Machine learning improves predictions of agricultural nitrous oxide (N 2 O) emissions from intensively managed cropping systems

Article *in* Environmental Research Letters · December 2020

DOI: 10.1088/1748-9326/abd2f3

CITATIONS
9

READS
137

3 authors, including:

Debasish Saha
University of Tennessee
24 PUBLICATIONS **452** CITATIONS

SEE PROFILE

G Philip Robertson
Michigan State University
366 PUBLICATIONS **29,117** CITATIONS

SEE PROFILE

ENVIRONMENTAL RESEARCH
LETTERS

ACCEPTED MANUSCRIPT • OPEN ACCESS

Machine learning improves predictions of agricultural nitrous oxide (N_2O) emissions from intensively managed cropping systems

To cite this article before publication: Debasish Saha *et al* 2020 *Environ. Res. Lett.* in press https://doi.org/10.1088/1748-9326/abd2f3

Machine learning improves predictions of agricultural nitrous oxide (N_2O) emissions from intensively managed cropping systems

Debasish Saha[*1,2,†], Bruno Basso[1, 2, 3], and G. Philip Robertson[1,2]

[1] W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060

[2] Department of Plant, Soil, and Microbial Sciences and Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48824

[3]Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI 48824

Corresponding Author: Debasish Saha[†] dsaha3@utk.edu

Running Title: Machine Learning improves N_2O predictions

Keywords: nitrogen, nitrous oxide, agriculture, greenhouse gas, corn, fertilizer, machine learning.

Author Contributions: D.S. and G.P.R. designed research; B.B. provided process model data; D.S. analyzed data; D.S. and G.P.R. wrote the paper to which all authors contributed.
The authors declare no competing financial interests.

[†] Current address: Department of Biosystems Engineering and Soil Science, University of Tennessee, Knoxville, TN 37996. Phone: 865-974-7003, Email: dsaha3@utk.edu

Abstract

The potent greenhouse gas nitrous oxide (N_2O) is accumulating in the atmosphere at unprecedented rates largely due to agricultural intensification, and cultivated soils contribute ~60% of the agricultural flux. Empirical models of N_2O fluxes for intensively managed cropping systems are confounded by highly variable fluxes and limited geographic coverage; process-based biogeochemical models are rarely able to predict daily to monthly emissions with > 20% accuracy even with site-specific calibration. Here we show the promise for machine learning (ML) to significantly improve field-level flux predictions, especially when coupled with a cropping systems model to simulate unmeasured soil parameters. We used sub-daily N_2O flux data from six years of automated flux chambers installed in a continuous corn rotation at a site in the upper U.S. Midwest (~3000 sub-daily flux observations), supplemented with weekly to biweekly manual chamber measurements (~1100 daily fluxes), to train an ML model that explained 65-89% of daily flux variance with very few input variables –soil moisture, days after fertilization, soil texture, air temperature, soil carbon, precipitation, and N fertilizer rate. When applied to a long-term test site not used to train the model, the model explained 38% of the variation observed in weekly to biweekly manual chamber measurements from corn, and 51% upon coupling the ML model with a cropping systems model that predicted daily soil N availability. This represents a 2-3 times improvement over conventional process-based models and with substantially fewer input requirements. This coupled approach offers promise for better predictions of agricultural N_2O emissions and thus more precise global models and more effective agricultural mitigation interventions.

1 INTRODUCTION

Atmospheric nitrous oxide (N_2O) concentrations are currently increasing by 0.98 ppb yr^{-1}, a rate 44% higher than during the period 2000-2005 (0.68 ppb yr^{-1}). This dramatic shift is largely driven by increased anthropogenic sources that lift current total global N_2O emissions to ~17 Tg N yr^{-1} (Syakila & Kroeze, 2011; Thompson et al., 2019). This represents a global warming impact of ~2 Pg $C_{equivalents}$ yr^{-1} based on N_2O's global warming potential (IPCC, 2013; Robertson, 2004). Approximately 60% of the contemporary N_2O increase comes from cultivated soils that receive nitrogen (N) fertilizer (IPCC, 2014; Robertson, 2014; Tian et al., 2019, 2020), but this value involves large uncertainty because the episodic nature of soil N_2O fluxes challenges our ability to quantify emissions accurately (Barton et al., 2015; Parkin, 2008; Saha et al., 2017).

Soil N_2O fluxes are of microbial origin, mainly from incomplete denitrification (the sequential reduction of nitrite or nitrate to N_2O then N_2) and nitrification (the oxidation of ammonium to nitrate). A number of environmental and management factors affect N_2O production, in particular soil N inputs and N availability, pH, soil moisture, tillage, and temperature (Butterbach-Bahl et al., 2013). Furthermore, N_2O can also be consumed in soil (Schlesinger, 2013), and controlling factors often interact in a non-linear fashion to add considerable spatial and temporal variability to fluxes (Dobbie & Smith, 2003; Jin et al., 2017).

Quantitative predictions of daily to monthly N_2O fluxes are especially challenging. Commonly used methods include i) empirical equation-based parametric models (Roelandt et al., 2005; Sozanska et al., 2002), ii) emission factor models based on the proportion of N inputs emitted as N_2O (De Klein et al., 2006), and iii) process-based biogeochemical models (Brilli et al., 2017; Ehrhardt et al., 2018; Gaillard et al., 2018). Each prediction method has its own

limitations with accompanying bias. For example, equation-based parametric models require N_2O fluxes to be normally distributed with homogeneous variance, conditions rarely met by N_2O flux data even when transformed. The IPCC's 1% emission factor is a useful but simple estimate of annual N_2O emissions based largely on annual fertilizer N inputs, thus ignoring intra-annual variation and the dynamic variable interactions that influence emissions.

The IPCC approach often underestimates local (Philibert et al., 2012; Shcherbak et al., 2014) and regional (Grace et al., 2011; Griffis et al., 2013) N_2O emissions, with variable success at the global scale (Crutzen et al., 2008; Del Grosso et al., 2008a; Tian et al., 2020). In contrast, process-based models predict daily to weekly fluxes, important for designing management interventions to mitigate emissions; however, they are complex and heavily rely on site and version-specific parameterizations that are sometimes ad hoc tunings (Gilhespy et al., 2014; Del Grosso et al., 2008b). Parameterizations add additional uncertainty when extending such models to sites without reliable model calibration data (Berardi et al., 2020; Fuchs et al., 2020; Ehrhardt et al., 2018; Gaillard et al., 2018; Jarecki et al., 2008; Rafique et al., 2013). Moreover, the N_2O algorithms of popular biogeochemical models such as DayCent (Del Grosso et al., 2000; Parton et al., 2001), DNDC (Li, 2000, 2007) and EPIC (Izaurralde et al., 2012), widely used for predicting regional N_2O budgets, are usually derived from laboratory-based responses of N_2O to individual environmental factors. This necessarily (and by design) simplifies the complex variable interactions typical of field settings but adds to the need for calibration with intensive chamber-based field measurements.

Although the lack of agreement between observed and predicted N_2O fluxes by current prediction methods are widely known (Ehrhardt et al., 2018), the development of improved alternative prediction approaches has been limited. While top-down models have benefited

annual N_2O flux predictions (Perlman et al., 2014; Philibert et al., 2013), such approaches have not been available for daily to weekly predictions due to a general lack of N_2O measurements at finer temporal scales. That said, the increasing availability of long-term high frequency observations from automated flux chambers (Grace et al., 2020) creates new opportunities to develop data-driven top-down models that can improve predictability and our understanding of the interacting factors and threshold conditions controlling daily N_2O fluxes.

Here we demonstrate a novel application of a data-driven machine learning (ML) model (Random Forest) to predict N_2O fluxes from two sites in the upper U.S. Midwest growing corn (*Zea mays* L.). Corn is responsible for ~60% of U.S. N fertilizer use, and ~60% of U.S. corn is grown in the Midwest, a region of high N_2O emissions (ERS, 2019a; 2019b; Larsen et al., 2007). We use multi-year automated N_2O flux data at one site together with observations from less dense manual chambers at both sites to deduce causal relationships among key variables to predict fluxes with known confidence. This hierarchically divergent approach requires no knowledge of underlying process-level relationships and thus is less able than process-based models to predict the effects of novel management changes (e.g. broadcast vs injected N fertilizer). Model training is also data intensive. Nevertheless, because the ML system learns relationships from available data it can make unbiased predictions, and as well provide insights into functional relationships by discriminating among different predictor variables as it learns patterns and identifies critical response thresholds. We further test a coupled model that uses ML to predict N_2O fluxes based on a process-level cropping systems model's provision of an additional predictor variable, soil N availability.

2 MATERIALS AND METHODS

2.1 Experimental systems

We used data from three experiments in two geographic locations with different soil types. The first experiment is a continuous no-till corn system in the Biofuel Cropping Systems Experiment (BCSE) at the W. K. Kellogg Biological Station (KBS-BCSE, 42°23′43″N, 85°22′24″W, 288-m elevation) in Michigan, USA. KBS-BCSE soils are well-drained Kalamazoo series Alfisols (fine-Loamy, mixed, semiactive, mesic Typic Hapludalfs) with 1.2% soil organic carbon (Table 1); climate at the site is humid continental with a mean annual precipitation and temperature of 1027 mm and 10°C, respectively. The second experiment is also continuous no-till corn, located in the BCSE at the Arlington Agricultural Research Station (ARL-BCSE, 43°17′45″N, 89°22′48″W, 315-m elevation) in Wisconsin, USA. The ARL-BCSE soils are well-drained Plano series Mollisols (fine-silty, mixed, superactive, mesic Typic Arguidolls) with 4% soil organic carbon; climate is humid continental with a mean annual precipitation and temperature of 869 mm and 6.8°C, respectively. Both BCSE experiments were established in 2008 as part of the US Department of Energy's Great Lakes Bioenergy Research Center and management, site histories, and other details are reported in detail elsewhere (Gelfand et al., 2020; Sanford et al., 2016). The third experiment is a no-till corn-soybean (*Glycine max* L.)-winter wheat (*Triticum aestivum* L.) rotation at the KBS Long-term Ecological Research (LTER) site (KBS-LTER; 42° 24' N, 85° 24' W, 288-m elevation), 2 km from KBS-BCSE, established in 1989 (Robertson & Hamilton, 2015). Details on agronomic management can be found in https://lter.kbs.msu.edu/datatables.

TABLE 1 Daily mean N_2O fluxes, soil, and environmental conditions in the studied sites. The values in parenthesis are standard deviation from mean. The data from KBS-BCSE and ARL-BCSE were used to train the model that was then tested on KBS-LTER.

Sites	Mean N_2O (g N ha^{-1} d^{-1})	Clay (g kg^{-1})	SOC (%)	NH_4^+-N (kg ha^{-1})	NO_3^--N (kg ha^{-1})	Air T (°C)	$SPrecip_{2d}$† (mm)	n*
KBS-BCSE (corn)	5.2 (15.9)	63	1.2	11.1 (7.6)	27.9 (17.0)	11.2(10.3)	4.9 (9.1)	1094
ARL-BCSE (corn)	18.3 (54.6)	261 (16.8)	4.0 (0.3)	16.1 (24.1)	36.2 (41.1)	12.4 (9.5)	6.5 (11.8)	482
KBS-LTER $_{Rotation}$	3.4 (7.5)	161 (20.2)	1.6 (0.3)	9.2 (7.9)	12.7 (12.5)	14.4 (7.6)	4.8 (10.2)	670
Corn	4.2 (8.8)	161	1.6	12.8 (10.3)	18.3 (16.5)	13.1 (8.7)	5.7 (12.4)	269
Soybean	2.5 (4.8)	161	1.6	6.0 (2.8)	9.5 (7.4)	15.0 (6.7)	3.7 (8.5)	204
Wheat	3.2 (7.8)	161	1.6	7.5 (5.1)	8.6 (6.3)	15.4 (6.7)	4.6 (8.1)	197

† Cumulative precipitation received in the past two days (see Table 2 for variable description)

* Total number of daily observations for each site

2.2 Data generation

N_2O fluxes from KBS-BCSE were measured using automated gas flux chambers (50 × 50 × 18 cm) as described in Ruan & Robertson (2017) from 2012 to mid-July of 2017, except not in 2015 due to instrument failure. Fluxes were measured in one replicate block from three chamber locations within a plot (27 × 43 m). Fluxes were measured from one location on any given day and locations were randomly re-located every 10-15 days to minimize bias due to small-scale spatial variability. Each chamber was sampled four times per day; the daily average flux was used for this analysis, previously shown to approximate diurnal flux variability in the Midwest (Parkin 2008). At each sampling time, the chamber was closed for 45 minutes and headspace samples of 1 L each were taken at 0, 15, 30, and 45 minutes after chamber closure. Headspace samples traveled through a Teflon sampling line to a nearby (< 80 m distant) trailer housing a gas chromatograph with an electron capture detector (350°C) (SRI 8610C with custom sample acquisition, Torrance, CA, USA) to measure N_2O concentrations, and an in-line infra-red gas analyzer (LI-820, LI-COR Biosciences, NE, USA) to measure CO_2 concentrations. N_2O fluxes from ARL-BCSE (five replicate blocks) and KBS-LTER (four replicate blocks) were measured using static chambers (28.5 cm diameter, ~17 cm height) on a weekly to biweekly sampling frequency as described elsewhere (Duncan et al., 2019; Gelfand et al., 2016; Oates et al., 2016).

The predictor variables avoid co-linearity as they were uncorrelated (Table 2, Figure S1), and their distribution is shown in Figure S2. The water-filled pore space (WFPS), an approximation of soil O_2 limitation, was estimated from volumetric water content (VWC) and bulk density in the top 25-cm soil layer. The VWC was continuously measured at the KBS-BCSE site by a data logger-controlled sensor network as reported elsewhere (Hamilton et al.,

2015); at ARL-BCSE and KBS-LTER, VWC was estimated from bulk density and gravimetric soil moisture values measured at each manual gas sampling event.

To test the additional power of adding to the model daily available soil N, measured only with manual sampling events, we used SALUS, a process-based model extensively validated at KBS (Basso & Ritchie, 2015; Hamilton et al., 2020; Hussain et al., 2019), to simulate values for soil NH_4^+-N and NO_3^--N in the top 25-cm layer at KBS-BCSE. SALUS is an easy to calibrate functional model designed to represent feedbacks and interactions among crop, soil, management, genotype, and climate; avoiding known problems with accurately simulating daily crop growth and soil water balances by many biogeochemical models such as DayCent and DNDC (Brilli et al., 2017; Fuchs et al., 2020; Jarecki et al., 2008). For ARL-BCSE and KBS-LTER, we used measured soil inorganic N data interpolated between the soil sampling dates to match the manual gas sampling dates. Static soil variables included soil organic carbon and clay content. Weather variables included precipitation and air temperature, obtained from weather stations within 1 km of each site. Missing data on predictor variables were imputed using the proximity matrix within the Random Forest algorithm.

2.3 Data processing

Each flux event was checked for linearity of headspace N_2O and CO_2 accumulations during the chamber closure period, and periods with leakage (< 3% of all fluxes data) were deleted. We converted fluxes between zero and the automated chamber's N_2O detection limit (±1.1 g N_2O-N $ha^{-1}d^{-1}$) to 50% of the detection limit. Approximately 11% of the total number of N_2O fluxes (2246) were negative and the lowest value was -7.5 g N_2O-N $ha^{-1}d^{-1}$. The highest measured flux was 593 g N_2O-N $ha^{-1}d^{-1}$.

TABLE 2 Abbreviations and description of the predictor variables used in the Random Forest modeling

Abbreviated variable	Variable category	Description	Measurement unit
WFPS	Soil†	Water-filled pore space	Fraction
NH_4^+-N	Soil	Ammonium N content	kg NH_4^+-N ha^{-1}
NO_3^--N	Soil	Nitrate N content	kg NO_3^--N ha^{-1}
Clay	Soil	Clay content	g kg^{-1}
SOC	Soil	Soil organic carbon	%
DAF	Management	Days after N fertilization	Days
N rate	Management	N fertilizer application rate	kg N ha^{-1}
Air T	Weather	Mean daily air temperature	°C
$SPrecip_{2d}$	Weather	Cumulative precipitation in last two days	mm

† Soil variables were measured at the top 25-cm soil layer

2.4 Machine Learning Algorithm

We used R statistical software (R Core Team, 2018) to implement Random Forest, a supervised ML algorithm for classification and regression based on the principle of recursive partitioning (Breiman, 2001), and independent of the assumption of functional relationships between the response and predictor variables. A detailed description of the Random Forest algorithm can be found in Hoffman et al. (2018). Briefly, Random Forest analysis ensembles numerous regression and classification trees following a process called "bootstrap aggregation" or "bagging." Classification trees are relatively uncorrelated in two ways. First, a random subset of the data space is drawn (with replacement) to grow a tree to its full length, and each node of the tree group's observations are characterized by certain conditions on the predictor variables to produce an average prediction for the response variable. Each tree growing process uses only two-thirds of the bootstrapped data and one-third of the observations (out-of-bag data, OOB) are used for

estimating the prediction errors. Second, each node split in a tree considers a random subset of predictor variables, usually a square root of the total number of predictor variables. The predictions from all the trees are averaged to make final predictions.

The variable importance function within the Random Forest algorithm ranks predictor variables based on the increase in model error by randomly permuting the values of the predictor variables. Briefly, the mean square error of the OOB data (OOB-MSE) for each tree is the average squared deviations of OOB observations from the predictions. The difference in OOB-MSE before and after random permutation of a predictor variable is averaged over all trees to compute variable importance. Top predictors were visualized using partial dependence and feature contribution plots using the Forest Floor package to unravel linear or non-linear interactions (Welling et al., 2016).

We used one dimensional Individual Conditional Expectation (ICE) plots that show the effect of a predictor feature on the response variable for all instances as well as a global average effect i.e. the partial dependence of the response on that feature (Friedman, 2001; Goldstein et al., 2015). The partial dependence plot allows us to identify whether the relationship between the predictor and response variable is linear, monotonic, or interactive (i.e., the pattern changes after a certain threshold of the predictor variable). The feature contribution which computes the contribution of a variable to Random Forest prediction with numerical observations, belongs to the same family of partial dependence plots. Unlike the variable importance, feature contribution is separately computed for each observation and allows us to apply a color gradient based on a predictor variable to identify latent interactions and correlations between predictor variables (Kuz'min et al., 2011; Welling et al., 2016).

2.5 Model building

Data used for model building included 1576 daily observations from continuous corn rotations at KBS-BCSE and ARL-BCSE; we used 670 daily observations collected from 2002 to 2014 in a corn-soybean-wheat rotation at KBS-LTER for independent model testing (Table 1, Figure S3). The N_2O fluxes within the model-building data were divided into eight bins by merging some of the bins in Figure S4 containing only a few high flux observations to assure representative sampling from the entire range of N_2O fluxes; 70% of the randomly selected observations within each bin were used to train the model (training data) and the model was validated with the remaining data. We used the Synthetic Minority Over Sampling Technique for regressions to correct multiclass data imbalance in the training set (Text S1; Chawla et al., 2002; Torgo et al., 2013); however, the number of observations in the highest domain of N_2O fluxes > 100 g N $ha^{-1}d^{-1}$ were still limited (Figure S4).

The balanced training data were used to train two Random Forest models based on their representation of soil N availability. A Coupled model included both measured and SALUS predicted soil available N (see above); the Standard model included measured data only and used Days after fertilizations (DAF) as a surrogate of soil N availability. For each model we optimized the number of trees used in the forest (n_{tree}) and the number of variables considered at each node (m_{try}). The n_{tree} = 800 best stabilized error and m_{try} = 4 resulted in the greatest error reduction for the out-of-bag observations (Figure S5). The Random Forest models were tested on KBS-LTER data and model performance was assessed by root mean square errors (RMSE), coefficients of determination (R^2), and correlation coefficients (r) between observed and predicted daily N_2O fluxes.

3 RESULTS

3.1 N_2O flux variability, soils, and environment

The three studied sites exhibited large inter-annual variability in N_2O fluxes during their respective measurement years (Figure 1). Long-term average daily N_2O fluxes were in the order ARL-BCSE (18.3) > KBS-BCSE (5.2) > KBS-LTER (3.4 g N ha^{-1} d^{-1}) and widely varied within sites from -0.5 to 593, -7 to 237, and -4 to 89 g N ha^{-1} d^{-1}, respectively (Figure 1, Table 1). The highest long-term daily average N_2O flux from corn at ARL-BCSE (18.3 g N ha^{-1} d^{-1}) was perhaps due to this site's greater soil N availability stemming from a high soil organic matter content and due to greater anoxia in its finer textured soils (Table 1). The Mollisols at the ARL-BCSE in Wisconsin had three times higher SOC content than the Alfisols at the KBS sites in Michigan. Unlike the highly fertilized continuous corn systems at the BCSE sites, diverse crop rotation including unfertilized soybean and low-input winter wheat at KBS-LTER site lowered daily average fluxes. Furthermore, bi-weekly sampling frequencies for manual chambers at ARL-BCSE and KBS-LTER may have missed peak emissions events, in particular at KBS-LTER with its relatively narrow range of fluxes.

3.2 Model performance

We trained two Random Forest models to predict N_2O fluxes. A Standard model used readily available soil, agronomic, and weather variables as model inputs and a Coupled model included additional high frequency soil N data generated by SALUS (Table 2). We expected simulated soil N pools to improve estimates of available N based otherwise on N fertilizer rate and DAF in the Standard model.

Overall, both the Standard and Coupled models performed well in predicting temporal N_2O fluxes from corn at the KBS-BCSE and ARL-BCSE validation sites using 70% of the observations for model training (Figure 1). The models explained 65 to 89% of the variability between observed and predicted N_2O fluxes that were highly correlated ($r > 0.80$, $p < 0.001$). However, the Standard model over-predicted N_2O fluxes on several occasions at the KBS-BCSE site, which resulted in higher prediction error than the Coupled model (RMSE of 10.2 vs 8.3). Surprisingly, however, the Standard model outperformed the Coupled model at ARL-BCSE (RMSE 15.8 vs 21), where the Coupled model over-predicted fluxes in 2010 and 2011. At the KBS-LTER test site, the Coupled model explained 51% variability of observed N_2O fluxes in corn at KBS-LTER, for which no data were used to train the model. The standard model explained 38% of observed N_2O fluxes from corn. For the entire corn-soybean-winter wheat rotation at KBS-LTER, the Coupled model explained 38% variability, whereas the Standard model predicted only 13% variability.

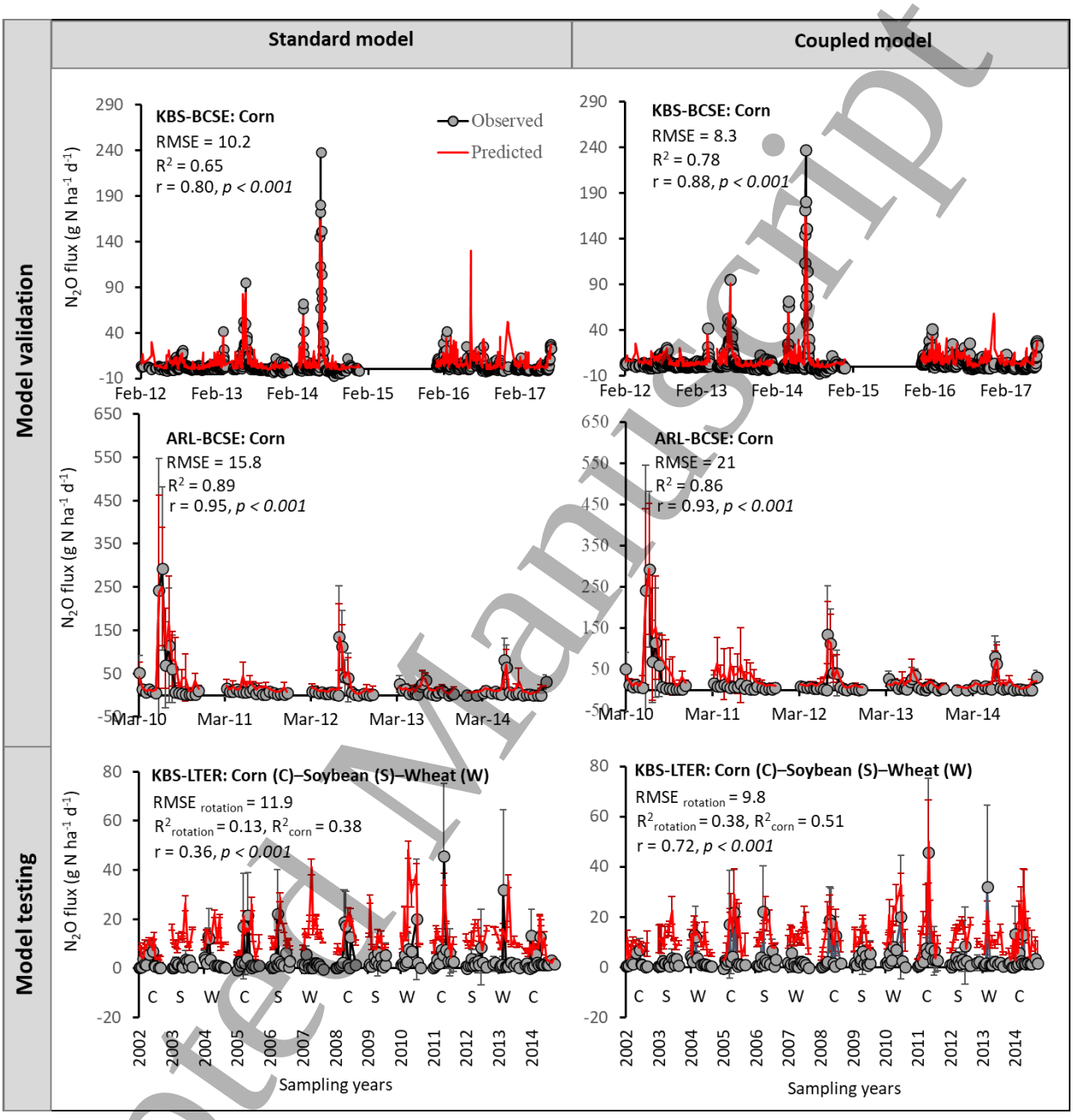

FIGURE 1 Observed vs. Random Forest predicted N_2O fluxes at the validation (KBS-BCSE and ARL-BCSE) and test (KBS-LTER) sites. The left and right columns show the predictions from the Standard and Coupled Random Forest models, respectively. Inorganic N availability is represented in the Standard model as Days after fertilization (DAF) and in the Coupled model as both measured and modeled (by SALUS) NH_4^+-N and NO_3^--N availability. 70% of the data from

KBS-BCSE and ARL-BCSE were used to build the Random Forest models and 30% were used for validation. KBS-LTER data were used for independent model testing.

3.3 Critical predictor variables for N_2O emissions

The most important variables influencing N_2O fluxes for both Standard and Coupled models were soil moisture (WFPS) and N availability (DAF for the Standard model, inorganic N for the Coupled model), which together explained > 80% of the N_2O flux variance in the out-of-bag samples (Figure 2). Other soil, environment, and management related variables had relatively smaller predictive values.

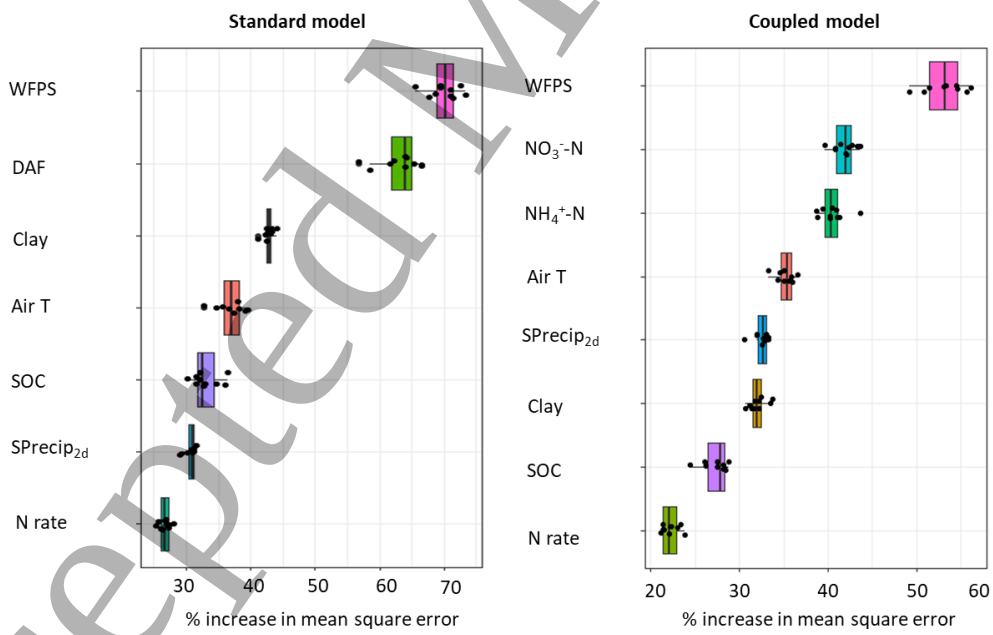

FIGURE 2 Importance of variables controlling N_2O fluxes as predicted by Standard and Coupled Random Forest models. The variables associated with the greatest % increase in mean square error are the most important variables. Inorganic N availability is represented in the

simple model as Days after fertilization (DAF) and in the Coupled model as simulated NH_4^+-N and NO_3^--N) availability. The Standard and Coupled models explained 86 and 84% variability in N_2O fluxes in the out-of-bag observations, respectively. Each box corresponding to a variable represents the uncertainty of variable importance scores computed by the Random Forest model built on ten random subsamples of the training data.

The Individual Conditional Expectation plot visualizes the average predictor variable effect on prediction (partial dependence) along with the spread for each instance in the data (Figure 3). Both models identified that for most instances, N_2O fluxes sharply increased after WFPS exceeded 0.70. The feature contribution plot showed that the Standard model predicted an exponential decrease in N_2O fluxes with increasing DAF in interactions with WFPS (Figure 4b). Occurrence of high WFPS within a month after N fertilization increased N_2O fluxes. Soil clay content > 250 g kg^{-1} increased N_2O for some instances, but lack of range on this variable in our data limits our interpretation (Figure S2h). Average N_2O fluxes showed threshold responses to NO_3^--N and NH_4^+-N contents with a much smaller value for NH_4^+-N (~30 kg NH_4^+-N ha^{-1}, Figure 4d, e). However, heterogeneous relationships between N_2O and NO_3^--N were observed for many instances (Figure 3e), indicating possible interactions with WFPS. Higher WFPS values variably influenced N_2O fluxes up to 125 kg NO_3^--N ha^{-1} and sharply increased beyond that threshold (Figure 4). At the same time, WFPS > 0.75 negatively contributed to N_2O fluxes when soil ammonium was below 30 kg NH_4^+-N ha^{-1}; however, both moderate to high soil moisture content (0.6 to 0.9 WFPS) positively contributed to fluxes beyond 30 kg NH_4^+-N ha^{-1}.

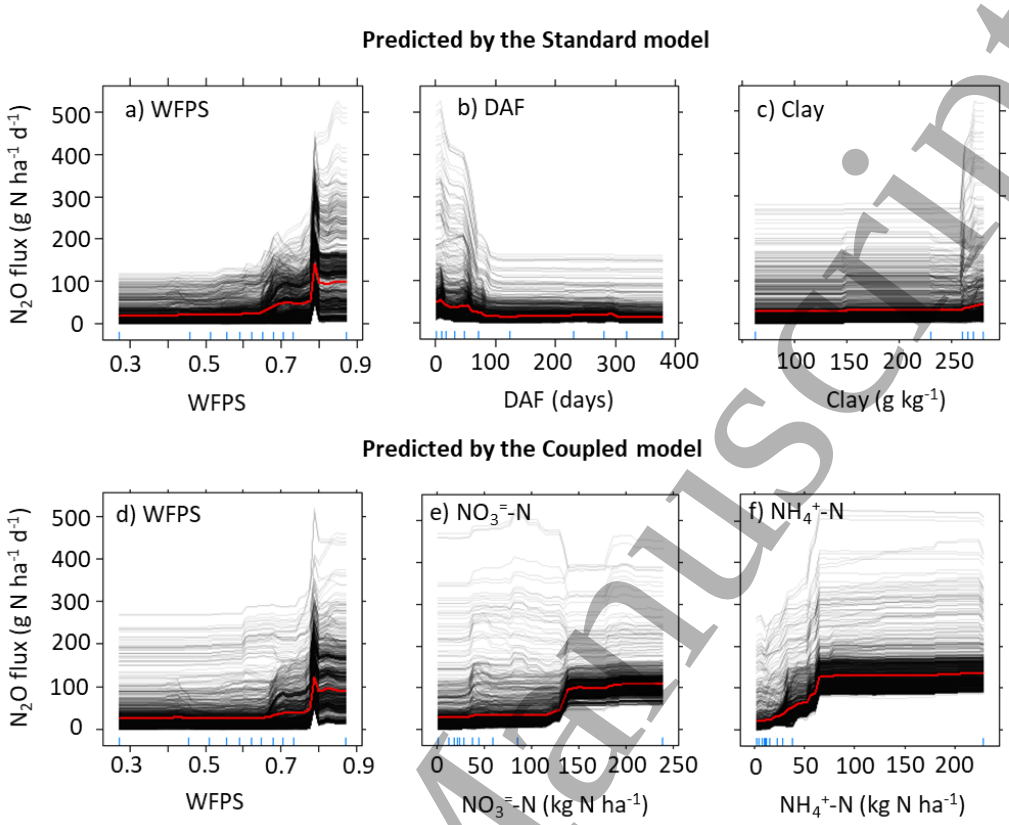

FIGURE 3 Individual Conditional Expectation plots of predicted N_2O fluxes in response to the top three predictor variables as identified by the Standard (a, b, c) and Coupled (d, e, f) model. In each panel, each dark line represents one data instance in the training set. The red line is the partial dependence line, an average of all the instances.

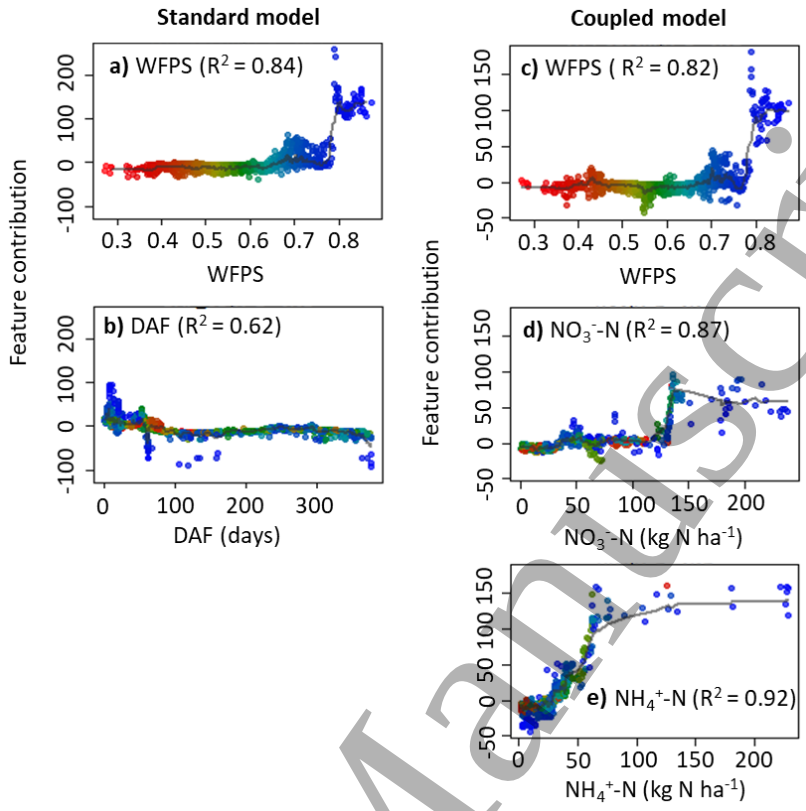

FIGURE 4 Percent change in the probability of accurately predicting N_2O fluxes (feature contribution) for water filled pore space (WFPS; a, c) and its interactions with N availability: Days after fertilization (DAF; b) and inorganic N pools (d, e) as predicted by the Standard (a, b) and Coupled (c, d, e) Random Forest models. The color gradient from red to blue represents dry to wet soil conditions based on WFPS (a, c). R^2 refers to leave-one-out goodness of fit of the average contribution line (denoted in black).

4 DISCUSSION

Our results demonstrate a high utility for using ML to predict agricultural soil N_2O emissions. The Random Forest models were developed using both automated and manual chamber based N_2O flux data from corn production at two sites with different soils and explained 65-89% of

daily flux variance at the training sites. When extended to a long-term test site with a very different cropping system it explained up to 51% of emissions variation in corn, especially when coupled with a cropping systems model that provided daily inorganic N values.

4.1 Predictability and interpretability of ML models

Our ability to assimilate data in traditional ways has not kept pace with the increasing availability of long-term N_2O data from automated chambers and advanced sensor technology. Machine learning provides the potential to efficiently use such data to generate new insights and derive predictive models that emulate high resolution fluxes. When tested against 13 years of weekly to biweekly flux measurements from an independent test site (KBS-LTER), our Coupled ML model accounted for 51% of the variability of daily average N_2O fluxes for corn phase emissions and 38% for the entire corn-soybean-wheat rotation. This effort represents a 2-3 times improvement over conventional process-based models and with substantially fewer input requirements.

The most commonly used process-based models DayCent (Del Grosso et al., 2000; Parton et al., 2001), DNDC (Li, 2000, 2007), and their variants (Table S1) explain, on average, 20% variability of temporal N_2O fluxes in 15 cropping system studies (71 observations) that we reviewed and that explicitly reported model performance (median of 17%, Figure S6). A recent modeling study using 24 process-based models showed equally large uncertainties for an ensemble model (Ehrhardt et al., 2018). The improved performance of the ML model is due to its depending on functional relationships between predictor and dependent variables as learned from the data (Breiman, 2001), rather than depending on an underlying process-level understanding of flux variability. This does not mean that ML models are inherently superior – indeed, their

predictive ability is strictly correlational, and thus of limited use for predicting the effects of novel conditions, an important feature of process-based models with their superior understanding of biophysical interactions. Rather, ML models might be used to better scale existing knowledge and perhaps help to optimize process-based models by better identifying the key variables interactions responsible for driving episodic fluxes.

Indeed, in our model, key predictor variables are few (Table 2, Figure 2): two static soil properties (soil organic carbon and clay content) and four dynamic properties (WFPS, N availability (expressed either as DAF or inorganic N pools), temperature, rainfall, and N fertilizer rate). All of these factors are well known drivers of N gas emissions (Conen et al., 2000; Firestone & Davidson, 1989; Groffman et al., 1988). These input variables are easy to measure proxies for soil biophysical processes.

WFPS, for example, is an integrative measure of water availability and soil gas diffusivity (Linn & Doran, 1984), and stands out here as the strongest individual predictor of N_2O fluxes (Figures 2, 3). The strong interaction of WFPS with N availability (either DAF or inorganic N pools; Figure 4) reflects the greater risk of significant N_2O emissions following precipitation events close to N fertilization, reported in many other studies (e.g., Parkin & Kaspar, 2006; Saha et al., 2017). The lower predictive value for fertilizer rate is surprising and probably reflects the relatively narrow range of fertilizer rates (170 to 200 kg N ha^{-1} yr^{-1}) used in our training data, which include only corn (Figure 2). We might expect a broader range of fertilizer managements to elevate the predictive capacity of N rate, well known for its ability to mitigate N_2O emissions from fertilized cropping systems (IPCC, 2014; Mikkelsen & Snyder, 2012; Millar et al., 2010; Shcherbak et al., 2014). On the other hand, inorganic N species availability (the second most important variable) may represent a sufficient proxy of the

integrative effect of management (fertilization) and microbial N transformations. Future development of our knowledge and data availability on N_2O response to microbial community composition and activity may provide additional predictive capacity for the ML models.

4.2 Challenges of Machine Learning models for predicting N_2O fluxes

Prior efforts to use ML to predict N_2O fluxes have used cumulative fluxes as estimated from literature reviews (Philibert et al., 2013) or process-level models (Perlman et al., 2014; Villa-Vialaneix et al., 2012). To our knowledge, the present effort is the first to use ML to predict highly resolved daily fluxes – especially important for designing management interventions to mitigate high fluxes and perhaps for building better process-based models. Yet there remain important challenges before ML models can be used more broadly to predict regional N_2O emissions.

The first challenge is to increase the model's generalizability. It would be naïve to believe that our ML model might be extrapolable to diverse soils, climates, and production systems given that the model was trained on data from a single type of cropping system with narrow ranges for predictor variables measured within a constrained biogeographic context. Machine learning predictions are bound to the data range in the training set, such that the model will, for example, underestimate N_2O fluxes beyond 593 g N ha^{-1} d^{-1}, the maximum value in our training data (Figure 1; Figure S4). Moreover, DAF is relevant only for N-fertilized systems. These limitations are illustrated by the model's lower predictability for fluxes from the whole rotation relative to the corn phase at KBS-LTER. For example, the Standard model identified DAF as the second most important variable, ranging from 1 to 378 days in the training data (annually fertilized corn); however, DAF ranged from 1 to 678 days for the KBS-LTER corn-

soybean-wheat rotation at the test site, with N fertilization not present during the soybean phase of this rotation. Likewise, overestimation of N_2O fluxes from the wheat phase of this rotation is likely due to the model's low temperature range – winter wheat is fertilized in early spring when soil temperatures are relatively low. The Coupled Model's use of soil inorganic N instead of DAF overcame this problem to some extent (Figure 1).

The second difficulty is imbalance in our training data. Only 2% of total observations were associated with N_2O fluxes > 50 g N ha^{-1} d^{-1}, an arbitrary threshold that represents high flux events (Figure S4). This limits the ML model's opportunity to learn critical variable interactions promoting episodic emission events. To avoid this limitation, we attempted to balance the distribution of N_2O flux data by dividing the training data into bins based on N_2O fluxes, and then respectively over- and under-sampled the minority (high N_2O flux observations) and majority (low N_2O flux observations) bins (Figure S4). However, this may not have eliminated the imbalance problem and may be a reason that the ML models over and under predicted N_2O fluxes that exhibited murky relationships with all predictor variables (Figure S7).

Both of these problems can be addressed by the addition of data sets from more diverse managements and geographies. Greater variability in training data (rather than more data from similar sites) will be key to developing more generalizable models. This can be achieved by coordinating N_2O research to combine cross-site observations that follow a consistent protocol for predictor variables (Almaraz et al., 2020; Borer et al., 2014).

4.3 Coupled machine learning and process-based models

Our results provide formative evidence for improving the efficiency and accuracy of N_2O predictions by integrating ML and process-based modeling approaches. Traditionally,

biogeochemical models are guided by process-level theory while ML models are data-driven. Their fusion is gaining attention in Earth and Geosciences (Brenowitz & Bretherton, 2018; Karpatne et al., 2017a, 2017b; Reichstein et al., 2019) but is in its infancy in terrestrial biogeochemistry. We show that ML predictions of a temporally dynamic soil biogeochemical processes such as N_2O fluxes can be improved by incorporating data produced by a well-validated process-based model simulating soil-plant-atmosphere processes – in our case, by predicting inorganic N pools.

This approach provides two important synergies for N_2O prediction: First, replacing the N_2O subroutine in the process-based model with an ML subroutine could provide a coupled model with greater predictive power (Figure 5). This, however, requires the ML model to be trained on diverse input data with a wide prediction range. For example, process-based models typically predict soil water, temperature, and N availability with acceptable fidelity in a wide variety of cropping systems and geographies. This information could be contributed to the ML sub-model as input variables to predict N_2O. This approach has been efficiently used in atmospheric and ocean science (de Bezenac et al., 2017; Karpatne et al., 2017b).

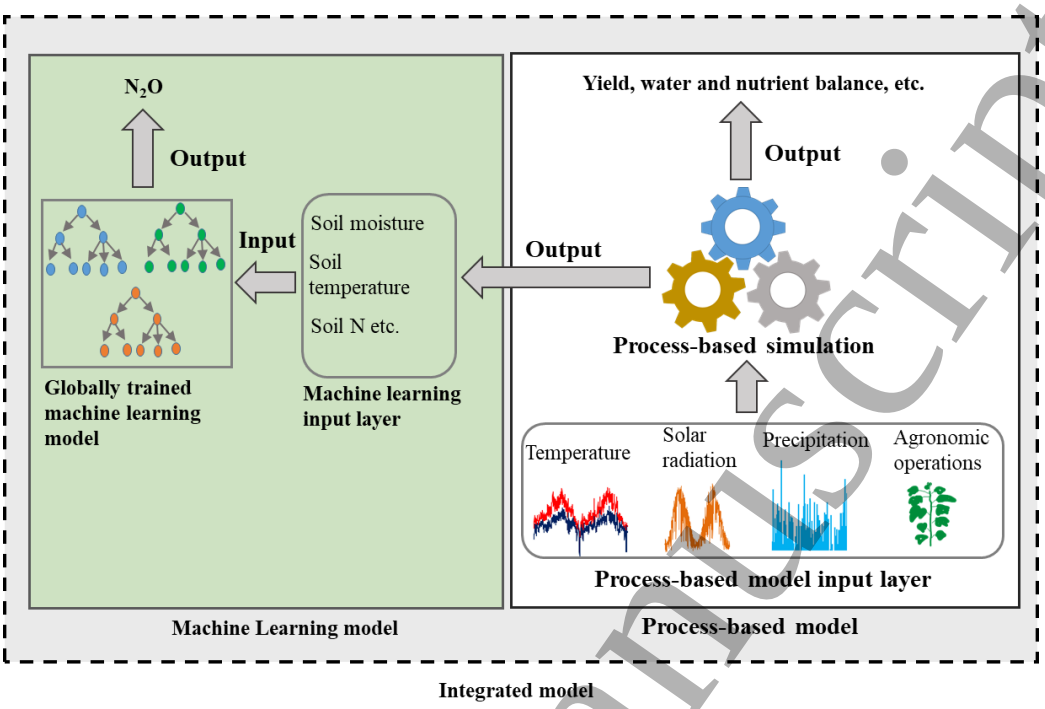

FIGURE 5 A conceptual strategy of integrated process-based and machine learning modeling for N_2O prediction. Daily time-step outputs on soil biophysical and biogeochemical variables predicted by the process-based models with a satisfactory level of confidence feed into a Random Forest model, trained on diverse soil, climate, and management scenarios, to predict N_2O fluxes.

Second, ML models can be used to analyze N_2O prediction bias by process-based models. That is, an ML model can automatically learn the pattern of prediction deviation from observed fluxes and identify the variables associated with the greatest contribution to the mismatch. This would lead to a better representation of the biogeochemical processes affecting the predictor variable through improved parameterization learned from the real-world variability of N_2O, and thereby help to correct model bias.

5 CONCLUSIONS

Results show that regression-based ML models such as Random Forest can, with limited input data, substantially improve temporal N_2O flux predictions from intensively managed cropping systems. Furthermore, ML facilitates interpreting non-linear interactions among N_2O predictor variables, which is often challenging with contemporary statistical methods. While our ML model predicted up to 51% variability of daily N_2O fluxes from corn at two upper Midwest sites, its application to other regions and crops requires further improvements in model training based on diverse data sources from various soils, climate, crop, and management conditions. Results also identify potential opportunities for integrating ML with process-based models to improve the overall predictability of soil N_2O emissions – a new paradigm, perhaps, for N_2O modeling.

6 DATA AVAILABILITY

Data and R code are available online at datadryad.org (https://doi.org/10.5061/dryad.bnzs7h493). The supplementary data provides supporting information.

7 ACKNOWLEDGMENTS

We thank K. Kahmark and S. Bohm for building and maintaining the automated chamber system and for data management, and H. Hsieh for helpful discussions during model development. We also thank D. Rowlings and P. Grace for help with the design and implementation of an earlier autochamber system. Financial support was provided by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Award DE-SC0018409), by the National Science Foundation Long-term Ecological

Research Program (DEB 1832042) at the Kellogg Biological Station, by the USDA Long-term Agroecosystem Research Program, and by Michigan State University AgBioResearch.

REFERENCES

Almaraz, M., Wong, M. Y., & Yang, W. H. (2020). Looking back to look ahead: a vision for soil denitrification research. *Ecology*, 101(1), e02917. doi:10.1002/ecy.2917.

Barton, L., Wolf, B., Rowlings, D. *et al.* (2015). Sampling frequency affects estimates of annual nitrous oxide fluxes. *Scientific Reports*, 5, 15912. doi:10.1038/srep15912.

Basso, B., & Ritchie, J. T. (2015). Simulating crop growth and biogeochemical fluxes in response to land management using the SALUS model. In S. K. Hamilton, J. E. Doll, & G. P. Robertson (Eds.), The Ecology of Agricultural Landscapes: Long-Term Research on the Path to Sustainability (pp. 252-274). New York, New York, USA: Oxford University Press.

Berardi, D., Brzostek, E., Blanc-Betes, E. *et al.* (2020). 21st-century biogeochemical modeling: Challenges for Century-based models and where do we go from here? *Global Change Biology Bioenergy*, 12(10), 774-788. doi.org/10.1111/gcbb.12730.

Borer, E. T., Harpole, W. S., Adler, P. B. *et al.* (2014). Finding generality in ecology: a model for globally distributed experiments. *Methods in ecology and evolution*, 5, 65-73. doi:10.1111/2041-210X.12125.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.

Brenowitz, N. D., & Bretherton, C. S. (2018). Prognostic validation of a neural network unified physics parameterization. *Geophysical Research Letters*, 45(12), 6289-6298. doi:10.1029/2018GL078510.

Brilli, L., Bechini, L., Bindi, M. *et al.* (2017). Review and analysis of strengths and weaknesses of agro-ecosystem models for simulating C and N fluxes. *Science of The Total Environment*, 598, 445-470. doi.org/10.1016/j.scitotenv.2017.03.208.

Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M. *et al.* (2013). Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Philosophical Transactions of the Royal Society B,* 368(1621), 20130122. doi:10.1098/rstb.2013.0122.

Chawla, N. V., Bowyer, K. W., Hall, L. O. *et al.* (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research,* 16(1), 321–357.

Conen, F., Dobbie, K. E., & Smith, K. A. (2000). Predicting N_2O emissions from agricultural land through related soil parameters. *Global Change Biology,* 6, 417-426.

Crutzen, P. J., Mosier, A. R., Smith, K. A. *et al.* (2008). N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. *Atmospheric Chemistry and Physics*, 8(2), 389-395. doi:10.5194/acp-8-389-2008.

de Bezenac, E., Pajot, A., & Gallinari, P. (2017). Deep learning for physical processes: incorporating prior scientific knowledge. arxiv:1711.07970. Retrieved from https://arxiv.org/abs/1711.07970

De Klein, C., Novoa, R. S. A., Ogle, S. *et al.* (2006). N_2O emissions from managed soils, and CO_2 emissions from lime and urea application. In H. S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.), 2006 IPCC guidelines for national greenhouse gas inventories. Volume 4. Agriculture, forestry and other land use (pp. 11.11-11.54). Kanagawa, Japan: Institute for Global Environmental Strategies (IGES).

Del Grosso, S. J., Wirth, T., Ogle, S. M. et al. (2008a). Estimating agricultural nitrous oxide emissions. *EOS,* 89, 529-530.

Del Grosso, S. J., Halvorson, A. D., & Parton, W. J. (2008b). Testing DAYCENT model simulations of corn yields and nitrous oxide emissions in irrigated tillage systems in Colorado. *Journal of Environmental Quality,* 37(4), 1383-1389. doi.org/10.2134/jeq2007.0292

Del Grosso, S. J., Parton, W. J., Mosier, A. R. *et al.* (2000). General model for N_2O and N_2 gas emissions from soils due to denitrification. *Global Biogeochemical Cycles*, 14, 1045-1060.

Dobbie, K. E., & Smith, K. A. (2003). Nitrous oxide emission factors for agricultural soils in Great Britain: the impact of soil water-filled pore space and other controlling variables. *Global Change Biology,* 9, 204-218.

Duncan, D. S., Oates, L. G., Gelfand, I. *et al.* (2019). Environmental factors function as constraints on soil nitrous oxide fluxes in bioenergy feedstock cropping systems. *Global Change Biology Bioenergy,* 11(2), 416-426. doi:10.1111/gcbb.12572

Ehrhardt, F., Soussana, J.-F., Bellocchi, G. *et al.* (2018). Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N_2O emissions. *Global Change Biology*, 24(2), e603-e616. doi:10.1111/gcb.13965

ERS (Economic Research Service). (2019a). Fertilizer use and price. Retrieved from https://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx

ERS (Economic Research Service). (2019b). Quickstats: Corn - acres planted. Retrieved from https://quickstats.nass.usda.gov

Firestone, M. K., & Davidson, E. A. (1989). Microbiological basis of NO and N_2O production and consumption in soil. In M. O. Andreae & D. S. Schimel (Eds.), Exchange of Trace

Gases between Terrestrial Ecosystems and the Atmosphere (pp. 7-21). Chichester U.K.: John Wiley and Sons.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232.

Fuchs, K., Merbold, L., Buchmann, N. *et al.* (2020). Multimodel evaluation of nitrous oxide emissions from an intensively managed grassland. *Journal of Geophysical Research Biogeosciences*, 125(1), e2019JG005261. doi.org/10.1029/2019JG005261

Gaillard, R. K., Jones, C. D., Ingraham, P. *et al.* (2018). Underestimation of N_2O emissions in a comparison of the DayCent, DNDC, and EPIC models. *Ecological Applications*, 28(3), 694-708. doi:10.1002/eap.1674

Gelfand, I., Hamilton, S. K., Kravchenko, A. N. *et al.* (2020). Empirical evidence for the potential climate benefits of decarbonizing light vehicle transport in the U.S. with bioenergy from purpose-grown biomass with and without BECCS. *Environmental Science & Technology*, 54, 2961-2974.

Gelfand, I., Shcherbak, I., Millar, N. *et al.* (2016). Long-term nitrous oxide fluxes in annual and perennial agricultural and unmanaged ecosystems in the upper Midwest USA. *Global Change Biology*, 22, 3594-3607. doi:10.1111/gcb.13426

Goldstein, A., Kapelner, A., Bleich, J. *et al.* (2015). Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, 24(1), 44-65. doi:10.1080/10618600.2014.907095

Gilhespy S. L., Anthony, S., Cardenas, L. *et al.* (2014). First 20 years of DNDC (DeNitrification DeComposition): Model evolution. *Ecological Modelling,* 292(24), 51-62. doi.org/10.1016/j.ecolmodel.2014.09.004

Grace, P., Robertson, G. P., Millar, N. *et al.* (2011). The contribution of maize cropping in the Midwest USA to global warming: a regional estimate. *Agricultural Systems,* 104(3), 292-296. doi:10.1016/j.agsy.2010.09.001

Grace, P., Ven der Weerden, T., Rowlings, D. *et al.* (2020). Considerations for automated flux measurement in the Global Research Alliance N_2O chamber methodology guidelines. *Journal of Environmental Quality,* 49(5), 1126-1140. doi.org/10.1002/jeq2.20124

Griffis, T. J., Lee, X., Baker, J. M. *et al.* (2013). Reconciling the differences between top-down and bottom-up estimates of nitrous oxide emissions for the U.S. Corn Belt. *Global Biogeochemical Cycles,* 27(3), 746-754. doi:10.1002/gbc.20066

Groffman, P. M., Tiedje, J. M., Robertson, G. P. *et al.* (1988). Denitrification at different temporal and geographical scales: proximal and distal controls. In J. R. Wilson (Ed.), Advances in nitrogen cycling in agricultural ecosystems (pp. 174-192). Wallingford, U.K.: CAB International.

Hamilton, S. K., Basso, B., Robertson, G. P. *et al.* (2020). Leaching losses of dissolved organic carbon and nitrogen from agricultural soils in the upper US Midwest. *Science of the Total Environment,* 734 (10), 139379. doi:10.1016/j.scitotenv.2020.139379

Hamilton, S. K., Hussain, M. Z., Bhardwaj, A. K. *et al.* (2015). Comparative water use by maize, perennial crops, restored prairie, and poplar trees in the US Midwest. *Environmental Research Letters,* 10, 064015. doi:10.1088/1748-9326/10/6/064015

Hoffman, A. L., Kemanian, A. R., & Forest, C. E. (2018). Analysis of climate signals in the crop yield record of sub-Saharan Africa. *Global Change Biology,* 24(1), 143-157. doi.org/10.1111/gcb.13901

Hussain, M. Z., Bhardwaj, A. K., Basso, B. *et al.* (2019). Nitrate leaching from continuous corn, perennial grasses, and poplar in the US Midwest. *Journal of Environmental Quality,* 48(6), 1849-1855. doi:10.2134/jeq2019.04.0156

IPCC (Intergovernmental Panel on Climate Change). (2013). Climate change 2013: the physical science basis. Retrieved from http://www.ipcc.ch/report/ar5/wg1/:

IPCC (Intergovernmental Panel on Climate Change). (2014). Climate change 2014: mitigation of climate change. Retrieved from https://www.ipcc.ch/report/ar5/wg3/:

Izaurralde, R. C., McGill, W. B., & Williams, J. R. (2012). Development and application of the EPIC Model for carbon cycle, greenhouse gas mitigation, and biofuel studies. In M. A. Liebig, A. J. Franzluebbers, & R. F. Follett (Eds.), Managing Agricultural Greenhouse Gases (Vol. Chapter 17, pp. 293-308). Waltham, MA: Academic Press.

Jarecki, M. K., Parkin, T. B., Chan, A. S. K. *et al.* (2008). Comparison of DAYCENT-simulated and measured nitrous oxide emissions from a corn field. *Journal of Environmental Quality,* 37(5), 1685-1690. doi:10.2134/jeq2007.0614

Jin, V. L., Schmer, M. R., Stewart, C. E. *et al.* (2017). Long-term no-till and stover retention each decrease the global warming potential of irrigated continuous corn. *Global Change Biology,* 23(7), 2848-2862. doi:10.1111/gcb.13637

Karpatne, A., Atluri, G., Faghmous, J. H. *et al.* (2017a). Theory-guided data science: a new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering,* 29(10), 2318-2331. doi:10.1109/TKDE.2017.2720168

Karpatne, A., Watkins, W., Read, J. *et al.* (2017b). Physics-guided neural netwoks (PGNN): an application in lake temperature modeling. arxiv:1710.11431. Retrieved from https://arxiv.org/abs/1710.11431

Kuz'min, V. E., Polishchuk, P. G., Artemenko, A. G. *et al.* (2011). Interpretation of qsar models based on random forest methods. *Molecular Informatics,* 30(6-7), 593–603.

Larsen, J., Damassa, T., & Levinson, r. (2007). Charting the Midwest: an inventory and analysis of greenhouse gas emissions in America's heartland. Retrieved from http://www.wri.org/publication/charting-the-midwest:

Li, C. (2000). Modeling trace gas emissions from agricultural ecosystems. *Nutrient Cycling in Agroecosystems,* 58, 259-276.

Li, C. (2007). Quantifying greenhouse gas emissions from soils: scientific basis and modeling approach. *Soil Science and Plant Nutrition,* 53, 344-352.

Linn, D. M., & Doran, J. W. (1984). Effect of water-filled pore space on CO_2 and N_2O production in tilled and non-tilled soils. *Soil Science Society of America Journal,* 48, 1267-1272.

Mikkelsen, R. L., & Snyder, C. S. (2012). Fertilizer nitrogen management to reduce nitrous oxide emissions in the U.S. In L. Guo, A. Gunasekara, & L. McConnell (Eds.), Understanding Greenhouse Gas Emissions from Agricultural Management. Washington, D.C.: American Chemical Society.

Millar, N., Robertson, G. P., Grace, P. R. *et al.* (2010). Nitrogen fertilizer management for nitrous oxide (N_2O) mitigation in intensive corn (Maize) production: an emissions reduction protocol for US Midwest agriculture. *Mitigation and Adaptation Strategies for Global Change,* 15(2), 185-204. doi:10.1007/s11027-010-9212-7

Oates, L. G., Duncan, D. S., Gelfand, I. *et al.* (2016). Nitrous oxide emissions during establishment of eight alternative cellulosic bioenergy cropping systems in the North Central United States. *Global Change Biology Bioenergy,* 8, 539-549. doi:10.1111/gcbb.12268

Parkin, T. B. (2008). Effect of sampling frequency on estimates of cumulative nitrous oxide emissions. *Journal of Environmental Quality,* 37(4), 1390-1395. doi:10.2134/jeq2007.0333

Parkin, T. B., & Kaspar, T. C. (2006). Nitrous oxide emissions from corn-soybean systems in the Midwest. *Journal of Environmental Quality,* 35(4), 1496-1506. doi:10.2134/jeq2005.0183

Parton, W. J., Holland, E. A., Del Grosso, S. J. *et al.* (2001). Generalized model for NO_x and N_2O emissions from soils. *Journal of Geophysical Research-Atmospheres,* 106, 17403-17419. doi:10.1029/2001JD900101

Perlman, J., Hijmans, R. J., & Horwath, W. R. (2014). A metamodelling approach to estimate global N_2O emissions from agricultural soils. *Global Ecology and Biogeography,* 23(8), 912-924. doi:10.1111/geb.12166

Philibert, A., Loyce, C., & Makowski, D. (2012). Quantifying uncertainties in N_2O emission due to N fertilizer application in cultivated areas. *PLoS ONE,* 7(11), e50950. doi:10.1371/journal.pone.0050950

Philibert, A., Loyce, C., & Makowski, D. (2013). Prediction of N_2O emission from local information with Random Forest. *Environmental Pollution,* 177, 156-163. doi:10.1016/j.envpol.2013.02.019

R Core Team. (2018). R: a language and environment for statistical computing (Version 3.5.2). https://www.r-project.org/: R Foundation for Statistical Computing.

Rafique, R., Fienen, M. N., Parkin, T. B. *et al.* (2013). Nitrous oxide emissions from cropland: a procedure for calibrating the DayCent biogeochemical model using inverse modelling. *Water Air and Soil Pollution,* 224(9), 1677. doi:10.1007/s11270-013-1677-z

Reichstein, M., Camps-Valls, G., Stevens, B. *et al.* (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195-204. doi:10.1038/s41586-019-0912-1

Robertson, G. P. (2004). Abatement of nitrous oxide, methane, and the other non-CO2 greenhouse gases: the need for a systems approach. In C. B. Field & M. R. Raupach (Eds.), The global carbon cycle (pp. 493-506). Washington, D.C., USA: Island Press.

Robertson, G. P. (2014). Soil greenhouse gas emissions and their mitigation. In N. Van Alfen (Ed.), Encyclopedia of agriculture and food systems (Vol. 5, pp. 185-196). San Diego, California, USA: Elsevier.

Robertson, G. P., & Hamilton, S. K. (2015). Long-term ecological research in agricultural landscapes at the Kellogg Biological Station LTER site: conceptual and experimental framework. In S. K. Hamilton, J. E. Doll, & G. P. Robertson (Eds.), The Ecology of Agricultural Landscapes: Long-Term Research on the Path to Sustainability (pp. 1-32). New York, New York, USA: Oxford University Press.

Roelandt, C., Wesemael, B. V., & Rounsevell, M. (2005). Estimating annual N_2O emissions from agricultural soils in temperate climates. *Global Change Biology,* 11(10), 1701-1711. doi.org/10.1111/j.1365-2486.2005.01025.x

Ruan, L., & Robertson, G. P. (2017). Reduced snow cover increases wintertime nitrous oxide (N_2O) emissions from an agricultural soil in the upper U.S. Midwest. *Ecosystems*, 20, 917-927. doi:10.1007/s10021-016-0077-9

Saha, D., Kemanian, A. R., Rau, B. M. *et al.* (2017). Designing efficient nitrous oxide sampling strategies in agroecosystems using simulation models. *Atmospheric Environment*, 155, 189-198. doi:10.1016/j.atmosenv.2017.01.052

Saha, D., Rau, B. M., Kaye, J. P. *et al.* (2017). Landscape control of nitrous oxide emissions during the transition from conservation reserve program to perennial grasses for bioenergy. *Global Change Biology Bioenergy*, 9(4), 783-795. doi:10.1111/gcbb.12395

Sanford, G. R., Oates, L. G., Jasrotia, P. *et al.* (2016). Comparative productivity of alternative cellulosic bioenergy cropping systems in the North Central USA. *Agriculture, Ecosystems & Environment*, 216, 344-355. doi:10.1016/j.agee.2015.10.018

Schlesinger, W. H. (2013). An estimate of the global sink for nitrous oxide in soils. *Global Change Biology*, 19, 2929-2931. doi:10.1111/gcb.12239

Shcherbak, I., Millar, N., & Robertson, G. P. (2014). Global metaanalysis of the nonlinear response of soil nitrous oxide (N_2O) emissions to fertilizer nitrogen. *Proceedings of the National Academy of Sciences USA*, 111(25), 9199-9204. doi:10.1073/pnas.1322434111

Sozanska, M., Skiba, U., & Metcalfe, S. (2002). Developing an inventory of N_2O emissions from British soils. *Atmospheric Environment*, 36(6), 987-998. doi.org/10.1016/S1352-2310(01)00441-1.

Syakila, A., & Kroeze, C. (2011) The global nitrous oxide budget revisited. *Greenhouse Gas Measurement and Management*, 1(1), 17-26. DOI: 10.3763/ghgmm.2010.0007

Thompson, R. L., Lassaletta, L., Patra, P. K. *et al.* (2019). Acceleration of global N2O emissions seen from two decades of atmospheric inversion. *Nature Climate Change*, 9(12), 993-998. doi:10.1038/s41558-019-0613-7

Tian, H., Yang, J., Xu, R. *et al.* (2019). Global soil nitrous oxide emissions since the preindustrial era estimated by an ensemble of terrestrial biosphere models: magnitude, attribution, and uncertainty. *Global Change Biology*, 25(2), 640-659. doi:10.1111/gcb.14514

Tian, H., Xu, R., Canadell, J. G. *et al.* (2020). A comprehensive quantification of global nitrous oxide sources and sinks. *Nature*, 586(7828), 248-256.

Torgo, L., Ribeiro, R. P., Pfahhringer, B. *et al.* (2013). SMOTE for regression. In L. Correia, L. P. Reis, & J. Csascalho (Eds.), Progress in Artificial Intelligence. EPIA 2013. *Lecture Notes in Computer Science*, vol 8154 (pp. 378-389). Berlin, Heidelberg: Springer.

Villa-Vialaneix, N., Follador, M., Ratto, M. *et al.* (2012). A comparison of eight metamodeling techniques for the simulation of N_2O fluxes and N leaching from corn crops. *Environmental Modelling & Software*, 34, 51-66.

Welling, S. H., Refsgaard, H. H. F., Brockhoff, P. B. *et al.* (2016). Forest floor visualizations of random forests. arxiv:1605.09196. Retrieved from https://arxiv.org/abs/1605.09196