Field Crops Research 260 (2021) 107985

Contents lists available at ScienceDirect

Field Crops Research

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

Assessing the uncertainty of maize yield without nitrogen fertilization

Adrian A. Correndo [a,*], Jose L. Rotundo [b], Nicolas Tremblay [c], Sotirios Archontoulis [d], Jeffrey A. Coulter [e], Dorivar Ruiz-Diaz [a], Dave Franzen [f], Alan J. Franzluebbers [g], Emerson Nafziger [h], Rai Schwalbert [a], Kurt Steinke [i], Jared Williams [j], Charlie D. Messina [b], Ignacio A. Ciampitti [a,*]

[a] Department of Agronomy, Kansas State University, 1712 Claflin Rd., Manhattan, KS, 66506, USA
[b] Corteva Agriscience, 7250 NW 62nd Ave., Johnston, IA, 50310, USA
[c] Agriculture and Agri-Food Canada, Horticultural R&D Center, St-Jean-sur-Richelieu, J3B 3E6, Canada
[d] Department of Agronomy, Iowa State University, Agronomy Hall, Ames, IA, 50011, USA
[e] Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, Minnesota, 55108, USA
[f] Department of Soil Science, School of Natural Resource Sciences, North Dakota State University, Dept. 7180, PO Box 6050, Fargo, ND, 58108, USA
[g] USDA-Agricultural Research Service, 101 Derieux Place, NCSU Campus Box 7620, Raleigh, NC, 27695, USA
[h] Department of Crop Sciences, University of Illinois, 1102 S. Goodwin Ave., Urbana, IL, 61821, USA
[i] Department of Plant, Soil, and Microbial Sciences, Michigan State University, East Lansing, MI, 48824, USA
[j] Department of Applied Plant Science, Brigham Young University-Idaho, 310K STC, Rexburg, ID, 83460, USA

A R T I C L E I N F O

Keywords:
Maize
Soil nitrogen supply
Yield forecast

A B S T R A C T

Maize (*Zea Mays* L.) yield responsiveness to nitrogen (N) fertilization depends on the yield under non-limiting N supply as well as on the inherent productivity under zero N fertilizer (Y_0). Understanding the driving factors and developing predictive algorithms for Y_0 will enhance the optimization of N fertilization in maize. Using a random forest algorithm, we analyzed data from 679 maize N fertilization studies (1031 Y_0 observations) conducted between 1999–2019 in the United States and Canada. Predictability of Y_0 was assessed while identifying determinant factors such as soil, crop management, and weather. The inclusion of weather variables as predictors improved the model efficiency (ME) from 51 up to 64 %, and reduced the root mean square error (RMSE) from 2.5 to 2.0 Mg ha^{-1}, 34 to 27 % in relative terms (RRMSE). The most relevant predictors of Y_0 were previous crop, irrigation, and soil organic matter (SOM), while the most influential weather data was linked to the radiation per unit of thermal time (Q quotient) around flowering and spring precipitations. The crop rotation effect resulted in Alfalfa (*Medicago sativa* L.) as the previous crop with the highest Y_0 level (IQR = 11.5–15.0 Mg ha^{-1}) as compared to annual legumes (IQR = 5.6–10.0 Mg ha^{-1}) and other previous crops (IQR = 3.6–7.8 Mg ha^{-1}). The Q quotient around flowering positively affected Y_0, while spring precipitations and extreme temperature events during grain filling showed a negative association to Y_0. Overall, these results reinforce the concept that yields are controlled not only by soil N supply but also by factors modifying plant demand and ability to capture N. Lastly, we foresee a promising future for the use of machine learning to address both prediction and interpretation of maize yield to obtain more reliable N guidelines.

1. Introduction

Decades of research on yield response to N application has not yet produced accurate algorithms to issue N recommendations for maize in North America. Addressing the uncertainty on N needs for maize (*Zea mays* L.) is still a major concern (Morris et al., 2018; Raun et al., 2019) because of the unintentional impacts of misuse of N and low N use efficiency (Sela et al., 2018a, 2018b). Estimations of N recovery efficiency in the region are typically below 50 % of the applied N, which may reflect a higher uptake efficiency from indigenous sources (soil) than for applied fertilizer (Cassman et al., 2002). This scenario is linked to the complex process of fertilizer N losses such as leaching, denitrification, and volatilization (Baker and Johnson, 1981; Francis et al., 1993; Bowles et al., 2018). Despite genetic improvement for N use efficiency

* Corresponding authors.
E-mail addresses: correndo@ksu.edu (A.A. Correndo), ciampitti@ksu.edu (I.A. Ciampitti).

https://doi.org/10.1016/j.fcr.2020.107985
Received 25 July 2020; Received in revised form 5 October 2020; Accepted 13 October 2020
Available online 3 November 2020
0378-4290/

(Mueller et al., 2019) there are further opportunities to develop prescription algorithms to improve N management and fertilizer recommendations.

For most of the twentieth century, N recommendations in North America have been mostly based on estimation of yield and production goals (Stanford et al., 1966; 1973), that is the N demand dictated the amount of N to be added as fertilizer after the estimation of a simplified N balance that considered N credits and other subtractions and additions (Morris et al., 2018). Refined N guidelines for maize has been addressed following different systems over time and across states (Heady and Pesek, 1954; Bundy and Andraski, 1995; Scharf et al., 2005; Kyveryga et al., 2007; Kitchen et al., 2010; Setiyono et al., 2011; Wortmann et al., 2011; Yost et al., 2014; Sindelair et al., 2015). Lory and Scharf (2003) have described an approach using delta yield, as the yield difference between non-N-limited and non-N-fertilized plots (Y_0), assuming the latter as a proxy of indigenous soil N supply (Cassman et al., 1996). More recently, utilizing a large database of N response trials, the "maximum return to N (MRTN) recommendation system represented an approach to adjust estimations of the economic optimum N rate (EONR) grouping response functions according to several factors of interest including management and soil features (Sawyer et al., 2006). Likewise, the integration of multiple site-years expanding combinations of soil, crop management and weather scenarios, might lead to the use of complementary predictive models (*e.g.*, supervised learning techniques) with more focus on forecasting the N needs for maize crop rather than an *ex-post* analysis.

The dissection of the yield response to N can inform decisions to manage a complex system such as the one governing the soil-plant N dynamics. For a given site-year, we may depict the Y_0 as the intercept of the function that along with non-N-limited yield (plateau) defines a yield response to N fertilization for a given curvature. Thus, defining realistic expectations for EORN predictions will inevitably rely on accurate predictions of Y_0. Recent attempts to address the problem of forecasting yield response to N have been pursued with limited datasets that restrict our inference space (Puntel et al., 2019) or used yield simulations that restrict the inference to the set of parameters and model assumptions (Shahhosseini et al., 2019; Archontoulis et al., 2020). Yield under non-N-limiting scenario is largely determined by temperature and solar radiation (van Ittersum et al., 2013) and it is adequately captured within dynamic crop growth model frameworks (Monteith, 1972; Messina et al., 2009). In contrast, soil processes governing N cycling and its interactions with the plant and environment system are complex and less well represented in models. Predicting N deficiency level and Y_0 poses a much difficult problem to solve than non-N-limited yield (Puntel et al., 2018; Archontoulis et al., 2020), in particular for experiments conducted in small plots (Tao et al., 2018). The combination of mechanistic models for predicting non-N-limited yield and data-driven machine learning models for predicting Y_0 could open up opportunities to increase the predictability of complex systems (Messina et al., 2020).

Methodologically, science is entering an entirely new phase that involves data-intensive practices (Tolle et al., 2011). Machine learning is one method, laying at the intersection of computer science and statistics (Jordan and Mitchell, 2015) useful to identify repeatable patterns in large datasets. Belonging to the family of supervised learning techniques, tree-based methods such as decision trees, boosting and random forest (RF) are robust and versatile techniques as demonstrated in remote sensing applications (Belgiu and Drăgut, 2016; Schwalbert et al., 2018) and more recently in agriculture (Khaki and Wang, 2019; Ramanantenasoa et al., 2019). For forecasting purposes, a minimum set of candidate predictors including as early as possible metadata during the crop growing season is desirable. Since most substantial uncertainties are inherent to weather, with very limited predictability beyond 10−15 days (Stern and Davidson, 2015; Zhang et al., 2019), then those variables are the main candidates to perform a sensitivity analysis. A model with no-weather, assuming it as completely unknown and stochastic, may serve as a reference prediction framework to later assess the value of adding weather information. On the other hand, spring weather is likely to be known by the time of planting and including weather predictors may be useful in forecasting applications for N availability in production fields (Puntel et al., 2016). Lastly, defined seasonal weather patterns could serve as model limits.

The main goal of this work is to describe properties of Y_0 on a large database of maize fertilization studies performed in the United States and Canada, an develop a prediction model with potential to improve N management systems. The specific goals for this manuscript are to i) rank and identify the main soil, management and weather features impacting Y_0, and ii) assess the prediction performance of different frameworks involving soil and management factors but varying the inclusion of weather features: a) no weather variables; b) spring weather known around planting; and c) weather known for the entire crop growing season.

2. Material and methods

2.1. Data collection

A database was built through meeting certain requirements as follows: i) experiments performed during the last two decades (1999–2019) in order to reduce the noise related to different hybrids eras (Woli et al., 2016); ii) only replicated field trials having N treatments either on small plots or strip-plots; iii) absolute yield data reported for the zero-N control treatment; iv) top-soil analysis results and/or soil series reported; v) data of previous crop and tillage system; vi) latitude and longitude coordinates, or nearest town reported in order to retrieve weather and missing soil data; vii) starter-N and manure treatments were excluded to minimize confounding effects; and viii) general crop management (*e.g.,* planting date, row spacing, other nutrients, weed and pest management) was assumed to have been set to maximize yield under each site-specific condition. Published manuscripts were the first source of data through an engine-search in Web of Science® filtering by the following keywords: "corn/maize" and "nitrogen fertilizer" or "nitrogen fertilization" and "United States" and/or "Canada". In order to reduce publication bias effect (Dickersin and Min, 1993), unpublished data (*e.g.*, dissertations, field reports, unpublished experiments) were also included in the database as long as they met the established criteria. After filtering and selection processes, 679 site-years resulting in 1031 treatments of maize without N fertilizer were gathered from 59 different data sources, including published and unpublished studies (Supp. Table 1).

2.2. Data analysis

2.2.1. Response and explanatory variables

Yield that resulted from treatments receiving zero-N application (Y_0, Mg ha^{-1}) was used as the response variable in the analysis. Grain yield was standardized at a water content of 155 g kg^{-1}. Average values (3–5 replications) were considered as an unbiased central tendency-values of Y_0.

A set of weather, soil, and crop management variables were considered as explanatory variables, predictors or features. Soil related variables were topsoil (0−15 cm) soil organic matter (SOM, %) and soil texture (clay, silt and sand, %). Soil data were collected from original sources, accessed from authors' records when not reported in manuscripts, or retrieved from gridded POLARIS soil data engine (Chaney et al., 2016), a raster optimization based on SSURGO data with a spatial resolution of 1 km^2. When SOM data were reported at 0–20 or 0−30 cm, values were standardized to 0−15 cm using stratification factors based on data from previous research on grain crops trials (Al-Kaisi et al., 2005; Varvel and Wilhelm, 2011; Franzluebbers, 2010; Villamil et al., 2015).

Daily weather data were accessed via the Google Earth Engine platform (Gorelick et al., 2017) using reported latitude-longitude coordinates of the trials or nearest town. With a spatial resolution of 1 km^2, precipitation (PP), temperature (T, ^oC, maximum and minimum) and vapor pressure deficit (vpd, kPa) were obtained from the Parameter-elevation Regressions on Independent Slopes Model -*PRISM*- (Daly et al., 2015); while incident shortwave solar radiation during daylight period –Rad, MJ m^{-2}- plus day-length were retrieved from *Daymet* (Thornton et al., 2018). Weather data were transformed into bi-monthly basis (as sum or average) following Carter et al. (2018a). We divided the weather data into three main periods: i) April-May (AM) as proxy of the early-growth period; ii) June-July (JJ) as proxy of the flowering period; and iii) and August-September (AS) as a proxy of grain filling period. In addition, we also considered PP and mean temperature of March as spring weather with the intent to represent typical weather data accessible to farmers when planting and N fertilizer decisions are made.

A series of additional weather variables were calculated in order to capture environmental differences that might not be captured by analyzing standard weather information. For example, the Shannon Diversity Index (SDI) as described by Tremblay et al. (2012) was included to describe the distribution of PP during each period. Extreme PP events were included as the number of days with precipitations greater than 25 mm as a proxy of excessive rainfall events (Puntel et al., 2019). Crop development was described by crop heat units (CHU; Tremblay et al., 2012). Extreme temperature events (ETE, defined as the number of days with mean maximum temperature over 30 °C) were also included as a proxy of heat stress risk (Butler and Huybers, 2013; Ye et al., 2017). The photo-thermal quotient (Q) was calculated as the ratio between cumulative Rad and CHU, as an indicator of the solar radiation available to the crop per unit of thermal time during each period, related to yield potential (Bannayan et al., 2018).

2.2.2. Prediction models

Three prediction models were tested with models differing in the weather features included:

i a "No-weather" model includes only management and soil features;
ii a "Spring-weather" model includes precipitations and mean temperature during March and April-May as proxy of pre-plant and early vegetative periods; and
iii a "Full-weather" model includes all features from April 1st through September 30th (Table 1). This model is descriptive and enables assessing the relevance of seasonal weather and interactions with soil properties and management on Y_0.

Table 1
Explanatory variables included for the prediction of maize yield under N omission (Y_0). *Periods: *AM = April-May, JJ = June-July, AS = August-September.*

Management		
Previous Crop	alfalfa; annual legumes; others (maize, sorghum, wheat, barley, rye, sunflower)	
Tillage system	Till; no-till	
Irrigation	Irrigated; Rainfed	
	Soil	
Variables	*Units*	*Depth*
SOM = Soil Organic Matter	%	0–15 cm
Clay		
Silt		
Sand		
	Weather	
Variables	*Units*	*Periods**
PP = Precipitations	mm	March, AM, JJ, AS
Tm = Mean Temperature	^oC	
SDI = Shannon Diversity Index	0–1 (uneven - even)	AM, JJ, AS
EPE = Extreme PP Events	# days PP > 25 mm	
vpd = Vapor Pressure Deficit (sum)	KPa	
Rad = Incident radiation (sum)	MJ m^{-2}	
CHU = Crop Heat Units	^oC	
Q = Photothermal quotient	MJ m^{-2} / CHU	
ETE = Extreme Temperature Events	# days T_{max} > 30 °C	JJ, AS

2.2.3. Machine learning algorithm

A tree-based algorithm was selected over other learning alternatives because as a non-parametric tool, it allows constructing prediction rules based on the simultaneous use of categorical and continuous predictors without making prior assumption on normality or on the form of associations with the response variable (Probst et al., 2019). While a single regression tree might be easier to interpret, its prediction power is normally low (and easy to overfit), so it is considered a "weak learner". As an ensemble of trees, the RF is considered as a "strong learner" being much more capable in terms of prediction power (Breiman, 2001). Random Forest is primarily used here for two purposes: i) as a prediction tool, and ii) to assess the relevance of features on prediction.

Among the RF alternatives, we used conditional inference trees to build the ensembles (forests) using the *party* package (Hothorn et al., 2006) for R software (R Core Team, 2019). The function *cforest()* from *party* implements safeguards at the tree level to ensure the feature selection is not biased towards continuous predictors and/or those with many possible splits (Strobl et al., 2009; Probst et al., 2019), which is not available in *randomForest()* and *ranger()* functions. The permutation variable importance measure (Breiman, 2001; Strobl et al., 2007) has been demonstrated to reduce bias as compared with other alternatives (Strobl et al., 2007; Boulesteix et al., 2012). Moreover, since our dataset includes correlated features (Supp. Figs. 2 and 3), we evaluated the variable importance with a "conditional" permutation test to minimize the overestimation on importance scores of correlated features (Strobl et al., 2008; Probst et al., 2019).

2.2.4. Cross-validation scheme

For each prediction model, a nested cross-validation (CV) scheme was applied to avoid over-fitting during the model selection process (Zhang and Yang, 2015). This type of CV encompasses the use of an inner-loop for optimization and an outer-loop to assess the generalization performance (Krstajic et al., 2014). Acknowledging our dataset as relatively small for machine learning purposes (Zhang and Ling, 2018), we increased the *k* value (folds) with respect to the traditional 5 or 10-folds as a safeguard to reduce potential bias on the generalization error (Cawley and Talbot, 2010). Thus, an outer 20-fold scheme was used, setting aside a different 5 % of observations at a time to be used later as the *testing* data. At the inner loop, a 10-fold-CV was applied over each outer-training set, dividing 90 % for training and 10 % for validation. A grid-search was performed to optimize model hyper-parameters of interest: i) *ntree,* as the number of trees in the forest, and ii) *mtry,* as the number of random variables considered at each tree node-split across the forests. Best combinations were selected based on average performance on the inner-validation set. With the optimized hyper-parameters, performance metrics and features importance were assessed using the outer-training sets (20) to predict the observations on the outer-testing sets.

Six complementary metrics were used to evaluate models performance: i) the mean absolute error (MAE, Mg ha^{-1}) as an average magnitude of the errors; ii) the root mean square error (RMSE, Mg ha^{-1}) as an average squared errors-based statistic that penalizes large residuals more heavily than MAE; iii) the normalized or relative RMSE (RRMSE, %) as a metric of percentage deviation from the average yield (Yang et al., 2014); iv) the mean bias error (MBE, Mg ha^{-1}) as the average difference of predicted values with respect to observed, for which positive values mean a systematic over-prediction while negative mean under-prediction; v) the Nash–Sutcliffe model efficiency (ME) as a normalized analogous statistic to the coefficient of determination (Nash and Sutcliffe, 1970; Krause et al., 2005); and vi) the concordance correlation coefficient (CCC) as a normalized metric that weighs the Pearson correlation coefficient (r) by an index of accuracy (Lin, 1989). The medians (50th percentile) of each metric based on the 20-folds-CV were selected as their unbiased central-tendency statistic.

3. Results

3.1. Database description

Maize experiments were distributed across 21 US states (AL, AR, IA, IL, IN, KS, KY, MI, MN, MO, NC, ND, NE, OH, OK, PA, SD, TN, TX, VA and WI) and two Canada provinces (ON and QC) (Fig. 1A). In temporal terms, 19.7 %, 31.0 %, 31.2 %, and 18.1 % were distributed between 1999–2004, 2005–2009, 2010–2014, and 2015–2019, respectively (**Supp.** Fig. 1). A total of 831 (81 %) and 200 trials (19 %) were conducted under rainfed and irrigated conditions, respectively. Under rainfed conditions, Y_0 ranged from 0.73 to 17.7 Mg ha^{-1}, with a mean of 6.97 Mg ha^{-1} and a median of 6.41 Mg ha^{-1} (interquartile range, IQR_{25-75} = 4.21−9.49 Mg ha^{-1}). Under irrigation, Y_0 varied from 1.29 to 16.1 Mg ha^{-1}, with a mean of 9.10 Mg ha^{-1} and a median of 9.50 Mg ha^{-1} (IQR_{25-75} = 6.84-11.65 Mg ha^{-1}). Based on available observations of above-ground plant N uptake at maturity (n = 279), estimations of apparent indigenous soil N supply varied from at least 23 kg N ha^{-1} to 411 kg N ha^{-1}, representing apparent N requirements from 11. 8–22.1 kg N Mg grain yield^{-1} (Fig. 1B). Complementary, observations of grain N uptake and grain dry biomass (n = 305) were used to estimate a grain N nutrition index (NNI) following the ear-N dilution curve (%Nc = 2.22 * Grain$^{-0.26}$; Zhang et al., 2020), which was able to portray the positive effect of alfalfa as previous crop on maize N nutrition (Fig. 1C).

In terms of soil, experiments represented 11 soil textural groups (Soil Survey Staff, 2014) (Fig. 1D). Soil organic matter at topsoil (%, 0−15 cm) ranged from 0.46 % to 11.3 %, with a mean of 3.49 % and a median of 3.40 % (IQR_{25-75} = 2.12 %–4.91 %). In terms of weather, studies were exposed to a wide range of mean seasonal temperatures (Fig. 1E) that ranged from 13.5 °C to 26.6 °C, with a mean of 18.9 °C and a median of 18.7 °C (IQR_{25-75} = 17.2–20.9 °C); and seasonal precipitations -April-September- (Fig. 1F) ranged from 165 mm to 1167 mm, with a mean of 613 mm and a median of 593 mm (IQR_{25-75} = 502–703 mm). A total of 630 (61 %) and 401 trials (39 %) were reported under conventional tillage (TI) and no-tillage (NT) systems, respectively. Previous crops were alfalfa (n = 83), soybean and annual legumes (n = 497), and cereals and others (n = 451). Planting dates were reported in 643 cases (62 % of database), which in 95 % of cases ranged between March-20th to May-28th and were centered around May-5th (IQR_{25-75} = April-23rd to May-11th).

Exploratory correlation matrix was calculated (**Supp.** Fig. 2) and principal components analysis (**Supp.** Fig. 3) conducted to understand the main relationship patterns between the continuous explanatory variables. The first component, explaining 36 % of variability, can be interpreted a temperature-dimension where temperature variables (Temp, CHU, ETE) showed a high correlation to each other and were

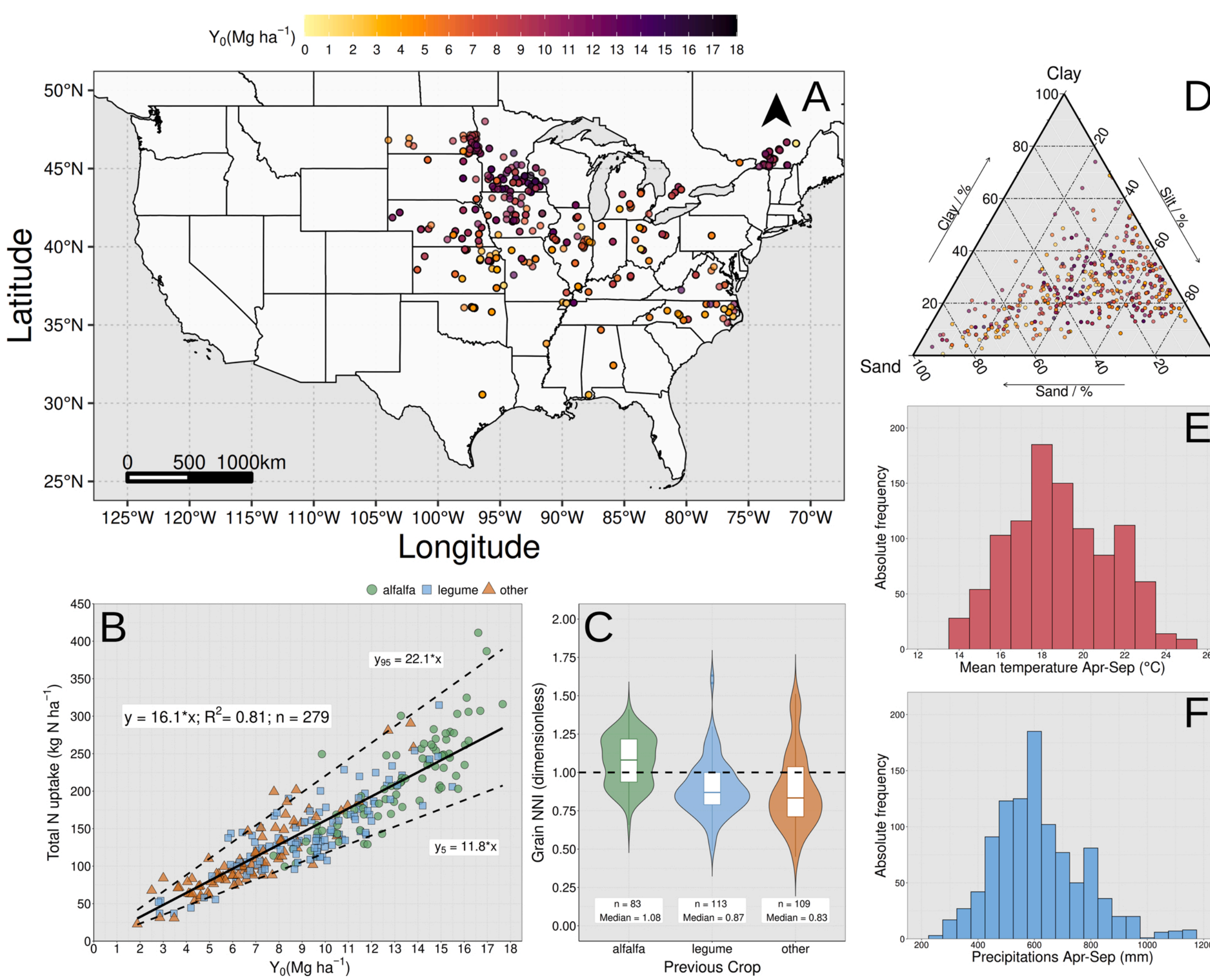

Fig. 1. A: Geographical distribution of maize nitrogen fertilization trials under study (1031 Y_0 observations from 679 site-years) performed in the USA and Canada during the period 1999-2019. B: Relationship between total above-ground N uptake at crop maturity (R6, n = 279) and yield under zero-N (Y_0). C: Estimated grain N Nutrition Index (NNI, n = 305) of zero-N maize for different previous using ear N dilution curve as reference (Zhang et al., 2020). D: variability of soil texture (0-15 cm), E: distribution of mean temperature, and F: total precipitation (mm) from April 1st to September 30th.

negatively correlated with latitude and Q index. The second component, explaining 11 % of variability, discriminates levels of precipitation and radiation. In total, five interpretable components explained about 70 % of variability in both, rainfed and irrigated conditions (**Supp.** Fig. 3).

3.2. Prediction performance

Performance metrics improved with increasing number of weather predictors accessible to the model (Full > Spring > No Weather, Fig. 2). The model "No weather" that did not include weather predictors accounted for roughly half of the variation in Y_0 (ME = 0.51), with CCC = 0.66, MAE = 1.94 Mg ha^{-1}, RMSE = 2.46 Mg ha^{-1}, RRMSE = 33.7 %, and MBE = -0.107 Mg ha^{-1}. The "spring model" improved the accuracy relative to "No weather" model by adding mean temperature and precipitations of March and April-May periods. Prediction metrics medians were ME = 0.59, CCC = 0.75, MAE = 1.72 Mg ha^{-1}, RMSE = 2.16 Mg ha^{-1}, RRMSE = 29.3 %, and MBE = -0.036 Mg ha^{-1}. The "Full weather" model accounted for 64 % the variation in Y_0 (ME = 0.64), with CCC = 0.77, MAE = 1.56 Mg ha^{-1}, RMSE = 2.01 Mg ha^{-1}, RRMSE = 27.1 %, and MBE = –0.043 Mg ha^{-1}.

3.3. Features importance

Conditional importance analysis indicated that the most important factors driving Y_0 variability were previous crop and irrigation for all models (Fig. 3). These factors were several times more relevant than the evaluated soil and weather features. Regarding the previous crop effect, Y_0 levels were the greatest with alfalfa as previous crop, followed by annual legumes and others, respectively (Fig. 4A). Irrigation positively influenced Y_0 of maize, especially with annual legumes as previous crop, increasing yields differences over other previous crops that did not reflect a positive effect of irrigation as annual legumes (Fig. 4A). Soil factors decreased in relative importance as weather features were introduced. However, SOM ranked as the most important soil variable for Y_0 regardless of the model (Fig. 3). Regarding soil texture, its relevance resulted inconsistent with no fraction resulting particularly relevant.

When the weather features were introduced to the model, they improved the prediction accuracy, reduced the relevance of soil factors, and increased the relevance of management factors (Fig. 3**B; C**). Precipitations and mean temperature during April-May ranked as the most important features for the Spring weather. Although including all weather variables still refined the prediction accuracy (Fig. 2), signs of redundant features with only a marginal effects on performance were observed. Since importance scores were estimated conditional to the presence of correlated features, general low scores and a considerable fragmentation was observed across all the weather variables. Thus, relative importance of weather in the Full-weather model did considerable not increase with respect to the Spring-weather model. Notwithstanding, it is noticeable that the Full model allowed better ensemble structures that increased the relevance of previous crop and Irrigation factors (Fig. 3C), which resulted in increased prediction accuracy (Fig. 2). Moreover, several important insights emerged from the ranking of weather predictors. The occurrence of extreme precipitation events (EPE_AM, daily PP>25 mm) during early-growth stages exhibited a negative effect on Y_0 (Fig. 4C). The amount of radiation per unit of thermal time (Q quotient) during April-May (Fig. 4D) but particularly during June-July (Fig. 4E) exhibited a positive effect on predicted Y_0 until reaching an optimum level (about 1.0 unit for Q_AM, and 0.6 units for Q_JJ). Likewise, a negative association of Y_0 with extreme temperature events (>30 °C) during August-September (ETE_AS) (Fig. 4F) as well as with the mean temperature of April-May (data not shown), as

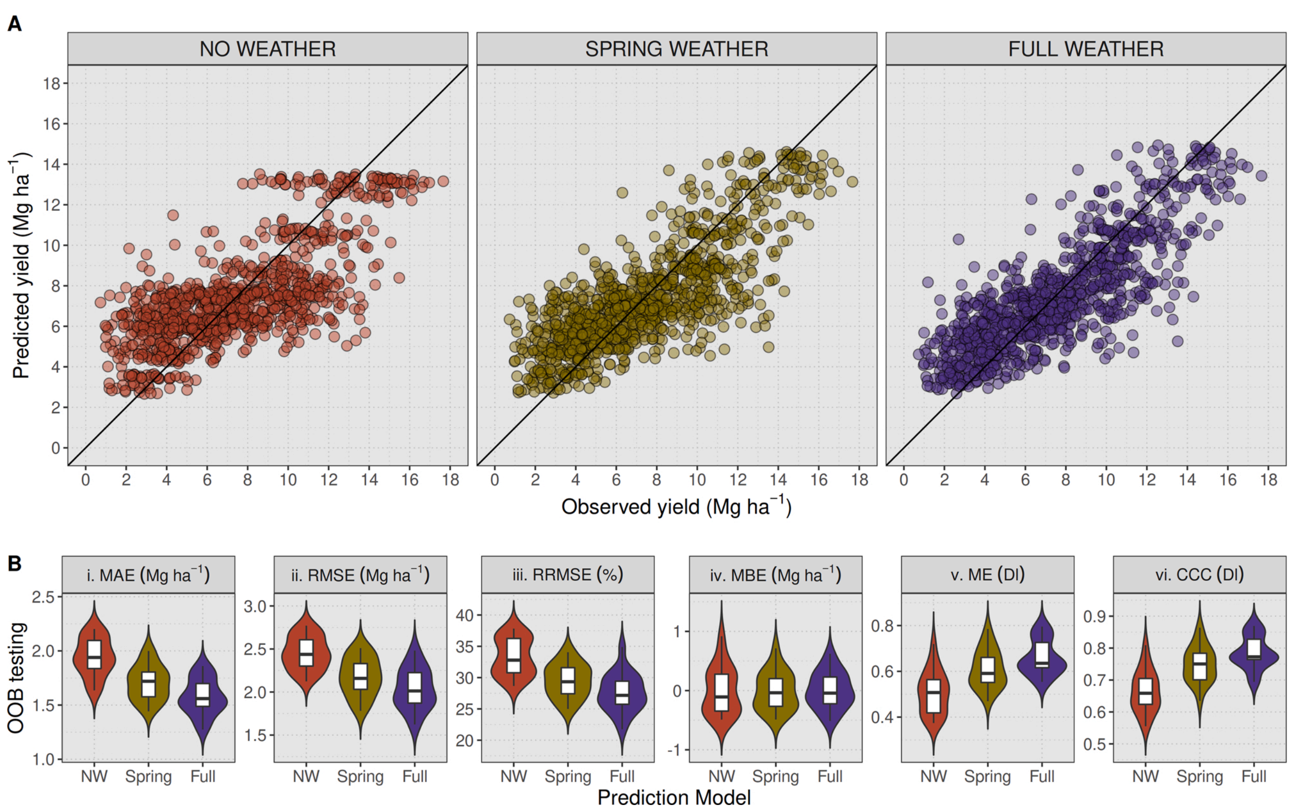

Fig. 2. Out of bag (OOB) prediction performance of conditional random forest considering three alternative models: NW – No weather, only soil and crop management features; Spring weather – including March, April and May mean temperature and precipitations; and Full weather – including all weather variables during the cropping season (April-September). Violin plots represent variability of performance metrics assessed on a 20-fold cross-validation scheme. Internal boxes represent the inter-quartile range (25th to 75th percentile) and whiskers the 5th to 95th percentiles. Model Efficiency (ME) and concordance correlation coefficient (CCC) are dimensionless (Dl) indices.

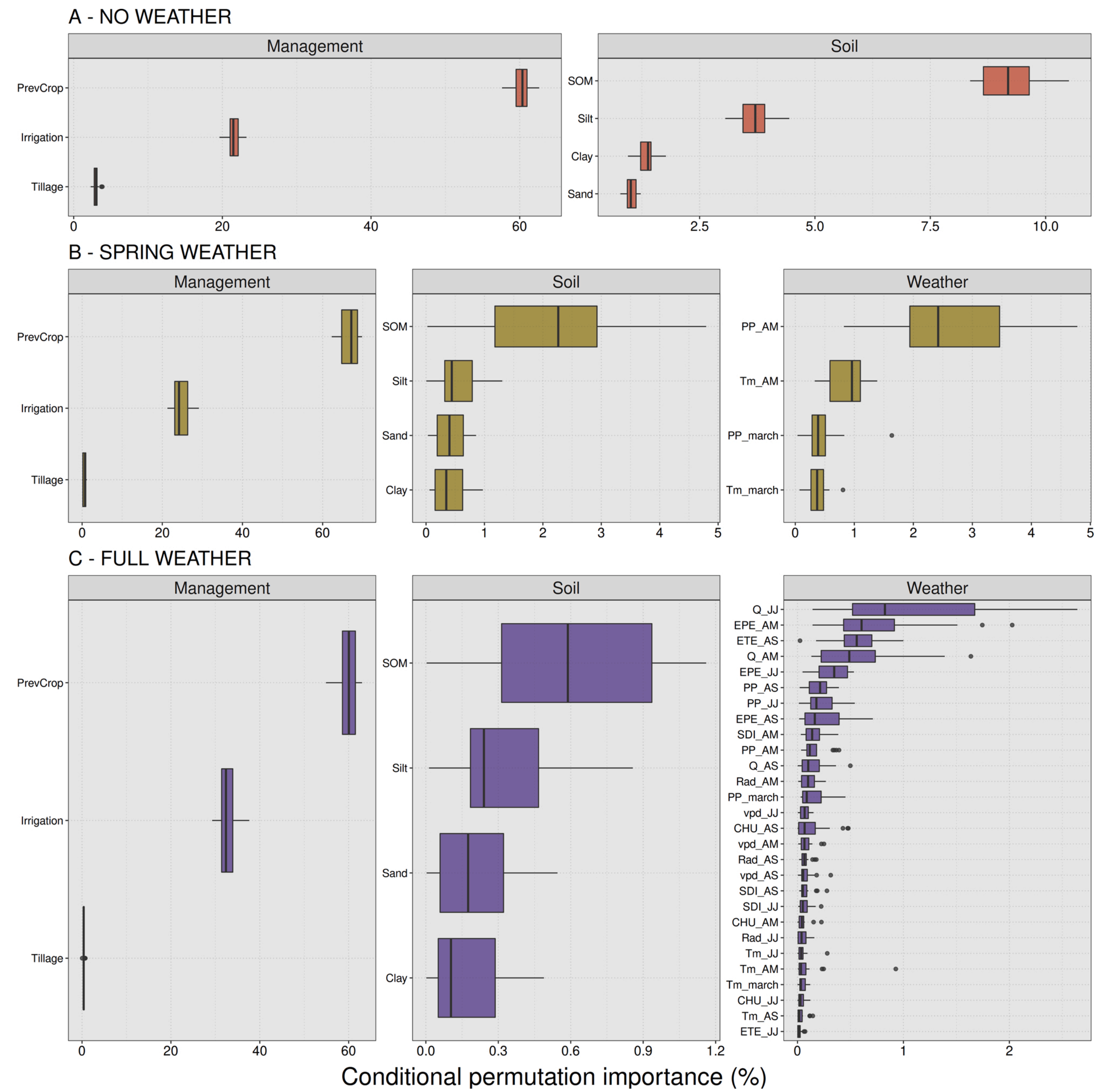

Fig. 3. Variable importance of management, soil, and weather features on the prediction of Y_0 at three alternative frameworks assessed via conditional permutations on random forest models (Strobl et al., 2008) re-scaled to percentage. Within each framework, boxes represent the inter-quartile range (25th to 75th percentile) and whiskers the 5th to 95th percentiles of conditional importance under a 20-fold cross-validation scheme. Abbreviations from Table 1.

both weather features are moderately correlated (**Supp.** Figs. 2 **and 3**). Although, only simple dependencies are shown, this did not preclude existence of significant higher-level interactions.

High-level interactions arose from this analysis. However, it is remarkable that two out of the five most important weather variables in the full model were from early stages (EPE_AM, and Q_AM), plus the high relevancy of PP_AM and Tm_AM for the Spring weather model. These results indicate that early spring weather data already provides relevant information relative to the interaction between plant N demand and soil N supply.

4. Discussion

This study combined a comprehensive collection of maize experiments and advanced analytics to: i) describe properties of Y_0 under a large variation of production conditions, and ii) to assess the importance of environmental and agronomic determinants of variation in this important descriptor of maize productivity. The outlined model could be used in combination with mechanistic models to improve prediction accuracy and decision making in N fertilization (Messina et al., 2020). This study also determined uncertainty levels for the forecast of Y_0 under alternative prediction frameworks, which defines limits of predictability. Awareness about uncertainty on Y_0 is crucial to set realistic expectations on prediction accuracy for yield response to N, EONR, and *ex-ante* N recommendations.

Further insights on the main driving factors of Y_0 have implications for its use as a proxy of indigenous soil N supply (Cassman et al., 2002) or as a metric of biological buffering capacity (Morris et al., 2018). Available data on plant N uptake at crop maturity (R6) on this database indicates that under zero-N fertilizer, a maize crop needed at least

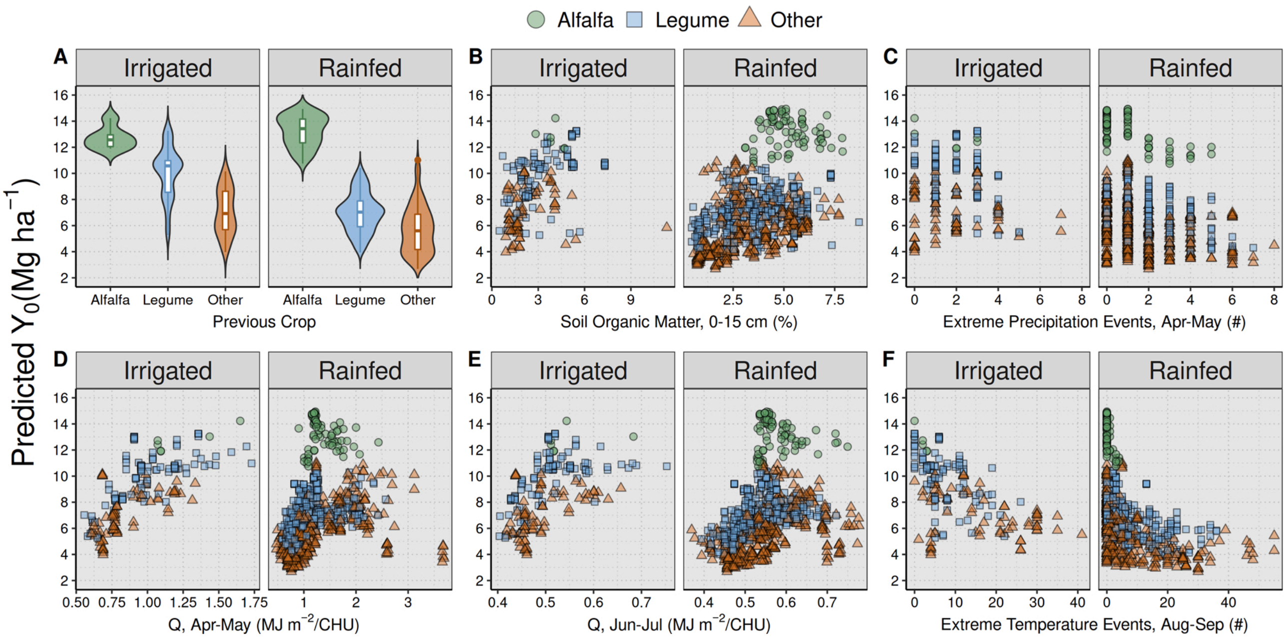

Fig. 4. Partial main dependencies of predicted maize grain yield under N omission (Y_0, Mg ha^{-1}) on the most relevant features related to management, soil, and weather (Fig. 3). In A, Boxes represent the inter-quartile range (25th to 75th percentile) and whiskers the 5th to 95th percentiles. Out-of-bag predictions from 20-fold cross-validation for the Full weather framework.

between 11.8–22.1 kg available N ha^{-1} per Mg of grain yield (Fig. 1B), acknowledging that the crop is not a merely passive sink for N (Fox and Piekielek, 1995; Vanotti and Bundy, 1994; Meisinger et al., 2008; Soufizadeh et al., 2018). Undoubtedly, addressing the soil-*N*-supply and plant-demand trade-offs (Briat et al., 2020) from complementary perspectives plays a key role for the design of N management strategies in maize crop. For a reduced portion of our dataset (<30 %), Fig. 1**C** shows that following the concept of N dilution curves (Plénet and Lemaire, 2000; Lemaire and Ciampitti, 2020), estimates of N nutrition index (NNI) could provide a mechanistic-approximation of N uptake satisfied by a given soil condition (Devienne-Baret et al., 2000). This estimation of grain NNI at harvest using ear-N dilution curve as reference (Zhang et al., 2020) was able to portray differences of zero-N maize under different previous crops. However, a major limitation at a regional scale relies on the lack available and relevant data (co-variables) such as on whole-plant biomass and plant N uptake at specific stages (e.g., flowering) in order to represent contrasting management, soil, and weather conditions.

For the above-mentioned purposes, it is noteworthy that collecting field data on Y_0 would be fairly scalable. Similarly, collecting initial soil data and obtaining precise spring weather data for building a simple but an effective prediction approach would also be fairly scalable. The reasonable performance of our data assessment framework across a wide geographic region suggests that cross-state guidelines could be pursued, a pending aspect for most of current N guidelines (Morris et al., 2018). Further efforts should recognize the value of combining collaborative research with increasing computational resources, data sources and type of models (Messina et al., 2020).

This study also offers an *ex-ante* approach using a large database of field studies to develop forecast models for Y_0. Past efforts were mostly focused on: i) describing N response curves *ex-post* (Morris et al., 2018); ii) predicting the EONR via simulation models (Melkonian, 2008; Setiyono et al., 2011; Puntel et al., 2018); or iii) predicting EONR via machine learning using datasets of limited size that constrain the generalization of outcomes (Qin et al., 2018; Ransom et al., 2019). The vast majority of models in literature use all the available data for training, but not out-of-sample data is used for testing how well they predict unseen observations. Predicting EONR faces the issue of defining a reference value, and its degree of uncertainty is generally overlooked (Hernandez and Mulla, 2008), highly depending on the best fitted model (Jaynes, 2011) and on the fertilizer to grain price ratio (Kim et al., 2013). Machine learning with small datasets (up to few hundred observations) is likely to suffer of high bias, limiting the detection of patterns and restricting the predictive ability in unexplored domains (Zhang and Ling, 2018). Still yet, limited efforts were focused specifically on the prediction of Y_0 (Puntel et al., 2019), also with constraints on data availability to explore benefits of machine learning-type models.

We acknowledge issues limiting the scope of this approach: i) achieving a balanced and more detailed dataset, ii) research plot data has limitations, and iii) the trade-off between prediction power and interpretability of machine learning. For the first point, our dataset suffered from unevenly reported metadata and a lack of relevant features such as soil N availability tests, plant biomass and N uptake, planting and maturity dates, among other data descriptors that could eventually result in improved performance. From the scalability perspective, yields in well-managed research experiments are generally greater than yield with the same practices applied by farmers in production fields (Cassman et al., 2002). Regarding the interpretability limitations, this is currently shared by most of the machine learning algorithms (Khaki and Wang, 2019). Nonetheless, as computing power and algorithms exponentially grow, we will likely overcome the "black-box" limitation in the foreseeable future with refined methods to assess features role on prediction (Springenberg et al., 2015). Meta-learning models as ensembles of learning algorithms (Makowski et al., 2015) coupled with simulations (Shahhosseini et al., 2019; Messina et al., 2020) and cross-scales models (Wu et al., 2019) may contribute to this process. Finally, Bayesian statistics are also likely to contribute to yield forecast models as they offer more inference options on dealing with yield uncertainty (Iizumi et al., 2009).

A noteworthy outcome of this study is that a large fraction of the Y_0 variability was explained just by management and soil factors (~50 %). Weather contributed to improving the overall performance (+15 %). The "Full weather" and the "Spring weather" models reduced the RRMSE by 7 % and 4 %, respectively, with respect to the "No weather" model. While the reduction in RRMSE of the "Spring weather" model is lower than the "Full weather" model, it could be utilized for prediction.

Prediction errors in the range of RMSE ~2 Mg ha^{-1} (RRMSE from 27 % to 34 %) still represent a moderate performance and significant remaining uncertainty (Liu et al., 2013). Taking into consideration the observed range of apparent N requirement to produce 1 Mg yield ha^{-1} (Fig. 1B), those values can be translated into an uncertainty in soil N supply of at least from 23 to 44 kg N ha^{-1} (considering an ideal, 100 %, N uptake efficiency). However, this also represents an opportunity for improvement. For example, a similar research approach on the prediction of rainfed maize yield using 2267 field studies across the US obtained a RRMSE up to 11 % using deep neural networks (Khaki and Wang, 2019), although encompassed more than 140,000 observations for training, as well as a much a more balanced and detailed database in terms soil, weather, in addition to the use of genetic markers data.

Across all models, the positive influence of legumes residues into crop rotations is clearly highlighted among management factors. The effect of alfalfa on the following maize N response has been well documented affecting soil N availability as well as soil physical conditions (Yost et al., 2012, 2013; 2014; Riedell, 2014). At the cropping system level, better coupling of C and N cycling processes can be achieved by relying more on organic rather than inorganic nutrient inputs (Drinkwater and Snapp, 2007). On the other hand, as one of the most limiting factors of maize yields (Mueller et al., 2012; Elliott et al., 2013; Meng et al., 2016), water supply was also a critical management factor for Y_0, particularly enhancing yields of annual legumes as previous crops more than for cereals (Fig. 4A), as the first group is comparatively less likely to suffer N-limitations. Counter-intuitively, our analysis did not show the expected influence of factors such as tillage on improving the estimation of Y_0. Nonetheless, a lack of differences in yield response was also noted from the MRTN database (Sawyer and Nafziger, 2005). At a regional scale of our analysis, marginal effects are likely distorted by higher level interactions and by systematic differences in experimental methods. At a field level, however, it is well documented that tillage can modify soil aggregation, water holding capacity, soil temperature, and consequently soil N mineralization (Bruce et al., 1990; Andraski and Bundy, 2008; Coulter and Nafziger, 2008).

Considered an essential part of the soil and farming systems (Lal, 2004), SOM played the most influential role among soil features. A recent global meta-analysis documented a positive trend of maize yields with SOM with leveling off at ~3.4 % (Oldfield et al., 2019). This study estimated that the same yield would be achievable with zero-N input in a soil with SOM of 3.4 % as with 50 kg N ha^{-1} with SOM of 0.9 %. However, N mineralization and the total organic carbon pool shows inconsistent relationship across the literature (Fox and Piekielek, 1984; Narteh and Sahrawat, 1997; Schomberg et al., 2009; Soon et al., 2007; Sainz Rozas et al., 2008), potentially related to differences in the most active of SOM fractions (Schmidt et al., 2011). In this sense, indices of soil N mineralization would theoretically improve the utilization of SOM and a simple index from soil-test biological activity appears noteworthy (Franzluebbers, 2018). Lastly, the soil texture is sometimes presented in association with soil N mineralization, but the relationship is variable across studies in the literature (Hassink, 1997; Franzluebbers et al., 1996; Yoo and Wander, 2006; Zhu et al., 2009; Dessureault-Rompré et al., 2010; Ros et al., 2011; Cai et al., 2016). For instance, a meta-analysis including 51 experiments in North-America have reported higher maize N responses under finer soil textures (Tremblay et al., 2012), while only marginal effects of spatial variability for soil texture relative to variation across years were also reported in other studies (van Es et al., 2005v; Tremblay and Bélec, 2006; Kyveryga et al., 2009).

Weather factors are determinants of both N supply and demand (Soufizadeh et al., 2018). In this study, the excess of rainfall early in the season enter in prediction models consistent with the negative impact of high precipitation on drainage, water-logging and increased N losses (Cameron et al., 2013; Wang et al., 2014). Spring precipitations have been reported to account for 74 % of inter-annual variation in mean soil residual N at pre-sidedress (Balkcom et al., 2003). Similarly, every 10 mm of April precipitation above historical average delayed planting date for 1 day in the main 12 central US states (Kucharik, 2008). Although it is unlikely that yields under N limitations were limited by solar radiation (DeBruin et al., 2013; Soufizadeh et al., 2018), radiation per unit of thermal time (Q quotient) during June-July (JJ) and early in the season (AM) positively affected yields (Andrade et al., 2000; Carter et al., 2018; Soufizadeh et al., 2018) until variable optimum levels, exhibiting the trade-off with the temperature effects on radiation use efficiency (Andrade et al., 1993) and biomass partitioning to the ear (Wilson et al., 1995). Regarding temperature, the occurrence of extreme temperatures during the reproductive period (ETE_AS) resulted in one of the most relevant features suggesting that the positive effect of temperature on soil N mineralization (Dalias et al., 2002; Wu et al., 2008; Fernández et al., 2017) could be offset by a negative impact of supra-optimal temperatures on plant growth (*e.g.*, shortening the grain filling duration) and plant N demand (Muchow et al., 1990; Soufizadeh et al., 2018). Overall, the high relevance of weather features at early stages (spring) appraises to invest more resources in the aggregation and analysis of massive databases that allow to further explore the development of prediction frameworks for Y_0 that can be applied in practice.

5. Conclusions

Management factors such as previous crop and irrigation in combination with top-soil SOM accounted for the largest portion of variation in Y_0, while the inclusion of weather features refined the prediction accuracy. In a practical sense, a simple framework including weather variables of spring (March-May) might result comparable in performance to a framework including all-season weather. Future attempts should assess alternative statistical and machine learning approaches offering performance and interpretability improvements. Refined prediction frameworks for Y_0 could provide new insights on N responsiveness and represent a step-forward towards more collaborative and regional-scale N recommendation guidelines.

CRediT authorship contribution statement

Adrian A. Correndo: Conceptualization, Methodology, Data curation, Formal analysis, Writing - original draft, Writing - review & editing, Visualization. **Jose L. Rotundo:** Conceptualization, Methodology, Data curation, Writing - review & editing. **Nicolas Tremblay:** Investigation, Data curation, Writing - review & editing. **Sotirios Archontoulis:** Investigation, Data curation, Writing - review & editing. **Jeffrey A. Coulter:** Investigation, Data curation, Writing - review & editing. **Dorivar Ruiz-Diaz:** Investigation, Data curation, Writing - review & editing. **Dave Franzen:** Investigation, Data curation, Writing - review & editing. **Alan J. Franzluebbers:** Investigation, Data curation, Writing - review & editing. **Emerson Nafziger:** Investigation, Data curation, Writing - review & editing. **Rai Schwalbert:** Methodology, Formal analysis, Writing - review & editing. **Kurt Steinke:** . **Jared Williams:** Investigation, Data curation, Writing - review & editing. **Charlie D. Messina:** Conceptualization, Methodology, Writing - review & editing. **Ignacio A. Ciampitti:** Conceptualization, Methodology, Data curation, Writing - review & editing, Project administration.

Declaration of Competing Interest

The authors reported no declarations of interest.

Acknowledgments

The authors gratefully acknowledge the financial support provided by Fulbright Program, Kansas Corn Commission, Corteva Agriscience, and Kansas State University for sponsoring A. Correndo's doctoral studies and Dr. I.A. Ciampitti's research program. This is contribution no. 21-092-J from the Kansas Agricultural Experiment Station.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fcr.2020.107985.

References

Al-Kaisi, M.M., Yin, X., Licht, M.A., 2005. Soil carbon and nitrogen changes as influenced by tillage and cropping systems in some Iowa soils. Agric. Ecosyst. Environ. 105, 635–647. https://doi.org/10.1016/j.agee.2004.08.002.

Andrade, F.H., Uhart, S.A., Cirilo, A., 1993. Temperature affects radiation use efficiency in maize. Field Crops Res. 32, 17–25. https://doi.org/10.1016/0378-4290(93)90018-I.

Andrade, F.H., Otegui, M.E., Vega, C., 2000. Intercepted radiation at flowering and kernel number in maize. Agron. J. 92, 92–97. https://doi.org/10.2134/agronj2000.92192x.

Andraski, T.W., Bundy, L.G., 2008. Corn residue and nitrogen source effects on nitrogen availability in no-till corn. Agron. J. 100, 1274–1279. https://doi.org/10.2134/agronj2008.0039.

Archontoulis, S.V., Castellano, M.J., Licht, M.A., et al., 2020. Predicting crop yields and soil-plant nitrogen dynamics in the US Corn Belt. Crop Sci. 60, 721–738. https://doi.org/10.1002/csc2.20039.

Baker, J.L., Johnson, H.P., 1981. Nitrate-nitrogen in tile drainage as affected by fertilization. J. Environ. Qual. 10, 519–522. https://doi.org/10.2134/jeq1981.00472425001000040020x.

Balkcom, K.S., Blackmer, A., Hansen, D., Morris, T., Mallarino, A., 2003. Testing soils and cornstalks to evaluate nitrogen management on the watershed scale. J. Environ. Qual. 32, 1015–1024. https://doi.org/10.2134/jeq2003.1015.

Bannayan, M., Hoogenboom, G., Crout, N.M.J., 2018. Photothermal impact on maize performance: a simulation approach. Ecol. Modell. 180 (2–3), 277–290. https://doi.org/10.1016/j.ecolmodel.2004.04.022.

Belgiu, M., Drăgut, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Boulesteix, A.-L., Janitza, S., Kruppa, J., König, I.R., 2012. Overview of Random Forest Methodology and Practical Guidance With Emphasis on Computational Biology and Bioinformatics. WIREs Data Mining Knowl Discov, 2N, pp. 493–507. https://doi.org/10.1002/widm.1072.

Bowles, T.M., Atallah, S.S., Campbell, E.E., Gaudin, A.C.M., Wiederm, W.R., 2018. Addressing agricultural nitrogen losses in a changing climate. Nat. Sustain. 1, 399–408. https://doi.org/10.1038/s41893-018-0106-0.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://link.springer.com/content/pdf/10.1023%2FA%3A1010933404324.pdf.

Briat, J.-F., Gojon, A., Plassard, C., Rouached, H., Lemaire, G., 2020. Reappraisal of the central role of soil nutrient availability in nutrient management in light of recent advances in plant nutrition at crop and molecular levels. Eur. J. Agron. 116, 126069. https://doi.org/10.1016/j.eja.2020.126069.

Bruce, R.R., Langdale, G.W., Dillard, A.L., 1990. Tillage and crop rotation effect on characteristics of a sandy surface soil. Soil Sci. Soc. Am. J. 54, 1744–1747. https://doi.org/10.2136/sssaj1990.03615995005400060039x.

Bundy, L.G., Andraski, T.W., 1995. Soil yield potential effects on performance of soil nitrate tests. J. Prod. Agric. https://doi.org/10.2134/jpa1995.0561, 8N561–568.

Butler, E.E., Huybers, P., 2013. Adaptation of US maize to temperature variations. Nat. Clim. Chang. 3N, 68–72. https://doi.org/10.1038/nclimate1585.

Cai, A., Xu, H., Shao, X., Zhu, P., Zhang, W., Xu, M., Murphy, D.V., 2016. Carbon and nitrogen mineralization in relation to soil particle-size fractions after 32 years of chemical and manure application in a continuous maize cropping system. PLoS One 11 (3), e0152521. https://doi.org/10.1371/journal.pone.0152521.

Cameron, K.C., Di, H.J., Moir, J.L., 2013. Nitrogen losses from the soil/plant system: a review. Ann. Appl. Biol. https://doi.org/10.1111/aab.12014, 162N145–173.

Carter, E., Melkonian, J., Steinschneider, S., Riha, S., 2018. Rainfed maize yield response to management and climate covariability at large spatial scales. Agric. Forest Meteorol. 256, 242–252. https://doi.org/10.1016/j.agrformet.2018.02.029.

Cassman, K.G., Dobermann, A., Walters, D.T., 2002. Agroecosystems, nitrogen-use efficiency, and nitrogen management. AMBIO A J. Hum. Environ. 31 (2), 132–140. https://doi.org/10.1579/0044-7447-31.2.132 (1 March 2002).

Cassman, K.G., Gines, G.C., Dizon, M.A., Samson, M.I., Alcantara, J.M., 1996. Nitrogen use efficiency in tropical lowland rice systems: contribution from indigenous and applied nitrogen. Field Crops Res. 47 (1), 1–12. https://doi.org/10.1016/0378-4290(95)00101-8.

Cawley, G.C., Talbot, N.L.C., 2010. On overofitting in model selection and subsequent selection bias in performance evaluation. J. Mach. Learn. Res. 11 (2079–2107), 2010. http://jmlr.org/papers/v11/cawley10a.

Chaney, N., Minasny, B., Herman, J., Nauman, T., Brungard, C., Morgan, C., Mcbratney, A., Wood, E., Yimam, Y., 2016. POLARIS soil properties: 30-meter probabilistic maps of soil properties over the contiguous United States. Water Resour. Res. https://doi.org/10.1029/2018WR022797.

Coulter, J.A., Nafziger, E.D., 2008. Continuous corn response to residue management and nitrogen fertilization. Agron. J. 100, 1774–1780. https://doi.org/10.2134/agronj2008.0170.

Dalias, P., Anderson, J.M., Bottner, P., Coûteux, M.M., 2002. Temperature responses on net nitrogen mineralization and nitrification in conifer forest soils incubated under standard laboratory conditions. Soil Biol. Biochem. 34, 691–701. https://doi.org/10.1016/S0038-0717(01)00234-6.

Daly, C., Smith, J.I., Olson, K.V., 2015. Mapping atmospheric moisture climatologies across the conterminous United States. PLoS One 10 (10), e0141140. https://doi.org/10.1371/journal.pone.0141140.

Debruin, J., Messina, C.D., Munaro, E., Thompson, K., Conlon-Beckner, C., Fallis, L., Sevenich, D.M., Gupta, R., Dhugga, K.S., 2013. N distribution in maize plant as a marker for grain yield and limits on its remobilization after flowering. Plant Breed. 132, 500–505. https://doi.org/10.1111/pbr.12051.

Dessureault-Rompré, J., Burton, D.L., Sharifi, M., Cooper, J., Grant, C.A., Drury, C.F., 2010. Relationships among mineralizable soil nitrogen, soil properties, and climatic indices. Soil Sci. Soc. Am. J. 74, 1218–1227. https://doi.org/10.2136/sssaj2009.0213.

Devienne-Baret, F., Justes, E., Machet, J.M., Mary, B., 2000. Integrated control of nitrate uptake by crop growth rate and soil nitrate availability under field conditions. Ann. Bot. 86, 995–1005. https://doi.org/10.1006/anbo.2000.1264.

Dickersin, K., Min, Y., 1993. Publication bias: the problem that won't go away. Ann. N. Y. Acad. Sci. 1993 (December 703), 135–146. https://doi.org/10.1111/j.1749-6632.1993.tb26343.x discussion 146-148.

Drinkwater, L.E., Snapp, S.S., 2007. Nutrients in agroecosystems: rethinking the management paradigm. Adv. Agron. 92, 163–186. https://doi.org/10.1016/S0065-2113(04)92003-2.

Elliott, J., et al., 2013. Constraints and potentials of future irrigation water availability on agricultural production under climate change. Proc. Natl. Acad. Sci. U.S.A. 111, 3239–3244. https://doi.org/10.1073/pnas.1222474110.

Fernández, F., Fabrizzi, K., Naeve, S., 2017. Corn and soybean's season-long in-situ nitrogen mineralization in drained and undrained soils. Nutr. Cycl. Agroecosyst. 107, 33–47. https://doi.org/10.1007/s10705-016-9810-1.

Fox, R.H., Piekielek, W.P., 1984. Relationships among anaerobically mineralized nitrogen, chemical indexes, and nitrogen availability to corn. Soil Sci. Soc. Am. J. 48, 1087–1090. https://doi.org/10.2136/sssaj1984.03615995004800050027x.

Fox, R.H., Piekielek, W.P., 1995. The relationship between corn grain yield goals and economic optimum nitrogen fertilizer rates. Agron. Ser. 136. Pennsylvania State Univ., University Park.

Francis, D.D., Schepers, J.S., Vigil, M.F., 1993. Post-anthesis nitrogen loss from corn. Agron. J. 85, 659–663. https://doi.org/10.2134/agronj1993.00021962008500030026x.

Franzluebbers, A., 2010. Depth distribution of soil organic carbon as a signature of soil quality. In: 19th World Congress of Soil Science, Soil Solutions for a Changing World. Brisbane, Australia, 1–6 August 2010. https://www.iuss.org/19th%20WCSS/Symposium/pdf/2164.pdf.

Franzluebbers, A.J., 2018. Soil-test biological activity with the flush of CO2: III. Corn yield responses to applied nitrogen. Soil Sci. Soc. Am. J. 82, 708–721. https://doi.org/10.2136/sssaj2018.01.0029.

Franzluebbers, A.J., Haney, R.L., Hons, F.M., Zuberer, D.A., 1996. Active fractions of organic matter in soils with different texture. Soil Biol. Biochem. 28, 1367–1372. https://doi.org/10.1016/S0038-0717(96)00143-5.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2017.06.031.

Hassink, J., 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. Plant Soil 191, 77–87. https://doi.org/10.1023/A:1004213929699.

Heady, E.O., Pesek, J., 1954. A fertilizer production function surface with specification of economic optima for corn grown on calcareous Ida silt loam. J. Farm Econ. 36 (3), 466–482. https://doi.org/10.2307/1233014.

Hernandez, J.A., Mulla, D.J., 2008. Estimating uncertainty of economically optimum fertilizer rates. Agron. J. 100, 1221–1229. https://doi.org/10.2134/agronj2007.0273.

Hothorn, T., Hornik, K., Zeileis, A., 2006. Unbiased recursive partitioning: a conditional inference framework. J. Comput. Graph. Stat. 15 (3), 651–674. https://doi.org/10.1198/106186006X133933. Illustrations, sources and a solution. BMC Bioinformatics 8:25.

Iizumi, T., Yokozawa, M., Nishimori, M., 2009. Parameter estimation and uncertainty analysis of a large-scale crop model for paddy rice: application of a Bayesian approach. Agric. For. Meteorol. 149 (2), 333–348. https://doi.org/10.1016/j.agrformet.2008.08.015.

Jaynes, D.B., 2011. Confidence bands for measured economically optimal nitrogen rates. Precis. Agric. 12 (2), 196–213. https://doi.org/10.1007/s11119-010-9168-3.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: trends, perspectives, and prospects. Science 349, 255–260. https://doi.org/10.1126/science.aaa8415.

Khaki, S., Wang, L., 2019. Crop yield prediction using deep neural networks. Front. Plant Sci. 10, 621. https://doi.org/10.3389/fpls.2019.00621.

Kim, K.I., Clay, D., Clay, S., Carlson, G., Trooien, T., 2013. Testing corn (Zea mays L.) preseason regional nitrogen recommendation models in South Dakota. Agron. J. 105, 1619–1625. https://doi.org/10.2134/agronj2013.0166.

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.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97. https://doi.org/10.5194/adgeo-5-89-2005.

Krstajic, D., Buturovic, L.J., Leahy, D.E., Thomas, S., 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. J. Cheminform. 2014 (6), 10. http://www.jcheminf.com/content/6/1/10.

Kucharik, C.J., 2008. Contribution of planting date trends to increased maize yields in the central United States. Agron. J. 100, 328–336. https://doi.org/10.2134/agrojnl2007.0145.

Kyveryga, P., Blackmer, A.M., Morris, T.F., 2007. Disaggregating model bias and variability when calculating economic optimum rates of nitrogen fertilization for corn. Agron. J. 99, 1048–1056. https://doi.org/10.2134/agronj2006.0339.

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/agronj2008.0168.

Lal, R., 2004. Soil carbon sequestration impacts on climate change and food security. Science 304, 1623–1627. https://doi.org/10.1126/science.1097396.

Lemaire, G., Ciampitti, I.A., 2020. Crop mass and N status as prerequisite covariables for unraveling nitrogen use efficiency across genotype-by-environment-by-management scenarios: a Review. Plants 9 (10), 1309. https://doi.org/10.3390/plants9101309, 2020.

Lin, L.I.-K., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45 (1), 255–268. https://doi.org/10.2307/2532051.

Liu, S., Yang Zhang, J.Y., Drury, X.Y., Reynolds, C.F., Hoogenboom, G., 2013. Modelling crop yield, soil water content and soil temperature for a soybean–maize rotation under conventional and conservation tillage systems in Northeast China. Agric. Water Manage. 123, 32–44. https://doi.org/10.1016/j.agwat.2013.03.001.

Lory, J.A., Scharf, P.C., 2003. Yield goal versus delta yield for predicting fertilizer nitrogen need in corn. Agron. J. 95, 994–999. https://doi.org/10.2134/agronj2003.0994.

Makowski, D., et al., 2015. A statistical analysis of ensembles of crop model responses to climate change factors. Agric. For. Meteorol. 214–215, 483–493. https://doi.org/10.1016/j.agrformet.2015.09.013.

Meisinger, J.J., Schepers, J.S., Raun, W.R., 2008. Crop nitrogen requirement and fertilization. In: Schepers, J.S., Raun, W.R. (Eds.), Nitrogen in Agricultural Systems. Agron. Monogr. 49. ASA, CSSA, and SSSA, Madison, WI, pp. 563–612. http://www.chesapeake.org/stac/presentations/63_Meisinger%20J%20et%20al_2008b_Crop%20nitrogen%20requirement%20and%20fertilization.pdf.

Meng, Q., Chen, X., Lobell, D., Cui, Z., Zhang, Y., Yang, H., Zhang, F., 2016. Growing sensitivity of maize to water scarcity under climate change. Sci. Rep. 6, 19605. https://doi.org/10.1038/srep19605.

Messina, C.D., Hammer, G.L., Dong, Z., Podlich, D., Cooper, M., 2009. Modelling crop improvement in a G*E*M framework via gene-trait-phenotype relationships. In: Sadras, V., Calderini, D. (Eds.), Crop Physiology: Interfacing With Genetic Improvement and Agronomy. Elsevier, p. 2009. https://doi.org/10.1016/B978-0-12-374431-9.00010-4.

Messina, C.D., Cooper, M., Reynolds, M., Hammer, G., 2020. Crop science: a foundation for advancing predictive agriculture. Crop Sci. 60, 544–546. https://doi.org/10.1002/csc2.20116.

Monteith, J.L., 1972. Solar radiation and productivity in tropical ecosystems. J. Appl. Ecol. 9 (3), 747–766. https://doi.org/10.2307/2401901.

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., 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.

Muchow, R.C., Sinclair, T.R., Bennett, J.M., 1990. Temperature and solar radiation effects on potential maize yield across locations. Agron. J. 82, 338–343. https://doi.org/10.2134/agronj1990.00021962008200020033x.

Mueller, N., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., Foley, J.A., 2012. Closing yield gaps through nutrient and water management. Nature 490, 254–257. https://doi.org/10.1038/nature11420.

Mueller, S.M., Messina, C.D., Vyn, T.J., 2019. Simultaneous gains in grain yield and nitrogen efficiency over 70 years of maize genetic improvement. Sci. Rep. 9, 9095. https://doi.org/10.1038/s41598-019-45485-5.

Narteh, L.T., Sahrawat, K.L., 1997. Potentially mineralizable nitrogen in West African lowland rice soils. Geoderma 76 (1–2), 145–154. https://doi.org/10.1016/S0016-7061(96)00097-3.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I - a discussion of principles. J. Hydrol. 10 (3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Oldfield, E.E., Bradford, M.A., Wood, S.A., 2019. Global meta-analysis of the relationship between soil organic matter and crop yields. Soil 5, 15–32. https://doi.org/10.5194/soil-5-15-2019.

Plénet, D., Lemaire, G., 2000. Relationships between dynamics of nitrogen uptake and dry matter accumulation in maize crops. Plant Soil 216, 65–82. https://doi.org/10.1023/A:1004783431055.

Probst, P., Wright, M.N., Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. WIREs Data Mining Knowl. Discov. 2019 (9), e1301. https://doi.org/10.1002/widm.1301.

Puntel, L.A., Sawyer, J.E., Barker, D.W., Dietzel, R., Poffenbarger, H., Castellano, M.J., 2016. Modeling long-term corn yield response to nitrogen rate and crop rotation. Front. Plant Sci. 7, 630. https://doi.org/10.3389/fpls.2016.01630.

Puntel, L.A., Sawyer, J.E., Barker, D.W., Thorburn, P.J., Castellano, M.J., Moore, K.J., VanLoocke, A., Heaton, E.A., Archontoulis, S.V., 2018. A systems modeling approach to forecast corn economic optimum nitrogen rate. Front. Plant Sci. 9, 436. https://doi.org/10.3389/fpls.2018.00436.

Puntel, L.A., Pagani, A., Archontoulis, S.V., 2019. Development of a nitrogen recommendation tool for corn considering static and dynamic variables. Eur. J. Agron. 105, 189–199. https://doi.org/10.1016/j.eja.2019.01.003.

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., 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.

R Core Team, 2019. R: a Language and Environment for Statistical Computing. URL. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Ramanantenasoa, M.M.J., Génermont, S., Gilliot, J.M., Bedos, D., Makowski, D., 2019. Meta-modeling methods for estimating ammonia volatilization from nitrogen fertilizer and manure applications. J. Environ. Manage. 236, 195–205. https://doi.org/10.1016/j.jenvman.2019.01.066.

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

Riedell, W.E., 2014. Nitrogen fertilizer applications to maize after alfalfa: grain yield, kernel composition, and plant mineral nutrients. J. Plant Nutr. 37 (12), 2026–2035. https://doi.org/10.1080/01904167.2014.911892.

Ros, G.H., Hanegraaf, M.C., Hoffland, E., van Riemsdijk, W.H., 2011. Predicting soil N mineralization: relevance of organic matter fractions and soil properties. Soil Biol. Biochem. 43 (8), 1714–1722. https://doi.org/10.1016/j.soilbio.2011.04.017.

Sainz Rozas, H., Calvino, P.A., Echeverría, H.E., Barbieri, P.A., Redolatti, M., 2008. Contribution of anaerobically mineralized nitrogen to the reliability of planting or presidedress soil nitrogen test in maize. Agron. J. 100, 1020–1025. https://doi.org/10.2134/agronj2007.0077.

Sawyer, J.E., Nafziger, E.D., 2005. Regional approach to making nitrogen fertilizer rate decisions for corn. In: Proceedings North Central Extension-Ind. Soil Fertility ConFerence, Des Moines, IA. 16–17 Nov. Potash and Phosphate Inst., Brookings, SD. https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1039&context=agron_conf.

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., Sripada, R.P., Beegle, D.B., Rotz, C.A., Hong, N., 2011. Within-field variability in potimum nitrogen rate for corn linked to soil moisture availability. Soil Sci. Soc. Am. J. 75, 306–316. https://doi.org/10.2136/sssaj2010.0184.

Schomberg, H.H., Wietholter, S., Griffin, T.S., Reeves, D.W., Cabrera, M.L., Fisher, D.S., Endale, D.M., Novak, J.M., Balkcom, K.S., Raper, R.L., Kitchen, N.R., Locke, M.A., Potter, K.N., Schwartz, R.C., Truman, C.C., Tyler, D.D., 2009. Assessing indices for predicting potential nitrogen mineralization in soils under different management systems. Soil Sci. Soc. Am. J. 73, 1575–1586. https://doi.org/10.2136/sssaj2008.0303.

Schwalbert, R., Amado, T., Nieto, L., Varela, S., Corassa, G., Horbe, T., Rice, C., Peralta, N., Ciampitti, I.A., 2018. Forecasting maize yield at field scale based on high-resolution satellite imagery. Biosyst. Eng. 171, 179–192. https://doi.org/10.1016/j.biosystemseng.2018.04.020.

Sela, S., van Es, H.M., Moebius-Clune, B.N., Marjerison, R., Kneubuhler, G., 2018a. Dynamic model-based recommendations increase the precision and sustainability of N fertilization in midwestern US maize production. Comput. Electron. Agric. 153, 256–265. https://doi.org/10.1016/j.compag.2018.08.010.

Sela, S., Woodbury, P.B., van Es, H.M., 2018b. Dynamic model-based N management reduces surplus nitrogen and improves the environmental performance of corn production. Environ. Res. Lett. 13 (2018), 054010. https://doi.org/10.1088/1748-9326/aab908.

Setiyono, T.D., Yang, H., Walters, D., Dobermann, A., Ferguson, D., Roberts, L., 2011. Maize-N: a decision tool for nitrogen management in maize. Agron. J. 103, 1276–1283. https://doi.org/10.2134/agronj2011.0053.

Shahhosseini, M., Martinez-Feria, R.A., Hu, G., Archontoulis, S.V., 2019. Maize yield and nitrate loss prediction with machine learning algorithms. Environ. Res. Lett. 14 (2019), 124026. https://doi.org/10.1088/1748-9326/ab5268.

Sindelair, A.J., Coulter, J.A., Lamb, J.A., Vetsch, J.A., 2015. Nitrogen, stover, and tillage management affect nitrogen use efficiency in continuous corn. Agron. J. 107 (843–850), 2015. https://doi.org/10.2134/agronj14.0535.

Soil Survey Staff, 2014. Keys to Soil Taxonomy. 12th Edition. USDA-Natural Resources Conservation Service, Washington DC, p. 360p. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/class/taxonomy/?cid=nrcs142p2_053580.

Soon, Y.K., Haq, A., Arshad, M.A., 2007. Sensitivity of nitrogen mineralization indicators to crop and soil management. Commun. Soil Sci. Plant Anal. 38, 2029–2043. https://doi.org/10.1080/00103620701548688.

Soufizadeh, S., Munaro, E., McLean, G., Massignam, A., van Oosterom, E.J., Chapman, S. C., Messina, C., Cooper, M., Hammer, G., 2018. Modelling the nitrogen dynamics of maize crops – enhancing the APSIM maize model. Eur. J. Agron. 100, 118–131. https://doi.org/10.1016/j.eja.2017.12.007.

Springenberg, J.T., Dosovitskiy, A., Brox, T., Riedmiller, M.A., 2015. Striving for Simplicity: The All Convolutional Net. CoRR. https://arxiv.org/abs/1412.6806.

Stanford, G., 1973. Rationale for optimum nitrogen fertilization in corn production. J. Environ. Qual. 2, 159–166. https://doi.org/10.2134/jeq1973.00472425000200020001x.

Stanford, G., et al., 1966. Nitrogen requirements of crops for maximum yield. In: McVickar, M.H. (Ed.), Agricultural Anhydrous Ammonia Technology and Use. SSSA, Madison, WI, pp. 237–257. https://doi.org/10.2134/1966.nh3agricultural.c14.

Stern, H., Davidson, N.E., 2015. Trends in the skill of weather prediction at lead times of 1–14 days. Q.J.R. Meteorol. Soc. 141, 2726–2736. https://doi.org/10.1002/qj.2559.

Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 8, 25. https://doi.org/10.1186/1471-2105-8-25.

Strobl, C.A., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9, 307. https://doi.org/10.1186/1471-2105-9-307.

Strobl, C.A., Hothron, T., Zeileis, A., 2009. Party on! A new, conditional variable-importance measure for random forests available in the party package. R J. 1/2 (14–17). December 2009. https://www.bibsonomy.org/bibtex/2f1f8129a72657282688b058cf7f67099/pillo.

Sawyer, J., Nafziger, E., Randall, G., Bundy, L., Rehm, G., Joern, B., 2006. Concepts and rationale for regional nitrogen rate guidelines for corn. PM 2015. Iowa State Univ. Ext., Ames, IA. https://store.extension.iastate.edu/product/Concepts-and-Rationale-for-Regional-Nitrogen-Rate-Guidelines-for-Corn.

Tao, H., Morris, T.F., Kyveryga, P., McGuire, J., 2018. Factors affecting nitrogen availability and variability in Cornfields. Agron. J. 110, 1974–1986. https://doi.org/10.2134/agronj2017.11.0631.

Thornton, P.E., Thornton, M.M., Mayer, B.W., Wei, Y., Devarakonda, R., Vose, R.S., Cook, R.B., 2018. Daymet: Daily Surface Weather Data on a 1-km Grid for North America. ORNL-DAAC, Oak Ridge, Tennessee, USA.

Tolle, K.M., Tansley, S., Hey, T., 2011. The fourth paradigm: data-intensive scientific discovery. Proc. IEEE 99, 1334–1337. https://doi.org/10.1109/JPROC.2011.2155130.

Tremblay, N., Bélec, C., 2006. Adapting nitrogen fertilization to unpredictable seasonal conditions with the least impact on the environment. HortTechnology 16, 408–412. https://doi.org/10.21273/HORTTECH.16.3.0408.

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.

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

van Es, H.M., Yang, C.L., Geohring, L.D., 2005v. Maize nitrogen response as affected by soil type and drainage variability. Precis. Agric. 6, 281–295. https://doi.org/10.1007/s11119-005-1387-7.

Vanotti, M.B., Bundy, L.G., 1994. An alternative rationale for corn nitrogen-fertilizer recommendations. J. Prod. Agric. 7, 243–249. https://doi.org/10.2134/jpa1994.0243.

Varvel, G.E., Wilhelm, W.W., 2011. Soil carbon levels in irrigated Western Corn Belt rotations. Agron. J. 100 (4), 1180–1184. https://doi.org/10.2134/agronj2007.0383.

Villamil, M.B., Little, J., Nafziger, E.D., 2015. Corn residue, tillage, and nitrogen rate effects on soil properties. Soil Tillage Res. 151, 61–66. https://doi.org/10.1016/j.still.2015.03.005.

Wang, G., Wu, B., Zhang, L., Jiang, H., Xu, Z., 2014. Role of soil erodibility in affecting available nitrogen and phosphorus losses under simulated rainfall. J. Hydrol. 514, 180–191. https://doi.org/10.1016/j.jhydrol.2014.04.028.

Wilson, D.R., Muchow, R.C., Murgatroyd, C.J., 1995. Model analysis of temperature and solar radiation limitations to maize potential productivity in a cool climate. Field Crops Res. 43, 1–18. https://doi.org/10.1016/0378-4290(95)00037-Q.

Woli, K.P., Boyer, M.J., Elmore, R.W., Sawyer, J.E., Abendroth, L.J., Barker, D.W., 2016. Corn era hybrid response to nitrogen fertilization. Agron. J. 108, 473–486. https://doi.org/10.2134/agronj2015.0314.

Wortmann, C.S., Tarkalson, D., Shapiro, C., Dobermann, A., Ferguson, R., Hergert, G., Walters, D., 2011. Nitrogen use efficiency of irrigated corn for three cropping systems in Nebraska. Agron. J. 103, 76–84. https://doi.org/10.2134/agronj2010.0189.

Wu, T.Y., Ma, B.L., Liang, B.C., 2008. Quantification of seasonal soil nitrogen mineralization for corn production in eastern Canada. Nutr. Cycl. Agroecosyst. 81, 279–290. https://doi.org/10.1007/s10705-007-9163-x.

Wu, A., Hammer, G., Doherty, A.I., Caemmerer, S., Farquhar, G., 2019. Quantifying impacts of enhancing photosynthesis on crop yield. Nat. Plants 5. https://doi.org/10.1038/s41477-019-0398-8.

Yang, J.M., Yang, J.Y., Liu, S., Hoogenboom, G., 2014. An evaluation of the statistical methods for testing the performance of crop models with observed data. Agric. Syst. 127 (May 81-89), 2014. https://doi.org/10.1016/j.agsy.2014.01.008.

Ye, Q., Lin, X., Adee, E., Min, D., Assefa Mulisa, Y., O'Brien, D., Ciampitti, I.A., 2017. Evaluation of climatic variables as yield-limiting factors for maize in Kansas. Int. J. Climatol. 37 (S1), 464–475. https://doi.org/10.1002/joc.5015.

Yoo, G., Wander, M.M., 2006. Influence of tillage practices on soil structural controls over carbon mineralization. Soil Sci. Soc. Am. J. 70, 651–659. https://doi.org/10.2136/sssaj2005.0036.

Yost, M.A., Coulter, J.A., Russelle, M.P., Sheaffer, C.C., Kaiser, D.E., 2012. Alfalfa nitrogen credit to first-year corn: potassium, regrowth, and tillage timing effects. Agron. J. 104, 953–962. https://doi.org/10.2134/agronj2011.0384.

Yost, M.A., Coulter, J.A., Russelle, M.P., 2013. First-year corn after alfalfa showed no response to fertilizer nitrogen under no-tillage. Agron. J. 105, 208–214. https://doi.org/10.2134/agronj2012.0334.

Yost, M.A., Morris, T.F., Russelle, M.P., Coulter, J.A., 2014. Second-year corn after alfalfa often requires No fertilizer nitrogen. Agron. J. 106, 659–669. https://doi.org/10.2134/agronj2013.0362.

Zhang, Y., Ling, C.A., 2018. Strategy to apply machine learning to small datasets in materials science. NPJ Comput. Mater 4, 25. https://doi.org/10.1038/s41524-018-0081-z.

Zhang, Y., Yang, Y., 2015. Cross-validation for selecting a model selection procedure. J. Econom. 187 https://doi.org/10.1016/j.jeconom.2015.02.006.

Zhang, F., Sun, Q., Lin, S.J., Chen, J.H., Emanuel, K., 2019. What is the predictability limit of mid-latitute weather? J. Atmos. Sci. 76, 1077–1091. https://doi.org/10.1175/JAS-D-18-0269.1.

Zhang, B., Niu, X., Ata-UI-Karim, S.T., Wang, L., Duan, A., Liu, Z., Lemaire, G., 2020. Determination of the post-anthesis nitrogen status using ear critical nitrogen dilution curve and its implications for nitrogen management in maize and wheat. Eur. J. Agron. 113, 125967. https://doi.org/10.1016/j.eja.2019.125967.

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.