2019

Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations

Curtis J. Ransom
University of Missouri, cjr8bf@mail.missouri.edu

Newell R. Kitchen
USDA, Agricultural Research Service

James J. Camberato
Purdue University

Paul R. Carter
Corteva Agrisciences

Richard B. Ferguson
University of Nebraska-Lincoln, rferguson1@unl.edu

Computers and Electronics in Agriculture 164 (2019) 104872

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations

Curtis J. Ransom[a,*], Newell R. Kitchen[b], James J. Camberato[c], Paul R. Carter[d], Richard B. Ferguson[e], Fabián G. Fernández[f], David W. Franzen[g], Carrie A.M. Laboski[h], D. Brenton Myers[d], Emerson D. Nafziger[i], John E. Sawyer[j], John F. Shanahan[k]

[a] University of Missouri, 269 Agric. Eng. Bldg., Columbia, MO 65211, USA
[b] USDA-ARS Cropping Systems and Water Quality Research Unit, 243 Agric. Eng. Bldg., Columbia, MO 65211, USA
[c] Purdue University, Lilly 3-365, West Lafayette, IN 47907, USA
[d] Corteva Agrisciences, 8325 NW 62nd Ave., Johnston, IA 50131, USA
[e] University of Nebraska, Keim 367, Lincoln NE 68583, USA
[f] University of Minnesota, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
[g] North Dakota State University, PO Box 6050, Fargo, ND 58108, USA
[h] University of Wisconsin-Madison, 1525 Observatory Dr., Madison, WI 53706, USA
[i] University of Illinois, W-301 Turner Hall, 1102 S. Goodwin, Urbana, IL 61801, USA
[j] Iowa State University, 3208 Agronomy Hall, Ames, IA 50011, USA
[k] Fortigen, 6807 Ridge Rd, Lincoln, NE 68512, USA

ARTICLE INFO

Keywords:
Corn
Machine learning
Nitrogen fertilizer recommendations
Soil
Weather

ABSTRACT

Nitrogen (N) fertilizer recommendation tools could be improved for estimating corn (*Zea mays* L.) N needs by incorporating site-specific soil and weather information. However, an evaluation of analytical methods is needed to determine the success of incorporating this information. The objectives of this research were to evaluate statistical and machine learning (ML) algorithms for utilizing soil and weather information for improving corn N recommendation tools. Eight algorithms [stepwise, ridge regression, least absolute shrinkage and selection operator (Lasso), elastic net regression, principal component regression (PCR), partial least squares regression (PLSR), decision tree, and random forest] were evaluated using a dataset containing measured soil and weather variables from a regional database. The performance was evaluated based on how well these algorithms predicted corn economically optimal N rates (EONR) from 49 sites in the U.S. Midwest. Multiple algorithm modeling scenarios were examined with and without adjustment for multicollinearity and inclusion of two-way interaction terms to identify the soil and weather variables that could improve three dissimilar N recommendation tools. Results showed the out-of-sample root-mean-square error (RMSE) for the decision tree and some random forest modeling scenarios were better than the stepwise or ridge regression, but not significantly different than any other algorithm. The best ML algorithm for adjusting N recommendation tools was the random forest approach (r^2 increased between 0.72 and 0.84 and the RMSE decreased between 41 and 94 kg N ha^{-1}). However, the ML algorithm that best adjusted tools while using a minimal amount of variables was the decision tree. This method was simple, needing only one or two variables (regardless of modeling scenario) and provided moderate improvement as r^2 values increased between 0.15 and 0.51 and RMSE decreased between 16 and 66 kg N ha^{-1}. Using ML algorithms to adjust N recommendation tools with soil and weather information shows promising results for better N management in the U.S. Midwest.

1. Introduction

To maximize profits and minimize N related environmental issues, N fertilizer recommendation decision tools are needed that closely match the EONR for a given field (Bandura, 2017; Hong et al., 2007; Kyveryga et al., 2009). The EONR for a given field and year is unknown at the time of N application. Determining the EONR typically requires an elaborate process, involving the establishment of on-farm research trials with varying N rates,

Abbreviations: EONR, economically optimal nitrogen rate; Lasso, least absolute shrinkage and selection operator; ML, machine learning; PCR, principal component regression, PLSR, partial least squares; RMSE, root-mean-square error; YG, yield goal

* Corresponding author.
E-mail address: cjr8bf@mail.missouri.edu (C.J. Ransom).

https://doi.org/10.1016/j.compag.2019.104872
Received 7 March 2019; Received in revised form 12 June 2019; Accepted 23 June 2019
Available online 01 August 2019
0168-1699/

measuring yield responses and finally an end of season calculation using current grain and fertilizer prices. Furthermore, EONR varies considerably from year-to-year and within a field, making it even more challenging to estimate (Kyveryga et al., 2009; Scharf et al., 2005; Shanahan et al., 2008). Both the temporal and spatial variability of EONR are driven by environmental, genetic, and management processes. More specifically, these factors include: N form, placement, timing, rate, rainfall distribution, soil texture, soil water-holding capacity, plant genetics, manure history, and the previous crop (Dinnes et al., 2002; Kay et al., 2006; Schmidt et al., 2009; Zhu et al., 2009; Tremblay et al., 2012; Morris et al., 2018).

Multiple corn N fertilizer rate decision tools have been developed over the last five decades in an attempt to incorporate many of these factors to help farmers make better N management decisions to optimize yield and/or profit. Historically, this began with yield goal-based recommendation approach (Stanford, 1973) which targeted plant N content at maturity based on predicted yield times an N factor [2.1×10^{-2} kg N (kg grain)$^{-1}$] minus an estimate of the N contribution from the soil all divided by an estimate of the efficiency of fertilizer N use. Later iterations of this approach adjusted the soil N estimate by tillage, previous crop, and other factors and/or subtracted measured levels of nitrate-N from the targeted N rate (Morris et al., 2018). Many of the major corn producing states no longer recommend using YG based tools, but this method underpins many of the crop growth models addressed below. Alongside the YG approach, soil nitrate tests have been utilized to measure the nitrate concentration prior to an at-planting or in-season application. Corn N recommendations based on soil nitrate measurements have been helpful in reducing over-application of N fertilizers in fields with a large residual nitrate concentration, such as manured fields or planting after alfalfa (Bundy et al., 1999; Bundy and Andraski, 1995; Sawyer and Mallarino, 2017). Crop canopy reflectance sensing is a more recent approach that assesses N stress based on the color and size of plants (Kitchen et al., 2010). This approach accounts for differences in growing conditions throughout a field by integrating the biotic and abiotic factors that affect the crop positively or negatively into an N recommendation at a very small spatial scale (1–5 m resolution; Kitchen et al., 2010). Recently, crop growth models (e.g., Encirca, Maize-N, Climate Fieldview, and Adapt-N) have been used to incorporate environmental, genetic, and management processes through mechanistic models to make an N recommendation (Moebius-Clune et al., 2013; Puntel et al., 2018; Setiyono et al., 2011). However, no one tool has consistently performed better than other tools across the U.S. Midwest with some research showing promising results (Scharf et al., 2006) and others showing poor performances for some of these tools (Bean et al., 2018a; Ransom, 2018; Thompson et al., 2015).

One way to improve N recommendation tools is to incorporate soil and weather information known to affect EONR into the N recommendation process. For example, the performance of a canopy reflectance sensing algorithm was improved when compared to EONR after incorporating site-specific soil and weather information into the algorithm (r^2 increased from 0.13 to 0.40) (Bean et al., 2018b). However, determining which soil and weather variables to incorporate into N recommendation tools requires filtering out the most useful variables from a large dataset. While identifying applicable variables can be done using any one of many mathematical and statistical procedures (Table 1) it is not known which procedure is the best. Some statistical approaches are computationally slow or result in adjustments that are agronomically incomprehensible due to multicollinearity, multiple interacting terms, and/or require many parameters inputs. Potential statistical and ML approaches that may be helpful for adjusting N recommendation tools are briefly discussed in the next section.

2. Statistical and machine learning algorithms

2.1. Stepwise regression

A standard method used in agricultural research is the least squares regression for estimating parameter coefficients. Stepwise regression is an example of this method. The utility of stepwise regression is the ability to add or remove variables from a model in controlled steps to better determine which explanatory variables relate to a response variable (Yamashita et al., 2007). However, stepwise regression predictions have been found to overestimate validation datasets as it puts a high bias on each of the parameters and relies heavily on the assumption of having a single best model (Thompson, 2001; Whittingham et al., 2006; Zou, 2006).

2.2. Penalization regression

Instead of using stepwise regression, penalty-based regression procedures [i.e., ridge regression or Lasso algorithms] could be used (Tibshirani, 1996; Zhao and Yu, 2006; McDonald, 2009). Ridge regression works as a continual shrinkage method in which the residual sum of squares is minimized as each parameter's coefficient is adjusted close to zero, thus reducing the importance or influence of any one parameter (Hoerl and Kennard, 1970). In contrast, the Lasso regression reduces coefficient parameters to zero, thus selecting essential variables and shrinking the number of model parameters simultaneously (Tibshirani, 1996). Lasso is particularly useful with large datasets as it computes efficiently and quickly (Friedman et al., 2010). Lasso can fail in the variable selection process when the number of observations is less than the number of parameters in the model or when there are many highly correlated variables (Zou and Hastie, 2005). To account for this weakness, the elastic net algorithm has been suggested as a way to determine the best combination of both the ridge regression and Lasso (Zou and Hastie, 2005). The balance between the two algorithms is accomplished by weighting the effects of ridge regression and Lasso.

2.3. Principal component regression and partial least squares regression

Apart from the penalization methods just described, other ML algorithms can account for the weakness associated with regression analysis. The PCR can overcome issues related to multicollinearity by transforming groups of explanatory variables into orthogonal basis vectors or principal components. Principal components are determined by iteratively finding linear combinations of input vectors that best capture the quantity and direction of the variance. Multiple approaches to this are available. The number of principal components extracted is limited to the number of explanatory variables. However, typically only the principal components that explain the most variance are retained for regression, resulting in a reduced number of new orthogonal explanatory variables. However, there is no guarantee that the newly devised principal components will be related to the response variable (Abdi and Williams, 2010; Jagadamma et al., 2008). A similar technique to the PCR is the PLSR. This method works by iteratively finding projected vectors with a minimum variance between the dependent and independent variables. The resulting set of vectors are the best relationship between the explanatory and response variables that explains the basic structure of the data. Relationships are established using linear regression models to fit pairs of explanatory and response variables. The best prediction functions are then regressed against the explanatory variable (Geladi and Kowalski, 1986). Both PCR and PLSR methods work well when the number of observations is less than the number of explanatory variables in the model.

2.4. Decision trees and random forests

Other ML algorithms that have the potential to identify and incorporate soil and weather data into N recommendation tools are decision tree based algorithms. Decision trees function by continually splitting a dataset based on some statistic from explanatory variables thereby creating a flowchart of decisions to predict a response variable (Quinlan, 1986). While the method is straightforward and easy to interpret, it often performs poorly on independent datasets as it overfits

Table 1
Summary of eight statistical and machine learning algorithms.

Algorithm	Background	Pros/Cons	Assumptions	Example Uses
Stepwise	Minimizing least square errors by controlled stepwise removal or addition of variables to a model.	**Pros:** Easy to run and interpret. **Cons:** Assumes that there is a perfect model. Results vary based on the order of variables. Affected by multicollinearity. Computationally very slow with lots of variables.	A linear relationship between the explanatory and independent variable. Normally distributed. No multicollinearity.	Used to explain geospatial variation in aboveground biomass and grain yield (Grassini et al., 2009).
Ridge Regression	Applies L2 penalty to minimize overfitting – less important variables' coefficients are reduced to close to zero	**Pros:** Easy to understand and explain. **Cons:** Requires an optimization of hyperparameters. Performs poorly with non-linear relationships.	A linear relationship between the explanatory and independent variable. Normally distributed. No multicollinearity.	Evaluate different forage traits using multisensory data (Xing et al., 2018).
Least absolute shrinkage and selection operator (Lasso)	Applies L1 penalty to minimize overfitting – less important variables' coefficients are reduced to zero.	**Pros:** Easy to understand and explain. **Cons:** Requires an optimization of hyperparameters. Performs poorly with non-linear relationships.	A linear relationship between the explanatory and independent variable. Normally distributed. No multicollinearity.	Predicting genomic traits associated with *Fusarium* head blight in wheat (Arruda et al., 2015).
Elastic Net	A combination of Ridge Regression and Lasso regression.	**Pros:** Easy to understand and explain. **Cons:** Requires an optimization of hyperparameters. Performs poorly with non-linear relationships.	A linear relationship between the explanatory and independent variable. Normally distributed. No multicollinearity.	Predicting genomic traits associated with *Fusarium* head blight in wheat (Arruda et al., 2015).
Principal Component Regression (PCR)	Combines multiple explanatory variables into new and fewer variables that are used to explain the variation of a dataset.	**Pros:** Able to account for multicollinearity. **Cons:** Variables need to be scaled and normalized. There is no guarantee that new variables are related to the response variable.	A linear relationship between the explanatory and independent variable. Normally distributed. No multicollinearity.	Determining leaf N concentration and fruit yield using vegetative indices (Wang et al., 2017).
Partial Least Squares Regression (PLSR)	A multivariate regression between explanatory and response variables.	**Pros:** Accounts for multicollinearity, small datasets, and missing samples **Cons:** Difficult interpreting latent variables. A lack of model test statistics.	A linear relationship between the explanatory and independent variable. No outliers.	Estimate N status in plants from hyperspectral reflectance data (Hansen and Schjoerring, 2003).
Recursive Partitioning Decision Tree	A decision tree is developed by splitting the response variable based on explanatory variables.	**Pros:** Easy to interpret, minimal data preparation. **Cons:** Often overfits data, small changes in training dataset could result in large changes in the final model.	No assumptions.	Using hyperspectral data for detecting water, nitrogen, or weed stress (Waheed et al., 2006).
Random Forest	An ensemble of multiple regression decision trees. Randomly uses a subsample of explanatory variables at each fit.	**Pros:** Learn non-linear relationships Robust to outliers and corrects overfitting observed with decision trees. **Cons:** Ensembling reduces the understanding of which variables are most important.	No assumptions.	Predict sugarcane leaf N concentration by selecting important spectral bands from satellite imagery (Abdel-Rahman et al., 2013).

the training datasets used to create the model. There are multiple ways to minimize overfitting, one of which is to resample the data to fit hundreds or thousands of decision trees. The average or the majority response from all the trees is then used as a robust production of the response variable. The random forest algorithm is an example of this approach; however, for each split in the dataset, it only uses a random subset of the explanatory variables to determine the split, thereby improving the diversity of tree types (Breiman, 2001; Grömping, 2009).

2.5. Improving nitrogen recommendation tools

Prior efforts showed promising results using ML algorithms to directly predict EONR using soil and weather information (Qin et al., 2018). However, these efforts did not utilize existing N recommendation methods which are already familiar to farmers. As such, a similar comparison of ML algorithms could also be employed to incorporate soil and weather information in an attempt to improve already existing N recommendation tools. The primary objective of this investigation was to expand the comparison of statistical and ML algorithms for incorporating soil and weather information into N recommendation tools. A secondary objective was to determine how well these ML algorithms performed with and without multicollinearity and two-way interaction terms.

3. Materials and methods

Data were obtained through a research collaboration between Corteva Agrisciences and eight U.S. Midwest universities (Iowa State University, University of Illinois Urbana-Champaign, University of Minnesota, University of Missouri, North Dakota State University, Purdue University, University of Nebraska-Lincoln, and University of Wisconsin-Madison). Each state conducted research on two sites each year from 2014 to 2016, with a third site in Missouri in 2016, totaling 49 site-years. About half the sites were on farmers' fields and the other half on University research stations. All states followed a similar protocol for plot research implementation including site selection, weather data collection, soil and plant sample timing and collection methodology, N application timing, N source, and N rates, with specific details described in Kitchen et al. (2017). Treatments included four replications of N fertilizer rates between 0 and 315 kg N ha^{-1} applied either all at-planting or split where 45 kg N ha^{-1} was applied at-planting and the remaining fertilizer N applied at the V9 ± 1 corn developmental stage (Abendroth et al., 2011).

3.1. Nitrogen recommendation tools and EONR

Three unique N recommendation tools were selected for this evaluation that ranged in their ability to predict EONR across a range of environmental and soil conditions (Ransom, 2018). They were: (1) farmer's N recommendation (Farmer's NR), (2) yield goal (YG), and (3) proximal active-optical canopy reflectance sensing. The farmer's NR was the rate the farmer or research station manager typically applied to the field site under ideal corn growing conditions. The information or recommendation system the farmer/station manager used to base the N rate was not recorded, but it was assumed to be based on crop response to N of the site over multiple years, and not necessarily on any particular decision tool.

The YG tool tested was previously promoted in the state of Indiana, Ohio, and Michigan (Tri-State YG; Vitosh et al., 1995), and was calculated as follows:

$$Tri - State \quad YG = [-30 + 0.025 \times YG - N_{credit}] \tag{1}$$

where YG was the expected yield in kg ha^{-1} and N_{credit} = 34 kg N ha^{-1} for corn following soybean. The YG for each site was determined using the average of the previous five-year county corn yields for the county where the site was located. The five-year average was then locally adjusted based on the soil productivity of the predominantly mapped soil of each site, similar to the approach of Laboski and Peters (2012). This soil adjustment classifies soil productivity as either low, medium, or high using soil texture, drainage class, depth to bedrock, available water capacity in the upper 150 cm of soil, average growing degree days, irrigation, and artificial tile drainage. The YG of a site was then adjusted by increasing the five-year average yield for low, medium, and high soil productivity by 10, 20, and 30%, respectively.

For canopy reflectance sensing, measurements were obtained using the RapidSCAN CS-45 (Holland Scientific, Lincoln NE, USA) at the same time as the split N application (i.e., generally ± 2 d of sensing). For most sites, this was done at the ~V8–V10 corn development stage. Measurement details are described in Kitchen et al. (2017). The Holland and Schepers algorithm (HS; Holland and Schepers, 2010) was used to calculate an N fertilizer recommendation derived from these reflectance measurements. This algorithm is based on a sufficiency index calculated using measurements from both well-fertilized corn ("N-Rich") and minimally-fertilized corn that is referred to here as the "target" corn:

$$SI = \frac{VI_{Target}}{VI_{N-Rich}} \tag{2}$$

where SI is the sufficiency index; VI_{Target} is the vegetative index obtained by averaging measurements from all plots that received 45 kg N ha^{-1} at-planting and where top-dress N fertilizer was to be applied, and $VI_{N\text{-}Rich}$ is the vegetative index obtained by averaging all plots of two high N treatments (225 and 270 kg N ha^{-1} applied all at-planting). The NDRE vegetative index was calculated using the red-edge (730 nm; RE) and near-infrared (780 nm; NIR) wavelengths as shown:

$$\text{NDRE} = \frac{NIR - RE}{NIR + RE} \tag{3}$$

Fertilizer N recommendations were then calculated as described in Holland and Schepers (2010) as follows:

$$N_{Rec} = (MZ_i{*}N_{Opt} - N_{PreFert} - N_{CRD} + N_{Comp}){*}\sqrt{\frac{(1 - SI)}{\Delta SI}} \tag{4}$$

where N_{Rec} is the calculated N fertilizer recommendation; MZ_i is a scaling value ($0 \geq MZ_i \leq 2$) used to adjust the N recommendation based on areas of high or low yield performance; N_{Opt} the base N rate, which is determined by the farmer; $N_{PreFert}$ is the amount of N already applied prior to sensing; N_{CRD} is N credits associated with the previous crop, NO_3–N in irrigation water, manure, or residual NO_3–N; N_{Comp} is an optional compensation factor for growth limiting conditions; SI is the sufficiency index, and ΔSI is a value to define the response range. For this analysis, MZ_i was left as the default value of 1.0, N_{opt} was set as the recorded farmer's N rate for each site, and $N_{PreFert}$ = 45 kg N ha^{-1}. With no supportive information relative to N_{CRD} and N_{Comp}, these two parameters were set to zero for all sites. The recommended value of 0.30 was used for ΔSI, which provides a response range between the measured vegetative index value between 0.70 and 1.00.

Grain yield in response to N fertilizer rate treatments was used to calculate the EONR on a site level by modeling all four replicates as described in Kitchen et al. (2017), using proven quadratic or quadratic-plateau modeling methods (Cerrato and Blackmer, 1990; Scharf et al., 2005). Economic optimal N rate values were calculated for all N fertilizer applied at-planting, and N split applied between planting and the single top-dress. For this study, the prices of N and grain were set at \$0.88 kg N^{-1} and \$0.158 kg grain^{-1} (equivalent to \$0.40 lb N^{-1} and \$4.00 bu^{-1}). The EONR was set to not exceed the maximum N rate (315 kg N ha^{-1}). Seven irrigated sites had additional N applied through irrigation between 1 and 53 kg N ha^{-1} which was included in models for determining the EONR of these sites. The EONR results were used as the standard for evaluating N recommendations developed from statistical and ML approaches.

3.2. Modeling scenarios

Ninety modeling scenarios [8 statistical and ML algorithms × 2 data processing scenarios (complete vs. a reduced number of soil and weather variables; explained below) × 2 interaction types (without vs. with interactions; described below) × 3 N recommendation tools (described above)] were considered. Six scenarios were excluded from these 90 scenarios as explained below. The essence of these analyses was to model the difference between a tool's N recommendation and the actual EONR as a function of soil and weather information using statistical and ML methods. Eight statistical and ML algorithms (hereafter all algorithms will be referred to as "ML algorithms" for simplicity) that were evaluated included: (1) stepwise regression using Akaike's information criteria, (2) ridge regression, (3) Lasso, (4) elastic net regression, (5) PCR, (6) PLSR, (7) recursive partitioning decision tree, and (8) random forest. Each of the ML algorithms was evaluated using a complete and reduced dataset. The complete dataset contained all available soil and weather variables (Table 2), while the reduced dataset excluded variables with a high pair-wise correlation ($|r| > 0.85$) (Table 3, Figs. 1 and 2). Correlated variables with the highest mean absolute correlation were removed from the dataset as identified with the findCorrelation function from the R 'caret' package (Kuhn, 2017). Additionally, for both the complete and reduced datasets the ML algorithms were evaluated with and without 2-way-interactions among all possible pairs of predictors. Lastly, six modeling scenarios (3 tools × 2 complete/reduced datasets) associated with using the stepwise regression with two-way interaction terms were excluded because of the process being computationally slow.

3.3. Statistical and machine learning parameters

All ML algorithms were fit using the 'caret' package in R Statistical Software (R Core Team, 2016). Seven of the eight ML algorithms required specific internal parameters to be tuned for optimal performance. A range of tuning parameter values was evaluated to ensure each ML algorithm was optimized (Table 4). Parameter tuning was accomplished by fitting each ML algorithm with a tenfold cross-validation repeated five times. Algorithms were trained on nine of the ten folds, and the accuracy was tested by calculating a RMSE by comparing predicted values to actual values on the tenth fold. The best tuning parameters were selected as the ones that produced the lowest average RMSE across all of the 50 cross-validation folds.

All explanatory variables were normalized for all ML algorithms except for the decision tree and random forest. Normalization was done by subtracting the mean and dividing by the standard deviation of each variable. After finding the optimal tuning parameter, all variable coefficients were "de-normalized" to their original scale.

For each modeling scenario, the response variable was the difference between each tool's N recommendation and the EONR value for each site. Where the EONR was calculated using N treatments applied all at-planting (for the Farmer's NR and Tri-State YG) or with split N treatments (for canopy reflectance sensing). Explanatory variables included measured physical and chemical soil properties and weather information. Soil properties were collected by sampling 1.2 m soil cores from each of the sites and analyzing each pedological soil horizon for texture, bulk density, pH salt, pH water, CEC, total N, total carbon, inorganic carbon, organic carbon, and organic matter as (Table 2). Soil

Table 2
Weather and soil variables used in the complete dataset with calculations, methods, and corresponding citations.

Complete Dataset			
Variables	Calculations and Sample Depths	Method	References
Weather			
Precipitation (PPT)	Sum of daily rainfall, mm.	Tipping bucket†	
Corn heat units (CHU)	Σ(Ymax + Ymin)/2; Ymax and Ymin are the daily maximum and minimum temperatures, °C.	Temperature sensor†	
Growing degree day (GDD)	$\Sigma((Ymax + Ymin)/2)-T_{base}$; Ymax, Ymin, T_{base} are the daily maximum, minimum, and base temperatures, respectively. T_{base} = 10 °C.	Temperature sensor	
Shannon diversity index (SDI)	$[-\Sigma pi\ ln(pi)]/ln(n)$; where pi = Rain/PPT (daily rainfall relative to total rainfall in a given time; n = total number of days.	Tipping bucket	(Tremblay et al., 2012)
Abundant and well-distributed rainfall (AWDR)	SDI × PPT	Tipping bucket	(Tremblay et al., 2012)
Soil			
Clay	0–30, 0–60, 0–90 cm	Pipette	Soil Survey Staff (2014) 3A1‡
Sand	0–30, 0–60, 0–90 cm	Pipette	Soil Survey Staff (2014) 3A1‡
Silt	0–30, 0–60, 0–90 cm	Pipette	Soil Survey Staff (2014) 3A1‡
Cation exchange capacity	0–30, 0–60, 0–90 cm	Ammonium acetate	Soil Survey Staff (2014) 4B1a1a1a1a-b1‡
Total N	0–30, 0–60, 0–90 cm	Dry combustion	Soil Survey Staff (2014) 4H2a1‡
Total carbon (C)	0–30, 0–60, 0–90 cm	Dry combustion	Soil Survey Staff (2014) 4H2a1‡
Total organic C	0–30, 0–60, 0–90 cm	Dry combustion	Nelson and Sommers (1996)
Total inorganic C	0–30, 0–60, 0–90 cm	Difference between Total C and total organic C	
Organic matter	0–30, 0–60, 0–90 cm	Loss-on-ignition	Soil Survey Staff (2014) 5A‡
pH (Salt)	0–30, 0–60, 0–90 cm	pH meter using $CaCl_2$ (0.01 *M*)	Soil Survey Staff (2014) 4C1a2a2‡
pH (Water)	0–30, 0–60, 0–90 cm	pH meter	Soil Survey Staff (2014) 4C1a2a1‡
Bulk Density	0–30, 0–60, 0–90 cm	Core	Soil Survey Staff (2014) 3B6a‡

† Daily temperature and precipitation measured using HOBO weather stations instrumentation (Onset Computer Corporation, Bourne, MA).
‡ Indicates the method code associated with the Kellogg Soil Survey Laboratory Methods Manual (Soil Survey Staff, 2014).

Table 3
Variable inputs for all algorithm models to modify three N recommendation tools. Within the table, ✓ indicates parameters used for modeling and X indicates parameters that were removed due to multicollinearity issues, and NA indicates not applicable.

	Reduced Dataset		
Parameter	Farmer's NR	Tri-State YG	Canopy Reflectance Sensing
Weather			
Total precipitation (Planting)†	X	X	X
Total precipitation (SD)‡	NA	NA	✓
Corn heat units (Planting)	X	X	X
Corn heat units (SD)	NA	NA	✓
Growing degree days (Planting)	✓	✓	✓
Growing degree days (SD)	NA	NA	X
Shannon Diversity Index (Planting)	✓	✓	✓
Shannon Diversity Index (SD)	NA	NA	✓
Abundant and Well Distributed Rainfall (Planting)	✓	✓	✓
Abundant and Well Distributed Rainfall (SD)	NA	NA	X
Soil			
Clay	✓ (0–90 cm)	✓ (0–90 cm)	✓ (0–90 cm)
Sand	✓ (0–90 cm)	✓ (0–90 cm)	X
Silt	X	X	✓ (0–60 cm)
Cation exchange capacity	X	X	X
Total N	X	X	X
Total carbon (C)	✓ (0–90 cm)	✓ (0–90 cm)	✓ (0–90 cm)
Total organic C	X	X	X
Total inorganic C	✓ (0–30 cm)	✓ (0–30 cm)	✓ (0–30 cm)
Organic matter	✓ (0–30 cm)	✓ (0–30 cm)	✓ (0–90 cm)
pH (Salt)	X	X	X
pH (Water)	✓ (0–30 cm)	✓ (0–30 cm)	✓ (0–30 cm)
Bulk density	✓ (0–30 cm)	✓ (0–30 cm)	✓ (0–30 cm)

† Planting indicates data used 30 days prior to planting up to the date of planting.
‡ SD indicates data used from the date of planting up to the date of sidedress N application.

properties were then depth weighted across three different depths of 0–0.30, 0–0.60, and 0–0.90 m. Weather data were collected using on-site weather stations (HOBO U30 Automatic Weather Station; Onset Computer Corporation, Bourne, MA). Daily values were calculated for the maximum and minimum temperature and precipitation. These values were then used to calculate total precipitation, growing degree days, corn heat units, Shannon diversity index of precipitation (SDI), and abundantly and well-distributed rainfall (SDI multiplied by total precipitation) as described by Tremblay et al. (2012), in increments of either 30 days before planting to the time of planting or from the date of planting to the time of sensing (Table 2).

3.4. Evaluating statistical and machine learning algorithms

Each ML algorithm was evaluated using three metrics: (1) the out-of-sample RMSE, (2) the performance of each N recommendation tool after incorporating soil and weather information, and (3) an assessment of tool improvement relative to the number of variables used in the model.

To calculate the first metric, a total of 50 RMSE values were computed using the optimized tuning parameters on the 50 testing datasets created with the previously mentioned cross-validation process. The same cross-validation folds were used for all modeling scenarios to compare across all of the ML algorithms' performance. To determine significant differences between ML algorithms, an analysis of variance was conducted using the 50 RMSE values as the response variable and the ML algorithm and N recommendation tool (Farmer's NR, Tri-State YG, canopy reflectance sensing) as the explanatory variables. Significant means separation between ML algorithms was determined using a Tukey Honest Significant Difference test ($\alpha = 0.05$).

For the second evaluation metric, a newly adjusted N recommendation was created by taking the original tool's recommendation and subtracting the predicted values generated using the final model's parameters. For each tool, adjustments were repeated for each modeling scenario resulting in 30 newly adjusted N recommendation tools [8 ML algorithms × 2 data processing scenarios × 2 interaction types (excluding 2 interaction scenarios)]. Performance of unadjusted and adjusted tools was evaluated by examining their accuracy for predicting EONR. Accuracy was measured using both the coefficient of determination and RMSE. The coefficient of determination was calculated using a simple linear regression model based on the observed (measured EONR) and predicted (adjusted tools).

$$r^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (5)$$

where y_i was the observed EONR value, $\hat{y}_i$ was the adjusted tool's predicted value, and $\bar{y}$ was the mean of EONR values. The RMSE was calculated based on the difference between the adjusted tool's recommendations rates and EONR values.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (6)$$

For the third metric, to compare the performance of each ML algorithm for optimizing the three N recommendation tools while also maintaining a simple model, a RMSE Improvement Index was calculated.

$$\text{RMSE Improvement Index} = \frac{RMSE_{unadjusted} - RMSE_{adjusted}}{k} \quad (7)$$

where the $\text{RMSE}_{\text{unadjusted}}$ was derived from the unadjusted N recommendation tool's performance, the $\text{RMSE}_{\text{adjusted}}$ was derived from the adjusted N recommendation tool's performance, and k was the number of important variables in the final model used to adjust the N recommendation tool. Important variables were identified using the varImp function in the R 'caret' package. The methods used to determine which variables were important varied for each ML algorithm. For regression-based algorithms (stepwise, Lasso, ridge regression, and elastic net) the varImp function calculates the absolute value of the t-statistic for each parameter in the model, with higher t-statistic values indicating greater importance. For the PCR a loess smoother was fit for each predictor and the outcome, and an r^2 was calculated from this model compared to the intercept (i.e., the null model) with higher r^2 values indicating greater importance. While for PLSR, the varImp function is based on the weighted sums of the absolute regression coefficients. The decision tree used a reduction in the loss function [e.g., mean square error], which was calculated for each variable at each split and summed. Variables with the least mean square error were the most important and were used in the final model. The random forest used the mean square error for both the "out-of-bag" prediction accuracy for each tree constructed and the "out-of-bag" prediction accuracy for each predictor variable permuted. The differences between the tree and predictor variable out-of-bag mean square errors were averaged and normalized using the standard error. Additional details can be found in the caret vignette under "15.1 Model Specific Metrics" (Kuhn, 2017). A discussion of which variables were most important was not included in this manuscript but is addressed in another manuscript (Ransom, 2018).

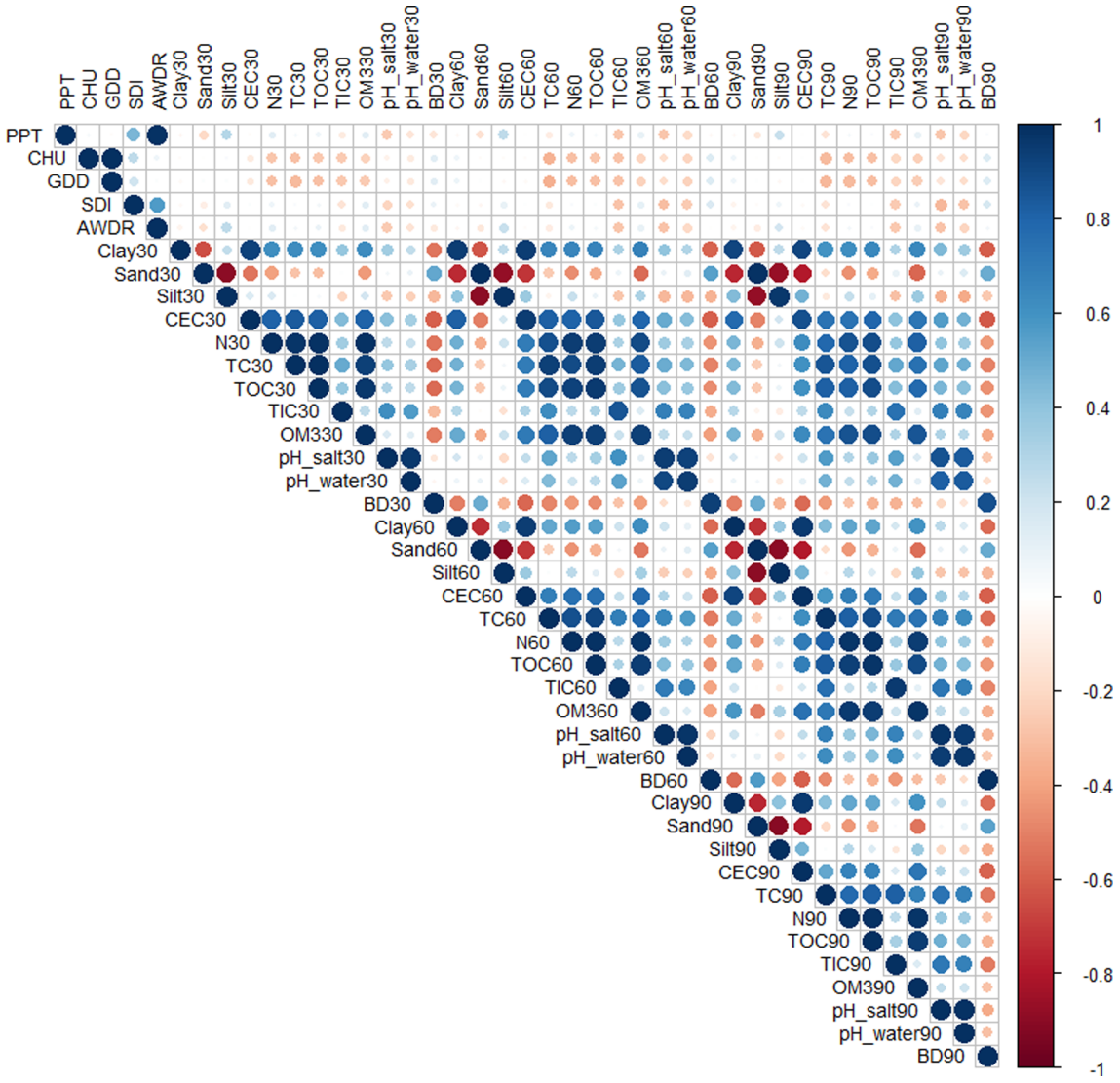

Fig. 1. Correlation matrix of variables used by algorithms to adjust tools used for the at-planting N fertilizer recommendation. The size and color of the circle indicate the sign and magnitude of the correlation between variables. Weather variables calculated from 30 days prior to planting up to the date of planting included cumulative precipitation (PPT), corn heat units (CHU), growing degree days (GDD), Shannon diversity index of precipitation (SDI), and abundant and well-distributed rainfall (AWDR). Soil variables included texture, cation exchange capacity (CEC), total nitrogen (N), total carbon (TC), total organic carbon (TOC), total inorganic carbon (TIC), organic matter (OM), pH in water and in salt (0.01 *M* $CaCl_2$), and bulk density (BD). Each soil measurement was averaged over three separate depth increments of 0–30, 0–60, and 0–90 cm as indicated by the numbers (30, 60, and 90) following the soil variable name. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Results and discussion

4.1. How well did the statistical and machine learning algorithms perform?

The first step in answering this question was to determine if there was a significant statistical difference between ML algorithms and tool type. An analysis of variance using the out-of-sample errors as a function of the ML algorithm and tool types resulted in a significant two-way interaction ($Pr < 0.001$). As such, each of the three tools is displayed separately in Figs. 3–5. Of note, the stepwise regression using the complete dataset resulted in out-of-sample RMSE values as much as 18 times greater than any other ML modeling scenario. As such, this model was not included in the analysis of variance in order to better determine significance between all other ML algorithm types and modeling scenarios. For the Farmer's NR, only the decision tree using the reduced dataset was found to have significantly lower out-of-sample RMSE values than the stepwise regression (Fig. 3). All other ML algorithm types and modeling scenarios for the Farmer's NR were not significantly different from each other. For the Tri-State YG, the biggest differences observed were that the decision tree and random forest resulted in some modeling scenarios with significantly lower out-of-sample RMSE values than the stepwise and some of the ridge regression scenarios (Fig. 4). Apart from these differences all other ML algorithm type and modeling scenarios did not have significantly different values. For the canopy reflectance sensing, there were no significant differences in the out-of-sample RMSE values between any of the ML algorithms (Fig. 5).

4.2. Tool improvement

Based on the out-of-sample RMSE values, one would expect that any one of the ML algorithms (except stepwise using the complete dataset) could satisfactorily improve any of the N recommendation tools. However, when evaluating these ML algorithms on the merit of improving the N recommendation tools' performance, there was a range in how well the ML algorithms performed. The random forest model

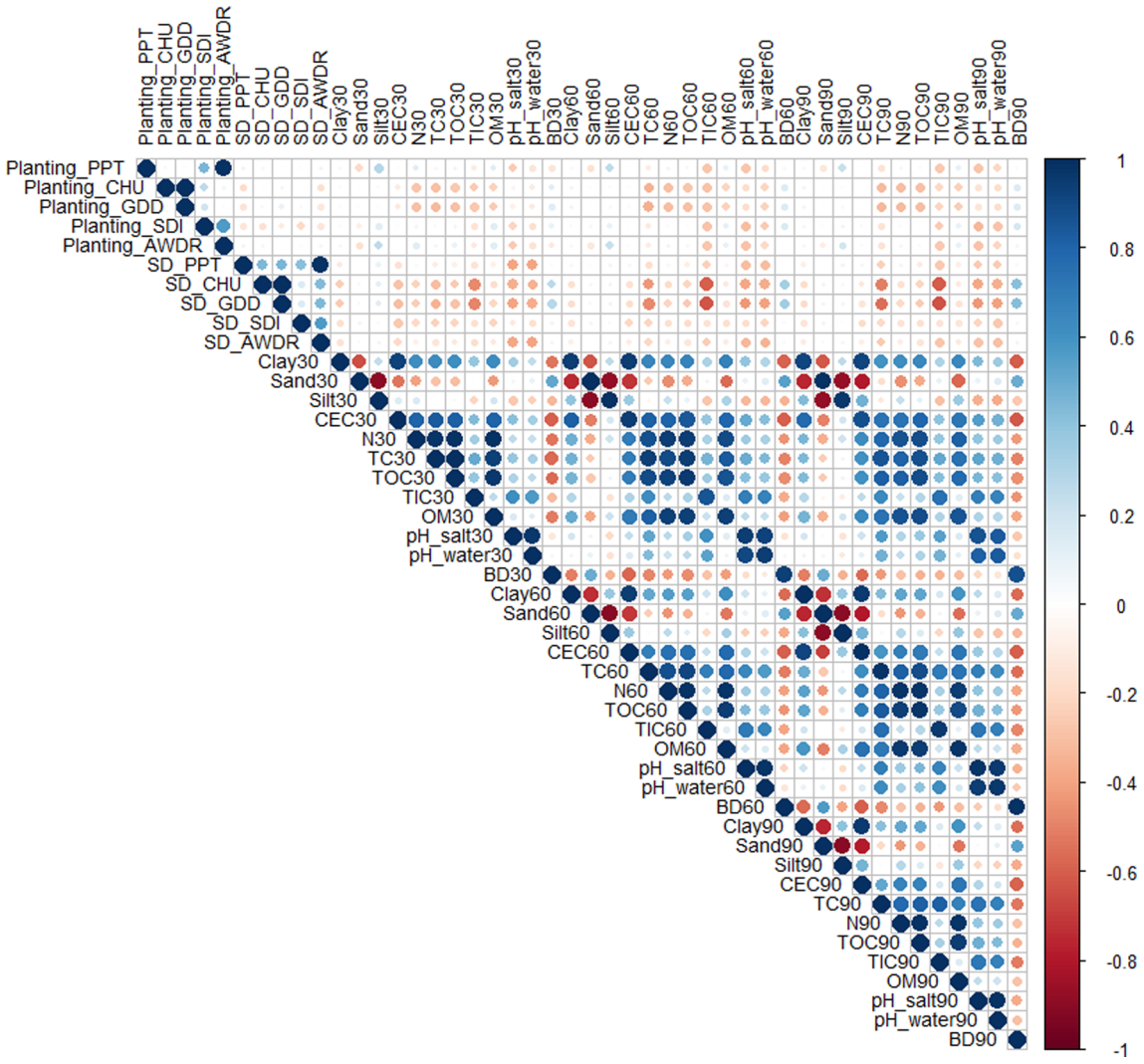

Fig. 2. Correlation matrix of variables used by algorithms to adjust tools used for the split N fertilizer application. The size and color of the circle indicate the sign and magnitude of the correlation between variables. Weather variables calculated from 30 days prior to planting up to the date of planting (planting) and from the date of planting to the date of a sidedress N fertilizer application (SD). They included cumulative precipitation (PPT), corn heat units (CHU), growing degree days (GDD), Shannon diversity index of precipitation (SDI), and abundant and well-distributed rainfall (AWDR). Soil variables included texture, cation exchange capacity (CEC), total nitrogen (N), total carbon (TC), total organic carbon (TOC), total inorganic carbon (TIC), organic matter (OM), pH in water and in salt (0.01 M $CaCl_2$), and bulk density (BD). Each soil measurement was averaged over three separate depth increments of 0–30, 0–60, and 0–90 cm as indicated by the numbers (30, 60, and 90) following the soil variable name. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
The type and range of tuning parameters used for optimizing each algorithm. Where 'ncomp' is the number of components, 'complexity parameter' is a pruning parameter, 'mtry' is the number of variables randomly sampled as candidates at each split, and 'tune length' is the number of levels used by each tuning parameter.

Algorithm	Tuning Parameters
Stepwise	None
Ridge Regression	Lambda = 0.001–25 (in increments of 0.25) Alpha = 0
Lasso	Lambda = 0.001–25 (in increments of 0.25) Alpha = 1
Elastic Net	Lambda = 0.001–25 (in increments of 0.25) Alpha = 0.1–1 (in increments of 0.009)
Principal Component Regression (PCR)	ncomp (tune length = 15)
Partial Least Squares Regression (PLSR)	ncomp (tune length = 15)
Decision Tree	Complexity parameter (tune length = 10)
Random Forest	mtry (tune length = 15)

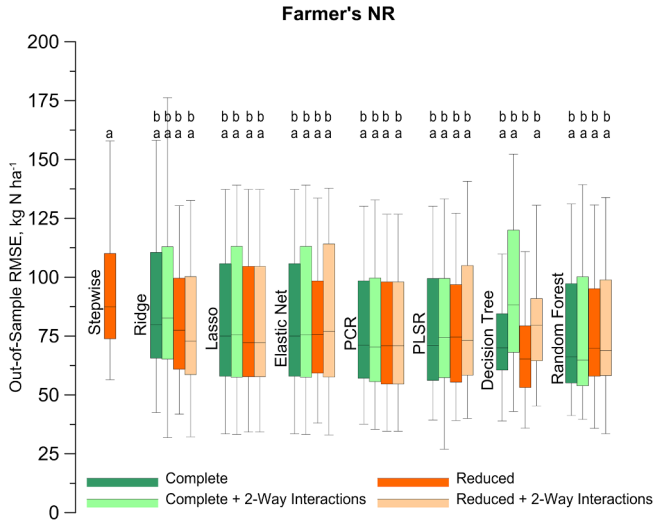

Fig. 3. The out-of-sample errors associated with each fold from 5 × 10 fold cross-validation (totaling 50 RMSE values for each model). Each of the eight models types was used with a complete and reduced dataset (multicollinearity removed) with and without 2-way interactions. The eight models include (1) stepwise linear regression, (2) Ridge regression parameter penalization, (3) Least absolute shrinkage and selection operator (Lasso), (4) Elastic Net, (5) Elastic Net and principal component regression (PCR), (6) partial least square regression (PLSR), (7) decision tree, and (8) random forest. The RMSE values were calculated for the Farmer's N recommendation system. Limits of the box indicate the 1st and 3rd quartile and whiskers indicate 1.5 × IQR. The significance between models is noted by lower case letters, using Tukey's Honest Significant Difference test (α = 0.05).

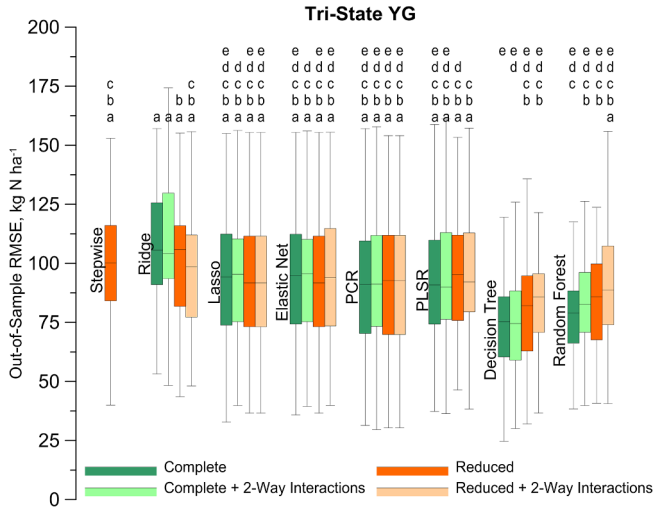

Fig. 4. The out-of-sample errors associated with each fold from 5 × 10 fold cross-validation (totaling 50 RMSE values for each model). Each of the eight models types was used with a complete and reduced dataset (multicollinearity removed) with and without 2-way interactions. The eight models include (1) stepwise linear regression, (2) Ridge regression parameter penalization, (3) Least absolute shrinkage and selection operator (Lasso), (4) Elastic Net, (5) Elastic Net and principal component regression (PCR), (6) partial least square regression (PLSR), (7) decision tree, and (8) random forest. The RMSE values were calculated for the Tri-State yield goal N recommendation tool. Limits of the box indicate the 1st and 3rd quartile and whiskers indicate 1.5 × IQR. The significance between models is noted by lower case letters, using Tukey's Honest Significant Difference test (α = 0.05).

showed the most improvement for predicting EONR, regardless of the modeling scenario. Adjusting tools with the random forest improved the relationship with EONR; resulting in r^2 values ≥ 0.82 for all N recommendation tools—an increase in r^2 values ≥ 0.72 when compared

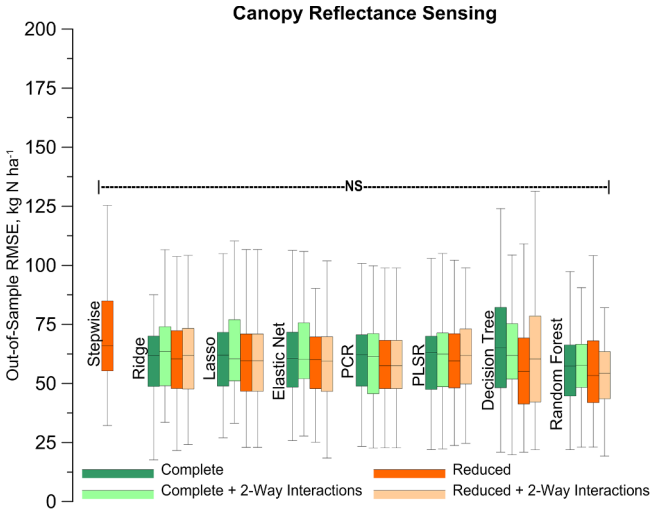

Fig. 5. The out-of-sample errors associated with each fold from 5 × 10 fold cross-validation (totaling 50 RMSE values for each model). Each of the eight models types was used with a complete and reduced dataset (multicollinearity removed) with and without 2-way interactions. The eight models include (1) stepwise linear regression, (2) Ridge regression parameter penalization, (3) least absolute shrinkage and selection operator (Lasso), (4) Elastic Net, (5) Elastic Net and principal component regression (PCR), (6) partial least square regression (PLSR), (7) decision tree, and (8) random forest. The RMSE values were calculated for the canopy reflectance sensing N recommendation tool. Limits of the box indicate the 1st and 3rd quartile and whiskers indicate 1.5 × IQR. The significance between models is noted by lower case letters, using Tukey's Honest Significant Difference test (α = 0.05).

to the unadjusted tools (Table 5). The RMSE values were also some of the lowest, ranging between 33 and 43 kg N ha^{-1}, which was a decrease of ≥41 kg N ha^{-1} compared to the unadjusted tools. The stepwise regression (using the complete dataset only) was the only other ML algorithm that was able to adjust all N recommendation tools similar to the random forest (r^2 value ≥ 0.86 and RMSE between 21 and 41 kg N ha^{-1}). Under conditions where multicollinearity was removed, the stepwise regression performance was similar to most other ML algorithms. The large discrepancy of results between the stepwise using the complete and reduced dataset proved unreliable and over-fit the data; indicating that it would also perform poorly when validated on independent datasets.

The overall performance of tools being adjusted using each of the ML algorithms depends on how well the unadjusted tool was initially related to EONR. For example, in the case of the Farmer's NR, where unadjusted estimates were not related to EONR but on average slightly overestimated EONR by 24 kg N ha^{-1} (Fig. 6), it showed greater improvement when soil and weather variables were incorporated with many of the ML algorithm types compared to the Tri-State YG. Still, not all ML algorithms were able to adequately improve the Farmer's NR relationship with EONR (e.g., Lasso regression showed no improvement). The ML algorithms that did positively affect the Farmer's NR performance (besides stepwise and random forest) still had a weak to a moderate association with EONR (r^2 ≤ 0.55) and a high RMSE value ≥ 55 kg N ha^{-1}.

The Tri-State YG, which had a negative linear relationship with EONR and on average overestimated EONR by 73 kg N ha^{-1} (Fig. 7), showed minimal improvement for the majority of the ML algorithms and modeling scenarios (Table 5). An observed improvement with the Tri-State YG, when adjusted for soil and weather information, was that the predicted N values were no longer negatively related to EONR but in contrast, were also not significantly associated with EONR. When evaluating the Tri-State YG without removing multicollinearity, the ridge regression, decision tree, and random forest were the only ML algorithms able to improve the Tri-State YG to where it had a significant

Table 5
The accuracy of each N recommendation tool compared to EONR that was unadjusted and adjusted with soil and weather variables as determined by each of the eight statistical and machine learning algorithms. Models were evaluated using a complete dataset or reduced dataset (multicollinearity removed) with and without 2-way interactions. The coefficient of determination was measured from a simple linear relationship between each tool and EONR with the corresponding relationship marked in parenthesis: (+) positive linear relationship, (−) negative linear relationship, and no parenthesis after the r^2 value is non-significant. The RMSE was calculated from the difference between a tool's N recommendation and EONR with units of kg N ha^{-1}.

	Complete Dataset						Reduced Dataset					
	Farmer's NR		Tri-State YG		Canopy Reflectance Sensing		Farmer's NR		Tri-State YG		Canopy Reflectance Sensing	
	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
Unadjusted	*0.01*	*88*	*0.10 (−)*	*127*	*0.13 (+)*	*85*	*0.01*	*88*	*0.10 (−)*	*127*	*0.13 (+)*	*85*
Stepwise	0.94 (+)	21	0.86 (+)	31	0.95 (+)	41	0.30 (+)	68	0.03	90	0.41 (+)	63
Ridge	0.32 (+)	67	0.08 (+)	83	0.66 (+)	52	0.27 (+)	70	0.03	89	0.43 (+)	62
+ Interactions	0.47 (+)	60	0.14 (+)	78	0.58 (+)	56	0.23 (+)	71	0.01	91	0.38 (+)	64
Lasso	0.03	82	0.00	92	0.40 (+)	63	0.03	82	0.00	92	0.39 (+)	63
+ Interactions	0.06 (+)	80	0.00	91	0.32 (+)	66	0.04	81	0.00	93	0.44 (+)	61
Elastic Net	0.03	82	0.00	91	0.39 (+)	63	0.24 (+)	71	0.00	92	0.38 (+)	64
+ Interactions	0.06 (+)	80	0.00	91	0.34 (+)	65	0.04	81	0.00	93	0.43 (+)	62
PCR	0.17 (+)	76	0.01	94	0.29 (+)	66	0.20 (+)	74	0.01	95	0.31 (+)	66
+ Interactions	0.04 (+)	78	0.01	95	0.32 (+)	66	0.15 (+)	77	0.00	98	0.29 (+)	67
PLSR	0.19 (+)	74	0.02	93	0.32 (+)	65	0.22 (+)	73	0.02	92	0.38 (+)	63
+ Interactions	0.12 (+)	78	0.01	94	0.30 (+)	66	0.20 (+)	74	0.02	94	0.34 (+)	65
Decision Tree	0.49 (+)	58	0.25 (+)	73	0.57 (+)	54	0.49 (+)	58	0.17 (+)	79	0.57 (+)	54
+ Interactions	0.55 (+)	55	0.47 (+)	61	0.59 (+)	55	0.23 (+)	72	0.17 (+)	79	0.64 (+)	50
Random Forest	0.85 (+)	38	0.87 (+)	33	0.90 (+)	42	0.84 (+)	39	0.83 (+)	37	0.90 (+)	43
+ Interactions	0.82 (+)	40	0.86 (+)	35	0.91 (+)	44	0.83 (+)	41	0.82 (+)	38	0.91 (+)	43

and positive linear relationship with EONR.

Unadjusted canopy reflectance sensing N recommendations had a significant positive linear relationship with EONR but underestimated EONR by 49 kg N ha^{-1} (Fig. 8). Nonetheless, this relationship was still improved when using any of the ML algorithms to adjust estimates. Apart from the random forest and stepwise regression, the ridge regression showed the most improvement using the complete dataset (Table 5). All other ML algorithms resulted in noticeable improvements, with r^2 values between 0.24 and 0.44 and RMSE values between 63 and 71 kg N ha^{-1}. When removing multicollinearity, the decision tree was able to improve the canopy reflectance sensing with r^2 values of 0.57 and 0.64 and RMSE values of 54 and 50 kg N ha^{-1} without and with interaction terms, respectively. These ML-based adjustments to the canopy reflectance sensing tool were some of the best when compared to other analyses done with this dataset. Bean et al. (2018) used general linear models to incorporate soil and weather information into three canopy reflectance algorithms resulted in r^2 values ≤ 0.40. Qin et al. (2018) reported that using a ridge regression with a range of weather and soil variables to directly predict EONR resulted in r^2 values ≤ 0.46. Comparably, using ridge regression to adjust the canopy reflectance sensing algorithms to better predict EONR resulted in r^2 values ≤ 0.66 (Table 5). While Qin et al. (2018) did not use reflectance information

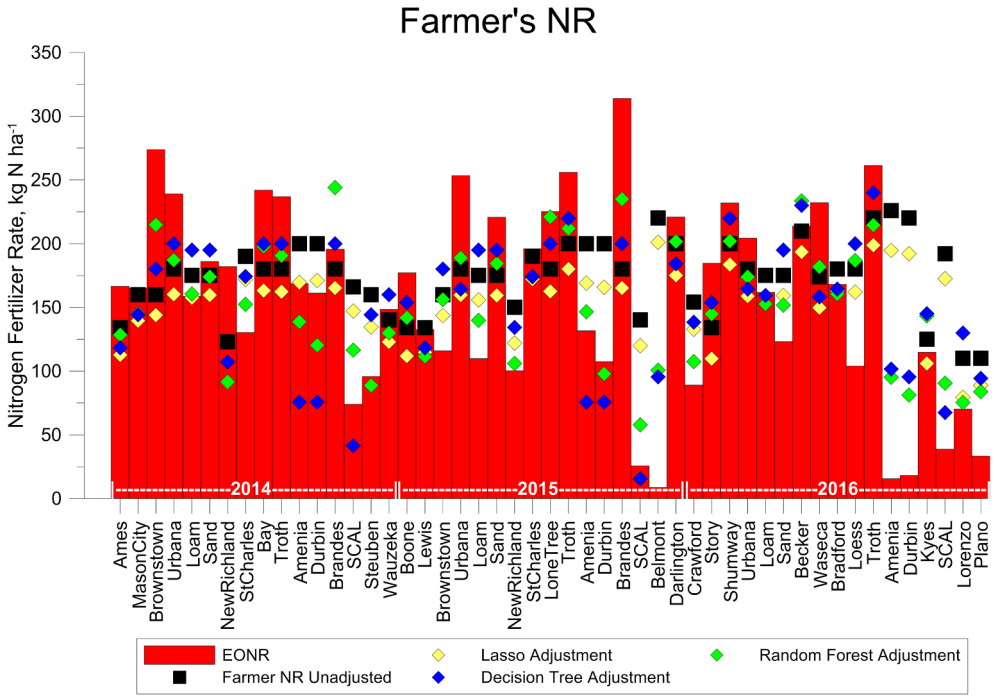

Fig. 6. The unadjusted Farmer's N rate and weather/soil adjusted Farmer's N rate using the reduced dataset with either (1) Lasso regression, (2) decision tree, or (3) random forest all compared to the measured economic optimal N rate (EONR).

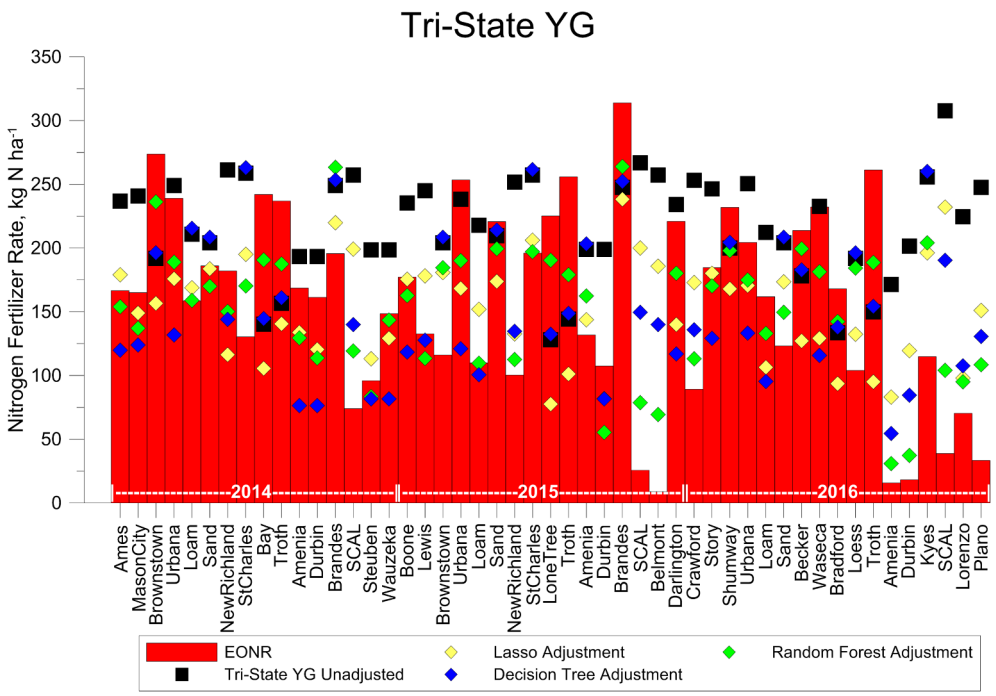

Fig. 7. The unadjusted Tri-State YG's N rate and weather/soil adjusted Tri-State YG's N rate using the reduced dataset with either (1) Lasso regression, (2) decision tree, or (3) random forest all compared to the measured economic optimal N rate (EONR).

(i.e., red, near-infrared, and red-edge, or vegetative indicies) in their ML models to directly predict EONR, efforts to do so did not lead to any additional improvement (data not shown).

While only three tools were tested in this analysis, results indicate that most ML algorithms would best improve tools that already have a significant positive linear relationship with EONR. Nitrogen recommendation tools that are negatively or not linearly related to EONR (e.g., Tri-State YG and Farmer's NR) are only best improved by using random forest or decision tree based approaches.

4.3. Reducing multicollinearity

Ideally, a ML algorithm that can select for variables using a dataset with highly correlated variables, and multiple interaction terms, would minimize the amount of data processing required. The ridge regression, PCR, PLSR, and random forest algorithms have been identified as suitable algorithms when multicollinearity and interaction terms are included in the model (Abdi and Williams, 2010; Geladi and Kowalski, 1986; Grömping, 2009; Lu and Petkova, 2014). However, these ML

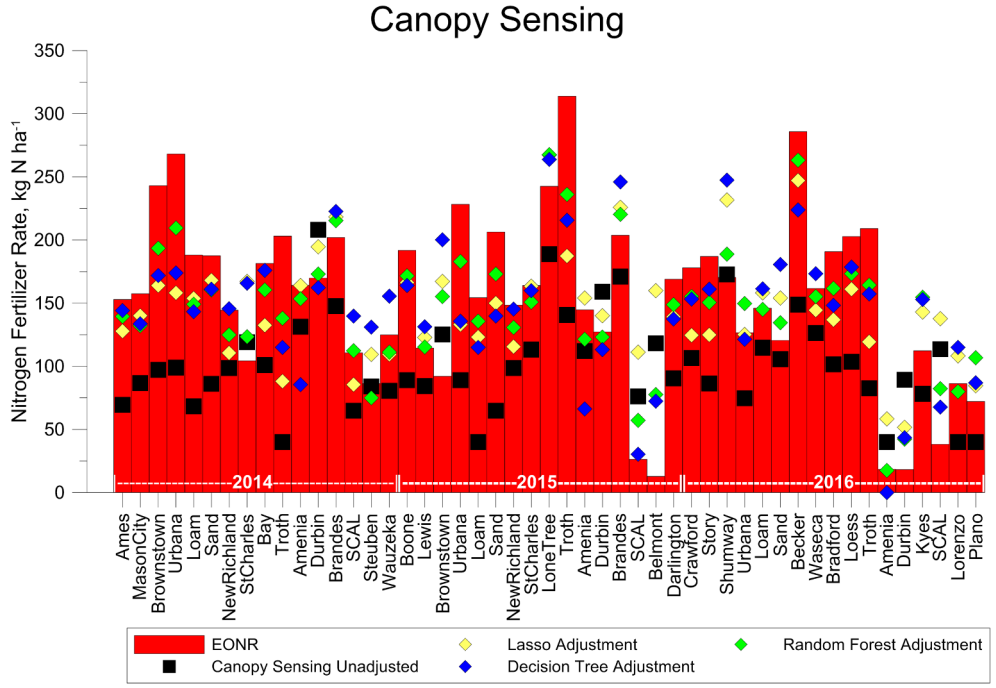

Fig. 8. The unadjusted canopy reflectance N rate and weather/soil adjusted canopy reflectance N recommendation using the reduced dataset with either (1) Lasso regression, (2) decision tree, or (3) random forest all compared to the measured economic optimal N rate (EONR).

Table 6
The number of soil and weather variables identified as important for adjusting three different nitrogen recommendation tools from eight different statistical and machine learning algorithms. Models were fit with and without two-way interactions, terms using either a complete dataset with all soil and weather variables or a reduced dataset with highly correlated variables removed (r > 0.85).

	Complete Dataset			Reduced Dataset		
	Farmer's NR	Tri-State YG	Canopy Reflectance Sensing	Farmer's NR	Tri-State YG	Canopy Reflectance Sensing
	Number of Important Variables			Number of Important Variables		
Stepwise	39	39	44	9	9	12
Ridge	40	40	45	9	9	12
+ Interactions	860	860	1080	54	54	90
Lasso	2	4	7	1	2	5
+ Interactions	2	3	4	3	2	7
Elastic Net	2	7	15	9	2	8
+ Interactions	2	8	12	3	2	7
PCR	39	39	44	9	9	12
+ Interactions	39	39	44	9	9	12
PLSR	40	40	45	9	9	12
+ Interactions	860	860	1080	54	54	90
Decision Tree	2	1	2	2	1	2
+ Interactions	2	2	2	1	1	2
Random Forest	40	40	45	9	9	12
+ Interactions	860	860	1080	54	54	90

algorithms did not produce a parsimonious model for all three N recommendation tools. To improve parsimony for the above mentioned ML algorithms, the number of variables can be prefiltered before modeling. Lowering the number of variables for the ML algorithms to use was successfully done when based on correlation tests. However, while the results were promising, models still contained many more variables than the elastic net and decision tree (Table 6).

When removing multicollinearity (i.e., reduced variables in the dataset), there were some observed changes in performance (explained as a change in $r^2 \geq 0.10$ and/or RMSE $\geq 10\,kg\,N\,ha^{-1}$) between the complete and reduced dataset. This varied by ML algorithm and N recommendation tool. The stepwise (for all tools) and the ridge regression (for canopy reflectance sensing) showed a decrease in performance when using the reduced dataset (Table 5). This is to be expected, as stepwise regression has been found to overestimate actual values and does not perform well with multicollinearity (Whittingham et al., 2006; Zou, 2006). In contrast, the elastic net (for the Farmer's NR) was the only ML algorithms that showed an improved tool performance (change in $r^2 \geq 0.10$ and/or RMSE $\geq 10\,kg\,N\,ha^{-1}$) with the reduced dataset (Table 6). For the remaining modeling scenarios, there was no observed improvement by using a reduced dataset.

4.4. No interactions versus with interactions

There were a few modeling scenarios in which including the interaction terms improved a tool's performance (r^2 increased ≥ 0.10 and/or RMSE decreased ≥ 10). Improvement occurred for the Farmer's NR using the ridge regression without removing multicollinearity and the Tri-State YG using the decision tree with multicollinearity (Table 5). However, there were four instances where including the interaction terms decreased the tool performance (r^2 decreased ≥ 0.10 and/or RMSE increased ≥ 10), but only when first removing multicollinearity. These instances occurred for the elastic net, PCR, and decision tree with the Farmer's NR and the PCR with the canopy reflectance sensing recommendation tool. For the remaining modeling scenarios, there was no change in performance for adjusting N recommendation tools by including the two-way interactions.

The observed decrease in performance with the decision tree is not surprising, as slight changes in the training data often affect this approach, especially with the small number of observations used in this study. The decrease in performance with the elastic net and PCR when including interaction terms is contrary to what has been reported by others (Wang et al., 2011; Wu et al., 2009). However, elastic net, similar to Lasso, could be affected by having too many parameters in the model which leads to including more two-way interactions and higher penalization of more important main effects (Bien et al., 2013). Likewise, PCR has often been found to select variables that are poorly related to the response variable (Abdi and Williams, 2010). With increased terms in the model, the propensity of PCR to select unimportant variables increases.

4.5. Balancing complexity versus accuracy

Determining the ML algorithm that best incorporates soil and weather information into an N recommendation tool depends on the investigatory priority. Machine learning algorithms can be chosen based on either optimizing accuracy and/or resulting in a parsimonious model. Tools that select and utilize numerous variables could improve the accuracy of the model, such as using the random forest. However, from an agronomic management standpoint, ML algorithms that select fewer variables have the advantage of being easier to interpret and more practical to implement. Implementing a model that requires more variables that are not readily available for farmers will increase costs associated with measurement, sampling, and analysis. From this standpoint, using the decision tree, Lasso, or elastic net under any modeling scenario would be preferred due to simplicity and interpretability of the parameters in the model (Table 6). However, as data becomes more readily available to farmers, selecting complex multivariable models will be more feasible in the future.

Evaluating the performance of the ML algorithms needs to take into account both the overall improvement and the number of variables used in the final model. Using the RMSE Improvement Index allows one to evaluate the ML algorithms based on the greatest improvement to accuracy (i.e., the greatest reduction to RMSE) with the fewest variables. Results showed the decision tree approach had the highest RMSE Improvement Index values regardless of the tool be adjusted. Following the decision tree, the random forest, Lasso, and elastic net regression consistently had the next highest values across all three N recommendation tools (Fig. 9). In general, the RMSE Improvement Index values were greater with algorithms using the reduced dataset, though there are a few exceptions (e.g., the decision tree had higher values using the complete dataset with the Farmer's NR and Tri-State YG).

Another factor that needs to be taken into consideration is the difficulty of interpreting the variables used in the model. While regression

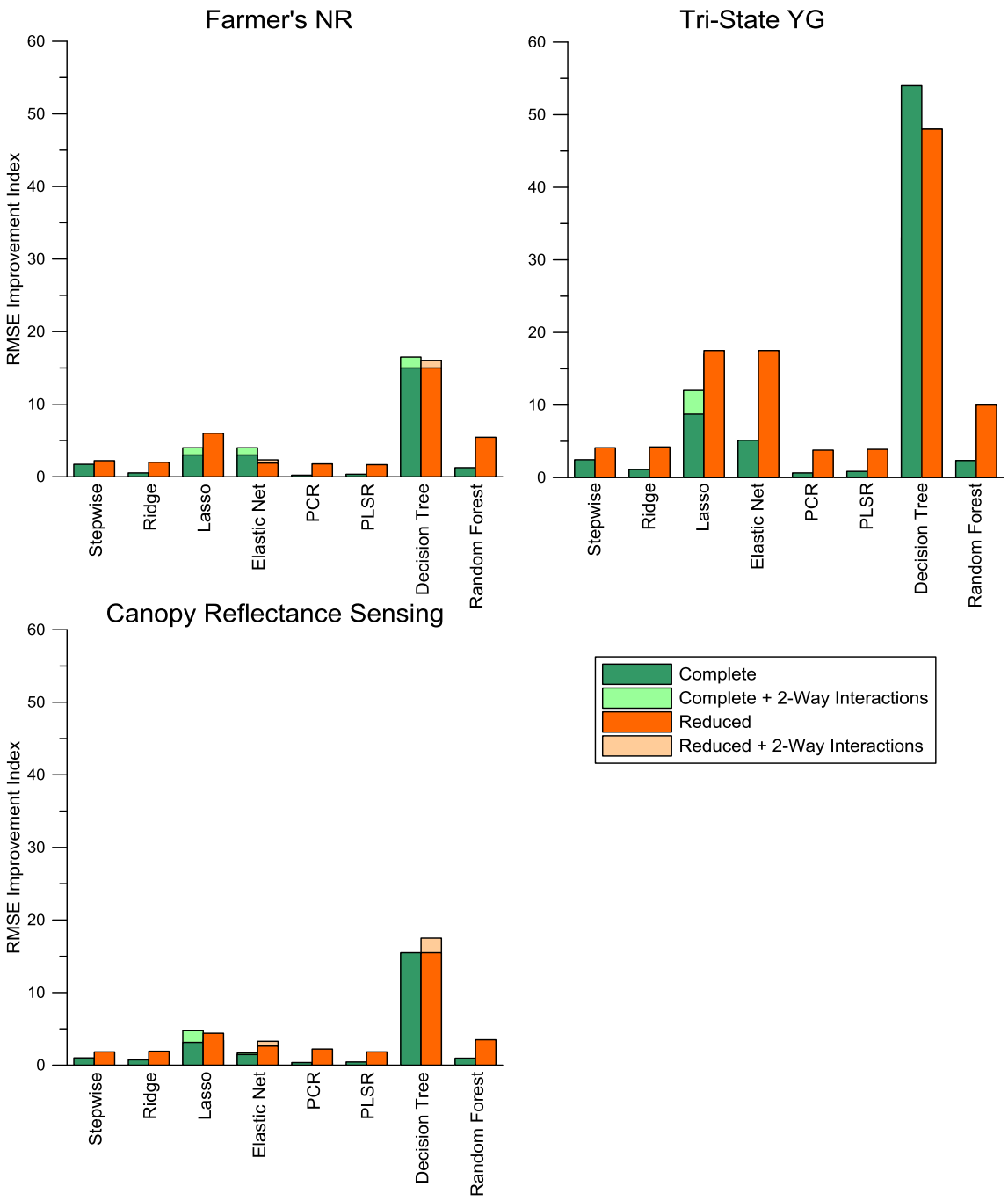

Fig. 9. The RMSE Improvement Index calculated for each N recommendation tool, machine learning algorithm, and modeling scenario based on Eq. (7) calculated using values from Tables 5 and 6.

and decision tree based algorithms can easily be interpreted, other ML algorithms are more complicated and thus require additional steps. For example, results from PCR and PLSR must be displayed in a figure in order to interpret underlining relationships in the data (Mueller et al., 2017; Zuber et al., 2017). Similarly, for the random forest, it is also difficult to interpret which of the variables are the most important. The random forest is able to incorporate complex interaction and highly correlated variables since it uses non-parametric methods to identify relationships between explanatory and response variables (Archer and Kimes, 2008; Strobl et al., 2008). The ability to model complex interactions could be helpful as soil and weather interaction terms have been found to explain EONR or a yield response (Schröder et al., 2000; Shahandeh et al., 2011, 2008; Tremblay et al., 2012; Qin et al., 2018). However, like PCR the interpretability of the random forest is complicated as it is the result of an ensemble of hundreds of decision trees. The random forest can be interpreted by using the most important explanatory variables identified to capture the most variability in the response variable. However, determining the specific effects of each parameter requires additional steps which include complex graphical interpretations or producing and interpreting partial dependence plots (Friedman, 2001; Meinshausen, 2011; Welling et al., 2016, 2015). An easier first step is to filter out unimportant variables selected using another technique (e.g., correlation tests or elastic net). Employing this method did not penalize the accuracy of the random forest.

The focus of this manuscript was not to discuss which parameters were selected by each algorithm and their agronomic importance instead focus on determining the best ML algorithm for improving N recommendation tools. However, the most important variables for

adjusting the tools are listed here with additional discussion found elsewhere (Ransom, 2018). The most important variables varied by ML algorithm and N recommendation tool type. For the majority of the ML algorithms, the most influential variables used to adjust the Farmer's NR included (1) total carbon (0–90 cm), (2) pH (0–60 cm), (3) organic matter (0–30 cm), and (4) the abundant and well-distributed rainfall (AWDR) calculated 30 days prior to planting up to the date of planting. For adjusting the Tri-State YG the most influential variables across all ML algorithms included: (1) organic matter (0–30 cm), (2) pH (0–60 cm), (3) total carbon (0–90 cm), and (4) the AWDR calculated 30 days prior to planting up to the date of planting. For adjusting the canopy reflectance sensing, the most influential variables found to be important across all ML algorithms included: (1) pH (0–60 cm), (2) percent clay (0–90 cm), (3) interaction between pH and clay, (4) bulk density (0–30 cm), (5) the interaction between bulk density and the Shannon Diversity Index (SDI) calculated from the time of planting to the time of sensing, and (6) the SDI calculated 30 days prior to planting up to the time of planting. For each of the N recommendation tools and ML modeling scenario, there were many more soil and weather variables and two-way interactions that helped to improve the models.

5. Conclusions

This study compared eight statistical and ML algorithms as a means for incorporating soil and weather data into three different N recommendation tools. General results showed that the performance of these ML algorithms could be improved slightly by removing multicollinearity. This was the case even for ML algorithms that were not impacted by multicollinearity. However, including two-way interactions in the models did not lead to further improvements. Not all ML algorithms were able to improve all N recommendation tools. Regression-based ML algorithms (e.g., Lasso, elastic net, PCR, and PLSR) did not work well at adjusting tools that were negatively or not significantly linearly related to EONR. The random forest models were best able to improve all three N recommendation tools, however, at the expense of including an extensive number of variables. Pre-filtering variables based on collinearity helped minimize the number of variables in the final model and did not decrease the accuracy of the random forest algorithm regardless of the N recommendation tool. Improved accuracy is a tradeoff of interpretability. With the random forest, one can only know which variables are the most important, not how the variables are used for making a prediction.

On the other hand, a recursive partitioning decision tree is easier to interpret than the random forest because it used only one or two variables in the final model. While it did not perform as well as the random forest it had some of the better adjustments to all three tools compared to other methods. From an agronomic standpoint, the decision tree approach appears to be best suited for improving N recommendation tools with soil and weather information. The ease of computing, interpreting, and performance observed from decision trees in all modeling scenarios demonstrates that researchers should consider utilizing this algorithm, in addition to traditional methods they already use.

There were some promising adjustments made to the studied N recommendation tools but further improvement could be observed by including additional factors known to impact N cycling and EONR such as management decisions, plant genetics, and finer detailed weather and soil information. However, additional site-years of N response trials are required that measure these factors. With larger datasets, other ML algorithms like neural networks could be utilized which allow for more modeling options.

Declaration of Competing Interest

None.

Acknowledgement

This project was funded by DuPont Pioneer (currently named Corteva Agriscience).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2019.104872.

References

Abdel-Rahman, E.M., Ahmed, F.B., Ismail, R., 2013. Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. Int. J. Remote Sens. 34, 712–728. https://doi.org/10.1080/01431161.2012.713142.

Abendroth, L.J., Elmore, R.W., Boyer, M.J., Marley, S.K., 2011. Corn growth and development. Ext. Publ. PM 1009. Ames.

Abdi, H., Williams, L.J., 2010. Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat. 2, 433–459. https://doi.org/10.1002/wics.101.

Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. Comput. Stat. Data Anal. 52, 2249–2260. https://doi.org/10.1016/j.csda.2007.08.015.

Arruda, M.P., Brown, P.J., Lipka, A.E., Krill, A.M., Thurber, C., Kolb, F.L., 2015. Genomic selection for predicting head blight resistance in a wheat breeding program. Plant Genome 8, 1–12. https://doi.org/10.3835/plantgenome2015.01.0003.

Bandura, C., 2017. Agronomic and environmental evaluation of nitrogen rate and timing for Midwestern corn production. Thesis. University of Wisconsin-Madison.

Bean, G.M., Kitchen, N.R., Camberato, J.J., Ferguson, R.B., Fernandez, F.G., Franzen, D.W., Laboski, C.A.M., Nafziger, E.D., Sawyer, J.E., Scharf, P.C., Schepers, J., Shanahan, J.S., 2018a. Active-optical reflectance sensing corn algorithms evaluated over the United States Midwest Corn Belt. Agron. J. 110, 2552–2565. https://doi.org/10.2134/agronj2018.03.0217.

Bean, G.M., Kitchen, N.R., Camberato, J.J., Ferguson, R.B., Fernandez, F.G., Franzen, D.W., Laboski, C.A.M., Nafziger, E.D., Sawyer, J.E., Scharf, P.C., Schepers, J., Shanahan, J.S., 2018b. Improving an active-optical reflectance sensor algorithm using soil and weather information. Agron. J. 110, 2541–2551. https://doi.org/10.2134/agronj2017.12.0733.

Bien, J., Taylor, J., Tibshirani, R., 2013. A lasso for hierarchical interactions. Ann. Stat. 41, 1111–1141. https://doi.org/10.1214/13-AOS1096.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Bundy, L.G., Andraski, T.W., 1995. Soil yield potential effects on performance of soil nitrate tests. J. Prod. Agric. 8, 561–568. https://doi.org/10.2134/jpa1995.0561.

Bundy, L.G., Walters, D.T., Olness, A.E., 1999. Evaluation of soil nitrate tests for predicting corn nitrogen response in the North Central Region. North Central Reg, Publication No. 342. Wisconsin Agricultural Experiment Station, University of Wisconsin, Madison.

Cerrato, M.E., Blackmer, A.M., 1990. Comparison of models for describing corn yield response to nitrogen fertilizer. Agron. J. 82, 138–143. https://doi.org/10.2134/agronj1990.00021962008200010030x.

Dinnes, D.L., Karlen, D.L., Jaynes, D.B., Kaspar, T.C., Hatfield, J.L., 2002. Review and interpretation: nitrogen management strategies to reduce nitrate leaching in tile-drained Midwestern soils. Agron. J. 94, 153–171. https://doi.org/10.2134/agronj2002.0153.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33, 1–22. https://doi.org/10.18637/jss.v068.c02.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Anal. Chim. Acta 185, 1–17. https://doi.org/10.1016/0003-2670(86)80028-9.

Grassini, P., Yang, H., Cassman, K.G., 2009. Limits to maize productivity in Western Corn-Belt: a simulation analysis for fully irrigated and rainfed conditions. Agric. For. Meteorol. 149, 1254–1265. https://doi.org/10.1016/j.agrformet.2009.02.012.

Grömping, U., 2009. Variable importance assessment in regression: linear regression versus random forest. Am. Stat. 63, 308–319. https://doi.org/10.1198/tast.2009.08199.

Hansen, P.M., Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. 86, 542–553. https://doi.org/10.1016/S0034-4257(03)00131-7.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12, 55–67. https://doi.org/10.1080/00401706.1970.10488634.

Holland, K.H., Schepers, J.S., 2010. Derivation of a variable rate nitrogen application model for in-season fertilization of corn. Agron. J. 102, 1415–1424. https://doi.org/10.2134/agronj2010.0015.

Hong, N., Scharf, P.C., Davis, J.G., Kitchen, N.R., Sudduth, K.A., 2007. Economically optimal nitrogen rate reduces soil residual nitrate. J. Environ. Qual. 36, 354–362. https://doi.org/10.2134/jeq2006.0173.

Jagadamma, S., Lal, R., Hoeft, R.G., Nafziger, E.D., Adee, E.A., 2008. Nitrogen

fertilization and cropping system impacts on soil properties and their relationship to crop yield in the central Corn Belt, USA. Soil Tillage Res. 98, 120–129. https://doi.org/10.1016/j.still.2007.10.008.
Kay, B.D., Mahboubi, A.A., Beauchamp, E.G., Dharmakeerthi, R.S., 2006. Integrating soil and weather data to describe variability in plant available nitrogen. Soil Sci. Soc. Am. J. 70, 1210–1221. https://doi.org/10.2136/sssaj2005.0039.
Kitchen, N.R., Shanahan, J.F., Ransom, C.J., Bandura, C.J., Bean, G.M., Camberato, J.J., Carter, P.R., Clark, J.D., Ferguson, R.B., Fernández, F.G., Franzen, D.W., Laboski, C.A.M., Nafziger, E.D., Qing, Z., Sawyer, J.E., Shafer, M., 2017. A public–industry partnership for enhancing corn nitrogen research and datasets: project description, methodology, and outcomes. Agron. J. 109, 2371–2388. https://doi.org/10.2134/agronj2017.04.0207.
Kitchen, N.R., Sudduth, K.A., Drummond, S.T., Scharf, P.C., Palm, H.L., Roberts, D.F., Vories, E.D., 2010. Ground-based canopy reflectance sensing for variable-rate nitrogen corn fertilization. Agron. J. 102, 71–84. https://doi.org/10.2134/agronj2009.0114.
Kuhn, M., 2017. caret: classification and regression training. R Package Version 6.0-76. https://doi.org/10.1126/science.1127647.
Kyveryga, P.M., Blackmer, A.M., Zhang, J., 2009. Characterizing and classifying variability in corn yield response to nitrogen fertilization on subfield and field scales. Agron. J. 101, 269–277. https://doi.org/10.2134/agronj2008.0168x.
Laboski, C.A.M., Peters, J.B., 2012. Nutrient application guidelines for field, vegetable, and fruit crops in Wisconsin. Univ. WI Cooperative Extension A2809.
Lu, F., Petkova, E., 2014. A comparative study of variable selection methods in the context of developing psychiatric screening instruments. Stat. Med. 33, 401–421. https://doi.org/10.1002/sim.5937.
McDonald, G.C., 2009. Ridge regression. Wiley Interdiscip. Rev. Comput. Stat. 1, 93–100. https://doi.org/10.1002/wics.14.
Meinshausen, N., 2011. Partition maps. J. Comput. Graph. Stat. 20, 1007–1028. https://doi.org/10.1198/jcgs.2011.10107.
Moebius-Clune, B., van Es, H., Melkonian, J., 2013. Adapt-N uses models and weather data to improve nitrogen management for corn. Better Crop. 97, 7–9.
Morris, T.F., Murrell, T.S., Beegle, D.B., Camberato, J.J., Ferguson, R.B., Grove, J., Ketterings, Q., Kyveryga, P.M., Laboski, C.A.M.M., McGrath, J.M., Meisinger, J.J., Melkonian, J., Moebius-clune, B.N., Nafziger, E.D., Osmond, D., Sawyer, J.E., Scharf, P.C., Smith, W., Spargo, J.T., van Es, H.M., Yang, H., 2018. Strengths and limitations of nitrogen rate recommendations for corn and opportunities for improvement. Agron. J. 110, 1–37. https://doi.org/10.2134/agronj2017.02.0112.
Mueller, S.M., Camberato, J.J., Messina, C., Shanahan, J., Zhang, H., Vyn, T.J., 2017. Late-split nitrogen applications increased maize plant nitrogen recovery but not yield under moderate to high nitrogen rates. Agron. J. 109, 2689–2699. https://doi.org/10.2134/agronj2017.05.0282.
Nelson, D.W., Sommers, L.E., 1996. Total carbon, organic carbon, and organic matter. p. 961-1010. In: Sparks, D.L., et al. (Ed.) Methods of soil analysis. Part 3 - chemical methods. SSSA Book Series No. 5, SSSA ASA, Madison, WI. https://doi.org/19971902103.
Puntel, L.A., Sawyer, J.E., Barker, D.W., Thorburn, P.J., Castellano, M.J., 2018. A systems modeling approach to forecast corn economic optimum nitrogen rate. Front. Plant Sci. 9, 1–15. https://doi.org/10.3389/fpls.2018.00436.
Qin, Z., Myers, D.B., Ransom, C.J., Kitchen, N.R., Liang, S.-Z., Camberato, J.J., Carter, P.R., Ferguson, R.B., Fernandez, F.G., Franzen, D.W., Laboski, C.A.M., Malone, B.D., Nafziger, E.D., Sawyer, J.E., Shanahan, J.F., 2018. Application of machine learning methodologies for predicting corn economic optimal nitrogen rate. Agron. J. 110, 2596–2607. https://doi.org/10.2134/agronj2018.03.0222.
Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn. 1, 81–106. https://doi.org/10.1023/A:1022643204877.
R Core Team, 2016. R development core team. R: A language and environment for statistical computing. < http://www.R-project.org > .
Ransom, C.J., 2018. Evaluating and improving corn nitrogen fertilizer recommendation tools across the U.S. Midwest. Dissertation. University of Missouri.
Sawyer, J.E., Mallarino, A.P., 2017. Use of the late-spring soil nitrate test in Iowa corn production. Extension and Outreach Publications. 497. https://lib.dr.iastate.edu/extension_pubs/497. (accessed 5 May 2018).
Scharf, P.C., Brouder, S.M., Hoeft, R.G., 2006. Chlorophyll meter readings can predict nitrogen need and yield response of corn in the north-central USA. Agron. J. 98, 655–665. https://doi.org/10.2134/agronj2005.0070.
Scharf, P.C., Kitchen, N.R., Sudduth, K.A., Davis, J.G., Hubbard, V.C., Lory, J.A., 2005. Field-scale variability in optimal nitrogen fertilizer rate for corn. Agron. J. 97, 452–461. https://doi.org/10.2134/agronj2005.0452.
Schmidt, J.P., Dellinger, A.E., Beegle, D.B., 2009. Nitrogen recommendations for corn: an on-the-go sensor compared with current recommendation methods. Agron. J. 101, 916–924. https://doi.org/10.2134/agronj2008.0231x.
Schröder, J.J., Neeteson, J.J., Oenema, O., Struik, P.C., 2000. Does the crop or the soil indicate how to save nitrogen in maize production? Reviewing the state of the art. F. Crop. Res. 66, 151–164. https://doi.org/10.1016/S0378-4290(00)00072-1.
Setiyono, T.D., Yang, H., Walters, D.T., Dobermann, A., Ferguson, R.B., Roberts, D.F., Lyon, D.J., Clay, D.E., Cassman, K.G., 2011. Maize-N: a decision tool for nitrogen management in maize. Agron. J. 103, 1276–1283. https://doi.org/10.2134/agronj2011.0053.
Shahandeh, H., Wright, A.L., Hons, F.M., 2011. Use of soil nitrogen parameters and texture for spatially-variable nitrogen fertilization. Precis. Agric. 12, 146–163. https://doi.org/10.1007/s11119-010-9163-8.
Shanahan, J.F., Kitchen, N.R., Raun, W.R., Schepers, J.S., 2008. Responsive in-season nitrogen management for cereals. Comput. Electron. Agric. 61, 51–62. https://doi.org/10.1016/j.compag.2007.06.006.
Soil Survey Staff, 2014. Kellogg Soil survey laboratory methods manual, soil survey investigations report No. 42, Version 5.0.
Stanford, G., 1973. Rationale for optimum nitrogen fertilizer in corn production. J. Environ. Qual. 2, 159–166. https://doi.org/10.2134/jeq1973.00472425000200020001x.
Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinf. 9. https://doi.org/10.1186/1471-2105-9-307.
Thompson, B., 2001. Significance, effect sizes, stepwise methods, and other issues: strong arguments move the field. J. Exp. Educ. 70, 80–93. https://doi.org/10.1080/00220970109599499.
Thompson, L.J., Ferguson, R.B., Kitchen, N., Franzen, D.W., Mamo, M., Yang, H., Schepers, J.S., 2015. Model and sensor-based recommendation approaches for in-season nitrogen management in corn. Agron. J. 107, 2020–2030. https://doi.org/10.2134/agronj15.0116.
Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc. B. 58, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
Tremblay, N., Bouroubi, Y.M., Bélec, C., Mullen, R.W., Kitchen, N.R., Thomason, W.E., Ebelhar, S., Mengel, D.B., Raun, W.R., Francis, D.D., Vories, E.D., Ortiz-Monasterio, I., 2012. Corn response to nitrogen is influenced by soil texture and weather. Agron. J. 104, 1658–1671. https://doi.org/10.2134/agronj2012.0184.
Vitosh, M.L., Johnson, J.W., Mengel, D.B., 1995. Tri-state fertilizer recommendations for corn, soybeans, wheat and alfalfa. Ext. Bull. E-2567 2567, 1–4.
Waheed, T., Bonnell, R.B., Prasher, S.O., Paulet, E., 2006. Measuring performance in precision agriculture: CART-A decision tree approach. Agric. Water Manag. 84, 173–185. https://doi.org/10.1016/j.agwat.2005.12.003.
Wang, D., Eskridge, K.M., Crossa, J., 2011. Identifying QTLs and epistasis in structured plant populations using adaptive mixed LASSO. J. Agric. Biol. Environ. Stat. 16, 170–184. https://doi.org/10.1007/s13253-010-0046-2.
Wang, J., Shen, C., Liu, N., Jin, X., Fan, X., Dong, C., Xu, Y., 2017. Non-destructive evaluation of the leaf nitrogen concentration by in-field visible/near-infrared spectroscopy in pear orchards. Sensors 17, 538. https://doi.org/10.3390/s17030538.
Welling, S.H., Clemmensen, L.K.H., Buckley, S.T., Hovgaard, L., Brockhoff, P.B., Refsgaard, H.H.F., 2015. In silico modelling of permeation enhancement potency in Caco-2 monolayers based on molecular descriptors and random forest. Eur. J. Pharm. Biopharm. 94, 152–159. https://doi.org/10.1016/j.ejpb.2015.05.012.
Welling, S.H., Refsgaard, H.H.F., Brockhoff, P.B., Clemmensen, L.H., 2016. Forest floor visualizations of random forests. *arXiv preprint arXiv:1605.09196*.
Whittingham, M.J., Stephens, P.A., Bradbury, R.B., Freckleton, R.P., 2006. Why do we still use stepwise modelling in ecology and behaviour? J. Anim. Ecol. 75, 1182–1189. https://doi.org/10.1111/j.1365-2656.2006.01141.x.
Wu, T.T., Chen, Y.F., Hastie, T., Sobel, E., Lange, K., 2009. Genome-wide association analysis by lasso penalized logistic regression. Bioinformatics 25, 714–721. https://doi.org/10.1093/bioinformatics/btp041.
Xing, L., Pittman, J.J., Inostroza, L., Butler, T.J., Munoz, P., 2018. Improving predictability of multisensor data with nonlinear statistical methodologies. Crop Sci. 58, 972–981. https://doi.org/10.2135/cropsci2017.09.0537.
Yamashita, T., Yamashita, K., Kamimura, R., 2007. A stepwise AIC method for variable selection in linear regression. Commun. Stat. - Theory Methods 36, 2395–2403. https://doi.org/10.1080/03610920701215639.
Zhao, P., Yu, B., 2006. On model selection consistency of Lasso. J. Mach. Learn. Res. 7, 2541–2563. https://doi.org/10.1109/TIT.2006.883611.
Zhu, Q., Schmidt, J.P., Lin, H.S., Sripada, R.P., 2009. Hydropedological processes and their implications for nitrogen availability to corn. Geoderma 154, 111–122. https://doi.org/10.1016/j.geoderma.2009.10.004.
Zou, H., 2006. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc. 101, 1418–1429. https://doi.org/10.1198/016214506000000735.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B (Statistical Methodology) 67, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.
Zuber, S.M., Behnke, G.D., Nafziger, E.D., Villamil, M.B., 2017. Multivariate assessment of soil quality indicators for crop rotation and tillage in Illinois. Soil Tillage Res. 174, 147–155. https://doi.org/10.1016/j.still.2017.07.007.