A Data-Driven Approach to Evaluating Soybean Best Management Practices

By

Emma Grace Matcham

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Agronomy)

at the

UNIVERSITY OF WISCONSIN-MADISON

2022

Date of final oral examination: 5 April 2022

The dissertation is approved by the following members of the Final Oral Committee:

Shawn P. Conley, Professor, Agronomy, University of Wisconsin-Madison
Steven Culman, Associate Professor, The Ohio State University
Brian D. Luck, Associate Professor, Biological Systems Engineering, University of Wisconsin-Madison
Matthew D. Ruark, Professor, Soil Science, University of Wisconsin-Madison
David E. Stoltenberg, Professor, Agronomy, University of Wisconsin-Madison

Acknowledgments

This work would not be possible without the support of many, many people. I have a wonderful community of folks filled with more people than I could possibly list, and for that I am truly grateful. I would like to thank my advisor, Dr. Shawn Conley, for his humor, kindness, and flexibility throughout my time with the Bean Team. John Gaska and Adam Roth have been immeasurably important to my success as UW both due to their research support, and probably more importantly, their truly excellent pep talks. I would also like to thank Haleigh Ortmeier-Clarke and Lindsay Malone for their writing support and friendship throughout my time at UW-Madison. Research plot maintenance and field data collection was made infinitely better thanks to Caleb Novak, Nathan Coulthurst, and Evan Robran.

My committee has also been incredibly supportive throughout my time at UW-Madison. Specifically, I'd like to thank Dr. Matt Ruark for his support, especially at the beginning of the COVID-19 pandemic. Thank you to Drs. Brian Luck and Steve Culman for their thoughtful conversations—you've both pushed me to think bigger and more conceptually. Also, thank you to Dr. Dave Stoltenberg for your kindness and grounding—your feedback has helped me connect my research to the broader cropping systems we work within.

Throughout graduate school I have the privilege of working with agronomists and researchers from across the eastern US, and cannot possibly name them all here, but please know I appreciate you all. Special thanks to Dr. Rachel Vann for trusting me with the foliar fertilizer project—I learned so much about project administration and myself throughout the process. I would also like to thank Dr. Laura Lindsey for being an unending source of guidance throughout my graduate school career. Additionally, thank you to the farmers who contributed management data and yield maps to these research projects-- your thoughtful questions have been so valuable to my research and education.

Finally, I'd like to thank my friends and family for their support throughout graduate school. To my wife Kaitlin, thank you for being you. I love you and am so grateful to have you in my life. I'd also like to thank our dogs, Fred and Barney, for being the best writing buddies I could ever ask for. Thank you to the Oppengard-O'Donnell family for your encouragement and humor over the last 3 years. And finally, thank you to Lori Hayes, Tracy Campbell, Dr. Spyros Mourtzinis, Dr. Greg McGlinch, Dr. Filipe Matias, Andrew Osburn, Fabiano Collet, and Wanderson Novais for improving my research and writing skills, and especially for your patient listening to my many, many rants.

Chapters 4, 5, and 6 were previously published in Agronomy Journal. Original publications are available at the sources below:

Matcham, E. G., Mourtzinis, S., Conley, S. P., Rattalino Edreira, J. I., Grassini, P., Roth, A. C., Casteel, S. N., Ciampitti, I. A., Kandel, H. J., Kyveryga, P. M., Licht, M. A., Mueller, D. S., Nafziger, E. D., Naeve, S. L., Stanley, J., Staton, M. J., and Lindsey, L. E. (2020). Management strategies for early and late-planted soybean in the North Central US. *Agronomy Journal 112(*4): 2928-2943*.* https://doi.org/10.1002/agj2.20289

Matcham, E. G., Vann, R., Lindsey, L. E., Gaska, J. M., Lilley, D. T., Ross, J., Wright, D. L., Knott, C., Lee, C. D., Mosely, D., Singh, M., Naeve, S., Irby, T., Weibold, W., Kandel, H., Lofton, J., Inman, M., Kleinjan, J., Holshouser, D. L., and Conley, S. P. (2021). Foliar Fertilizers Rarely Increase Yield in US Soybean. *Agronomy Journal 113*(6): 5246-5253. https://doi.org/10.1002/agj2.20889

Matcham, E. G., Matias, F.I., Luck, B., and Conley, S. P. (2022). Filtering, editing, and cropping yield maps in R environment with the package cleanRfield. Accepted for publication at *Agronomy Journal*.

Table of Contents

Abstract

Regularly evaluating best management practices for soybean is important to maintaining agronomic crop production as the climate and seed varieties change over time. Many phosphorous and potassium fertilizer recommendations in the North Central US are based on the build-maintain framework and were developed in 1970s and 80s and are due to be reevaluated. To estimate the yield-maximizing soil test potassium level (YMK) under current growing conditions, nutrient management records and yield maps from southern WI were analyzed via quadratic quantile regression to estimate both overall YMK and determine if YMK varied across the study space. The overall YMK was 76 ppm, and lower buffer pH and organic matter levels were associated with higher YMK. Some fertilizer recommendations include leaf tissue K concentrations in addition to soil test K levels. Results of a 2021 on-farm trial indicate that the critical K concentration in soybean leaf tissue is 2.04%. The relationship between K soil test results from Bray-1 extraction and Mehlich-3 extraction for silty loam soils was represented by the linear regression line Bray = 0.77 * Mehlich - 0.75.

Management decisions that increase soybean yield are region-specific and vary between planting dates, so larger multi-state research projects are valuable for developing best management practices. In a survey study of soybean farmers in ten North Central US states, late-planted fields had higher yields associated with tillage and using both a PRE and POST herbicide application. Early-planted fields had higher yields associated with artificial drainage, insecticide seed treatment, and lower seeding rates. Less variation between sites was observed in a small-plot study of foliar fertilizers across 46 site-years in 16 eastern US states. Foliar fertilizers did not increase soybean

yield in the absence of visual symptoms of nutrient deficiency. In multi-state and on-farm research, efficient processing of yield maps represents a research bottleneck. A new R package, cleanRfield, allows for more efficient processing of yield maps. Together, these projects represent ways for multistate and multidisciplinary teams to leverage technology and improve best management practices for soybean production.

Chapter 1: A Review of Current Soybean Fertility Recommendations

Soybean Production

The United States (US) produces 34% of the world's soybeans, more than any other country (American Soybean Association, 2019). Of the roughly 2 million farms in the US, over 300,000 grew soybeans in 2017. In total, 36.5 million hectares of soybean were harvested, producing 117.3 billion kg of soybean seed (USDA NASS, 2019).

Wisconsin (WI) produced 2.8 billion kg of soybean seed in 2017 (USDA NASS, 2019). In 2019, there were 890,000 hectares of soybeans planted in WI, the 14th most of any state (American Soybean Association, 2019). Average WI soybean yield in 2018 was 3,233 kg ha^{-1}, slightly lower than the national average of 3,476 kg ha^{-1} (American Soybean Association, 2019; USDA NASS, 2020)

While management practices vary between fields and years, WI producers typially plant soybean in the second or third week of May (Rattalino Edreira et al., 2019). In southern WI, seeding rates generally range from 333k to 408k seeds ha^{-1} In northern WI, seeding rates can be as high as 445k seeds/acre (Rattalino Edreira et al., 2019). Most fields in WI are planted with treated seed, and most soybeans are planted following a corn crop. The most common row spacing in WI is 38 cm, although parts of eastern WI have more hectarage with 76 cm row spacing. Fewer than half of WI soybean fields have artificial drainage (Rattalino Edreira et al., 2019).

Current Fertilizer Recommendations

Potassium (K) and phosphorous (P) are essential plant macronutrients. Potassium is used in plants to activate enzymes and open and close stomata, regulating transpiration (Tisdale et al., 1985). Phosphorous is crucial for storing chemical energy in plants, such as in adenosine triphosphate (ATP) (Tisdale et al., 1985).

Plant available soil nutrients are estimated using soil sampling. Potassium in the soil solution is available for plant uptake. Exchangeable K, or K^+ ions from the surface of clay particles, can enter the soil solution to replenish soil solution K (Franzen & Bu, 2018). Non-exchangeable K, tightly held between layers of clays, is sometimes available to plants when held between illite particles but not when held by other clays. Mineral K, or K that is a part of the micas and feldspars clays form from does not enter the soil solution within a given growing season (Franzen & Bu, 2018).

When soybeans are harvested, nutrients are removed with the grain. Without fertilization, available nutrients in the soil will become depleted over time (Tisdale et al., 1985). State fertilizer rate recommendations for P and K generally follow either a sufficiency or a build-maintain approach (Macnack, 2017). The sufficiency approach sets a soil test level at which soils are considered “sufficient” for a given crop. When soils are at 80% sufficiency, they can reach 80% of their yield potential. Nutrients are applied annually to meet crop need; however, additional nutrients are not applied to keep soil test levels sufficient for future growing seasons (Hoskins, 1997; Macnack, 2017).

Unlike the sufficiency approach, the build-maintain approach accounts for nutrients removed in crop harvest. Those nutrients are replaced using fertilizer prior to, or during the following growing season. Soils are divided into three categories based on soil test values: build, maintain, and drawdown (Heckman, 2012; Macnack, 2017; Vitosh et al., 1995). In WI and IA, soils are split into six or five categories, respectively, but the same general build-maintain approach is followed (Laboksi & Peters, 2012; A. Mallarino, 2013). When soils are above the soil test critical level for a given crop, nutrients are applied to replace the nutrientslost in crop removal. When soils are below the soil test critical level, fertilizers that exceed crop removal are applied over a period of four or more years to spread out fertilizer cost as soil test values slowly increase (Macnack, 2017). Applications for multiple seasons can be made at once (Macnack, 2017). Critical levels from the build-maintain approach are the same as the 100% sufficiency level in the sufficiency approach (Macnack, 2017).

Methods to Calculate Critical Levels

Soils above the critical level for a given nutrient have sufficient available nutrients for a given crop (Cox, 1992). Crops grown on soils below the critical level are likely to respond to fertilizer (Heckman, 2012; Vitosh et al., 1995). In a survey of Ohio fields, fields below the STK critical level had an average yield of 296 kg ha^{-1} lower than fields above the critical level for STK (Brooker et al., 2017).

Typically, critical levels are estimated using fertilizer rate trials. For each trial, the relative grain yield is calculated by dividing the yield of the unfertilized plots by the yield of the fertilized plots. The Cate-Nelson graphical method was historically used to estimate the critical level from the relative grain yield and soil test values (Nelson & Anderson, 1977). The responsive trials, or trials where yield was higher in fertilized plots than in unfertilized plots, were plotted with one symbol and the non-responsive trials were planted with another color. A clear sheet of plastic divided into four quadrants was laid over a scatter plot of relative grain yield and soil test value. The sheet of plastic was moved across the scatter plot to maximize the number of responsive trials falling in the lower-left quadrant and the number of non-responsive trials falling in the upper-right quadrant while keeping the quadrant lines square with the scatter plot axes. Once the plastic sheet was appropriately placed, the intersection of the vertical line on the plastic sheet with the x-axis represented the critical level (Nelson & Anderson, 1977). Cate and Nelson later released an updated method that uses an iterative process to divide the data based on soil test value. The final estimate of the critical value is the soil test value where the R^2 value of the prediction of whether there will be a response to fertilizer is highest (Cate & Nelson, 1971). The R package rcompanion can implement this procedure without the need to physically manipulate a plastic sheet and has been used by Ohio researchers looking at both soil test critical level and critical leaf nutrient concentrations (Fulford & Culman, 2018; Mangiafico, 2017). The Cate-Nelson method, as implemented by the NLIN procedure in SAS, has also been used to relate critical K concentration to anthracnose severity for annual bluegrass (Schmid et al., 2018).

Other methods based on plotting relative grain yield against soil test value can also be used to estimate the critical level. A regression can predict relative yield from soil test value. Early regressions for determining critical levels used a modified Mitscherlich growth curve (Nelson & Anderson, 1977). Today, a quadratic plateau model is commonly used (Dodd & Mallarino, 2005; Van Scoyoc, 2004). The soil test level where the model plateaus and relative yield no longer increases is considered the soil test critical level (Dodd & Mallarino, 2005; Singh et al., 2019; Van Scoyoc, 2004; Williams et al., 2017, 2018). Similar methods are used to predict critical concentration ranges of nutrients in plant tissues using linear plateau models (Stammer & Mallarino, 2018; Williams et al., 2018).

Studies to estimate the critical level will place fertilizer trials across ten or more site-years. Instead of estimating the critical level using quadratic plateau models, some studies determine whether there is a difference in grain yield between fertilized and unfertilized plots for each site-year using an ANOVA (Clover & Mallarino, 2013; Culman, S et al., 2017; Fulford & Culman, 2018; Nelson & Anderson, 1977). The highest soil test value where the site still maintained higher yield in fertilized plots than in unfertilized plots is considered an estimate of the critical level (Antonangelo et al., 2019; Clover & Mallarino, 2013).

Boundary line analysis (BLA) is a useful approach for estimating the yield potential at a range of soil nutrient values (Shatar & McBratney, 2004). This technique is an extension of Liebig’s Law of the Minimum, which proports that yield is most limited by the nutrient that is available in the lowest quantity relative to the plant’s total need for that nutrient. Early implementations of BLA were performed by hand-drawing curves on scatter plots with soil nutrient status on the X axis and yield on the Y axis such that the curve approximated the maximum yield across the domain of soil nutrient status. Shatar and McBratney (2004) have since developed a standardized method for implementing BLA in S-PLUS, incorporating a spline model so that the regression can fit the maximum possible yield across a wider range of yield distributions. Boundary line analysis has also been used to evaluate soil nutrient concentration recommendations and to quantify yield gaps for wheat at both whole-field and within-field scales (Hajjarpoor et al., 2018; Lark et al., 2020).

Measuring Soil Nutrients

Soil sampling is used to determine if the level of nutrients in a field is above or below the soil test critical level. In WI, state recommendations indicate that one sample should be taken for every 2 hectares. Ten cores from a uniform depth of at least 15 cm should be taken for every composite sample, and composite samples should be dried and sent to a reputable lab (Laboksi & Peters, 2012).

Nutrient levels in soil samples can be measured using a variety of lab methods. In WI, Bray extractant is used for both P and K (Laboksi & Peters, 2012). Iowa P fertilizer recommendations are available for Bray, Mehlich-3, and Olsen extractants, and K recommendations are available for ammonium acetate and Mehlich-3 tests (Mallarino, 2013). Ohio, Indiana, and Michigan fertilizer recommendations had been based on a Bray extractant for P and ammonium acetate for K until 2020 when new recommendations were released based on Mehlich-3 extractant for both P and K (Culman et al., 2019, 2020; Vitosh et al., 1995). North and South Dakota both use the Olsen method for measuring P since it is less sensitive than Bray extraction to calcium carbonate in soils. Mehlich-3 or ammonium acetate can both be used to extract K, but most state recommendations are based on ammonium acetate since it extracts similar amounts of K from samples regardless of soil pH (Eliason et al., 2015). As an alternative to measuring extractable K, a study of Italian ryegrass grown in nine different soils indicates that measuring soil solution K^+ concentration may be a more accurate way to predict plant-available K and estimate critical levels due to the close relationship between K^+ buffer power and soil solution K^+ concentration (Mengel & Busch, 1982).

In addition to measuring soil nutrients, some fertilizer rate trials also measure leaf K concentration. Increases in leaf K concentrations due to fertilizer applications are observed even when K fertilizer does not result in increased yield (Farmaha et al., 2011; Fulford & Culman, 2018). In an Ohio fertilizer rate trial, leaf K concentration was better predicted by fertilizer rate than by STK levels (Fulford & Culman, 2018). Current estimates of soybean plant critical K concentration in Iowa were 18.8-22.7 g K kg^{-1}, and leaf tissue critical K concentration was slightly lower at 15.6-19.9 g K per kg. Critical P concentrations for soybean plants and leaves, respectively, were 3.3- 4.1 and 3.5-4.7 g P per kg (Stammer & Mallarino, 2018).

Farmers in the UK use the Agriculture and Horticulture Development Board Nutrient Management Guide (RB209) to make fertilizer rate and timing decisions, which uses tissue nutrient concentration to determine sufficient field fertility for S, Zn, and other nutrients. But, soil nutrient concentration is used for P and K fertilizer rate recommendations (Agriculture and Horticulture Development Board, 2021, p. 209). Other European countries also use tissue nutrient concentration instead of or in addition to the soil nutrient critical level methods that are common in the US. The nutrient management guide for the UK, RB209, states that arable crop yield is maximized when soil P and K concentrations are 16-25 and 121-180 mg l^{-1}, respectively (Agriculture and Horticulture Development Board, 2021). In 2020, this range of critical concentrations was confirmed using boundary line analysis by Lark, et al.

Estimating Nutrient Removal Rates

Accurate estimates of nutrient removal are crucial for implementing a build-maintain strategy, since fertilizer recommendations are intended to return the nutrients removed through harvest so that soil test values remain stable over time (Laboksi & Peters, 2012). Soybean accumulates 172 kg ha^{-1} K to produce 3.5 Mg ha^{-1} of grain (Bender et al., 2015). Soybean K concentration in WI is estimated to be 0.016 kg K kg^{-1} grain and 0.0054 kg P kg^{-1} grain (Gaspar et al., 2017). Potassium uptake is highest during late vegetative growth, and 75-100% of K is taken up before grain fill (Bender et al., 2015; Gaspar et al., 2017). Phosphorous accumulation is generally lower (21 kg ha^{-1} P) and occurs later in the season (Bender et al., 2015; Gaspar et al., 2017). To grow 3,435 kg ha^{-1} soybeans in WI (state average) on optimum soils and replace the nutrients removed, an annual application of 79 kg ha^{-1} of K_2O is recommended (Laboksi & Peters, 2012).

Removal rates on an area basis are more affected by grain yield than potassium concentration (Clover & Mallarino, 2013; Gaspar et al., 2017); thus, nutrient removal rates are presented as pounds of K_2O per bushel of grain, or as kilograms of K_2O per kilogram of grain (Culman et al., 2020; Mallarino, 2013). Yield level did not change grain K concentration (Gaspar et al., 2017). Ohio recommendations from 1995 estimate 0.53 kg of K_2O are removed with every bushel of soybeans (Culman et al., 2020). A study taking place from 2006 to 2014, soil test values fell even when fertilizer was applied at the estimated crop removal rate, and soil test values did not rise when fertilizer was applied at twice the estimated crop removal rate (Fulford & Culman, 2018). Kansas has also observed a long-term reduction in STK levels and a subsequent increase in K deficiency symptoms due to nutrient applications under the KSU sufficiency recommendations being below the nutrient removal rate (Matz, 2012). Both Ohio and Kansas may need to increase crop removal estimates to maintain STK and STP levels.

In Illinois, updated estimates of nutrient removal for soybeans based on samples from 2,620 fields over three years indicate soybeans remove 0.34 kg P_2O_5 and 0.53 kg K_2O per bushel, which is roughly 11% lower than past estimates. Nutrient removal per bushel did not vary based on field average yield or location (Nafziger, 2017). Iowa researchers estimate nutrient removal rates for soybean are 0.33 kg P_2O_5 and 0.55 kg K_2O per bushel of grain, which is very similar to the revised IL estimates (Mallarino, 2013).

Spatial Variability in Soil

Soils vary in STK due to both natural anthropogenic processes. Steeper slopes have lower STK, likely due to higher erosion rates (Jiang & Thelen, 2004; Kravchenko & Bullock, 2000). Backslopes have greater STK than other landscape positions. In areas with a greater depth to claypan, K buffering is also decreased (Conway et al., 2018). Crop rotation can impact STK due to previous crops removing different amounts of K in their grain (Conway et al., 2018; Singer et al., 2004).

Quantifying expected variation in STK is important for planning sampling regimes. Reported coefficients of variation (CV) for STK in an Iowa grid soil sampling study range from 19-43% (Mallarino & Wittry, 2004). In cultivated fields in Australia, the CV for STK ranged from 24-64% (Bolland & Allen, 1998). In Mississippi, the CV for STK was 22-85% (Cox et al., 2003). Coefficients of variation for STP have a narrower observed range than STK, with CVs ranging from 30-55% in Iowa and from 32-44% in Australia (Bolland & Allen, 1998; Mallarino & Wittry, 2004).

Predicting where and how STK levels will vary is important for planning effective soil sampling. The spatial structure of variability can be illuminated through semivariogram analysis. The range of a variogram is the maximum distance where spatial autocorrelation is observed between samples (Avendaño et al., 2004). Semivariograms of STK in eight Iowa fields and two Michigan fields had ranges between 25 and 100m (Avendaño et al., 2004; Mallarino, 1996). Current 1-3 ha grid sampling may not be dense enough to see spatial structure in STK levels due to samples being spaced further apart than the range of their variograms (Mallarino, 1996).

In addition to the range distance determined through semivariograms, the direction of spatial autocorrelation is also important to understanding the spatial structure of soil properties. Anisotrophic properties have spatial autocorrelation patterns that differ with both distance and direction between sample locations (*Esri Support GIS Dictionary*, 2021). The spatial structures of soil properties including STK, Mg, pH, and electrical conductivity exhibit anisotrophy (Fu et al., 2010; Kitchen et al., 2005; Piepho et al., 2011). Yield monitor observations from two soybean fields was also observed to be anisotrophic (Avendaño et al., 2004). In this same study, soil properties including STK, STP, pH, Ca, and Mg were not anisotrophic, and other studies have observed that electrical conductivity was not anisotrophic for some fields (Avendaño et al., 2004; Landrum et al., 2015). Testing properties for anisotrophy is important because it varies between fields and properties, and it impacts the types of interpolation methods that can be used on the data (Hengl et al., 2007).

Within-Field Soybean Yield Patterns

Soil properties vary between and within fields, but the relationship between soybean yield and different soil properties is inconsistent at the within-field scale. STP and STK were important predictors of between- and within-field soybean yield variation in WI (Smidt et al., 2016). In three of four MI site years, STK was correlated with soybean yield when STK varied between 50 and 550 kg/ha (Avendaño et al., 2004). In a Mississippi study of within-field yield variation, STK was negatively correlated with yield and observed correlation coefficient values between -0.33 and -0.48 for each field. Authors hypothesize this inverse relationship between STK and yield could be due to low-yielding areas of the field accumulating K over time due to lower crop removal rates (Cox et al., 2003).

Other studies did not observe correlation between STK and soybean yield. When STK levels are above the critical level, it is not a predictor of soybean yield (Culman et al., 2017; Kravchenko & Bullock, 2000). In a MN study where yield was inversely related to pH and OM over 6 site-years, there was no relation between yield and STK. Seed protein concentration was negatively related to STK (Anthony et al., 2012).

Terrain can also impact soybean growth. Slope was correlated with yield at three out of eight sites in IL and IN. Correlation coefficient values ranged from 0.11 to 0.39 (Kravchenko & Bullock, 2000). The most important factor for predicting within-field soybean yield across 22 WI site years was elevation (Smidt et al., 2016).

Variable Rate Fertilizer Application

Variable rate fertilizer allows producers to apply fertilizer to parts of the field that are below the critical level without applying fertilizer to parts of the field that are above the critical level and unlikely to see a yield response to fertilizer. Iowa research on the level of STK variation indicates that variable rate fertilizer could improve nutrient use efficiency and farm profitability (Wittry & Mallarino, 2004). No-till sites had higher variability in the spatial structure of their STK levels, indicating that they might experience greater benefit of VR fertilizer (Mallarino, 1996). Sixty one percent of Minnesota fields have some areas below the optimum STK level, at the optimum STK range, and above the STK level. In these fields, it is economical to apply fertilizer only to the areas below the STK optimum level and conserve nutrients in high STK areas (McGraw, 1994).

Adoption of variable rate technology varies by farm size. Larger farms are more likely to adopt precision agriculture technologies than smaller farms (Castle et al., 2016; Robertson et al., 2012). Thirty to 40% of US corn producers with over 1,175 ha in production use some type of variable rate technology, and 70-80% use mapping technologies. This is double the national average for VRT and mapping technology adoption across all farm sizes (Schimmelpfennig, 2016). In 2020, 81% of surveyed dealers reported using autosteer and other GPS technology for within-field navigation and 89% offer variable rate fertilizer application (Erickson & Lowenberg-DeBoer, 2020).

Effectively implementing variable rate fertilizer application can be challenging due to variation in fertilizer products and equipment limitations. Variation in particle size accounts for most variation in fertilizer application rates, with small and medium particles traveling shorter distances from the spreader-disc (Fulton & Port, 2016; Virk et al., 2013). In fertilizer with high variation in particle size, the short movement of small particles leads to high application rates directly behind the spreader (Virk et al., 2013). Particle size differences are common with potassium fertilizer, since potash is irregular in shape (Fulton &

Port, 2016). Terrain and wind speed can also impact the accuracy of disc-spreaders, with wind tending to move dust and fine particles further than heavy particles (Fulton & Port, 2016). All fertilizer applicators have inherent error, and additional error can be introduced due to differences in applicator speed and distance between equipment passes (Lawrence & Yule, 2007). Variable rate fertilizer applicators also introduce error when changing between rates.

Past studies have compared fertilizer prescriptions, as-applied maps, and field-collected fertilizer quantities to quantify spreader error. Fulton et al. (2013) found that as-applied maps correlated with field-collected fertilizer quantities and accurately represented where rate changes occurred for VR fertilizer applications. Fertilizer applicators were within 10% of their target rate 25-45% of the time (Fulton et al., 2013). A similar study found an R^2 value of 0.47 relating as-applied maps to prescriptions (Fulton et al., 2001).

Variable Crop Nutrient Need

Most fertilizer recommendations are based off a single estimate of the soil test critical level and nutrient removal, but there are some studies that indicate that the critical level is not constant. In Wisconsin, optimum soil test K (STK) for growing soybeans on loamy soil is 101-130 mg kg^{-1}. On sandy or organic soils, the optimum STK is 66-90 mg kg^{-1} (Laboksi & Peters, 2012). The optimum range differs between soil textures because finer texture soils can supply more P and K than coarser soils. Iowa also varies state fertilizer recommendations based on soil texture, and OH, MI, and IN use CEC instead of soil texture (Mallarino, 2013; Vitosh et al., 1995).

North Dakota has observed differences in the STK critical level based on clay minerology of soils. Out of 25 sites between 2014 and 2016, those with a smectite/illite ratio of greater than 3.5 had STK critical levels roughly 70 mg kg^{-1} higher than those sites with a smectite/illite ratio of less than 3.5 (Breker et al., 2019). Crop response to K fertilizer predictions on sites with STK between 130-200 mg kg^{-1} were more accurate when accounting for clay minerology (Breker et al., 2019). Soils that are high in illite minerals hold more K that is not detected on a standard exchangeable K soil test, accounting for the differences in exchangeable STK critical levels (Franzen & Bu, 2018). Crops may need higher exchangeable STK levels when soil conditions are dry, due to dry soils having tighter clay interlayer spaces and not releasing non-exchangeable K into the soil solution as readily (Franzen & Bu, 2018). Differences in STK critical level based on minerology could be due to differences in how different clay minerals fix and release K ions. Clay minerals with more negative layer charges, such as vermiculites, tend to fix more K than minerals with less negative charges, such as smectites. Potassium fixation also increases when iron on clay layers is reduced, particularly when the iron is in tetrahedral layers as compared to octahedral layers (Florence et al., 2017). Since K fixation varies between clay minerals and is impacted by iron levels and oxidation states, minerology may change optimum STK levels for crop production.

Soybeans also have variable STP and zinc needs. Across six MN site-years, Olsen STP critical levels consistently varied based on pH. Moderately acidic and neutral soils maximized soybean yield between 15-19 mg kg^{-1}, while moderately alkaline soils maximized yield at 36 mg kg^{-1}. These values correspond to an economic optimum STP rate of 15 for acidic and neutral soils and of 30 for alkaline soils (Anthony et al., 2012). Optimum zinc rates were also higher in alkaline soils than acidic soils (Anthony et al., 2012).

Analyzing data across many response trials from different environments has helped identify site characteristics that predict differences in soil test critical level or optimum fertilizer rate. A meta-analysis of approximately 2000 site-years of P rate trials in Germany and Austria found that crop responses to phosphorous are variable but related to STP, SOM, and soil texture (Buczko et al., 2018). In a similar decision tree analysis of approximately 9000 central European P and K fertilizer response trials, K response was best predicted by crop and STK (Kuchenbuch & Buczko, 2011). For high STK sites planted with cereal crops, the final predictor of yield response was soil texture (Kuchenbuch & Buczko, 2011). The most important predictor of yield response to P fertilizer was STP. On low STP sites, yield increase was predicted by pH. On high STP sites, yield increase was predicted by soil texture and pH. Crop was relatively unimportant for predicting response to P fertilizer as compared to predicting response to K fertilizer (Kuchenbuch & Buczko, 2011).

In a study of 43 soybean fields in China, soil properties accounted for only 24% of yield variability using a generalized linear model. Soil organic carbon was only an important variable when predicting yield in a CART analysis of soil properties and in-season management when less than 22.37 kg ha^{-1} P was applied on non-manured fields (Zheng et al., 2009). Between field subsections, total soil potassium was the most important predictor of soybean yield. Available soil potassium and soil EC were important on sites with high total soil potassium (Zheng et al., 2009).

Similar meta-analyses have been performed in corn for nitrogen (N) rate analysis using a variety of statistical approaches. Ridge regression using data from 47 N response trials has been used for N rate trials across the Corn Belt, demonstrating that soil hydrology was importing for predicting differences in optimum rate between trials (Qin et al., 2018). Comparing methods for predicting optimum N rate using a similar dataset of 49 rate trials in the Corn Belt indicated that random forest had the highest predictive ability, but that decision trees could have comparable predictive strength while using fewer variables. Multiple regression methods, including LASSO and elastic net, were limited by multicollinearities (Ransom et al., 2019).

Current state nutrient recommendations, individual fertilizer rate trials, and meta-analyses all indicate that soybean nutrient need varies based on soil test level, pH, and soil texture (Anthony et al., 2012; Buczko et al., 2018; Kuchenbuch & Buczko, 2011; Laboksi & Peters, 2012). Classification and regression trees and random forest analysis are able to accurately predict crop response to fertilizer for large datasets including yield, management, soil, and terrain variables (Buczko et al., 2018). Multiple linear regressions provide more information about how continuous independent variables impact the dependent variable compared to decision trees, greatly aiding the interpretation of CARTs and random forest analysis (Kuchenbuch & Buczko, 2011; Ransom et al., 2019; Zheng et al., 2009).

Moving to Reevaluate Fertilizer Recommendations

It is important to regularly evaluate the accuracy of fertilizer recommendations since changes in climate, average yields, and soybean genetics can all influence when soybean yield is likely to increase in response to potassium fertilizer application. Many states use a similar nutrient management framework to Wisconsin, including AR, IA, IN, MI, MN, and OH, and recent trials in some of these states has indicated that that fertilizer rates that were effective when written may no longer be appropriate. In Ohio, current fertilizer rates do not effectively build STK, possibly due to crop removal being underestimated (Fulford & Culman, 2018). Current rate recommendations in KS are too low to maintain STK levels, indicating that their crop removal rates may also be underestimates (Matz, 2012). In IL, the opposite is true—a survey of grain at elevators indicated that potassium crop removal estimates are too high (Nafziger, 2017). Arkansas fertilizer response trials indicate that current recommendations predict that more fields will respond to fertilizer than are responsive, leading to many “false positive” recommendations for both soybeans and rice (Fryer, Slaton, Roberts, Hardke, et al., 2019; Fryer, Slaton, Roberts, & Ross, 2019). Wisconsin has not extensively reevaluated their soybean fertilizer recommendations in at least 3 decades. Evaluating the efficacy of current potassium recommendations in the state is important to maintaining soybean yields and profitability for farmers throughout the state.

Works Cited

Agriculture and Horticulture Development Board. (2021). Nutrient Management Guide (RB209) | AHDB. https://ahdb.org.uk/nutrient-management-guide-rb209

American Soybean Association. (2019). 2019 SoyStats. https://soygrowers.com/wp-content/uploads/2019/10/Soy-Stats-2019_FNL-Web.pdf

Anthony, P., Malzer, G., Sparrow, S., & Zhang, M. (2012). Soybean Yield and Quality in Relation to Soil Properties. *Agronomy Journal*, 104(5), 1443–1458. https://doi.org/10.2134/agronj2012.0095

Antonangelo, J. A., Firmano, R. F., Alleoni, L. R. F., Oliveira, A., & Zhang, H. (2019). Soybean Production under Continuous Potassium Fertilization in a Long-Term No-Till Oxisol. *Agronomy Journal*, 111(5), 2462–2471. https://doi.org/10.2134/agronj2019.02.0084

Avendaño, F., Pierce, F. J., & Melakeberhan, H. (2004). Spatial analysis of soybean yield in relation to soil texture, soil fertility and soybean cyst nematode. *Nematology*, 6(4), 527–545. https://doi.org/10.1163/1568541042665313

Bender, R. R., Haegele, J. W., & Below, F. E. (2015). Nutrient Uptake, Partitioning, and Remobilization in Modern Soybean Varieties. *Agronomy Journal,* 107(2), 563–573. https://doi.org/10.2134/agronj14.0435

Bolland, M. D. A., & Allen, D. G. (1998). Spatial variation of soil test phosphorus and potassium, oxalate-extractable iron and aluminum, phosphorus-retention index, and organic carbon content in soils of Western Australia. *Communications in Soil Science and Plant Analysis*, 29(3–4), 381–392. https://doi.org/10.1080/00103629809369952

Breker, J. S., DeSutter, T., Rakkar, M. K., Chatterjee, A., Sharma, L., & Franzen, D. W. (2019). Potassium Requirements for Corn in North Dakota: Influence of Clay Mineralogy. *Soil Science Society of America Journal*, 83(2), 429–436. https://doi.org/10.2136/sssaj2018.10.0376

Brooker, A. P., Lindsey, L. E., Culman, S. W., Subburayalu, S. K., & Thomison, P. R. (2017). Low Soil Phosphorus and Potassium Limit Soybean Grain Yield in Ohio. *Crop, Forage & Turfgrass Management*, 3(1), cftm2016.12.0081. https://doi.org/10.2134/cftm2016.12.0081

Buczko, U., van Laak, M., Eichler-Löbermann, B., Gans, W., Merbach, I., Panten, K., Peiter, E., Reitz, T., Spiegel, H., & von Tucher, S. (2018). Re-evaluation of the yield response to phosphorus fertilization based on meta-analyses of long-term field experiments. *Ambio,* 47(S1), 50–61. https://doi.org/10.1007/s13280-017-0971-1

Castle, M. H., Lubben, B. D., & Luck, J. D. (2016). Factors Influencing the Adoption of Precision Agriculture Technologies by Nebraska Producers. *Presentations, Working Papers, and Gray Literature: Agricultural Economics*, 1-25.

Cate, R. B., & Nelson, L. A. (1971). A Simple Statistical Procedure for Partitioning Soil Test Correlation Data Into Two Classes. *Soil Science Society of America Journal,* 35(4), 658–660. https://doi.org/10.2136/sssaj1971.03615995003500040048x

Clover, M. W., & Mallarino, A. P. (2013). Corn and Soybean Tissue Potassium Content Responses to Potassium Fertilization and Relationships with Grain Yield. *Soil Science Society of America Journal*, 77(2), 630–642. https://doi.org/10.2136/sssaj2012.0223

Conway, L. S., Yost, M. A., Kitchen, N. R., Sudduth, K. A., & Veum, K. S. (2018). Cropping System, Landscape Position, and Topsoil Depth Affect Soil Fertility and Nutrient Buffering. *Soil Science Society of America Journal*, 82(2), 382–391. https://doi.org/10.2136/sssaj2017.08.0288

Cox, F. R. (1992). Range in Soil Phosphorus Critical Levels with Time. *Soil Science Society of America Journal* 56(6), 1504-1509.

Cox, M. S., Gerard, P. D., Wardlaw, M. C., & Abshire, M. J. (2003). Variability of Selected Soil Properties and Their Relationships with Soybean Yield. *Soil Science Society of America Journal,* 67(4), 1296–1302. https://doi.org/10.2136/sssaj2003.1296

Culman, S., Fulford, A., Camberato, J., Steinke, K., Lindsey, L., LaBarge, G., Watters, H., Lentz, E., Haden, R., Richer, E., Herman, B., Hoekstra, N., Thomison, P., & Warncke, D. (2020). Tri-State Fertilizer Recommendations. 1-56.

Culman, S., Mann, M., Sharma, S., Saeed, M. T., Fulford, F., Lindsey, L. E., Brooker, A., Dayton, L., Braig, E., Warden, R., Steinke, K., Camberto, J., & Joern, B. (2019). Converting between Mehlich-3, Bray P, and Ammonium Acetate Soil Test Values. https://ohioline.osu.edu/factsheet/anr-75

Culman, S, Saeed, M. T., & Fulford, A. (2017). Ohio Data that Shaped the Tri-State Fertilizer Recommendations. https://ohioline.osu.edu/factsheet/agf-518

Dodd, J. R., & Mallarino, A. P. (2005). Soil-Test Phosphorus and Crop Grain Yield Responses to Long-Term Phosphorus Fertilization for Corn-Soybean Rotations. *Soil Science Society of America Journal*, 69(4), 1118–1128. https://doi.org/10.2136/sssaj2004.0279

Eliason, R., Goos, R. J., Hoskins, B., Joern, B., Mallarino, A., Mengel, D., Dahl, J., Nathan, M., Shaver, T., Franzen, D., Culman, S, Gelderman, R., & Peters, J. (2015). Recommended chemical soil test procedures for the North Central Region, the USA. North Central Regional Research Publication No. 221 (Revised), 1–76.

Erickson, B. J., & Lowenberg-DeBoer, J. (2020). 2020 precision agriculture dealership survey. 1-21. https://ag.purdue.edu/digital-ag-resources/wp-content/uploads/2020/11/CropLife-Report-2020.pdf

Esri Support. (2021). GIS Dictionary. https://support.esri.com/en/other-resources/gis-dictionary

Farmaha, B. S., Fernández, F. G., & Nafziger, E. D. (2011). No-Till and Strip-Till Soybean Production with Surface and Subsurface Phosphorus and Potassium Fertilization. *Agronomy Journal*, 103(6), 1862–1869. https://doi.org/10.2134/agronj2011.0149

Florence, A., Ransom, M., & Mengel, D. (2017). Potassium Fixation by Oxidized and Reduced Forms of Phyllosilicates. *Soil Science Society of America Journal,* 81(5), 1247–1255. https://doi.org/10.2136/sssaj2016.12.0420

Franzen, D. W., & Bu, H. (2018). North Dakota Clay Mineralogy Impacts Crop Potassium Nutrition and Tillage Systems. NDSU Extension. 1-12.

Fryer, M. S., Slaton, N. A., Roberts, T. L., Hardke, J. T., & Norman, R. J. (2019). Validation of Soil-Test-Based Phosphorus and Potassium Fertilizer Recommendations for Flood-Irrigated Rice. *Agronomy Journal*, 111(5), 2523–2535. https://doi.org/10.2134/agronj2019.03.0159

Fryer, M. S., Slaton, N. A., Roberts, T. L., & Ross, W. J. (2019). Validation of Soil-Test-Based Phosphorus and Potassium Fertilizer Recommendations for Irrigated Soybean. *Soil Science Society of America Journal*, 83(3), 825–837. https://doi.org/10.2136/sssaj2019.02.0032

Fu, W., Tunney, H., & Zhang, C. (2010). Spatial variation of soil nutrients in a dairy farm and its implications for site-specific fertilizer application. *Soil and Tillage Research*, 106(2), 185–193. https://doi.org/10.1016/j.still.2009.12.001

Fulford, A. M., & Culman, S. W. (2018). Over-Fertilization Does Not Build Soil Test Phosphorus and Potassium in Ohio. *Agronomy Journal*, 110(1), 56–65. https://doi.org/10.2134/agronj2016.12.0701

Fulton, J. P., Shearer, S. A., & S.F. Higgins. (2001). Pattern Assessment of a Spinner Disc Variable-Rate Fertilizer Applicator. 2001 Sacramento, CA July 29-August 1,2001. https://doi.org/10.13031/2013.7350

Fulton, J. P., Shearer, S. A., Higgins, S. F., & McDonald, T. P. (2013). A method to generate and use as-applied surfaces to evaluate variable-rate fertilizer applications. *Precision Agriculture*, 14(2), 184–200. https://doi.org/10.1007/s11119-012-9286-1

Fulton, J., & Port, K. (2016). Physical Properties of Granular Fertilizers and Impact on Spreading. https://ohioline.osu.edu/factsheet/fabe-5501

Gaspar, A. P., Laboski, C. A. M., Naeve, S. L., & Conley, S. P. (2017). Phosphorus and Potassium Uptake, Partitioning, and Removal across a Wide Range of Soybean Seed Yield Levels. *Crop Science*, 57(4), 2193–2204. https://doi.org/10.2135/cropsci2016.05.0378

Hajjarpoor, A., Soltani, A., Zeinali, E., Kashiri, H., Aynehband, A., & Vadez, V. (2018). Using boundary line analysis to assess the on-farm crop yield gap of wheat. *Field Crops Research*, 225, 64–73. https://doi.org/10.1016/j.fcr.2018.06.003

Heckman, J. R. (2012). Soil Fertility Test Interpretation Phosphorus, Potassium, Magnesium, and Calcium. Rutgers Cooperative Extension, FS719, 1-4.

Hengl, T. (2007). A practical guide to geostatistical mapping of environmental variables. Publications Office, Luxembourg. ISBN #: 978-92-79-06904-8

Hoskins, B. R. (1997). Soil testing handbook for professionals in agriculture, horticulture, nutrient and residuals management (Third). University of Maine. https://umaine.edu/soiltestinglab/wp-content/uploads/sites/227/2016/07/handbook.pdf

Jiang, P., & Thelen, K. D. (2004). Effect of Soil and Topographic Properties on Crop Yield in a North-Central Corn–Soybean Cropping System. *Agronomy Journal*, 96(1), 252–258. https://doi.org/10.2134/agronj2004.0252

Kitchen, N. R., Sudduth, K. A., Myers, D. B., Drummond, S. T., & Hong, S. Y. (2005). Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity. *Computers and Electronics in Agriculture*, 46(1–3), 285–308. https://doi.org/10.1016/j.compag.2004.11.012

Kravchenko, A. N., & Bullock, D. G. (2000). Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties*. Agronomy Journal* 92, (9).

Kuchenbuch, R. O., & Buczko, U. (2011). Re-visiting potassium- and phosphate-fertilizer responses in field experiments and soil-test interpretations by means of data mining. *Journal of Plant Nutrition and Soil Science*, 174(2), 171–185. https://doi.org/10.1002/jpln.200900162

Laboksi, C. A. M., & Peters, J. B. (2012). Nutrient Application Guidelines for Field, Vegetable, and Fruit Crops in Wisconsin (A2809). UW Cooperative Extension. 1-94.

Landrum, C., Castrignanò, A., Mueller, T., Zourarakis, D., Zhu, J., & De Benedetto, D. (2015). An approach for delineating homogeneous within-field zones using proximal sensing and multivariate geostatistics. *Agricultural Water Management*, 147, 144–153. https://doi.org/10.1016/j.agwat.2014.07.013

Lark, R. M., Gillingham, V., Langton, D., & Marchant, B. P. (2020). Boundary line models for soil nutrient concentrations and wheat yield in national-scale datasets. *European Journal of Soil Science*, 71(3), 334–351. https://doi.org/10.1111/ejss.12891

Lawrence, H. G., & Yule, I. J. (2007). Estimation of the in-field variation in fertiliser application. *New Zealand Journal of Agricultural Research*, 50(1), 25–32. https://doi.org/10.1080/00288230709510279

Macnack, N. (2017). Fertilization Based on Sufficiency, Build-up and Maintenance Concept. Oklahoma Cooperative Extension, PSS-2266, 1-4.

Mallarino, A. (2013). A General Guide for Crop Nutrient and Limestone Recommendations in Iowa. Iowa State University Extension and Outreach, 1-14.

Mallarino, A. P. (1996). Spatial Variability Patterns of Phosphorus and Potassium in No-Tilled Soils for Two Sampling Scales. *Soil Science Society of America Journal,* 60(5), 1473–1481. https://doi.org/10.2136/sssaj1996.03615995006000050027x

Mallarino, A. P., & Wittry, D. J. (2004). Efficacy of Grid and Zone Soil Sampling Approaches for Site-Specific Assessment of Phosphorus, Potassium, pH, and Organic Matter. *Precision Agriculture*, 5(2), 131–144. https://doi.org/10.1023/B:PRAG.0000022358.24102.1b

Mangiafico, S. (2017). rcompanion: Functions to support extension education program evaluation (R Package Version 1.5.7.) [Computer software]. Rutgers Coop. Ext., New Brunswick, NJ. https://cran.r-project.org/package=rcompanion

Matz, J. D. (2012). Correction of potassium deficiency in soybean and corn production in southeast Kansas [Thesis, Kansas State University]. https://krex.k-state.edu/dspace/handle/2097/13415

McGraw, T. (1994). Soil Test Variability in Southern Minnesota. *Better Crops,* 78(4), 24–25.

Mengel, K., & Busch, R. (1982). The importance of the potassium buffer power on the critical potassium level in soils. *Soil Science*, 133(1), 27–32. https://doi.org/10.1097/00010694-198201000-00005

Nafziger, E. (2017, September 7). New Grain Phosphorus and Potassium Numbers. *Farmdoc*. https://farmdoc.illinois.edu/field-crop-production/uncategorized/new-grain-phosphorus-and-potassium-numbers.html

Nelson, L. A., & Anderson, R. L. (1977). Partitioning of Soil Test-Crop Response Probability. In T. R. Peck, J. T. Cope, & D. A. Whitney (Eds.), ASA Special Publications (pp. 19–38). American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. https://doi.org/10.2134/asaspecpub29.c2

Piepho, H.-P., Richter, C., Spilke, J., Hartung, K., Kunick, A., & Thöle, H. (2011). Statistical aspects of on-farm experimentation. *Crop and Pasture Science,* 62(9), 721. https://doi.org/10.1071/CP11175

Qin, Z., Myers, D. B., Ransom, C. J., Kitchen, N. R., Liang, S., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernandez, F. G., Franzen, D. W., Laboski, C. A. M., Malone, B. D., Nafziger, E. D., Sawyer, J. E., & Shanahan, J. F. (2018). Application of Machine Learning Methodologies for Predicting Corn Economic Optimal Nitrogen Rate. *Agronomy Journal,* 110(6), 2596–2607. https://doi.org/10.2134/agronj2018.03.0222

Ransom, C. J., Kitchen, N. R., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernández, F. G., Franzen, D. W., Laboski, C. A. M., Myers, D. B., Nafziger, E. D., Sawyer, J. E., & Shanahan, J. F. (2019). Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations. *Computers and Electronics in Agriculture,* 164, 104872. https://doi.org/10.1016/j.compag.2019.104872

Rattalino Edreira, J. I., Andrade, J., Mourtzinis, S., Roth, A. C., Grassini, P., & Conley, S. P. (2019). Benchmarking Soybean Production Systems in the North Central US. https://coolbean.info/wp-content/uploads/sites/3/2019/02/2019_Soybean_Benchmarking_ReviewFinal.pdf

Robertson, M. J., Llewellyn, R. S., Mandel, R., Lawes, R., Bramley, R. G. V., Swift, L., Metz, N., & O'Callaghan, C. (2012). Adoption of variable rate fertiliser application in the Australian grains industry: Status, issues and prospects. *Precision Agriculture*, 13(2), 181–199. https://doi.org/10.1007/s11119-011-9236-3

Schimmelpfennig, D. (2016). Farm Profits and Adoption of Precision Agriculture. USDA Economic Research Report Number 217, 1-46.

Schmid, C. J., Clarke, B. B., & Murphy, J. A. (2018). Potassium Nutrition Affects Anthracnose on Annual Bluegrass. *Agronomy Journal*, 110(6), 2171–2179. https://doi.org/10.2134/agronj2018.03.0147

Shatar, T. M., & Mcbratney, A. B. (2004). Boundary-line analysis of field-scale yield response to soil properties. *The Journal of Agricultural Science*, 142(5), 553–560. https://doi.org/10.1017/S0021859604004642

Singer, J. W., Kohler, K. A., Liebman, M., Richard, T. L., Cambardella, C. A., & Buhler, D. D. (2004). Tillage and compost affect yield of corn, soybean, and wheat soil fertility. *Agronomy Journal* 96 (7).

Singh, S., Savoy, H. J., Yin, X., Schneider, L., & Jagadamma, S. (2019). Phosphorus and Potassium Fertilizer Rate Verification for a Corn–Wheat–Soybean Rotation System in Tennessee. *Agronomy Journal*, 111(4), 2060–2068. https://doi.org/10.2134/agronj2018.12.0749

Smidt, E. R., Conley, S. P., Zhu, J., & Arriaga, F. J. (2016). Identifying Field Attributes that Predict Soybean Yield Using Random Forest Analysis. *Agronomy Journal*, 108(2), 637–646. https://doi.org/10.2134/agronj2015.0222

Stammer, A. J., & Mallarino, A. P. (2018). Plant Tissue Analysis to Assess Phosphorus and Potassium Nutritional Status of Corn and Soybean. *Soil Science Society of America Journal,* 82(1), 260–270. https://doi.org/10.2136/sssaj2017.06.0179

Tisdale, S. L., Nelson, W. L., & Beaton, J. D. (1985). Soil Fertility and Fertilizers (Fourth). Macmillon Publishing Company.

USDA NASS. (2019). 2017 Census of Agriculture United States Summary and State Data. AC-17-A-51, 1(51).

USDA NASS. (2020). USDA/NASS 2020 State Agriculture Overview for Wisconsin. Quick Stats. https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=WISCONSIN

Van Scoyoc, G. E. (2004). Predictability of Yield Response to P and K Application Based on Soil Test Levels. Indiana Crop Advisor Conference. https://www.agry.purdue.edu/CCA/2004/PDF/VanScoyoc.pdf

Virk, S. S., Mullenix, D. K., Sharda, A., Hall, J. B., Wood, C. W., Fasina, O. O., McDonald, T. P., Pate, G. L., & Fulton, J. P. (2013). Case Study: Distribution Uniformity of a Blended Fertilizer Applied Using a Variable-Rate Spinner-Disc Spreader. *Applied Engineering in Agriculture*, 627–636. https://doi.org/10.13031/aea.29.9774

Vitosh, M. L., Johnson, J. W., & Mengel, D. B. (1995). Tri-state Fertilizer Recommendations for Corn, Soybeans, Wheat and Alfalfa. Ohio State University Extension, 1-24.

Williams, A. S., Parvej, Md. R., Holshouser, D. L., Frame, W. H., & Reiter, M. S. (2017). Correlation of Field-Moist, Oven-Dry, and Air-Dry Soil Potassium for Mid-Atlantic USA Soybean. *Soil Science Society of America Journal, 81(6), 1586–1594. https://doi.org/10.2136/sssaj2016.10.0324*

Williams, A. S., Parvej, Md. R., Holshouser, D. L., Frame, W. H., & Reiter, M. S. (2018). Correlation and Calibration of Soil-Test Potassium from Different Soil Depths for Full-Season Soybean on Coarse-Textured Soils. *Agronomy Journal*, 110(1), 369–379. https://doi.org/10.2134/agronj2017.06.0344

Wittry, D. J., & Mallarino, A. P. (2004). Comparison of Uniform- and Variable-Rate Phosphorus Fertilization for Corn–Soybean Rotations. *Agronomy Journal* 96 (8).

Zheng, H., Chen, L., Han, X., Zhao, X., & Ma, Y. (2009). Classification and regression tree (CART) for analysis of soybean yield variability among fields in Northeast China: The importance of phosphorus application rates under drought conditions. *Agriculture, Ecosystems & Environment*, 132(1–2), 98–105. https://doi.org/10.1016/j.agee.2009.03.004

Chapter 2: Analyzing Historic Management Records to Evaluate the Soil Test Potassium – Soybean Grain Yield Relationship

Abstract

Potassium is a critical nutrient for plant growth and soybean production, and potassium availability can vary based on the physical and chemical soil properties. The objectives of this study were to (1) estimate the yield-maximizing (YMK) soil test potassium (STK) level for soybeans in the study area and (2) determine whether the YMK varies across the study area in relation to management or environmental factors. A database of yield maps, management history, and soil test results was collected from fields in southern WI, totaling 1080 observations. Quantile regression was used to estimate the YMK across all observations in the database, which was 76 ppm. Variables of interest (VOI) that may be associated with differences in YMK value were identified and used to cluster the database into two groups per each identified variable of interest. The YMK was estimated for each cluster, and higher YMK was associated with lower buffer pH, lower organic matter, higher normal height, and lower temperature. Normal height is a scaled measure of elevation where locations with higher normal height are at higher elevations. When planning future trials to update nutrient management recommendations, placing trials on at sites with different organic matter and buffer pH levels could help to better predict the impact of potassium fertilizer application on soybean yield and STK levels.

Abbreviations

Available water storage, AWS; cation exchange capacity, CEC; diammonium phosphate, DAP; digital elevation models, DEM; national commodity crop productivity index, NCCPI; soil organic carbon, SOC; soil test potassium, STK; variable of interest, VOI; yield-maximizing soil test potassium, YMK

Introduction

Potassium is a critical nutrient for plant growth and soybean production, and soybean fields regularly receive potassium fertilizer application in Wisconsin. Soils of different textures and cation exchange capacities have different optimum soil test potassium (STK) levels. Current potassium fertilizer recommendations in WI classify loamy soils with STK levels of 101-130 ppm as within the optimum range, whereas the optimum STK range for sandy soils is lower, only 66-90 ppm (Laboksi & Peters, 2012).

Recommendations state that fields with soil test results within the optimum range should replace the nutrients lost through crop removal, or approximately 0.023 kg K_2O_5 per kg of soybean grain. On these fields within the optimum range, a yield response to fertilizer application is expected 20% of the time. These current WI soybean fertilizer recommendations have not been extensively updated or evaluated in at least three decades. These recommendations may be agronomically sound, as other states have similar recommendations (Culman et al., 2020; Kaiser et al., 2011; A. Mallarino, 2013). However, as other states reevaluate their recommendations there is increasing evidence that some older fertilizer recommendations may no longer be appropriate.

The Ohio State University has changed their potassium critical level to 100 ppm regardless of soil texture or cation exchange capacity (CEC) (Culman et al., 2020). Older nutrient management guidelines in OH varied the critical level between 88 and 150 ppm based on soil CEC (Vitosh et al., 1995). North Dakota State University updated statewide potassium recommendations in 2018 to vary between fields based on their ratio of smectite clays to illite clays. Fields with a higher percentage of smectite clays have a higher estimated STK critical level (Franzen, 2018; Franzen & Bu, 2018). In Arkansas, current fertilizer recommendations predict that a higher percentage of soybean fields will respond to potassium application than has been observed through recent field trials, which suggests that current state nutrient management guidelines lead to fertilizer being applied to fields that are unlikely to respond to fertilizer application (Fryer, Slaton, Roberts, Hardke, et al., 2019; Fryer, Slaton, Roberts, & Ross, 2019).

Reevaluating nutrient management recommendations in Wisconsin helps ensure profitability for soybean farmers under current crop management conditions and a changing climate. Factors such as crop rotation, tillage, texture, CEC, pH, minerology, and terrain may impact K need (Buczko et al., 2018; Conway et al., 2018; Culman et al., 2020; D. W. Franzen & Bu, 2018; Kuchenbuch & Buczko, 2011; Vitosh et al., 1995). Evaluating recommendations using traditional small-plot trials would be time intensive and costly, and it would limit the number of environmental covariates investigated.

In this trial we use a novel data mining approach to understand which environmental factors impact the relationship between soybean yield and soil test potassium. The objectives of this study were to (1) estimate the yield-maximizing STK level for soybeans in the study area and (2) determine whether the yield-maximizing STK level varies across the study area in relation to management or environmental factors.

Methods

Data Collection

Nutrient management plans, soybean yield maps, soil test results, and variable rate application prescriptions for lime, diammonium phosphate (DAP), and potassium chloride (potash) were collected from farmers through their local co-op, which manages the custom fertilizer and pesticide applications for the fields included in this study. Yield data were collected between 2014 and 2018, and management data were collected from 2013 through 2018. Soil test potassium and soil test phosphorous were measured using Bray extraction methods, and other soil properties were measured using methods in accordance with (Eliason et al., 2015).

The 1/3 arc-second resolution digital elevation models (DEMs) were downloaded from USGS National Map for the entire study region (USGS, 2019). Additional terrain properties, listed in Table 2.1, were calculated from the DEMs using the SAGA plugin for QGIS (QGIS Development Team, 2020).

Table 2.1. Definitions of the terrain properties used in this study. All terrain properties were calculated from 1/3 arcsecond digital elevation models (DEMs).

Terrain Property	Description	Reference
Elevation	Vertical distance from sea level, calculated from a digital elevation model that represents the surface of the earth	(Hengl et al., 2007)
Slope	Rate of change of elevation	(Peckham, 2011)
Aspect	Cardinal direction, measured in degrees, that the prevailing slope faces	(Peckham, 2011)

Hillshade	Measure of how much light hits the soil surface when the light originates from the northwest at an azimuth of 45◦ from Earth's surface	(Conrad, 2015)
Slope Height	Measure of elevation compared to the lowest point within its channel network	(Conrad, 2015)
Midslope Position	Identifies the mean elevation within a channel network	(Conrad, 2015)
Valley Depth	Measure of elevation at a certain location as compared to the highest ridge within its channel network	(Conrad, 2015)
Standard Height	Rescales elevation within a neighborhood such that mean = 0 and a standard deviation = 1	(Conrad, 2015)
Normal Height	Rescales elevation within a neighborhood so that all values fall between 0 and 1	(Conrad, 2015)
Length and Steepness Factor	A value that combines a slope's total length with its steepness to estimate erosion	(Conrad, 2015)
General Curvature	Difference between profile curvature and plan curvature	(Conrad, 2015)
Minimum Curvature	Smallest value of curvature in any plane at that location	(*Geomorphic Curvature*, 2021; Olaya, 2009)
Maximum Curvature	Largest value of curvature in any plane at that location	(*Geomorphic Curvature*, 2021; Olaya, 2009)
Plan Curvature	Measure of concavity (rate of change of slope) along the contour line, also described as change of surface aspect	(Blaga, 2012; *Geomorphic Curvature*, 2021; Peckham, 2011)
Profile Curvature	Measure of concavity (rate of change of slope) along the same direction as the steepest slope	(Conrad, 2015; *Geomorphic Curvature*, 2021)
Tangential Curvature	Measure of concavity (rate of change of slope) perpendicular to the direction of steepest slope	(Blaga, 2012)
Longitudinal Curvature	Measure of concavity along the plane normal to both slope and aspect direction	(*Geomorphic Curvature*, 2021)
Cross Curvature	Measure of concavity along the plane normal to slope and perpendicular to aspect direction	(*Geomorphic Curvature*, 2021)
Flow Line Curvature	Describes the rate of flow line twisting along the horizontal plane; also known as streamline curvature	(Wu et al., 2020)
Total Curvature	Sum of the second derivative of elevation in the horizontal plane and two planes normal to the horizontal	(Conrad, 2015; Minár et al., 2020)
Topographic Wetness Index	An index that incorporates slope, upstream drainage area, and width of the upstream area	(Conrad, 2015)

Soil map unit data for the state of Wisconsin, including available water storage (AWS), soil organic carbon (SOC), national commodity crop productivity index (NCCPI), maximum rooting depth, soil taxonomy, temperature regime, drainage class, and water regime, were downloaded from the gSSURGO database using the USDA Geospatial Data Gateway in November, 2021 (USDA NRCS, 2021).

Monthly total precipitation and average maximum daily temperature were downloaded from the Daymet THREDDS server as 1km grid summaries for May through August of each year where yield was collected (2014 – 2018) (Thornton et al., 2020).

Yield Data Processing

Farmers provided their yield data to researchers as shapefiles, after performing pre-processing as recommended by their yield monitor and combine manufacturers. Pre-processing steps include correcting grain flow, combine header-up, and start-pass delays. More information on these types of delays and the importance of correcting for them can be found in Simbahan et al. (2004). Farmers also provided field boundary shapefiles for each field.

Shapefiles of yield data were further processed in two steps in QGIS 3.4.14 (QGIS Development Team, 2020) to (1) remove yield monitor observations within 38 m of the field boundary and (2) remove observations where combine speed was outside of 3 standard deviations of the field average or yield was outside of 4 standard deviations of the field average.

Tenth-hectare (0.1 ha) squares were centered over each grid soil sampling location. Mean yield and mean fertilizer application rate was calculated within each square. Values for terrain attributes, temperature, and precipitation were also averaged within each 0.1 ha square.

Database Structure

Each location in this database has soil test potassium measured at exactly one timepoint, and between one and four years of soybean yield data. The database is formatted so that each observation (row) is a single yield observation, and multiple observations (rows) correspond to the same physical location in space. Both, the calendar year soil sampling and yield data was collected and were included as independent variables (columns) in this database. Management decisions including fertilizer application and tillage are coded within the database based on their proximity in time to the yield data collection. For instance, three columns in this database are yield (Yield), year of yield observation (Yield Year), and potash application in the fall immediately preceding the yield observation (Potash_YY). A yield observation that took place in 2016 would have fall 2015 potash rates in column Potash_YY, and a yield observation that took place in 2017 would have fall 2016 potash rates in column Potash_YY. For all variables, the suffix "_YY" indicates that the management took place during the same growing season as the yield observation or in the immediately preceding fall. The suffix "_YYM1" indicates that the management practice took place one year before management practices in columns with the suffix "_YY." Similarly, the suffixes "_YYM2" and "_YYM3" indicate management practices from 2 or 3 years, respectively, before practices in columns with the suffix "_YY." Temperature and precipitation data was only included for the year of each yield observation, not previous growing seasons.

There are 1080 observations in the complete database. In total, the database included 76 continuous variables (pH, buffer pH, soil test potassium, soil test phosphorous, soil test calcium, soil test magnesium, soil cation exchange capacity, soil sampling year, organic matter loss on ignition, soybean yield, year yield was collected, year soil samples were collected, multi-year average soybean yield, Lime_YY, Lime_YYM1, Lime_YYM2, Lime_YYM3, Lime_YYM4, DAP_YY, DAP_YYM1, DAP_YYM2, DAP_YYM3, DAP_YYM4, Potash_YY, Potash_YYM1, Potash_YYM2, Potash_YYM3, Potash_YYM4, StarterFertilizer_YY, StarterFertilizer_YYM1, StarterFertilizer_YYM2, elevation, slope, aspect, hillshade , slope height, midslope position, valley depth, standard height, normal height, length and steepness factor, general curvature, minimum curvature, maximum curvature, plan curvature, profile curvature, tangential curvature, longitudinal curvature, cross curvature, flow line curvature, total curvature, terrain wetness index, latitude, longitude, available water storage (AWS) 0-5 cm, AWS 5 -20 cm, AWS 20-50 cm, AWS 50-100 cm, AWS 0-20 cm, AWS 0-30 cm, AWS 0 -100 cm; AWS 0 – 999cm, soil organic carbon (SOC) 0-5cm, SOC 5-20 cm, SOC 20-50 cm, SOC 50 -100 cm, SOC 0-20 cm, SOC 0-30 cm, SOC 0-100 cm, SOC 0 – 999 cm, NCCPI corn, NCCPI soybean, NCCPI small grains, maximum rooting depth, summer precipitation, summer temperature), ten binary variables (recent tillage; FallTillage _YY, FallTillage_YYM1, FallTillage_YYM2, FallTillage_YYM3, FallTillage_YYM4, SpringTillage_YY, SpringTillage_YYM1, SpringTillage_YYM2, SpringTillage_YYM3), and 20 categorical variables (rotation, map unit, farmer, Crop_YYM1, Crop_YYM2, Crop_YYM3, Crop_YYM4, , drainage class, taxonomic class, taxonomic order, taxonomic suborder, taxonomic great group, taxonomic subgroup, taxonomic particle size, taxonomic particle size mod, taxonomic CEC activity class, taxonomic reaction, taxonomic temperature class, taxonomic moisture class, taxonomic temperature regime). Note that starter fertilizers include any combination of P, K, or N applied at plant, and NCCPI stands for national commodity crop productivity index.

Analysis Methods

Fourteen observations where STK values fell outside of three standard deviations from the database mean STK and were removed from further analysis. 1066 observations remained in the database for analysis after outlier removal (Figure 1). For binary and categorical variables, independent variables where 90% or more of fields had the same treatment were excluded from further analysis. Excluded variables include rotation, Crop_YYM2, Crop_YYM3, Crop_YYM4, FallTillage_YYM2, FallTillage_YYM3, FallTillage_YYM4, SpringTillage_YY, SpringTillage_YYM1, SpringTillage_YYM2, SpringTillage_YYM3, taxonomic particle size mod, taxonomic reaction, taxonomic temperature class, taxonomic moisture class, taxonomic temperature regime.

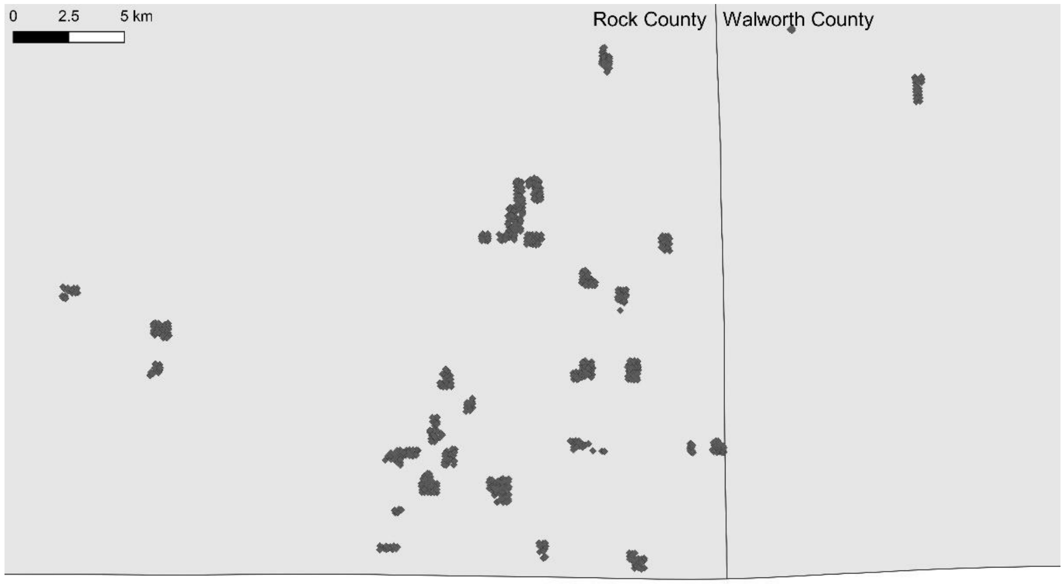

Figure 2.1. Map of Rock and Walworth counties in southern Wisconsin with locations of database observations shown as dark grey makers.

Data analysis in this trial included four main steps: (1) estimating the yield-maximizing STK (YMK) value across all observations in the database, (2) identify variables of interest that may be associated with differences in YMK value, (3) cluster the database into two groups for each identified variable of interest, and (4) estimate the YMK for each cluster of observations to determine which variables of interest were associated with differences in YMK.

Quadratic regression and boundary regression analysis were performed using quantile regression (95th percentile) to determine the relationship between soybean yield and STK. Boundary regression analyses, including quantile regression, are useful tools for understanding the yield-STK relationship in data sets without treatment effects since they can reduce the effect of other environmental factors on yield and better identify the impact of a single environmental factor on yield in a complex data set with many potential yield-limiting conditions (Cade & Noon, 2003; Lark et al., 2020). The x-coordinate of the maxima of the regression line was considered the YMK, and quantile regressions were calculated in R 4.1.2 using the package quantreg (Koenker et al., 2022; R Core Team, 2021).

Conditional inference (CI) trees and random forest (RF) analysis were used to identify variables of interest that may be associated with differences in YMK. Six models were run via both CI trees and RF analysis to identify and rank variables that may have an impact on yield and potassium availability (Table 2.2). Models were selected based on the categories of independent variables, where model A includes all possible independent variables, model B includes all independent variables except latitude and longitude, model C includes only independent variables originally sourced from the gSSURGO database, model D includes only independent variables related to elevation and terrain, model E includes only independent variables from soil test results, and model F includes independent variables related to management history. Using multiple models and two different methods allowed us to better identify variables that had a consistent impact on yield, without the effect of the most important handful of variables masking the effect of independent variables that had a smaller but consistent association with yield and reduced the impact individual algorithms on the variable selection process.

Table 2.2. Models used for variable selection via conditional inference trees and random forest analysis. The dependent variable for all models was soybean yield.

Model	Independent Variables
A	pH, buffer pH, soil test potassium, soil test phosphorous, soil test calcium, soil test magnesium, soil cation exchange capacity, soil sampling year, organic matter loss on ignition, soybean yield, year yield was collected, year soil samples were collected, multi-year average soybean yield, Lime_YY, Lime_YYM1, Lime_YYM2, Lime_YYM3, Lime_YYM4, DAP_YY, DAP_YYM1, DAP_YYM2, DAP_YYM3, DAP_YYM4, Potash_YY, Potash_YYM1, Potash_YYM2, Potash_YYM3, Potash_YYM4, StarterFertilizer_YY, StarterFertilizer_YYM1, StarterFertilizer_YYM2, elevation, slope, aspect, hillshade , slope height, midslope position, valley depth, standard height, normal height, length and steepness factor, general curvature, minimum curvature, maximum curvature, plan curvature, profile curvature, tangential curvature, longitudinal curvature, cross curvature, flow line curvature, total curvature, terrain wetness index, latitude, longitude, available water storage

	(AWS) 0-5 cm, AWS 5 -20 cm, AWS 20-50 cm, AWS 50-100 cm, AWS 0-20 cm, AWS 0-30 cm, AWS 0 -100 cm; AWS 0 – 999cm, soil organic carbon (SOC) 0-5cm, SOC 5-20 cm, SOC 20-50 cm, SOC 50 -100 cm, SOC 0-20 cm, SOC 0-30 cm, SOC 0-100 cm, SOC 0 – 999 cm, NCCPI corn, NCCPI soybean, NCCPI small grains, maximum rooting depth, summer precipitation, summer temperature, recent tillage, map unit, farmer, FallTillage _YY, FallTillage_YYM1, drainage class, taxonomic class, taxonomic order, taxonomic suborder, taxonomic great group, taxonomic subgroup, taxonomic particle size, taxonomic CEC activity class
B	pH, buffer pH, soil test potassium, soil test phosphorous, soil test calcium, soil test magnesium, soil cation exchange capacity, soil sampling year, organic matter loss on ignition, soybean yield, year yield was collected, year soil samples were collected, multi-year average soybean yield, Lime_YY, Lime_YYM1, Lime_YYM2, Lime_YYM3, Lime_YYM4, DAP_YY, DAP_YYM1, DAP_YYM2, DAP_YYM3, DAP_YYM4, Potash_YY, Potash_YYM1, Potash_YYM2, Potash_YYM3, Potash_YYM4, StarterFertilizer_YY, StarterFertilizer_YYM1, StarterFertilizer_YYM2, elevation, slope, aspect, hillshade , slope height, midslope position, valley depth, standard height, normal height, length and steepness factor, general curvature, minimum curvature, maximum curvature, plan curvature, profile curvature, tangential curvature, longitudinal curvature, cross curvature, flow line curvature, total curvature, terrain wetness index, available water storage (AWS) 0-5 cm, AWS 5 -20 cm, AWS 20-50 cm, AWS 50-100 cm, AWS 0-20 cm, AWS 0-30 cm, AWS 0 -100 cm; AWS 0 – 999cm, soil organic carbon (SOC) 0-5cm, SOC 5-20 cm, SOC 20-50 cm, SOC 50 -100 cm, SOC 0-20 cm, SOC 0-30 cm, SOC 0-100 cm, SOC 0 – 999 cm, NCCPI corn, NCCPI soybean, NCCPI small grains, maximum rooting depth, summer precipitation, summer temperature, recent tillage, map unit, farmer, FallTillage _YY, FallTillage_YYM1, drainage class, taxonomic class, taxonomic order, taxonomic suborder, taxonomic great group, taxonomic subgroup, taxonomic particle size, taxonomic CEC activity class
C	available water storage (AWS) 0-5 cm, AWS 5 -20 cm, AWS 20-50 cm, AWS 50-100 cm, AWS 0-20 cm, AWS 0-30 cm, AWS 0 -100 cm; AWS 0 – 999cm, soil organic carbon (SOC) 0-5cm, SOC 5-20 cm, SOC 20-50 cm, SOC 50 -100 cm, SOC 0-20 cm, SOC 0-30 cm, SOC 0-100 cm, SOC 0 – 999 cm, NCCPI corn, NCCPI soybean, NCCPI small grains, maximum rooting depth, map unit, drainage class, taxonomic class, taxonomic order, taxonomic suborder, taxonomic great group, taxonomic subgroup, taxonomic particle size, taxonomic CEC activity class
D	elevation, slope, aspect, hillshade , slope height, midslope position, valley depth, standard height, normal height, length and steepness factor, general curvature, minimum curvature, maximum curvature, plan curvature, profile curvature, tangential curvature, longitudinal curvature, cross curvature, flow line curvature, total curvature, terrain wetness index
E	pH, buffer pH, soil test potassium, soil test phosphorous, soil test calcium, soil test magnesium, soil cation exchange capacity
F	year soil samples were collected, multi-year average soybean yield, Lime_YY, Lime_YYM1, Lime_YYM2, Lime_YYM3, Lime_YYM4, DAP_YY, DAP_YYM1, DAP_YYM2, DAP_YYM3, DAP_YYM4, Potash_YY, Potash_YYM1, Potash_YYM2, Potash_YYM3, Potash_YYM4, StarterFertilizer_YY, StarterFertilizer_YYM1, StarterFertilizer_YYM2,

Conditional inference tree analysis was implemented using the R package partykit (Hothorn et al., 2021). The independence-test criterion for splits was univariate p value ($\alpha = .05$). Interior nodes were required to maintain at least 100 observations. At minimum, terminal nodes included 10 fields. Overfitting was prevented by constraining trees at a maximum depth of 10 nodes.

Random forest analysis was implemented using the R package randomForest (Cutler & Wiener, 2022). The number of trees in each RF model was 1000, and the number of independent variables are used to generate each individual tree was algorithmically tuned for each RF model using the command tuneRF within the package randomForest. Random forest analysis can only be performed on datasets without missing values. The full data set was reduced to only include complete cases using two steps: (1) independent variables that were missing for more than 10% of observations were excluded and (2) observations that still had missing values were removed (n = 1006, independent variables = 60). Within each RF model, independent variables were ranked using importance, a measure of node purity used for evaluating which independent variables contribute to model stability and predictive power.

The output CI trees and importance rankings from RF analysis were used to assign points to independent variables. A variable received a point for every time A) it has an importance ranking within 20% of the most important variable for that RF model or B) is as an internal node in a CI tree. The top 50% of point-earning variables were considered variables of interest for paired quantile regression analysis.

K-means clustering was used to separate the database into two clusters based on the value of each variable of interest that was identified via CI trees and RF analysis. K-means clustering was performed in R 4.1.2 and was run separately for each variable of interest. Variables of interest where k-means clustering resulted in a cluster that contains 20% or less than the database overall were excluded from further analysis.

After clustering the database using k-means, quadratic quantile regression was used to estimate YMK for each cluster of data, resulting in two YMK estimates associated with each variable of interest. When the two YMK estimates differed from each other by more than 13 ppm (10% of the current STK critical level in WI; Laboski and Peters, 2012), the variable of interest was considered to be associated with differences in YMK within the database.

Results

Across all observations in the database, the relationship between yield and STK for the 95th quantile was described by the parabola Yield = 82 + 0.03758K - 0.00025K^2 (Figure 2.2).

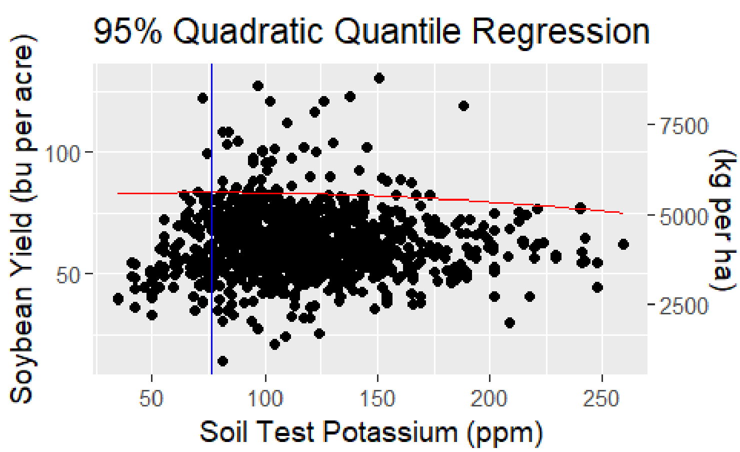

Figure 2.2. Scatterplot showing soybean yield at various soil test potassium (STK) levels. The red regression line represents the quadratic relationship between yield and STK for the 95% quantile of data, and the blue vertical line is at STK = 76 ppm, the yield maximizing STK level for this data set.

Through conditional inference trees and random forest analysis, 16 variables of interest were identified, 14 of which were continuous (Table 2.3). Two variables that did not have continuous distributions, farmer and recent tillage, were summarized separately (Table 2.4).

Table 2.3. Mean [max, min] of yield (kg ha^{-1}), soil test potassium (STK; ppm), and the variable of interest (VOI) for each cluster based on continuous variables of interest. Where appropriate, the unit for each VOI is listed parenthetically following the VOI, and n denotes the number of observations from the total database included in each cluster. For lime and potash variables, the suffix "_YY" indicates that the management took place during the same growing season as the yield observation or in the immediately preceding fall. The suffix "_YYM1" indicates that the management practice took place one year before management practices in columns with the suffix "_YY." Similarly, the suffixes "_YYM2" and "_YYM3" indicate management practices from 2 or 3 years, respectively, before practices in columns with the suffix "_YY."

		Cluster 1				Cluster 2			
Variable of Interest Category	**Variable of Interest (unit)**	**n**	**Mean yield [min yield, max yield]**	**Mean STK [min STK, max STK]**	**Mean of VOI [min of VOI, max of VOI]**	**n**	**Mean yield [min yield, max yield]**	**Mean STK [min STK, max STK]**	**Mean of VOI [min of VOI, max of VOI]**
Soil Test Results	Buffer pH	427	4307 [1412, 8794]	118 [41, 229]	6.75	639	3985 [951, 7295]	113 [35, 259]	7.15
	Organic Matter (%)	432	4411 [1993, 8793]	132 [61, 259]	3.71	634	4037 [951, 7303]	109 [35, 248]	2.23
	Year soil samples were collected	848	4271 [951, 8793]	117 [41, 248]	2018.03	218	3868 [1412, 5561]	124 [35, 259]	2014.73
gSSURGO Data	National commodity crop productivity index—soybean*	154	-	-	-	912	-	-	-
Terrain	Flow line curvature*	6	-	-	-	1060	-	-	-
	Normal height	488	4236 [951, 8594]	116 [46, 242]	0.25	578	4063 [1412, 8793]	115 [35, 259]	0.76
	Slope height	870	-	-	-	196	-	-	-
Management	Lime_YY*	125	-	-	-	941	-	-	-
	Potash_YY(kg ha^{-1})	568	4182 [951, 8793]	119 [35, 248]	21.72	498	4196 [1412, 7303]	117 [35, 259]	258.18

	Potash_YYM1(kg ha^{-1})	817	4029 [951, 5991]	122 [35, 259]	22.72	249	4713 [2342, 8793]	107 [41, 248]	264.39
	Potash_YYM2 (kg ha^{-1})	699	4263 [951, 8793]	122 [35, 259]	13.78	367	4048 [2230, 6864]	111 [35, 241]	316.55
	Potash_YYM3*	933	-	-	-	133	-	-	-
Other	Temperature (°C)	697	3937 [951, 8793]	119 [35, 259]	65306.37	369	4664 [2458z, 7303]	117 [55, 240]	69163.69
	Year of yield observation	702	4167 [951, 7303]	116 [35, 248]	2016.82	364	4233 [1412, 8793]	122 [35, 259]	2014.58

* Denotes variables of interest where one cluster contains 20% or less of the overall database, and these variables of interest were excluded from further analysis

Table 2.4. Mean [max, min] of yield (kg ha^{-1}) and soil test potassium (STK; ppm), for each cluster based on the categorical variables of interest, farmer and recent tillage. n denotes the number of observations from the total database included in each cluster.

Cluster for categorical variables of interest	**n**	**Mean yield [min yield, max yield]**	**Mean STK [min STK, max STK]**
Farmer H Cluster	240	4654 [1993, 7303]	115 [41, 248]
Farmer L Cluster	328	3767 [1412, 5991]	101 [35, 224]
Farmer M Cluster	344	3983 [2369, 5542]	135 [55, 259]
No-Till Cluster	298	4054 [951, 7010]	122 [55, 248]
Tilled Cluster	604	4369 [1412, 8793]	117 [35, 259]

The quantile regressions for VOI had YMK values that varied by cluster, although not all clusters had a quadradic regression with a maxima value for estimating YMK. For the VOI with only one cluster that has a YMK value, the YMK value of the cluster was compared to the YMK value for the dataset overall.

The fields with higher buffer pH (Cluster 2; Figure 2.3 Panel D) had a QR model without a yield maximizing STK value. However, fields with lower pH (Cluster 2; Figure 2.3 Panel E) had a 46 ppm higher YMK than the dataset overall (Figure 2.2). The fields with higher organic matter (Cluster 1; Figure 2.3 Panel B) had an 11 ppm lower YMK than the fields with lower organic matter (Cluster 2; Figure 2.3 Panel E). Both YMK values for organic matter clusters were within 10 ppm of the YMK for the database overall. The YMK for observations where STK was measured in 2017 – 2019 was 113 ppm (Cluster 1; Figure 2.3 Panel C), and the YMK for STK observations where was measured in 2014-2016 had a YMK of 173 ppm (Cluster 2, Figure 2.3 Panel F).

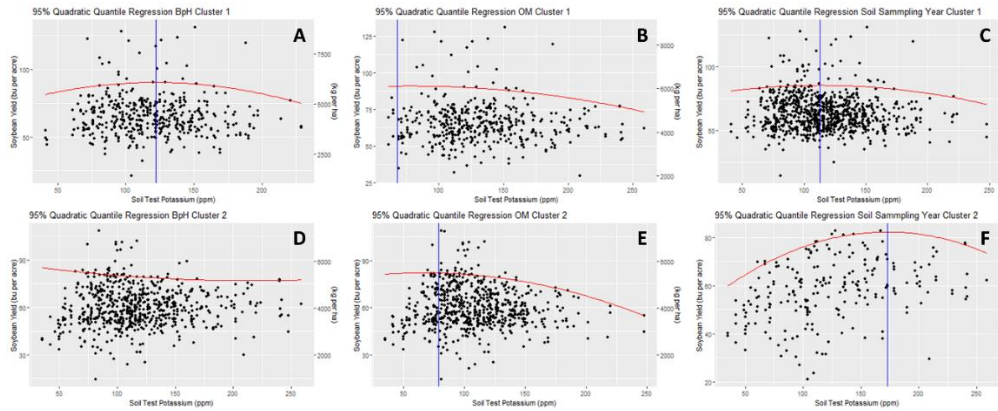

Figure 2.3. Scatterplot showing soybean yield at various soil test potassium (STK)levels for different clusters of data. Cluster information can is available in Table 1. The red regression line represents the quadratic relationship between yield and STK for the 95% quantile of data, and the blue vertical line is the yield maximizing STK level (YMK) for this data set.

Normal height is a terrain metric that compares elevation at a given raster cell to neighboring measures of elevation, and locations with lower normal height represent the low-elevation locations within a field (Conrad, 2015). Lower normal height (Cluster 1; Figure 2.4 Panel A) was associated with a 31 ppm lower YMK than higher normal height (Cluster 2; Figure 2.4 Panel B).

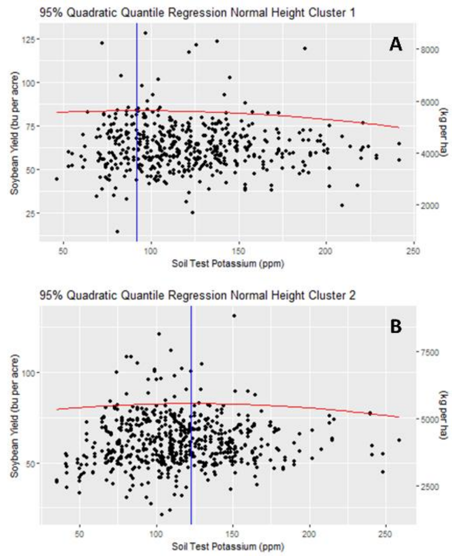

Figure 2.4. Scatterplot showing soybean yield at various soil test potassium (STK)levels for the two clusters of data determined by normal height, a terrain characteristic. Cluster information is available in Table 1. The red regression line represents the quadratic relationship between yield and STK for the 95% quantile of data, and the blue vertical line is the yield maximizing STK level (YMK) for this data set.

The YMK for Farmer M was 58 ppm, and for Farmer L it was 201 ppm (Figure 2.5). Due to the shape of the parabola for Farmer H, a YMK could not be calculated.

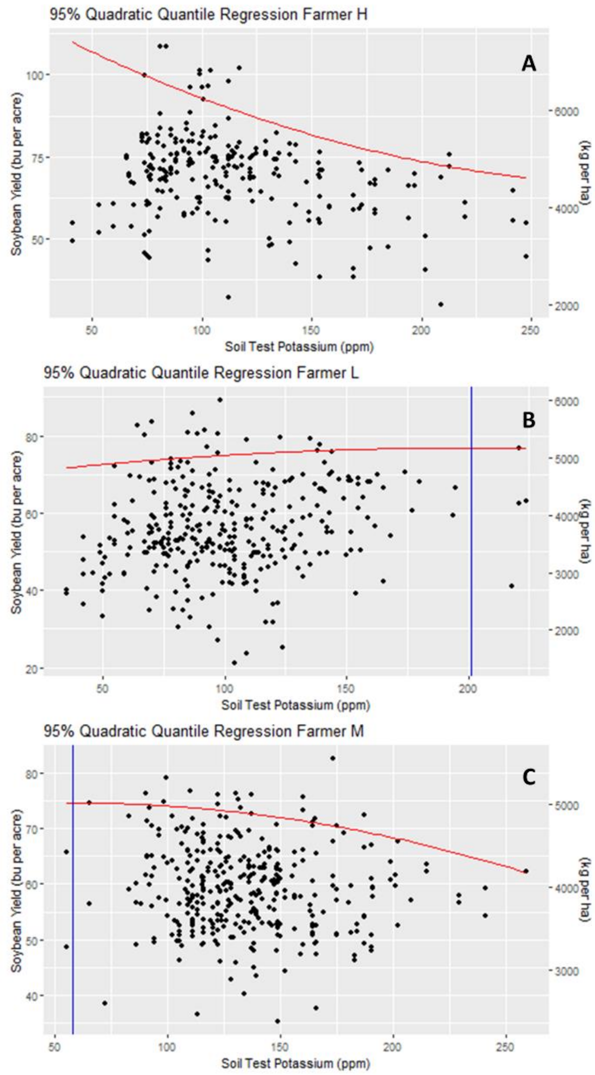

Figure 2.5. Scatterplot showing soybean yield at various soil test potassium (STK) levels for the clusters of data determined by Farmer. Cluster information can is available in Table 1. The red regression line represents the quadratic relationship between yield and STK for the 95% quantile of data, and the blue vertical line is the yield maximizing STK level (YMK) for this data set.

The YMK for no-till locations it was 143 ppm, and for tilled locations was 119 ppm (Figure 2.6, Panels A and B, respectively). Approximately 160 locations did not have tillage information, so not all observations in the database are encompassed by these two figures. Additional information about the number and distribution of observations is available in Table 2.2.

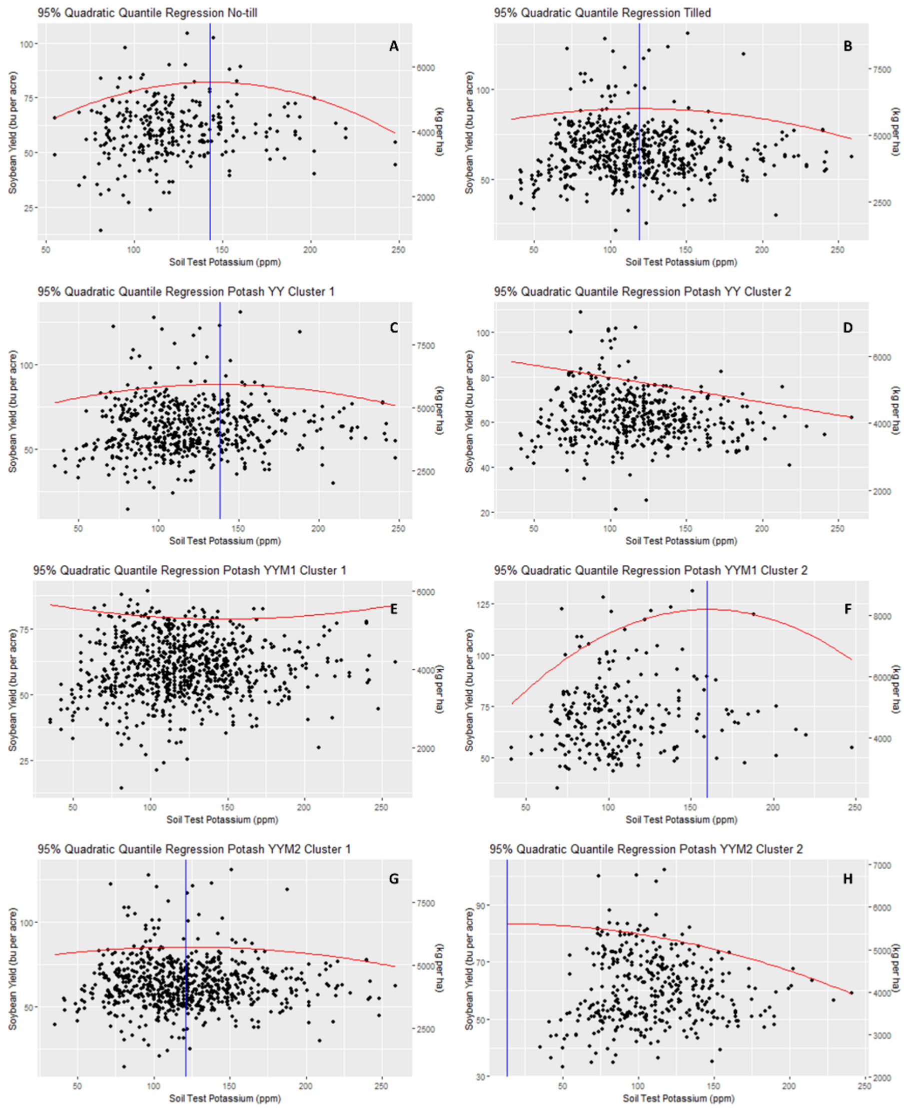

Figure 2.6. Scatterplot showing soybean yield at various soil test potassium (STK) levels for the clusters of data determined by past potash application and recent tillage. Cluster information is available in Table 1. The red regression line represents the quadratic relationship between yield and STK for the 95% quantile of data, and the blue vertical line is the yield maximizing STK level (YMK) for this data set.

For potash application in the fall immediately preceding yield data collection (YY), YMK at low potash locations (Cluster 1; Figure 2.6 Panel C) was 132 ppm, which is much higher than the YMK for the dataset overall (Figure 2.2). YMK could not be calculated for higher potash locations (Cluster 2; Figure 2.6 Panel D). When higher rates of potash were applied two falls preceding yield data collection (YYM1), the calculated YMK was 160 ppm (Cluster 2; Figure 2.6

Panel F)—nearly double the YMK of the dataset overall. YMK could not be calculated for fields with lower potash application (Cluster 1; Figure 2.6 Panel E).

Fields with lower potash application 3 falls before yield data collection took place (YYM2; Cluster 1; Figure 2.6 Panel G) have a YMK estimate of 121 ppm. For fields with higher potash application in that year, YMK was estimated to be 13 ppm. Since the lowest value of STK in this dataset was 35 ppm, the YMK estimate of 13 ppm would be extrapolation and is not considered an accurate estimate of YMK (Cluster 2; Figure 2.6 Panel H).

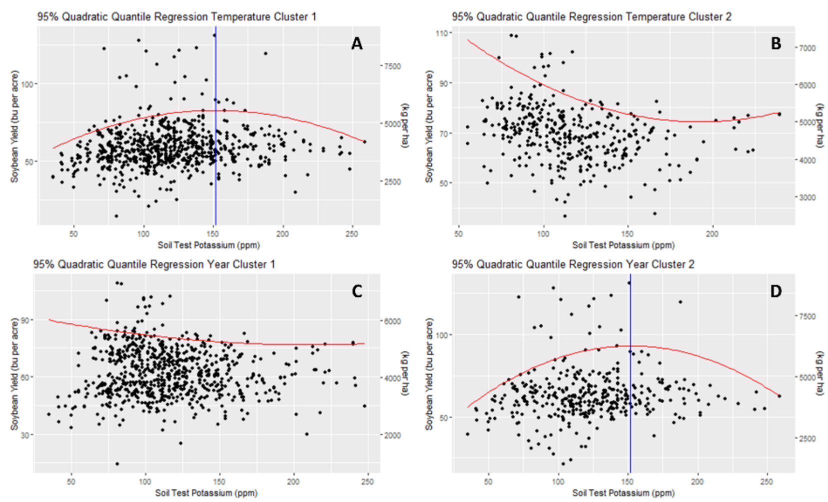

Figure 2.7. Scatterplot showing soybean yield at various soil test potassium (STK) levels for the clusters of data determined by temperature and year of yield observation. Cluster information is available in Table 1. The red regression line represents the quadratic relationship between yield and STK for the 95% quantile of data, and the blue vertical line is the yield maximizing STK level (YMK) for this data set.

The difference in cumulative summer temperature between Cluster 1 and Cluster 2 was less than 5% (Table 1). The YMK of cluster one, which had lower temperatures, was 152 ppm (Figure 2.7, Panel A). Yield observations that were collected in 2014-2016 also had a YMK of 152 ppm (Cluster 2; Figure 2.7, Panel D). The trends in temperature and year were likely similar due to annual trends in weather being a larger driver than spatial variation in temperature across the fields in this study.

Discussion

The YMK in this data set was slightly lower than the current soil test critical level in WI, which indicates that reviewing the critical level may be appropriate for the state. In this data set, environmental conditions that were associated with higher YMK include lower buffer pH, lower organic matter, higher normal height, and lower temperature. Most fields in this study had well-managed pH and had received lime in the last five years, so changes in pH management strategies are unlikely to change YMK. The differences in YMK between observations with different buffer pH levels are likely more related to variations in clay content and other physical properties that resist pH change than it is to pH concentration or lime history. The temperature range in this data set was narrow, and the inverse relationship between YMK and temperature may be different at temperatures higher or lower than those represented in this data set. Tilled fields had a lower YMK than no-till fields, which may be related to no-till fields having higher potassium ion concentration in shallow soil samples.

Lower potash application in the fall immediately preceding yield data collection (YY) was associated with higher YMK than the dataset overall, and higher rates of potash two falls preceding yield data collection (YYM1) was also associated with higher YMK than the dataset overall. These differences could be due to soil sampling timing relative to the yield observations, or due to changes in potassium variability based on whether it was applied in the fall immediately before soybean or in the fall previous to corn in a corn-soybean rotation. Small plot potassium fertilizer timing trials had previously focused on fall vs. spring application, but future studies over multiple growing seasons in WI would be valuable when planning nutrient management recommendations.

Results indicate that quadratic quantile regression is sensitive to changes in the input data set and can estimate differences in YMK without overfitting issues. Only one of the YMK estimates agronomically improbable (Figure 6 Panel H) at 13 ppm. Given that this database did not include sandy soils, the actual YMK is likely much higher. Even so, the fact that only one YMK estimate was unreasonable is a good indication that analysis through quadratic quantile regression is a promising method.

As is common with data sets that are collected without imposing treatments, it is difficult to separate the effect of different environmental conditions or management practices that are observed together. For instance, it is difficult to separate the impact of soil sampling year from the impact of specific fields being sampled at different intervals. A good example of this is the clustering variable Farmer (Figure 5), which is unlikely to be biologically important on its own, but clustering fields based on Farmer also groups fields based on management practices such as their year of soil sampling and whether pH was addressed with a single large lime application or multiple years of smaller lime applications. Additionally, Farmer can be a proxy for unmeasured variables that can impact data quality (e. g. combine operator, soil sampler, and yield monitor) and variables not measured that impact yield but not necessarily data quality (e. g. preferred herbicide program and brand of seed). These are common challenges with observational data analysis and emphasize the importance of follow up field trials to determine which of the associations identified in this study are due to causal relationships between environmental and management factors and the YMK.

Conclusions

Yield maximizing STK level was 76 ppm across this data set in southern WI, or around half of the current soil test critical level in WI. Differences in YMK were associated with soil and terrain properties such as buffer pH, organic matter, and normal height, which indicates that the critical level may not be consistent across the whole state. Fields with a higher YMK may respond to potassium fertilizer applications when STK is at a higher concentration than fields with a lower YMK. On high-YMK fields, applying potassium fertilizer rates lower than crop removal rates may lower yields, even if STK is in the above-optimum range. When planning future trials to update nutrient management recommendations, placing trials on at sites with different organic matter and buffer pH levels could help to better predict the relationship between soybean yield and STK application on fields that receive potash applications.

Works Cited

Blaga, L. (2012). Aspects regarding the significance of the curvature types and values in the studies of geomorphology assisted by GIS*. Analele Universităţii Din Oradea, Seria Geografie,* 327–337.

Buczko, U., van Laak, M., Eichler-Löbermann, B., Gans, W., Merbach, I., Panten, K., Peiter, E., Reitz, T., Spiegel, H., & von Tucher, S. (2018). Re-evaluation of the yield response to phosphorus fertilization based on meta-analyses of long-term field experiments. *Ambio*, 47(S1), 50–61. https://doi.org/10.1007/s13280-017-0971-1

Conrad. (2015). System of Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geoscientific Model Development, 8, 1991–2007.

Conway, L. S., Yost, M. A., Kitchen, N. R., Sudduth, K. A., & Veum, K. S. (2018). Cropping System, Landscape Position, and Topsoil Depth Affect Soil Fertility and Nutrient Buffering. *Soil Science Society of America Journal*, 82(2), 382–391. https://doi.org/10.2136/sssaj2017.08.0288

Culman, S., Fulford, A., Camberato, J., Steinke, K., Lindsey, L., LaBarge, G., Watters, H., Lentz, E., Haden, R., Richer, E., Herman, B., Hoekstra, N., Thomison, P., & Warncke, D. (2020). Tri-State Fertilizer Recommendations. Ohio State University Extension, 1-56.

Eliason, R., Goos, R. J., Hoskins, B., Joern, B., Mallarino, A., Mengel, D., Dahl, J., Nathan, M., Shaver, T., Franzen, D., Culman, S, Gelderman, R., & Peters, J. (2015). Recommended chemical soil test procedures for the North Central Region, the USA. North Central Regional Research Publication No. 221 (Revised), 1–76.

Franzen, D. (2018). North Dakota Fertilizer Recommendation Tables and Equations—Publications. NDSU Extension. https://www.ag.ndsu.edu/publications/crops/north-dakota-fertilizer-recommendation-tables-and-equations#section-14

Franzen, D. W., & Bu, H. (2018). North Dakota Clay Mineralogy Impacts Crop Potassium Nutrition and Tillage Systems. NDSU Extension, 1-12.

Fryer, M. S., Slaton, N. A., Roberts, T. L., Hardke, J. T., & Norman, R. J. (2019). Validation of Soil-Test-Based Phosphorus and Potassium Fertilizer Recommendations for Flood-Irrigated Rice. *Agronomy Journal*, 111(5), 2523–2535. https://doi.org/10.2134/agronj2019.03.0159

Fryer, M. S., Slaton, N. A., Roberts, T. L., & Ross, W. J. (2019). Validation of Soil-Test-Based Phosphorus and Potassium Fertilizer Recommendations for Irrigated Soybean. *Soil Science Society of America Journal*, 83(3), 825–837. https://doi.org/10.2136/sssaj2019.02.0032

Geomorphic Curvature. (2021). US Naval Academy. https://www.usna.edu/Users/oceano/pguth/md_help/html/geomorph_curvature.htm

Hengl, T. (2007). A practical guide to geostatistical mapping of environmental variables. Publications Office, Luxembourg. ISBN #: 978-92-79-06904-8

Kaiser, D., Lamb, J., & Eliason, R. (2011). Fertilizer guidelines for agronomic crops in Minnesota. University of Minnesota Extension.

Kuchenbuch, R. O., & Buczko, U. (2011). Re-visiting potassium- and phosphate-fertilizer responses in field experiments and soil-test interpretations by means of data mining. *Journal of Plant Nutrition and Soil Science*, 174(2), 171–185. https://doi.org/10.1002/jpln.200900162

Laboksi, C. A. M., & Peters, J. B. (2012). Nutrient Application Guidelines for Field, Vegetable, and Fruit Crops in Wisconsin (A2809). UW Cooperative Extension, 1-94.

Mallarino, A. (2013). A General Guide for Crop Nutrient and Limestone Recommendations in Iowa. Iowa State University Cooperative Extension, 1-14.

Minár, J., Evans, I. S., & Jenčo, M. (2020). A comprehensive system of definitions of land surface (topographic) curvatures, with implications for their application in geoscience modelling and prediction. *Earth-Science Reviews*, 211, 103414. https://doi.org/10.1016/j.earscirev.2020.103414

Olaya, V. (2009). Chapter 6 Basic Land-Surface Parameters. In Developments in Soil Science (Vol. 33, pp. 141–169). Elsevier. https://doi.org/10.1016/S0166-2481(08)00006-8

Peckham, S. D. (2011). Profile, Plan and Streamline Curvature: A Simple Derivation and Applications. 1-4.

QGIS Development Team. (2020). QGIS (3.4.14) [Computer software]. https://www.qgis.org/en/site/index.html

Simbahan, G. C., Dobermann, A., & Ping, J. L. (2004). Screening Yield Monitor Data Improves Grain Yield Maps. *Agronomy Journal*, 96(4), 1091–1102. https://doi.org/10.2134/agronj2004.1091

Thornton, M. M., Shrestha, R., Wei, Y., Thornton, P. E., Kao, S., & Wilson, B. E. (2020). DaymetDaymet: Monthly Climate Summaries on a 1-km Grid for North America, Version 4 [NetCDF, GTiff]. 0 MB. https://doi.org/10.3334/ORNLDAAC/1855

USDA NRCS. (2021). Gridded soil survey geographic (gSSURGO) database for Wisconsin. https://gdg.sc.egov.usda.gov/

USGS. (2019). The National Map. https://www.usgs.gov/core-science-systems/national-geospatial-program/national-map

Vitosh, M. L., Johnson, J. W., & Mengel, D. B. (1995). Tri-state Fertilizer Recommendations for Corn, Soybeans, Wheat and Alfalfa. 24.

Wu, Q., Chen, Y., Zhou, H., Chen, S., & Wang, H. (2020). A New Algorithm for Calculating the Flow Path Curvature (C) from the Square-Grid Digital Elevation Model (DEM). *ISPRS International Journal of Geo-Information*, 9(9), 510. https://doi.org/10.3390/ijgi9090510

Chapter 3: Field Evaluation of Soybean Response to Potassium Fertilizer

Abstract

Potassium fertilization is critical for soybean production in southern WI. On-farm potassium rate trials were established at 3 sites in southern WI during the 2021 growing season to compare soybean yield between treatments of 0 or 213 kg ha^{-1} potash (potassium chloride; 0-0-60). Yield did not differ between treatments or sites. Analysis via the Cate-Nelson procedure indicated that the critical K concentration in soybean leaf tissue was 2.04%. The relationship between soil test results from Bray-1 extraction and Mehlich-3 extraction for silty loam soils was represented by the linear regression line Bray = 0.77 * Mehlich - 0.75. The estimated soil test potassium critical level for soybean was 15 ppm lower than is currently recommended in WI nutrient management guidelines, which indicates that potassium fertilizer is being spread on some soybean fields that may have a low response or nonresponse to potash fertilizer application. These results suggest that a larger review of nutrient management guidelines would be timely in the state, especially as fertilizer costs continue to rise.

Abbreviations

Cation exchange capacity, CEC; soil test potassium, STK

Introduction

Potassium (K) is a critical nutrient for plant growth and supports transpiration and other plant functions (Tisdale et al., 1985). Soybean plants uptake K primarily from the soil solution, and nutrients are removed from the field with the soybean grain at harvest. Without adequate fertilization, available nutrients in the soil will become depleted over time (Tisdale et al., 1985).

Current K fertilizer recommendations in WI classify loamy soils with soil test potassium (STK) levels of 101-130 ppm as within the optimum range. The optimum STK range for sandy soils is 66-90 ppm in WI (Laboksi & Peters, 2012). Within the optimum range, a yield response to fertilizer application is expected 20% of the time. Recommended fertilizer rates for fields within the optimum soil test range are expected to replace the nutrients lost through crop removal, or approximately 0.023 kg K_2O_5 per kg of soybean grain (Laboksi & Peters, 2012).

Wisconsin fertilizer guidelines have not been extensively evaluated or updated in at least three decades (Laboksi & Peters, 2012). Many states have similar recommendations, including OH, IN, MI, IA, MN, and ND, although WI is the only US state that recommends using Bray-1 extraction methods to measure STK (Franzen, 2018; Fulford & Culman, 2018; Kaiser et al., 2011; Laboksi & Peters, 2012; Mallarino, 2013).

Three US states have found that some fertilizer application recommendations that were effective when recommendations were written may no longer be appropriate due to changes in STK critical level. The Ohio State University has changed their potassium critical level to 100 ppm regardless of soil texture or cation exchange capacity (CEC), and now recommends using Mehlich-3 extraction to analyze soil samples (Culman et al., 2020). Older nutrient management guidelines in OH varied the critical level between 88 and 150 ppm based on soil CEC and used ammonium acetate extraction for analyzing soil samples (Vitosh et al., 1995). North Dakota State University updated statewide K recommendations in 2018 to vary between fields based on the ratio of smectite clays to illite clays. Fields with a higher percentage of smectite clays have a higher estimated STK critical level (Franzen, 2018; Franzen & Bu, 2018).

In Arkansas, current fertilizer recommendations predict that a higher percentage of soybean fields will respond to K application than has been observed through recent field trials, which suggests that current state nutrient management guidelines lead to “false positive” K fertilizer application recommendations (Fryer, Slaton, Roberts, Hardke, et al., 2019; Fryer, Slaton, Roberts, & Ross, 2019).

The objectives of this research were to (1) calculate the soil test K critical level for soybean production in southern WI, (2) calculate the difference in critical level between soil samples extracted with Bray-1 and Mehlich-3 extraction, and (3) estimate the critical K concentration for soybean leaf tissue.

Methods

Field Trial Design

Field trials were established at three Rock County, WI sites in 2020 for growing season 2021. All sites were established on silty loam soils, and the number of replications varied based on field size (Table 3.1).

Table 3.1. Soil properties and number of replications for each site.

	Site A	**Site B**	**Site C**
Replications	9	5	8
Predominant Soil Series	Plano Silt Loam	Mahalasville Silt Loam	Plano Silt Loam
Taxonomy	Fine-silty, mixed, superactive, mesic typic Argiudolls	Fine-silty, mixed, superactive, mesic typic Argiaquolls	Fine-silty, mixed, superactive, mesic typic Argiudolls

pH	6.3	6.1	6.1
Organic Matter (%)	4.1	5.3	4.7

The trial included two treatments: (1) 213 kg ha^{-1} potash (potassium chloride; 0-0-60) application or (2) no potash applied. Plots were arranged in a randomized complete block design where all plots were 146 m long. Treatment (1) plots (213 kg ha^{-1} potash applied) were 110 m wide to allow for 3 passes of the fertilizer applicator within each plot. Treatment (2) plots (no potash applied) were 37 m wide (Figure 3.1).

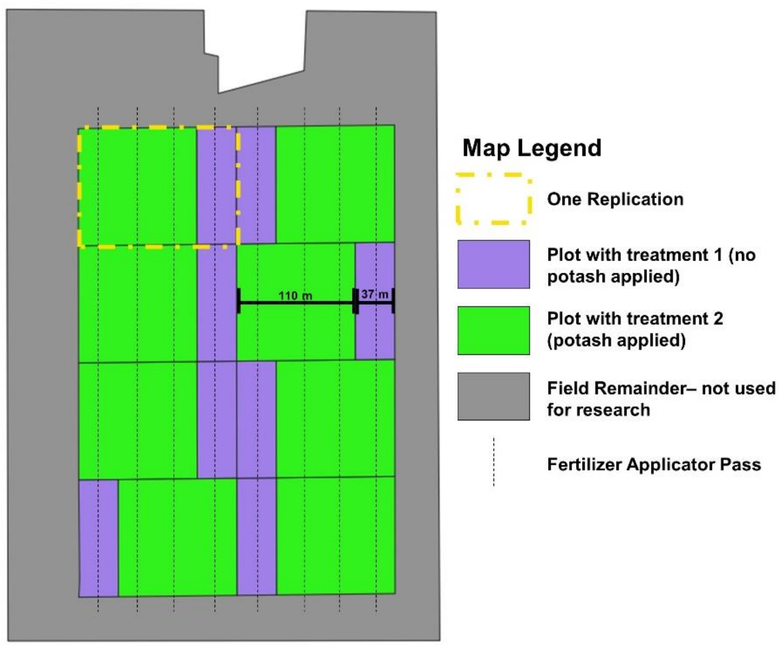

Figure 3.1. Field map showing 8 replications of the 2-treatment potash trial layout at Site C. Plot size and fertilizer applicator pass width were the same at all sites.

Potash was applied in November 2021 using a commercial 37 m spinner spreader that is owned and operated by the local co-op. All fields were planted within a 76-cm row spacing in the first week of May 2022 by farmer cooperators in each field. Site B experienced freeze damaged 4 weeks after planting and was replanted in the first week of June. Farmer cooperators managed each field with herbicides and crop protectant products throughout the season. All replications within the same field received the same pesticide applications.

In November 2021, composite soil samples were collected from two locations within every plot before fertilizer was applied (Figure 3.2). Each composite sample included 10, 1.9-cm wide, 20.3-cm deep soil cores. Soil samples were air dried and shipped to A&L Great Lakes Laboratories in Fort Wayne, IN for analysis. Soil test potassium was measured using both Bray-1 and Mehlich-3 extraction methods. Organic matter was measured using loss-on-ignition. Full method details can be found in the Recommended Chemical Soil Test Procedures for the North Central Region (Eliason, et al., 2015).

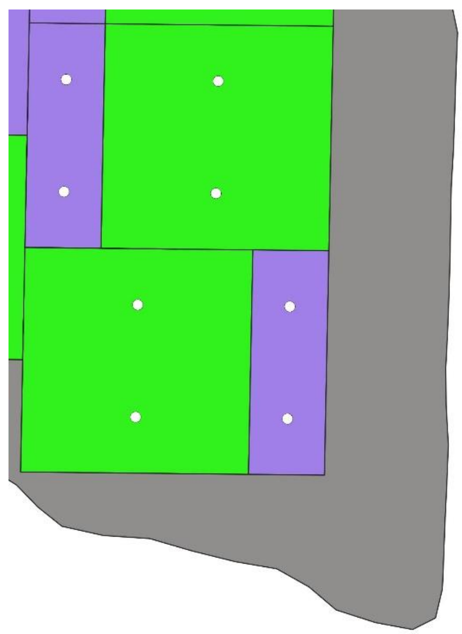

Figure 3.2. White circles represent the two soil and tissue sampling locations within each plot.

Tissue samples were collected at the R1 growth stage at each of the locations used for soil sampling. At each sampling location, the newest fully-expanded trifoliate leaf was collected from 20 plants. Samples were dried in paper bags (dryer temperature 38–54 ˚C) until constant mass was achieved and shipped to A&L Great Lakes Laboratories for nutrient concentration analysis.

Yield Data Collection

Yield data were collected by farmer cooperators using combine-mounted yield monitors with moisture sensors, and moisture values were used to adjust wet yield measurements to 130 g kg^{-1} moisture concentration. Farmers exported the data as a shapefile to share with researchers. Yield monitor flow delays were set during the export process and visually verified when the shapefile was opened in QGIS3 for further processing (QGIS Development Team, 2020). Yield observations that fell within 30.5 m of the field boundary were removed from further analysis, and further filtered to remove yield observations that were outside of three standard deviations of combine speed or dry yield. Less than 3% of observations were removed by the standard deviation filters.

The mean of yield monitor observations within the center 24 m by 116 m of each plot was calculated, and that mean value was used to represent the whole plot in further analyses (Figure 3.3). This area is smaller than the full plot size so that observations within 6 m of plot borders in the direction of combine travel or within 15 m of fertilizer applicator rate changes are not included in the plot average yield. This procedure ensures that no yield observations were collected from areas with known causes of potassium rate variation (partial rates from the spinner-spreader changing rates or from observations falling near the outer ranges of the spinner-spreader application width).

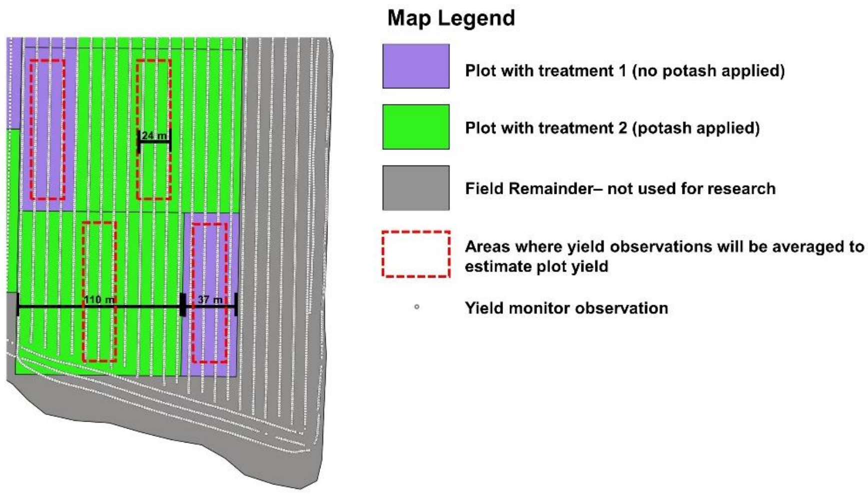

Figure 3.3. Map of an example field trial with yield monitor observations displayed as grey circles. Observations are averaged within the red dashed rectangles to estimate plot yields.

Analysis Methods

Yield data were collected from 44 plots in total, and data from all sites were analyzed together. Mixed-model ANOVA was performed using R 4.1.2 (R Core Team, 2021) and the package *lme4* to assess yields differed between treatments. Treatment and field were considered fixed variables, and replication nested within field was considered a random variable. Degrees of freedom were estimated using Kenward–Rogers approximation to account for unequal replication among fields. Data were not transformed, and residuals were plotted to assess for normality.

The relationship between Bray-1 and Mehlich-3 extraction was quantified using linear regression in R 4.2.1 (R Core Team, 2021). To calculate relative grain yield within each replication, the yield of the 0 kg ha^{-1} potash treatment plot was divided by the yield of the fertilized treatment plot within each experimental block, and then the quotient was multiplied by 100. Critical STK level and critical tissue K concentration were calculated by comparing relative grain yield to the STK and tissue K concentration, respectively, using the Cate-Nelson procedure in the package rcompanion (Mangiafico, 2017). Soil test potassium values and tissue K concentrations from all four sampling locations within each replication were averaged before proceeding with critical STK level and critical tissue K concentration calculations. In cases where the Cate-Nelson procedure identified more than one critical level that minimizes sum of squares, the highest value was selected.

Results and Discussion

Soybean Yield

Soybean yield did not differ among sites, nor did yield between potash and no potash treatments (Table 3.2) ($p = 0.826$). Since yield did not vary among sites and there was potash treatment by site interaction, data from all sites were pooled for subsequent analyses.

Table 3.2. Analysis of variance to predict difference of yield among potassium rates (K), site (S), and their interaction.

	F-value	p-value
Potassium (K)	0.05	0.826
Site (S)	3.40	0.056
K × S	0.12	0.883

Soil Test Results

The relationship between STK results from Bray-1 extraction and STK results from Mehlich-3 extraction was linear ($p < 0.001$; Adj. $R^2 = 0.91$). For each ppm of K^+ ions removed from the soil sample by Mehlich-3 extraction, Bray-1 removed approximately 0.77 ppm of K^+ ions (Figure 3.4).

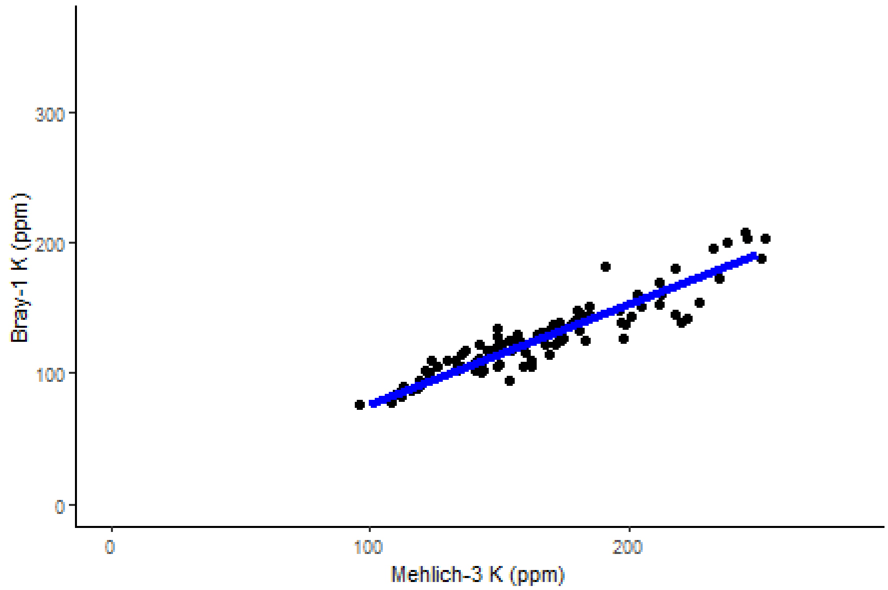

Figure 3.4. Scatter comparing two potassium extraction methods, Bray-1 and Mehlich-3. The linear regression line Bray = 0.77 * Mehlich - 0.75 ($p < 0.001$; Adj. $R^2 = 0.91$) .

The close correlation between Bray-1 and Mehlich-3 extraction methods suggests that in the future, WI nutrient management recommendations could be updated to utilize the more common Mehlich-3 test without reducing the accuracy of the nutrient management recommendations for silty loam soils with pH between 6 and 6.5, which is common in southern WI. Before recommending statewide adoption of Mehlich-3 extraction for K, additional studies should be performed to determine whether the same linear relationship holds for other soil textures or pH levels.

The soil test critical level, when STK was measured using Bray-1 extraction procedures, was estimated to be 115 ppm (p = 0.48) (Figure 3.5). This is slightly lower than the current critical level 0f 130 ppm for soybean grown on loamy soils in WI, but it does fall within the current “optimum” soil test category (101-130 ppm) (Laboksi & Peters, 2012).

When Mehlich-3 extraction was used to measure K concentration, the soil test critical level was estimated to be 145 ppm (p = 1). Since Mehlich-3 extraction tended to remove more K^+ ions from soil samples than Bray-1 extraction (Figure 3.4), it makes sense that the estimated critical level was higher when calculated using STK values measured using Mehlich-3 extraction procedures.

The Cate-Nelson plots for both Bray-1 and Mehlich-3 extraction (Figures 3.5 and 3.6, respectively) are split into four quadrants. Observations that fall into quadrants II or IV (filled circles) are accurately classified in Cate-Nelson plots, and observations that are in quadrants I or III (open circles) are considered misclassified. Observations in quadrant I suggest increased risk of soybean nonresponse to applied potash. Quadrant IV observations suggest increased risk that potash would not be applied to responsive soybean. The Cate-Nelson plot for Bray-1 samples has six misclassified observations, five of which fall into quadrant I and the remaining one in quadrant III. Cate-Nelson plot for Mehlich-3 samples has five misclassified observations in quadrant I and eight in quadrant III.

Since the critical level estimate that minimizes sum of squares fell within the current optimum category based on current state nutrient management guidelines and neither Cate-Nelson plot used to predict STK critical level was statistically significant, this trial does not provide evidence that the current STK critical level is no longer appropriate for the three sites studied. A larger study testing a wider range of environments would be needed to update the guidelines or set a new state critical level.

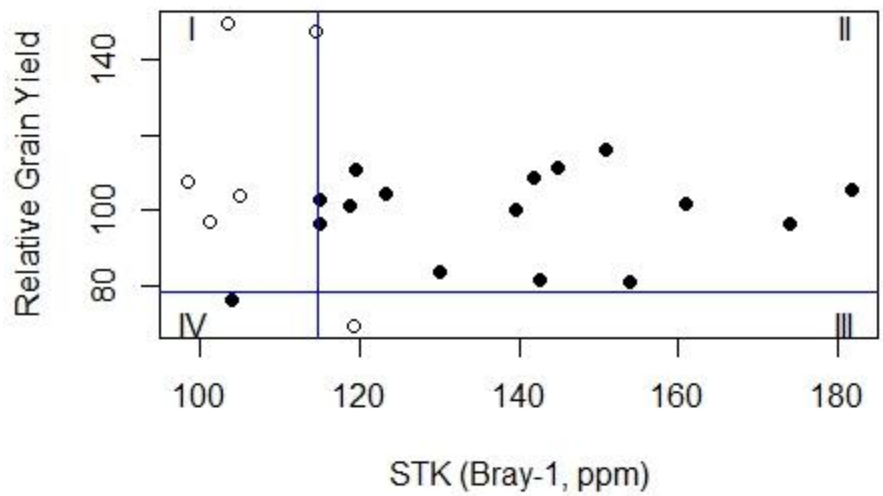

Figure 8. Cate-Nelson figure that shows the estimated soil test potassium (STK) Critical Level to be 115 ppm for soybeans grown in Southern WI in 2021 (p = 0.481). Soil samples were processed using Bray-1 extraction procedures, the current state standard in WI.

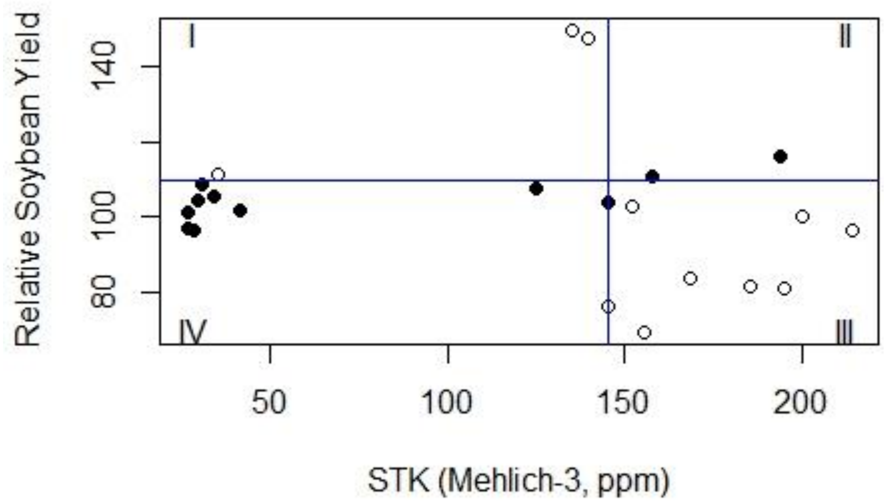

Figure 3.6. Cate-Nelson figure that shows the estimated soil test potassium (STK) Critical Level to be 145 ppm for soybeans grown in southern WI in 2021 (p = 1). Soil samples were processed using Mehlich-3 extraction procedures, which are the current standard in many commercial soil laboratories.

Tissue Nutrient Concentration

Across these three sites, tissue K concentration varied less than in the soil test values. The critical K concentration in soybean leaf tissue was estimated to be 2.04 % (p = 0.004), and all observations were classified into quadrants II and IV by the Cate-Nelson plot (Figure 3.7). Current WI nutrient management recommendations do not include a critical K concentration for soybean tissue, but similar tissue K critical concentrations (15.6 – 19.9 g K kg^{-1} or 1.56 - 1.99%, respectively) were observed by Stammer and Mallarino (2018).

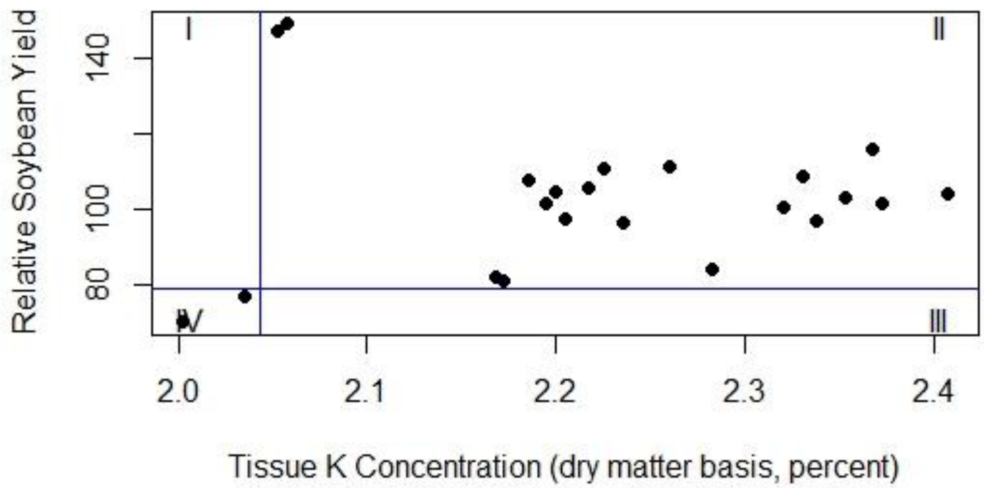

Figure 3.7. Cate-Nelson figure that shows the estimated critical K concentration in leaf tissue of 2.04 % for soybeans grown in southern WI in 2021 (p = 0.004).

Conclusions

A major limitation of this study is the sample size and variation in field properties. Although the three fields included in this study are reflective of southern WI soybean growing environments, they are not necessarily reflective of the state at large. Soybean grown in the Central Sands region would likely have a lower soil test critical level and may have a different relationship between STK values measured via Bray-1 and Mehlich-3 extraction.

The estimated STK critical level for soybean was 15 ppm lower in this study than is currently recommended in the WI state nutrient management guidelines, which indicates that there may be risk that K fertilizer is being spread on soybean fields that are likely nonresponsive to applied K. A larger review of nutrient management guidelines would be timely in the state, especially as fertilizer costs continue to rise.

Works Cited

Culman, S., Fulford, A., Camberato, J., Steinke, K., Lindsey, L., LaBarge, G., Watters, H., Lentz, E., Haden, R., Richer, E., Herman, B., Hoekstra, N., Thomison, P., & Warncke, D. (2020). Tri-State Fertilizer Recommendations. Ohio State University Extension, 1-56.

Franzen, D. (2018). North Dakota Fertilizer Recommendation Tables and Equations—Publications. NDSU Extension. https://www.ag.ndsu.edu/publications/crops/north-dakota-fertilizer-recommendation-tables-and-equations#section-14

Franzen, D. W., & Bu, H. (2018). North Dakota Clay Mineralogy Impacts Crop Potassium Nutrition and Tillage Systems. NDSU Extension, 1-12.

Fryer, M. S., Slaton, N. A., Roberts, T. L., Hardke, J. T., & Norman, R. J. (2019). Validation of Soil-Test-Based Phosphorus and Potassium Fertilizer Recommendations for Flood-Irrigated Rice. *Agronomy Journal*, 111(5), 2523–2535. https://doi.org/10.2134/agronj2019.03.0159

Fryer, M. S., Slaton, N. A., Roberts, T. L., & Ross, W. J. (2019). Validation of Soil-Test-Based Phosphorus and Potassium Fertilizer Recommendations for Irrigated Soybean. *Soil Science Society of America Journal*, 83(3), 825–837. https://doi.org/10.2136/sssaj2019.02.0032

Fulford, A. M., & Culman, S. W. (2018). Over-Fertilization Does Not Build Soil Test Phosphorus and Potassium in Ohio. *Agronomy Journal*, 110(1), 56–65. https://doi.org/10.2134/agronj2016.12.0701

Kaiser, D., Lamb, J., & Eliason, R. (2011). Fertilizer guidelines for agronomic crops in Minnesota. University of Minnesota Extension.

Laboksi, C. A. M., & Peters, J. B. (2012). Nutrient Application Guidelines for Field, Vegetable, and Fruit Crops in Wisconsin (A2809). UW Cooperative Extension, 1-94.

Mallarino, A. (2013). A General Guide for Crop Nutrient and Limestone Recommendations in Iowa. Iowa State University Cooperative Extension, 1-14.

Tisdale, S. L., Nelson, W. L., & Beaton, J. D. (1985). Soil Fertility and Fertilizers (Fourth). Macmillon Publishing Company.

Vitosh, M. L., Johnson, J. W., & Mengel, D. B. (1995). Tri-state Fertilizer Recommendations for Corn, Soybeans, Wheat and Alfalfa. 24.

Chapter 4: Management Strategies for early and late-planted soybean in the North Central US

This chapter was previously published in Agronomy Journal.

Matcham, E. G.[1], Mourtzinis, S.[2], Conley, S. P.[1], Rattalino Edreira, J. I.[3], Grassini, P.[3], Roth, A. C.[1], Casteel, S. N.[4], Ciampitti, I. A.[5], Kandel, H. J.[6], Kyveryga, P. M.[7], Licht, M. A.[8], Mueller, D. S.[9], Nafziger, E. D.[10], Naeve, S. L.[11], Stanley, J.[6], Staton, M. J.[12], and Lindsey, L. E.[13] (2020). Management strategies for early and late-planted soybean in the North Central US. Agronomy Journal 112(4): 2928-2943. https://doi.org/10.1002/agj2.20289

[1] Department of Agronomy, University of Wisconsin-Madison, Madison, WI 53706, USA
[2] Agstat Consulting, Athens, Greece
[3] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583-0915, USA
[4] Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA
[5] Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA
[6] Department of Plant Sciences, North Dakota State University, Fargo, ND 58108-6050, USA
[7] Iowa Soybean Association, Ankeny, IA 50023, USA
[8] Department of Agronomy, Iowa State University, Ames, IA 50011-1010, USA
[9] Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA 50011, USA
[10] Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA
[11] Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA
[12] Department of Crop and Soil Sciences, Michigan State University, East Lansing, MI 48824, USA
[13] Department of Horticulture and Crop Science, The Ohio State University, Columbus, OH 43210, USA

Abstract

It is widely recognized that planting soybean *[Glycine max* (L.) *Merr.*] early is critical to maximizing yield, but the influence of changing management factors when soybean planting is delayed is not well understood. The objectives of this research were to (a) identify management decisions that increase seed yield in either early- or late-planted soybean scenarios, and (b) estimate the maximum break-even price of each management factor identified to influence soybean seed yield in early- or late-planted soybean. Producer data on seed yield and management decisions were collected from 5682 fields planted with soybean during 2014−2016 and grouped into 10 technology extrapolation domains (TEDs) based on growing environment. A subsample of 1512 fields was classified into early and late-planted categories using terciles. Conditional inference trees were created for each TED to evaluate the effect of management decisions within the two planting date timeframes on seed yield. Management strategies that maximized yield and associated maximum break-even prices varied across TEDs and planting date. For early-planted fields, higher yields were associated with artificial drainage, insecticide seed treatment, and lower seeding rates. For late-planted fields, herbicide application timing and tillage intensity were related to higher yields. There was no individual management decision that consistently increased seed yield across all TEDs.

Abbreviations

AI, aridity index; AOSR, agronomic optimum seeding rate; CI, conditional inference; GDD, growing degree day; POST, post-emergence; PRE, pre-emergence; RM, relative maturity; RSS, residual sum of squares; RZWHC, rhizosphere water holding capacity; ST, seed treatment; TED, technology extrapolation domain

Core Ideas

- Management decisions that increased soybean yield were region specific.

- No single management decision consistently increased seed yield across the entire study region.
- Integrated pest management principles should be followed when deciding the use of pesticide inputs.

Introduction

Timely planting of soybean [*Glycine max* (L.) *Merr.*] is extremely important to maximize seed yield in the north-central United States. Several field experiments have shown seed yield reduction when planting date is delayed beyond early- to mid-May (Hu & Wiatrak, 2012; Robinson, Conley, Volenec, & Santini, 2009). For example, in Iowa, a seed yield reduction of 0.13 Mg ha^{-1} wk^{-1} (−0.02 Mg ha^{-1} d−1) was observed for soybean planted from early May to late May and −0.40 Mg ha^{-1} wk^{-1} (−0.06 Mg ha^{-1} d^{-1}) for planting dates from late May to early June (De Bruin & Pedersen, 2008). In Nebraska and Ohio, delayed planting after 1 May resulted in yield declines that ranged from −0.02 to −0.04 Mg ha^{-1} d^{-1} (Bastidas et al., 2008; Hankinson, Lindsey, & Culman, 2015). Apart from the aforementioned region-specific studies, a U.S.-wide study estimated a 10% increase in yield and approximately US$9 billion in monetary gains could be realized if soybean was planted at the optimal time across the United States (Mourtzinis, Specht, & Conley, 2019b).

Recent studies using producer data identified planting date as the most important management practice explaining field-to-field variation across regions with similar weather and soil condition in the north-central United States (Mourtzinis et al., 2018; Rattalino Edreira et al., 2017). These studies showed maximum seed yield reductions of −0.34 Mg ha^{-1} d^{-1} for each day soybean was planted after the last week of April. In regions where planting date was the most important factor influencing soybean yield, additional factors that explained a large percentage of field-to-field yield variation were topographic wetness index, subsoil pH, row width, foliar fungicide, and foliar insecticide (Mourtzinis et al., 2018).

In this study, the dataset described in Mourtzinis et al. (2018), which included data from 2014−2015, was expanded to include fields from 2014−2016 and used to identify agronomic management decisions to optimize soybean yield in early- and late-planted situations. However, unlike the previously conducted analyses, this work focused on management practices and not factors beyond producers’ control such as topographic wetness index and subsoil pH. Furthermore, this research provided an estimate of the break-even price point for inputs identified as significant predictors of yield. The objectives of this research were to: (a) identify management decisions that increase seed yield in early- or late-planted soybean scenarios, and (b) estimate the maximum break-even price of each management factor identified to influence soybean seed yield in early- or late-planted soybean.

Methods

Data Collection and Database Description

Between 2014 and 2016, researchers, extension educators, and crop consultants from 10 north-central U.S. states (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Ohio, North Dakota, Nebraska, and Wisconsin) collected data on seed yield and management decisions from 5682 producer soybean fields. The resulting database was described in Rattalino Edreira et al. (2017) and Mourtzinis et al. (2018). Self-reported management practices included planting date, cultivar relative maturity (RM), seeding rate, row width, tillage type, artificial drainage, seed treatments, fertilizer type and amount, and lime, manure, and pesticide application. Year-specific conditions such as pest pressure, Fe deficiency chlorosis incidence, and weather adversities were reported by producers. A few fields with extremely low yield due to unmanageable production site adversities (hail, waterlogging, wind, and frost) were excluded from the analyses. The procedure to exclude these fields consisted of three steps: (a) grouping fields within regions with similar soil and climate (further described in Soybean Field Classification), (b) selecting fields within the 25th percentile of yield data distribution within each region–year, and (c) excluding fields affected by any of the aforementioned adversities reported by producers. Fields that were both affected by reported adversities and fell within the 25th percentile of yield within their region were excluded from further analysis. Fields planted after 15 June that also had wheat (*Triticum aestivum* L.) as a previous crop were removed from further analyses to exclude double crop soybean production systems, which are rare in the majority of the study area.

Soybean Field Classification

Fields were grouped into technology extrapolation domains (TEDs) according to growing conditions, as characterized by growing degree days (GDDs), aridity index (AI), and root zone water holding capacity (RZWHC). Growing degree days is a measure of heat accumulation and is used to predict crop development, and it was calculated using a base temperature of 0°C. Aridity index is a measure of how dry an area is and is calculated as the ratio of mean annual precipitation and mean annual potential evapotranspiration. Root zone water holding capacity is a measure of how much water the soil can hold within the rootable depth. More information about TEDs and the calculation of GDDs and AI is available in Rattalino Edreira et al. (2018).

Technology extrapolation domains were selected for this study when more than 180 fields were located within a TED, as that number balanced having a diversity of environments included while still having a sufficient number of fields to detect differences in yield due to management (see more information in Section 1.3 Statistical analysis). The 10 TEDs included in this study contained 1512 of the 5682 total fields. Some soybean-producing regions were not included in this study due to an insufficient number of fields.

The geographic distribution of the 10 TEDs is available in Figure 4.1. The six-digit numbers following the TED numbers in the legend of Figure 4.1 are the reference numbers to locate these TEDs in the global database at yieldgap.org. All TEDs were rainfed except for TED 2, which was irrigated. Within each TED, fields were classified as early- or late-planted when falling within the first or the third terciles of planting date data distribution, respectively (Table 4.1). Some TEDs have a different number of fields in the early- and late-planted tercile due to many fields being planted on the first or last day within each timeframe.

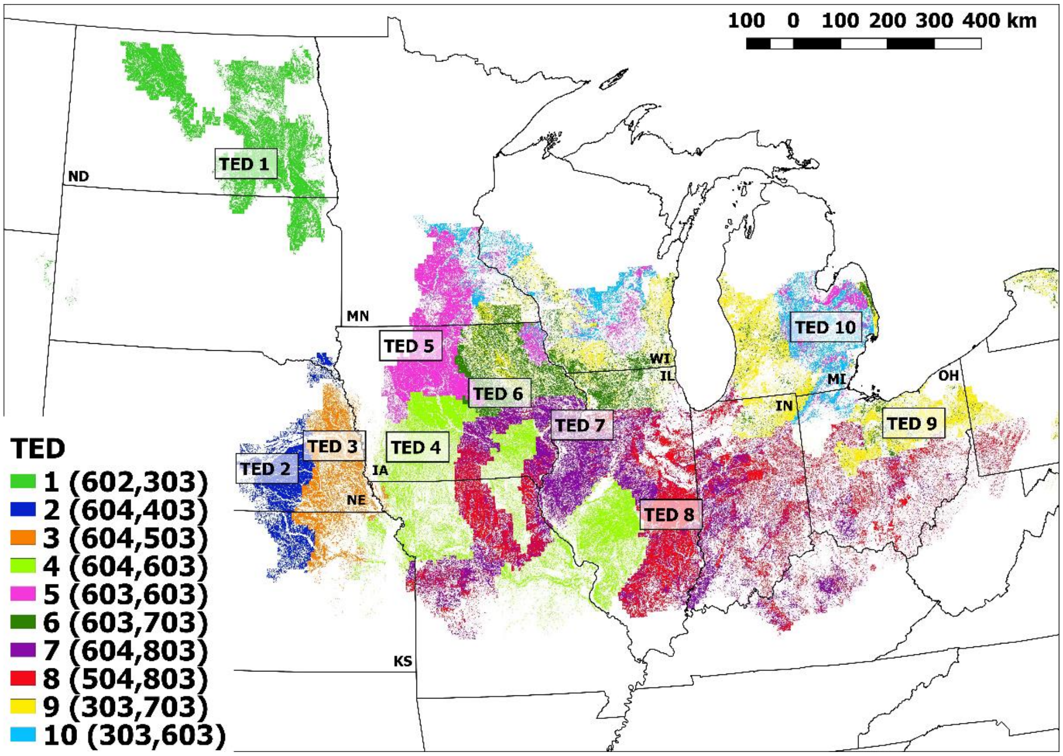

Figure 4.1. Technology extrapolation domains (TEDs) distributed across the north-central U.S. region. The six-digit numbers following the TED numbers are the reference numbers to locate these TEs in the global database at yieldgap.org

Table 4.1. Range of planting dates and number of fields for early and late-planted soybean fields within each technology extrapolation domain (TED).

	Early		Late		
TED	**Dates**	**Number of fields**	**Dates**	**Number of fields**	**Minimum difference between early- and late-planted fields†**
					days
1	24 Apr-18 May	65	26 May-14 Jun	71	8
2	6 Apr-6 May	105	17 May-6 Jul	111	11
3	21 Apr-14 May	59	22 May-1 Jul	64	8
4	22 Apr-7 May	79	20 May-30 Jun	85	13
5	18 Apr-11 May	90	22 May-13 Jun	84	11
6	17 Apr-8 May	89	21 May-10 Jun	99	13
7	10 Apr-7 May	54	22 May-23 Jun	59	15
8	10 Apr-8 May	89	22 May-19 Jun	84	14
9	29 Apr-15 May	56	26 May-16 Jun	62	11
10	26 Apr-16 May	53	25 May-15 Jun	54	9

†Minimum difference between early and late-planted fields is the number of days between the last early-planted field and the first late-planted field.

Statistical Analysis

To explore the relationship between seed yield and management decisions within the two planting date timeframes, two conditional inference (CI) trees were created for each TED−one for early-planted fields and one for lateplanted fields (Table 2). Conditional inference trees were used to identify and visualize interactions among independent variables with less risk of overfitting than other recursive decision tree algorithms (Hothorn, Hornick, & Zeileis, 2006). Significance testing was used to perform splits within CI trees, with the lowest p value determining each split. The null hypothesis for each split was that the dependent variable (seed yield) was independent of the management decision variable.

The above described CI tree analysis was implemented using the package partykit within R 3.2.4 statistical software (Hothorn & Zeileis, 2015; R Development Core Team, 2016). The independence-test criterion for splits was univariate p value ($\alpha = .05$). Interior nodes were required to maintain at least 33% of the data. At minimum, terminal nodes included 10 fields. Overfitting was prevented by constraining trees at a maximum depth of 10 nodes. To quantify the minimum detectible yield difference given the number of trees used to create each CI tree, power analysis was performed using the package pwr within R 3.2.4 (Champley et al., 2018). One-way ANOVA tests were performed to determine the effect size (f) when the significance (alpha) level was .05 and the power level was 0.80. The average standard deviation of yield within each TED and planting date timeframe was 0.264 Mg ha^{-1}, and was used to calculate minimum detectable difference from the effect size (f). Effect size as measured by Cohen's f is a standardized, unitless measure. Under the range of sample sizes present between planting date timeframes and TEDs (Table 1) and the possible unevenness of splits in the CI trees, the effect sizes ranged from 0.27 to 0.66. After converting f to Cohen's d, d was divided by standard deviation to estimate the minimum detectable difference in-yield. The range of effect size (f) of 0.27−0.66 corresponds to a range in minimum detectable difference in yield of 0.033−0.088 Mg ha^{-1}.

The following variables were considered binary (yes/no): artificial drainage, fungicide seed treatment, insecticide seed treatment, inoculant seed treatment, nematicide seed treatment, starter fertilizer (all possible fertilizer sources and placements), starter P, foliar fungicide, foliar insecticide, and manure application. The following variables were considered categorical: tillage (minimal or intense), herbicide (none, pre-emergence only, post-emergence only, or both), row width (narrow, medium, or wide), and previous crop (corn [*Zea mays* L.], soybean, wheat, sunflower [*Helianthus annuus* L.], sorghum [*Sorghum bicolor* L.], cereal rye [*Secale cereal* L.], sugarbeet [*Beta vulgaris* L.], popcorn, alfalfa [*Medicago sativa* L.], oat [*Avena sativa* L.], barley [*Hordeum vulgare* L.], hay, potato [*Solanum tuberosum* L.], or corn silage) Minimal tillage included no-tillage, strip-tillage, ridge tillage, or harrow while intense tillage included chisel plow, moldboard plow, disk, field cultivator, and/or soil finisher implements. For row width, <25 cm, 25−56 cm, and >56 cm were considered, narrow, medium, and wide, respectively. Seeding rate and RM were considered continuous variables. For each TED and planting date combination, independent variables where 90% of fields had the same treatment were excluded from the analysis, such as artificial drainage in early-planted fields in TED 2. If the management decision for more than half of the fields in a TED was not available from our survey form for a particular management decision, the management decision was also excluded from analysis, such as inoculant seed treatment in late-planted fields in TED 8. A summary of management decisions within each TED and planting date timeframe is shown in Table 4.2.

Table 4.2. Summary of management decisions within each technology extrapolation domain (TED) for early (E) and late (L) planting timeframes displayed as percent of fields with that treatment, except for average seeding rate (1000 seeds ha^{-1}) and average yield (Mg^{-1}).

		TED 1		TED 2		TED 3		TED 4		TED 5		TED 6		TED 7		TED 8		TED 9		TED 10	
		E †	L ‡	E	L	E	L	E	L	E	L	E	L	E	L	E	L	E	L	E	L
Artificial drainage		43	23	- §	-	20	20	-	75	89	77	80	84	83	71	78	46	66	73	66	80
Seed treatment	Fungicide	55	68	46	-	59	50	71	69	61	64	71	63	-	59	-	61	-	73	53	56
	Insecticide	42	46	44	-	-	48	63	69	54	61	54	55	-	54	58	49	-	69	49	52
	Inoculant	65	79	-	-	-	-	-	-	10	-	-	-	19	-	-	-	20	-	-	-
	Nematicide	-	-	-	-	-	-	30	26	14	-	24	-	20	-	-	-	-	-	-	-
Starter fertilizer	Starter fertilizer	28	44	16	-	15	14	-	-	-	-	-	10	-	15	-	10	-	-	23	22
	P fertilizer	37	39	-	-	14	11	-	-	-	-	-	-	-	-	-	-	-	-	21	19
Manure	Manure	-	-	-	-	-	-	-	14	-	17	-	-	-	-	11	1	14	8	15	2
Foliar application	Fungicide	12	-	23	-	25	13	20	31	61	19	45	48	65	19	34	-	29	-	26	-
	Insecticide	40	31	19	-	20	11	19	29	39	29	43	35	61	24	33	21	23	19	19	17
Average seeding rate (1,000 ha^{-1}) ¶		420	417	415	408	375	367	366	378	368	378	380	373	385	410	395	395	378	408	395	398
Row width	Narrow	5	13	4	5	0	0	15	5	9	19	3	6	9	8	18	17	11	23	25	31
	Medium	78	66	28	7	53	45	38	56	47	39	74	55	57	86	66	77	57	58	58	30
	Wide	17	20	67	86	46	53	47	39	44	42	22	39	33	5	11	5	32	19	15	39
Tillage	Intense	22	17	16	39	-	-	47	-	13	17	20	15	17	25	13	15	20	15	21	20
	Minimal	35	27	82	57	-	-	28	-	64	70	69	80	56	66	56	74	64	65	58	41
Herbicide application timing	PRE + POST	58	37	87	76	-	89	-	-	80	74	78	79	83	83	82	89	80	65	57	67
	POST only	22	28	9	14	-	3	-	-	20	20	12	13	7	8	15	5	18	29	43	30
	PRE only	9	1	0	0	-	5	-	-	0	6	6	3	0	2	1	6	2	6	0	4
	None	11	5	5	10	-	3	-	-	0	0	4	5	9	7	2	0	0	0	0	0

Previous crop	Corn	43	33	93	92	-	80	100	95	91	86	99	93	96	92	92	87	84	85	83	74
	Soybean	22	13	-	-	-	2	-	-	-	4	-	-	-	-	-	5	4	13	6	9
Average yield (Mg ha^{-1})		2.8	2.4	5	4.7	3.9	3.6	4.5	4	4.3	3.7	4.3	3.7	4.5	4.1	4.3	4	3.8	3.5	3.9	3.6

† E is for early planted soybean, as determined using the first third of planted soybeans within each TED.

‡ L is for late planted soybean, as determined using the last third of planted soybeans within each TED.

§Variable excluded from analysis for that TED and planting date timeframe due to 90% or more of fields being treated identically or greater than 50% of fields not having adequate data for that particular variable.

¶Average seeding rate for each TED and planting date timeframe is presented in 1,000 seeds ha^{-1}, differing from other treatments that are displayed as percentages

For in-season management decisions that increased yield, the maximum break-even price was calculated. The maximum break-even price is the highest price a producer can pay for a treatment and still expect a profit, or in other words, have a positive return on investment. Grain yield benefit was calculated using the CI trees by subtracting the average yield from the node without the yield-improving treatment from the average yield from the node with the yield-improving treatment. Grain yield was multiplied by grain price to calculate the maximum break-even price under three different grain price scenarios: \$297, \$333, and \$368 Mg^{-1}. These three values represent conservative, but realistic, price scenarios, given that between January 2015 and June 2019, the lowest observed grain price was \$297 Mg^{-1}, and the median observed price was \$368 Mg−1 (USDA NASS, 2019). Costs for implementing each decision includes both products and their application. Product costs were estimated in 2017 using a phone survey of retailers in the 10 participating states (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Ohio, North Dakota, Nebraska, and Wisconsin), and application costs were averaged from state custom application budgets.

Results

Early-planted Fields

Among early-planted fields, management factors that were associated most consistently with changes in soybean seed yield within several TEDs included artificial drainage (TEDs 1, 6, and 10), insecticide seed treatment (TEDs 1, 6, and 10), and seeding rate (TEDs 7, 8, and 10) (Figure 4.2; Table 4.3). In TED 6, maximum average seed yield for early-planted soybean was 4.8 Mg ha^{-1} and was associated with fields without artificial drainage (Figure 4.2). In TED 1, soybean not treated with insecticide seed treatment yielded 0.39 Mg ha^{-1} more when artificial drainage was present, compared to yields in fields without artificial drainage (Table 4.3). There was also an increase in soybean seed yield in TED 10 when artificial drainage was present (Table 4.3)

Table 4.3. Summary of conditional inference trees for early and late-planted fields in technology extrapolation domains (TEDs) 1, 2, 3, 4, 5, 7, 9, and 10. Bracketed values are the number of fields (n) and average yield (Y, Mg ha^{-1}), and RSS is the residual sum o squares. Conditional inference trees did not identify any significant decisions for late-planted fields in TED 3.

	Early					**Late**			
TED	**Decision 1**	**Decision 2**	**Decision 3**	**[n,Y]**	**RSS**	**Decision 1**	**Decision 2**	**[n,Y]**	**RSS**
1	Insecticide ST (No)	--	--	[33, 2.56]	--	Starter P (No)	--	[43, 2.56]	--
	--	Artificial Drainage (No)	--	[26, 2.51]	3.6	--	Fungicide ST (No)	[11, 2.16]	3.4
	--	Artificial Drainage (Yes)	--	[10, 2.90]	2.9	--	Fungicide ST (Yes)	[32, 2.62]	6.8
	Insecticide ST (Yes)	--	--	[27, 2.96]	--	Starter P (Yes)	--	[28, 2.18]	11.4
	--	Relative Maturity ≤0.9	--	[17, 2.98]	5.4	--	--	--	--
	--	Relative Maturity >0.9	--	[12, 3.06]	2.5	--	--	--	--
2	Starter Fertilizer (No)	--	--	[88, 4.91]	18.6	Herbicide (Both)	--	[84, 4.84]	--
	Starter Fertilizer (Yes)	--	--	[17, 5.41]	5.2	--	Relative Maturity ≤2.7	[31, 5.02]	3.5
	--	--	--	--	--	--	Relative Maturity >2.7	[53, 4.74]	7.1
	--	--	--	--	--	Herbicide (None or POST)	--	[27, 4.33]	12.2
3	Foliar Fungicide (No)	--	--	[44, 3.65]	22.3	--	--	--	--
	Foliar Fungicide (Yes)	--	--	[15, 4.63]	6.3	--	--	--	--
4	Relative Maturity ≤3	--	--	[45, 4.33]	7.3	Relative Maturity ≤3.4	--	[57, 4.10]	11.8
	Relative Maturity >3	--	--	[34, 4.70]	8.2	Relative Maturity >3.4	--	[28, 3.79]	15.5

5	Foliar Insecticide (No)	--	--	[55, 4.16]	21.5	Row Width (Medium or Narrow) †	--	[49, 3.53]	--
	Foliar Insecticide (Yes)	--	--	[35, 4.52]	12.0	--	Herbicide (Both)	[37, 3.70]	18.0
	--	--	--	--	--	--	Herbicide (Post or Pre)	[12, 2.98]	5.6
	--	--	--	--	--	Row Width (Wide)	--	[35, 4.00]	10.9
7	Seeding Rate ≤403k seeds/ha	--	--	[36, 4.81]	19.7	Seeding Rate ≤358k seeds/ha	--	[10, 4.94]	4.3
	Seeding Rate >403k seeds/ha	--	--	[18, 3.91]	8.5	Seeding Rate >358k seeds/ha	--	[49, 3.97]	--
	--	--	--	--	--	--	Seeding Rate ≤432k seeds/ha	[33, 4.17]	9.7
	--	--	--	--	--	--	Seeding Rate >432k seeds/ha	[16, 3.57]	3.2
9	Inoculant ST (No)	--	--	[38, 4.00]	--	Foliar Insecticide (No)	--	[50, 3.42]	22.9
	--	Foliar Fungicide (N)	--	[26, 3.71]	16.8	Foliar Insecticide (Yes)	--	[12, 3.98]	2.7
	--	Foliar Fungicide (Y)	--	[12, 4.62]	2.0	--	--	--	--
	Inoculant ST (Yes)	--	--	[18, 3.47]	8.5	--	--	--	--
10	Artificial Drainage (No)	--	--	[18, 3.42]	6.1	Herbicide (Both, PRE)	--	[38, 3.83]	20.6
	Artificial Drainage (Yes)	--	--	[35, 4.09]	--	Herbicide (POST)	--	[16, 3.08]	3.2

		--	Insecticide ST (No)	--	[13, 3.72]	3.8				
		--	Insecticide ST (Yes)	--	[22, 4.31]		--	--	--	--
		--	--	Seeding Rate ≤383 seeds/ha	[12, 4.57]	4.2	--	--	--	--
		--	--	Seeding Rate >383 seeds/ha	[10, 4.00]	1.6	--	--	--	--

† Narrow rows were < 25 cm, medium rows were 25 to 56 cm, and wide rows were > 56 cm in width.
‡ Seed treatment

In TED 1, the highest seed yield (3.06 Mg ha^{-1}) was achieved when insecticide seed treatment was applied to soybean cultivars with a MG > 0.9. On average, TED 1 fields with insecticide seed treatment yielded 0.5 Mg ha^{-1} greater than fields without insecticide seed treatment (Table 4.3). In TED 6, fields with artificial drainage and both herbicide timings, but lacking nematicide seed treatment, had 0.09 Mg ha^{-1} greater yield with insecticide seed treatments compared to fields without insecticide seed treatments (Figure 4.2). Technology extrapolation domain 10 also had higher yield in fields that had artificial drainage. Seed yield was further associated with insecticide seed treatment, resulting in lower seed yield when seed was not treated with an insecticide compared to seed treated with an insecticide (Table 4.3).

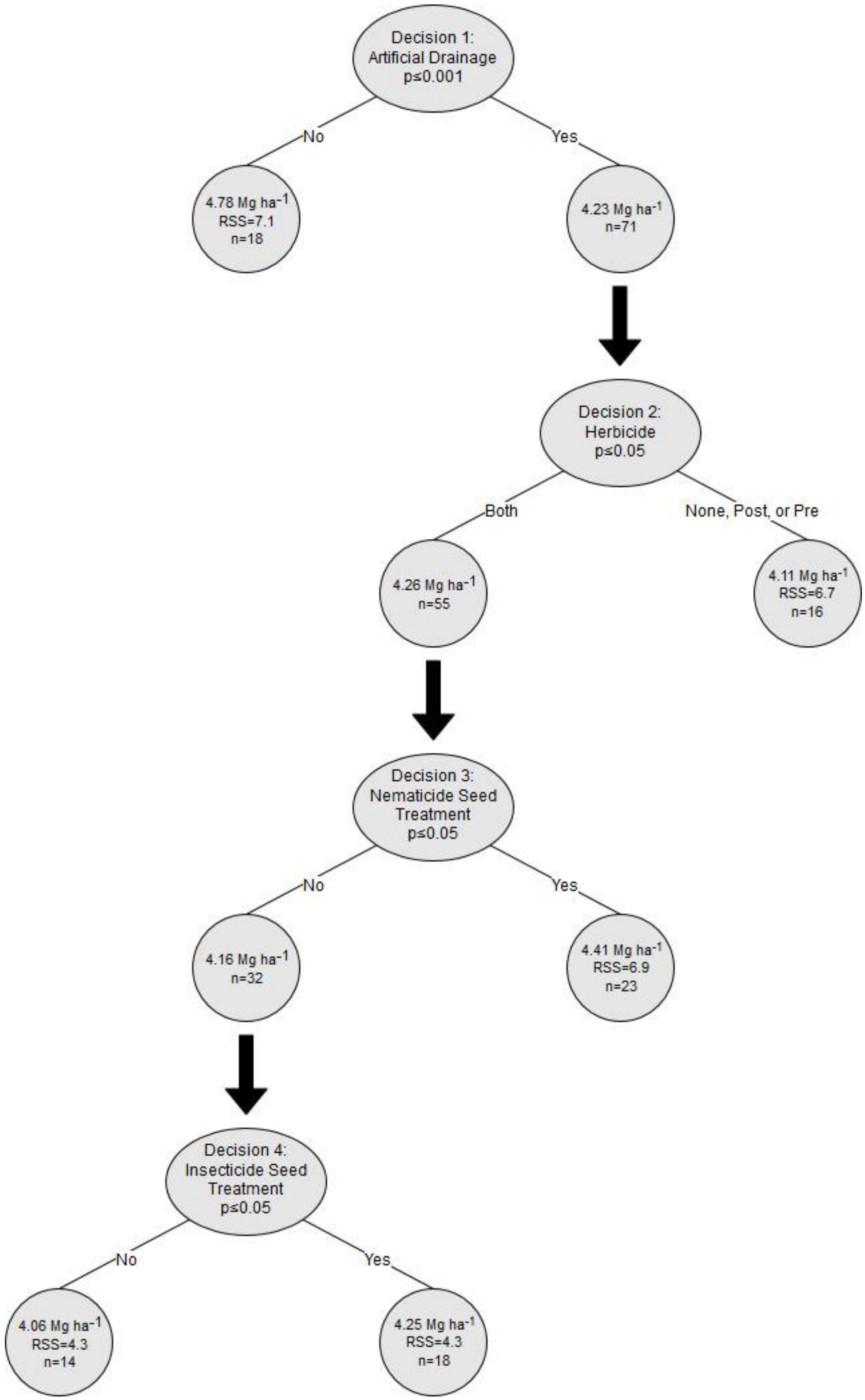

Figure 4.2. Conditional inference tree for technology extrapolation domain (TED) 6 showing significant management decisions for predicting yield in early-planted soybean fields where RSS is the residual sum of squares for each terminal node, and n is the number of fields present in each node.

Of the TEDs with a significant difference in yield corresponding to seeding rate, higher yields were consistently observed where seeding rate was lower. Among TED 10 fields with artificial drainage and where insecticide seed treatment was applied, soybean yield was greater when seeding rate was ≤383,000 seeds ha^{-1} (Table 4.3). Other TEDs with higher yield at lower seeding rates were TEDs 7 and 8. In TED 7, fields planted early at ≤403,000 seeds ha^{-1} resulted in a soybean seed yield 0.90 Mg ha^{-1} greater than fields planted at >403,000 seeds ha^{-1} (Table 3). In TED 8, fields with seeding rates ≤383,000 seeds ha^{-1} showed greater seed yield than fields with higher seeding rates (Figure 4.3).

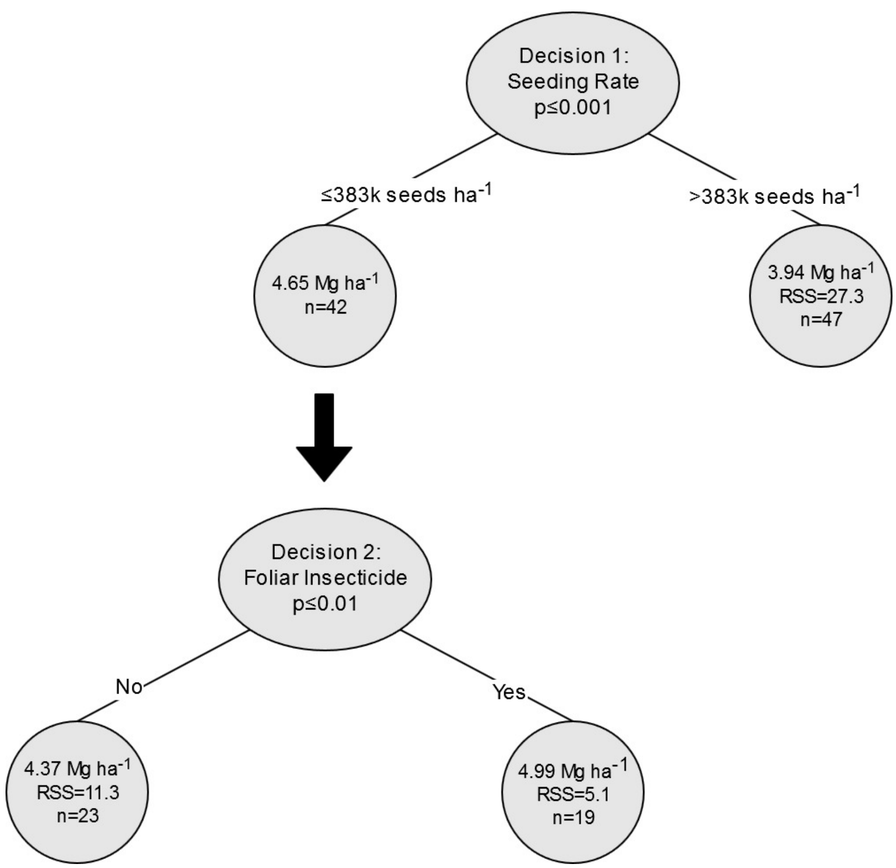

Figure 4.3. Conditional inference tree for technology extrapolation domain (TED) 8 showing significant management decisions for predicting yield in early-planted soybean fields where RSS is the residual sum of squares for each terminal node, and n is the number of fields present in each node.

Late-planted Fields

Among late-planted fields, management factors that were correlated with changes in soybean seed yield within several TEDs included herbicide application timing (TEDs 2, 5, and 10) (Table 4.3) and tillage intensity (TEDs 6 and 8) (Figures 4.4 and 4.5, respectively). In TED 2, fields that received no herbicide application or only a POST-herbicide application were associated with the lowest soybean seed yield (4.33 Mg ha−1). In TED 5 fields where soybean was planted in narrow or medium row widths, seed yield was correlated with herbicide application. Greater soybean seed yield (0.72 Mg ha−1) was associated with fields that received a PRE- and POST-herbicide application compared to fields that only received a PRE or POST herbicide application (Table 4.3). Across late-planted fields in TED 10, when a PRE and POST or only a PRE herbicide was applied, soybean seed yield was greater compared to fields that only received a POST-herbicide application (Table 4.3).

In TED 6, late-planted fields receiving intense tillage were associated with the greatest seed yield at 4.1 Mg ha^{-1} (Figure 4.4). In fields with minimal tillage, foliar fungicide increased yield by 0.29 Mg ha^{-1} compared with fields without a foliar fungicide application. In TED 8, the highest yields were observed in fields with intensive tillage when corn or sorghum was the previous crop and there was no artificial drainage. Minimally tilled fields had 0.21 Mg ha^{-1} higher yield for cultivars of ≤3.8 RM compared to cultivars of >3.8 RM (Figure 4.5).

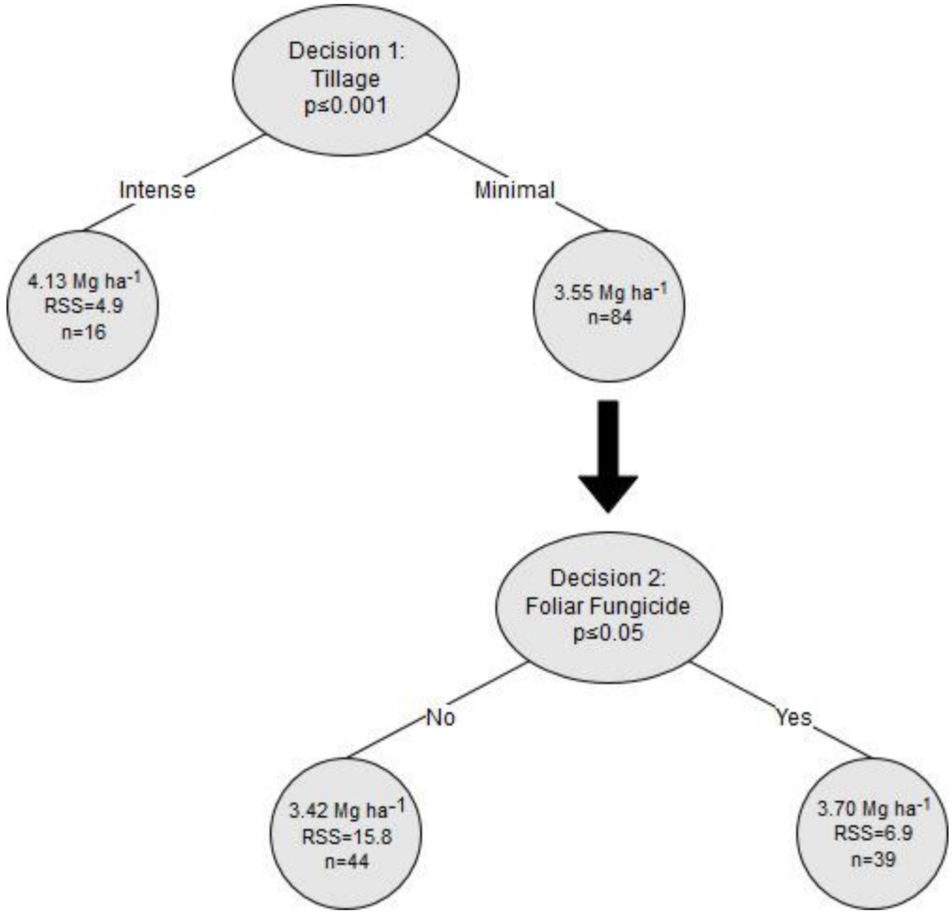

Figure 4.49. Conditional inference tree for technology extrapolation domain (TED) 6 showing significant management decisions for predicting yield in late-planted soybean fields where RSS is the residual sum of squares for each terminal node, and n is the number of fields present in each node.

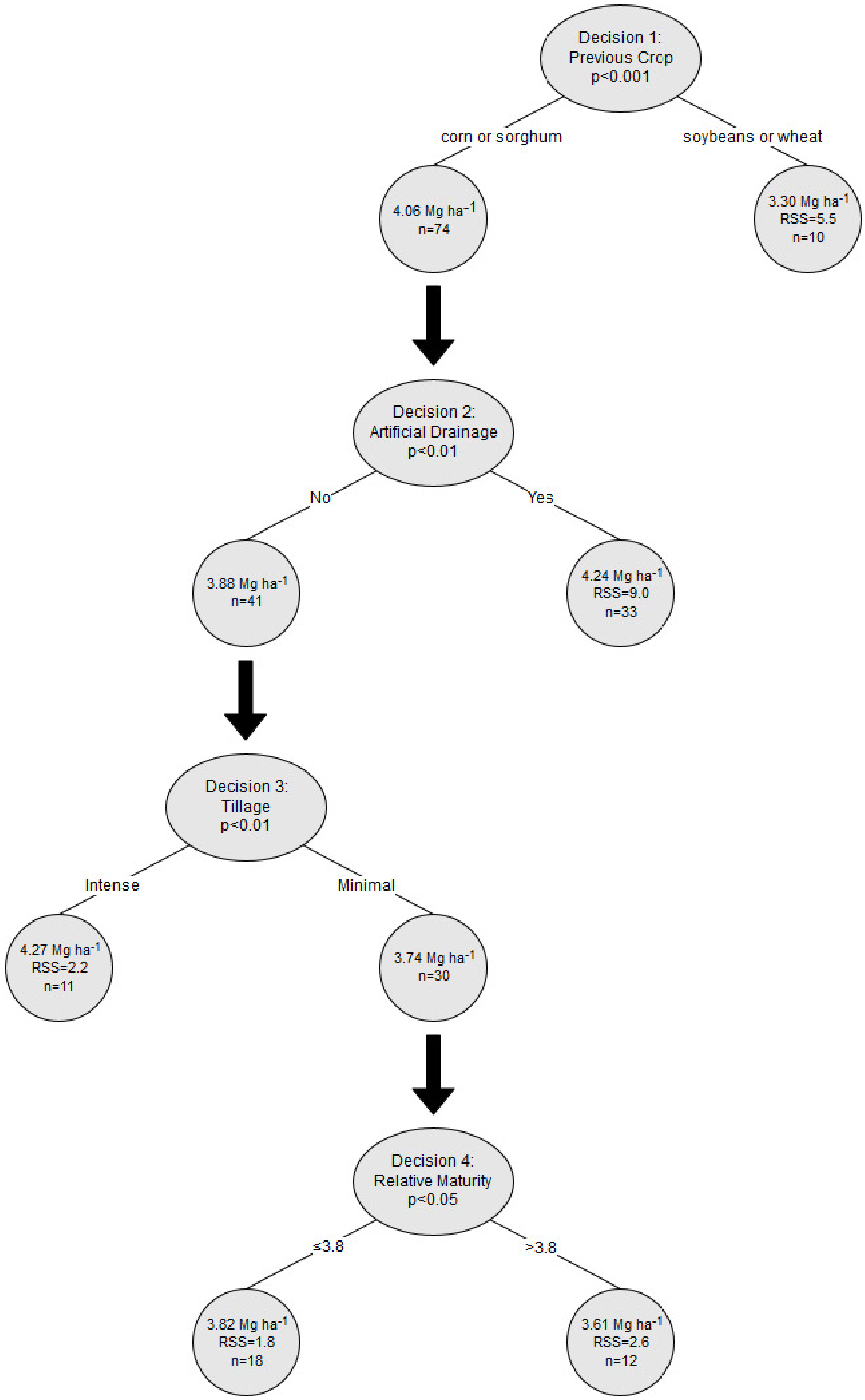

Figure 4.5. Conditional inference tree for technology extrapolation domain (TED) 8 showing significant management decisions for predicting yield in late-planted soybean fields where RSS is the residual sum of squares for each terminal node, and n is the number of fields present in each node.

Foliar fungicides and insecticides improved yield for late-planted fields in three TEDs. In minimally tilled TED 6 fields, seed yield was 0.29 Mg ha^{-1} greater with an application of foliar fungicide compared to yields in fields without foliar fungicide (Figure 4.4). In TED 9 fields where foliar insecticide was applied, there was an increase in yield of 0.56 Mg ha−1 (Table 4.3).

Economics

Maximum break-even price for insecticide seed treatment ranged from \$63 to \$196 ha^{-1} at a grain price of \$333 Mg^{-1} (Table 4.4). More frequent herbicide applications improved yield for late-planted soybean in TEDs 2, 5, and 10, and for early planted soybean in TED 6 (Table 4.3), with the maximum break-even price for herbicide ranging from \$50 to \$250 ha^{-1} at a grain price of \$333 Mg^{-1} (Table 4.4). For late-planted soybean in TEDs 2, 5, and 10, this maximum break-even price at a grain price of \$333 Mg^{-1} covers the cost of moving from an herbicide program with only a POST application to a program with both a PRE and a POST. The maximum break-even price was not high enough to cover the cost of implementing a PRE and POST program for early-planted soybean in TED 6 (Table 4.4). Foliar insecticide improved yields in TEDs 5 and 8 for early-planted soybean, and in TED 9 for late-planted soybean. The maximum break-even price for foliar insecticide ranged from \$120 to \$206 ha^{-1} at a grain price of \$333 Mg^{-1}. For early-planted soybean in TEDs 5 and 8, and late-planted soybean in TED 9, the estimated cost of applying foliar insecticide is lower than the maximum breakeven price (Table 4.4). Foliar fungicide improved yield for early planted soybean in TEDs 3 and 9, and for late-planted soybean in TED 6, with a maximum break-even price of \$326 ha^{-1}. The cost of applying foliar fungicide was lower than the maximum break-even price at a grain price of \$333

Mg^{-1} for early-planted soybean in TEDs 3 and 9. The cost of applying foliar fungicide was higher than the maximum break-even price for late-planted fields in TED 6.

Table 4.4. Maximum break-even price a producer should pay for specific management decisions or inputs that improved yield at three grain prices ($297, $333, and $368 Mg^{-1}), where TED is the technology extrapolation domain and PD is the planting date timeframe (E = early, L = late). Yield benefit was taken by subtracting the average yield in fields without that treatment from fields with that treatment in the conditional inference trees from Table 3. Technology extrapolation domains 4 and 7 (both planting windows) and 3 for late planting did not have in-season decisions that would have an associated break-even price. Costs for implementing each decision includes both products and their application. Product costs were estimated using a phone survey of retailers in the 10 participating states (IL, IN, IA, KS, MI, MN, OH, ND, NE, and WI), and application costs were averaged from state custom application budgets.

				Maximum break-even price at the given grain price			
TED	PD	Decision	Yield benefit	297 USD Mg^{-1}	333 USD Mg^{-1}	368 USD Mg^{-1}	Estimated cost of implementation
			Mg ha^{-1}	-------USD ha^{-1}-------			
1	E	Insecticide ST †	0.40	119	133	147	37
1	L	Fungicide ST	0.46	137	153	169	37
2	E	Starter Fertilizer	0.50	149	167	184	81
2	L	Herbicide	0.51	151	170	188	123*
3	E	Foliar Fungicide	0.98	291	326	361	117
5	E	Foliar Insecticide	0.36	107	120	132	65
5	L	Herbicide	0.72	214	240	265	123*
6	E	Herbicide	0.15	45	50	55	123*
6	E	Nematicide ST	0.25	74	83	92	46
6	E	Insecticide ST	0.19	56	63	70	37
6	L	Foliar Fungicide	0.28	83	93	103	117
8	E	Foliar Insecticide	0.62	184	206	228	65
9	E	Foliar Fungicide	0.91	270	303	335	117
9	L	Foliar Insecticide	0.56	166	186	206	65
10	E	Insecticide ST	0.59	175	196	217	37
10	L	Herbicide	0.75	223	250	276	123*

†ST: Seed treatment
*Cost of adding a PRE-emergence herbicide

Discussion

While each TED had a different combination of treatments that maximized yield under different planting date timeframes, there were some commonalities among TEDs. Among early-planted fields, management factors that influenced soybean seed yield within a few TEDs included artificial drainage (TEDs 1, 6, and

10), insecticide seed treatment (TEDs 1, 6, and 10), and seeding rate (TEDs 7, 8, and 10). Improved yield in fields with artificial drainage as compared to fields without artificial drainage is likely due to a combination of reduced plant damage from flooding and improved timeliness of farm operations such as tillage, planting, and spraying (Aldabagh & Beer, 1975; Kanwar, Johnson, Schult, Fenton, & Hickman, 1983). Improved planting conditions, particularly in wet springs, could be part of the reason there was an association between artificial drainage and higher yields for early-planted fields in three TEDs (1, 6, and 10), but the same association was only seen in one TED 8 for late-planted fields.

While insecticide seed treatments were associated with higher yields in three TEDs (1, 6, and 10) for early-planted soybean, they were not associated with a change in yield for any late-planted soybean. In Wisconsin, combined insecticide−fungicide seed treatments improved yield by 4−12% (Gaspar, Mitchell, & Conley, 2015). However, Mourtzinis et al. (2019a) recently reported a minimal yield increase (0.13 Mg ha−1) across 14 states due to combined insecticide–fungicide seed treatments. While insecticide seed treatments were not associated with a change in yield for late-planted soybean, higher yields for late-planted fields treated with foliar insecticides were observed in TED 9. Insect pest pressure can vary by soybean-planting date (Hammond, Higgins, Mack, Pedigo, & Bachinski, 1991; Zeiss & Klubertanz, 1994). Technology extrapolation domains with an association between insecticides and soybean yield had maximum break-even prices that were higher than the estimated cost of implementing the insecticide seed treatments or foliar sprays, which indicates that insecticides may be an economically feasible treatment for producers (Table 4.4).

Among early-planted fields in TEDs 8 and 10, fields with seeding rates greater than 383,000 seeds ha^{-1} yielded significantly less than fields with lower seeding rates. Early-planted fields in TED 7 yielded less when their seeding rate was in excess of 403,000 seeds ha^{-1}. Past studies indicate that the agronomic optimum seeding rate (AOSR) for soybean in the north-central United States is variable. For May-planted soybean in Iowa and Ohio, AOSR has been observed to vary between 157,000 and 211,800 seeds ha^{-1} and 345,800 and 481,650 seeds ha^{-1}, respectively (Barker et al., 2017; De Bruin & Pedersen, 2008). In Wisconsin, seeding rates between 296,400 and 345,800 seeds ha^{-1} yielded similarly (Gaspar et al., 2015). In a regional study, the AOSR for the Midwest was 365,000 seeds ha^{-1} (Gaspar et al., 2020). The seeding rate value selected in the CI tree analysis is likely near or in excess of the AOSR for each TED given past seeding rate studies, so the lower yield in fields with higher seeding rates in TEDs 7, 8, and 10 was likely due at least in part to high seeding rate and not just an artifact of farmers selecting higher seeding rates for fields with lower yield potential. Fields in these TEDs had similar use of tillage, foliar fungicide, foliar insecticide, and seed treatment regardless of seeding rate.

Foliar fungicides were associated with increased yield in early-planted fields in two TEDs (3 and 9) and late-planted fields in one TED 6. In TED 6 where foliar fungicide was associated with higher yields in late-planted fields, it was only in minimally tilled fields. Minimally tilled fields yielded less than intensively tilled fields, but foliar fungicide helped recover part of the difference in yield between tillage regimes in late-planted fields.

This dataset did not include information on scouting practices or insect and disease pressure. Since insect and pathogen pressure vary annually, the association between greater yields and pesticide application could change among growing seasons. It is recommended to follow state guidelines for insect and disease management based on an integrated pest management (IPM) approach. Prophylactic applications of foliar insecticide and fungicide are not recommended as they are generally not associated with an economic benefit (Bluck, Lindsey, Dorrance, & Metzger, 2015; Mourtzinis, Marburger, Gaska, & Conley, 2016; Ng, Lindsey, Michel, & Dorrance, 2018). Similarly, prophylactic use of fungicide- and/or insecticide-treated seed does not provide a consistent economic benefit for different combinations of consequential management practices, such as seeding rate (Mourtzinis et al., 2019a). Market prices and pest pressure both play an important role in determining where insecticide and fungicide applications are likely to be profitable (Gaspar et al., 2015)

Among late-planted fields, management factors that were associated with soybean seed yield within several TEDs included herbicide application timing (TEDs 2, 5, and 10) and tillage intensity (TEDs 6 and 8). Response to herbicide could be related to delayed planting resulting in the soybean canopy fully closing later in the growing season, and in some cases, never completely closing (Steele & Grabau, 1997). Full canopy closure is necessary to minimize weed pressure, especially from weeds with an extended emergence period, such as Palmer amaranth (*Amaranthus palmeri* S. Wat.) (Hock, Knezevic, Martin, & Lindquist, 2005; Jha & Norwsorthy, 2009). Of total Palmer amaranth germination throughout the growing season, more than 90% occurred prior to soybean canopy closure (Jha & Norsworthy, 2009). In TED 5, herbicide timing was associated with increase yield only when medium or narrow rows were used.

Management decisions that best correlated with soybean yield differed between early- and late-planted fields in every TED. In TED 4, RM was the decision most strongly associated with yield for both early- and late-planted fields; however, in early-planted fields longer RMs yielded better, whereas in late-planted fields the opposite was true. The management decision best correlated with yield was seeding rate for both early- and late-planted fields in TED 7, but the binary split occurred at different seeding rates

In TED 2, starter P was associated with lower yield and in TED 6, artificial drainage was associated with lower yield. This could be due to treatments being selected by producers for specific fields, not randomly applied. Producers likely applied starter P or installed artificial drainage on fields with lower yield potential due to known fertility or drainage issues, respectively. The decrease in yield at higher seeding rates observed in early-planted soybean in TEDs 7, 8, and 10 and late-planted soybean in TED 7 could be due to producers selecting higher rates for fields with lower yield potential, since lower yield potential areas have higher agronomic optimum plant densities (Carciochi et al., 2019).

Conclusions

The challenges associated with treatments being nonrandomly assigned to fields and applied in combination were outweighed by the effectiveness of survey data collection. Surveys allowed for data to be collected on 16 different management factors applied in varying combinationsacross 10 different states over three growing seasons. Small plot research studying a similar number of treatments and combinations in multiple environments would be cost prohibitive. Conditional inference trees did not identify all potentially significant decisions, but were useful for identifying interactions among management decisions, such as herbicide and row width in late-planted fields in TED 5 or tillage and foliar fungicide in late-planted fields in TED 6. Since producers used a combination of management decisions on each field, identifying interactions was important for this work.

Across all TEDs, early-planted soybean fields yielded higher than late-planted soybean fields. Our results showed no single management factor was responsible for higher yields across TEDs and planting windows, thus decisions need to be both region and planting date specific. These results confirm the importance of and continued need for locally driven data and IPM practices from which research-based best management practices can be developed. Our results also suggest the use of producer survey data can complement and expand the interpretative reach of in-field replicated research.

Acknowledgements

We acknowledge the North-Central Soybean Research Program (NCSRP), Nebraska Soybean Board, and Wisconsin Soybean Marketing Board for their support to this project. We also thank UNL Extension Educators, Nebraska Natural Resource Districts, OSU Extension Educators, and Iowa Soybean Association for helping collect the producer data; and Doug Alt for helping gather input price data. Finally, we thank Lim Davy, Agustina Diale, Juan Pedro Erasun, Laurie Gerber, Clare Gietzel, Mariano Hernandez, Ngu Kah Hui, Caleb Novak, Juliana de Oliveira Hello, Pedro Rocha Pereira, Matt Richmond, and Paige Wacker for help inputting and cleaning the survey data.

Conflict of Interest

The authors declare no conflict of interest.

Works Cited

Aldabagh, A. S. Y., & Beer, C. E. (1975). Economics of increased mobility from tile drainage. Transactions of the ASAE, 18, 116–121. https://doi.org/10.13031/2013.36536

Barker, D. J., Culman, S., Dorrance, A., Fulton, J., Haden, R., Lentz, E., ... Witter, J. (2017). Ohio agronomy guide (Bulletin 472, 15th ed.). Columbus, OH: Ohio State University Extension.

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

Bluck, G. M., Lindsey, L. E., Dorrance, A. E., & Metzger, J. D. (2015). Soybean yield response to rhizobia inoculant, gypsum, manganese fertilizer, insecticide, and fungicide. Agronomy Journal, 107, 1757–1765. https://doi.org/10.2134/agronj15.0094

Carciochi, R., Schwalbert, R., Andrade, F., Corassa, G. M., Carter, P., Gaspar, A. P., ... Ciampitti, I. A. (2019). Soybean seed yield response to plant density by yield environment in North America. Agronomy Journal, 111, 1923–1932. https://doi.org/10.2134/agronj2018.10.0635

Champley, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... De Rosario, H. (2018). pwr: Basic functions for power analysis. CRAN. Wirtschaftsuniversität Wien, Austria: Institute for Statistics and Mathematics of WU (Wirtschaftsuniversität Wien). Retrieved from https://github.com/heliosdrm/pwr

De Bruin, J. L., & Pedersen, P. (2008). Soybean seed yield response to planting date and seeding rate in the Upper Midwest. Agronomy Journal, 100, 696–703. https://doi.org/10.2134/agronj2007.0115

Gaspar, A. P., Mitchell, P. D., & Conley, S. P. (2015). Economic risk and profitability of soybean fungicide and insecticide seed treatments at reduced seeding rate. Crop Science, 55(2), 924–933. https://doi.org/10.2135/cropsci2014.02.0114

Gaspar, A. P., Mourtzinis, S., Kyle, D., Galdi, E., Lindsey, L. E., Matcham, ... Conley, S. P. (2020). Defining optimal soybean seeding rates and associated risk across North America. Agronomy Journal, 112(3), 2103–2114. https://doi.org/10.1002/agj2.20203

Hammond, R. B., Higgins, R. A., Mack, T. P., Pedigo, L. P., & Bachinski, E. J. (1991). Soybean pest management. In D. Pimentel (Ed.), CRC handbook of pest management in agriculture (2nd ed., pp. 341– 472). Boca Raton, FL: CRC Press.

Hankinson, M. W., Lindsey, L. E., & Culman, S. W. (2015). Effect of planting date and starter fertilizer on soybean grain yield. Crop, Forage & Turfgrass Management, 1(1), 1–6. https://doi.org/10.2134/cftm2015.0178

Hock, S. M., Knezevic, S. Z., Martin, A. R., & Lindquist, J. L. (2005). Soybean row spacing and weed emergence time influence weed competitiveness and competitive indices. Weed Science, 54, 38–46. https://doi.org/10.1614/WS-05-011R.1

Hothorn, T., Hornick, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. Journal of Computational and Graphical Statistics, 15(3), 651–674. https://doi.org/10.1198/106186006X133933

Hothorn, T., & Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. Journal of Machine Learning Research, 16, 3905–3909.

Hu, M., & Wiatrak, P. (2012). Effect of planting date on soybean growth, yield, and grain quality: A review. Agronomy Journal, 104, 785–790. https://doi.org/10.2134/agronj2011.0382

Jha, P., & Norsworthy, J. K. (2009). Soybean canopy and tillage effects on emergence of Palmer amaranth (Amaranthus palmeri) from a nature seed bank. Weed Science, 57, 644–651. https://doi.org/10.1614/WS-09-074.1

Kanwar, R. S., Johnson, H. P., Schult, D., Fenton, T. E., & Hickman, R. D. (1983). Drainage needs and returns in North-Central Iowa. Transactions of the ASAE, 26, 457–0464. https://doi.org/10.13031/2013.33958

Mourtzinis, S., Krupke, C. H., Esker, P. D., Varenhorst, A., Arneson, N. J., Bradley, C. A., ... Conley, S. P. (2019a). Neonicotinoid seed treatments of soybean provide negligible benefits to US farmers. Scientific Reports, 9, article 11207. https://doi.org/10.1038/s41598-019-47442-8

Mourtzinis, S., Marburger, D. A., Gaska, J. M., & Conley, S. P. (2016). Characterizing soybean yield and quality response to multiple prophylactic inputs and synergies. Agronomy Journal, 108, 1337–1345. https://doi.org/10.2134/agronj2016.01.0023

Mourtzinis, S., Rattalino Edreira, J. I., Grassini, P., Roth, A. C., Casteel, S. N., Ciampitti, I. A., ... Conley, S. P. (2018). Sifting and winnowing: Analysis of farmer field data for soybean in the US northcentral region. Field Crops Research, 221, 130–141. https://doi.org/10.1016/j.fcr.2018.02.024

Mourtzinis, S., Specht, J. E., & Conley, S. P. (2019b). Defining optimal soybean sowing dates across the US. Scientific Reports, 9, article 2800. https://doi.org/10.1038/s41598-019-38971-3

Ng, S. J., Lindsey, L. E., Michel, A. P., & Dorrance, A. E. (2018). Effect of mid-season foliar fungicide and foliar insecticide applied alone and in-combination on soybean yield. Crop, Forage & Turfgrass Management, 4(1), 1–6. https://doi.org/10.2134/cftm2017.09.0067

R Development Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from http://www.r-project.org

Rattalino Edreira, J. I., Mourtzinis, S., Conley, S. P., Roth, A. C., Ciampitti, I. A., Licht, M. A., ... Grassini, P. (2017). Assessing causes of yield gaps in agricultural areas with diversity in climate and soils. Agricultural and Forest Meteorology, 247, 170–180. https://doi.org/10.1016/j.agrformet.2017.07.010

Rattalino Edreira, J. I., Cassman, K. G., Hochman, Z., van Ittersum, M. K., van Bussel, L. G. J., Claessens, L., & Grassini, P. (2018). Beyond the plot: Technology extrapolation domains for scaling out agronomic science. Environmental Research Letters, 13(5). https://doi.org/10.1088/1748-9326/aac092

Robinson, A. P., Conley, S. P., Volenec, J. J., & Santini, J. B. (2009). Analysis of high yielding, early-planted soybean in Indiana. Agronomy Journal, 101, 131–139. https://doi.org/10.2134/agronj2008.0014x

Steele, C. C., & Grabau, L. J. (1997). Planting date for early-maturing soybean cultivars. Agronomy Journal, 89, 449–453. https://doi.org/10.2134/agronj1997.00021962008900030013x

USDA NASS. (2019). Prices received for soybeans by month. USDA NASS. Retrieved from https://www.nass.usda.gov/Charts_and_Maps/Agricultural_Prices/pricesb.php

Zeiss, M. R., & Klubertanz, T. H. (1994). Sampling programs for soybean arthropods. In L. P. Pedigo & G. D. Buntin (Eds.), Handbook of sampling methods for arthropods in agriculture (pp. 539–601). Boca, Raton, FL: CRC Press.

Chapter 5: Foliar Fertilizers Rarely Increase Yield in US Soybean

This chapter was previously published in Agronomy Journal.

Matcham, E. G.[1], Vann, R. A. [2], Lindsey, L. E. [3], Gaska, J. M. [1], Lilley, D. T. [2], Ross, W. J. [4], Wright, D. L. [5], Knott, C. [6], Lee, C. D.[7], Mosely, D.[8], Singh, M.[9], Naeve, S.[10], Irby, J.T. [11], Weibold, W.[12], Kandel, H. [13], Lofton, J.[14], Inman, M.[15], Kleinjan, J.[16], Holshouser, D. L.[17], and Conley, S. P.[1] (2021). Foliar Fertilizers Rarely Increase Yield in US Soybean. *Agronomy Journal* 113(6): 5246 – 5253. https://doi.org/10.1002/agj2.20889

[1] Dep. of Agronomy, Univ. of Wisconsin–Madison, Madison, WI, 53706, USA
[2] Dep. of Crop and Soil Sciences, NC State Univ., Raleigh, NC, 27695, USA
[3] Dep. of Horticulture and Crop Science, The Ohio State Univ., Columbus, OH, 43210, USA
[4] Cooperative Extension Service, Univ. of Arkansas System, Division of Agriculture, Little Rock, AR, 72204, USA
[5] Dep. of Agronomy, Univ. of Florida, North Florida Research and Education Center, Quincy, FL, 32351, USA
[6] Dep. of Plant and Soil Sciences, Univ. of Kentucky, Princeton, KY, 42445, USA
[7] Dep. of Plant and Soil Sciences, Univ. of Kentucky, Lexington, KY, 40506, USA
[8] Dean Lee Research and Extension Center, LA State Univ., Alexandria, LA, 71302, USA
[9] Dep. of Plant, Soil and Microbial Sciences, MI State Univ., Lansing, MI, 48824, USA
[10] Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, St. Paul, MN, 55108, USA
[11] Dep. of Plant and Soil Sciences, MS State Univ., Mississippi State, MS, 39762, USA
[12] Division of Plant Sciences, Univ. of Missouri, Columbia, MO, 65211, USA
[13] Dep. of Plant Sciences, ND State Univ., Fargo, ND, 58105, USA
[14] Dep. of Plant and Soil Sciences, OK State Univ., Stillwater, OK, 74078, USA
[15] Pee Dee Research and Education Center, Clemson Univ., Florance, SC 29506, USA
[16] Dep. of Agronomy, Horticulture, and Crop Science, SD State Univ., Brookings, SD, 57007, USA
[17] Dep. of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State Univ., Tidewater Research Station, Suffolk, VA, 23437, USA

Abstract

Farmers have been interested in using foliar-applied nutrient products to increase soybean [Glycine max (L.) Merr.] yield since at least the 1970s, despite limited evidence that these products offer consistent yield increases when used prophylactically. Recently, interest in foliar fertilizer products for soybean production has been renewed, likely related to elevated soybean prices. Over the 2019 and 2020 growing seasons (46 site-years), agronomists in 16 states collaborated to test six foliar nutrient treatments (commercial mixtures of macro- and micro-nutrients) on soybean grain yield and composition. Soybean grain yield and composition differed among sites but not among foliar fertilizer treatments. Results show that prophylactic foliar fertilization is likely to decrease the profitability of soybean production. Foliar fertilizer products tested in this study and similar products should not be recommended to U.S. soybean farmers in the absence of visual symptoms of nutrient deficiency.

Abbreviations

NIR, near-infrared spectroscopy

Core Ideas

- The tested prophylactic foliar fertilizers did not increase soybean yield.
- Foliar fertilizers did not change grain composition.
- Prophylactic foliar fertilizers tested decreased profitability.

Introduction

Annual soybean [*Glycine max* (L.) Merr.] production in the United States varied between 97 and 121 billion kilograms between 2015 and 2020 (USDA-NASS, 2021b). Many soybean farmers are interested in foliar products that apply a mixture of micronutrients and macronutrients and can be tank-mixed with insecticides and fungicides and applied during reproductive growth. This timeline corresponds with a period of high nutrient uptake for soybean (Gaspar et al., 2017). There has been interest in testing different fertilizer methods that may increase soybean yield as the United States has reached record high soybean yields, since some farmers are concerned that fields with higher yields may need nutrients supplied at different times or in different forms. Recently, questions regarding foliar fertilizers have been increasing. Thus, interest in foliar fertilizer products for soybean production has been renewed.

Past foliar fertilizer research has shown inconsistent impacts on soybean yield. In the 1970s, a study in Iowa associated up to 538 kg ha^{-1} yield increases to foliar application of N, P, K, and S in combination, while a similar study in Wisconsin reported no yield increase in soybean yield with P, K, and S foliar applications and a smaller yield increase when N was applied foliarly (Garcia & Hanway, 1976; Syverud et al., 1980). A contemporaneous study in Minnesota showed a yield benefit to N–P–K–S foliar fertilization in only 1 out of 16 trial site-years, and no yield benefit to micronutrient application (Poole et al., 1983).

Larger studies in the 1990s in Iowa showed small, inconsistent increases in yield with early-season prophylactic foliar fertilizer application. Treatments contained N, P, and K and increased yield as compared to untreated controls by 30–60 kg ha^{-1} at 10 of the 48 site-years (Haq & Mallarino, 1998). In a subsequent on-farm strip trial, comparing an untreated control to soybean treated with 1.2 kg N, 3.1 kg P, and 5.9 kg K (elemental rate per hectare) during reproductive growth, there was a 35 kg ha^{-1} increase in soybean yield at one out of eight sites (Mallarino et al., 2001). The associated small-plot trial tested a wider range of nutrient rates and had two responsive locations out of 18 with a

93–360 kg ha^{-1} increase in soybean yield when N, P, and K were applied (Mallarino et al., 2001).

Agronomists in Michigan have performed extensive foliar fertilizer trials in soybean since 2000. Out of the 51 location N–P–K product trials, four locations had increased yield and the fertilized plots had lower profitability than the unfertilized control at all locations. Foliarly applied N alone in 18 Michigan trials resulted in higher yield in three trial locations (Staton, 2019).

Prophylactic application of micronutrients has shown similarly minimal effects on soybean yield. Between the 1980s and today, trials in Iowa, Minnesota, and Michigan have not shown a yield increase in soybean associated with Fe, Zn, B, Co, Cu, Zn, Mn, or Mo foliar prophylactic application (Mallarino et al., 2001; Poole et al., 1983). Rare response to micronutrients has been observed in Ohio, where <2% of Mn trials have seen an increase in yield when fertilizer was applied and <5% of trials treated with a mixture of Mn, Fe, Cu, Mo, and B fertilizers had an observed soybean yield increase (Sharma et al., 2018). In Michigan fields with high pH lakebed soils that are likely to respond to Mn application, foliar Mn application only increased yield when it was applied after visual symptoms of nutrient deficiency began, but not when Mn was applied prophylactically (Staton, 2019).

One challenge to assessing the efficacy of foliar fertilization in soybean is that when yield increases have been observed, the magnitude of yield improvement is relatively small and generally does not pay for the cost of application. Additionally, it is difficult to identify field conditions where agronomists should recommend foliar fertilizer application in soybean because past studies have shown that soybean yield response to foliar fertilizer is inconsistent. Despite the lack of evidence that soybean yield and farm profit increase with prophylactic foliar nutrient application in the United States, these products are still commonly marketed for soybean in the United States.

Past foliar fertilization in soybean research in the United States has been isolated to a few states in the upper Midwest (Iowa, Wisconsin, Michigan, and Ohio). This study is a coordinated effort across 16 states (Arkansas, Florida, Kentucky, Louisiana, Michigan, Minnesota, Missouri, Mississippi, Ohio, Oklahoma, North Carolina, North Dakota, South Carolina, South Dakota, Wisconsin, Virginia) that allowed us to test the effects of macronutrient and micronutrient foliar fertilization throughout the primary soybean-producing region of the United States and includes a broad range of commercially available foliar fertilizer products to assess the efficacy of both macronutrient and micronutrient applications. The objectives of this study were to (a) identify soybean grain yield response to prophylactic foliar fertilizer application across a broad range of environments, (b) determine if foliar fertilizer application changes soybean grain composition, and (c) conduct economic analyses on the value of these products in U.S. soybean-growing environments.

Methods

Field Methods

In 2019 and 2020, small-plot trials were established at a total of 46 sites in 16 states (Figure 5.1). Six foliar nutrient products (Table 5.1) and the untreated control were applied in a randomized complete block design with four to eight replications depending on site. Products were selected with the input of industry professionals to identify foliar fertilizers that are nationally marketed to soybean producers. Products were applied at soybean growth stage R3 to align with commonly used fungicide and insecticide application timing. Growth stage R3 was defined by one pod of at least 5 mm in length on one or more of the top four nodes of the plant (University of Wisconsin-Extension, 2017). Due to lack of product availability, HarvestMore UreaMate was not applied in Lexington, KY, in 2019 and Smart Quatro Plus was not applied at any 2019 Wisconsin sites and the Arlington, WI, site in 2020.

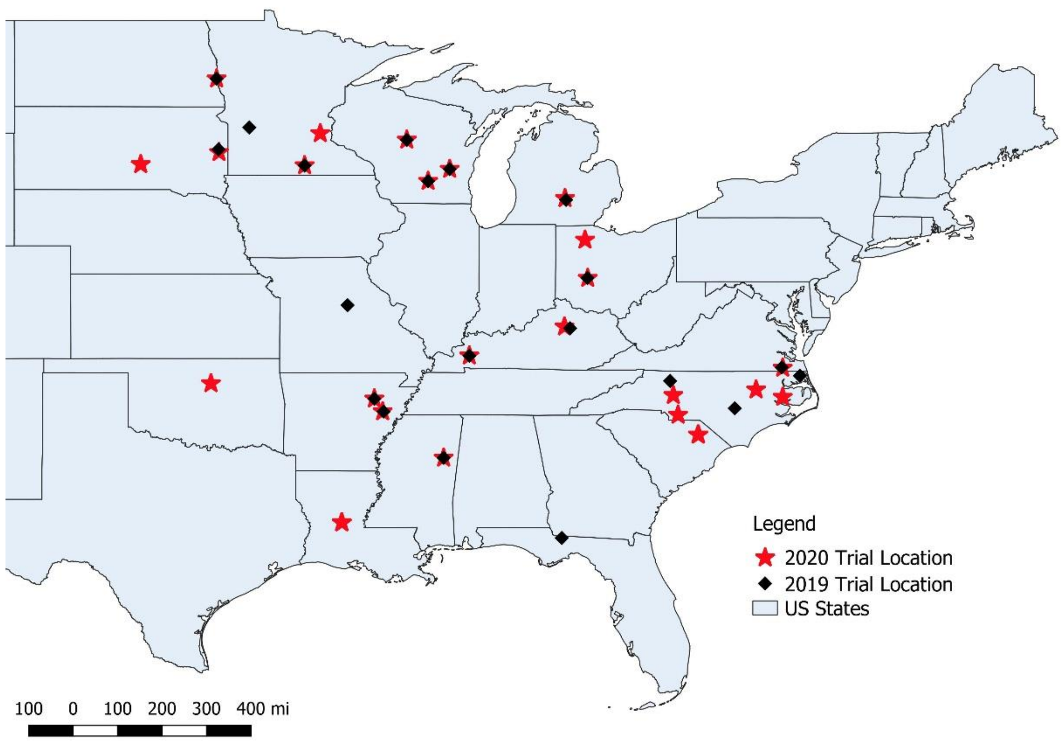

Figure 5.1. Trial locations in 2019 and 2020, displayed with red stars and black diamonds, respectively. South Carolina and Louisiana have two nearby sites each that appear as a single marker due to the scale of this map.

Table 5.1. List of foliar products names, application rate, cost of product, and nutrients applied for each treatment.

Treatment Name	**Manufacturer**	**Application Rate**	**Cost of Product**	**N**	**P**	**K**	**S**	**Mn**	**Fe**	**Mo**	**Zn**	**B**	**Other**
			USD ha^{-1}	**---------- kg ha^{-1} ----------**									
FertiRain	AgroLiquid	28.0 l ha^{-1}	$55	3.1	1.0	1.0	0.6	0.02	0.03	-	0.03	-	-
Sure-K	Agroliquid	28.0 l ha^{-1}	$48	0.7	0.3	1.0	-	-	-	-	-	-	-
HarvestMore Ureamate	Stoller	2.8 kg ha^{-1}	$12	0.1	0.3	-	-	0.01	-	0.002	0.01	-	Ca, Mg, B, Co, Cu
Smart B-Mo	Brandt	1.2 l ha^{-1}	$9	-	-	-	-	-	-	0.007	-	0.08	-
Smart Quatro Plus	Brandt	4.7 l ha^{-1}	$16	-	-	-	0.04	0.09	-	0.003	0.09	0.07	-
Maximum NPact K	Nutrien	14.0 l ha^{-1}	$52	2.1	-	2.1	-	-	-	-	-	-	-
Untreated Control	-	-	-	-	-	-	-	-	-	-	-	-	-

Composite soil samples were taken from each replication at each site in the spring. Samples were air-dried, and soil physical and chemical properties were measured by A&L Great Lakes (Fort Wayne, IN). Soil sample results and site management practices can be found in Supplemental Table S1.

Backpack sprayers were used to to apply foliar fertilizers at the R3 growth stage. Visual symptoms of nutrient deficiency were not present at any site prior to foliar fertilizer application. Selected application rates (Table 5.1) were within the range of rates recommended on each product's label. Leaf tissue samples were taken before foliar products were applied at R3 and 2 wk following application. At both sampling time points, the newest fully-expanded trifoliate leaf was collected from 20 plants per plot. Samples were dried in paper bags (dryer temperature 38–54 ˚C) until constant weight was achieved and shipped to the North Carolina Department of Agriculture & Consumer Services Agronomic Division (Raleigh, NC) for analysis of N, P, K, Ca, Mg, S, Fe, Mn, Zn, Cu, and B. The North Carolina Department of Agriculture Agronomic Division's Plant Tissue lab measures N using oxygen combustion with gas chromatography, NO_3^-–N using an electrode, and all other nutrients using HNO3 closed vessel microwave digestion followed by inductively coupled plasma (ICP). Tissue samples were taken from all sites in 2020 but were not collected at seven sites in 2019: Newport, AR; Pine Tree, AR; Florida; Princeton, KY; Missouri; Minnesota Lake, MN; and Danvers, MN.

Yield data were collected using plot combines at each site and adjusted to 130 g kg^{-1} moisture concentration. Grain samples were taken at harvest from all sites except in Missouri in 2019; Oklahoma in 2020; and Hoytville, OH, in 2020. In 2019, grain protein and oil concentration were analyzed via near-infrared spectroscopy (NIR) using a Perten DA7520 machine. The NIR calibration curve was developed from hundreds of soybean samples with known composition values (Soybean NIR Consortium). In 2020, grain protein and oil were determined using the Perten Instruments Inframatic 9500 NIR Grain Analyzer. Calibration curves were provided and validated by Perten and were normalized using a polystyrene reference standard. Grain protein and oil concentration were reported at a standard moisture of 130 g kg^{-1}.

Cost of foliar fertilizer products were assessed by calling retailers in the study region in 2019 and averaging the cost of product per hectare at the application rate used in the study (Table 1). Partial profits were calculated by multiplying yield by the price of soybean grain and subtracting the cost of the foliar fertilizer product. Application costs were not considered since these products are frequently applied by farmers as part of a tank-mix with foliar fungicides and insecticides. Calculations were performed at \$0.550 and \$0.367 kg^{-1} to be reflective of recent soybean prices (USDA-NASS, 2021a).

Analysis Methods

Change in tissue nutrient concentration was calculated by subtracting nutrient concentration from the pre-application samples from the nutrient concentration from the 2 wk post application samples. Yield, protein, oil, and change in tissue nutrient concentration values that fell outside of three standard deviations of each site's mean value were considered outliers and removed from further analysis. Yield data was collected or 1,868 plots in total, and 34 of those observations (<2%) were considered outliers and removed from further analysis because they fell outside of three standard deviations of each site's mean yield.

Mixed-model ANOVA was performed using R 3.6.2 and the package lme4. All site-years were analyzed together with treatment and site-year considered fixed variables, and replication nested within site-year being considered a random variable. Throughout the manuscript, site-years will be referred to as “site.” Degrees of freedom were estimated using Kenward–Rogers approximation to account for unequal replication among site years. Data were not transformed, and residuals were plotted to assess for normality. Means comparisons were performed using Bonferroni adjustments.

Results

Soybean Grain Yield

In 2019, the highest-yielding site was Arlington, WI (5,513 kg ha^{-1}) and the lowest-yielding site was Yadkin, NC (1,824 kg ha^{-1}). Yields were overall higher in 2020, with the highest yields observed at Arlington, WI (5,592 kg ha−1). Figure 5.2 compares site average yield for each treatment to the control at each site, and additional summaries of site mean yield are available in Supplemental Table S2. Most observations fall near or on the 1:1 line (Figure 5.2), indicating that the treated plots and untreated control plots yielded similarly. The few points that fell above the 10% yield increase line tended to have yields near 4,000 kg ha^{-1}. All sites with yields higher than 5,000 kg ha^{-1} had mean treated plot yield within 10% of the untreated control plots for all foliar fertilizer products (Figure 5.2). Observed differences in yield among treatments were not statistically significant ($F = 0.23$, $p = .9663$), although there was a significant difference in yield among sites. There was not a significant interaction between site and treatment (Table 5.2).

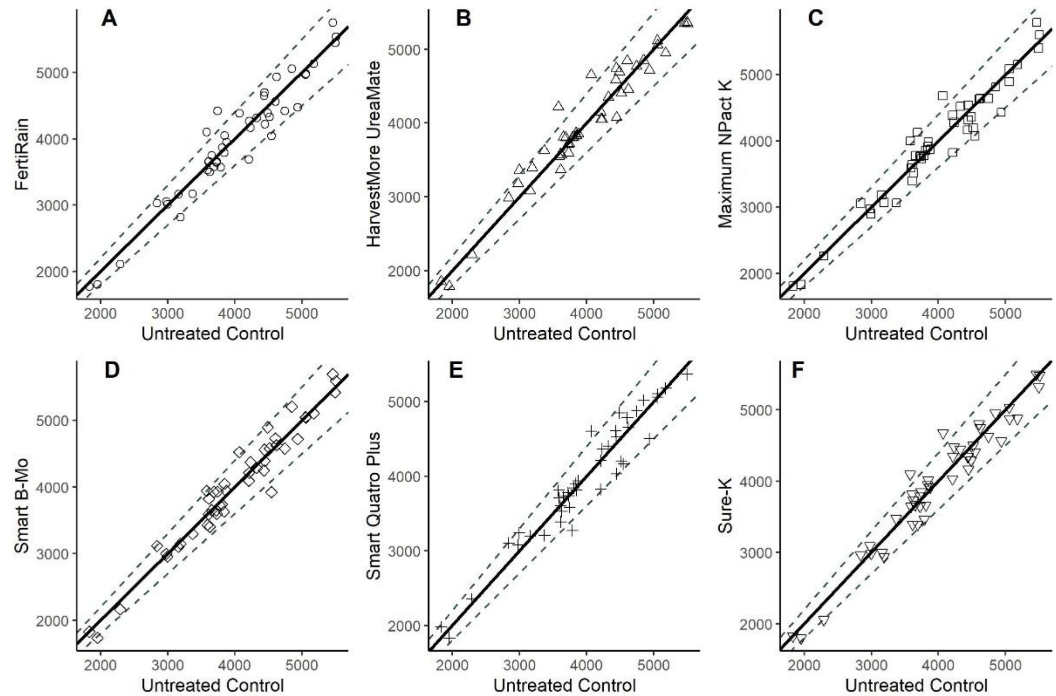

Figure 5.2. Average yield (kg ha−1) at each site for each treatment plotted against the average yield of the untreated control at the same site. Solid lines represent x = y, and the dashed lines represent ±10% of yield. (a) Average yield of plots treated with FertiRain compared to Untreated Control plots, (b) average yield of plots treated with HarvestMore UreaMate compared to Untreated Control plots, (c) average yield of plots treated with Maximum NPact K compared to Untreated Control plots, (d) average yield of plots treated with Smart B-Mo compared to Untreated Control plots, (e) average yield of plots treated with Smart Quatro Plus compared to Untreated Control plots, and (f) average yield of plots treated with Sure-K compared to Untreated Control plots.

Table 5.25. Results from Analysis of Variance used to identify differences in yield, protein, and oil based on treatment, site, and their interaction.

		F-value	p-value
	Treatment (T)	0.23	0.9663
Yield	**Site (S)**	61.05	<0.001
	T × S	1.00	0.4812
	Treatment (T)	1.37	0.2248
Protein	**Site (S)**	557.92	<0.001
	T × S	1.15	0.0703
	Treatment (T)	1.62	0.1382
Oil	**Site (S)**	392.72	<0.001
	T × S	1.17	0.0490

An additional ANOVA model was run to determine whether low- (<3,000 kg ha^{-1}), medium- (3,000–4,000 kg ha^{-1}), or high-yielding (>4,000 kg ha^{-1}) sites responded to treatment differently, with sites grouped into yield environments based on the average yield of the untreated control. All site-years were analyzed together with treatment and yield environment considered fixed variables, and site-year nested within yield environment and replication nested within site-year and yield environment being considered random variables. This model confirmed that there were neither differences in yield among treatments ($F = 0.44$, $p = .8532$), nor an interaction between treatment and yield environment ($F = 0.89$, $p = .5540$).

The sites tested in this trial include a wide range of soil chemical and physical properties (Supplemental Table S1). Even at sites such as Princeton, KY (2019 and 2020) and North Dakota (2019) where soil test P concentration was below 15 mg kg^{-1}, there was not a yield response to treatment. Site soil pH ranged from 4.7 to 8.3, but sites did not have significant differences in response to nutrient application even though high pH can reduce micronutrient availability. Given the uniformity of the response across these 46 sites, there is no evidence that foliar fertilizers increase soybean yield in the absence of visual symptoms of nutrient deficiency. Similar results were observed in a smaller geographic area in past trials from Iowa and Michigan, where micronutrient and macronutrient foliar fertilization did not consistently increase soybean grain yield (Mallarino et al., 2001; Staton, 2019).

Grain Composition

Grain samples from each plot were collected in 19 sites in 2019 and 24 sites in 2020. Average protein and oil content across all sites and treatments was 376 and 206 g kg^{-1}, respectively. Differences in grain protein and oil content were observed among sites but not treatments (Table 5.2). Most sites had similar oil content across all treatments, but there was a treatment × site interaction related to two differences between sites: the Ohio 2019 site had approximately 0.5% higher average oil content in the untreated control and FertiRain-treated plots and the Sampson, NC, 2019 site had slightly lower oil content in the plots treated with Sure-K as compared with other treatments. At nutrient application rates currently recommended by foliar fertilizer manufacturers, there is no evidence that fields that receive foliar fertilizer should be expected to have different grain protein or oil content as compared to fields that do not receive foliar fertilizer.

Leaf Nutrient Content

Across all sites and treatments, average leaf tissue Ca, Mn, and B concentration increased by 1.5, 78, and 19 g kg^{-1}, respectively, between the pre-application sampling timepoint and the 2 wk after application timepoint (Supplemental Table S3). Leaf tissue S concentration did not change between sampling timepoints. Concentration of N, P, K, Mg, Fe, and Cu decreased slightly (<10 g kg^{-1}) between the preapplication sampling timepoint and the 2 wk after application timepoint, likely due to soybean plants partitioning an increasing proportion of their nutrient uptake to seeds relative to other plant parts after R4 (Gaspar et al., 2017). Observed decreases in tissue nutrient concentrations were <10 g kg^{-1} on average, with the exception of Fe which decreased by an average of 70 g kg^{-1} between the sampling timepoints.

Across all nutrients tested (N, P, K, Ca, Mg, S, Fe, Mn, Cu, and B), there was a significant difference in leaf tissue nutrient content among sites (Table 5.3). Leaf tissue Mn, Cu, and B content varied among treatments (Table 5.3). While past studies indicate that fields with low leaf tissue P concentration may be more likely to see a yield response to foliar fertilization (Haq & Mallarino, 1998), foliar fertilizer treatments in our study and others did not necessarily cause differences in leaf tissue nutrient concentrations for most nutrients. Application of micronutrients such as Cu and B are more likely to result in differences in leaf tissue micronutrient concentration. Application of P frequently does not change leaf tissue P concentration (Alt et al., 2018; Haq & Mallarino, 1998; Nelson et al., 2012).

Table 5.3. Results from analysis of variance used to identify differences in leaf nutrient concentration based on treatment, site, and the interaction of site and treatment.

	Nitrogen		**Phosphorus**		**Potassium**		**Calcium**		**Manganese**	
	F-value	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**
Treatment (T)	1.20	0.3037	0.89	0.5029	1.28	0.2614	0.63	0.7026	0.94	0.4666
Site (S)	31.37	<0.001	3.28	<0.001	5.92	<0.001	39.39	<0.001	47.42	<0.001
T × S	0.82	0.9673	1.18	0.0565	1.19	0.0422	0.98	0.5522	1.18	0.0489
	Sulfur		**Iron**		**Manganese**		**Copper**		**Boron**	
	F-value	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**
Treatment (T)	0.55	0.7728	1.62	0.1368	2.58	0.0174	6.86	<0.001	40.16	<0.001
Site (S)	27.29	<0.001	14.50	<0.001	16.56	<0.001	21.84	<0.001	52.65	<0.001
T × S	1.00	0.4994	1.00	0.5019	0.78	0.9875	1.24	0.0168	2.28	<0.001

Cost of Foliar Fertilzer Products

Cost of foliar fertilizer products ranged from US$9 to $55 ha^{-1} (Table 5.1). Partial profits were different among treatments and sites at both tested soybean grain prices ($0.550and $0.367 kg^{-1}), and there was no interaction between treatment and site at either tested soybean grain price (Table 5.4). At $0.550 kg^{-1}, plots treated with Maximum NPact K had $60 ha^{-1} lower profits than the untreated control and at $0.367 kg^{-1}, plots treated with Maximum NPact K or FertiRain had lower profits than the untreated control by $58 and $53 ha^{-1}, respectively (Table 5.5). While other treatments did not have statistically lower profits than the untreated control at the tested grain prices, application of foliar fertilizer products included in this study would not increase profit since foliar fertilizer treatments did not statistically increase soybean grain yield. Further reductions in profit may occur when applying foliar fertilizer using a ground-based applicator since wheel damage can reduce soybean yield by 3–5% after R1 (Hanna et al., 2008).

Table 5.46. Results from analysis of variance used to identify differences in partial profits based on treatment, site, and the interaction of site and treatment.

		F-value	p-value
Profit at soybean grain price of $0.550 kg^{-1}	**Treatment (T)**	5.74	<0.001
	Site (S)	59.31	<0.001
	T x S	1.01	0.4396
Profit at soybean grain price of $0.367 kg^{-1}	**Treatment (T)**	5.74	<0.001
	Site (S)	59.31	<0.001
	T x S	1.01	0.4396

Table 5.5. Mean partial profit at two soybean grain prices and mean grain yield, oil concentration, and protein concentration among foliar fertilizer treatments.

Treatment	**Mean partial profit at soybean grain price of $0.550 kg^{-1}**	**Mean partial profit at soybean grain price of $0.367 kg^{-1}**	**Mean yield**	**Mean grain oil concentration**	**Mean grain protein concentration**
	USD kg^{-1}	**USD kg^{-1}**	**kg ha^{-1}**	**g kg^{-1}**	**g kg^{-1}**

Untreated Control	2202 a[a]	1470 a	4004[b]	20.6	37.5
Smart B-Mo	2198 ab	1464 a	4013	20.6	37.6
HarvestMore UreaMate	2193 ab	1459 a	4008	20.5	37.6
Smart Quatro Plus	2168 ab	1442 ab	3972	20.6	37.6
FertiRain	2151 ab	1417 b	4012	20.6	37.5
Sure-K	2149 ab	1418 b	3994	20.6	37.6
Maximum NPact K	2142 b	1412 b	3990	20.6	37.6

[a]Means not sharing common letters within each column denote statistical differences among treatments (α = .05). Bonferroni adjustments were used to adjust for multiplicity.
[b]Means separation was not performed for yield or grain composition (oil and protein) due to no significant differences among treatments.

Conclusions

Prophylactic foliar fertilizer applications did not consistently increase soybean yield or alter grain composition when applied at rates recommended by their manufacturer, and foliar fertilizer application may decrease farm profitability. None of the tested foliar fertilizer treatments had higher partial profits than the untreated control. Agronomists and farmers interested in increasing soybean yield or farm profitability are unlikely to see benefit from foliar fertilizer application in the absence of visual symptoms of nutrient deficiency.

Acknowledgements

This research was funded by Soybean Checkoff funds through the Qualified State Soybean Board in each state. Additional support was provided by the United Soybean Board through the Science For Success initiative.

Author Contributions

Emma G. Matcham: Formal analysis; Methodology; Project administration; Writing-original draft; Writing-review & editing. R. Atwell Vann: Conceptualization;

Data curation; Funding acquisition; Methodology; Project administration; Resources; Supervision; Validation; Writing-original draft; Writing-review & editing.; Laura E. Lindsey: Conceptualization; Funding acquisition; Methodology; Resources; Writing-review & editing; John M. Gaska: Project administration; Resources; Writing-review & editing. Dylan T. Lilley: Data curation; Writing-review & editing. W. Jeremy Ross: Funding acquisition; Resources; Writing-review & editing. David L. Wright: Funding acquisition; Resources; Writing-review & editing. Carrie Knott: Funding acquisition; Resources; Writing-review & editing. Chad D. Lee: Funding acquisition; Resources; Writing-review & editing. David Moseley: Funding acquisition; Resources; Writingreview & editing. Maninder Singh: Funding acquisition; Resources; Writing-review & editing. Seth Naeve; Funding acquisition; Resources; Writing-review & editing. J. Trenton Irby: Funding acquisition; Resources; Writing-review & editing. William Wiebold: Funding acquisition; Resources; Writing-review & editing. Hans Kandel: Funding acquisition; Resources; Writing-review & editing. Josh Lofton: Funding acquisition; Resources; Writing-review & editing. Matthew Inman: Funding acquisition; Resources; Writingreview & editing. Jonathon Kleinjan: Funding acquisition; Resources; Writing-review & editing. David L. Holshouser: Funding acquisition; Resources; Writing-review & editing. Shawn P. Conley: Funding acquisition; Project administration; Resources; Supervision; Writing-review & editing. All authors managed plots and data collection within their respective states and contributed to writing and editing the manuscript. Dylan T. Lilley collected price data. Emma G. Matcham, R. Atwell Vann, Laura E. Lindsey, and Shawn P. Conley coordinated data collection and planned the trial.

Conflict of Interest

The authors declare no conflicts of interest.

Supplemental Materials

Supplemental Tables S1 and S2 are available as a downloadable .docx file from the *Agronomy Journal* website: https://doi.org/10.1002/agj2.20889

Works Cited

Alt, D., Ng, S. J., Grusenmeyer, J., & Lindsey, L. E. (2018). Seed yield and quality of transgenic high-oleic and conventional soybean as influenced by foliar manganese application. *Crop Science*, 58(2), 874–879. https://doi.org/10.2135/cropsci2017.06.0391

Garcia, L., R., & Hanway, J. J. (1976). Foliar fertilization of soybeans during the seed-filling period. *Agronomy Journal*, 68(4), 653–657. https://doi.org/10.2134/agronj1976.00021962006800040030x

Gaspar, A. P., Laboski, C. A. M., Naeve, S. L., & Conley, S. P. (2017). Phosphorus and potassium uptake, partitioning, and removal across a wide range of soybean seed yield levels. *Crop Science*, 57(4), 2193–2204. https://doi.org/10.2135/cropsci2016.05.0378

Hanna, S. O., Conley, S. P., Shaner, G. E., & Santini, J. B. (2008). Fungicide application timing and row spacing effect on soybean canopy penetration and grain yield*. Agronomy Journal*, 100(5), 1488–1492. https://doi.org/10.2134/agronj2007.0135

Haq, M. U., & Mallarino, A. P. (1998). Foliar fertilization of soybean at early vegetative stages. *Agronomy Journal*, 90(6), 763–769. https://doi.org/10.2134/agronj1998.00021962009000060008x

Mallarino, A. P., Haq, M. U., Wittry, D., & Bermudez, M. (2001). Variation in soybean response to early season foliar fertilization among and within fields. *Agronomy Journal,* 93(6), 1220–1226. https://doi.org/10.2134/agronj2001.1220

Nelson, K. A., Motavalli, P. P., Stevens, W. E., Kendig, J. A., Dunn, D., & Nathan, M. (2012). Foliar potassium fertilizer additives affect soybean response and weed control with glyphosate. *International Journal of Agronomy,* 2012, 1–10. https://doi.org/10.1155/2012/461894

Poole, W. D., Randall, G. W., & Ham, G. E. (1983). Foliar fertilization of soybeans. I. Effect of fertilizer sources, rates, and frequency of application. *Agronomy Journal,* 75(2), 195–200. https://doi.org/10.2134/agronj1983.00021962007500020010x

Sharma, S., Culman, S., Fulford, A., Lindsey, L., Alt, D., & Looker, G. (2018). *Corn, soybean, and alfalfa yield responses to micronutrient fertilization in Ohio, AGF-519.* Ohio State University Extension. https://ohioline.osu.edu/factsheet/agf-519

Staton, M. J. (2019). *Foliar fertilizer applications to soybeans are rarely profitable.* Michigan State University Extension.

Syverud, T. D., Walsh, L. M., Oplinger, E. S., & Kelling, K. A. (1980). Foliar fertilization of soybeans (Glycine max L.). *Communications in Soil Science and Plant Analysis,* 11(6), 637–651. https://doi.org/10.1080/00103628009367069

University of Wisconsin-Extension. (2017). *A visual guide to soybean growth stages.* http://ipcm.wisc.edu/apps/beancam/

USDA-NASS. (2021a). Prices received: Soybean prices received by month, US. USDA-NASS.

https://www.nass.usda.gov/Charts_and_Maps/Agricultural_Prices/prices.php

USDA-NASS. (2021b). Soybean production. Charts and maps. USDANASS. https://www.nass.usda.gov/Charts_and_Maps/graphics/soyprod.pdf

Chapter 6: Filtering, editing, and cropping yield maps in an R environment with the package cleanRfield

This chapter has been accepted for publication in Agronomy Journal.

Matcham, E. G.[1], Matias, F.I.[2], Luck, B.[3], and Conley, S. P.[1] (2022). Filtering, editing, and cropping yield maps in R environment with the package cleanRfield. Accepted for publication at Agronomy Journal.

[1] Dep. of Agronomy, Univ. of Wisconsin–Madison, Madison, WI, 53706, USA
[2] Dep. of Plant Sciences, North Dakota State Univ., Fargo, ND, 58102, USA
[3] Dep. of Biological Systems Engineering, Madison, WI, 53706, USA

Abstract

Cleaning yield monitor observations to remove erroneous points can improve the accuracy of yield estimates used for farm record keeping or on-farm research data collection, but current practices are time-intensive and cumbersome. cleanRfield is an open-source R package to improve the efficiency of processing spatial agricultural data such as yield maps. Compared to current standard yield monitor data cleaning solutions, cleanRfield can read and interpret a broader range of input data formats. Other key features of cleanRfield include automatic field boundary delineation and batch processing of data from multiple fields. In this Scientific note, we overview functions within the cleanRfield package and introduce an integrative pipeline to evaluate and visualize yield monitor data. The package is being distributed under the GNU General Public License 2, and a more detailed tutorial including downloading instructions is available at https://github.com/filipematias23/cleanRfield.

Abbreviations

comma separated values, CSV; Global Navigation Satellite System, GNSS; GNU's Not Unix, GNU; standard deviation, sd

Core Ideas

- Open-source tool for cleaning and filtering yield maps using R language
- Easily builds vector polygon files and rasters of field boundaries from yield maps
- Reads and interprets shapefiles, CSVs, or textfiles
- An integrative pipeline to evaluate and visualize yield monitor data

Introduction

Current combine and yield monitor technology coupled with global navigational satellite system (GNSS) technology allows farmers to utilize yield maps to measure yield within management zones or conduct precision agriculture experiments with on-farm trials. Over half of US soybean farms use yield monitors on combines to map within-field yield patterns (Schimmelpfennig, 2016). Yield monitors typically measure grain flow using volume samples or pressure plates within the clean grain elevators of combines at regular intervals, commonly every one to three seconds (Casady & Shannon, 1998). Yield estimates are recorded alongside other metrics, including grain moisture, combine speed, location, and time (Casady & Shannon, 1998).

Raw yield files are stored on monitors using manufacturer-specific, proprietary formatting and must be converted to shapefile (.shp) or text file (.txt) format for further processing by using agricultural software (Griffin, Brown, &

Lowenberg-DeBoer, 2007). This initial conversion usually corrects for grain flow delay, which helps place each yield monitor observation in the correct geographic location within the field (Kleinjan et al, 2002; Simbahan, Doberman, & Ping, 2004). Once the data is stored in a general file format, it can be further processed, mapped, or summarized.

Processing and filtering of yield monitor observations improves the accuracy of whole-field and within-field yield estimates (Kleinjan et al., 2002; Simbahan, Dobermann & Ping, 2004; Griffin, Brown, & Lowenberg-DeBoer, 2007; Kharel et al., 2019), and cleaning yield data can improve the accuracy of conclusions from on-farm research and farmer decision support tools (Griffin, Brown, & Lowenberg-DeBoer, 2007; Kharel et al., 2019). Current standard practice for filtering yield maps includes using Yield Editor, a free software program written by USDA (Sudduth, Drummond, & Myers, 2012; USDA Ag Data Commons, 2021). Automatic filtering is available for corn, soybean, wheat, and other crops in Yield Editor, and generally results in similar final maps to manual filtering performed by experienced Yield Editor users (Kharel et al., 2019). Removing low-moisture data is not currently a part of Yield Editor's automatic filtering, but adding a low-moisture filter does improve the quality of final cleaned yield maps and indicates that there is benefit to studying additional yield filtering options (Kharel et al., 2019).

Errors in yield maps can be systematic or random, and calibrating yield monitors and processing the data before mapping or summarizing can improve the accuracy of the yield estimates (Simbahan, Dobermann & Ping, 2004). Calibrating yield monitors before harvesting a field using procedures recommended by the monitor's manufacturer helps reduce errors related to combine speed or grain flow. Other inaccurate yield monitor observations are more likely to occur on steep slopes, places where the combine speed is fluctuating, in point rows where the full header width is not being utilized, or near field borders where the combine is turning. Cleaning maps can reduce the impact of these errors on yield estimates.

Automatic filtering using Yield Editor is a commonly accepted method for cleaning yield maps, but it has many limitations. Within Yield Editor, all yield maps must be formatted as AgLeader Advanced or Greenstar text files and must include longitude (decimal degrees), latitude (decimal degrees), flow (lbs s^{-1}), GNSS time (s), logged interval (s), distance (in), swath width (in), moisture (percentage), header status (up or down), and pass number without flexibility for processing files with data in different units (metric units are not supported within Yield Editor), or with columns in a different order. Some combines also record other information like elevation and combine speed that may be useful for filtering.

The package cleanRfield is a compilation of functions to clean and filter observations from yield monitors or other agricultural spatial point data. Yield monitors are prone to error, and filtering the observations or removing observations from near field boundaries can improve estimates of whole-field yield, combine speed, grain moisture, or other parameters. In this package, users can easily select filters thresholding for one or more attributes and prepare a smaller dataset for analysis or decision making.

In this Scientific Note, we illustrate key features of the cleanRfield package using soybean yield maps from Wisconsin, USA. The package is being distributed under the GNU General Public License 2. A more detailed tutorial is available at https://github.com/filipematias23/cleanRfield.

Materials and Methods

Data Preparation

Yield data is collected using combine-mounted yield monitors. Data quality is significantly improved by properly calibrating the yield monitor. Calibration instructions vary by monitor make and model, but multipoint calibration improves calibration accuracy (Bergman, 2020).

Monitors using mass-flow, weight, optical, and nuclear sensors all collect point data observations that are compatible with filtering data in cleanRfield. After data are collected on yield monitors, the raw yield maps (.yld files or other machinery-specific file type) must be exported off the monitor as shapefiles (.shp), comma separated value (.csv) or other delimited text format. If data are exported off the monitor as .yld files or another unsupported file type, they may be converted to a supported file type (.shp, .csv, or delimited text) using an intermediate software that is compatible with the combine and monitor before opening the data within the R environment. Yield maps must be formatted as vector data when exported, as clearRfield does not currently support filtering of raster data.

Cropping Fields to Areas of Interest

The function cropField() allows drawing one or more polygons within the dataset to be used as a boundary file (Table 6.1). The output of cropField() is a List of three items within the R working environment that includes the drawn boundary saved as a SpatialPolygonDataFrame, the yield observations from the input map that fall within the drawn boundary saved as a SpatialPointsDataFrame, and points that define the corners of the boundary file saved as a SpatialPoints object. This approach is useful for selecting or removing a specific region in the field that was affected by biotic or abiotic stresses during the season or evaluate regions with different yield performance. It allows overlaying multiple data layers and making informative visualizations. Users can apply different filters or cleaning criteria to different field regions to better evaluate and understand yield and other metrics at the sub-field scale. The output polygon from cropField() also can be used as a boundary for other point datasets collected at the same location in different years, allowing cleanRfield users to evaluate the field performance across time. cropField() facilitates measuring the impact of different management approaches (e.g., fertilization, pesticide application, irrigation, etc.) between locations and crop years.

Table 6.1. Functions within the cleanRfield package.

Function Name	Formal Class of Input Objects	Formal Class of Output Objects	Function Description
cropField()	LargeSpatialPointsDataFrame, number of polygons to draw, number of points that define the boundary of each polygon	List of 3: LargeSpatialPointsDataFrame, SpatialPolygonDataFrame, and a SpatialPoints object per polygon	Opens a map of the input data set and allows users to use point-and-click functionality to define regions to be evaluated
boundaryField()	LargeSpatialPointsDataFrame or RasterStack	SpatialPolygonsDataFrame or raster	Allows users to manually draw a boundary around a field using point-and-click functionality or automatically draws a boundary around a field. Requires a RasterStack input for automatic boundary drawing
rasterField()	LargeSpatialPointsDataFrame, trait of interest within the lLargeSpatialPointsDataFrame, and the resolution to use for the raster	RasterStack	Converts the point data from the user-supplied yield map to a raster file with the user-defined resolution
bufferField()	LargeSpatialPointsDataFrame, SpatialPolygonsDataFrame, linear distance to reduce the field size	List of 2: SpatialPolygons and SpatialPointsDataFrame	Removes point observations that fall within the user-defined distance from the border of the input SpatialPolygonsDataFrame, and provides a new SpatialPolgyon that matches the extent of the output SpatialPointsDataFrame
sampleField()	LargeSpatialPointsDataFrame, percentage of observations to randomly sample	SpatialPointsDataFrame	Randomly samples a user-defined percentage of the observations from the input data set

filterField()	LargeSpatialPointsDataFrame, trait of interest within the LargeSpatialPointsDataFrame, value of the trait that will become the threshold, and a logical statement to denote whether the threshold value represents the highest or lowest value to remove from the LargeSpatialPointsDataFrame	LargeSpatialPointsDataFrame	Removes observations from the input LargeSpatialPointsDataFrame based on the threshold values set via the user-defined input values
sdField()	LargeSpatialPointsDataFrame, trait of interest within the LargeSpatialPointsDataFrame, maximum number of standard deviations the value of the trait can vary from the mean value of the trait before that observation is removed from the LargeSpatialPointsDataFrame	LargeSpatialPointsDataFrame	Removes observations from the input LargeSpatialPointsDataFrame based on the standard deviation threshold set via the user-defined input values

In the example below only the central part of the field was selected to be evaluated and filtered (Figure 6.1). After cropping the field to select data from only a portion of the field, users can continue to process data using further steps in the cleanRfield pipeline (see full pipeline in Figure 6.4). If users are interested in selecting 2 or more polygons within cropField(), the parameter nPolygon should be increased to reflect the number of desired polygon. When running cropField() for multiple polygons, the help text within the R Console will inform viewers when they are done selecting points for a given polygon and should begin selecting points for the subsequent polygon. Utilizing cropField() for multiple polygons will save all polygons to the same output SpatialPolygons object. Users who wish to save each polygon as separate SpatialPolygons objects would be better served by using boundaryField() instead of cropField(). Summarized use of boundaryField() is shown in Figure 6.2, and example code for drawing multiple polygons from the same field and saving them as separatem SpatialPolygons objects is provided in full on the Github tutorial.

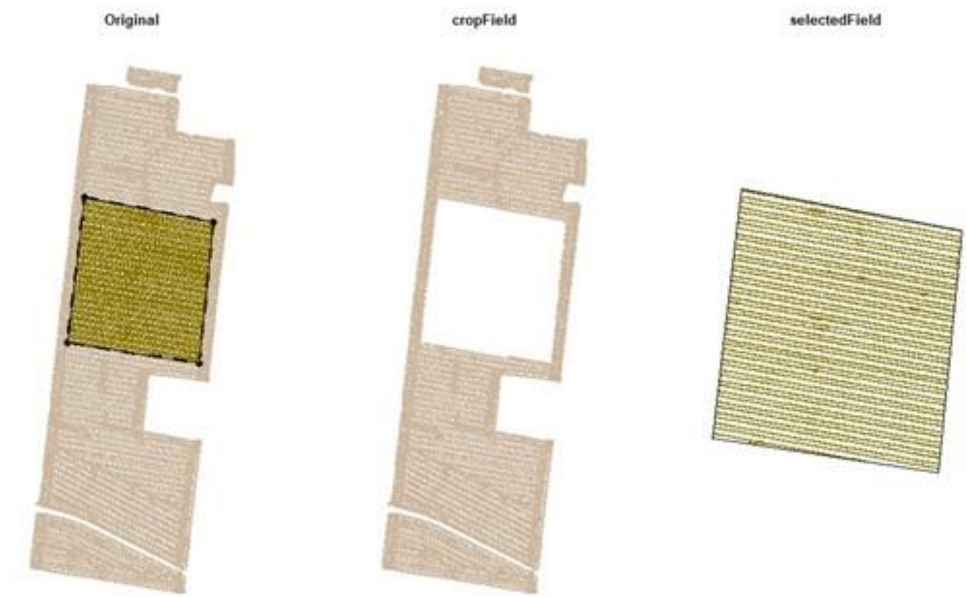

Figure 6.1. Using function cropField() from cleanRfield R package to select and extract data from a rectangular area of interest in the center of the field. Selected data can be further filtered and analyzed within the cleanRfield pipeline.

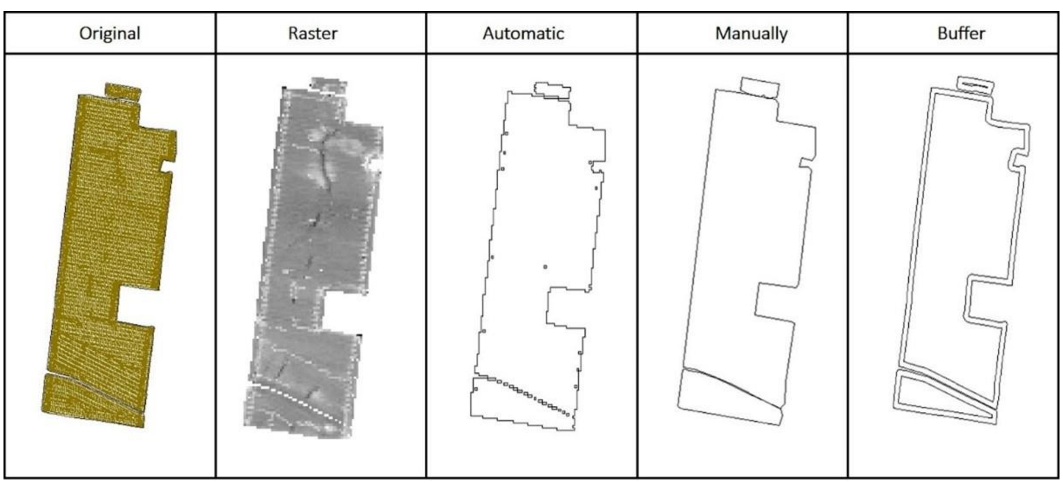

Figure 6.2. The cleanRfield step by step pipeline to build the boundary shapefile and reduce the boundary with a buffer. The function rasterField() was used to transform the original data to a raster. The function boundaryField() was used to draw the boundary shapefile automatically or manually. And the function bufferField() to reduce the manually drawn boundary using a buffer of fixed width.

Building Boundaries and Buffering

Managing field boundary files can be very time consuming for researchers or crop consultants that cover large geographic ranges. Therefore, we developed a function called boundaryField() that allows users to build a boundary polygon by (1) drawing the boundary manually or (2) automatically generating the boundary from the yield data. In the first method, the user utilizes the cursor to click within the plot area of RStudio and draw the border around the field data visualization. The process for drawing boundaries on multiple fields and then merging boundaries together for subsequent cleaning steps within the cleanRfield pipeline is described in full on the Github tutorial (https://github.com/filipematias23/cleanRfield), and boundaries can be exported as shapefiles or other vector data files such as geopackages for use in other software programs.

The automatic boundary method transforms the dataset from data points to a raster using the function rasterField(). This function allows the transformation of data point layers on raster layers from where the boundaries can be identified automatically by boundaryField(). The resolution chosen during rasterization impacts the appearance of the final map, especially near the edges of the field, and can impact the accuracy of automatic boundary delineation. Optimum resolution varies based on both the coordinate reference system of the input data set and the header width of the combine used while collecting the yield monitor observations, and more information on resolution suggestions is available in the Github tutorial (https://github.com/filipematias23/cleanRfield).

Once the boundaries are created the user can apply a buffer to eliminate the observations within the point data set that are located near the edge of the field, since yield monitor observations near the edges of fields frequently contain errors due to the motion of the combine and/or monitor errors (e.g., combine turns, stops and starts, changing speed, etc.). The function bufferField() allows users to buffer the entire field boundary according to the buffer value. Buffer values must be negative, and the function bufferField() works with boundary polygons delineated using any linear unit (units: in, m, us_feet, etc.). The steps described above are summarized in Figure 6.2, (1) starting by transforming the original dataset into a raster, (2) followed by building the boundary automatically or manually, and (3) ending by applying the buffer reduction. Note the differences between automatic and manual boundary delineation: jagged field edges and occasional holes or patches within the field boundary are typical of automatic field boundary delineation.

Filtering Data

The package cleanRfield has different methods to clean, filter, and evaluate datasets. (1) The first and simplest method is the reduction of data by randomly sampling points using the function sampleField(). This function allows the user to select any percentage of points either sampling the entire spatial extent of the input file or using polygons from cropField() to sample only a subset of the field. The sampleField() function may be useful for exploratory analysis of very large data sets, or for users looking to build smaller datasets for learning spatial data processing in R. Additionally, sampling large data sets can help users determine the optimum number yield monitor observations during power analysis and trial planning. (2) The second method is filtering by using collected data values with the function filterField(). This function allows the user to choose one or more layers in the dataset (e.g., dry yield, speed, etc.) in order of priority and filtering for minimum or maximum permitted values for each attribute. (3) The third method uses the standard deviation (sd) of each data layer to select points that are inside the thresholding value. The function sdField() determines the minimum and the maximum values to select only the points inside the sd interval. All filtering methods can be combined in a pipeline where the output dataset from one filtering step will be the input of the next filtering step. For example, on Figure 6.3 the first filter method used was the buffering to eliminate boundary effects followed by filtering data points based on collected data values. For instance, the buffered shapefile with 5 meters reduced from the border was used in the function filterField() associated with a second step of filtering where only data points with “Dry_Yield” greater than 70 bu/acre and speed lower than 5 mph were retained.

Figure 6.3. Different filtering steps used on cleanRfield pipeline. Initially the buffered boundary from bufferField() function was used to remove the borders effects followed by filtering data points to retain observations where Dry_Yield>70 and Speed<5 using the function filterField. The second and third panels display the points retained by each filter in brown—white space represents areas where points were removed via filtering. These

filtering parameters were used only for example purposes and are not necessarily the appropriate thresholds for most cleaning procedures.

Parallel Processing for Multiple Fields

Processing yield data in parallel is an important step when there are different field trials or datasets to be evaluated. This is a common practice in agriculture where many crop trials are harvested in a short window of time and the data must be evaluated as soon as possible for decision making. In the online tutorial available on GitHub we developed a sequence of code as a suggestion about how to use cleanRfield functions in parallel to evaluate three different yield data trials at the same time. For instance, for this specific case the same criteria of filtering were applied in all fields. However, it is possible including conditions (e.g., if and else) for applying different filtering steps for each trial according to the user's needs.

Discussion

This manuscript describes functions available in cleanRfield R package and potential applications for using the package to analyze agricultural point data, such as yield maps collected with combines. However, the pipeline can be easily adapted for filtering different sources of point data (e.g., demographic, political health, economic, etc.). Following steps in Figure 6.4, the cleanRfield pipeline starts by collecting the data (e.g., combine) and preparing it to upload in R (e.g., .csv, .txt, .shp, etc.). From this point, users have different options as (1) making rasters, (2) using rasters to define field boundaries, (3) manually drawing polygon boundaries, or (4) importing preexisting boundary shapefiles to determine the geographic extent for the software to evaluate and filter data points. The next step is using different strategies available on cleanRfield for cleaning data such as (1) buffering the boundary, (2) filtering by setting thresholds on data values, or (3) determining threshold limits with standard deviation. As far as we are aware, this is the first package in R with a detailed pipeline to process multiple yield maps at the same time, improving the standardization and reproducibility of yield monitor data cleaning procedures. This package helped to increase the efficiency of our team’s daily activities due to its relative speed over point-and-click software alternatives and easy adaptability to different applications.

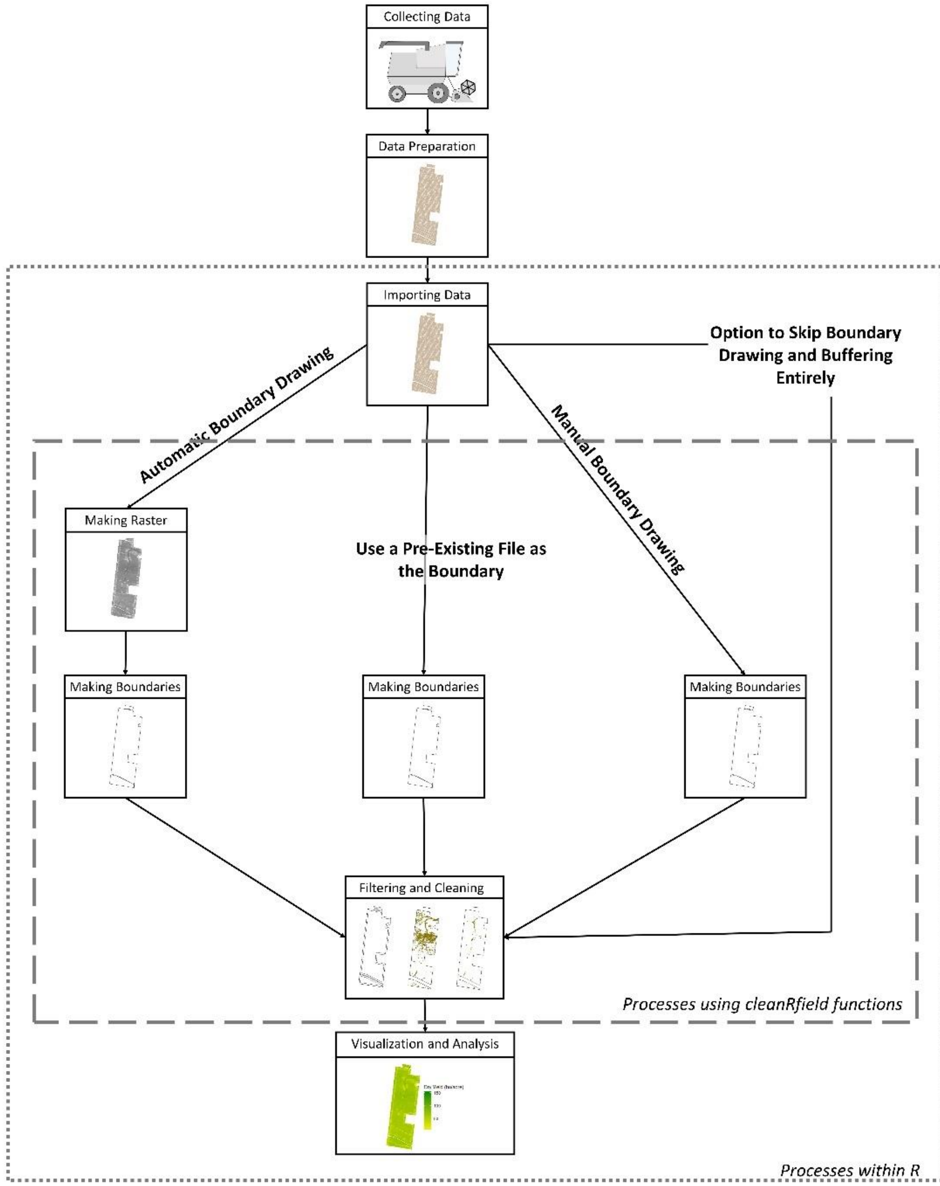

Figure 6.4. Workflow illustrating the main steps of cleanRfield pipeline, starting with data collection using combines with yield monitors, then preparing the yield monitor data for import into R as a .shp, .txt, or .csv file. The remaining steps, including data importation, making field boundaries, filtering observations, and data visualization can all occur efficiently within R.

Limitations

While cleanRfield will continue to improve functionality by addressing user feedback, at the time of publication it has some noteworthy limitations. Firstly, cleanRfield cannot read the proprietary file formats that combines and yield monitors use to store raw data unless those yield maps are converted to a more general file format such as shapefiles or delimited text before being used as input data for cleanRfield functions. Additionally, there are not functions within cleanRfield that can be used to adjust sensor delay values. While these limitations do constrain the useage of cleanRfield, this package still represents an improvement over other currently available yield cleaning software options.

Conclusions

The cleanRfield package offers crop advisors and researchers a convenient set of tools to evaluate and filter data points in a friendly and open-source way using R. We developed a detailed and straightforward online tutorial to illustrate each step of this software with more detail at https://github.com/filipematias23/cleanRfield. For instance, users can download data examples and follow the pipeline to understand and adapt the code for different applications. cleanRfield will help the agricultural community quickly and accurately clean yield monitor data which will improve time to decision making, and the package tutorial also provides users with open feedback channels to further improve the features and the quality of this tool.

Acknowledgements

The authors thank the Wisconsin Soybean Marketing Board for providing partial funding for this research and the sample data used to develop cleanRfield.

Conflict of Interest

The authors declare no conflict of interest.

Works Cited

Bergman, R. (2020). Tips for calibrating your combine's yield monitor. *Iowa State University Extension*. Retrived from https://crops.extension.iastate.edu/blog/levi-powell-matt-darr-ryan-w-bergman/tips-calibrating-your-combines-yield-monitor

Casady, W. & Shannon, K. (1998) Precision agriculture: yield monitors. WQ 451. *University Extension, University of Missouri-System*. Retrieved from https://extension.missouri.edu/publications/wq451

Griffin, T.W., Brown, J.P., & Lowenberg-DeBoer, J. (2007). Yield Monitor Data Analysis Protocol: A Primer in the Management and Analysis of Precision Agriculture Data. Site Specific Management Center Publication. Purdue University, West Lafayette, Indiana.

Kharel, T. P., Swing, S. N., Maresma, A., Yongerman, C., Kharel, D., Czymmek, K. J., & Ketterings, Q. M. (2019). Yield monitor data cleaning is essential for accurate corn grain and silate yield determination. *Agronomy Journal*. 111: 509-516.

Kleinjan, J., Change, J., Wilson, J., Hamburg, D., Carlson, G., Clay, D., & Long, D. (2002). Cleaning Yield Data. Retrieved from http://static1.1.sqspcdn.com/static/f/891472/12242856/1305568816010/Kleinjan+et+al.+2002+Cleaning+yield+data.pdf?token=uq%2BesCvXYB1n289M6z%2BtXON0CLY%3D

Schimmelpfennig, D. (2016). Farm profits and the adoption of precision agriculture. *USDA Economic Research Report Number 217*.

Simbahan, G., Dobermann, A, & Ping, J. (2004). Screening Yield monitor data improves grain yield maps. *Agronomy Journal.* 96: 1091-1102.

Sudduth, K.A., Drummond, S.T. & Myers, D. B. 2012. Yield editor 2.0: Software for automated removal of yield map errors. Paper No. 121338243. ASABE Ann. Intl. Mtg., Dallas, TX.

USDA Ag Data Commons. (2021). Yield Editor 2.0.7. Retrieved from https://data.nal.usda.gov/dataset/yield-editor-207