Published October 11, 2018

Optimum Soybean Seeding Rates by Yield Environment in Southern Brazil

Geomar M. Corassa,* Telmo J. C. Amado, Mércio L. Strieder, Rai Schwalbert, João L. F. Pires, Paul R. Carter, and Ignacio A. Ciampitti*

ABSTRACT

Optimizing seed inputs while increasing farming profit is the main purpose of variable rate seeding (VRS) technology adoption. Previous studies in corn (*Zea mays* L.) suggested that optimal seeding rates increase as yield productivity level increased. For soybean [*Glycine max* (L.) Merr.], optimal yield-to-seeding rate by yield level has not been fully investigated, representing a scientific knowledge gap. Therefore, a dataset was collected from 109 replicated field trials from Southern Brazil (2180 experimental units) presenting the following objectives: (i) identify the optimum seeding rate at varying yield levels (herein termed as yield environments), and (ii) explore the contribution of management factors (i.e., seeding rate, planting date, row spacing, maturity groups, growing season, yield environment, and ecological region) on soybean seed yield. Hierarchical modeling and Bayesian statistical inference were used to predict optimum seeding rate at varying yield environments, while conditional inference tree analysis was explored to identify and rank factors contributing to yield variation. The main results were: (i) soybean seeding rate increased from high- to low-yielding environments; (ii) seeding rate could be reduced by 18% in high-yielding (>5 Mg ha^{-1}) relative to the low-yielding (<4 Mg ha^{-1}) environments, without penalizing yields. For improving site-specific soybean seeding rate prescriptions, future studies should focus on the physiological mechanisms underpinning yield formation and on understanding the main factors (soil × plant × weather) contributing to the differential optimum seeding rate response.

Core Ideas

- Soybean yield response to seeding rate was dependent on yield environment.
- Optimum seeding rate increased as yield environments were reduced.
- Seeding rate could be reduced by 18% for high-yielding relative to low-yielding environments, without penalizing yields.
- Planting date interacts with seed yield response to seeding rate, optimum seeding rates increase with late planting.
- For high-yielding environment, late planting time decreased yields regardless of the seeding rate.

Published in Agron. J. 110:2430–2438 (2018)
doi:10.2134/agronj2018.04.0239

Globally, soybean [*Glycine max* (L.) Merr.] is one of the most cultivated field crops, planted on 120 million hectares (FAO, 2016). Among the primary producing countries, Brazil is the second largest producer with nearly 115 million metric tons (CONAB, 2017). Seed yield potential is associated with genetic attributes, environmental conditions (i.e., geographical position, soil, weather), management practices (i.e., plant density, row spacing), and their interactions (van Ittersum and Rabbinge, 1997; Evans and Fisher, 1999; Vanlauwe et al., 2003; Rowntree et al., 2013; Van Roekel et al., 2015). At the field-level, management practices are applied as a strategy to reduce the gap between current and attainable yields (i.e., yield under optimal management) (van Ittersum et al., 2013; Bunselmeyer and Lauer, 2015).

From the standpoint of management practices, seeding rate is one of the main factors controlled by growers (Egli, 1988; Lee et al., 2008; Walker et al., 2010; Cox and Cherney, 2011; Mueller et al., 2014; Thompson et al., 2015). Consequently, many studies have been conducted globally on the effect of seeding rate on soybean yields (Lee et al., 2008; De Bruin and Pedersen, 2008; Walker et al., 2010; Coulter et al., 2011; Rahman et al., 2011; Cox and Cherney, 2011; Thompson et al., 2015; Ferreira et al., 2016). Conceptually, soybean yield response to plant density can be separated into three phases (Duncan, 1986): (i) yield-density model without plant competition; yield mainly depending on the individual plant contribution; (ii) yield-density model at canopy-scale, community of plants increasing light interception on a unit-area basis until yield reaches a plateau; and (iii) yield-density model after yield has plateaued, further seeding rate increase does not improve yield.

Environmental conditions such as yield potential could play an important role on the optimum seeding rate prescription (Egli, 1988; De Bruin and Pedersen, 2008; Lee et al., 2008; Walker et al., 2010; Van Roekel and Coulter, 2011; Rowntree

G.M. Corassa, R. Schwalbert, Dep. of Agricultural Engineering, Federal Univ. of Santa Maria, Rural Science Center, Santa Maria, RS, Brazil; G.M. Corassa, R. Schwalbert, I.A. Ciampitti, Dep. of Agronomy, Kansas State Univ., 2004 Throckmorton Plant Science Center, Manhattan, KS 66506; T.J.C. Amado, Dep. of Soil, Federal Univ. of Santa Maria, Rural Science Center, Santa Maria, RS, Brazil; M.L. Strieder, Embrapa, Secretaria de Pesquisa e Desenvolvimento, Brasília, DF 70770-90, Brazil; J.L.F. Pires, Embrapa Trigo, BR-285, Km 294, Passo Fundo, RS 99050-970, Brazil; P.R. Carter, DuPont Pioneer, 7100 NW 62nd Ave., PO BOX 1000, Johnston, IA 50131. Received 8 Apr. 2018. Accepted 15 July 2018. *Corresponding authors (geomarmateus@hotmail.com; ciampitti@ksu.edu).

Abbreviations: DOY, day of the year; HY, high yielding; ICC, interclass correlation coefficient; LY, low yielding; MY, medium yielding; VRS, variable rate seeding.

Table 1. Site name and location, soybean ecological region, growing season, number of genotypes, maturity groups, and seeding rates evaluated for each site-year in Southern Brazil.

Site	Soybean ecological region†	Growing season	Number of genotypes	Maturity groups	Seeding rate (× 1000 ha^{-1})
Campo Novo, RS	102	2014/2015	10	9 (from 4.2 to 6.3)	100, 230, 300, 360, 490
(27°39′ S; 53°49′ W)		2015/2016	8	5 (from 5.3 to 6.2)	100, 230, 300, 360, 490
		2016/2017	7	5 (from 5.6 to 6.2)	100, 230, 300, 360, 490
Gentil, RS	102	2015/2016	8	5 (from 5.3 to 6.2)	100, 230, 300, 360, 490
(28°26′ S; 52°02′ W)		2016/2017	6	4 (from 5.7 to 6.2)	100, 230, 300, 360, 490
Passo Fundo, RS	102	2012/2013	2	1 (5.6)	100, 200, 300, 400, 500
(27°14′ S; 52°24′ W)		2013/2014	2	1 (5.6)	100, 230, 300, 360, 490
		2014/2015	12	9 (from 4.2 to 6.3)	100, 200, 300, 400, 500
		2015/2016	8	5 (from 5.3 to 6.3)	100, 230, 300, 360, 490
		2016/2017	8	5 (from 5.6 to 6.2)	100, 230, 300, 360, 490
São Luiz Gonzaga, RS	102	2015/2016	7	4 (from 5.7 to 6.2)	100, 230, 300, 360, 490
(28°23′ S; 54°59′ W)		2016/2017	6	4 (from 5.7 to 6.2)	100, 230, 300, 360, 490
Vacaria, RS	103	2015/2016	8	5 (from 5.3 to 6.2)	100, 230, 300, 360, 490
(28°27′ S; 50°56′ W)		2016/2017	7	5 (from 5.6 to 6.2)	100, 230, 300, 360, 490
Guarapuava, PR	103	2014/2015	10	9 (from 4.2 to 6.2)	100, 230, 300, 360, 490
(25°25′ S; 51°31′ W)					

† Ecological region 102 cover partially the states of Paraná, Santa Catarina and Rio Grande do Sul, represents medium to high altitudes (from 150 to 900 m), and Cfa and Cfb as the Köppen's climate classification (Alvares et al., 2013); Ecological region 103 cover partially the states of São Paulo (south) Paraná (northeast), Santa Catarina (central) and Rio Grande do Sul (northeast), represents high altitudes (>600 m), and Cfb as the Köppen's climate classification (Alvares et al., 2013).

et al., 2013; Thompson et al., 2015). Currently, little is known about the opportunity of adjusting optimum seeding rate according to yield levels or yield potential for soybean. Many farmers increase soybean seeding rates in lower yielding zones of fields (Lowenberg-DeBoer, 1999), but this practice has not been well-documented from a research standpoint. Improved understanding on this topic could shed light on optimizing seed input use by productive zone within a field, as well as increasing the return of investment. The current study provides a science-based foundation for the adoption of variable rate seeding (VRS), a precision agriculture technology available for modern planters (McBratney et al., 2005; Khosla et al., 2008; Gebbers and Adamchuk, 2010; Hörbe et al., 2013; Shearer and Pitla, 2014), for soybean. Yield-density models by yield level (herein termed as yield environment) facilitating the implementation of VRS technology were recently published for corn (Assefa et al., 2016; Schwalbert et al., 2018) and canola (Assefa et al., 2017). Thus, the main goal of this study was to identify the optimum soybean seeding rate at varying yield environments to provide a science-based foundation for adoption of VRS technology. Following this rationale, Bayesian statistical inference models were applied as the main approach to predict the probability of changing seeding rates across yield environments optimizing or without penalizing yields. Lastly, a conditional inference tree analysis was explored to identify and rank the main management factors contributing to variation on the soybean seed yield and seeding relationship at varying yield environments.

MATERIALS AND METHODS

Data Description

Soybean seeding rate data were aggregated from a combination of 15 site-years for soybean seeding rate trials performed by Embrapa between 2012 to 2013 and 2016 to 2017 growing seasons, in six dryland sites from Southern Brazil (2,180 experimental units) (Table 1). Soybean seeding rate trials were placed in two contrasting ecological regions (Fig. 1) based on the adaptability of soybean cultivars in the region (Kaster and Farias, 2012). The database had 109 site-years by cultivar combinations. All research trials were performed in a split-plot design with a randomized block arrangement with four replications, in a plot of 3-m width by 5-m length. Cultivars were the main-plot, and five seeding rates, ranging from 100,000 to 500,000 seed ha^{-1}, were the sub-plot level.

Experimental units were uniformly fertilized with all recommended nutrients following regional prescriptions. Weed, insect and disease control were accomplished according to the best management practices for soybean. For each site-year combination in addition to seeding rates, three main management variables were considered for the current analysis: planting date (ranging from 5 October to 15 December), row spacing (ranging from 20 to 45 cm), and maturity group (ranging from 4.2 to 6.3). Seed yield was obtained from the central two-rows for each plot and adjusted to 130 g kg^{-1} moisture. Not all planting date, row spacing, and maturity groups were tested in each site-year. The final plant density (number of plants harvested) was not available in our dataset; thus, seeding rate at planting time was considered as the main factor for evaluating the yield-density models.

Data Analysis

The yield data (Fig. 2A) were divided in three yield environments following the percentiles of data distribution (<33%, 33–66%, and >66%) for low (LY), medium (MY), and high (HY) yielding levels (Fig. 2B). The average yield, across all seeding rates at each site-year combination, was used as the approach to classify yield environment (Assefa et al., 2016). This method represents the interaction between site by environmental conditions within a year and the yield variation is only due to the experimental treatments (Assefa et al., 2016). The classification was based on frequency of the yield data distribution. A motivation behind this classification was to obtain balanced number of observations across yield environments.

To identify seed yield variation accounted by known factors such as yield environment, growing season, ecological region,

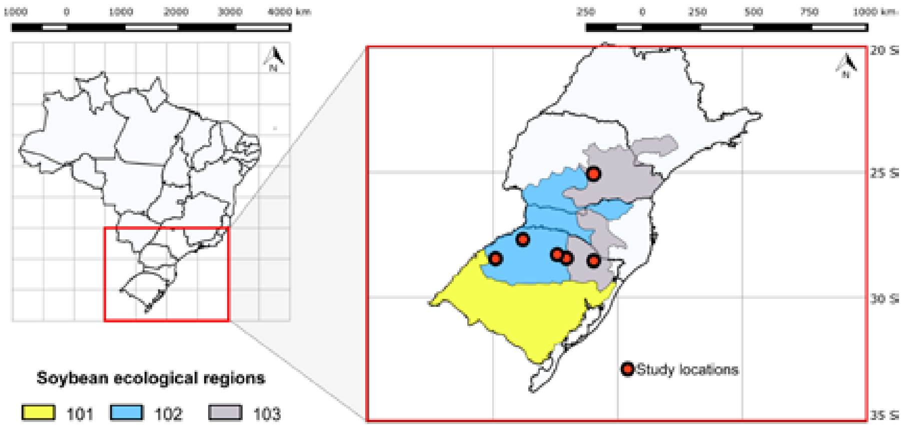

Fig. 1. Map of locations where experimental seeding rate trials were performed. The ecological regions are highlighted with different colors (101, yellow; 102, blue, and 103, gray) as proposed by Kaster and Farias (2012). The current classification is large-scale adopted to test the adaptability of soybean cultivars; the main characteristics are presented as follows: ecological region 101) located in the state of Rio Grande do Sul, represents the highest latitudes in the country, low altitudes (~100 m or less) and climate as Cfa according to Köppen's classification (Alvares et al., 2013). Ecological region 102 cover a 24 to 29.5° S latitude range (states of Paraná, Santa Catarina and Rio Grande do Sul), medium to high altitudes (from 150 to 900 m), and Cfa and Cfb as the Köppen's climate classification (Alvares et al., 2013). Lastly, the ecological region 103 cover partially the states of São Paulo (south), Paraná (northeast), Santa Catarina (central) and Rio Grande do Sul (northeast), and superior altitudes (>600 m), and Cfb as the Köppen's climate classification (Alvares et al., 2013).

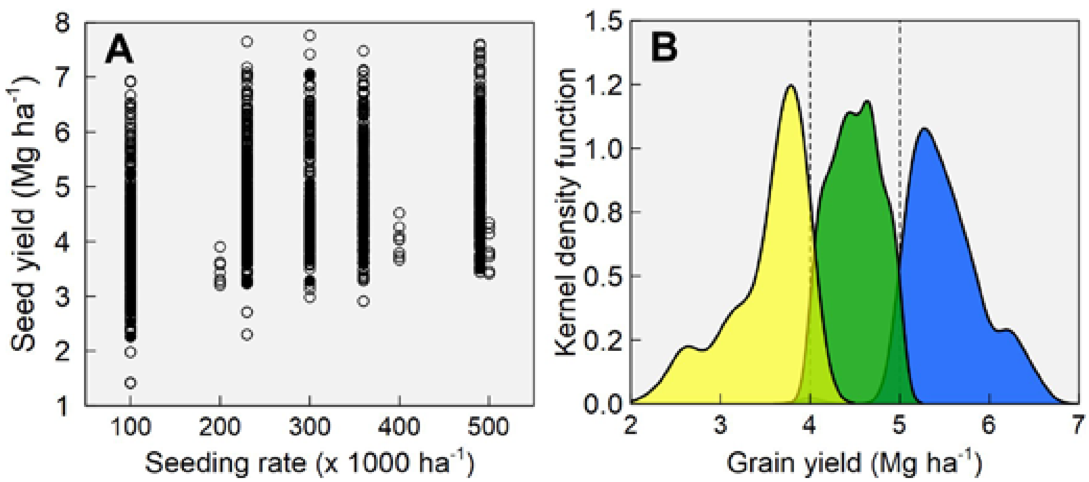

Fig. 2. Dataset of soybean seeding rate versus seed yield aggregated from a combination of 15 site-years (A) and frequency distribution classification of soybean yield for three yield environments: low (LY, <4 Mg ha^{-1}; yellow), medium (MY, 4–5 Mg ha^{-1}; green), and high (HY, >5 Mg ha^{-1}; blue) (B). Yield environments were delineated by average of site-year yield approach. Yield environments were classified by percentiles (<33%, LY; 33–66%, MY; >66%, HY).

seeding rate, planting date (DOY, day of year), row spacing, maturity group, and their respective interactions, the variance was estimated using the *nlme* procedure and *VarComb* package in R program (R Development Core Team, 2013). The variance components were represented by yield environment, ecological region within yield environment, growing season within yield environment, and the interaction ecological region × growing season within yield environment. Management practices such as DOY, plant density, row spacing, and maturity group were analyzed separately. These management factors were treated as fixed effects, whereas yield environment (or yield level), growing season, and ecological region were considered as random variables. The interclass correlation coefficient (ICC) was defined as the yield variance effect divided by the yield variance total.

Hierarchical Modeling and Bayesian Statistical Inference

Hierarchical Bayesian models were implemented to quantify soybean yield response to seeding rate. Large agronomic data sets coming from complex biological systems usually present an elevated degree of uncertainty, variability, and correlation (e.g., spatial and temporal) observed at different levels (e.g., within- and between-fields). Among the most recent proposed solutions for dealing with data portraying the aforementioned characteristics are the Hierarchical Bayesian models (Cressie et al., 2009; Gelman et al., 2004; Gelman and Hill, 2007). Those models can address multiple sources of variability at different scales or levels (Kyveryga et al., 2013). In contrast with standard frequentist approaches, Bayesian Hierarchical inference models provide better information about the source of variations and the difference in means than a simple *p*-value (Meredith and Kruschke, 2018). The Bayesian Hierarchical models can provide complete distributions of credible values for the effect size, group means and their difference, allowing researchers access to a more complete understanding of the effect size of the studied treatments. Hierarchical models represent the environmental interactions using a series of conditional probability distributions (Kyveryga and Blackmer, 2014).

The model comprised three hierarchical levels: field-level (site-year), yield environment-level (low, medium, and high), and regional-level. First, regression models were fitted to the field-level. As a second step, those models were aggregated to a higher hierarchical level (i.e., environment and regional level). Thus, three statistical models were tested individually for all the *n* fields to identify the yield–seeding rate relationship: linear with plateau, quadratic, and quadratic with plateau. These models were selected based on the typical soybean seed yield to seeding rate relationship (Popp et al., 2006; De Bruin and Pedersen, 2008; Cox and Cherney, 2011; Thompson et al., 2015). Soybean yield is assumed to increase at a decreasing rate until a specific seeding rate, at which point yield is expected to either plateau or decrease (Thompson et al., 2015). All the parameters of the response model (intercept, linear, angular coefficients and breakpoint) were assumed to follow a normal distribution with μ_n and precision λ_n unique for each field. The precision parameter was defined as the reciprocal of variance (i.e., higher precision with lower variation) (Kyveryga and Blackmer, 2014). The yield environment coefficients were expressed as conditional distribution of field means μ_n, given regional means and regional precisions distributions. Finally, for the regional model (global model), the yield environment precision parameters λ_n were assumed to follow a γ distribution with parameters α and β (Kyveryga and Blackmer, 2014).

All prior distributions were assumed to be "diffuse" (presenting small precision, large variances), having little influence on the analysis relative to the observed data (Kyveryga et al., 2013). The prior distributions were set accordingly the magnitude of the coefficient to be estimated since the intercept was estimated at the scale of the dependent variable (i.e., seed yield [Mg ha^{-1}]). Breakpoint is estimated at the scale of the independent variable (i.e., seeding rate [1000 × seed ha^{-1}]), and linear and angular coefficients were estimated at the smallest scale since they control the rate of change in soybean seed yield as function a of seeding rate at the scale units abovementioned. A Markov-chain Monte Carlo simulation was used for this approach (Gelman and Hill, 2007) following a Gibbs sampling algorithm with 15,000 random draws after a warm up period of 5,000 interactions. The *rjags* package (Plummer, 2016) was used to build the models in the R program. The models were parametrized using precision parameters, that is the default option for the package. All models were run in the Beocat Research Cluster at Kansas State University due to the high demand for computing power. Based on prior distributions, that were built to represent the possible values of observations using Bayesian analysis, we updated these values in a posterior predictive-probabilities distribution (Kyveryga et al., 2013) for each yield environment. Since the main focus of the work was to investigate the yield and seeding rate response models in a given yield environment, the hierarchical regional-level was not explored.

Conditional Inference Trees

A conditional inference regression tree analysis was executed to examine relevant interactions as well as significant sources of variation for yield and seeding rate factors. This approach is an alternative to overcome bias since it does not imply statistical assumptions relative to the data distribution. The conditional inference regression tree can be implemented using categorical and continuous explanatory variables and is robust for outliers, missing data, exposing variable interactions (Hothorn et al., 2006; Tittonell et al., 2008; Hastie et al., 2009). In addition, this approach has been recently implemented to identify yield constraints in field crops (Lobell et al., 2005; Ferraro et al., 2009; Mourtzinis et al., 2018). A more detailed explanation of the benefits of use conditional inference regression tree methodology was recently described by Mourtzinis et al. (2018). The *partykit* package (Zeileis and Hothorn, 2015) in the R program was used. The criterion for the independence test was based on univariate *p*-values ($\alpha = 0.05$). The number of intermediate, terminal nodes, and the maximum tree depth were set according *partykit* package default (Zeileis and Hothorn, 2015). To build a more robust decision tree a training/validation approach was implemented. Sampling without replacement was performed aiming at selecting 70% of the total data points for training and the remaining 30% for validation purposes. Separation of the dataset into training and validation was required to self-test model replicability with an independent dataset. The accuracy of the model was accessed using the root mean square error (RMSE) from the cross-validation. Lastly, training and validation datasets were pooled together to build the universal (global) model.

Table 2. Estimation of soybean yield variance components in an environmental-based (yield environment, ecological region, and growing season) and management factors-based (DOY, seeding rate, row spacing, maturity group).

Covariance effect	Variance	ICC†
Yield environment	0.862	0.69
Ecological region (yield environment)	0.000	0.00
Growing season (yield environment)	0.000	0.00
Ecological region × growing season (yield environment)	0.000	0.00
DOY (yield environment growing season ecological region)	0.082	0.07
Seeding rate (yield environment growing season ecological region)	0.063	0.05
Row spacing (yield environment growing season ecological region)	0.020	0.01
Maturity group (yield environment growing season ecological region)	0.161	0.01
High level interaction and residual	0.209	0.17

† ICC, interclass correlation coefficient, defined as yield variance effect/yield variance total.

RESULTS

Yield variability across site-years was largely (ICC = 0.69) explained by yield environment (e.g., HY, MY, and LY) (Table 2). The DOY and seeding rate (both within yield environment, growing season, and ecological region effects) were the second and third main factors in order of importance, accounting for 7 and 5% of yield variation, respectively (Table 2). Across all factors, growing season (year) and soybean ecological region (e.g., 102 and 103) accounted for a small yield variation relative to the other sources of variability. Similar results were found for row spacing and maturity group, both (combined) accounted for 2% of yield variability. The amount of variability accounted for unexplained factors (high level interaction and residual) was 17% (ICC = 0.17) (Table 2).

Among the statistical models evaluated, the linear with plateau model best explained the yield-seeding rate relationship; however, the average seeding rate at yield-plateau (breakpoint) varied across yield environments (Fig. 3).

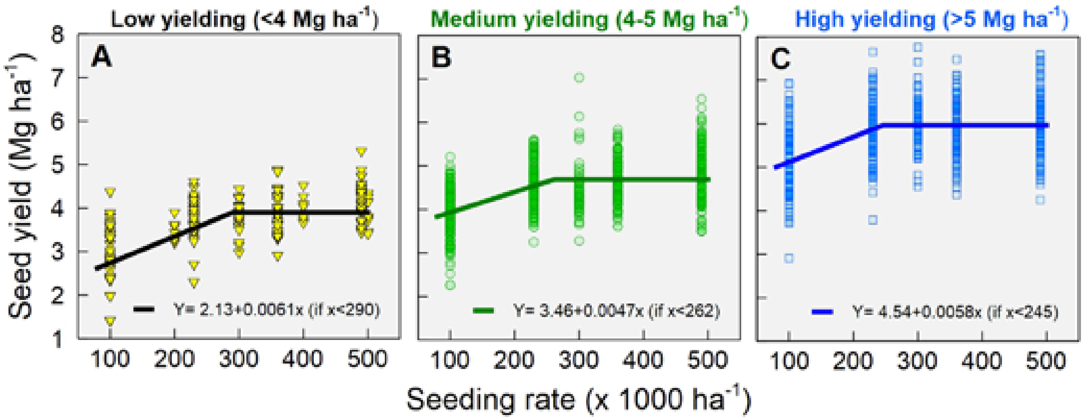

Fig. 3. Bayesian regression models from soybean seed yield relative to the seeding rate for low (yield <4 Mg ha^{-1}; yellow) (A), medium (yield 4–5 Mg ha^{-1}; green) (B), and high (yield >5 Mg ha^{-1}; blue) (C) yield environments. The model represents the most probable response across site-years × cultivar combination evaluated.

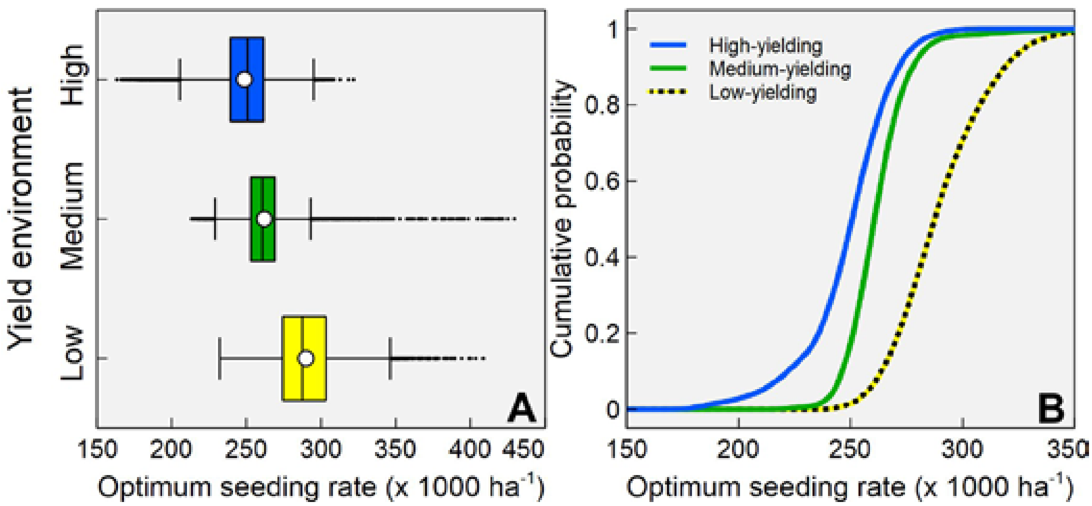

Fig. 4. Panel A denotes optimum seeding rate range obtained from site-years × cultivar combination to attain the yield plateau for low (yield <4 Mg ha^{-1}; yellow), medium (yield 4–5 Mg ha^{-1}; green), and high (yield >5 Mg ha^{-1}; blue) yield environments. In each boxplot, the central rectangle extends from the first to third quartile (percentiles 25 and 75). The circle inside the rectangle represents the mean value of seeding rate to attain the yield plateau across site-years and for each yield environment. Whiskers extend between the smallest and the largest non-outlier values. Black points before and after whiskers denote outliers. Panel B is the posterior predictive probabilities of optimum seeding rate to achieve the yield plateau at low, medium, and high yield environments. Yield environments were delineated by average of site-year yield approach.

Using the Bayesian Hierarchical inference models within each yield environment, the average yield at the plateau followed the order: LY > MY > HY (Fig. 3). Average seeding rate at the plateau was 10% greater for the LY (290 thousand seeds ha^{-1}) than MY (262 thousand seeds ha^{-1}), and 18% greater for the LY than HY (245 thousand seeds ha^{-1}) (Fig. 3). The slope for the linear function was also slightly superior for the LY (0.061) relative to MY (0.047) and HY (0.058) (Fig. 3).

Considering all the site-specific effects (plot-level data from the site-years by cultivar combination), the 50% interquartile range (between 25 and 75 quartiles) for the optimal seeding rate (yield plateau) ranged between 274 and 303 thousand seeds ha^{-1} for LY, 252 and 269 thousand seeds ha^{-1} for MY, and 238 and 262 thousand seeds ha^{-1} for HY environments (Fig. 4A). Based on prior distributions, we updated these values in a form of posterior cumulative distributions as an approach to predict the probabilities of optimal seeding rate across yield environments (Fig. 4B). For HY environment, there is a 90% chance of the optimal seeding rate being smaller than 270 thousand seeds ha^{-1}. For MY environment, this probability was attained with seeding rates smaller than 280 thousand seeds ha^{-1} (Fig. 4B). Lastly, for the LY environment, a 90% chance to attain the yield plateau was documented for seeding rates smaller than 320 thousand seeds ha^{-1} (Fig. 4B). The cumulative probability level of 90% could be considered as a threshold for on-farm seeding rate prescriptions in the region, presenting a low probability of improving yield with further increase in seeding rate (Fig. 4B). Reduction in seeding rate (<250 thousand seeds ha^{-1}) had less influence on seed yield for HY compared to MY and LY (Fig. 4B).

The conditional inference tree was fitted to identify and rank sources of yield variation for the factors collected within the dataset. Thus, all the known measured factors (i.e., yield environment, ecological region, growing season, DOY, seeding rate, row spacing, and maturity group) were included in the model. All the exploratory variables were treated as continuous factors, which means that the criteria to "node establishment" was based on the model and their respective significance ($\alpha = 0.05$). The results suggested that seeding rate and DOY were significant

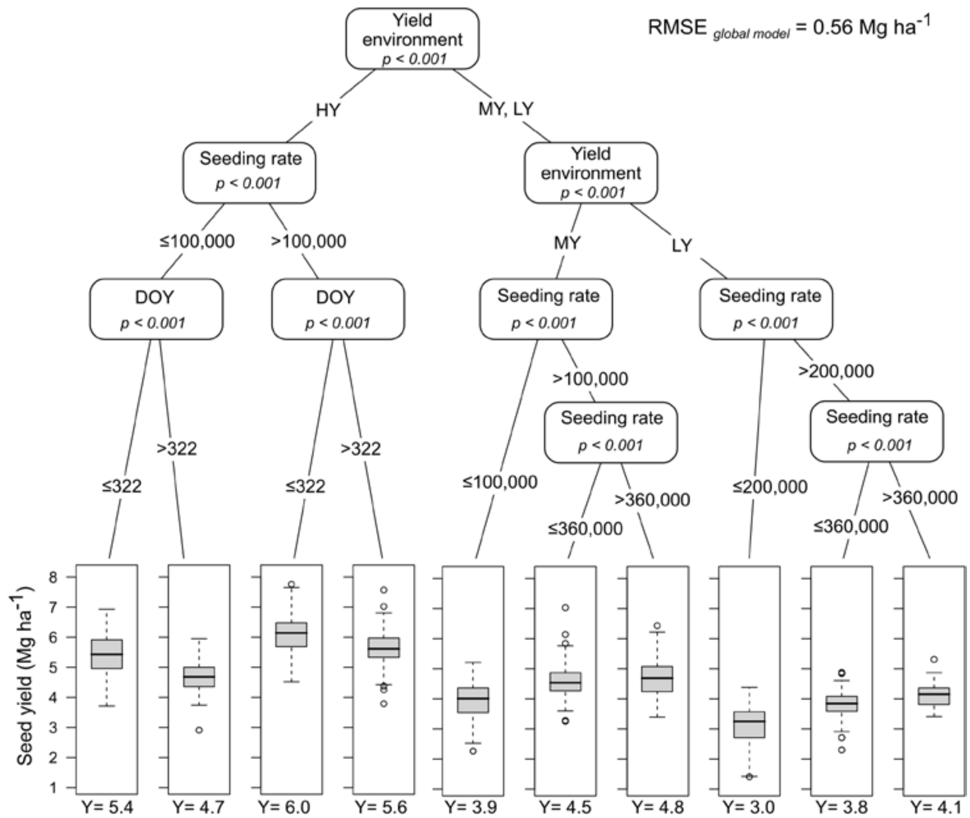

Fig. 5. Conditional inference tree across the 15 site-years evaluated. Boxplots in the bottom of the figure represent the soybean seed yield. In each boxplot, central rectangle extends the first to the third quartile. The solid line inside the rectangle represent the mean yield (numerical value is shown at the boxplot bottom). The vertical lines above and below the rectangle denote the maximum and minimum, respectively. Circles represent outliers. The criterion for the independence test was based on univariate *p*-values ($\alpha = 0.05$). Global model RMSE = 0.56 Mg ha^{-1}, Cross validation RMSE = 0.68 Mg ha^{-1}. HY, high-yielding environment; MY, medium-yielding environment; LY, low-yielding environment; DOY, day of the year.

factors influencing yield and seeding rate relationships across yield environments. Other important management factor such as maturity group was not significant, reflecting an opportunity to provide more universal recommendations.

Based on the regression tree model (Global model RMSE = 0.56 Mg ha^{-1}; Cross validation RMSE = 0.68 Mg ha^{-1}), results revealed that under the HY environment, seeding rates of approximately 100 thousand seeds ha^{-1} represented a slight reduction in yield than other rates evaluated (Fig. 5). However, as mentioned above, seeding rates greater than 250 thousand seeds ha^{-1} were likely to represent an unnecessary cost for HY due to the low probability of expected yield increase (Fig. 4B). For the HY environment, the model portrayed that late planting dates (after 18 November, DOY = 322) resulted in lower yields relative to earlier dates regardless of the seeding rate level (Fig. 5). Yield reduction due to late planting (15%) was greater for the lower seeding rate (≤100 thousand seeds ha^{-1}) relative to the other seeding rates (7%) (Fig. 5). The DOY was not a significant factor for MY and LY environments, but seeding rate was a critical factor. For MY, the use of lower seeding rates (≤100 thousand seeds ha^{-1}) represented a yield decrease of about 18% compared with higher seeding rates (Fig. 5). At LY, a linear increase in yield was documented with the increase in seeding rate from 200 to 360 thousand seeds ha^{-1} (yield gain = 23%) (Fig. 5). Overall, these results indicate that a limited number of management practices (seeding rate and planting date) can represent the main effects on soybean yield response across yield environments.

DISCUSSION

Modifying seeding rate by yield environment could represent an opportunity to increase soybean profitability. Since a large database with a satisfactory degree of variability was utilized, the hierarchical modeling and Bayesian inference represented a powerful statistical approach (Kyveryga et al., 2013; Kyveryga and Blackmer, 2014) to obtain credible values of soybean yield response to seeding rates. This type of statistical approach includes the most recent improvements to analyze complex datasets (Cressie et al., 2009; Gelman et al., 2004; Gelman and Hill, 2007). At the regional-level, a low probability of increasing yield was recorded for seeding rates above 330 thousand seed ha^{-1}. For the entire database, a high probability of the yield plateau was obtained in the seeding rate range from 170 to 320 thousand seeds ha^{-1}. The current findings are in agreement with several studies published around the globe (De Bruin and Pedersen, 2008; Epler and Staggenborg, 2008; Lee et al., 2008; Cox et al., 2010; Cox and Cherney, 2011; Luca and Hungría, 2014; Luca et al., 2014; Thompson et al., 2015; Ferreira et al., 2016). Furthermore, conditional inference trees revealed that seeding rate presented more influence than maturity group and row spacing across yield environments.

Advances in precision agriculture technologies are allowing growers to use site-specific management (Shanahan et al., 2004; McBratney et al., 2005; Khosla et al., 2008; Gebbers and Adamchuk, 2010) such as VRS prescriptions on-the-go to optimize yields and input costs (Shanahan et al., 2004; Hörbe et al., 2013; Butzen, 2016; Smidt et al., 2016). Yield-seeding rate response models across yield environments have been recently documented for some major crops, such as corn (Hörbe et al., 2013; Assefa et al., 2016; Schwalbert et al., 2018) and canola (*Brasica napus* L. 'Canola') (Assefa et al., 2017). Theoretical models recently published for corn, revealed that increasing seeding rate at LY should result in flat or negative yield response (Assefa et al., 2016); while at greater yield environments, a greater number of seeds per unit area benefitted yields. Overall, for VRS in corn the optimum seeding rate should follow the order HY > MY > LY environments (Hörbe et al., 2013; Assefa et al., 2016; Schwalbert et al., 2018). For canola, yield-to-plant density relationship showed a smaller effect for HY (>2.5 Mg ha^{-1}) and MY (1.5–2.5 Mg ha^{-1}) environments, but a quadratic model as the best fit for the LY (<1.5 Mg ha^{-1}) environment (Assefa et al., 2017). For soybean, due to lack of a clear relationship between productivity level, fields are often seeded at a single rate (Smidt et al., 2016). The current study helps to provide a science-based foundation for the yield-seeding rate response for soybean, with the optimum seeding rate following the order LY > MY > HY environments. In agreement, Smidt et al. (2016) found similar responses for soybean seed yield to seeding rate when a yield-seeding rate model was obtained within their datasets. These outcomes are similar to the response presented in canola (Assefa et al., 2017), but opposite of that for corn (Hörbe et al., 2013; Assefa et al., 2016; Schwalbert et al., 2018).

A few hypotheses can be postulated for the soybean yield-seeding rate response at the LY environment. One of them (i) involves the reduced "reproductive" ability at plant-level to compensate for low final stands with more pods and seeds per plant such that yields depend on individual production per plant (i.e., poor ability of the plants to compensate for the lack of resources). One possible but less likely factor is that this condition can be aggravated by self-thinning of plants during the growing season due to factors limiting growth. In other words, a LY environment impairs the plants ability to grow faster and reduces inter- and intra-specific competition. The potential results are shorter plants, with reduced canopy coverage, and lower attainable yield. Another important hypothesis (ii) is increased risk of stand failure at LY environments, limiting stand establishment requiring increased seeding rates to compensate for the lower germination and emergence efficiency in those production environments.

In soybean, the compensation mechanism is activated by red/far-red light ratios within the canopy during early stages, increasing the dry mass partitioning to branches and consequently, benefiting the pod production per plant (Board, 2000; Carpenter and Board, 1997; Cox et al., 2010; Kasperbauer, 1987; Norsworthy and Shipe, 2005; Weber et al., 1966). Overall, plants compensate by developing more seeds per plant with fewer plants. The opposite response occurs at supra-optimal plant densities (Luca and Hungría, 2014; Corassa et al., 2018). Thus, based on the first hypothesis, it is probable that such yield compensation is strongly manifested at HY environments due to a greater availability of resources, while at LY environments, the reproductive ability of the plant (e.g., less seeds per plant) is limited, increasing the need of increased seeding rate, more plants, to improve yields.

Recent studies performed at high-yielding environments in Brazil showed that soybean was able to maintain yields even under low densities (Luca et al., 2014; Werner et al., 2016). A reduction in the number of plants by 75% resulted in a yield decrease of 16%, but in two out of three growing seasons yield losses did not occur (Luca et al., 2014). Similarly, a recent study documented that at lower seeding rates (88 thousand seeds ha^{-1}), soybean showed a potential to quadruple both photosynthesis and biological nitrogen fixation, resulting in similar yield per unit area than when greater seeding rates were evaluated (362 thousand seeds ha^{-1}) (Luca and Hungría, 2014). Studies with low densities attaining the yield plateau were also found for the United States (Thompson et al., 2015).

The second relevant hypothesis to be considered is the greater risk of stand failure at LY environments; thus, more seed is required to attain a satisfactory stand. Several field and growing season factors, not assessed in this current study, might be related to the poor emergence, germination, establishment, or plant survivability at LY relative to HY environments. Potential factors such as soil temperature and moisture, compaction, and fertility (Butzen, 2016; Smidt et al., 2016; Sivarajan et al., 2018), as well as early-season plant diseases and weed pressure (Gaspar and Conley, 2014; Thompson et al., 2015; Butzen, 2016) are among those related to the stand establishment challenges that could be greater concern in the LY environments. Thus, future research should be pursued to better understand the soil, plant, weather and other factors behind the higher seeding rate need at LY relative to HY environments across regions and to provide more precise data layers to VRS prescription on soybean (Smidt et al., 2016).

Overall, our findings showed an opportunity for within-field VRS prescriptions across yield environments. Due to the low probability of increased yield with seeding rates above 330 thousand seeds ha^{-1} for modern soybean cultivars, main opportunities behind VRS in soybean are based on reducing seeding rate in the VRS prescription for HY environments, without penalizing yields, and in some occasions, slightly increased seeding rates above current levels for LY environments may enhance yields. Additionally, conditional inference tree models reveled that planting date was a factor that interacted with seed yield response to seeding rate, with increasing optimum seeding rates with late planting. Lastly, adjustments in seeding rates to achieve desired final stand densities should be assumed for environments with high risk of stand losses.

CONCLUSION

To the extent of our knowledge, this is the first assessment of a large dataset of soybean seed yield response to seeding rate in southern Brazil. The hierarchical modeling and Bayesian inference statistical approach were implemented to derive that seeding rate can be optimized differently at varying yield environments. The most probable optimum seeding rate should follow the trend LY > MY > HY environments. Overall, seeding rate could be reduced by 18% at HY relative to LY environments, without penalizing yields (an opportunity for seed savings). There may also be instances where growers should slightly increase seeding rates above current levels for LY environments. Overall, a low probability of yield increase was reported when seeding rates were above 330 thousand seeds ha^{-1} regardless of yield environments. Also, conditional inference tree models showed that planting date interacted with seed yield response to seeding rate, with increasing seeding rate as a potential strategy to mitigate the negative planting delay effects on attainable yield. This study provides a science-based foundation for improving profits by adopting VRS technology for soybean for specific production conditions.

Future research studies should investigate the physiological mechanisms underpinning the yield to seeding rate response related to the yield environments, with the primary role of improving the understanding of the main factors (soil × plants × weather) causing the differential optimum seeding rate response for soybean.

ACKNOWLEDGMENTS

The authors wish to thank CNPq (process numbers 140508/2015-5, 140275/2016-9, and 141144/2017-3) and CAPES (process number 88881.132762/2016-01) for funding this research. Thanks to Embrapa Trigo (project number 02.12.01.001.00.00- New seeding systems and plant arrangements to increase soybean yield and sustainability) for providing the data for this research. This is contribution no. 18-381-J from the Kansas Agricultural Experiment Station.

REFERENCES

Alvares, C.A., J.L. Stape, P.C. Sentelhas, J.L. De Moraes Gonçalves, and G. Sparovek. 2013. Köppen's climate classification map for Brazil. Meteorol. Zeitschrift 22:711–728. doi:10.1127/0941-2948/2013/0507

Assefa, Y., P.V.V. Prasad, C. Foster, Y. Wright, S. Young, P. Bradley, M. Stamm, and I.A. Ciampitti. 2017. Major management factors determining spring and winter canola yield in North America. Crop Sci. 58:1–16. doi:10.2135/cropsci2017.02.0079

Assefa, Y., P.V. Vara Prasad, P. Carter, M. Hinds, G. Bhalla, R. Schon, M. Jeschke, S. Paszkiewicz, and I.A. Ciampitti. 2016. Yield responses to planting density for US modern corn hybrids: A synthesis-analysis. Crop Sci. 56:2802–2817. doi:10.2135/cropsci2016.04.0215

Board, J.E. 2000. Light interception efficiency and light quality affect yield compensation of soybean at low plant populations. Crop Sci. 40:1285–1294. doi:10.2135/cropsci2000.4051285x

Bunselmeyer, H.A., and J.G. Lauer. 2015. Using corn and soybean yield history to predict subfield yield response. Agron. J. 107:558–562. doi:10.2134/agronj14.0261

Butzen, S. 2016. Variable-rate seeding considerations for soybeans. Crop focus- DuPont Pioneer Agron. Sci., Pioneer Hi-Bred International, Inc., Johnston, IA.

Carpenter, A.C., and J. Board. 1997. Branch yield components controlling soybean yield stability across plant populations. Crop Sci. 37:885–891. doi:10.2135/cropsci1997.0011183X003700030031x

Companhia Nacional de Abastecimento (CONAB). 2017. Acompanhamento da safra brasileira de grãos (Nono levantamento- junho 2017) (In Portuguese.). Monit. Agrícola 4:161. Companhia Nacional de Abastecimento, Brazil. http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17_06_08_09_02_48_boletim_graos_junho_2017.pdf (verified 15 June 2017).

Corassa, G.M., T.J.C. Amado, T. Liska, A. Sharda, J. Fulton, and I.A. Ciampitti. 2018. Planter technology to reduce double-planted area and improve corn and soybean yields. Agron. J. 110:300–310. doi:10.2134/agronj2017.07.0380

Coulter, J.A., C.C. Sheaffer, M.J. Haar, D.L. Wyse, and J.H. Orf. 2011. Soybean cultivar response to planting date and seeding rate under organic management. Agron. J. 103:1223–1229. doi:10.2134/agronj2011.0086

Cox, W.J., and J.H. Cherney. 2011. Growth and yield responses of soybean to row spacing and seeding rate. Agron. J. 103:123–128. doi:10.2134/agronj2010.0316

Cox, W.J., J.H. Cherney, and E. Shields. 2010. Soybeans compensate at low seeding rates but not at high thinning rates. Agron. J. 102:1238–1243. doi:10.2134/agronj2010.0047

Cressie, N., C.A. Calder, J.S. Clark, J.M. Ver Hoef, and C.K. Wikle. 2009. Accounting for uncertainty in ecological analysis: The strength and limitations of hierarchical modeling. Ecol. Appl. 19:553–570. doi:10.1890/07-0744.1

De Bruin, J.L., and P. Pedersen. 2008. Soybean seed yield response to planting date and seeding rate in the upper Midwest. Agron. J. 100:696–703. doi:10.2134/agronj2007.0115

Duncan, W.G. 1986. Planting patterns and soybean yields. Crop Sci. 26:584–588. doi:10.2135/cropsci1986.0011183X002600030033x

Egli, D.B. 1988. Plant density and soybean yield. Crop Sci. 28:977–981. doi:10.2135/cropsci1988.0011183X002800060023x

Epler, M., and S. Staggenborg. 2008. Soybean yield and yield component response to plant density in narrow row systems. Crop Manage. 7:1–13. doi:10.1094/CM-2008-0925-01-RS

Evans, L.T., and R.A. Fisher. 1999. Yield potential: Its definition, measurement, and significance. Crop Sci. 39:1544–1551. doi:10.2135/cropsci1999.3961544x

FAO. 2016. FAO statistics. FAO of the United Nations. http://www.fao.org/faostat/en/#data/QC (verified 15 June 2017).

Ferraro, D.O., D.E. Rivero, and C.M. Ghersa. 2009. An analysis of the factors that influence sugarcane yield in Northern Argentina using classification and regression trees. Field Crop. Res. 112:149–157. doi:10.1016/j.fcr.2009.02.014

Ferreira, A.S., A. Antonio, B. Junior, F. Werner, C. Zucareli, J.C. Franchini, and H. Debiasi. 2016. Plant density and mineral nitrogen fertilization influencing yield, yield components and concentration of oil and protein in soybean grains. Bragantia 75:362–370. doi:10.1590/1678-4499.479

Gaspar, A.P., and S.P. Conley. 2014. Responses of canopy reflectance, light interception, and soybean seed yield to replanting suboptimal stands. Crop Sci. 55:377–385. doi:10.2135/cropsci2014.03.0200

Gebbers, R., and V.I. Adamchuk. 2010. Precision agriculture and food security. Science 327:828–831. doi:10.1126/science.1183899

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. Bayesian data analysis. CRC Press, Boca Raton, FL.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge Univ. Press, New York.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The elements of statistical learning: Data mining, inference, and Prediction. Springer, New York. doi:10.1007/978-0-387-84858-7

Hörbe, T.A.N., T.J.C. Amado, A.O. Ferreira, and P.J. Alba. 2013. Optimization of corn plant population according to management zones in Southern Brazil. Precis. Agric. 14:450–465. doi:10.1007/s11119-013-9308-7

Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. J. Comput. Graph. Stat. 15:651–674. doi:10.1198/106186006X133933

Kasperbauer, M.J. 1987. Far-red light reflection from green leaves and effects on phytochrome-mediated assimilate partitioning under field conditions. Plant Physiol. 85:350–354. doi:10.1104/pp.85.2.350

Kaster, M., and J.R.B. Farias. 2012. Regionalização dos testes de valor de cultivo e uso e da indicação de cultivares de soja-Terceira aproximação (Regionalization of cultivation and use tests and soybean cultivars indication - Third Approach). Embrapa Soja Doc. No. 69. Embrapa Soja, Brazil.

Khosla, R., D. Inman, D.G. Westfall, R.M. Reich, M. Frasier, M. Mzuku, B. Koch, and A. Hornung. 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. 9:85–100.

Kyveryga, P.M., and T.M. Blackmer. 2014. Probability of profitable yield response to nitrification inhibitor used with liquid swine manure on corn. Precis. Agric. 15:133–146. doi:10.1007/s11119-013-9307-8

Kyveryga, P.M., P.C. Caragea, M.S. Kaiser, and T.M. Blackmer. 2013. Predicting risk from reducing nitrogen fertilization using hierarchical models and on-farm data. Agron. J. 105:85–94. doi:10.2134/agronj2012.0218

Lee, C.D., D.B. Egli, and D.M. TeKrony. 2008. Soybean response to plant population at early and late planting dates in the Mid-South. Agron. J. 100:971–976. doi:10.2134/agronj2007.0210

Lobell, D.B., J.I. Ortiz-Monasterio, G.P. Asner, R.L. Naylor, and W.P. Falcon. 2005. Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. Agron. J. 97:241–249.

Lowenberg-DeBoer, J. 1999. Economics of variable rate planting for corn. In: P.C. Robert, R.H. Rust, and W.E. Larson, editors, Precision agriculture. ASA, CSA, and SSSA, Madison, WI. p. 1643–1651.

Luca, M., and M. Hungría. 2014. Plant densities and modulation of symbiotic nitrogen fixation in soybean. Sci. Agric. 71:181–187. doi:10.1590/S0103-90162014000300002

Luca, M.J., M.A. Nogueira, and M. Hungria. 2014. Feasibility of lowering soybean planting density without compromising nitrogen fixation and yield. Agron. J. 106:2118–2124. doi:10.2134/agronj14.0234

McBratney, A., B. Whelan, T. Ancev, and J. Bouma. 2005. Future directions of precision agriculture. Precis. Agric. 6:7–23.

Meredith, M. and J. Kruschke. 2018. Bayesian estimation supersedes the t-Test. R package. R Development Core Team. https://cran.r-project.org/web/packages/BEST/vignettes/BEST.pdf.

Mourtzinis, S., G. Kaur, J.M. Orlowski, C.A. Shapiro, C.D. Lee, C. Wortmann, D. Holshouser, E.D. Nafziger, H. Kandel, J. Niekamp, W.J. Ross, J. Lofton, J. Vonk, K.L. Roozeboom, K.D. Thelen, L.E. Lindsey, M. Staton, S.L. Naeve, S.N. Casteel, W.J. Wiebold, and S.P. Conley. 2018. Soybean response to nitrogen application across the United States: A synthesis-analysis. Field Crop. Res. 215:74–82. doi:10.1016/j.fcr.2017.09.035

Mueller, T., P. Reeg, and P. Kyveryga. 2014. Increasing profitability in soybean production by optimizing planting rates. In: Integrated crop management conference. Iowa State Univ., Ames, IA. p. 59–63.

Norsworthy, J.K., and E.R. Shipe. 2005. Effect of row spacing and soybean genotype on mainstem and branch yield. Agron. J. 97:919–923. doi:10.2134/agronj2004.0271

Plummer, M. 2016. rjags: Bayesian graphical models using MCMC. R package. R Development Core Team. https://cran.r-project.org/web/packages/rjags/rjags.pdf.

Popp, M.P., J.T. Edwards, L.C. Purcell, and P.M. Manning. 2006. Profit-maximizing seeding rates and replanting thresholds for soybean: Maturity group interactions in the Mid-South. Agric. Syst. 91:211–228. doi:10.1016/j.agsy.2006.02.011

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. R Dev. Core Team 55:275–286. https://www.r-project.org/ (verified 20 May 2017).

Rahman, M.M., M.M. Hossain, M.P. Anwar, and A.S. Juraimi. 2011. Plant density influence on yield and nutritional quality of soybean seed. Asian J. Plant Sci. 10:125–132. doi:10.3923/ajps.2011.125.132

Rowntree, S.C., J.J. Suhre, N.H. Weidenbenner, E.W. Wilson, V.M. Davis, S.L. Naeve, S.N. Casteel, B.W. Diers, P.D. Esker, J.E. Specht, and S.P. Conley. 2013. Genetic gain × management interactions in soybean: I. Planting date. Crop Sci. 53:1128–1138. doi:10.2135/cropsci2012.03.0157

Schwalbert, R., T.J.C. Amado, T.A.N. Horbe, L.O. Stefanello, Y. Assefa, P.V.V. Prasad, C.W. Rice, and I.A. Ciampitti. 2018. Corn yield response models to plant density and N rates: Understanding data distribution. Agron. J. 110:1–13. doi:10.2134/agronj2017.07.0425

Shanahan, J.F., T.A. Doerge, J.J. Johnson, and M.F. Vigil. 2004. Feasibility of site-specific management of corn hybrids and plant densities in the great plains. Precis. Agric. 5:207–225. doi:10.1023/B:PRAG.0000032762.72510.10

Shearer, S.A., and S.K. Pitla. 2014. Precision planting and crop thinning. In: S.L. Young, and F.J. Pierce, editors, Automation: The future of weed control in cropping systems. Springer, the Netherlands. p. 99–124. doi:10.1007/978-94-007-7512-1_6

Sivarajan, S., M. Maharlooei, S.G. Bajwa, and J. Nowatzki. 2018. Impact of soil compaction due to wheel traffic on corn and soybean growth, development and yield. Soil Tillage Res. 175:234–243. doi:10.1016/j.still.2017.09.001

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. 108:637–646. doi:10.2134/agronj2015.0222

Thompson, N.M., J.A. Larson, D.M. Lambert, R.K. Roberts, A. Mengistu, N. Bellaloui, and E.R. Walker. 2015. Mid-south soybean yield and net return as affected by plant population and row spacing. Agron. J. 107:979–989. doi:10.2134/agronj14.0453

Tittonell, P., K. Shepherd, B. Vanlauwe, and K. Giller. 2008. Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya—An application of classification and regression tree analysis. Agric. Ecosyst. Environ. 123:137–150. doi:10.1016/j.agee.2007.05.005

van Ittersum, M.K., K.G. Cassman, P. Grassini, J. Wolf, P. Tittonell, and Z. Hochman. 2013. Yield gap analysis with local to global relevance-A review. Field Crop. Res. 143:4–17. doi:10.1016/j.fcr.2012.09.009

van Ittersum, M.K., and R. Rabbinge. 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crop. Res. 52:197–208. doi:10.1016/S0378-4290(97)00037-3

Vanlauwe, B., J. Mukalama, R. Abaidoo, and N. Sanginga. 2003. Soybean varieties, developed in West Africa, retain their promiscuity and dual- purpose nature under Western Kenyan conditions. African Crop Sci. Conf. Proc. 6: 58–68.

Van Roekel, R.J., and J.A. Coulter. 2011. Agronomic responses of corn to planting date and plant density. Agron. J. 103:1414. doi:10.2134/agronj2011.0071

Van Roekel, R.J., L.C. Purcell, and M. Salmerón. 2015. Physiological and management factors contributing to soybean potential yield. Field Crop. Res. 182:86–97. doi:10.1016/j.fcr.2015.05.018

Walker, E.R., A. Mengistu, N. Bellaloui, C.H. Koger, R.K. Roberts, and J.A. Larson. 2010. Plant population and row-spacing effects on maturity group III soybean. Agron. J. 102:821–826. doi:10.2134/agronj2009.0219

Weber, C.R., R.M. Shibles, and D.E. Byth. 1966. Effect of plant population and row spacing on soybean development and production. Agron. J. 58:99–102. doi:10.2134/agronj1966.00021962005800010034x

Werner, F., A.A.B. Junior, A.S. Ferreira, M.A.D.A. Silva, H. Debiasi, J.C. Franchini, L. Glycine, and L. Glycine. 2016. Soybean growth affected by seeding rate and mineral nitrogen. Rev. Bras. Eng. Agric. Ambient. 20:734–738. doi:10.1590/1807-1929/agriambi.v20n8p734-738

Zeileis, A., and T. Hothorn. 2015. partykit : A toolkit for recursive partytioning. J. Mach. Learn. Res. 16:3905–3909.