**Variable Rate Seeding: Impacts on Yield, Weeds, and White Mold** - **2019 Annual Report**

Leanna Leverich and Seth Naeve

## Introduction

Soybean seeding rate and its relationship with soybean yield has been a primary interest and research topic for growers and researchers to optimize yield and economic returns. As the industry pushes to maximize yield, inputs have increased with growers applying more products, making more passes in the field, and planting more seeds (Castrignanò et al., 2017). The introduction of precision agriculture provides a pathway for growers to maintain yield maximization, while lowering inputs and lowering environmental impact. The implementation of precision agriculture practices allows for growers to save money on input and application costs and promotes a more sustainable approach to farming by preventing wasteful and expensive overapplications (Khosla et al., 2008).

With the introduction of precision agriculture technology, many questions have arisen around the topic of variable rate seeding. Seeding rate and other management practices have long been focused on a per field basis. However, with the release of variable rate technology (VRT), there is now an opportunity for growers to manage soybean on a per acre basis. This revolutionizes the way farmers can manage inputs and can result in profit gains for growers in cost savings through reduction of inputs such as seed and fertilizer. The release of variable rate planters in recent years has offered growers new opportunities to further evaluate seeding rates to optimize returns on the farm. This includes identifying both agronomically and economically optimal input rates, including optimal seeding density (Jaynes, 2010). Yet much of past seeding density research has been conducted in small-plot studies in portions of larger production fields leading to information about variation in optimal seeding rates across states or regions. Missing is information about variation of the seeding rate response within individual fields.

Recent research by Gaspar et al. (2020) pointed to the opportunity for profit gains by growers through optimization of seeding rates. Researchers found that in areas of lower productivity, higher seeding rates were agronomically optimal, whereas in higher productivity areas, lower seeding rates were optimal. This has been further validated by other studies in Brazil and in the U.S. Midwest (Corassa et al., 2018; Carciochi et al., 2019), and it highlights the possibility for farmers in Minnesota to save on seed costs by optimizing seeding rates across their fields.

Past research and current recommendations from University of Minnesota Extension indicate 100K plants per acre is fully sufficient for obtaining yield goals. Yet oftentimes growers plant well beyond the recommended 125K minimum seeding rate, with averages ranging between 120-190K to ensure 100 percent of the yield goal is reached. This increase in seeding rate may slightly increase yield, but it follows a diminishing returns model, with each incremental increase leading to smaller yield gains on a whole-field basis. However, in low productivity areas of the field, there may be an increase in yield making the increased seeding density worthwhile. The temptation to maximize yield on every acre may be the key factor leading farmers to increase their seeding rates in soybean across the entirety of the field.

While some growers are increasing seeding densities to maximize yield on every acre, others are questioning seeding rate increases due to its effect on disease pressure. Seeding density has implications for disease rates, with higher seeding densities leading to increased disease severity (Jaccoud-Filho et al., 2016). White mold (*Sclerotinia sclerotiorum*) infestations throughout Minnesota have heightened grower interest in soybean seeding rates. The need to maximize yields yet control disease has further contributed to interest and questions surrounding variable rate seeding.

Turning to an acre-by-acre approach to management, there is potential for substantial profit gains by increasing the seeding density on acres classified as lower productivity and maintaining lower seeding densities throughout the remainder of the field considered higher or medium productivity. To benefit from the use of variable rate seeding, growers must first know if there is a benefit to varying seeding density and then where their fields need optimizing. Areas of the field must first be categorized as marginal, adequate, or optimal yielding. To determine the classification of the field, soil, climate, and landscape factors must be identified for their influence on soybean yield and soybean seeding success. Determining which factors influence soybean yield by seeding density can allow for a methodology to be developed for recommending soybean seeding densities across farmer fields (Smidt et al., 2016).

The objectives of this study were to 1) evaluate spatial variability of soybean yield response to seeding rates at the field scale, 2) evaluate the influence of soybean planting density on pest pressure at the field scale, and 3) evaluate key field properties influencing spatial variability within farmer fields.

## Methods

### Site Description and Experimental Design

This experiment was conducted on two farmer fields during the 2018 growing season and three fields for the 2019 growing season. Sites for 2019 were continued on the two farms from 2018 in central and southwest Minnesota. The additional site for 2019 was implemented in western Minnesota. Sites were chosen to represent the dominant agricultural regions of the state while capturing diversity in climate and soil properties (Table 1). Western Minnesota is known for its alkaline soils and drier climate, whereas central Minnesota growing regions have primarily acidic soils and greater amounts of precipitation. The contrast in features across the state was desirable to evaluate potential influences on yield at various seeding densities.

Soybean seeding density strip trials were conducted at each site to assess within-field variability and the effects of soybean seeding rate on soybean yield. Farmers had variable rate technology (VRT) capabilities and field size ranged between approximately 20 to 50 acres. Five seeding densities were used in each field: 75,000; 100,000; 125,000; 150,000; and 175,000 seeds per acre. A novel experimental design was implemented to assess seeding density at a fine geographical resolution. A traditional randomized strip trial with four to five replications (depending on field size) was implemented in each farmer field with perpendicular subdivisions (Figure 1). The subdivisions were created to capture the spatial variability across the field, making approximately 1-acre blocks containing each of the five seeding densities. This allowed for assessment of the effects of seeding density and other field characteristics on soybean yield across each individual acre of the field. As-planted data and GPS plot boundaries were used to guide the creation of the 1-acre subdivisions using QGIS geospatial software.

### Sampling

After planting at each site, soil samples and stand counts were collected on the 1-acre grid created by the subdivisions. Stand counts were collected for each of the five seeding densities within each acre grid. Approximately eight 1-inch soil cores (0-6” depth) were collected from the center of each block. Soil samples were collected at centers of the block to allow for an interpolation to provide predicted values between sampling points and to mimic a typical grower practice for a wider implementation of these methods. Soil samples were analyzed using a routine soil test for phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), pH, cation exchange capacity (CEC), and organic matter (OM). At approximately the R3 growth stage, each site was visually evaluated for white mold and weeds for each seeding density in each 1-acre block. At each site a weather station was installed to record rainfall and temperature.

### Geospatial Data

Several geospatial data layers were collected from the farmer, state, or national information sources to further identify soil and landscape properties that could affect soybean yield by soybean density. Lidar data was gathered from MNTopo application provided by the Minnesota Department of Natural Resources as a 1-meter digital elevation model (DEM). The DEM raster layer was then used to create several geospatial DEM derivative layers to be compared to soybean yield. This included elevation, slope, topographic wetness index, upslope, and aspect. These layers were used to assess how topography across each field influenced yield, and to identify if this varied by seeding density. Topographic layers were made using the maptools and raster packages in R statistical software.

Data from the Soil Survey Geographic (SSURGO) Database was also used to identify soil characteristics that may influence soybean yield for each seeding density. Data from SSURGO was used to look within sites, but given its resolution, it was primarily used across the greater study to evaluate influence of soil properties across all sites pooled. Soil properties evaluated for their influence on yield were soil type, water table depth, available water storage, drainage class, hydraulic conductivity, total sand, total clay, and available water capacity. The eight properties were chosen with the intention of reflecting water dynamics and dominant soil textures within fields to assess their effects on yield, and to capture if these effects varied by seeding density.

Historical yield maps were also collected from growers detailing approximately 10 years of historical yield for each field. The maps were used to assess fields for stable and unstable areas over time and space and to evaluate the potential to predict high, medium, and low productivity areas. The stability of the productivity could then be used to evaluate the predictive ability of determining where high versus lower seeding densities might be applied.

### Data Preparation and Analysis

Yield data was collected with calibrated yield monitors at the three sites and cleaned to reduce noise due to continuous data collection of the combine. This includes noise created by the combine stopping in the middle of a pass (farmer-induced error), harvest points at the end of passes where the combine may inaccurately record low yield when meeting headlands, and edge rows that contain inaccurate yield data. Each yield file was cleaned using approximately three standard deviations or the 99th and 1st percentile to remove erroneous points (Vega et al., 2019). Yield cleaning was conducted using the geopandas, numpy, and matplotlib packages in the Python Programming Language.

*Identifying Optimal Seeding Density Within Fields*

An initial assessment was conducted using ANOVA and Tukey’s HSD mean separation test to determine the optimal seeding density on a whole-field basis. To further assess spatial variability within the field, a more in-depth analysis was used to determine the agronomically optimal seeding density (OSD) for each 1-acre grid block at each site based on yield. This was achieved by conducting an ANOVA for each 1-acre block to determine if soybean seeding density significantly affected soybean yield. Tukey’s HSD test was then used on each 1-acre block to determine 1) which seeding rate(s) was statistically optimal (the highest yielding seeding densities), and 2) the number of seeding rates that were statistically different from one another. This revealed which areas of the field responded to seeding density, and which did not. Spatial analyses to determine spatial variability were conducted in R Statistical Software to determine OSDs across fields.

For 2019 yield data at site two, harvest passes were not correctly aligned with planter passes causing greater variability in the yield data. To better evaluate the yield data, fully compromised yield passes were removed. Psuedo-seeding rates were developed according to harvest passes to represent the true seeding density for each pass, rather than the five original densities. A second approach utilized interpolation of the yield data to assess yield for each seeding densities strip. Both the pseudo-seeding rate approach and the interpolation resulted in similar results for overall yield averages. Reporting on the original seeding densities for site 2 in 2019 utilized the interpolated data.

### Identifying Key Factors influencing Yield

All study rasters, designs, and polygons were prepared using QGIS Geospatial Software. To determine which field properties may be influencing yield, all layers were spatially joined to each seeding density in the 1-acre grid and summarized using the zonal.stats function from the spatialEco package in R. The layers joined to yield data included topography layers, SSURGO layers, soil sample data, white mold and weed counts, and seeding density (Table 2). For a 50-acre field, with 50 1-acre blocks, this results in a data frame with 250 summarized observations, five observations for each seeding density in a block.

Field properties were then identified for their importance using the Random Forest machine learning algorithm. The algorithm was not used for yield prediction, but rather for feature selection to determine the most important field factors influencing yield. Key Factors and the Random Forest machine learning algorithm also were conducted in R.

### Economic Assessment

A brief economic evaluation was conducted using yield data from the two sites. Comparisons were made using the farmers’ typical seeding rates, current cost of seed, and the market price. Two approaches were taken to evaluate 1) the economic impacts of the study on a whole-field basis, and 2) the economic outcome had the optimal seeding rate been planted in retrospect. To evaluate the economic effects of the various seeding densities from the study, the costs and profits of each seeding rate were compared to the farmer rate. In this method, the economic outcome was based on field averages. To derive the average yield for the grower rate, which was not included in the trial, models were fit to the trial data to estimate grower yield. Two models were fit, one including the first and last seeding density yields, and the other including all seeding densities. The more advantageous yield outcome from the two models was used for the grower seeding density yield. This yield estimate was used to compare against the other trial seeding densities.

The second approach provided an estimate of cost savings should a grower carry out variable rate seeding. For each 1-acre block in each field, average yield was calculated across the five seeding densities. This yield was then paired with the lowest seeding density that was considered statistically and economically optimal (see section on calculating OSD). In other words, the yield of the low seeding density was statistically the same as the other seeding density yields. For example, in non-responsive areas of the field (where all seeding density yields were statistically the same), the lowest density would be selected. This is referred to as the economically optimal seeding density (EOSD). In addition, the seeding density with the highest value (not statistically significant) was also selected and economically evaluated. This was labeled the highest yielding optimal seeding density (HOSD). Seed costs were estimated by the appropriate seeding density for each 1-acre at a cost of $53 per 140K seeds. The soybean market price that was used to calculate profit was $9.00. This was then summarized for each field and compared to the farmer seeding density.

## Results

*Results from 2019 will be reported for two of the three locations. Due to unforeseeable events, the collection of the yield data from the third site has been delayed until University travel is again permitted more widely. Once the raw data has been retrieved, analyzed, and interpreted, an amended final report to the MSR&PC will be submitted.*

### Optimal Seeding Densities and Spatial Variability

Each site was evaluated for the overall best performing seeding rate in both years. On average, across the whole field, evaluation of each block at each site revealed which seeding densities were considered optimal. The whole-field averages for site 1 revealed that higher seeding densities performed better overall. At site 1 in 2019, the highest two seeding densities were statistically different from the two lowest seeding density, and in 2018 the two highest seeding densities were statistically higher than the three lower densities (Table 3). For both years at site 1, the lowest seeding density was statistically the lowest yielding seeding density on a whole-field average. At site 2, yield was greater in 2018 on average (64.9 bu/ac) than in 2019 (47.5 bu/ac) likely due to pressure from white mold, late planting, and other agronomic factors. The lowest seeding rates did best at site 2. Planting 75K seeds per acre resulted in the greatest yields in 2018 and planting 100K seeds per acre resulted in the highest yields on average in 2019. The yield response to seeding rate was not linear at site 2 either year (Figure 2). Analysis of the pseudo-seeding densities followed a similar trend in 2019, with the lower seeding densities achieving higher yields. The spread in the data at higher seeding densities further indicated the greater variability in yield response at the higher densities, whereas the lower densities had a more uniform response that was generally greater on average (Figure 4).

To further evaluate spatial variability across fields, average yield was determined for each 1-acre grid block in the field (Figure 3). A range of spatial variability in yield was evident at site 2.

Seeding density success at each site was determined using the 1-acre subdivisions across the entirety of the field. An ANOVA and Tukey’s HSD test conducted within each block revealed which seeding densities were most commonly the statistically highest yielding. At site 1, the highest seeding densities (150K and 175K seeds per acre) were the most common optimal seeding densities (OSD) and lowest seeding densities were infrequently considered an optimal seeding density across the field in 2019. However, areas where the 75K seeds per acre was an optimal seeding density indicated where there was no yield response to seeding density and likely an opportunity for growers to lower their seeding rate. At site 2, 75K seeds per acre was the most frequent OSD in 2018 and 125K seeds per acre was the least frequent OSD (data not shown).

Analysis of data from year one showed a clear within-field variability at both locations. With the fields summarized by 1-acre grid, mean yields varied between 42 to 67 bu/ac at site 1, and between 59 to 76 bu/ac at site 2 (Figure 2). This indicated both fields contained spatial variability that could be captured and quantified. At both sites, often more than one of the five seeding rates was optimal at each 1-ac grid cell through the field. Considering there is often more than one optimal seeding density (OSD) in each 1-acre grid, this led to more total OSDs than grid cells at each site (Figure 5). In some areas of the field there was no response to seeding rate. Areas with no response to seeding rate may be candidate areas for lowering the seeding rate for growers to capitalize on seed savings. At site 1, the most common OSD was the highest seeding rate: 175K seeds/ac. However, in contrast, the most common OSD at site 2 was 75K seeds/ac, the lowest seeding density.

### Economic Impacts of Variable Seeding

A broad economic assessment was conducted for both farmers for both years to evaluate the general benefit or reduction to profits in utilizing various seeding densities in the study. Similar to methods used in 2018, yield and seed costs were compared to predicted farmer yield for the farmer's typical seeding rate. A second analysis evaluated the benefits of variable rate seeding with economically optimal seeding densities for each acre.

In both years at site 1, the use of lower seeding densities resulted in yield loss across the study area. The largest loss at site 1 in 2018 was approximately 4 bushels per acre at the 75K seeding density, compared to the predicted farmer yield, at a seeding density of 160K seeds per acre. The 175K seeding density led to a gain of approximately 0.5 bushels per acre. In 2019, using the five seeding densities resulted in a loss of approximately 29 bushels total from the study area, and a loss of approximately $6.42 per acre in 2018 and $12.11 per ac in 2019 from the 75K seeding density. This low yielding seeding density (75K) was the lowest performing density in the trial at site 1 and resulted in the largest yield losses. In considering seed costs and the savings in planting the lowest density, 75K, losses due to yield become negligible, and this results in a net positive both years of the study. At site 2 in both years, there were substantial gains in comparison to the farmer rate of 135K seeds per acre. Both in 2018 and 2019, the highest yielding density on average was lower than the farmer rate (75K in 2018 and 100K in 2019). This resulted in profit gains for the grower in both yield increases and seed savings.

A deeper evaluation after the study allowed for an economical optimal seeding density (EOSD) to be selected for each 1-acre block for each site each year. The EOSD was selected as the lowest seeding density that had the highest yield or yield that was considered statistically optimal according to Tukey’s HSD mean separation test. Using a specific EOSD in every acre of the field led to an average gain or cost savings of $18.55 per acre at site 1 in 2018 in comparison to the farmer's flat rate (Table 5). In 2019, site 1 would’ve had an average gain of $21.77 per ac in using EOSD for each acre compared to the farmer’s flat rate. At site 2 in 2018, use of EOSD for each acre would have resulted in an average cost savings of $17.04 per acre.

The highest yielding optimal seeding density (HOSD) was calculated using the OSD with the highest yield value. Though it was the highest yielding in value, the HOSD difference in yield was often negligible in comparison to the EOSD. Oftentimes the HOSD was a higher seeding density than the EOSD, and the use of the HOSD would ultimately result in a loss due to the cost of seed. Use of the HOSD is seen on Minnesota farms today. The University of Minnesota Extension recommendation is >125K seeds per acre, but commonly 120-190K seeds are planted. This can be attributed to the aim of increasing yields, even marginally, or to compensate for low productivity areas of the field. This highlights the importance of finding not only an agronomic optimum seeding density, but also an economic optimum seeding density for greatest on-farm gains. The benefit of savings on seed costs may lead growers to closely evaluate implementing variable rate seeding across their fields. This economic assessment of costs illustrates the potential for greater savings for growers should they chose to implement various seeding densities. The highest yielding seeding density is not always the best economic choice, when there are not statistical differences in yield between densities. Realizing the economic benefit of variable rate seeding and the variability of OSD throughout the field, there is a clear need to identify factors influencing this OSD and yield. Identifying factors of the field that have the most impact on yield and seeding density response can lead to the development of prescriptions for variable rate planting in Minnesota farm fields.

### Key Factors Influencing Yield

The clear difference between the two sites in yield performance within seeding density indicated there were likely soil and landscape factors contributing to the difference. A Random Forest (RF) machine learning algorithm was employed to identify important key factors contributing to variability in the data. The factors inputted to the model to compare with yield included routine soil test numbers, lidar derivatives, SSURGO data, and data from the experiment (i.e. yield) (Table 2). The RF model was run for each site in 2018 and with the sites pooled. The output of interest was a variable importance plot and the variable ratings. This method was used to conduct feature selection to determine the most important variables, rather than for the development of a model which is a future step. Feature selection will undergo cross-validation in future efforts for the development of a model to predict yield based on field factors. This initial testing was to provide an overall view into the most influential factors. Overall, the RF (untrained and untuned at this point) was able to characterize approximately 57% of the variability at site 1, 42% of the variability at site 2, and 62% of the variability for the pooled sites.

Through examination of the measures of variable importance, there were clear differences in factors contributing to yield variability at both sites (Figure 6). Mean Square Error (MSE) increase and Node Purity Inclusion both measure variable importance, with higher values indicating a more important variable. However, soil test numbers were highly important at each individual site (Table 4). Potassium was especially important at site 2 and phosphorus was important at site 1. Both , P and K, were also considered important variables (within the top five) across the sites. Similarly, elevation (DEM) was of specific importance at both sites and for the pooled sites.

Calcium was also of particular importance at site 1 (Figure 6) indicating the influence of the nutrient in soybean systems. In the western side of the state where site 1 is located, calcium carbonate commonly plays a role in nutrient availability. The higher pH levels and amounts of calcium carbonate could potential be contributing to differences in yield at site 1, which may be reflected in the RF model. Even without visual iron deficiency chlorosis (IDC) stress, Ca levels may be an important indicator for soybean population response.

Unsurprisingly, white mold was considered one of top five factors at site 2. This was evident where we saw disease pressure in the higher seeding densities at site 2 (Figure 7) in both years. The white mold was even more evident in 2019, and most likely will be an important factor yet again for 2019 in the RF model. This is further evidence of the influence of disease pressure at higher seeding densities seen in other studies.

The pooled sites showed two SSURGO factors were important in explaining yield variability, soil type and clay total. The use of SSURGO data for within-field research is oftentimes too coarse in resolution to provide a benefit in capturing variability. However, in using the RF model to assess factors contributing to variability, SSURGO stands out in its ability to segregate the sites from one another to explain initial variability between the two locations. In future development of properly tuned and trained models, SSURGO will be an important input to allow the model to first segregate sites and then to continue to parse variability in yield with other factors. The clay value collected from the SSURGO database was especially interesting due to its linear relationship with yield for the pooled sites (Figure 8). This was one of the few linear relationships and indicates that underlying SSURGO data can be beneficial in determining a baseline as to how successful soybean yields may be. This indicated the RF model used factors like clay content most often to first differentiate the sites and then the other factors to find differences within sites.

Interestingly, seeding density was not one of the top factors for site 1, site 2, or the combined sites. However, it was one of the top 10 variables influencing yield at site 2 in 2018 (Figure 6, listed as Treatment.x). This was likely due to its relationship to white mold and the greater differences between yield averages at site 2 in 2018 (Table 3). Its absence from the other two model’s importance variables indicated that yield variability is less dependent on seeding rate in comparison to other field properties. While this may indicate seeding density is not an important management area to consider for maximize yield, it actually suggested potential cost savings gains that may be incurred. The RF model results showed seeding rate is relatively insignificant at site 1 and at the pooled sites in determining yield variability, which indicated the potential for identifying areas of the field to lower seeding densities. With unresponsive areas identified, there is great potential for growers to reduce costs in seed.

While these measures are telling, it is important to realize that this model is not yet trained and not tuned, and therefore, these rankings can slide between positions. These rankings were used as an initial baseline in exploration of the data. Future tuning of the model and feature selection will reveal a more complete picture of which variables consistently influence yield response.

### Next Steps

Similar analyses to 2018 will be conducted on 2019 data using the random forest statistical model to further evaluate potential key factors most influencing yield and seeding density. To advance, two RF models may be employed to first differentiate fields, and secondly to characterize variability within each field. Furthermore, the RF models may be improved by removing variables that do not contribute to explaining yield. Eliminating some of these variables may reduce extra noise and error produced by the model. Additional factors will be added to the model as well, including weather data, spatial and temporal historical yield stability maps, and canopy closure estimations derived from drone imagery. Drone imagery will also be used to determine ability to detect white mold and to develop yield predictions. With initial assessments conducted with a RF model, a full cross-validation technique will be applied for feature selection (variables of importance) and the building of predictive model for yield. In future years, the predictive model will be used to predict yield based on field variables. Future studies can be implemented using the important variables to guide seeding rate densities across fields and can be validated with yield success.

**Conclusions**

Based on data from this study there is a unmistakable opportunity to deploy variable rate seeding on Minnesota soybean fields to increase yields and grower profits, while lowering inputs. This precision agriculture practice has great opportunity for implementation to help soybean growers improve yields and manage pests and disease like white mold. There is clear within-field variability shown in all four site-years that has potential to be managed by optimizing seeding rates throughout the field. There is also a clear economic incentive for implementing variable rate seedings. In utilizing EOSD on a per-acre basis, there were clear economic gains for growers both years. Determining the underlying factors influencing yield and OSD will lead to generation of maps for optimizing seeding densities in farmer fields. In initial exploration of key factors, it was evident that soybean yield was not strongly related to seeding density when evaluated using the random forest algorithm. Other field factors provided clues as to what factors are controlling yield variability, such as soil nutrient content and topography. The identification and mapping of these feature will lead to an understanding of which areas of the field are the most and the least productive. Further analysis and research to tune and refine the random forest model can lead to the development of recommendations for devising VRT seeding prescriptions.

## Tables

Table 1. Site location, predominant and secondary soil series information and classification for three sites in Minnesota

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Predominant Soil | |  | Secondary Soil | |
| Site | **County** | Series | Classification |  | Series | Classification |
|  |  |  |  |  |  |  |
| 1 | Lyon | Barnes | Fine-loamy, mixed, superactive, frigid Calcic Hapludolls |  | Hokans | Fine-loamy, mixed, superactive, frigid Calcic Hapludolls |
| 2 | Wright | Cordova | Fine-loamy, mixed, superactive, mesic Typic Argiaquolls |  | Lester | Fine-loamy, mixed, superactive, mesic Mollic Hapludalfs |
| 3 | Lac qui Parle | Parle | Fine-loamy, mixed, superactive, calcareous, frigid Cumulic Endoaquolls |  | Glynden | Coarse-silty, mixed, superactive, frigid Aeric Calciaquolls |

Table 2. Factors input into the random forest model for identification of key factors influencing spatial variability

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model Element | Model Factors | | | | |
|  |  |  | |  | |
| *Routine Soil Test* | pH | | Potassium | | Phosphorus |
|  | Calcium | | Organic Matter (OM) | | Magnesium |
|  | Texture | |  | |  |
|  |  | |  | |  |
| *Lidar Derivatives* | Elevation (DEM) | | Upslope | | Topographic wetness index (TWI) |
|  | Average Slope | | Max Slope | |  |
|  |  | |  | |  |
| *SSURGO* | Soil Type | | Water Table Depth | | Total Clay |
|  | Drainage Class | | Total Sand | | Available Water Storage (AWS) |
|  | Hydraulic Conductivity | | Available water capacity | |  |
|  |  | |  | |  |
| *Experimental* | Yield | | Weed counts | | White Mold Counts |

Table 3. Average yield for five seeding densities for two years at two sites in Minnesota. Letters denote statistical differences.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Site 1** | |  | **Site 2** | | |
|  | **2018** | **2019** |  | **2018** | **2019** |
| **Seeding Density** | *Yield (bu/ac)* | |  | *Yield (bu/ac)* | | |
| 75 | 51.8c | 50.9c |  | 64.9a | 47.5b |
| 100 | 54.7b | 55.5b |  | 61.5c | 48.9a |
| 125 | 53.9b | 54.6ab |  | 60.7c | 46.7bc |
| 150 | 56.0a | 55.7a |  | 61.5c | 46.3c |
| 175 | 57.1a | 55.7a |  | 62.7b | 44.8d |

Table 4. Variables of importance ranked from the use of the random forest machine learning algorithm. The top five variables were determined by MSE increase and node purity increase measurements obtained from the model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *Variables of Importance* | | | | |
|  | **Site 1** |  | **Site 2** |  | **Pooled Sites** |
|  | Phosphorus |  | Potassium |  | Soil Type |
|  | Calcium |  | Elevation (DEM) |  | Elevation (DEM) |
| **Top Five Variables** | Elevation (DEM) |  | White Mold |  | Clay Total |
|  | Organic Matter (OM) |  | Upslope |  | Phosphorus |
|  | Soil Type |  | Calcium |  | Potassium |

Table 5. Economic assessment on a per acre basis of variable rate seeding versus a flat farmer rate for three farmer fields over two years.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | 2018 | | |  | 2019 |
|  |  |  | Site 1 |  | Site 2 |  | Site 1 |
|  |  | |  |  |  |  |  |
| **Farmer Seeding Rate (K seeds/ac)** | | | 160 |  | 135 |  | 160 |
| **Average Gross Profit ($/ac)\*** | | | $ 486.70 |  | $ 561.40 |  | $ 490.31 |
|  |  |  |  |  |  |  |  |
| **Seed Cost ($/ac)\*\*** | |  |  |  |  |  |  |
|  | Economic OSD† | | $ 42.02 |  | $ 34.07 |  | $ 38.80 |
|  | Highest yielding OSD‡ | | $ 56.22 |  | $ 40.93 |  | $ 51.42 |
|  | Farmer Rate | | $ 60.57 |  | $ 51.11 |  | $ 60.57 |
|  |  |  |  |  |  |  |  |
| **Profit ($/ac)** | | |  |  |  |  |  |
|  | Economic OSD | | $ 444.68 |  | $ 527.33 |  | $ 451.50 |
|  | Highest yielding OSD | | $ 430.48 |  | $ 520.47 |  | $ 438.88 |
|  | Farmer Rate | | $ 426.13 |  | $ 510.29 |  | $ 429.73 |
|  |  |  |  |  |  |  |  |
| **Economic Gain: Economic OSD v Farmer Rate per ac** | | | | | | |  |
|  | |  | $ 18.55 |  | $ 17.04 |  | $ 21.77 |

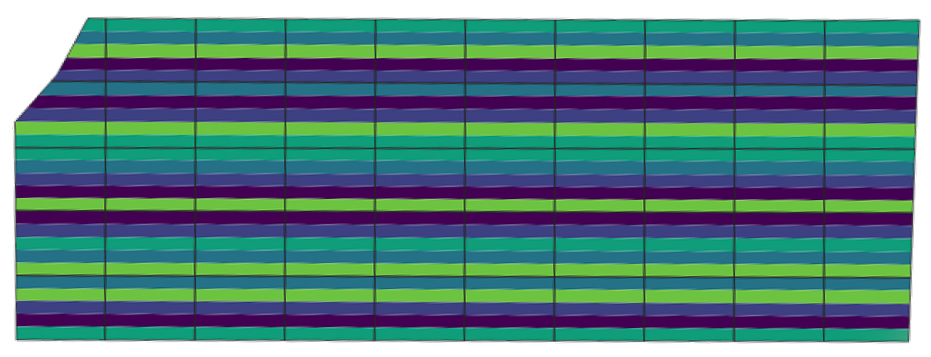
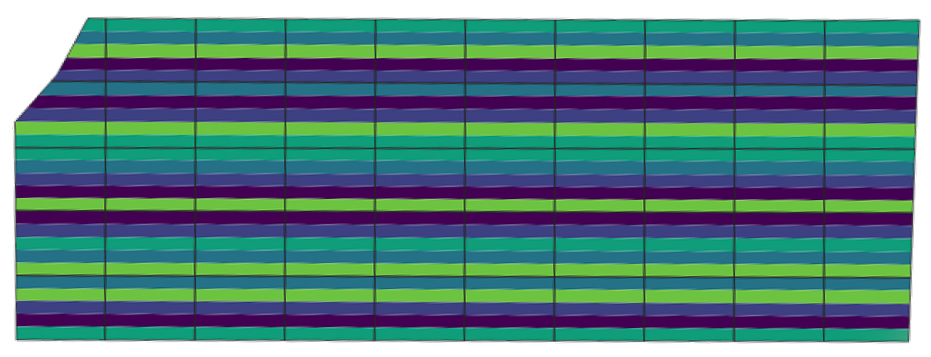
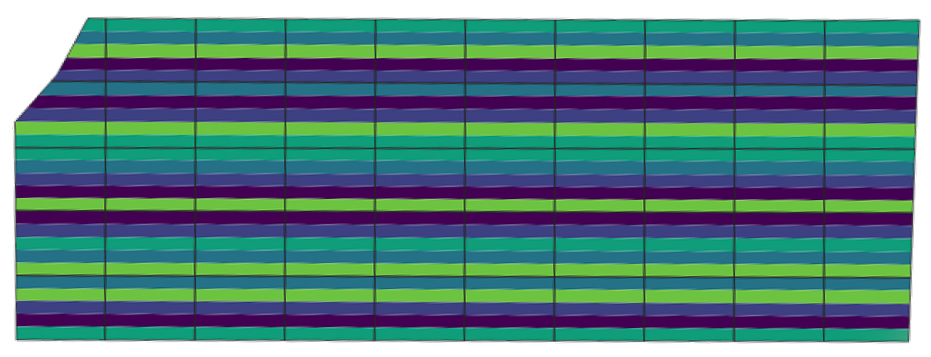
\*Average Gross profit per acre, calculated from the average yield of each 1-ac block and a soybean market price of $9

\*\*Average seed cost per acre, calculated from each 1-acre block at a seed cost of $53 per 140K seeds

† Economic Optimal Seeding Density; the lowest seeding density that was statistically the highest yielding or not statistically different from the value of the highest yielding seeding density

‡Highest Yielding Optimal Seeding Density; the highest yielding seeding density based on value. Not statistically different from other seeding densities.

## Figures



|  |  |
| --- | --- |
| Seeding Rate  (K seeds per ac) | |
|  | **75** |
|  | **100** |
|  | **125** |
|  | **150** |
|  | **175** |

Figure . Experimental design for variable rate seeding of soybean for a 50-acre farmer field in Southwest Minnesota. Five seeding rates were used for the experiment: 75,000; 100,000; 125,000; 150,000; and 175,000 seeds per acre. The north to south subdivisions

Figure 2. Average yield as influenced by five seeding densities for two sites and two years in Minnesota.

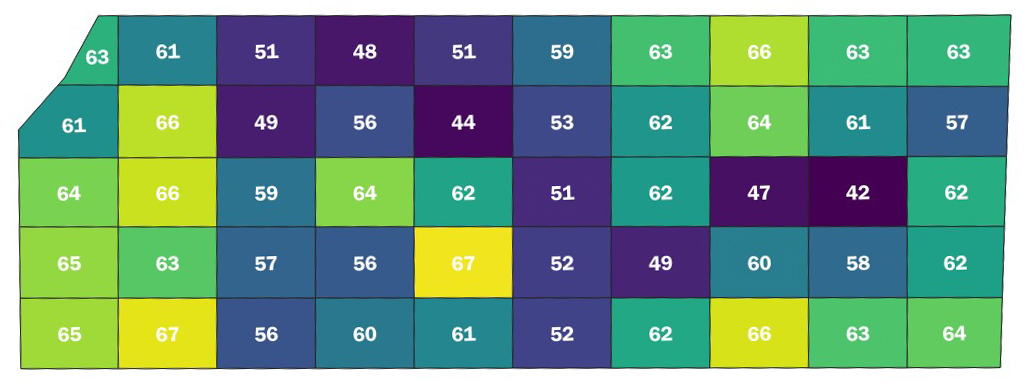


Figure 3. Within-field variability of soybean yield at a farmer field (site 1) in 2018. Yield ranged from 42 bushels per acre (bu/ac) to 67 bu/ac.

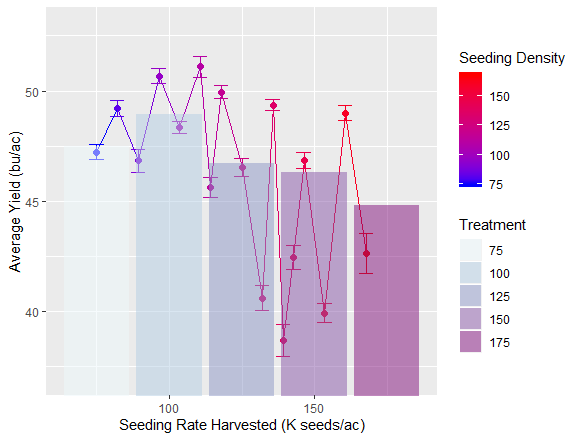


Figure 4. Yield response to various seeding rates at a site in central Minnesota. Two approaches were used to evaluate seeding density success. The bar graph indicates results derived from interpolated yield data. The line represents Pseudo seeding densities derived from each harvest pass.

23

21

27

25

5

0

10

20

75

100

125

150

175

a)

b)

29

30

34

40

17

0

10

20

30

40

75

100

125

150

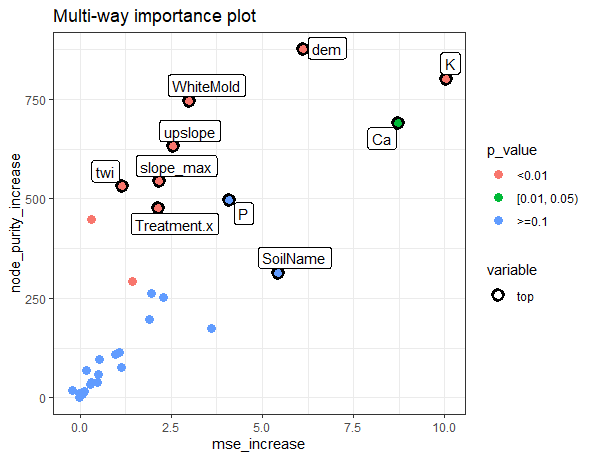
175

Seeding Rate (K seeds/ac)

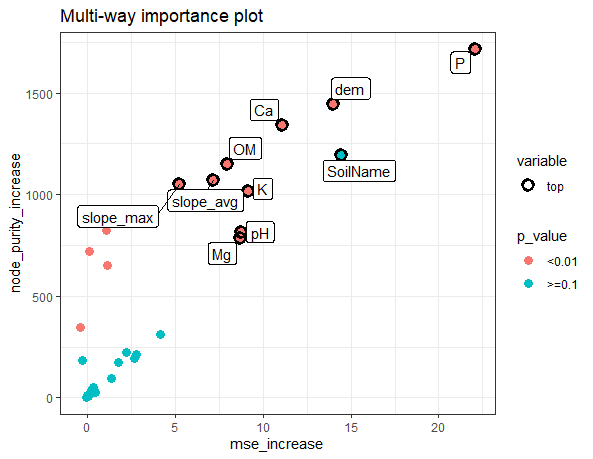
Count of Statistically highest OSD

Seeding Rate (K seeds/ac)

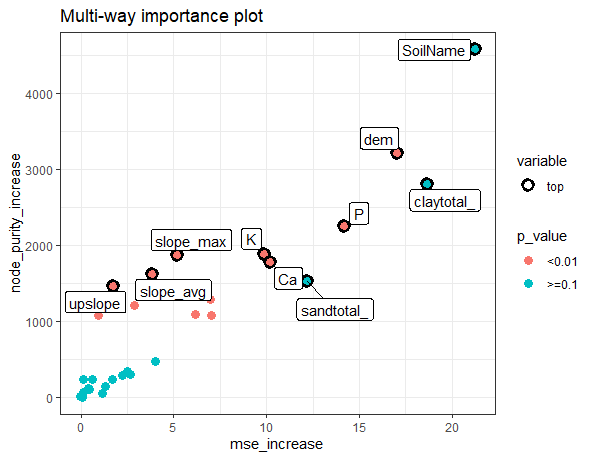
Figure . Count of times each seeding density (75, 100, 125, 150, and 175) was considered an optimal seeding density for site 1 for the 2018 (a) and 2019 (b) growing seasons using Tukey’s HSD test.



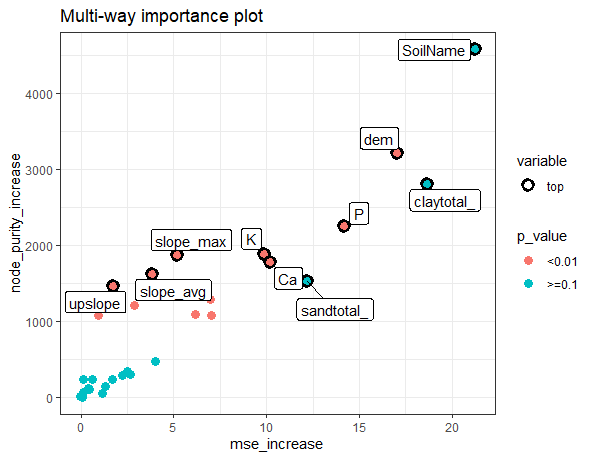
***Site 2***



***Site 1***

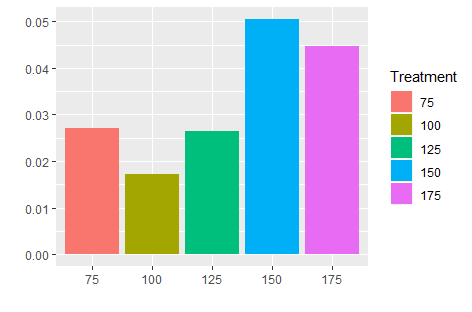
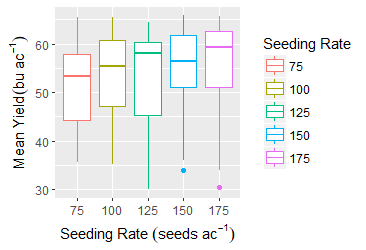


***Pooled Sites***



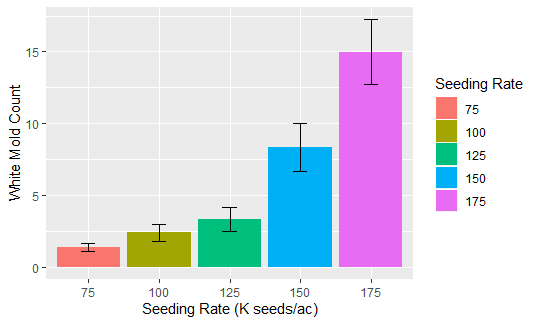
***Pooled Sites***

Figure . Importance of variables determined by the Random Forest model for Site 1, 2, and pooled sites. Higher MSE increase indicates a more informative variable. Higher node purity describes the structure of the tree, with higher values indicating splits based on the variable.



Percent of Disease Incidence

2018



2019

SSSSeeding Rate (K seeds/ac)

Figure . Disease Incidence (White Mold) across five seeding densities at a central Minnesota farm for 2 years.

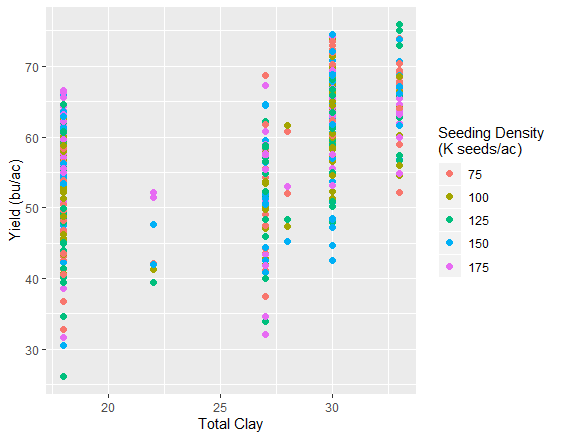


Figure 8. Relationship of yield with soil clay content for two farm fields in Minnesota planted at five seeding densities.

**Bibliography**

Carciochi, W.D., R. Schwalbert, F.H. Andrade, G.M. Corassa, P. Carter, et al. 2019. Soybean Seed Yield Response to Plant Density by Yield Environment in North America. Agron. J. 0(0): 0. doi: 10.2134/agronj2018.10.0635.

Castrignanò, A., G. Buttafuoco, R. Quarto, C. Vitti, G. Langella, et al. 2017. A combined approach of sensor data fusion and multivariate geostatistics for delineation of homogeneous zones in an agricultural field. Sensors (Switzerland) 17(12). doi: 10.3390/s17122794.

Corassa, G.M., T.J.C. Amado, M.L. Strieder, R. Schwalbert, J.L.F. Pires, et al. 2018. Optimum soybean seeding rates by yield environment in southern Brazil. Agron. J. 110(6): 2430–2438. doi: 10.2134/agronj2018.04.0239.

Gaspar, A.P., S. Mourtzinis, D. Kyle, E. Galdi, L.E. Lindsey, et al. 2020. Defining optimal soybean seeding rates and associated risk across North America. Agron. J. (August 2019): 1–12. doi: 10.1002/agj2.20203.

Jaccoud-Filho, D. de S., F.F. Sartori, M. Manosso-Neto, C.M. Vrisman, M.L. d. C. Pierre, et al. 2016. Influence of row spacing and plant population density on management of “white mould” in soybean in southern Brazil. Aust. J. Crop Sci. 10(2): 161–168.

Jaynes, D.B. 2010. Confidence bands for measured economically optimal nitrogen rates. Precis. Agric. doi: 10.1007/s11119-010-9168-3.

Khosla, R., D. Inman, D.G. Westfall, R.M. Reich, M. Frasier, et al. 2008. A synthesis of multi-disciplinary research in precision agriculture: Site-specific management zones in the semi-arid western Great Plains of the USA. Precis. Agric. 9(1–2): 85–100. doi: 10.1007/s11119-008-9057-1.

Smidt, E.R., S.P. Conley, J. Zhu, and F.J. Arriaga. 2016. Identifying field attributes that predict soybean yield using random forest analysis. Agron. J. doi: 10.2134/agronj2015.0222.

Vega, A., M. Córdoba, M. Castro-Franco, and M. Balzarini. 2019. Protocol for automating error removal from yield maps. Precis. Agric. (0123456789). doi: 10.1007/s11119-018-09632-8.