96	42.04
Giles	66.47	53.68	45.33
Haywood	66.33	51.57	40.71
Henderson	57.23	47.93	39.73
Lauderdale	61.47	51.00	41.33
Lincoln	66.07	53.40	41.84
Madison	63.53	50.67	38.69
Marshall	62.17	50.43	40.86
Montgomery	63.60	50.66	42.78
Obion	66.00	54.03	45.24
Sumner	64.63	44.30	39.97
Tipton	63.20	51.46	40.30
Weakley	65.50	50.29	39.59
White	69.90	49.80	47.01
Wilson*	60.33	41.89	38.41

* 2017 Yield Values used to calculate 5-year average

Table 7. Tennessee Corn Olympic Average Yield Summary Stats, 2018-2022

Yield	Average	St Dev.	Min	Max	Count
NASS	169.67	8.56	150.30	187.87	31
FSA	180.04	17.93	119.33	207.17	31
APH	146.65	14.89	114.41	170.44	31

Table 8. Tennessee Soybean Yield Olympic Average Summary Stats, 2018-2022

Yield	Average	St Dev.	Min	Max	Count
NASS	63.62	3.66	55.93	70.53	22
FSA	49.64	5.49	33.70	60.32	22
APH	41.57	3.53	37.03	49.84	22

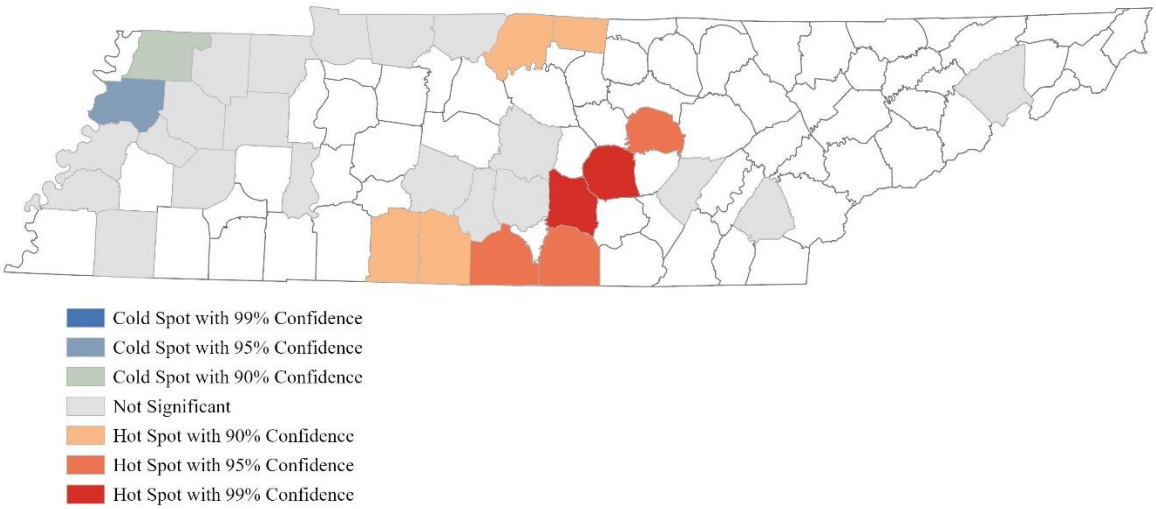


Figure 7. Hot Spot Analysis of NASS Corn Olympic Yield Averages

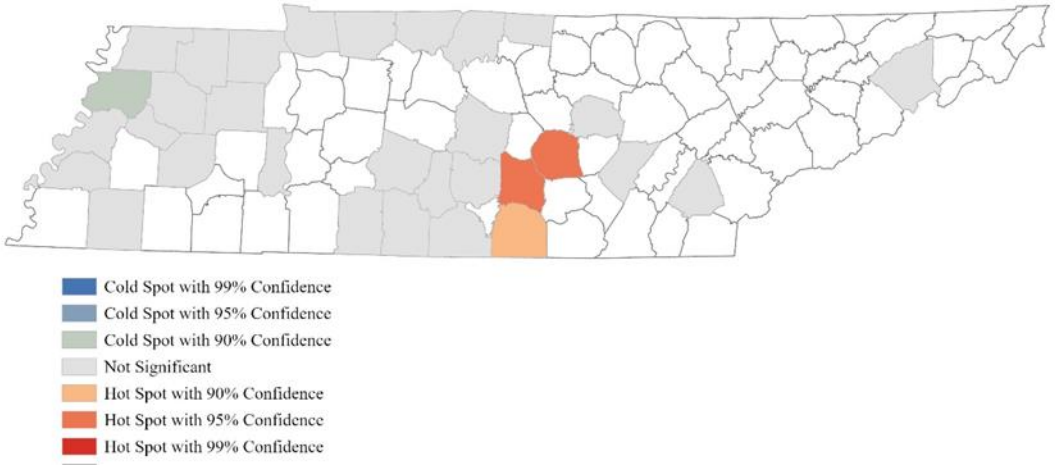


Figure 8. Hot Spot Analysis of FSA Corn Olympic Yield Averages

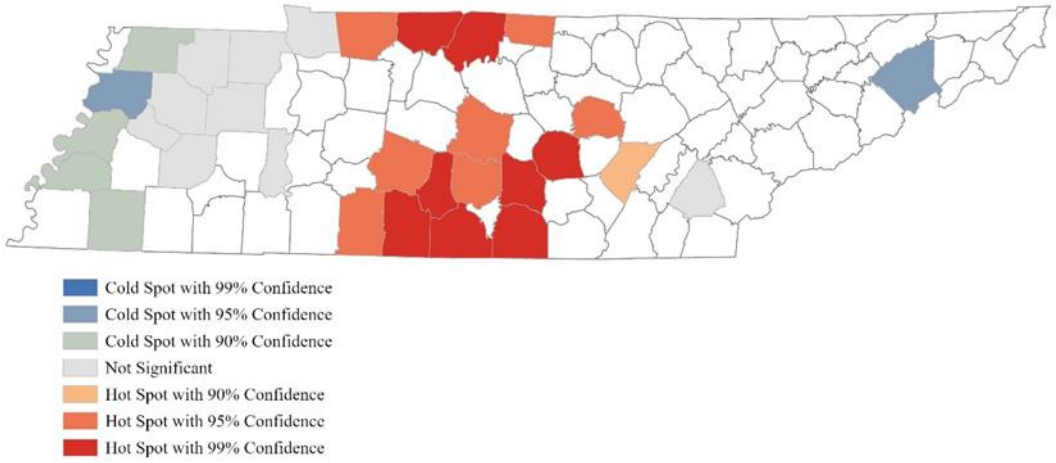


Figure 9. Hot Spot Analysis of APH Corn Yield Olympic Averages

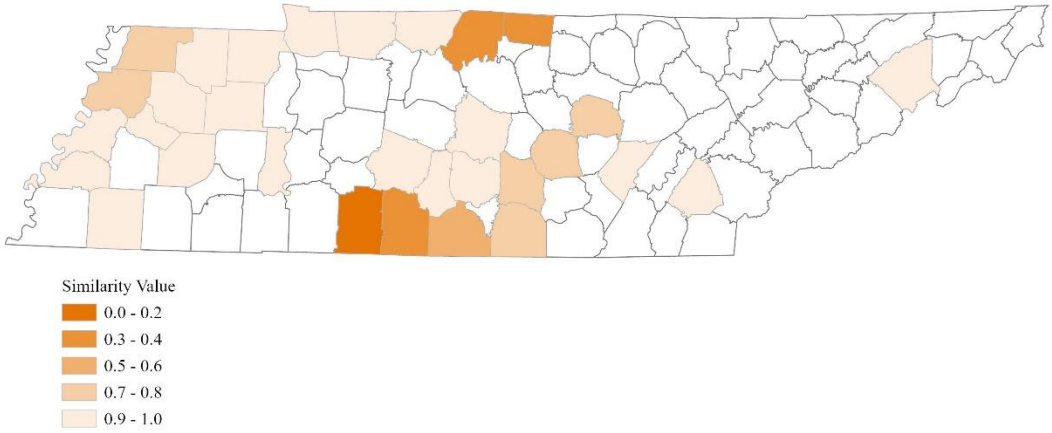


Figure 10. Hot Spot Comparison of the Olympic Averages of NASS and FSA Corn Yields from 2018 to 2022

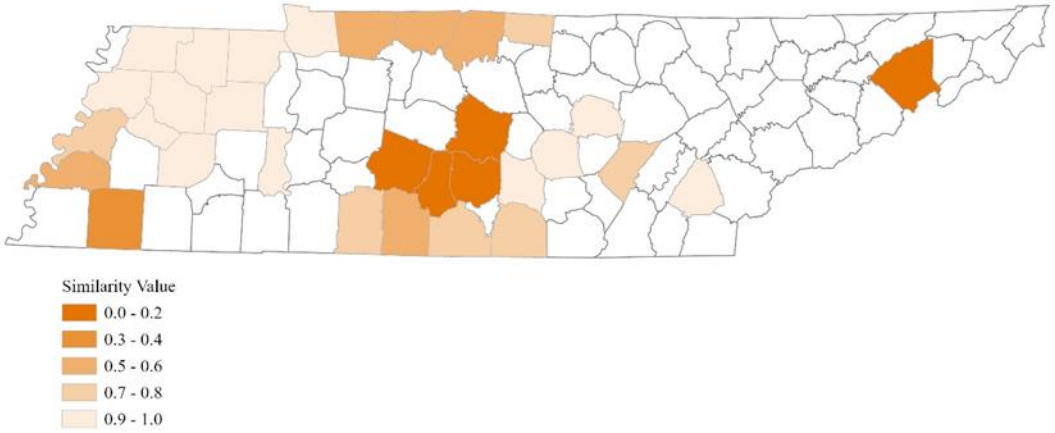


Figure 11. Hot Spot Comparison of Olympic Averages of NASS and APH Corn Yields from 2018 to 2022

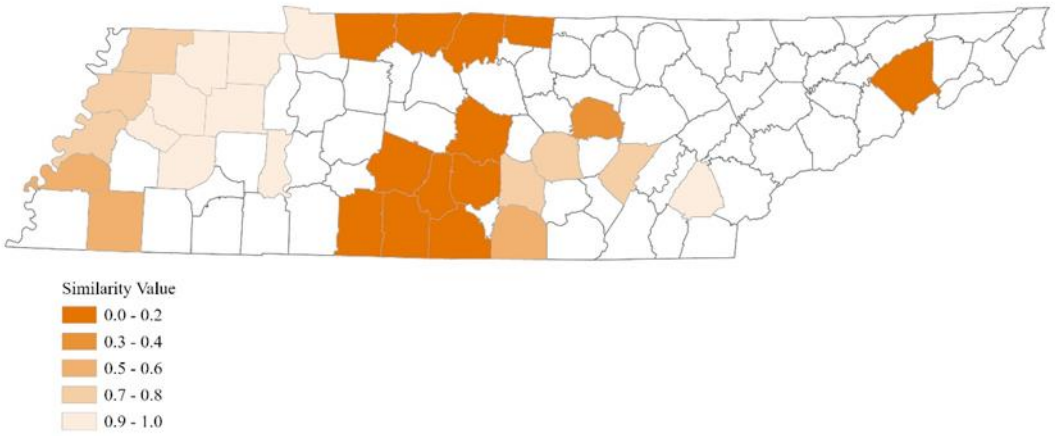


Figure 12. Hot Spot Comparison of Olympic Averages of FSA and APH Corn Yields from 2018 to 2022

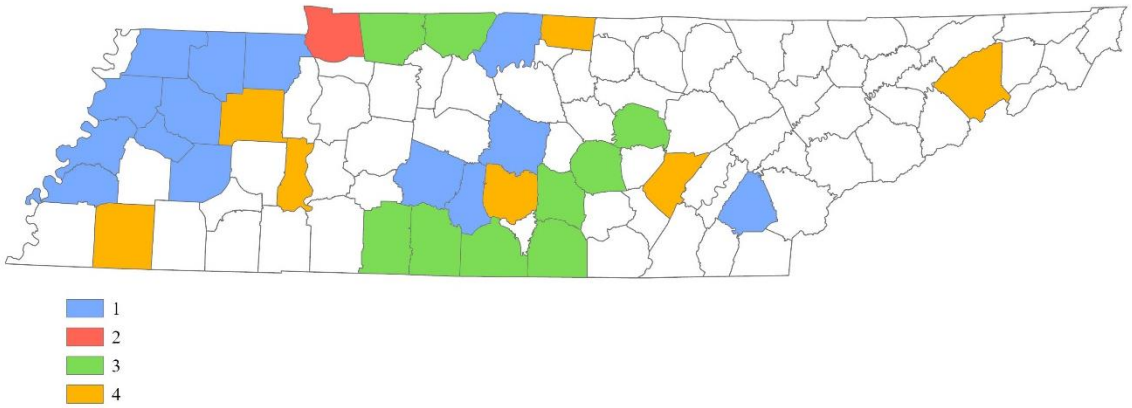


Figure 13. Multivariate Cluster Analysis of the Olympic Averages of Corn NASS Yields and FSA Yields from 2018 to 2022

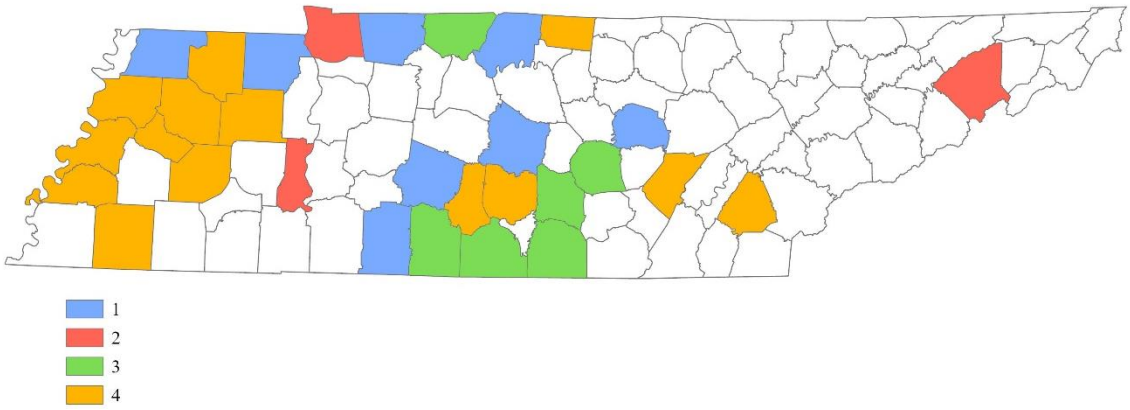


Figure 14. Multivariate Cluster Analysis of the Olympic Averages of Corn NASS Yields and APH Yields from 2018 to 2022

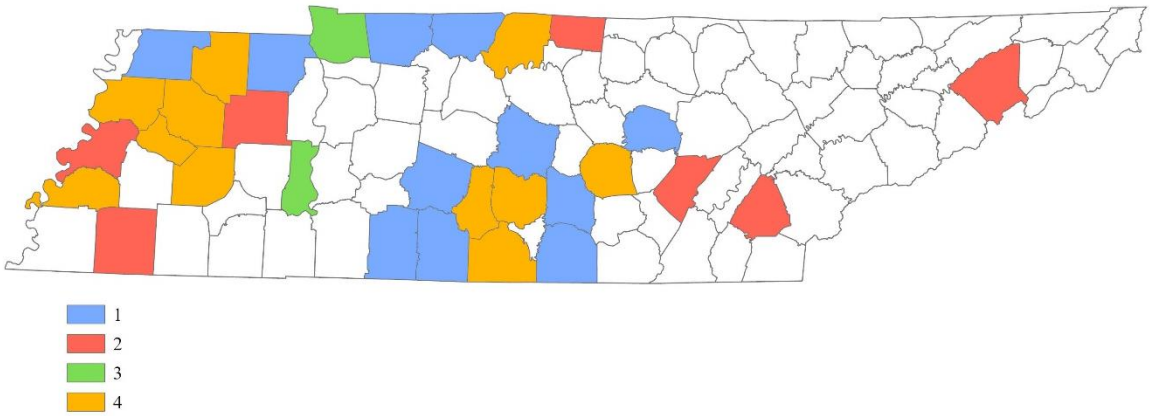


Figure 15. Multivariate Cluster Analysis of the Olympic Averages of Corn FSA Yields and APH Yields from 2018 to 2022

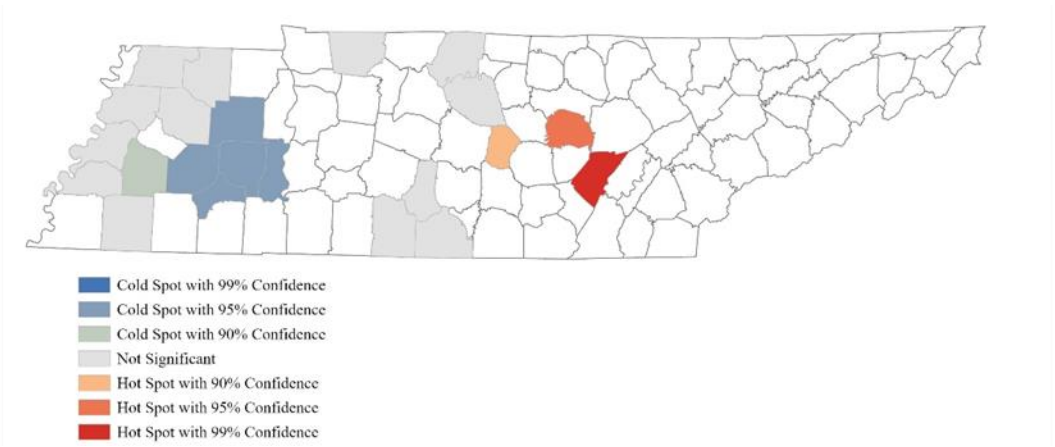


Figure 16. Hot Spot Analysis of NASS Soybean Olympic Yield Averages

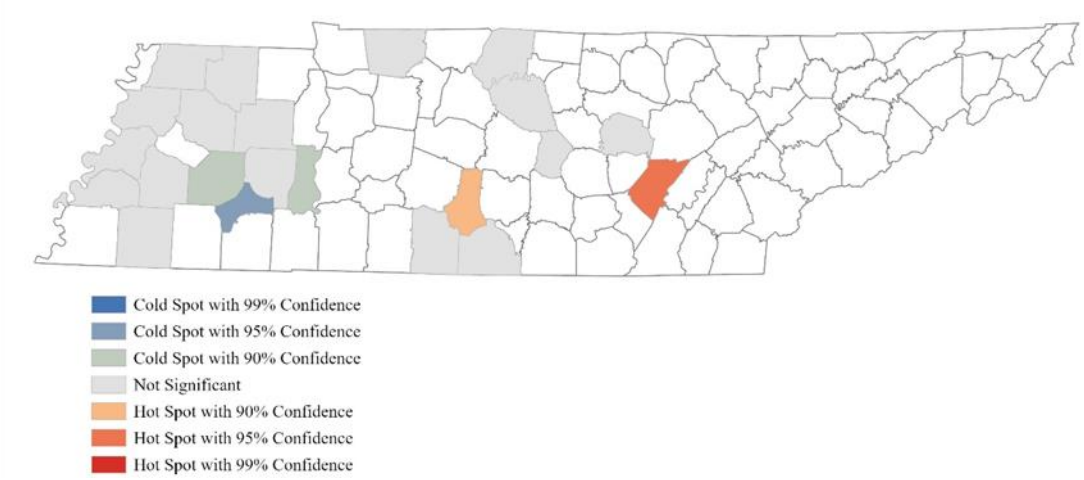


Figure 17. Hot Spot Analysis of FSA Soybean Olympic Yield Averages

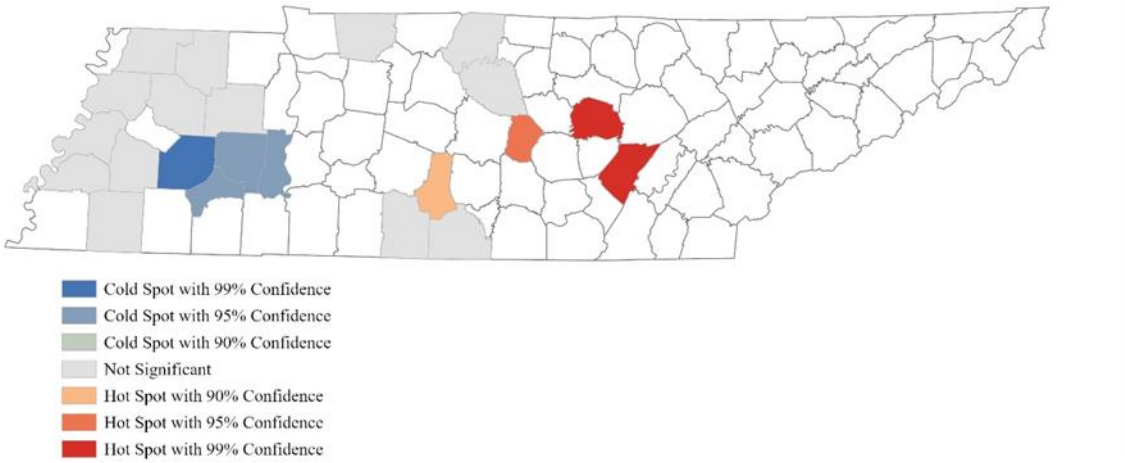


Figure 18. Hot Spot Analysis of APH Soybean Olympic Yield Averages

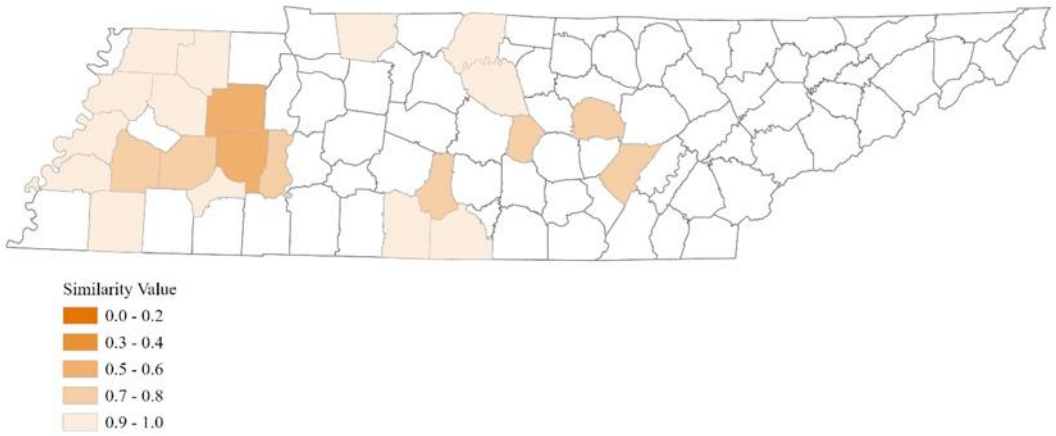


Figure 19. Hot Spot Comparison of Olympic Averages of NASS and FSA Soybean Yields from 2018 to 2022

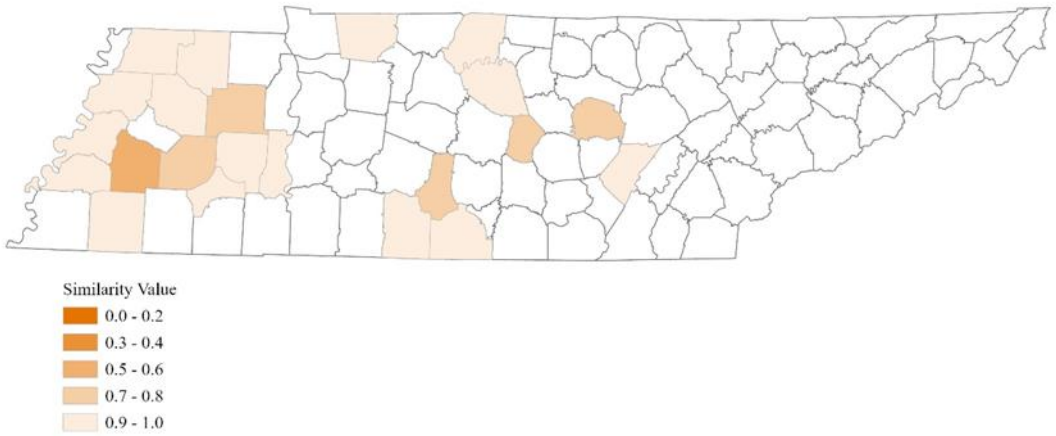


Figure 20. Hot Spot Comparison of Olympic Averages of NASS and APH Soybean Yields from 2018 to 2022

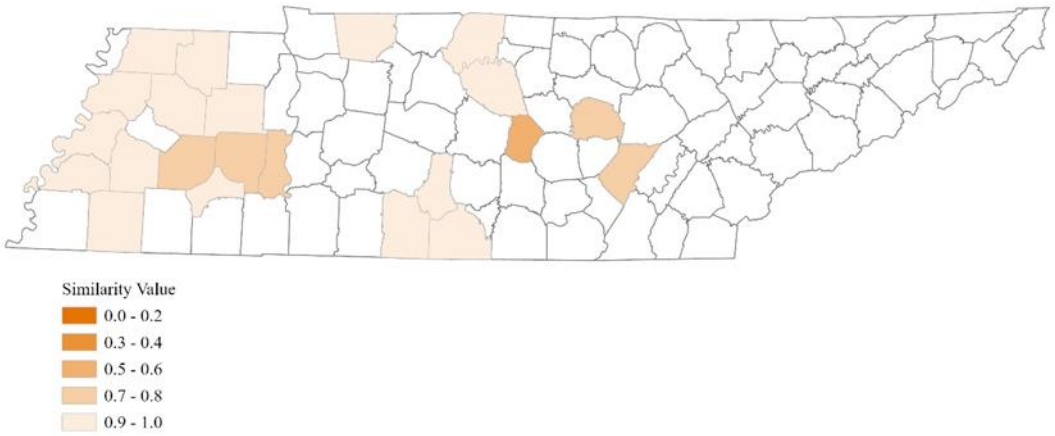


Figure 21. Hot Spot Comparison of Olympic Averages of FSA and APH Soybean Yields from 2018 to 2022

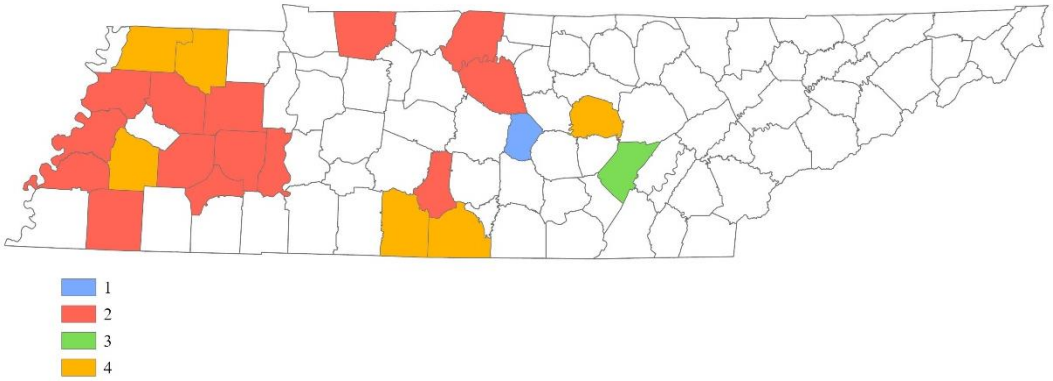


Figure 22. Multivariate Cluster Analysis of the Olympic Averages of Soybean NASS Yields and FSA Yields from 2018 to 2022

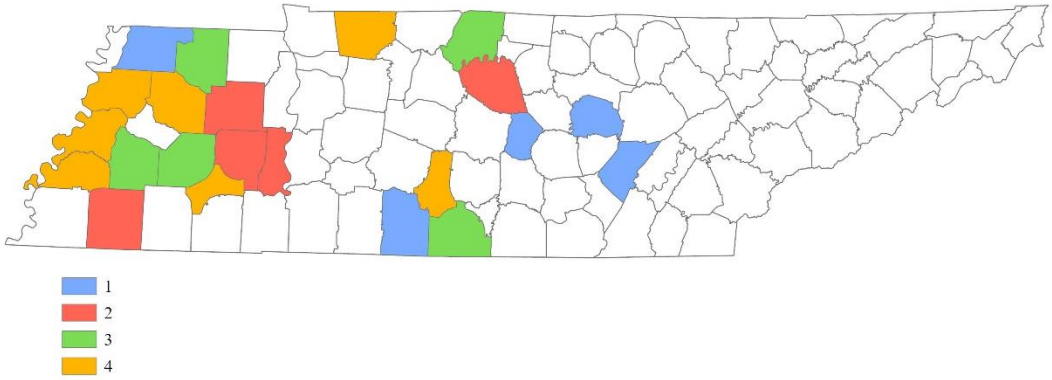


Figure 23. Multivariate Cluster Analysis of the Olympic Averages of Soybean NASS Yields and APH Yields from 2018 to 2022

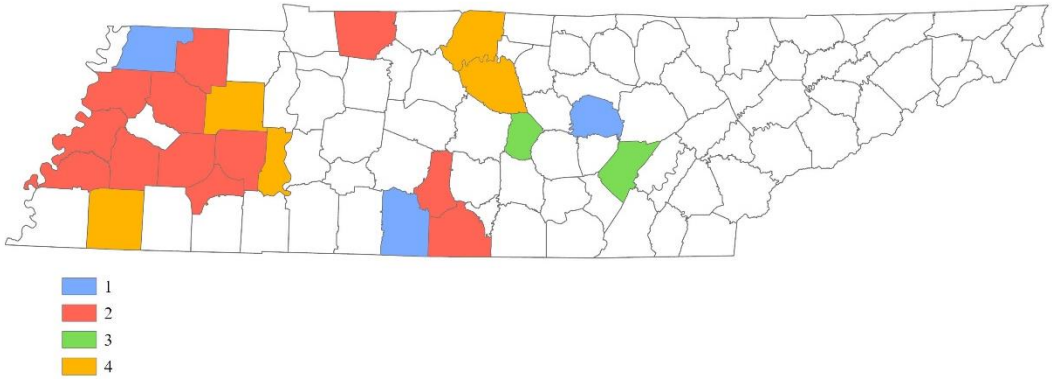


Figure 24. Multivariate Cluster Analysis of the Olympic Averages of Soybean FSA Yields and APH Yields from 2018 to 2022

Table 9. University of Tennessee Corn Variety Trial Average Yields, 2018 to 2022

County	Yield
Cannon	174.02
Coffee	208.17
Crockett	165.96
Dyer	247.00
Fayette	167.25
Gibson	172.60
Giles	206.17
Grainger	183.00
Henry	195.66
Lauderdale	121.75

Table 10. University of Tennessee Soybean Variety Trial Average Yields, 2018 to 2022

County	Yield
Cannon	50.75
Carroll	32.33
Chester	26.67
Coffee	38.67
Crockett	52.63
Dyer	44.50
Fayette	38.00
Franklin	39.44
Gibson	51.86
Giles	34.25
Haywood	48.30
Henderson	39.50
Henry	53.88
Hickman	48.25
Jefferson	47.00
Lauderdale	37.78
Madison	53.36
Marion	62.00
Maury	41.67
Obion	59.00
Tipton	39.63
Warren	43.27
Weakley	48.00

Table 11. University of Tennessee Variety Trial Summary Stats, 2018-2022

Crop	Average	St Dev.	Min	Max	Count
Corn	184.16	31.45	121.75	247.00	10
Soybeans	44.81	8.50	26.67	62.00	23

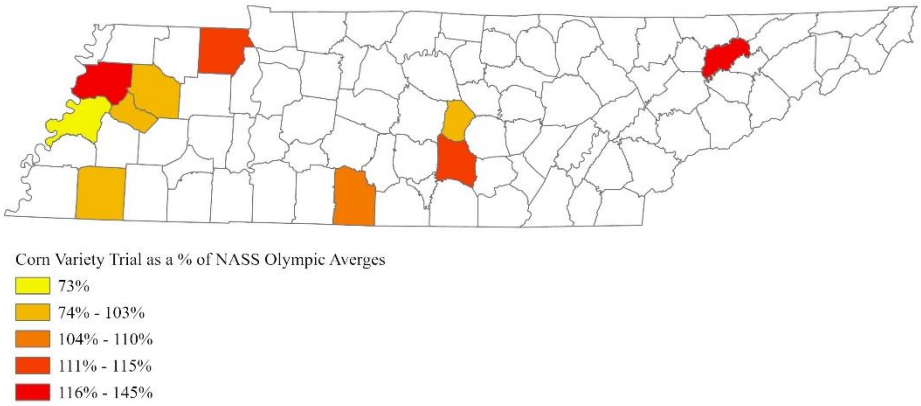


Figure 25. Corn Variety Trial Yields as a Percentage of Corn NASS Olympic Yield Averages, 2018-2022

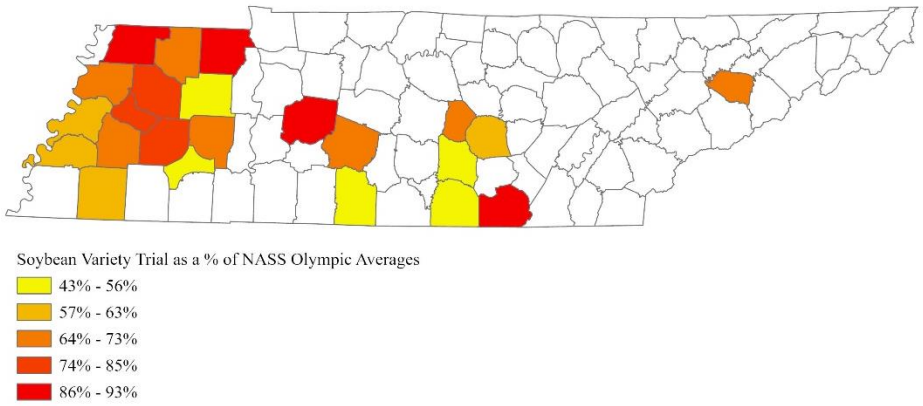


Figure 26. Soybean Variety Trial Yields as a Percentage of NASS Soybean Olympic Yield Averages, 2018-2022

VITA

William Hence Duncan is from Decherd, Tennessee. He graduated from Franklin County High school in 2018 then continued his education at the University of Tennessee. Hence graduated cum laude with a bachelor's degree in Agricultural Economics in May of 2022. In August of 2022 he decided to continue his education at the University of Tennessee receiving a master’s degree in Agricultural Economics in May of 2024. Hence will continue to study Agricultural Economics at Oklahoma State University in Agricultural Economics in August of 2024 as a doctoral student.