Received: 19 February 2024 | Accepted: 19 June 2024 | Published online: 11 September 2024

DOI: 10.1002/csc2.21336

Crop Science

ORIGINAL ARTICLE

Crop Breeding & Genetics

Leveraging soil mapping and machine learning to improve spatial adjustments in plant breeding trials

Matthew E. Carroll[1] | **Luis G. Riera**[2] | **Bradley A. Miller**[1] | **Philip M. Dixon**[3] | **Baskar Ganapathysubramanian**[2] | **Soumik Sarkar**[2] | **Asheesh K. Singh**[1]

[1]Department of Agronomy, Iowa State University, Ames, Iowa, USA

[2]Department of Mechanical Engineering, Iowa State University, Ames, Iowa, USA

[3]Department of Statistics, Iowa State University, Ames, Iowa, USA

Correspondence
Soumik Sarkar, Department of Mechanical Engineering, Iowa State Univdersity, Ames, IA, USA.
Email: soumiks@iastate.edu

Asheesh K. Singh, Department of Agronomy, Iowa State University, Ames, Iowa, USA. Email: singhak@iastate.edu

Assigned to Associate Editor Francisco Ernesto Gomez.

Funding information
North Central Soybean Research Program; Iowa Soybean Association; USDA CRIS project IOW04714; AI Institute for Resilient Agriculture, Grant/Award Number: USDA-NIFA #2021-67021-35329; COALESCE: COntext Aware LEarning for Sustainable CybEr-Agricultural Systems, Grant/Award Number: CPS Frontier # 1954556; Smart Integrated Farm Network for Rural Agricultural Communities (SIRAC), Grant/Award Number: NSF S&CC #1952045; RF Baker Center for

Abstract

Spatial adjustments are used to improve the estimate of plot seed yield across crops and geographies. Moving means (MM) and P-Spline are examples of spatial adjustment methods used in plant breeding trials to deal with field heterogeneity. Within the trial, spatial variability primarily comes from soil feature gradients, such as nutrients, but a study of the importance of various soil factors including nutrients is lacking. We analyzed plant breeding progeny row (PR) and preliminary yield trial (PYT) data of a public soybean breeding program across 3 years consisting of 43,545 plots. We compared several spatial adjustment methods: unadjusted (as a control), MM adjustment, P-spline adjustment, and a machine learning-based method called XGBoost. XGBoost modeled soil features at: (a) the local field scale for each generation and per year, and (b) all inclusive field scale spanning all generations and years. We report the usefulness of spatial adjustments at both PR and PYT stages of field testing and additionally provide ways to utilize interpretability insights of soil features in spatial adjustments. Our work shows that using soil features for spatial adjustments increased the relative efficiency by 81%, reduced the similarity of selection by 30%, and reduced the Moran's I from 0.13 to 0.01 on average across all experiments. These results empower breeders to further refine selection criteria to make more accurate selections and select for macro- and micro-nutrients stress tolerance.

Plain Language Summary

Plant breeding trials are a key component of crop improvement for yield, quality, and stress resistance. Breeding trials typically are grown on small plots of land and are highly affected by the area in the field where they are planted due to field trends. We

Abbreviations: ML, machine learning; MM, moving means; PYT, preliminary yield trial; PR, progeny row; XGBoost, extreme gradient boosting.

Matthew E. Carroll and Luis G. Riera, as co-first authors, made equal contributions to the paper.

Plant Breeding; Plant Sciences Institute, Iowa State University; NRT-DESE: P3 - Predictive Phenomics of Plants, Grant/Award Number: #1545453

investigated if using the soil features in a field could explain some of the variability in the early stages of a breeding program and used machine learning techniques to estimate the soil effects on observed yields. We found that by using the soil features for spatial adjustments, we could increase the accuracy of selections and improve the outcomes of decisions made by a breeder. This could have great impacts on increasing the accuracy of selection of early generation breeding trials, resulting in better lines being selected for yield, quality, and stress resistance traits, helping to make agricultural production more resilient and improve genetic gain.

1 | INTRODUCTION

Plant breeders make selection decisions within their programs to advance lines with the highest genetic value for the target population of environments (Cooper et al., 2021). Breeders must address non-uniform field conditions (i.e., field heterogeneity) in the selection decision-making process. Field trends did not account for bias selections toward the environmental value and not the genetic value of a line, resulting in an incorrect decision. These incorrect decisions cause economic strain on the program and limits success (D. P. Singh et al., 2021). Spatial adjustment methods have been proposed to alleviate the challenges of non-uniform field conditions. These methods allow breeders to make more appropriate comparisons of entries to each other, as well as to checks in yield plot testing. Spatial adjustments set up an effective and efficient selection process in the plot testing stages. Spatial adjustments are particularly applicable in unreplicated trials such as early-stage yield testing of hybrids in cross-pollinating crop species and for pure lines in progeny row (PR) and preliminary yield trial (PYT) stages in self-pollinating crops.

Reducing the size of the error associated with each genotype (i.e., pureline) gives breeders the ability to discriminate between yield levels of competing genotypes for more accurate selections, and therefore, several methods have been proposed. Augmented designs with replicated checks (Federer, 1961) give some form of local control within a field, giving the breeder a reasonable estimate of the trends that unreplicated genotypes may be experiencing. Augmented designs offer breeders the possibility of estimating the standard errors based on the replicated check performance. Check plot methods can be used by placing checks throughout the field so that a breeder can adjust for field trends based on known checks and their relative performance (Kempton, 1984). An advantage of the check plot method is that there is no limit to the number of lines used in the trial; on the other hand, it requires a large number of experimental plots to be allocated to the check plots (Kempton, 1984). The grid method was introduced for a mass selection experiment in maize [*Zea mays L.*] (Gardner, 1961), where small grids of 40 plants across the field were set up and the top-yielding 10% of plants in each grid were selected to minimize the spatial trends within the field. Another popular method is the moving means (MM) method (Richey, 1924), where the use of an MM coefficient to adjust yields resulted in a significant decrease in the error when estimating genotypic effects. More advanced spatial modeling techniques have been developed based on the two-dimensional auto-regressive model (Cullis & Gleeson, 1991) and expanded by Gilmour et al. (1997). These methods have been used widely in breeding programs, for example, to aid in multi-environment trial evaluations (Smith et al., 2001), multi-trait selections (De Faveri et al., 2017), and genomic selection (Oakey et al., 2016). Tensor product P-splines have also been proposed as an alternative to the two-dimensional auto-regressive model (Velazco et al., 2017; Verbyla et al., 2018). These methods are simple to run, and an open-source R package has been developed (Rodriguez-Alvarez et al., 2018) making the barrier to entry for a breeding program low. These spatial adjustment methods correct for field trends but do not directly quantify the soil features for each plot. Instead of directly quantifying soil variables and their effect on yield, the spatial adjustment methods either model the yield trends or reduce the area of the selection into smaller homogeneous blocks. These methods reduce the standard error of the difference between estimated genotypic values that are being evaluated and increase the precision in which yield differences can be compared within breeding trials (Qiao et al., 2000).

A common axiom in plant breeding trials is that spatial adjustments reduce the effects of soil heterogeneity, thereby facilitating an unbiased comparison of plot values. While there is a general acceptance of this notion, limited or no information is available that directly quantifies the effects of different soil characteristics and use these values to adjust plot yields. Soil is an integral feature of all breeding testing but is not typically utilized when addressing infield variability. While the literature is scant in this area, researchers have recently included soil features in the breeding selection model, by using principal components to reduce the dimensions of the soil features and directly add these features

into the linear model (Cursi et al., 2021) to aid in increased selection accuracy.

The utilization of soil features can provide an association between genotype, soil, and yield opening up avenues for improvement and utilization of spatial adjustments. Despite the use of soil characterization in production fields (Bakhsh, Jaynes et al., 2000; Bakhsh, Colvin et al., 2000; Dodd & Mallarino, 2005; McGrath et al., 2013; Mallarino et al., 1991; Sánchez T et al., 2011), they have not been utilized in breeding trials due to the resolution needed for soil maps, as well as the complexity of parsing out environmental and genotypic variance within a model. Furthermore, in breeding trials, the scale of experiments (no. of entries, locations, years, etc.) necessitates machine learning (ML) approaches to generate relationships and learn trends along with improved interpretability.

ML is a powerful tool in plant science research, with wide-ranging trait investigations such as biotic stress identification and quantification (Ghosal et al., 2018; Kar et al., 2023; Nagasubramanian et al., 2019, 2018, 2020, 2022; Rairdin et al., 2022; Singh et al., 2016, 2021, 2018; Tetila et al., 2019), organ detection (Falk, Jubery, O'Rourke et al., 2020; Ghosal et al., 2019; Jubery et al., 2021; Miao et al., 2020; Riera et al., 2021), microscopic objects (Akintayo et al., 2018; Fudickar et al., 2021), abiotic stresses (Chiranjeevi et al., 2021; Dobbels & Lorenz, 2019; Naik et al., 2017; Zhang et al., 2017), and crop yield (Parmley, Nagasubramanian et al., 2019; Shook, Gangopadhyay et al., 2021; Sagan et al., 2021). These ML methods provide an opportunity to integrate multiple variables for spatial adjustment while parsing out the trends and role of soil features on plot yield in tests. Among the extensive diversity of methods in ML, "Extreme Gradient Boosting" (XGBoost) (Chen & Guestrin, 2016) is well suited for the spatial adjustment problem since it is a hierarchical ensemble modeling tree structure that derives its outputs utilizing classification rules. XGBoost is scalable and gives accurate models, allows for the utilization of multiple features in prediction, and has been shown to be very useful in regression and classification problems (Chen & Guestrin, 2016; Herrero-Huerta et al., 2020). XGBoost, like other decision tree methods, is an algorithm that can effectively use multiple variables included in a dataset for prediction. Decision tree structures do not assume a linear relationship between variables, or the predicted values, which can help to predict nonlinear multivariate problems. This structure allows for variables to have multiple weights depending on the values of other variables. This is particularly applicable to soil plant growth interactions, as it is known that plant growth and development are dependent on multiple interacting soil features, weather, and genetics (Shook, Gangopadhyay et al., 2021).

The motivation for the study was to create a paradigm shift from the current spatial adjustments of yield plot data, primarily made on the plot yield values of neighbors, to a method that includes plot yield and soil features. To achieve our goals, digital soil maps and ML were leveraged to model the effect of soil features on plot yield in a soybean breeding program. XGBoost made it possible to adjust for the growing conditions experienced by each genotype. Furthermore, we present the use of Shapley values, which can be utilized by breeders to integrate soil features into the breeding decision-making process. Finally, we propose the usefulness of this method to consciously select for nutrient deficiency tolerance in field trials that experience abiotic stresses.

Core Ideas

- Spatial adjustments utilizing soil maps perform better than traditional methods for spatial adjustments of trials.
- Soil-based spatial adjustments can be used to better understand the spatial variability in breeding trials.
- Site-specific machine learning models for spatial adjustments perform better than large generalized models.

2 | MATERIALS AND METHODS

2.1 | Field experiments and data collection

Soybean [*Glycine max* (L.) Merr.] breeding trial plot data were collected in three growing seasons (2019, 2020, and 2021). Plot yield data were from PRs (coded as A-test; 2019, 2020, 2021) and PYTs (coded as B-test; 2019, 2020, 2021). Germplasm was sourced from crosses made within the breeding program; original material came from crossing promising Plant Introductions from the USDA collection (Oliveira et al., 2010), as well as shared germplasm from other public University breeding programs, received under material transfer agreements.

This study defines a field as a unique test, year, and location combination. A- and B-tests define the generation of the progeny that are being tested. A-test lines are F_5 lines derived from a single plant selection at the F_4 generation and were grown in progeny tests. The B-test consists of F_6 lines that are the selections from the A-test and are the first year of PYTs, which are grown in three to five locations in unreplicated tests. These two stages of testing were chosen because the number of yield trial plots in the program is the largest and, therefore, more impacted by spatial variation. Since these lines are grown in unreplicated tests, there is a need for accurate selections, which can be achieved by minimizing spatial variation. To minimize inter-plot competition, each field consisted of multiple experiments that grouped lines with similar crop maturity, phenology, and genetic background (see Figure S1 for a visual representation of experiments in a field). Experi-

ments were set up as single replication randomized complete block designs within each field. Selection decisions were made within each experiment within a field. A brief description of each field can be found in Table 1. Data were collected from the central Iowa research locations, centered around the Agricultural Engineering and Agronomy research farm (Boone, IA). No border rows were used between plots. The centers of plots in the same row were 1.52 m apart. Plot lengths are recorded as the harvested length of each plot; all fields had a 0.91-m alleyway between plots in each column giving a plot to plot center distance of 3.04 and 6.09 m for plots in the same column for the A and B tests, respectively. Plots were harvested with an ALMACO (Nevada, IA) small plot combine, and yields were adjusted to 13% moisture and converted to kilograms/hectare (kg/ha) for each harvested plot. Across 3 years, 238 experiments were examined across nine individual fields. Experiment sizes ranged from 50 to 370 plots with a mean of 183.0 plots per experiment. The A- and B-test plot number means were 209.7 and 124.9 plots, respectively. In these 238 experiments, 282 unique breeding populations were assessed. Breeding population sizes ranged from 1 to 263 genotypes with a mean size of 70.6. The average population size in the A and B test was 133.9 and 42.3, respectively. The average number of genetically distinct checks across all fields was ~10 and the average number of times each were replicated in the field was ~46.

2.2 | Soil data collection

Soil data were collected on a 25-m grid pattern for each field at a coring depth of 15cm and then sent to Midwest Labs (Omaha, NE) for analysis. Soil maps were created from the 25-m grid samples to a resolution of 3×3 m grids for soil nutrients and particle size fractions (Calcium [Ca], cation exchange capacity [CEC], phosphorous [P1], soil pH [PH], magnesium [MG], potassium [K], organic matter [OM], clay, silt, and sand), using the Cubist ML algorithm (Khaledian & Miller, 2020). Soil values were extracted for each plot using Esri's ArcGIS Python library (Redland, CA), arcpy, and Zonal Statistics toolbox. Plot areas defined were aligned with the plots described in Section 2.1. The extracted soil values for each plot were then used to calculate mean soil values on a per plot basis using a 3×3 grid of neighboring plots, which is represented by Equation (1), where $\overline{S}_{r,c}$ is the mean soil feature value of the r^{th} row and c^{th} column in each field, n is the number of neighboring plots used to calculate the mean soil feature, j is the row value of each neighboring plot used to calculate the mean, and k is the column value of each neighboring plot to calculate the mean. $\overline{S}_{r,c}$ values had higher correlation coefficients with individual plot yields than the raw plot-level soil values resulting in all further analysis using

TABLE 1 Listing of A-test (i.e., progeny row tests) and B-tests (i.e., preliminary yield trials) that were used in this study.

Year	Test	Location	Latitude	Longitude	No. of plots	Avg. no. of plots per check	No. of checks	Experiment count	Plot rows	Row spacing (m)	Plot length (m)	Plot area (m^2)	Planting density (100k seeds/hectare)
2019	A	Marsden	42.010N	93.780W	10,870	78	10	53	2	0.76	2.13	3.25	346
2019	B	Marsden	42.010N	93.778W	3062	31	10	21	2	0.76	5.18	7.90	346
2020	A	AEA	42.015N	93.768W	1458	26	7	8	2	0.76	2.13	3.25	346
2020	A	Burkey	42.012N	93.788W	3806	25	11	19	2	0.76	2.13	3.25	346
2020	A	Burkey_CH	42.015N	93.792W	2920	33	10	13	2	0.76	2.13	3.25	346
2020	B	CAD	42.047N	93.739W	3272	34	12	24	2	0.76	5.18	7.90	346
2021	A	Marsden_East	42.011N	93.778W	11,644	93	12	54	2	0.76	2.13	3.25	346
2021	A	Marsden_West	42.010N	93.781W	3479	60	5	16	2	0.76	2.13	3.25	346
2021	B	Burkey	42.015N	93.788W	3034	30	12	30	2	0.76	5.18	7.90	346

Note: Tests in each year, location, geographical location, and plots details are listed.

$\overline{S}_{r,c}$ values as input.

$$\overline{S}_{r,c} = 1/n \sum_{j=r-1}^{j=r+1} \sum_{k=c-1}^{k=c+1} S_{j,k} \quad (1)$$

2.3 | Baseline spatial adjustments

2.3.1 | Moving means

Plot seed yields were adjusted using MM with a grid size of 5×5 plots and were calculated using the R package "mvnGrad" (Technow, 2015). This method excluded the plot of interest when calculating the mean for each plot. MM for each experiment were calculated separately. The MM method makes an assumption that neighboring plots do not exhibit interplot competition and that the nonrandom error observed in a field is due to field environmental trends (Bos & Caligari, 2007). Breeders can reduce the interplot competition by subdividing fields into smaller experiments that are grouped by similar maturities and genetic backgrounds. The process for calculating MM can be found in the mvnGrad documentation (Technow, 2015). Results from the MM model from here on will be referred to as ADJ_5_MM. The first step is to calculate the regression coefficient of MM on the observed yield using Equation (2) (Technow, 2015):

$$P_{i,\text{obs}} = a + bx_i, \quad (2)$$

where x_i is the mean seed yield of the plots surrounding the i^{th} plot. $P_{i,\text{obs}}$ is the observed yield of the i^{th} plot. The adjusted yield is then calculated using Equation (3) (Technow, 2015):

$$P_{i,\text{adj}} = P_{i,\text{obs}} - \hat{b}(x_i - \bar{x}), \quad (3)$$

where $\hat{b}$ was calculated using Equation (2), $\bar{x}$ is the average of all of the moving means, and the adjusted phenotypic yield is $P_{i,\text{adj}}$.

2.3.2 | P-splines

P-splines are an alternative method to MM, commonly used in spatial adjustments for field trials. The R package "statgenSTA" (van Rossum et al., 2021) was used to correct for field variations using P-splines as described in Rodriguez-Alvarez et al. (2018). For this study, genotypes were treated as a random factor. The design was specified as a row-column design, which treats the row and column as random effects, and used the "SpATS" engine. Results from the P-Spline model from here on will be referred to as Spline. The P-spline method attempts to model both the large-scale observed spatial trends as well as the smaller deviations that can be missed when only row and column data are modeled directly. The large scale trends are modeled as polynomials, and the smaller trends are the smoothed trends, which model the deviations from the linear trends in Equation (4) (Rodriguez-Alvarez et al., 2018):

$$f(u,v) = 1_n\beta_0 + u\beta_1 + v\beta_2 + u * v\beta_3 + f_u(u) + f_v(v) + u * h_v(v) + v * h_u(u) + f_{u,v}(u,v), \quad (4)$$

where β_0 is the intercept, β_1 is the linear trend along the row, β_2 is the linear trend along the column, and β_3 is the linear interaction between the row and column. $f_u(u)$ is the smooth trend along the rows. $f_v(v)$ is the smooth trend along the columns, $u * h_v(v)$ is the linear row by smooth interaction along the columns, $v * h_u(u)$ is the linear column interaction by the smooth interaction along the rows, and $f_{u,v}(u,v)$ is the smooth by smooth interaction.

2.4 | Soil-based adjustments

2.4.1 | Statistical learning

XGBoost training process requires tuning parameter coefficients, usually denoted by θ. Among all the parameters, the number of estimators, maximum depth, learning rate, subsample size, and minimum child weights are commonly tuned using a greedy search algorithm. Most XGBoost models are configured using a relatively shallow depth or a small number of trees because of their sequential characteristic, where each new tree corrects the errors made by previous trees, quickly reaching a point of diminishing returns. Models were trained using the tuning parameters as listed in Table S1, with a regression with squared loss set as the objective loss function. Tuning parameters for the XGBoost algorithm are previously defined in detail (Chen & Guestrin, 2016). This training structure works well for soil-based spatial adjustments, because it is able to train models quickly and model complex data with multiple interactions. Our model used 240 estimators and a depth of 9 for training. Model parameter tuning was done using *Sklearn grid search API* (Buitinck et al., 2013) with 10-fold cross-validation. Two methods were compared for spatial adjustments using XGBoost. Model 1, from here on referred to as XGB_Global, combined data from all nine fields and attempted to make a generalized model. Model 2, from here on referred to as XGB_Local, was trained on each field individually to create a more specific model for each field location. Models were also assessed using Lin's concordance correlation coefficient, which is a measure of repeatability that is more robust than Pearson's correlation coefficient (Lin, 1989).

2.5 | Model overview

Plant breeders routinely express yield as $P = G + E$, where P is the observed phenotypic value of a line, for example, seed yield. G is the genotypic value of the line, which can be estimated when lines are replicated and grown in multiple locations (Tabery, 2008). Last, E is the environmental effect that breeders try to minimize, as it can confound the estimated genetic value of a line. Historically, E has been a macro-scale environment effect such as a single field, and micro-scale plot level environmental effects have been reduced using various spatial adjustment methods (Bernardo, 2002). Furthermore, genotypic value can be better estimated by using phenotypic values of related lines. If the mean of the population is included, the new equation is represented as $P = G + E + \mu$, where μ is the phenotypic mean of the population. In this context, we define the population to consist of pure lines from the same breeding family; that is, they have the same pedigree. Our proposed method uses soil features to estimate the value of E for each plot and to increase the correlation between the adjusted phenotypic values and the estimated genotypic values.

The first step was to collect the yield data for each plot in a given field, which was in a row and column configuration within each field. Each experiment has multiple lines that are from the same or related populations, and the mean of each population within each experiment was calculated and used as the estimate for the population mean within the experiment. Population statistics for each field can be found in Table S2. Equation (5) represents the basis for the soil-based adjustment methods. $P_{r,c,j,k}$ is the seed yield of the plot in the r^{th} row of the c^{th} column from the j^{th} population in the k^{th} experiment. $\mu_{j,k}$ is the mean of all plots in the j^{th} population of the k^{th} experiment. $\delta_{r,c}$ is the deviation from the population mean of each plot. Checks are unique in this study because they are the only lines that are replicated throughout each field. They do not share known familial relationships with each other, as they do not have a known common pedigree. $\mu_{j,k}$ for check varieties was calculated by taking the mean seed yield from each plot of that variety within a field.

$$P_{r,c,j,k} = \mu_{j,k} + \delta_{r,c} \tag{5}$$

$\delta_{r,c}$, also referred to as the ground truth, is partially a function of the soil feature values $S_{r,c}$. $S_{r,c}$ contains all of the soil feature values, described in Section 2.2 (Ca, CEC, P1, PH, MG, K, OM, clay, silt, and sand). XGBoost regression was used to predict the deviation, $\hat{\delta_{r,c}}$, from the population mean based on the soil features of each plot. This model is represented by Equation (6):

$$\hat{\delta_{r,c}} = f(S_{r,c}). \tag{6}$$

The genotypic value is not modeled in Equation (6) because for this work the focus is on approximating the microenvironment effect of each plot (E). The observed deviation with the model data inputs should not capture the genotypic effect of individual lines. When the genotypes are not replicated, a better approximation of the genetic value of the lines should be achieved by attributing some of the unexplained deviation from the population mean to the environmental conditions that the genotype is experiencing.

The last step estimates the adjusted phenotypic value of each plot by taking the observed phenotypic value and subtracting the estimated effect of the soil, $\hat{\delta}_{r,c}$, from the observed phenotypic value, $P_{r,c,j,k}$, which is represented as $P^{\text{adj}}_{r,c,j,k}$ in Equation (7). $P^{\text{adj}}_{r,c,j,k}$ was used for further analysis and comparison to other spatial adjustment techniques.

$$P^{\text{adj}}_{r,c,j,k} = P_{r,c,j,k} - \hat{\delta_{r,c}} \tag{7}$$

2.6 | Model interpretability

"SHAP" (SHapley Additive exPlanations) (Lundberg & Lee, 2017) was used to assist interpreting the results from XGBoost models. SHAP uses traditional Shapley values from game theory, and it has shown to be very successful in explaining the output of any ML model (Shrikumar et al., 2017; Štrumbelj & Kononenko, 2014). Shapley values estimate the features' importance in terms of their contribution to the outcome.

2.7 | Model performance for selection

Three evaluation metrics were used to assess the performance of the different methods. The first metric was the relative efficiency using the average standard error of the difference (SED) between the check lines in each field. The efficiency of different spatial adjustment methods was calculated using Equation (8), where $\text{SED}_{\text{Yield}}$ is the SED using no spatial adjustments, and SED_{Adj} is the SED using one of the tested spatial adjustment methods:

$$\text{RE}_{\text{SED}} = \frac{100 \times \text{SED}_{\text{Yield}}}{\text{SED}_{\text{Adj}}} \tag{8}$$

In the A-test, entries are not replicated due to seed limitation and in the B-test, breeders prefer to test in more than one location in an unreplicated manner; therefore, checks were used to estimate the standard error of difference. Previous works evaluating the efficiency of spatial and experimental designs in plant breeding trials have used this as a method for the evaluation of different selection methods (Gilmour et al., 1997; Magnussen, 1990; Qiao et al., 2000).

As the standard error of the difference decreases, more of the observed phenotypic yield variance can be attributed to the genotypic value.

The second metric used to evaluate the models was a similarity coefficient called the Czekanowski (CZ) coefficient, also known as the Sorensen Dice coefficient. This coefficient gives the percent of lines selected by both selection methods at a given selection intensity. The Czekanowski coefficient was calculated according to Qiao et al. (2000), where a is the number of lines selected by both selection methods, b is the number of lines selected only by model 1, and c is the number of lines selected only by model 2. CZ is a ratio given by Equation (9). If CZ has a value of 1, it would mean that both methods selected the same lines; if CZ has a value of 0, it would mean that there was no overlap in the lines that were selected. For our purposes, model 1 used no spatial adjustment, and model 2 used the adjusted phenotypic values from one of the listed spatial adjustment methods:

$$\mathrm{CZ} = \frac{2a}{2a + b + c}. \tag{9}$$

The last metric is the Moran's I statistic, a measure of the spatial autocorrelation present within a user-defined grid (Bivand & Wong, 2018). Our implementation uses a grid of 5×5 plots to define the plot neighbors given the row column field design of each test. Neighbor distances used relative coordinates to determine plot neighbors, and plot widths and lengths were considered to have a unit of 1. Plots that were within a distance of $2\sqrt{2}$ of a plot were considered neighbors. All other plots were not considered to be neighbors. Moran's I interpretation is similar to a regular correlation coefficient with values ranging from approximately -1 to 1. A value of 1 is a perfect spatial autocorrelation, and a value of -1 is a negative spatial autocorrelation. The null hypothesis expected values for the Moran's I statistic are close to zero but are slightly negative. p-Values are also reported with each statistic to assist in the interpretation of the significance of the correlation. Moran's I statistics were calculated in R version 4.0.3 with the package "spdep" (Bivand et al., 2015) and the Moran.test function.

3 | RESULTS

3.1 | Genetic variation

Population mean seed yields ranged from 74.0 to 6032.3 kg/ha with a mean yield of 4250.2 kg/ha. The average population seed yields for the A and B tests were 4391.4 and 3624.8 kg/ha, respectively. While B-tests have one additional generation of selection in the previous year, the lower seed yield is likely due to the shorter plot size in the A-test (2.3 m) compared to the B-test (5.18 m) giving a larger plot edge effect (Stelling et al., 1990). Table S2 gives a more detailed summary of the yield data across all nine fields.

3.2 | Soil variability

Soil core samples were taken for each of the nine fields with 10 soil features measured (see the *Materials and Methods* section). Organic matter ranged from 1.1% to 7.4% with a mean of 3.9%. PH ranged from 3.9 to 8.5 with a mean of 6.1. Clay had a range from 0.1% to 33.9% with a mean of 27.4%. Sand had a range from 12.2% to 42.7% with a mean of 30.4%. Silt had a range from 18.9% to 50.5% with a mean of 41.6%. Calcium parts per million (ppm) ranged from 507.1 to 6360.5 with a mean of 2774.9. CEC (meq/100 g soil) ranged from 9.8 to 36.5 with a mean of 21.1. Potassium (ppm) had a range from 85.4 to 286.1 with a mean of 169.1. Magnesium (ppm) ranged from 107.3 to 768.9 with a mean of 399.7. Phosphorous (ppm) had a range from 5.5 to 73.5 with a mean of 21.7. Table S3 shows the range of each soil parameter for each field to better represent the variation that was present across locations. It should be noted that due to a small-sized 2020_A_AEA field test, there was a lack of variability in the samples taken in the clay content, and this value was constant for that test.

3.3 | XGBoost and model interpretability

XGB_Local model's predictions, $\hat{\delta_{r,c}}$, compared to $\delta_{r,c}$, and the concordance correlation coefficient are shown in Figure 1. These predictions are on the centered variables, as described in Equation (6). The Concordance Correlation Coefficient between the ground truth and predicted values ranged from 0.71 to 0.86. Centered values ranged from -81.5 to 48.8. When looking at the two stages of the breeding program, that is, PR A-test and PYT B-test, $\delta_{r,c}$ values in A-tests ranged from -81.5 to 48.8 and in B-tests ranged from -47.7 to 32.5. This demonstrates that the more advanced stage of testing has less phenotypic variability. Predicted values, which were obtained from XGB_Local model and reflect estimated effects of soil features on seed yield, ranged from -62.7 to 36.3. These values represent $S'_{r,c}$ as defined in Equation (6). The A-test predicted values ranged from -62.7 to 36.3 and the B-test predicted values ranged from -40.2 to 21.6.

Figure 2 shows results as an example for the 2021_A_Marsden_East test. 2021_A_Marsden_East test was chosen because it contained the most plots of all tests, but the figures for the other tests can be found in Figures S2–S4. In feature selection, correlation provides a measure of independence between input variables; the higher the correlation, the greater the linear dependency. The heatmap plot in Figure 2a shows the correlation between the features, and the dendrogram on top provides a clustering

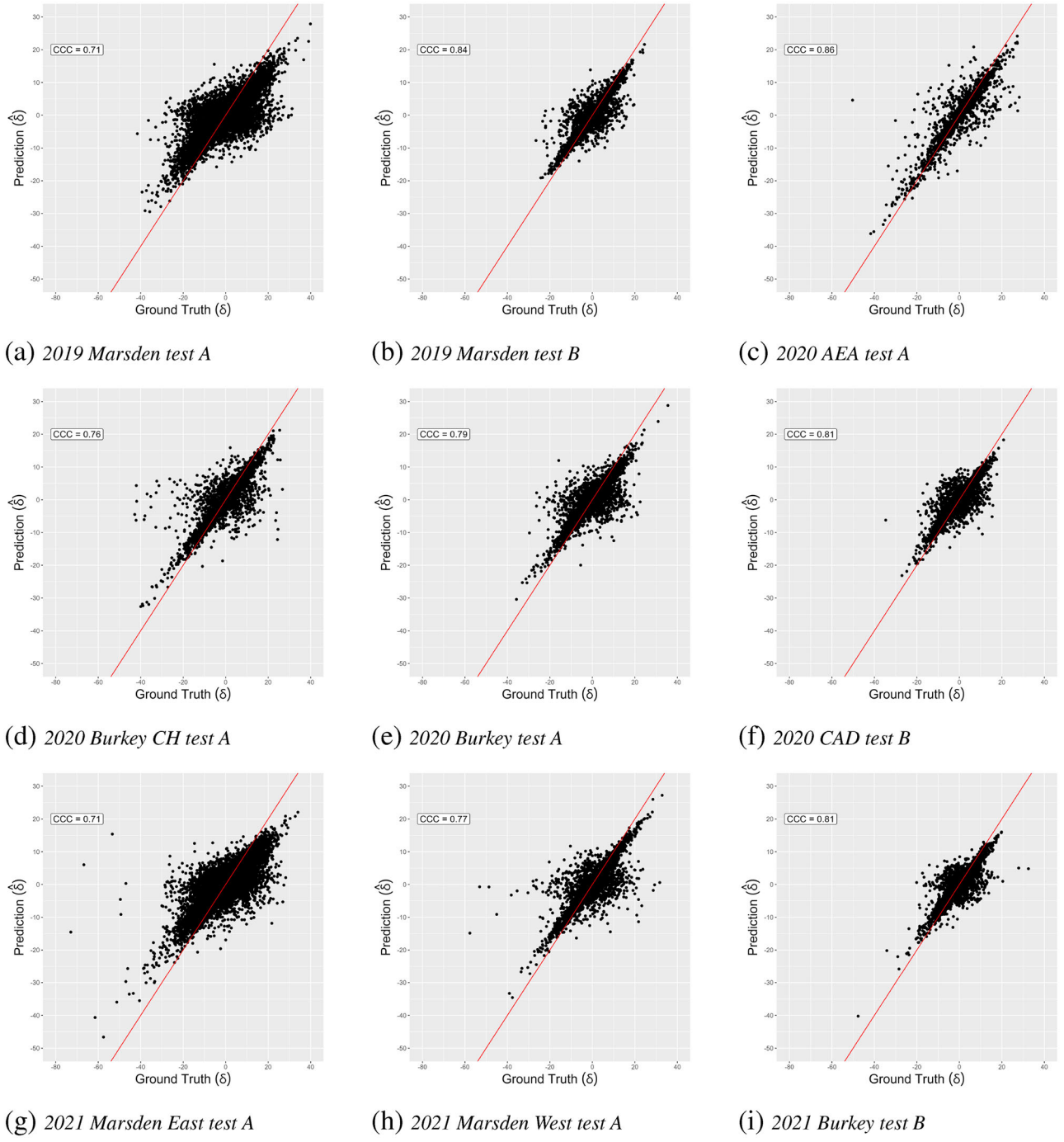
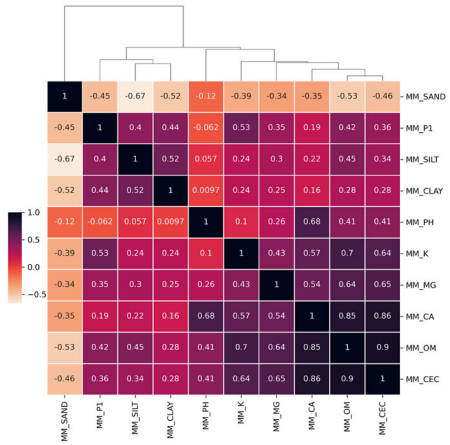


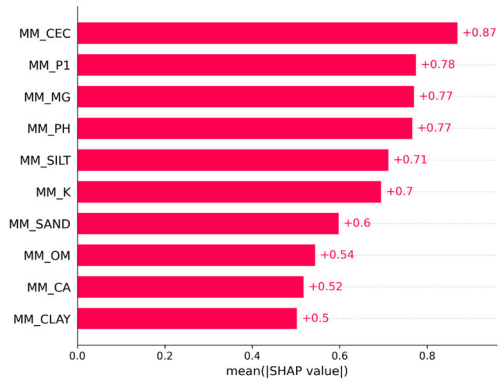
FIGURE 1 This figure shows the XGB_Local Ground truth (δ) versus prediction (δ') scatter plots with their respective concordance correlation coefficient. These charts show how well the model learned the ground truth data distribution. The red line shows a 1:1 relationship between predicted and observed values.

mechanism of these features according to their correlation values. The inverse "u" shape-like links group elements into clusters; the lower the "u" shape is in the figure, the higher the correlation of those elements is in relation to the others. The MM_OM and MM_CEC correlation is 0.9, the highest between elements, and is clustered in the dendrogram first. Then, the MM_OM MM_CEC cluster is merged with MM_CA to form another cluster. All features were kept to explore their outcomes using Shapley values. As described in Equation (1), the soil features used for analysis were a result of averaging neighboring plot soil feature values, variables with a "MM" preceding the soil feature indicating the moving mean average.

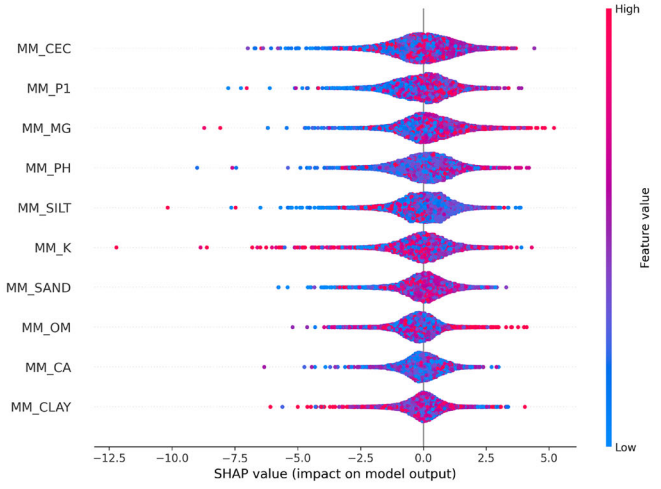
Figure 2b provides a summary of the features of importance for the model. The higher the value on the bars in this plot, the greater the information gained by the model when performing



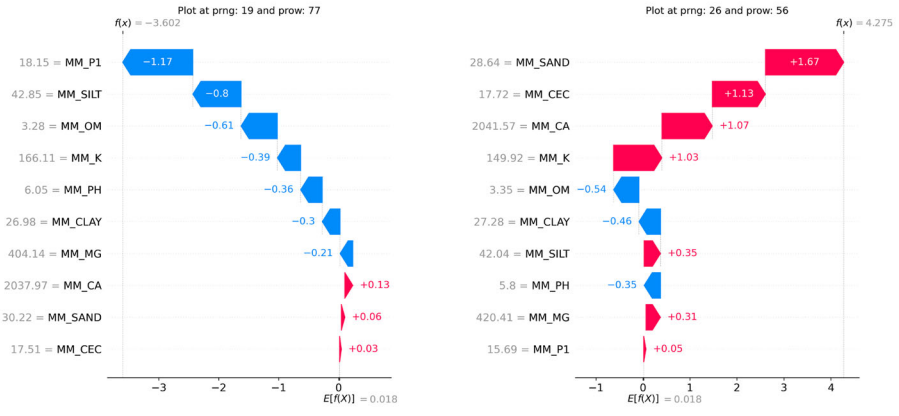
(a) *Heatmap*



(b) *Feature importance*



(c) *Shapley Summary plot*



(d) *Shapley values waterfall plot*

FIGURE 2 XGBoost (Extreme Gradient Boosting) model explainability plots using data from 2021 Marsden East Test A. The heatmap (a) shows the correlation between the features used to train the XGBoost model and their respective clustering hierarchy by the dendrogram. The feature importance (b) illustrates a measure of the prediction power of each feature for the model with this dataset. Each dot in the Shapley value Summary plot (c) represents a single plot. Its colors are normalized per feature using its maximum and minimum, respectively. The summary plot helps illustrate the relationship between the individual feature value range and their impact over the predicted yield. The Shapley values waterfall plots (d) provide an insight into how feature values contribute to the final expected value in two plots (Identities "IA2102") at different locations in the field where the feature values are substantially different. Ca, calcium; CEC, cation exchange capacity; P1, phosphorous; PH, soil pH; MG, magnesium; K, potassium; OM, organic matter.

the tree branch split during training. In the 2021 Marsden A test East, MM_CEC provided the most significant information gain, and the other top four driving prediction features were MM_P1, MM_MG, and MM_PH.

To aid in interpreting results, we investigated the use of Shapley values. Figure 2c describes how each plot reacted to the input features. Each dot is a plot, and its color represents the plot feature value relative to the feature's range scale, where red is high and blue is low. The accumulation of the dots across the *x*-axis characterizes the global behavior of the features, similar to a histogram. For the 2021 Marsden A test East, we see that high values of MM_CEC, MM_P1, and MM_MG tend to have positive Shapley values. However, it is not conclusive for the other soil features, where we see high feature values on both tails. For our model, we cannot draw any causal inference conclusions because we suspect the existence of many confounding factors.

The waterfall plots in Figure 2d describe the reaction of the the check genotype (IA2102) compared to the soil composition in distinct locations of the field used on the 2021 Marsden A test East. A check variety was used so that the comparison is only looking at the soil features and not including additional genetic differences between two lines. The model's mean output prediction for the entire field is 0.018. If we start from the mean predicted value of 0.018 and add the respective Shapley values, we obtain the model prediction of 4.275 for the right image, which is defined as the plot in column 56 and row 26. In the left image, which is defined as the plot in column 77 and row 19, we obtain a model prediction of -3.602. These plots were shown to illustrate how Shapley values could be used to understand model weights for soil features in specific plots. The model predicted values represent the relative quality of the growing conditions for each individual plot and the expected contribution of each soil feature, as determined by the model.

3.4 | Spatial adjustments

Mean relative efficiencies were calculated for each model, with mean values of 115.5%, 134.7%, 168.1%, and 181.0% across all fields for ADJ_5_MM, XGB_Global, Spline, and XGB_Local. Figure 3 shows the distribution of the SED across all locations and methods. No spatial adjustments was the base model to which all models were compared. The models are ranked from lowest to highest relative efficiency, where the MM method has the lowest relative efficiency and XGB_Local has the highest average relative efficiency across all nine fields that were tested. All of the models presented have a higher relative efficiency than the unadjusted yields. This indicates that XGBoost using plot-level soil fertility variables can be used to increase the precision of the genotypic estimates of the checks. These results infer that the estimated genotypic values for unreplicated lines will be more precise as well. Increasing the precision of the estimates of the genetic values of a line will lead to higher selection accuracy and increases in genetic gain.

Similarity coefficients were calculated at three selection intensities (0.1, 0.3, and 0.4) and the differences in the selected lines examined based on the method used for spatial adjustments. All models were compared to the unadjusted yield. Figure 4 shows the variation in the CZ value across different selection intensities and the methods of spatial adjustment. The *x*-axis is the adjustment method, and the *y*-axis is the CZ coefficient. The median CZ across all three selection intensities was 0.82, 0.74, 0.88, and 0.70 for MM, Spline, XGB_Global, and XGB_Local, respectively. To put this in perspective, 18% of the lines that were selected by the MM method were dissimilar from no-adjustment while 30% of the lines that were selected by the XGB_Local model were dissimilar from no-adjustment. These results show that utilizing spatial adjustments results in differences in the lines that are selected for advancement within a breeding program. The lowest median CZ value observed across all selection intensities was the XGB_Local model, meaning that the least amount of overlap between selected lines with no spatial adjustment would be achieved by using this method. However, it should be clarified that spatial adjustments are only needed when there is spatial autocorrelation in yield present within the field. In the absence of yield spatial autocorrelation, the use of unadjusted plot mean in statistical analysis will give the same outcome as spatial adjusted means.

Within breeding trials where lines have similar genetic backgrounds and maturities, it can be expected that there will be similar performance among lines, and because the lines within a trial are randomized it is highly unlikely that there should be evidence of spatial autocorrelation if the environmental effect has been accounted for. The reduction in the spatial trends observed in trials is a good indication that spatial adjustments were effective at removing non-random field effects. Figure 5 shows an example of an experiment before and after the yield has been adjusted using the XGB_Local method. The Moran's I is reduced from 0.58 to 0.09

Figure 6 shows the range of Moran's I values across all 238 experiments that were tested. The first two methods, ADJ_5_MM and Spline, are used by breeders for selection decision making. The next two methods are variations of the XGBoost method that used soil data instead of neighbor-based adjustments. The last column labeled as Yield is the Moran's I values calculated based on no spatial adjustment and serves as the baseline for all of the comparisons that are made. All methods reduce the spatial autocorrelation within trials, with median Moran's I values of -0.04, -0.03, 0.04, 0.01, and 0.13 for MM, Spline, XGB_Global, XGB_Local, and Yield spatial adjustment methods, respectively. Figure 6 shows the differing levels of spatial correlation across fields

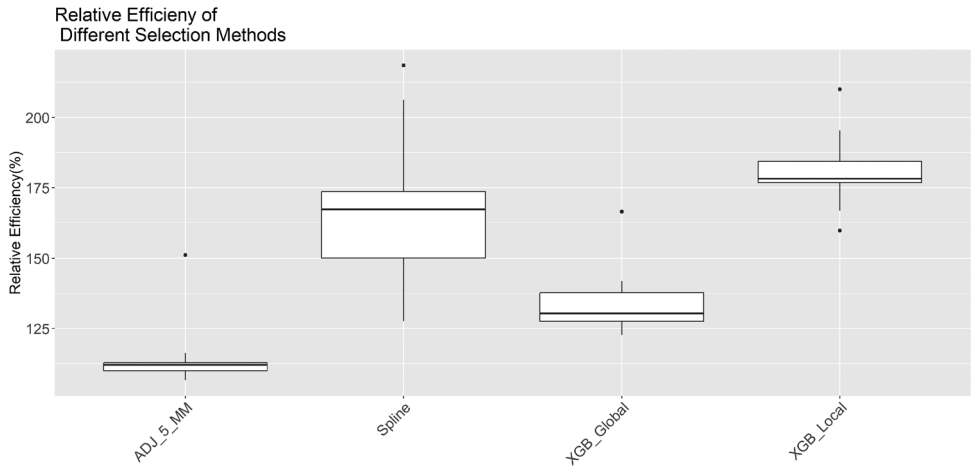


FIGURE 3 Standard error of the difference (SED) box plots for the different spatial adjustment methods showing improved relative efficiency. The XGB_Global model exhibited the highest SED.

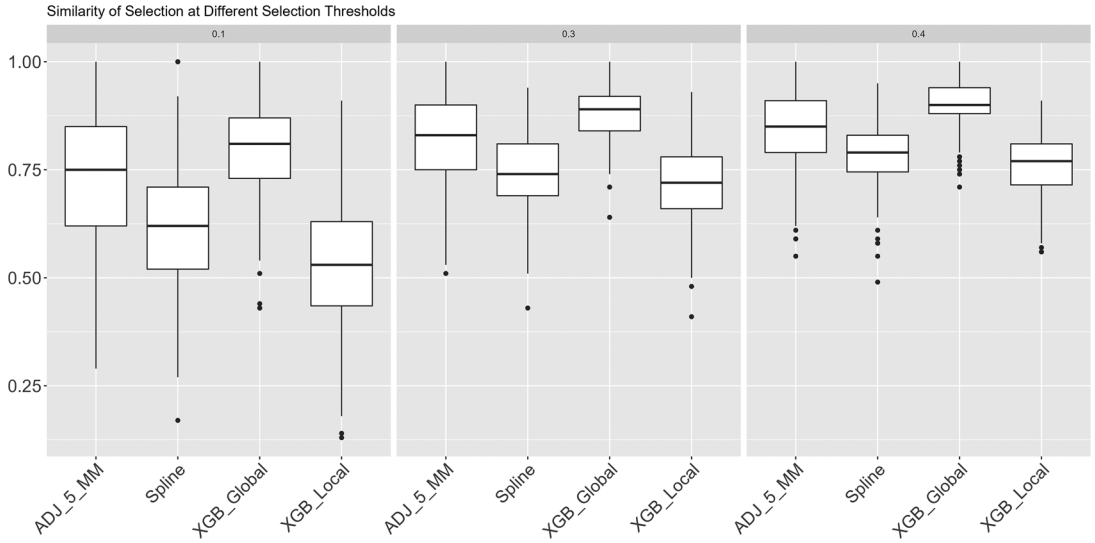


FIGURE 4 Czekanowski (CZ) box plots for the different methods. This graph shows the similarity in which lines would be selected between the different models using selection intensities of 0.1, 0.3, and 0.4, in comparison to no spatial adjustment. The XGB_Global model exhibited the highest CZ coefficient while XGB_Local the lowest. The higher the CZ is, the greater the selection overlapping between the models and the selections made based on no spatial adjustment.

and provides strong evidence that both traditional neighbor-based adjustments and soil-based adjustments reduced the overall Moran's I statistic in the majority of trials. It should be noted that the XGB_Local method had a median Moran's I statistic value (0.01) that was closest to 0 compared to all other methods. All methods were statistically different at an alpha of 0.05 based on the least significant difference (LSD).

4 | DISCUSSION

Directly modeling the effect of soil on a plot level basis has not been done for plant breeding field trials. The reason for this is that the labor and cost required to collect these types of data are cost prohibitive. As newer technologies arise from the combination of remote sensing and ML, this type of data can be obtained at a much lower cost than previously possible (Ferhatoglu & Miller, 2022; Khaledian & Miller, 2020; Minasny et al., 2018; Minasny & McBratney, 2016). Traditional spatial adjustments in field trials have focused on reducing the effect of environmental trends that are evident in most large scale breeding trials. Removing or accounting for the environmental effect of a plot will lead to more accurate selection and an increase in genetic gain for each breeding program (Hazel & Lush, 1942; Krause et al., 2023; Rutkoski, 2019). Soil-based adjustments give breeders

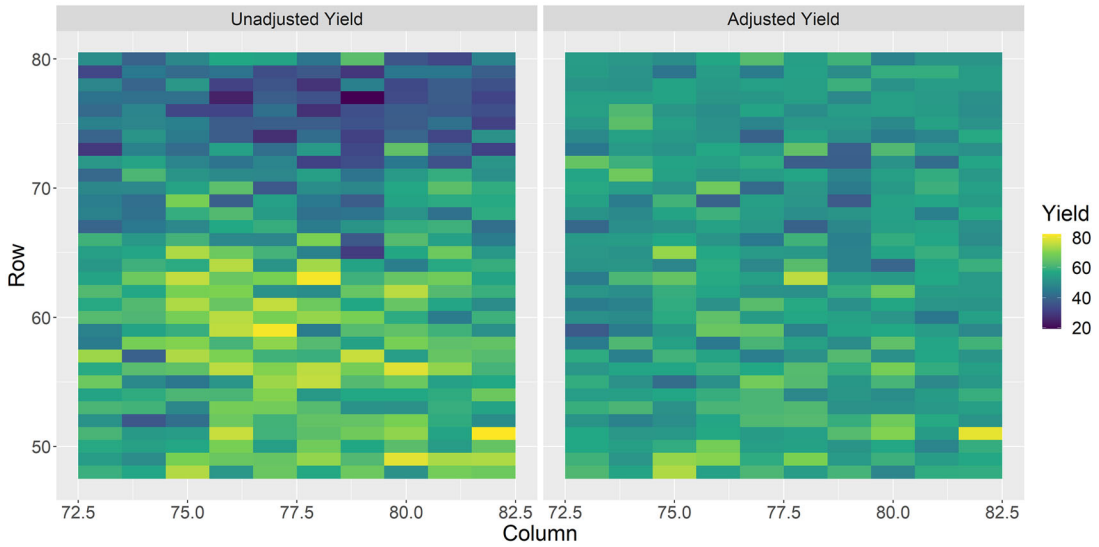


FIGURE 5 This figure shows the observed yield for both the unadjusted and adjusted (XGB_Local) yield in an experiment. The Moran's I for the left image is 0.58 and 0.09 for the right.

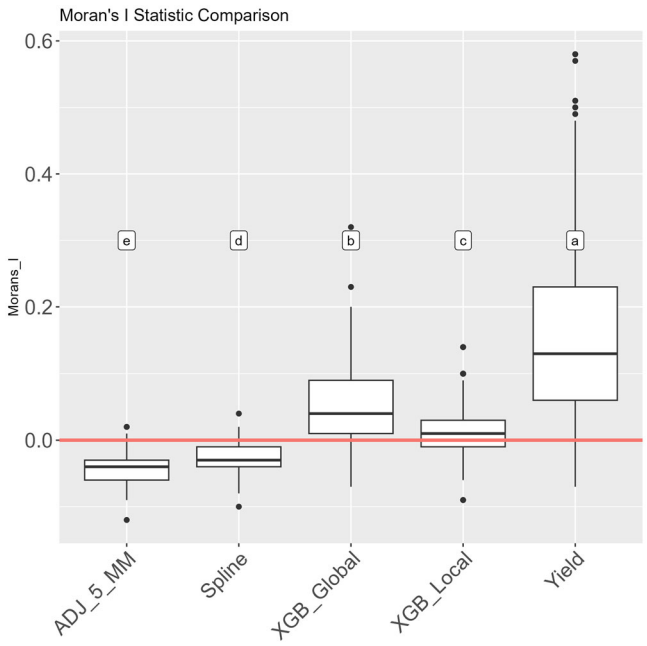


FIGURE 6 This figure shows the variation of the Moran's I statistic that was calculated for each experiment. The closer the value is to zero the better evidence there is to suggest that there is a minimal field trend. Letters above each boxplot shows the significant groups differences based on the least significant difference (LSD) using an alpha of 0.05.

additional tools when optimizing for selection of multiple traits (Akdemir et al., 2019), such as yield, and biotic stress-induced nutrient deficiencies.

Soil fertility recommendations for soybean production in the Midwest have primarily been focused on P, K, and pH. Optimal recommended P level is to target ≥ 16 ppm, optimal K levels are recommended to be ≥ 161 ppm (Mallarino & Sawyer, 2013), and optimal pH range is between 6.5 and 7.5 (McGrath et al., 2013). Based on these recommendations, three of the nine fields had an average P level that was below optimum, seven fields had a mean soil pH outside of the optimal zone, and four fields had a mean K level below the optimum. It should be noted that these recommendations are not based on maximizing yield but based on maximizing profitability. Other factors such as CEC have been reported to have an effect on soil nutrient availability and its interactions with plant growth and development (Brooker et al., 2017). These soil nutrients have well-established univariate effects on soybean growth and development (Dodd & Mallarino, 2005; Mallarino et al., 1991), but when they are combined within multiple interacting complex systems, it becomes difficult to quantify the effect of any individual factor. Previous studies have looked at the effects of multiple soil factors and their effects on predicting yield using Random Forest in production fields and found that some of the top-ranking predictor variables were P, K, OM, and pH (Smidt et al., 2016). Our findings show that these factors had an effect on the quality of the growing conditions of each plot. While feature importance measurements are useful for determining how often a variable is used to split a decision tree, it does not provide any additional insight into the effect that a variable has on a modeled outcome.

The three metrics we used to evaluate the different selection methods examine the utility of different spatial adjustment methods. Neighbor-based adjustments make an assumption that neighboring genotypes should be phenologically similar to reduce the effects of interplot competition (Bos & Caligari, 2007). Neighbor-based adjustments are based on linear models, which follow the assumption that the samples come from

the same distribution. If this assumption is no longer valid, it can result in poor performance of spatial adjustment methods. This is why experiments usually consist of similar maturity groups and genetic backgrounds, which helps to reduce interplot competition as genotypes invariably have similar phenological/architectural traits. The XGB_Local model had the best metrics for reducing the spatial autocorrelation for seed yield within tests, the greatest increase in the mean relative efficiency, and also had the lowest median CZ coefficient across all tests and selection thresholds. For caution, it is important to clarify that dissimilar selection decisions stemming from unadjusted or any of the spatial adjustment methods require further replicated field trials in the following year to confirm the accuracy of the selection of different methods. Since we used an active breeding pipeline for this research, which used MM in the A-test in the previous year, we do not have a completely independent set to compare the efficiency of selection across methods. The CZ coefficient serves as a way to showcase that using this method results in different varieties being selected and advanced. Using a more complex methodology for selection decision-making will not be adopted by breeders if the outcomes are the same or highly similar to a simpler model. The advantage of using soil features with XGBoost is that the model is based on the feature space of each field whereas neighbor-based adjustments utilize only the location space of each dataset. Utilizing the feature space becomes more important when breeders want to grow dissimilar genotypes in the same test. This happens because the nearest neighbor methods have a proclivity for better estimates for similar phenotypes. The other advantage of XGBoost is the interpretability of the adjustments giving increased confidence in breeders' decision-making.

The overall results suggest that using a locally trained model using the proposed methodology can increase the efficiency of selection within the breeding program. Interestingly, the local model consistently performed better than a global model for applications of spatial adjustments within a breeding program. We hypothesize that this is due to the variation in environments and that the local models can pick up on environment-specific factors. While the global model was generalized across fields, the non-generalizable models were effective because they were able to determine the local field effects and interactions. The global model became a statistical average of all models and removed the sensitivity that was obtained in the local model that was refined for each individual field.

There are several avenues for the integration of ML-based spatial adjustment models in other on-going areas of research in crop science. For example, crop modeling has been shown to be an effective tool for parsing out the effects of genotype, environment, and management for crop production (Ojeda et al., 2022). The fusion of better environmental data and increased modeling capabilities, along with high throughput phenotyping (HTP) methods for genotype specific calibrations, can be used by breeders to make selections for both current and future breeding environments (Cooper et al., 2021; Chattopadhyay et al., 2023; Gupta & Singh, 2023; Krause et al., 2022). High-resolution soil maps and spatial adjustments can be invaluable for these modeling projects, and it has been shown that spatial adjustments can be useful for increasing genomic prediction model accuracies (Lado et al., 2013; Mao et al., 2020). Comprehensive weather and genetics data have been used to predict complex phenotypic traits (Li et al., 2018; Shook, Gangopadhyay et al., 2021; Shook, Lourenco et al., 2021) and integration of soil maps and interpretable spatial adjustments can improve the explainability of genotype response. These advantages extend to the study of component traits in seed yield and other physiological end-trait predictions (Chiozza et al., 2021; Moreira et al., 2019; Riera et al., 2021; Xavier et al., 2017), including ML-based predictions and prescriptive breeding approaches (Parmley, Higgins et al., 2019; Parmley, Nagasubramanian et al., 2019). The integration of explainability features can also lead to an enhanced understanding and usage of 3D based approaches for yield improvement (Young et al., 2023).

Soil mapping and interpretable spatial adjustments could be useful in studying abiotic and biotic stress responses, particularly in conjunction with HTP and ML (Guo et al., 2021; Herr et al., 2023; Reynolds et al., 2020; A. K. Singh et al., 2021). Using digital soil maps to test varieties in the field may allow breeders to have the benefit of exploiting the levels of nutrient stress encountered by each variety (Byron & Lambert, 1983). For example, in a large test, breeders can utilize the soil map and Shapley values to demarcate regions that have a specific nutrient stress. This essentially leads to selection in this subregion for nutrient stress response, providing an opportunistic selection outcome. Similarly, breeders can select for disease tolerance in conjunction with soil features. Additionally, it enhances the value of field testing of each plot because we can better and strategically utilize soil variables without setting up large specialized nurseries. However, it is paramount that we realize that opportunistic selection does not change the value of specialized nurseries, which are essential for breeding for stress tolerance (D. P. Singh et al., 2021). Furthermore, advances in root trait studies (Carley et al., 2022; Falk, Jubery, O'Rourke et al., 2020; Falk, Jubery, Mirnezami et al., 2020; Jubery et al., 2021; Lynch & Brown, 2001) can be complemented with the use of soil maps and spatial adjustments for a more holistic interpretation of plant response accounting for both above- and below-ground traits.

There are four main avenues to further improve this work. (a) The setup of this experiment does not allow us to make causal inferences about the true effect of various soil parameters, but it does open the door for future work where researchers can start to investigate nutrient use efficiency for different genotypes (Baligar et al., 2001), and also the

complex interactions of soil, genotypes, weather, and management for optimal plant growth and development. (b) Future work can investigate the use of soil, weather, genotype, and remote sensing data to develop a generalized model for yield prediction. We controlled for variations in genetic background by centering populations around their mean values, but more work looking into additional crops should be done to determine the effectiveness of this methodology. More recent work that studied the effect of including genetic relationships in spatial adjustment methods has shown that including SNP and pedigree information can help to improve spatial adjustments; however, this was based on simulated data (Borges da Silva et al., 2021). As breeders continue to have more tools at their disposal, modeling the genetic and environmental factors could help to make more informed selections. This can remove non-genetic effects when evaluating lines and also start to investigate the genetic components that affect nutrient use efficiency and help to create more prescriptive cultivars across regions (Parmley, Higgins et al., 2019). (c) The analysis methods presented in our study were not used to compare performance in the following years, to see how different selection methods performed. Future work should investigate the repeatability of methods across years and locations. (d) Variable soil fertility was observed across all testing sites, but the locations that were tested are generally considered to be high-yielding environments. Future work investigating the utility of soil mapping in lower productive soil is needed to see the utility of this method in a wider range of environments.

5 | CONCLUSION AND FUTURE WORK

We demonstrate the usefulness of spatial adjustment models to improve the selection efficiency in plant breeding programs by removing environmental trends that can bias selection decision-making. While spatial adjustments do not give any advantage over unadjusted plot seed yield-based analysis when no field variation exists, such a scenario is extremely rare. Therefore, unreplicated trials will benefit from the use of spatial adjustment models. In this paper, we demonstrate the usefulness of an ML method for spatial adjustment in plot experiments. This method models soil fertility trends within a field and allows breeders to increase their knowledge about the quality of the growing conditions of each plot within a field trial. Compared to unadjusted, MM, and Spline methods, the XGB_Local model had the greatest increase in the mean relative efficiency, the lowest median CZ coefficient across all tests and selection thresholds, and showed a consistent ability to reduce field gradient effects on the yield of plots. The XGBoost method we demonstrate is based on soil features, which becomes more important when breeders want to grow genetically diverse and phenologically dissimilar genotypes in the same test where the soil map indicates variability in nutrients, soil texture properties, pH, CEC, and OM. XGBoost provides interpretability of the adjustments that can increase breeders' confidence in the application of spatial adjustment methods. Additionally, soil-based adjustments provide opportunities to select for genotypes that respond to soil features, for example, nutrient deficiency response. As crop breeding and production innovations continue to bring in and integrate advanced data analytics and computer science approaches, for example, cyber-agricultural systems (Singh et al., 2024; Sarkar et al., 2023), we envision crop improvement will effectively utilize plant production variables, including soil features, for variety development.

ACKNOWLEDGMENTS

We thank staff and student members of Singh Soybean group at ISU, particularly Brian Scott, Will Doepke, Jennifer Hicks, and Ryan Dunn for their assistance with field experiments and phenotyping. We thank Sarah Jones and Ashlyn Rairdin for help in editing the manuscript. We thank Yones Khaledian and Caner Ferhatoglu for generating the soil maps and members of the Geospatial Laboratory for Soil Informatics group for their assistance in soil core sampling. We are thankful to Dr. Walter Fehr, Grace Welke, and Susan Johnson for providing IA2102 seed.

Open access funding provided by the Iowa State University Library.

AUTHOR CONTRIBUTIONS

Matthew E. Carroll: Conceptualization; data curation; formal analysis; investigation; methodology; validation; visualization; writing—original draft; writing—review and editing. **Luis G. Riera**: Formal analysis; investigation; writing—original draft. **Bradley A. Miller**: Investigation; methodology; writing—review and editing. **Philip M. Dixon**: Formal analysis; methodology; writing—review and editing. **Baskar Ganapathysubramanian**: Funding acquisition; methodology; writing—review and editing. **Soumik Sarkar**: Conceptualization; formal analysis; funding acquisition; investigation; methodology; supervision; writing—original draft; writing—review and editing. **Asheesh K. Singh**: Conceptualization; funding acquisition; investigation; methodology; project administration; resources; supervision; visualization; writing—original draft; writing—review and editing.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data and codes will be made available publicly after acceptance of the paper through corresponding authors' GitHub pages.

ORCID

Matthew E. Carroll https://orcid.org/0000-0002-5224-7326

Bradley A. Miller https://orcid.org/0000-0001-8194-123X
Baskar Ganapathysubramanian https://orcid.org/0000-0002-8931-4852
Asheesh K. Singh https://orcid.org/0000-0002-7522-037X

REFERENCES

Akdemir, D., Beavis, W., Fritsche-Neto, R., Singh, A. K., & Isidro-Sánchez, J. (2019). Multi-objective optimized genomic breeding strategies for sustainable food improvement. *Heredity*, *122*(5), 672–683.

Akintayo, A., Tylka, G. L., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). A deep learning framework to discern and count microscopic nematode eggs. *Scientific Reports*, *8*(1), 1–11.

Bakhsh, A., Colvin, T. S., Jaynes, D. B., Kanwar, R. S., & Tim, U. S. (2000). Using soil attributes and GIS for interpretation of spatial variability in yield. *Transactions of the ASAE*, *43*(4), 819–828.

Bakhsh, A., Jaynes, D. B., Colvin, T. S., & Kanwar, R. S. (2000). Spatio-temporal analysis of yield variability for a corn-soybean field in Iowa. *Transactions of the ASAE*, *43*(1), 31–38.

Baligar, V., Fageria, N., & He, Z. (2001). Nutrient use efficiency in plants. *Communications in Soil Science and Plant Analysis*, *32*(7-8), 921–950.

Bernardo, R. (2002). *Breeding for quantitative traits in plants* (Vol. 1). Stemma Press.

Bivand, R., Altman, M., Anselin, L., Assunção, R., Berke, O., Bernat, A., & Blanchet, G. (2015). *Package 'spdep'* [Computer software]. The Comprehensive R Archive Network.

Bivand, R. S., & Wong, D. W. (2018). Comparing implementations of global and local indicators of spatial association. *Test*, *27*(3), 716–748.

Borges da Silva, É. D., Xavier, A., & Faria, M. V. (2021). Joint modeling of genetics and field variation in plant breeding trials using relationship and different spatial methods: A simulation study of accuracy and bias. *Agronomy*, *11*(7), 1397.

Bos, I., & Caligari, P. (2007). *Selection methods in plant breeding*. Springer Science & Business Media.

Brooker, A. P., Lindsey, L. E., Culman, S. W., Subburayalu, S. K., & Thomison, P. R. (2017). Low soil phosphorus and potassium limit soybean grain yield in Ohio. *Crop, Forage & Turfgrass Management*, *3*(1), 1–5. https://doi.org/10.2134/cftm2016.12.0081

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., & Varoquaux, G. (2013). API design for machine learning software: Experiences from the scikit-learn project. In *ECML PKDD workshop: Languages for data mining and machine learning*. https://doi.org/10.48550/arXiv.1309.0238

Byron, D., & Lambert, J. (1983). Screening soybeans for iron efficiency in the growth chamber. *Crop Science*, *23*(5), 885–888. https://doi.org/10.2135/cropsci1983.0011183X002300050017x

Carley, C., Zubrod, M., Dutta, S., & Singh, A. K. (2022). Using machine learning enabled phenotyping to characterize nodulation in three early vegetative stages in soybean. *Crop Science*, 63, 204–226. https://doi.org/10.1002/csc2.20861

Chattopadhyay, S., Gupta, A., Carroll, M., Raigne, J., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2023). A comprehensive study on soybean yield prediction using soil and hyperspectral reflectance data. *Preprints*. https://doi.org/10.20944/preprints202310.1232.v2

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on knowledge discovery and data mining*, KDD '16 (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Chiozza, M. V., Parmley, K. A., Higgins, R. H., Singh, A. K., & Miguez, F. E. (2021). Comparative prediction accuracy of hyperspectral bands for different soybean crop variables: From leaf area to seed composition. *Field Crops Research*, *271*, 108260.

Chiranjeevi, S., Young, T., Jubery, T. Z., Nagasubramanian, K., Sarkar, S., Singh, A. K., Singh, A., & Ganapathysubramanian, B. (2021). Exploring the use of 3d point cloud data for improved plant stress rating. In *AI for agriculture and food systems*. AAAI.

Cooper, M., Powell, O., Voss-Fels, K., Messina, C., Gho, C., Podlich, D., Technow, F., Chapman, S., Beveridge, C., Ortiz-Barrientos, D., & Hammer, G. L. (2021). Modelling selection response in plant-breeding programs using crop models as mechanistic gene-to-phenotype (cgm-g2p) multi-trait link functions. *in silico Plants*, *3*(1), diaa016.

Cullis, B., & Gleeson, A. (1991). Spatial analysis of field experiments-an extension to two dimensions. *Biometrics*, *47*(4), 1449–1460.

Cursi, D. E., Gazaffi, R., Hoffmann, H. P., Brasco, T. L., do Amaral, L. R., & Dourado Neto, D. (2021). Novel tools for adjusting spatial variability in the early sugarcane breeding stage. *Frontiers in Plant Science*, *12*, 749533.

De Faveri, J., Verbyla, A. P., Cullis, B. R., Pitchford, W. S., & Thompson, R. (2017). Residual variance–covariance modelling in analysis of multivariate data from variety selection trials. *Journal of Agricultural, Biological and Environmental Statistics*, *22*, 1–22.

Dobbels, A. A., & Lorenz, A. J. (2019). Soybean iron deficiency chlorosis high-throughput phenotyping using an unmanned aircraft system. *Plant Methods*, *15*(1), 1–9.

Dodd, J. R., & Mallarino, A. P. (2005). Soil-test phosphorus and crop grain yield responses to long-term phosphorus fertilization for corn-soybean rotations. *Soil Science Society of America Journal*, *69*(4), 1118–1128. https://doi.org/10.2136/sssaj2004.0279

Falk, K. G., Jubery, T. Z., Mirnezami, S. V., Parmley, K. A., Sarkar, S., Singh, A., Ganapathysubramanian, B., & Singh, A. K. (2020). Computer vision and machine learning enabled soybean root phenotyping pipeline. *Plant Methods*, *16*(1), 5.

Falk, K. G., Jubery, T. Z., O'Rourke, J. A., Singh, A., Sarkar, S., Ganapathysubramanian, B., & Singh, A. K. (2020). Soybean root system architecture trait study through genotypic, phenotypic, and shape-based clusters. *Plant Phenomics*, *2020*, 1925495. https://doi.org/10.34133/2020/1925495

Federer, W. T. (1961). Augmented designs with one-way elimination of heterogeneity. *Biometrics*, *17*(3), 447–473.

Ferhatoglu, C., & Miller, B. A. (2022). Choosing feature selection methods for spatial modeling of soil fertility properties at the field scale. *Agronomy*, *12*(8), 1786.

Fudickar, S., Nustede, E. J., Dreyer, E., & Bornhorst, J. (2021). Mask r-cnn based *C. elegans* detection with a DIY microscope. *Biosensors*, *11*(8), 257.

Gardner, C. (1961). An evaluation of effects of mass selection and seed irradiation with thermal neutrons on yield of corn. *Crop Science*, *1*(4), 241–245. https://doi.org/10.2135/cropsci1961.0011183X000100040004x

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences*, *115*(18), 4613–4618.

Ghosal, S., Zheng, B., Chapman, S. C., Potgieter, A. B., Jordan, D. R., Wang, X., Singh, A. K., Singh, A., Hirafuji, M., Ninomiya, S., Ganapathysubramanian, B., Sarkar, S., & Guo, W. (2019). A weakly supervised deep learning framework for sorghum head detection and counting. *Plant Phenomics*, *2019*, 1525874.

Gilmour, A. R., Cullis, B. R., & Verbyla, A. P. (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological, and Environmental Statistics*, *2*(3), 269–293.

Guo, W., Carroll, M. E., Singh, A., Swetnam, T. L., Merchant, N., Sarkar, S., Singh, A. K., & Ganapathysubramanian, B. (2021). UAS-based plant phenotyping for research and breeding applications. *Plant Phenomics*, *2021*, 9840192. https://doi.org/10.34133/2021/9840192

Gupta, A., & Singh, A. (2023). Agri-gnn: A novel genotypic-topological graph neural network framework built on graphsage for optimized yield prediction. *arXiv*, arXiv:2310.13037. https://arxiv.org/abs/2310.13037

Hazel, L., & Lush, J. L. (1942). The efficiency of three methods of selection. *Journal of Heredity*, *33*(11), 393–399.

Herr, A. W., Adak, A., Carroll, M. E., Elango, D., Kar, S., Li, C., Jones, S. E., Carter, A. H., Murray, S. C., Paterson, A., Sankaran, S., Singh, A., & Singh, A. K. (2023). Unoccupied aerial systems imagery for phenotyping in cotton, maize, soybean, and wheat breeding. *Crop Science*, *63*(4), 1722–1749. https://doi.org/10.1002/csc2.21028

Herrero-Huerta, M., Rodriguez-Gonzalvez, P., & Rainey, K. M. (2020). Yield prediction by machine learning from UAS-based multi-sensor data fusion in soybean. *Plant Methods*, *16*(1), 1–16.

Jubery, T. Z., Carley, C. N., Singh, A., Sarkar, S., Ganapathysubramanian, B., & Singh, A. K. (2021). Using machine learning to develop a fully automated soybean nodule acquisition pipeline (SNAP). *Plant Phenomics*, *2021*, 9834746. https://doi.org/10.34133/2021/9834746

Kar, S., Nagasubramanian, K., Elango, D., Carroll, M. E., Abel, C. A., Nair, A., Mueller, D. S., O'Neal, M. E., Singh, A. K., Sarkar, S., Ganapathysubramanian, B., & Singh, A. (2023). Self-supervised learning improves agricultural pest classification. *The Plant Phenome Journal*, *6*(1), e20079. https://doi.org/10.1002/ppj2.20079

Kempton, R. (1984). The design and analysis of unreplicated field trials. *Vortraege fuer Pflanzenzuechtung*, *7*, 219–242.

Khaledian, Y., & Miller, B. A. (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling*, *81*, 401–418.

Krause, M. D., das Gracas Dias, K. O., Singh, A. K., & Beavis, W. D. (2022). Using large soybean historical data to study genotype by environment variation and identify mega-environments with the integration of genetic and non-genetic factors. *bioRxiv*. https://doi.org/10.1101/2022.04.11.487885

Krause, M. D., Piepho, H.-P., Dias, K. O. G., Singh, A. K., & Beavis, W. D. (2023). Models to estimate genetic gain of soybean seed yield from annual multi-environment field trials. *TAG. Theoretical and Applied Genetics. Theoretische und angewandte Genetik*, *136*(12), 252. https://doi.org/10.1007/s00122-023-04470-3

Lado, B., Matus, I., Rodriguez, A., Inostroza, L., Poland, J., Belzile, F., del Pozo, A., Quincke, M., Castro, M., & von Zitzewitz, J. (2013). Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3: Genes, Genomes, Genetics*, *3*(12), 2105–2114.

Li, X., Guo, T., Mu, Q., Li, X., & Yu, J. (2018). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences*, *115*(26), 6679–6684.

Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, *45*(1), 255–268.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc.

Lynch, J. P., & Brown, K. M. (2001). Topsoil foraging–an architectural adaptation of plants to low phosphorus availability. *Plant and Soil*, *237*(2), 225–237.

Magnussen, S. (1990). Application and comparison of spatial models in analyzing tree-genetics field trials. *Canadian Journal of Forest Research*, *20*(5), 536–546.

Mallarino, A. P., & Sawyer, J. E. (2013). *Interpretation of soil test results*. Iowa State University Extension and Outreach. https://store.extension.iastate.edu/product/Interpretation-of-Soil-Test-Results

Mallarino, A., Webb, J., & Blackmer, A. (1991). Soil test values and grain yields during 14 years of potassium fertilization of corn and soybean. *Journal of Production Agriculture*, *4*(4), 560–567.

Mao, X., Dutta, S., Wong, R. K., & Nettleton, D. (2020). Adjusting for spatial effects in genomic prediction. *Journal of Agricultural, Biological and Environmental Statistics*, *25*, 699–718.

McGrath, C., Wright, D., Mallarino, A. P., & Lenssen, A. W. (2013). *Soybean nutrient needs*. Iowa State University Extension and Outreach. https://store.extension.iastate.edu/product/13954

Miao, C., Xu, Z., Rodene, E., Yang, J., & Schnable, J. C. (2020). Semantic segmentation of sorghum using hyperspectral data identifies genetic associations. *Plant Phenomics*, *2020*, 4216373.

Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, *264*, 301–311.

Minasny, B., Setiawan, B. I., Saptomo, S. K., & McBratney, A. B. (2018). Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands. *Geoderma*, *313*, 25–40.

Moreira, F. F., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2019). Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. *Plant Methods*, *15*(1), 1–9.

Nagasubramanian, K., Jones, S., Sarkar, S., Singh, A. K., Singh, A., & Ganapathysubramanian, B. (2018). Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant Methods*, *14*, 86.

Nagasubramanian, K., Jones, S., Singh, A. K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. (2019). Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods*, *15*, 98.

Nagasubramanian, K., Singh, A., Singh, A., Sarkar, S., & Ganapathysubramanian, B. (2022). Plant phenotyping with limited annotation: Doing more with less. *The Plant Phenome Journal*, *5*(1), e20051. https://doi.org/10.1002/ppj2.20051

Nagasubramanian, K., Singh, A. K., Singh, A., Sarkar, S., & Ganapathysubramanian, B. (2020). Usefulness of interpretability methods to explain deep learning based plant stress phenotyping. *arXiv*, arXiv:2007.05729. https://arxiv.org/abs/2007.05729

Naik, H. S., Zhang, J., Lofquist, A., Assefa, T., Sarkar, S., Ackerman, D., Singh, A., Singh, A. K., & Ganapathysubramanian, B. (2017). A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. *Plant Methods*, *13*, 23.

Oakey, H., Cullis, B., Thompson, R., Comadran, J., Halpin, C., & Waugh, R. (2016). Genomic selection in multi-environment crop trials. *G3: Genes, Genomes, Genetics*, *6*(5), 1313–1326.

Ojeda, J. J., Hammer, G., Yang, K.-W., Tuinstra, M. R., DeVoil, P., McLean, G., Huber, I., Volenec, J. J., Brouder, S. M., Archontoulis, S., & Chapman, S. C. (2022). Quantifying the effects of varietal types× management on the spatial variability of sorghum biomass across us environments. *GCB Bioenergy*, *14*(3), 411–433.

Oliveira, M. F., Nelson, R. L., Geraldi, I. O., Cruz, C. D., & de Toledo, J. F. F. (2010). Establishing a soybean germplasm core collection. *Field Crops Research*, *119*(2-3), 277–289.

Parmley, K., Nagasubramanian, K., Sarkar, S., Ganapathysubramanian, B., & Singh, A. K. (2019). Development of optimized phenomic predictors for efficient plant breeding decisions using phenomic-assisted selection in soybean. *Plant Phenomics*, *2019*, 5809404.

Parmley, K. A., Higgins, R. H., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2019). Machine learning approach for prescriptive plant breeding. *Scientific Reports*, *9*, 17132. https://doi.org/10.1038/s41598-019-53451-4

Qiao, C., Basford, K., DeLacy, I., & Cooper, M. (2000). Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theoretical and Applied Genetics*, *100*(1), 9–16.

Rairdin, A., Fotouhi, F., Zhang, J., Mueller, D., Ganapathysubramanian, B., Singh, A., Dutta, S., & Sarkar, S. (2022). Deep learning-based phenotyping for genome wide association studies of sudden death syndrome in soybean. *Frontiers in Plant Science*, *13*, 966244.

Reynolds, M., Chapman, S., Crespo-Herrera, L., Molero, G., Mondal, S., Pequeno, D. N., Pinto, F., Pinera-Chavez, F. J., Poland, J., Rivera-Amado, C., Saint Pierre, C., & Sukumaran, S. (2020). Breeder friendly phenotyping. *Plant Science*, *295*, 110396.

Richey, F. D. (1924). Adjusting yields to their regression on a moving average, as a means of correcting for soil heterogeneity. *Journal of Agricultural Research*, *27*(1-13), 79–90.

Riera, L. G., Carroll, M. E., Zhang, Z., Shook, J. M., Ghosal, S., Gao, T., Singh, A., Bhattacharya, S., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2021). Deep multiview image fusion for soybean yield estimation in breeding applications. *Plant Phenomics*, *2021*, 9846470. https://doi.org/10.34133/2021/9846470

Rodriguez-Alvarez, M. X., Boer, M. P., van Eeuwijk, F. A., & Eilers, P. H. (2018). Correcting for spatial heterogeneity in plant breeding experiments with p-splines. *Spatial Statistics*, *23*, 52–71.

Rutkoski, J. E. (2019). A practical guide to genetic gain. *Advances in Agronomy*, *157*, 217–249.

Sagan, V., Maimaitijiang, M., Bhadra, S., Maimaitiyiming, M., Brown, D. R., Sidike, P., & Fritschi, F. B. (2021). Field-scale crop yield prediction using multi-temporal worldview-3 and planetscope satellite data and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *174*, 265–281.

Sánchez T, J. D., Ligarreto M, G. A., & Leiva, F. R. (2011). Spatial variability of soil chemical properties and its effect on crop yields: a case study in maize (*Zea mays* L.) on the Bogota plateau. *Agronomía Colombiana*, *29*(2), 456–466.

Sarkar, S., Ganapathysubramanian, B., Singh, A., Fotouhi, F., Kar, S., Nagasubramanian, K., Chowdhary, G., Das, S. K., Kantor, G., Krishnamurthy, A., Merchant, N., & Singh, A. K. (2023). Cyber-agricultural systems for crop breeding and sustainable production. *Trends in Plant Science*, *29*(2), 130–149.

Shook, J. M., Gangopadhyay, T., Wu, L., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2021). Crop yield prediction integrating genotype and weather variables using deep learning. *PLoS ONE*, *16*(6), e0252402.

Shook, J. M., Lourenco, D., & Singh, A. K. (2021). PATRIOT: A pipeline for tracing identity-by-descent for chromosome segments to improve genomic prediction in self-pollinating crop species. *Frontiers in Plant Science*, *12*, 676269. https://doi.org/10.3389/fpls.2021.676269

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning important features through propagating activation differences. In *International conference on machine learning* (pp. 3145–3153). PMLR.

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science*, *21*(2), 110–124.

Singh, A., Jones, S., Ganapathysubramanian, B., Sarkar, S., Mueller, D., Sandhu, K., & Nagasubramanian, K. (2021). Challenges and opportunities in machine-augmented plant stress phenotyping. *Trends in Plant Science*, *26*(1), 53–69.

Singh, A. K., Balabaygloo, B. J., Bekee, B., Blair, S. W., Fey, S., Fotouhi, F., Gupta, A., Jha, A., Martinez-Palomares, J. C., Menke, K., Prestholt, A., Tanwar, V. K., Tao, X., Vangala, A., Carroll, M. E., Das, S., DePaula, G., Kyveryga, P., Sarkar, S., … Valdivia, C. (2024). Smart connected farms and networked farmers to improve crop production, sustainability and profitability. *Frontiers in Agronomy*, *6*, 1410829. https://doi.org/10.3389/fagro.2024.1410829

Singh, A. K., Ganapathysubramanian, B., Sarkar, S., & Singh, A. (2018). Deep learning for plant stress phenotyping: trends and future perspectives. *Trends in Plant Science*, *23*(10), 883–898.

Singh, A. K., Singh, A., Sarkar, S., Ganapathysubramanian, B., Schapaugh, W., Miguez, F. E., Carley, C. N., Carroll, M. E., Chiozza, M. V., Chiteri, K. O., Falk, K., Jones, S. E., Jubery, Z., Mirnezami, S. V., Nagasubramanian, K., Parmley, K. A., Rairdin, A. M., Shook, J. M., Van der Laan, L., … Zhang, J. (2021). High-throughput phenotyping in soybean. In *High-throughput crop phenotyping* (pp. 129–163). Springer.

Singh, D. P., Singh, A. K., & Singh, A. (2021). *Plant breeding and cultivar development*. Academic Press.

Smidt, E. R., Conley, S. P., Zhu, J., & Arriaga, F. J. (2016). Identifying field attributes that predict soybean yield using random forest analysis. *Agronomy Journal*, *108*(2), 637–646. https://doi.org/10.2134/agronj2015.0222

Smith, A., Cullis, B., & Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*, *57*(4), 1138–1147.

Stelling, D., Ismail, M., Ebmeyer, E., Fraukn, M., & Robbelen, G. (1990). Selection in early generations of dried peas, *Pisum sativum* L. III. Plot size and plot type. *Plant Breeding*, *105*(3), 238–247.

Štrumbelj, E., & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, *41*(3), 647–665.

Tabery, J. (2008). Ra fisher, lancelot hogben, and the origin (s) of genotype–environment interaction. *Journal of the History of Biology*, *41*, 717–761.

Technow, F. (2015). *R package mvnggrad: Moving grid adjustment in plant breeding field trials* (R package Version 0.1.5) [Computer software]. The Comprehensive R Archive Network.

Tetila, E. C., Machado, B. B., Menezes, G. K., Oliveira, A. d. S., Alvarez, M., Amorim, W. P., Belete, N. A. D. S., Da Silva, G. G., & Pistori, H. (2019). Automatic recognition of soybean leaf diseases using uav

images and deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, *17*(5), 903–907.

van Rossum, B.-J., van Eeuwijk, F., & Boer, M. (2021). *Package 'statgensta'*. The Comprehensive R Archive Network. https://cran.r-project.org/src/contrib/Archive/statgenSTA/

Velazco, J. G., Rodríguez-Álvarez, M. X., Boer, M. P., Jordan, D. R., Eilers, P. H., Malosetti, M., & Van Eeuwijk, F. A. (2017). Modelling spatial trends in sorghum breeding field trials using a two-dimensional p-spline mixed model. *Theoretical and Applied Genetics*, *130*, 1375–1392.

Verbyla, A. P., De Faveri, J., Wilkie, J. D., & Lewis, T. (2018). Tensor cubic smoothing splines in designed experiments requiring residual modelling. *Journal of Agricultural, Biological and Environmental Statistics*, *23*, 478–508.

Xavier, A., Hall, B., Hearst, A. A., Cherkauer, K. A., & Rainey, K. M. (2017). Genetic architecture of phenomic-enabled canopy coverage in *Glycine max*. *Genetics*, *206*(2), 1081–1089.

Young, T. J., Jubery, T. Z., Carley, C. N., Carroll, M., Sarkar, S., Singh, A. K., Singh, A., & Ganapathysubramanian, B. (2023). "canopy fingerprints" for characterizing three-dimensional point cloud data of soybean canopies. *Frontiers in Plant Science*, *14*, 1141153.

Zhang, J., Naik, H. S., Assefa, T., Sarkar, S., Reddy, R., Singh, A., Ganapathysubramanian, B., & Singh, A. K. (2017). Computer vision and machine learning for robust phenotyping in genome-wide studies. *Scientific Reports*, *7*, 44048.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Carroll, M. E., Riera, L. G., Miller, B. A., Dixon, P. M., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2024). Leveraging soil mapping and machine learning to improve spatial adjustments in plant breeding trials. *Crop Science*, *64*, 3135–3152. https://doi.org/10.1002/csc2.21336