Received: 30 May 2024 | Accepted: 12 September 2024

DOI: 10.1002/jeq2.20637

Journal of Environmental Quality

TECHNICAL REPORT

Special Section: Outcomes of the Long-Term Agroecosystem Research Network

Machine learning reveals dynamic controls of soil nitrous oxide emissions from diverse long-term cropping systems

Jashanjeet Kaur Dhaliwal[1] | **Dinesh Panday**[1,2] | **G. Philip Robertson**[3,4] | **Debasish Saha**[1]

[1]Biosystems Engineering and Soil Science, University of Tennessee, Knoxville, Tennessee, USA

[2]Rodale Institute, Kutztown, Pennsylvania, USA

[3]W. K. Kellogg Biological Station, Michigan State University, Hickory Corners, Michigan, USA

[4]Department of Plant, Soil, and Microbial Sciences, Michigan State University, East Lansing, Michigan, USA

Correspondence
Debasish Saha, Biosystems Engineering and Soil Science, University of Tennessee, Knoxville, TN, USA.
Email: dsaha3@utk.edu

Assigned to Associate Editor Ann-Marie Fortuna.

Funding information
NSF Long-term Ecological Research Program, Grant/Award Number: DEB 2224712; National Institute of Food and Agriculture, Grant/Award Number: 2021-67019-34247; USDA Long-Term Agroecosystem Research Program

Abstract

Soil nitrous oxide (N_2O) emissions exhibit high variability in intensively managed cropping systems, which challenges our ability to understand their complex interactions with controlling factors. We leveraged 17 years (2003–2019) of measurements at the Kellogg Biological Station Long-Term Ecological Research (LTER)/Long-Term Agroecosystem Research (LTAR) site to better understand the controls of N_2O emissions in four corn–soybean–winter wheat rotations employing conventional, no-till, reduced input, and biologically based/organic inputs. We used a random forest machine learning model to predict daily N_2O fluxes, trained separately for each system with 70% of observations, using variables such as crop species, daily air temperature, cumulative 2-day precipitation, water-filled pore space, and soil nitrate and ammonium concentrations. The model explained 29%–42% of daily N_2O flux variability in the test data, with greater predictability for the corn phase in each system. The long-term rotations showed different controlling factors and threshold conditions influencing N_2O emissions. In the conventional system, the model identified ammonium (>15 kg N ha^{-1}) and daily air temperature (>23°C) as the most influential variables; in the no-till system, climate variables such as precipitation and air temperature were important variables. In low-input and organic systems, where red clover (*Trifolium repens* L.; before corn) and cereal rye (*Secale cereale* L.; before soybean) cover crops were integrated, nitrate was the predominant predictor of N_2O emissions, followed by precipitation and air temperature. In low-input and biologically based systems, red clover residues increased soil nitrogen availability to influence N_2O emissions. Long-term data facilitated machine learning for predicting N_2O emissions in response to differential controls and threshold responses to management, environmental, and biogeochemical drivers.

Abbreviations: CT, conventionally tilled; C–S–W, corn–soybean–winter wheat; GWC, gravimetric water content; NT, no-till; OOB, out-of-bag; PD, partial dependence; RF, random forest; RMSE, root mean square error; UAN, urea ammonium nitrate; WFPS, water-filled pore space.

1 | INTRODUCTION

Global terrestrial emissions of nitrous oxide (N_2O), a potent greenhouse gas, have increased from 6.3 Tg N_2O-N year^{-1} in the preindustrial era to 10 Tg N_2O-N year^{-1} in the last decade (2007–2016), which represents an increase in atmospheric N_2O load of ~60% (Tian et al., 2019). This increase is largely governed by the increase in cropland soil emissions from 0.3 to 3.3 Tg N_2O-N year^{-1} during the same period. Increases in fertilizer nitrogen (N) application rates drive N_2O emissions from agricultural soils—the largest anthropogenic source of atmospheric N_2O that contributes nearly 10% of global anthropogenic radiative forcing (Butterbach-Bahl et al., 2013; Shcherbak et al., 2014; Syakila & Kroeze, 2011; Zhang et al., 2020).

Nitrous oxide emissions from agricultural soils result from biogeochemical processes primarily dominated by microbial nitrification and denitrification (Dobbie & Smith, 2003; Robertson & Groffman, 2024). These processes are highly sensitive to environmental conditions including soil moisture/water-filled pore space (WFPS), temperature, pH, and availability of organic carbon (C) and inorganic N (Baral et al., 2022; Giltrap et al., 2010; Oertel et al., 2016)—further influenced by management practices. Net N_2O emissions measured in situ are the outcomes of complex interactions among the driving factors and their threshold conditions that are often nonlinear and spatially discontinuous (Robertson, 2023). For this reason, the prediction of N_2O emissions based on simultaneously observed environmental factors and N substrate concentrations shows very weak to no correlations in most studies (e.g., Gelfand et al., 2016; Maharjan & Venterea, 2013; Wanyama et al., 2018). This complexity is often manifested in the highly dynamic and variable nature of soil N_2O emissions characterized as "hot spots" and "hot moments" (Groffman et al., 2009; Saha et al., 2018; Venterea et al., 2012). Only a fraction of this spatial and temporal variability may be attributed to applied fertilizer N, with much of the remainder attributed to the soil, climatic, and management factors influencing total N_2O emissions (de Klein et al., 2020; Deng et al., 2022; Groffman et al., 2009). Both direct observations and meta-analyses show a non-linear relationship between N fertilization rate and N_2O emissions (e.g., Hoben et al., 2011; Kim et al., 2013; Millar et al., 2018; Scheer et al., 2016; Shcherbak et al., 2014), which further highlights the confounding impacts of multiple driving factors on N_2O emissions.

Long-term management practices such as no-till (NT) and cover cropping alter soil biophysical and biogeochemical conditions to modify the shape and threshold response of N_2O emissions to environmental and biogeochemical drivers. For example, long-term NT improves soil physical properties such as soil aeration and moisture retention and reduces soil temperature (Nouri et al., 2019). While lower soil temperature and improved aeration under macropore-dominated NT soils may reduce N_2O emissions (Ussiri et al., 2009; Van Kessel et al., 2013), moist soils can promote N_2O emissions from denitrification and its rapid escape to the atmosphere due to greater diffusivity (Wang & Zou, 2020). Therefore, N_2O response to soil moisture and temperature may differ between NT and conventionally tilled (CT) soils, which makes it difficult to predict management impacts on N_2O emissions under inter-annual weather variability. Literature inconsistently shows that N_2O emissions under NT either decrease (Grandy et al., 2006; Six et al., 2004; Van Kessel et al., 2013) or increase (Ball et al., 2008; Mei et al., 2018) compared to CT. Similarly, the quantity and quality of cover crop biomass influence soil N availability, which in turn affects N_2O emissions and their response to fertilizer N application (Finney et al., 2015; Panday et al., 2022). Following cover crop termination, legume residues can significantly increase N_2O emissions due to fast N release (Davis et al., 2019; Saha, Kaye, et al., 2021). Accelerated cover crop decomposition and heterotrophic respiration rapidly consume soil oxygen (O_2) to promote anoxia and large N_2O emissions from denitrification regardless of soil moisture conditions (Lussich et al., 2024). These findings are in contrast with the generalized conclusions about negligible impacts of cover crops on soil N_2O emissions (Basche et al., 2014; Kaye & Quemada, 2017). Such inconsistencies further highlight our limited understanding of dynamic variable controls on N_2O emissions in response to environmental and management differences.

Additionally, long-term N_2O emissions are greatly influenced by inter-annual variability in weather conditions (Baral et al., 2022; Burchill et al., 2014), and in particular variability in rainfall distributions (Rowlings et al., 2015). However,

Core Ideas

- Long-term (2003–2019) data from an LTAR/LTER site was used to understand dynamic controls of N_2O emissions.
- The RF model showed different drivers and threshold conditions influencing N_2O emissions in long-term rotations.
- Soil ammonium and daily temperature were identified as the most influential variables in the Conventional system.
- Climate variables—precipitation and daily temperature—were important variables in the No-till system.
- In low-input and organic systems, nitrate from legume cover crops strongly influenced N_2O emissions.

many studies investigating spatial and temporal N_2O emission controls are based on short-term measurements spanning from one to two growing seasons and capturing only a snapshot within a growing season (Dorich et al., 2020). The time-consuming and resource-intensive nature of chamber-based N_2O flux measurements is a key limitation for individual research projects in implementing spatially and temporally extensive flux monitoring. Coordinated efforts by long-term research network sites can be useful in overcoming such limitations by providing multi-year data capturing weather variability and management legacies (e.g., crop rotation, NT, and cover cropping), which often take time to emerge (Cusser et al., 2020; Six et al., 2004).

Several quantitative tools have been traditionally employed to understand the complexity of soil, climate, and crop management practices influencing N_2O emissions with varying degrees of success. For example, parametric regression models, by design, do not represent nonlinear variable interactions influencing N_2O emissions (Kim et al., 2013). Similarly, the emission factor (EF) approach by the Intergovernmental Panel on Climate Change (IPCC) is insensitive to dynamic interactions between the environmental factors and management conditions. These models often fail to predict how N_2O emissions may change at a finer temporal and spatial scale (Ramírez-Melgarejo et al., 2020; Richards et al., 2016). Furthermore, these approaches are limited to provide insights on critical values of predictor variables differentially influencing N_2O under different management practices, such as crop diversification, NT, cover crop, and so forth. Unlike the empirical models, process-based biogeochemical models can simulate feedback and interactions that can be difficult to distinguish in the field (Giltrap et al., 2010). Process-based models, such as the DayCent (Del Grosso et al., 2000; Parton et al., 2001) and DeNitrification-DeComposition (Li, 2000, 2007), have considered important regulating factors to support the prediction of N_2O emissions and thus have been recognized as useful tools to evaluate the effects of management practices on N_2O emissions from agricultural soils (Deng et al., 2018; Jarecki et al., 2008). However, heavy parametrization and site-specific calibration need of the process-based models often limit their extensive use (Ehrhardt et al., 2018; Fuchs et al., 2020; Gaillard et al., 2018; Gilhespy et al., 2014).

Machine learning models such as decision trees and random forest (RF) treat the output variable (e.g., N_2O) as an implicit function of input features (e.g., soil, environment, and management), and can capture complex nonlinear relationships as learned from the data and not by predefined process-based relationships as in the case of biogeochemical models (Breiman, 2001; Huang et al., 2010). The RF appears to be a more promising technique than classical regression-based methods and other machine learning algorithms due to its ability to rank predictors using internal measures of variable importance and to provide valuable insights through partial dependence (PD) plots (Djiemon et al., 2019; Saha et al., 2017).

Machine learning has been increasingly used to predict N_2O emissions in recent years (Glenn et al., 2021; Joshi et al., 2024; Liao et al., 2023; Philibert et al., 2013; Saha, Basso, & Robertson, 2021; Yin et al., 2022). However, predictive ability of machine learning models is correlational as learned from the data, hence has limited power in representing a novel scenario, which is a key pursuit of process-based models. Nonetheless, machine learning models can be resource efficient in scaling our existing knowledge and provide insights on key variable interactions controlling N_2O emissions to optimize process-based models. The availability of long-term data creates novel opportunities for using machine learning models to understand differential controls of N_2O emissions under diverse management practices.

By leveraging 17 years of long-term observations, we used RF and decision tree models to infer the controls and drivers of soil N_2O emissions from four corn (*Zea mays* L.)–soybean (*Glycine max* L.)–wheat (*Triticum aestivum* L.) rotations employing diverse tillage, fertilization, and cover cropping practices in the US upper Midwest. Our objectives are to identify critical management, environmental, and biogeochemical drivers of N_2O emissions and their differential relationships and threshold conditions for emissions under diverse long-term cropping rotations.

2 | MATERIALS AND METHODS

2.1 | Site description

Our study tracks a 17-year (2003–2019) long-term data stream from the Main Cropping System Experiment (MCSE) of the Kellogg Biological Station (KBS) Long-Term Ecological Research (LTER) site, which is also one of 18 Long-Term Agroecosystem Research sites across the United States. Historical data on yearly N_2O emissions, soil properties, agricultural management practices, and weather were obtained from the KBS LTER data catalog (https://lter.kbs.msu.edu/datatables). The KBS LTER site is located in the northeast portion of the US corn belt in southwest Michigan (42°24′ N, 85°24′ W, and 288 masl) and was originally established in 1987 to examine the ecology of intensively managed field crops and the landscape in which they reside (Robertson & Hamilton, 2015). Soils at the site are well-drained Typic Hapludalfs of the Kalamazoo (fine-loamy, mixed, mesic) and Oshtemo (coarse-loamy, mixed, mesic) series, formed from glacial till and outwash with some intermixed loess (Crum & Collins, 1995; Luehmann et al., 2016). Surface soils exhibit an average of 43% sand and 17% clay contents (Robertson & Hamilton, 2015), with an average organic C of 11.9 g kg^{-1},

TABLE 1 Agronomic management practices under the four annual cropping systems studied (2003–2019).

Cropping system	Crop phase (N fertilizer, kg N ha^{-1})				
Conventional	Corn (137 ± 20)		Soybean	Wheat (77 ± 17)	
No-till	Corn (137 ± 20)		Soybean	Wheat (77 ± 17)	
Reduced input	Corn (30 ± 3)	Cereal rye	Soybean	Wheat (40 ± 13)	Red clover
Biologically based/organic	Corn	Cereal rye	Soybean	Wheat	Red clover

total N of 1.2 g kg^{-1}, and pH of 6.5. The local weather is humid continental with hot and wet summers. Annual air temperature (30-year average) at KBS is 9.9°C, and precipitation averages 1027 mm year^{-1} evenly distributed seasonally with a snowfall of about 1.4 m and an average snow depth of 148 mm for days when snow is present (Robertson & Hamilton, 2015). Details about weather conditions during the crop-growing season period are given in Figure S1.

2.2 | Cropping systems and management

The MCSE includes seven treatments arranged in a randomized complete block design with six replications; we used four annual crop treatments (Table 1), including: (1) a conventional system with chisel tillage and standard chemical inputs, (2) an NT system with standard chemical inputs, (3) a reduced input system with chisel tillage, low fertilizer inputs, and cover crops, and (4) a biologically based/organic system managed organically using chisel tillage, cover crops, and no synthetic chemical inputs. The NT system was identical to the conventional system except for the lack of tillage. Chisel tillage in conventional, reduced input, and biologically based/organic systems was conducted to a depth of 15–18 cm, followed by secondary tillage operations such as disking. Since 1993, all of the systems have been maintained as corn–soybean–winter wheat (C–S–W) rotations according to best management practices (Robertson, 2015; Robertson & Hamilton, 2015). Corn and soybean were planted in late April or May and winter wheat was planted in late September or early October. The conventional and NT systems received recommended rates of N fertilizer at 137 ± 20 kg N ha^{-1} year^{-1} during the corn phase and 77 ± 17 kg N ha^{-1} year^{-1} during the wheat phase of each rotation (Gelfand et al., 2016). Corn was managed with split fertilizer application ∼30 kg N ha^{-1} at planting, with the remainder side-dressed at the V6 stage around June 28 (±9 days), while wheat was fertilized on April 19 (±7 days). The reduced input system received N fertilizer at an average rate of 30 ± 3 kg N ha^{-1} year^{-1} on May 13 (±7 days) during the corn phase and 40 ± 13 kg N ha^{-1} year^{-1} in April during the wheat phase of the rotation. No N fertilizer was applied to the soybean phase, but it received minor N fertilizer inputs as part of phosphorus (P), potassium (K), and herbicide applications, which were applied as needed in some years according to the Michigan State University recommendations (Warncke et al., 2009). The biologically based/organic system is a USDA-certified organic treatment. No manure, compost, or insecticide was applied in any of the cropping systems. Nitrogen fertilizer was added as urea ammonium nitrate (UAN) injected at 10-cm depth between crop rows at planting and as side-dressing. Soybean and corn were harvested in October and November, respectively, and winter wheat was harvested in July. In the low-input and biologically based/organic systems, the winter cover crop cereal rye (*Secale cereale* L.) was planted in October following corn and before soybean, and red clover (*Trifolium pratense* L.) was frost-seeded into winter wheat in March and terminated just before planting corn the following spring.

Plots in conventional, reduced input, and biologically based/organic were chisel plowed, while those in NT were not tilled. Cereal rye was sown following corn before soybean and red cover was frost-seeded into winter wheat in reduced input and biologically based/organic cropping systems.

2.3 | Data collection

Nitrous oxide flux measurements were made in four out of the six replicates in each treatment using static chambers at weekly to monthly intervals when the soils were not frozen. The manual open-bottom chambers, equipped with rubber septa and measuring 29 cm × 29 cm × 14 cm, were made from opaque polycarbonate sheeting and placed on semipermanent aluminum bases (28 cm × 28 cm × 10 cm), which were removed only during agronomic activities. Each chamber had a volume of approximately 12 L. Gas samples were collected using a headspace-flushed syringe every 15 min over a 1-h chamber closure period. The samples were stored in 5.9-mL Exetainer vials (Labco Limited) and analyzed using gas chromatography. The gas flux was calculated using the equation:

$$F = \frac{a * M * P * V}{A * R * T}$$

where F is gas flux (g N cm^{-2} h^{-1}), a is the average rate of change of gas concentration (ppm h^{-1}), M is molecular weight of N in N_2O (28 µg N µmol N_2O^{-1}), P is assumed atmospheric pressure (1 atm), R is universal gas constant (0.0821 L-atm mol-K^{-1}), T is field temperature (˚K = ˚C + 273), V is volume of gas in chamber (cm^3), and A is soil surface area covered by chamber (cm^2).

More details on N_2O flux measurements can be found in Gelfand et al. (2016) and at https://lter.kbs.msu.edu/protocols/113. Long-term gas sampling frequency was designed to capture treatment differences. During each gas sampling event, soil cores from the 0- to 25-cm depth were extracted from near the chamber for determination of gravimetric water content (GWC). WFPS was determined using the GWC and a consistent bulk density of 1.44 g cm^{-3} across all treatments. Soil samples (0- to 25-cm depth) from five random locations in each plot were collected biweekly each year for inorganic N (NH_4^+ and NO_3^-) concentrations. Soils were extracted with 2 M KCl and analyzed for NH_4^+ and NO_3^- concentrations in a continuous flow analyzer (Alpkem 3550; O.I. Analytical). Soil NH_4^+, NO_3^-, and WFPS values were linearly interpolated between two sampling dates to match the N_2O sampling dates when soil samples were not collected on the gas sampling days due to management and weather reasons. Cumulative gas fluxes were calculated by linear interpolation between the successive sampling dates. The 17-year database includes five full W–C–S rotations under each treatment.

2.4 | Statistical methods

RF and regression tree analyses were performed using the packages "*randomForest*" and "*rpart*," respectively, in R statistical software v. 4.2.3 (R Core Team, 2023) to examine dynamic controls of N_2O emissions and threshold response to critical drivers across diverse cropping systems (Figure 1). RF is an ensemble learning algorithm that combines numerous decision trees (n_{tree}) and bagging (Breiman, 2001), wherein each tree is constructed using a bootstrap sample (called "in-bag") of dataset, and a random subset of total predictors (m_{try}) is considered for node splits. The final RF prediction is the mean fitted response from all tree predictions. In each tree, about one-third of data are left out and these are called out-of-bag (OOB) data, which are used to estimate percentage variation explained—a measure that indicates the goodness of OOB predictions explaining the target variance in the training dataset. More details on the RF algorithms can be found in Hoffman et al. (2018) and Saha, Basso, and Robinson (2021). The optimal number of predictor variables was determined using Pearson correlation analysis to remove the highly correlated variables ($r > 0.75$) to avoid overfitting (Figure S2). The final model includes daily N_2O fluxes (log transformed values) as the response variable and average daily air temperature (T_{avg}), cumulative precipitation in last 2 days ($\sum ppt_{2d}$), WFPS, NH_4^+, NO_3^-, and crop as predictor variables (Table 2).

Total 70% of the observations from the total observations (n = 3374) from each cropping system were randomly selected to construct the training set (n = 2362), and the remaining 30% were used for testing (n = 1012). The crop variable was used as a categorical variable with three levels in the data and encoded to dummy numbers to enhance the efficiency of the model algorithms. A 10-fold cross-validation scheme was applied to the training data to optimize the hyperparameters at $m_{try} = 2$ and $n_{tree} = 500$ using $seed = 123$ to get reproducible results. The model performance was evaluated by coefficient of determination (r^2), root mean square error (RMSE), and mean absolute error (MAE) between the observed and predicted daily N_2O fluxes.

We used two inbuilt RF functions, variable importance metric (package "*vip*") and PD (package "*pdp*"), and a decision tree (package "*rpart*") to understand cropping system-specific critical drivers, their interactions, and threshold conditions controlling N_2O emissions. The variable importance measures the increase in model error in the OOB data in response to random permutation of input variables (Breiman, 2001). Larger error before and after permutation means greater importance of the variable and its contribution to the model's predictive accuracy. Top predictors were visualized using one- and two-dimensional (2D) PD plots to identify the nature of the relationship between the predictor and response variables. Additionally, a single decision tree for each rotation was constructed using the *rpart* package in R to identify threshold conditions beyond which large changes in N_2O flux behavior occur.

The linear mixed model using the package "*lmerTest*" from R statistical software v. 4.2.3 (R Core Team, 2023) was employed to analyze cumulative N_2O emissions from five cycles of each rotation phase and the total cumulative emissions for the entire rotation (W–C–S) from 2004 to 2018. The N_2O data were tested for normality using the Shapiro–Wilk test and were transformed using Box-Cox transformations when needed. The treatments, crop phase, and their interactions were considered as fixed effects and blocks were treated as random effects. When main and interaction effects were significant at $\alpha = 0.05$, pairwise comparisons between treatments were performed with the estimated marginal mean function and a post hoc Tukey test using the package "*emmeans*."

3 | RESULTS

3.1 | N_2O emissions

Daily N_2O fluxes from these annual cropping systems exhibited wide variability, ranging from −0.7 to 55.4, −0.69 to 45.6, −0.17 to 118, and −0.31 to 144 g N ha^{-1} day^{-1} in conventional, NT, reduced input, and biologically based/organic

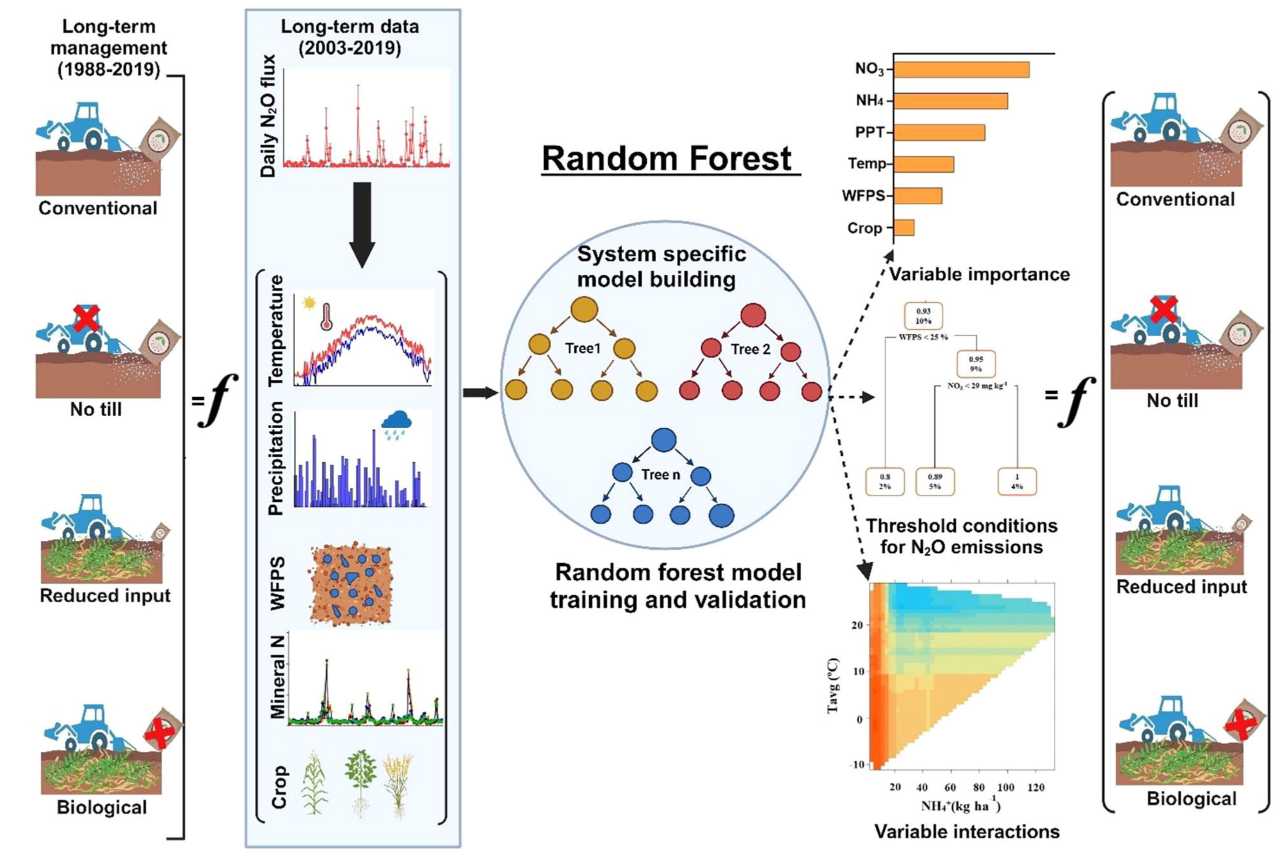

FIGURE 1 Schematic overview of random forest modeling to identify critical drivers of N_2O emissions, their differential relationships, and threshold conditions for emissions under diverse long-term cropping rotations.

TABLE 2 Predictor variables supplied to the random forest model.

Variable	Variable category	Description	Unit
T_{avg}	Climate	Growing season average temperature	˚C
$\sum ppt_{2d}$	Climate	Growing season cumulative precipitation in last 2 days	mm
WFPS	Soil	Water-filled pore space	%
NH_4^+	Soil	Soil ammonium	kg N ha^{-1}
NO_3^-	Soil	Soil nitrate	kg N ha^{-1}
Crop phase	Management	Growing main crop (corn, soybean, and wheat)	-

systems, respectively (Figure 2). Across five cycles of W–C–S rotations (2004–2018), total N_2O emissions from 3-year rotations were highest in biologically based (4.2 kg N ha^{-1}), followed by NT (3.5 kg N ha^{-1}), conventional (3.1 kg N ha^{-1}), and reduced input systems (2.9 kg N ha^{-1}), with no significant differences observed (Figure 3). In conventional system, soybean exhibited lower cumulative emissions compared to corn ($p < 0.05$), while in NT, reduced input, and biologically based systems, no discernible differences in emissions between corn and soybean were noted. Winter wheat emissions were significantly lower than emissions in corn in the reduced input system and lower than emissions in corn and soybean in the biologically based/organic system. Regardless of cropping system management, emissions from corn phase were significantly higher than emissions from soybean and wheat (1.58 vs. 1.03 vs. 0.83 kg N ha^{-1}, respectively).

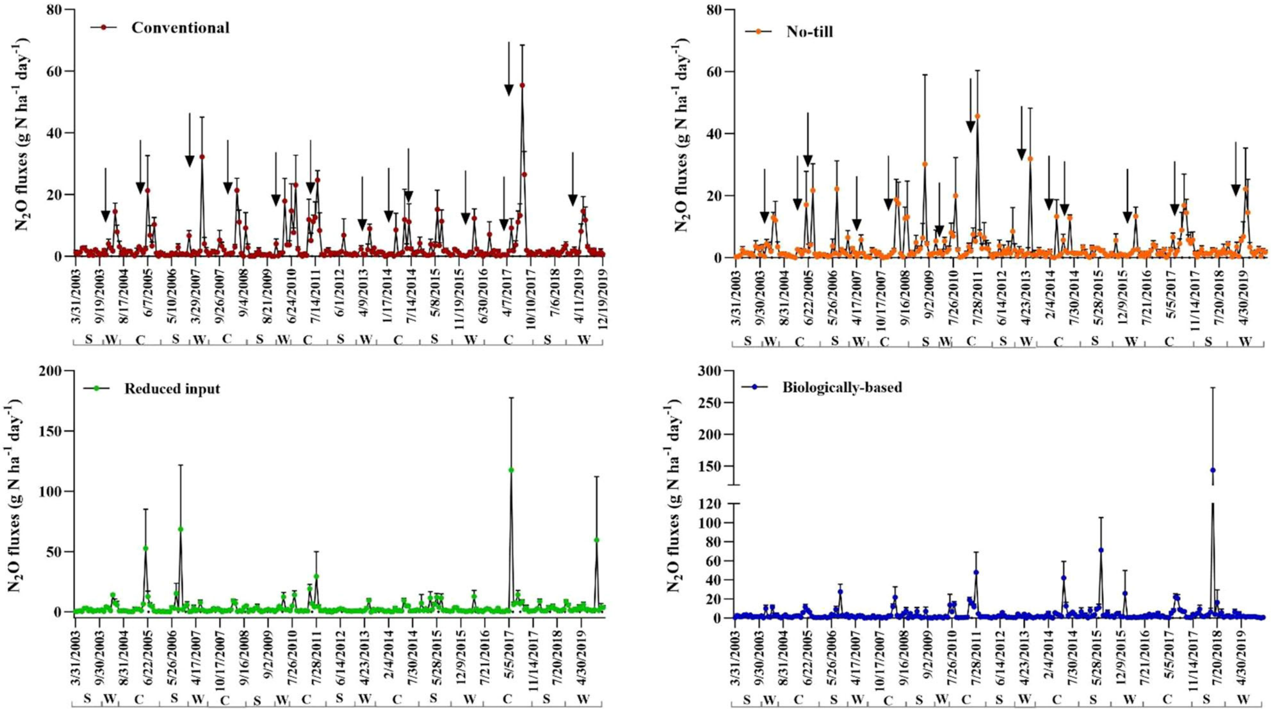

FIGURE 2 Daily N_2O fluxes in annual systems over the period of 2003–2019. C, S, and W represent corn, soybean, and winter wheat phases, respectively. Arrows indicate times of fertilizer application in conventional and no-till systems. Scales of the graph panels are different

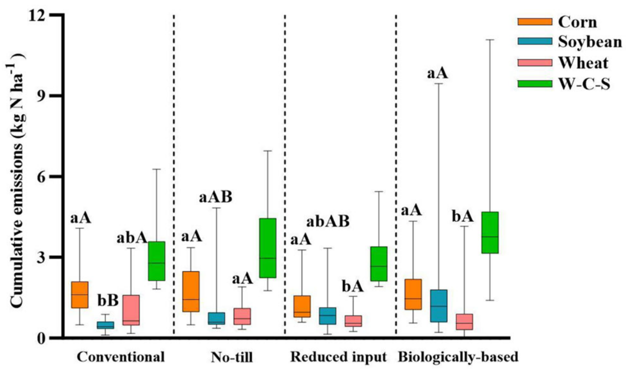

FIGURE 3 Average cumulative N_2O emissions of five cycles of each rotation phase within each annual system and the average cumulative emissions for the entire rotation (W–C–S) from 2004 to 2018. Significant differences ($p < 0.05$) in crop phases within the same annual system are represented by different lowercase letters. Uppercase letters represent differences ($p < 0.05$) in annual systems within same crop phase. C, S, and W represent corn, soybean, and winter wheat phases, respectively.

3.2 | Model performance

For each annual cropping system, the RF model was trained using T_{avg}, $\sum ppt_{2d}$, WFPS, NH_4^+, NO_3^-, and crop phase as model inputs (Table 2). For the entire C–S–W rotation, the model accounted for 29%–42% of the variability between observed and predicted N_2O fluxes, utilizing 30% of the observations for testing, which were not included in the model training (Figure S3). The highest proportion of variability in N_2O emissions was explained by the RF model in the conventional system (42%), followed by the biologically based/organic systems (40%), the reduced input system (37%), and the NT system (29%). The RF model underpredicted N_2O fluxes in the NT, reduced input, and biologically based systems on some occasions, with an RMSE of 0.17 in NT and biologically based systems, and 0.18 in reduced input system. The model predicted greater variability in corn phases than in soybean and wheat phases across all systems, with the highest variability predicted in biologically based (60%), followed by conventional (48%), reduced input (43%), and NT (32%) systems (Table 3). The least amount of variability was accounted for in the soybean phase in conventional (16.7%) and NT (1.46%), whereas in wheat, this was evident in reduced input (24%) and biologically based (6.1%) systems.

3.3 | Critical variables for N_2O emissions

Variable importance measures identified NH_4^+ and T_{avg} as the most influential variables in the conventional cropping system, with each variable accounting for more than 25% increase in mean square error of the OOB samples if randomly permuted (Table 3). This trend was also evident in the corn and wheat phases of the conventional system. As evident from PD plots, the model predicted high N_2O losses when NH_4^+

TABLE 3 Importance of variables controlling N_2O emissions as predicted by random forest for the entire rotation (C–S–W) of each annual system and different crop phases within each system

Variable	Conventional				No-till				Reduced input				Biologically based			
	C–S–W	C	S	W	C–S–W	C	S	W	C–S–W	C	S	W	C–S–W	C	S	W
	% increase in mean square error															
T_{avg}	28.3 (2)	23.2	14.2	11.6	21.6 (2)	16.1	6.4	10.7	22.1 (3)	19.4	18.0	8.0	22.3 (2)	29.4	14.6	15.0
$\sum ppt_{2d}$	14.5 (4)	13.4	8.9	5.2	25.4 (1)	15.1	8.1	5.3	25.6 (2)	12.8	20.0	9.7	21.6 (3)	18.7	20.8	7.0
WFPS	13.4 (6)	9.2	10.3	5.9	8.9 (6)	10.4	2.7	2.0	13.6 (5)	9.4	9.6	4.5	20.7 (4)	25.4	12.8	3.1
NO_3^-	13.0 (5)	13.3	11.2	−1.9	9.9 (5)	14.3	3.7	3.6	26.2 (1)	28.9	21.5	11.1	24.1 (1)	33.3	14.5	6.5
NH_4^+	31.7 (1)	29.4	3.7	17.5	14.3 (3)	13.9	1.6	17.7	20.2 (4)	9.5	8.1	18.9	12.6 (6)	13.8	6.2	5.3
Crop	17.1 (3)				12.8 (4)				10.9 (6)				14.0 (5)			
	% variability explained by random forest															
	40.6	48.1	16.7	30.3	26.0	32	1.46	15.1	36.2	43.0	31.5	24.0	39.1	59.8	30.1	6.1

Note: The number in parentheses represents the ranking of variables in descending order for each annual system.
Abbreviations: C, corn; NH_4^+, ammonium content; NO_3^-, soil nitrate content; S, soybean; $\sum ppt_{2d}$, cumulative 2-day precipitation; T_{avg}, average air temperature; WFPS, water-filled pore space; W, winter wheat.

levels exceeded 15 kg N ha^{-1}, concurrent with T_{avg} surpassing 20 ˚C (Figures 4a, 4e, and 5), which was also highlighted by the decision tree analysis (Figure S4). In the NT system, which differs from conventional only in tillage, climate variables $\sum ppt_{2d}$ and T_{avg} were ranked higher than other soil and management variables (Table 3). These two variables were repeatedly used for node splitting in the decision tree (Figure S5), further highlighting their importance in controlling N_2O fluxes from the NT system. These climate factors explained the high variability in corn phases, while NH_4^+ along with T_{avg} emerged as primary influencers in the winter wheat phase (Table 3). The 2D plot showed high emissions when $\sum ppt_{2d}$ reached 55 mm and T_{avg} exceeded 20˚C (Figure 5). In the reduced input and biologically based/organic systems, where red clover precedes corn and cereal rye precedes soybean in the rotation, NO_3^- emerged as the predominant variable, with climate variables and NH_4^+ following in the reduced input system, and climate variables and WFPS in the biologically based/organic system (Table 3; Figures S6 and S7). The predictability of N_2O emissions is higher in the corn phase, with NO_3^- serving as a key predictor (Table 3), supported by the increased availability of NO_3^- resulting from the legume cover crop's being tilled into the soil before corn planting (Figure S8b). The emissions increase with an increase in NO_3^-, particularly evident at approximately 10 kg N ha^{-1}, with a much higher increase observed in the biologically based/organic system compared to the reduced input system (Figure 4d). A positive interaction between NO_3^- and WFPS was illustrated by the 2D plot, indicating higher N_2O losses when WFPS exceeded ∼40% and NO_3^- levels exceeded ∼17 kg ha^{-1} (Figure 5; Figure S7).

4 | DISCUSSION

Our findings highlight that, despite the lack of statistically significant differences in cumulative rotational N_2O emissions for the five cycles of W–C–S rotations among annual systems, different variables emerge as the most influential factors for N_2O fluxes in each annual system. The RF model, developed utilizing biweekly N_2O flux manual chamber measurements from the annual cropping systems of KBS-LTER site from 2003 to 2019, explained 29%–42% of daily flux variance in the testing dataset of the four annual systems. The climate (T_{avg} and $\sum ppt_{2d}$) and soil variables (WFPS, NO_3^-, and NH_4^+) employed in model development are widely recognized as drivers of N_2O emissions (Firestone & Davidson, 1989; Gelfand et al., 2016; Saha, Basso, & Robinson, 2021) and serve as easily measurable proxies for soil biophysical and biogeochemical processes. The RF model has been extensively validated and demonstrated its reliability in predicting N_2O emissions in croplands, with r^2 values ranging from 0.38 to 0.73 (Glenn et al., 2021; Joshi et al., 2024; Liao et al., 2023;

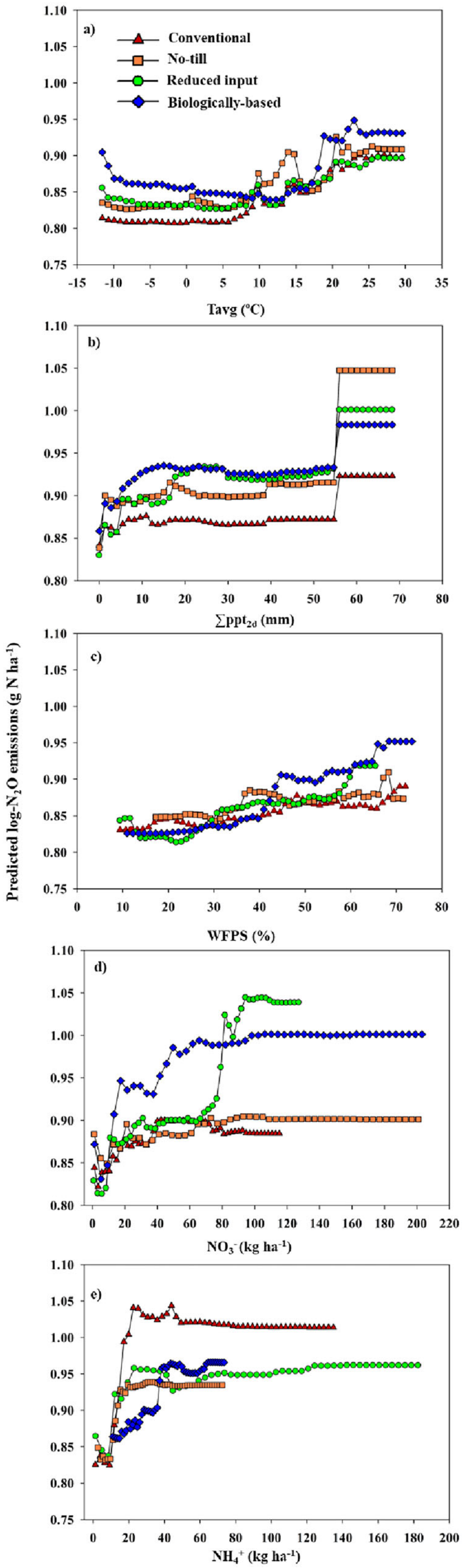

FIGURE 4 One-dimensional partial dependence of predictor variables (a) average air temperature (T_{avg}), (b) cumulative 2-day precipitation ($\sum ppt_{2d}$), (c) water-filled pore space (WFPS), (d) soil nitrate content (NO_3^-), and (e) soil ammonium content (NH_4^+) on N_2O emissions as predicted by the random forest model under conventional, no-till, reduced input, and biologically based/organic systems.

Philibert et al., 2013; Saha, Basso, & Robinson, 2021; Yin et al., 2022).

In the conventional cropping system, soil NH_4^+ and T_{avg} emerged as particularly significant predictors for N_2O fluxes. The model predicted a higher risk of N_2O emissions following N fertilization application ($NH_4^+ > 15$ kg N ha^{-1}) during periods of high air temperatures ($T_{avg} > 20°C$) (Figure 5; Figure S4). Emissions typically peak following the side-dress application of UAN during the corn and wheat phases, coinciding with high air temperatures (Figure 2), also reported in many other studies in temperate climates (Adviento-Borbe et al., 2007; Gasche & Papen, 2002; Kitzler et al., 2006; Ma et al., 2010). This observation might elucidate the relatively higher predictability for fluxes modeled during the corn (48%) and winter wheat (32%) phases compared to the soybean phase (16.7%) (Table 3), with NH_4^+ and T_{avg} emerging as the topmost variables within the corn and wheat phases. Microbial activities during nitrification and denitrification tend to be more active under higher temperatures (Kätterer et al., 1998), suggesting that air temperature plays an important role in regulating the rate of N_2O emissions (Rashti et al., 2015). Our findings, which highlight the primary influence of NH_4^+ and T_{avg} in conventional system, suggest that the emissions might be linked to strong nitrification activity in this system. High NH_4^+ levels have elsewhere also been associated with elevated N_2O emissions (Breitenbeck et al., 1980; Peyrard et al., 2016). However, Liang and Robertson (2021) conclude that nitrification is a minor source of N_2O emissions in this system based on combining soil-specific kinetics of nitrification-derived N_2O with 25 years of N_2O flux measurements. In that study, the maximum potential contributions from nitrification to in situ N_2O fluxes were found to be 13%–17%, with actual contributions likely only 1%–2%. Nitrification is rapid in these soils (Millar & Robertson, 2015), such that high NH_4^+ levels can simultaneously indicate high NO_3^- availability despite lower soil NO_3^- levels if NO_3^- pools are rapidly depleted by plant uptake, denitrification, or leaching (Gelfand et al., 2016; Syswerda et al., 2012). Inorganic N availability might be better (less ambiguously) assessed as a driver of N_2O fluxes by combining NH_4^+ and NO_3^- into a single soil mineral N predictor.

In the NT system, where all other management aspects remain identical to the conventional system except for the adoption of NT practices, the RF model had the least predictability (29%), reflecting the complexity of the underlying processes and drivers of N_2O production in NT systems. Despite emissions' rising similarly to those in the conventional system after fertilizer application (Figure 2), climate variables took precedence over the effects of soil mineral N variability. Changing climate factors can regulate soil O_2 dynamics, serve as proxies for soil biophysical processes, and impact N_2O emissions (Song et al., 2019). Precipitation primarily changes soil O_2 concentrations by displacing

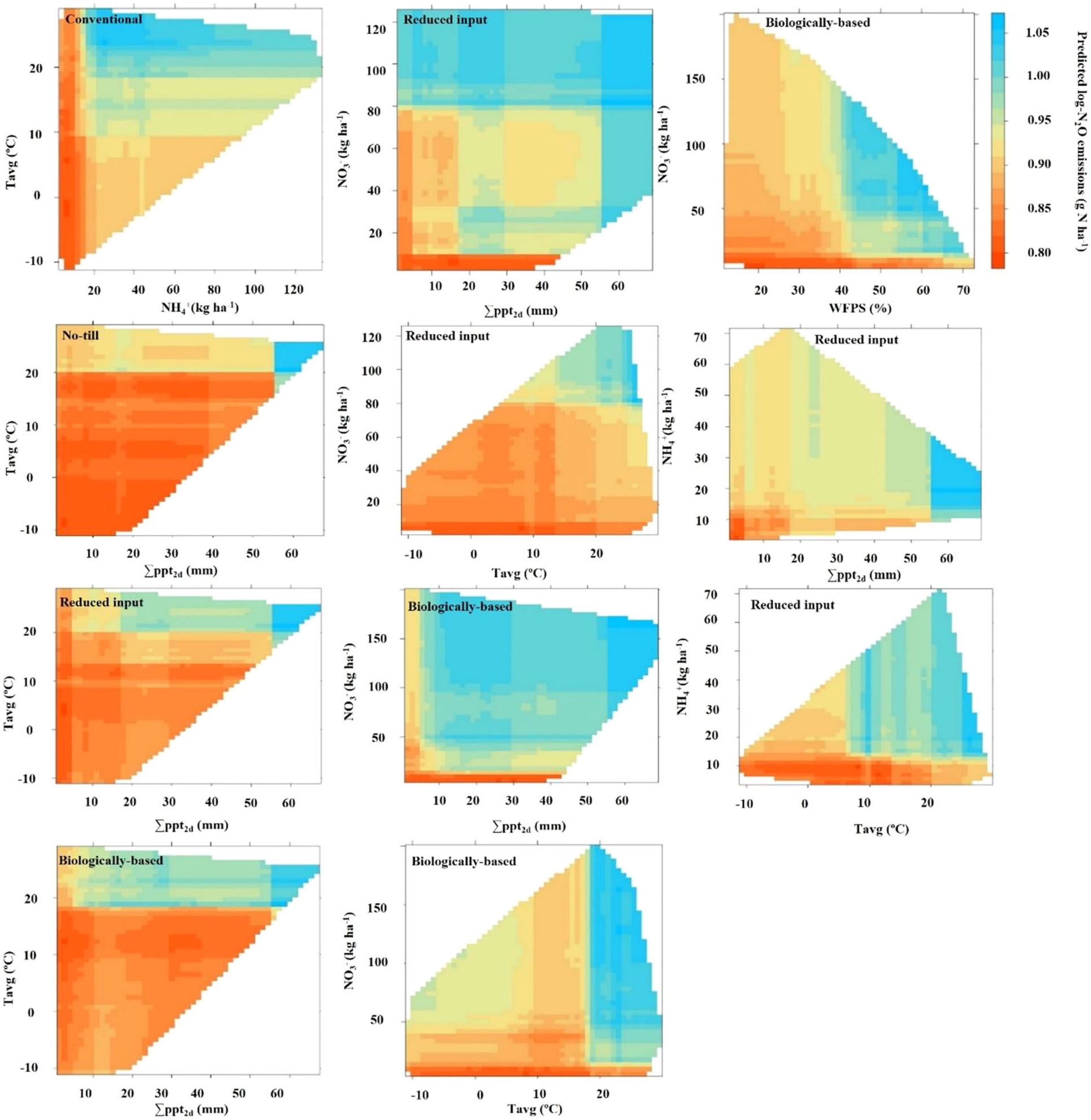

FIGURE 5 Two-dimensional partial dependence of selected predictor variables on N_2O emissions as predicted by the random forest model under conventional, no-till, reduced input, and biologically based systems.

soil air with water and serves as a reliable indicator of soil redox potential, affecting conditions that govern soil mineral N transformations and leading to N_2O production (Linn & Doran, 1984; Rochette et al., 2018). High T_{avg} (>~20°C) coupled with high $\sum ppt_{2d}$ (~55 mm) (Figure 5) can simultaneously promote microbial O_2 consumption via enhanced microbial activity and inhibited O_2 diffusion. Rochette et al. (2018) in their study on soil N_2O emissions and their controls in temperate climates of Canada reported that precipitation plays a primary role in determining N_2O emissions, and that environmental conditions can mask the impact of soil N content under well-aerated conditions. Grandy et al. (2006) documented increased aggregation and enhanced soil structure in the same NT system described here. Microsites or pores within these stable aggregates under long-term NT can lead to low and varying O_2 levels across the aggregate radius (Sexstone et al., 1985; Song et al., 2019), a dynamic not captured by WFPS, which typically serves as a proxy for soil O_2 fluctuation influenced by soil moisture levels (Dobbie & Smith, 2001, 2003). Likewise, differences in soil pores in well-structured soils can lead to microsite differences in water and O_2 levels that can drive differences in N_2O production (Kravchenko et al., 2017). Such heterogeneity in O_2 under well-structured soils could strongly impact N transformations and N_2O production. Thus, soil pore structure and soil O_2 parameters, which may serve as better proxies than WFPS and reliable predictors of N_2O emissions, could enhance the predictive capability of the RF model in the NT system. However, we acknowledge the challenges associated with accurately capturing high-resolution soil O_2 consumption within pore spaces. Furthermore, the availability of data on SOC may provide additional predictive capacity for the model. This is particularly significant, as NT is proposed as one of the main measures to reduce N_2O emissions and increase C sequestration (Van Kessel et al., 2013). A better understanding of controls of N_2O emissions in NT soils is required.

In contrast to the conventional and NT systems, the cover-cropped reduced input and biologically based/organic systems revealed soil NO_3^- as the primary variable explaining the largest portion of the variation in N_2O emissions (24%–26%; Table 3). This is particularly evident after incorporating red clover before corn planting, where there is a notable increase in NO_3^- availability (Figure S8b) and, concomitantly, in emissions (Figure 2). The model's high predictability of N_2O fluxes in the corn phase (43% in the reduced input and 60% in the biologically based/organic system), with NO_3^- as the top variable, further demonstrates this. Furthermore, the lower predictive value observed for wheat in these systems, along with the lower ranking of NO_3^- within the wheat system, suggests that there is a limited carryover effect of the decomposition of leguminous cover crop biomass into the wheat phase. Increased N_2O associated with legume crops could be attributed to enhanced N release from decomposing leguminous residues (Abalos et al., 2022; Rochette & Janzen, 2005). In these systems, chisel tillage may enhance N mineralization by incorporating legume cover crops into the soil when temperatures are sufficiently warm to support active decomposition. This is further supported by the prediction of an increase in emissions with rising air temperature (Figure 4a). The simultaneous availability of easily degraded N and C from organic inputs increases the risk of high N_2O emissions by enhancing biological activity, leading to soil O_2 depletion through enhanced soil respiration and increased denitrification (Hansen et al., 2019; Lussich et al., 2024).

The heightened risk of significant N_2O emissions following precipitation events when WFPS exceeds 53% and NO_3^- availability surpasses 17 kg ha^{-1} (Figure 5) indicates N_2O likely originated from denitrification. This threshold value for WFPS aligns with the findings of Peyrard et al. (2016), although when denitrification is involved, the WFPS threshold is often higher, ranging from 60% to 80% (Davidson, 1991). The wetness-independent anoxia created by decomposing legume residues might partly explain N_2O production, a phenomenon not captured by WFPS. Respiration-induced anoxia caused by decomposing cover crop residues can promote N_2O emissions, even under suboptimal WFPS (50%) conditions for denitrification (Lussich et al., 2024). This could also hold true for the reduced input system, where the predictive value of WFPS is lower and no interaction of NO_3^- with $\sum ppt_{2d}$ was observed. Future advancements in our understanding and data availability regarding the response of N_2O to soil O_2 consumption during the decomposition of cover crop residues may enhance the predictive capacity of models in cover crop-based cropping systems.

5 | CONCLUSIONS

Results underscore the efficacy of a decision tree-based nonlinear machine learning model for identifying key variables, their threshold conditions, and complex interactions in influencing N_2O emissions in intensively managed annual cropping systems. Our findings leveraging long-term data reveal that differential controls of N_2O emissions are important under different cropping system managements. In the conventional system, soil ammonium and air temperature emerged as the primary influencers of N_2O emissions, while in the NT system, climatic conditions—particularly precipitation and air temperature—exerted the greatest impact on emissions. Nitrate availability from legume cover crops drove N_2O emissions in the reduced input and biologically based/organic systems.

Although our RF model effectively predicted 29%–42% of the daily variability in N_2O fluxes from intensively managed cropping systems, the model can be further improved by incorporating long-term high-frequency observations from automated flux chambers and by including soil organic carbon data, as NT and cover crop systems can influence N_2O emissions by enhancing soil carbon content. Considering the challenge posed by the generalizability of the RF model, its application to other regions and crops necessitates further enhancements in model training based on diverse data sources encompassing various soils, climates, crops, and management conditions.

AUTHOR CONTRIBUTIONS

Jashanjeet Kaur Dhaliwal: Conceptualization; formal analysis; investigation; methodology; software; validation; visualization; writing—original draft; writing—review and editing. **Dinesh Panday**: Data curation; formal analysis; methodology; visualization; writing—original draft; writing—review and editing. **G. Philip Robertson**: Data curation; funding acquisition; methodology; resources; supervision; writing—review and editing. **Debasish Saha**: Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

The funding support for this research was provided by the USDA NIFA (Award# 2021-67019-34247), the NSF Long-term Ecological Research Program (DEB 2224712) and USDA Long-term Agroecosystem Research program at the Kellogg Biological Station, and by Michigan State University AgBioResearch.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data supporting the findings and R codes used in this study are openly available in Dryad at https://doi.org/10.5061/dryad.9cnp5hqv1.

ORCID

Dinesh Panday https://orcid.org/0000-0001-8452-3797
G. Philip Robertson https://orcid.org/0000-0001-9771-9895
Debasish Saha https://orcid.org/0000-0001-9425-675X

REFERENCES

Abalos, D., Recous, S., Butterbach-Bahl, K., De Notaris, C., Rittl, T. F., Topp, C. F. E., Petersen, S. O., Hansen, S., Bleken, M. A., Rees, R. M., & Olesen, J. E. (2022). A review and meta-analysis of mitigation measures for nitrous oxide emissions from crop residues. *Science of the Total Environment*, *828*, 154388.

Adviento-Borbe, M. A. A., Haddix, M. L., Binder, D. L., Walters, D. T., & Dobermann, A. (2007). Soil greenhouse gas fluxes and global warming potential in four high-yielding maize systems. *Global Change Biology*, *13*, 1972–1988.

Ball, B., Crichton, I., & Horgan, G. (2008). Dynamics of upward and downward N_2O and CO_2 fluxes in ploughed or no-tilled soils in relation to water-filled pore space, compaction and crop presence. *Soil and Tillage Research*, *101*, 20–30.

Baral, K. R., Jayasundara, S., Brown, S. E., & Wagner-Riddle, C. (2022). Long-term variability in N_2O emissions and emission factors for corn and soybeans induced by weather and management at a cold climate site. *Science of the Total Environment*, *815*, 152744.

Basche, A. D., Miguez, F. E., Kaspar, T. C., & Castellano, M. J. (2014). Do cover crops increase or decrease nitrous oxide emissions? A meta-analysis. *Journal of Soil and Water Conservation*, *69*, 471–482.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Breitenbeck, G., Blackmer, A., & Bremner, J. (1980). Effects of different nitrogen fertilizers on emission of nitrous oxide from soil. *Geophysical Research Letters*, *7*, 85–88.

Burchill, W., Li, D., Lanigan, G. J., Williams, M., & Humphreys, J. (2014). Interannual variation in nitrous oxide emissions from perennial ryegrass/white clover grassland used for dairy production. *Global Change Biology*, *20*, 3137–3146.

Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M., Kiese, R., & Zechmeister-Boltenstern, S. (2013). Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*, 20130122.

Crum, J., & Collins, H. (1995). *KBS soils*. https://doi.org/10.5281/zenodo.2581504

Cusser, S., Bahlai, C., Swinton, S. M., Robertson, G. P., & Haddad, N. M. (2020). Long-term research avoids spurious and misleading trends in sustainability attributes of no-till. *Global Change Biology*, *26*, 3715–3725.

Davidson, E. A. (1991). Fluxes of nitrous oxide and nitric oxide from terrestrial ecosystems. In J. E. Rogers & W. B. Whitman (Eds.), *Microbial production and consumption of greenhouse gases:* Methane*, nitrogen oxides, and halomethanes* (pp. 219–235). ASM Press.

Davis, B. W., Mirsky, S. B., Needelman, B. A., Cavigelli, M. A., & Yarwood, S. A. (2019). Nitrous oxide emissions increase exponentially with organic N rate from cover crops and applied poultry litter. *Agriculture, Ecosystems & Environment*, *272*, 165–174.

de Klein, C. A., Alfaro, M. A., Giltrap, D., Topp, C. F., Simon, P. L., Noble, A. D., & van der Weerden, T. J. (2020). Global Research Alliance N_2O chamber methodology guidelines: Statistical considerations, emission factor calculation, and data reporting. *Journal of Environmental Quality*, *49*, 1156–1167.

Del Grosso, S., Parton, W., Mosier, A., Ojima, D., Kulmala, A., & Phongpan, S. (2000). General model for N_2O and N_2 gas emissions from soils due to dentrification. *Global Biogeochemical Cycles*, *14*, 1045–1060.

Deng, J., Guo, L., Salas, W., Ingraham, P., Charrier-Klobas, J. G., Frolking, S., & Li, C. (2022). A decreasing trend of nitrous oxide emissions from California cropland from 2000 to 2015. *Earth's Future*, *10*, e2021EF002526.

Deng, J., Li, C., Burger, M., Horwath, W. R., Smart, D., Six, J., Guo, L., Salas, W., & Frolking, S. (2018). Assessing short-term impacts of management practices on N_2O emissions from diverse Mediterranean agricultural ecosystems using a biogeochemical model. *Journal of Geophysical Research: Biogeosciences*, *123*, 1557–1571.

Djiemon, A., Gasser, M.-O., & Gallichand, J. (2019). Random forests to detect subsoiling and subsurface drainage effects on corn plant height and water table depth. *Soil and Tillage Research*, *192*, 240–249.

Dobbie, K. E., & Smith, K. A. (2001). The effects of temperature, water-filled pore space and land use on N_2O emissions from an imperfectly drained gleysol. *European Journal of Soil Science*, *52*, 667–673.

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.

Dorich, C. D., Conant, R. T., Albanito, F., Butterbach-Bahl, K., Grace, P., Scheer, C., Snow, V. O., Vogeler, I., & van der Weerden, T. J. (2020). Improving N2O emission estimates with the global N2O database. *Current Opinion in Environmental Sustainability*, *47*, 13–20.

Ehrhardt, F., Soussana, J.-F., Bellocchi, G., Grace, P., Mcauliffe, R., Recous, S., Sándor, R., Smith, P., Snow, V., De Antoni Migliorati, M., Basso, B., Bhatia, A., Brilli, L., Doltra, J., Dorich, C. D., Doro, L., Fitton, N., Giacomini, S. J., Grant, B., … Zhang, Q. (2018). Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N_2O emissions. *Global Change Biology*, *24*, e603–e616.

Finney, D. M., Eckert, S. E., & Kaye, J. P. (2015). Drivers of nitrogen dynamics in ecologically based agriculture revealed by long-term, high-frequency field measurements. *Ecological Applications*, *25*, 2210–2227.

Firestone, M. K., & Davidson, E. A. (1989). Microbiological basis of NO and N2O production and consumption in soil. In M. O. Andreae & D. S. Schimel (Eds.), *Exchange of trace gases between terrestrial ecosystems and the atmosphere* (Vol. 47, pp. 7–21). Wiley.

Fuchs, K., Merbold, L., Buchmann, N., Bretscher, D., Brilli, L., Fitton, N., Topp, C. F., Klumpp, K., Lieffering, M., & Martin, R. (2020). Multimodel evaluation of nitrous oxide emissions from an intensively managed grassland. *Journal of Geophysical Research: Biogeosciences*, *125*, e2019JG005261.

Gaillard, R. K., Jones, C. D., Ingraham, P., Collier, S., Izaurralde, R. C., Jokela, W., Osterholz, W., Salas, W., Vadas, P., & Ruark, M. D. (2018). Underestimation of N_2O emissions in a comparison of the DayCent, DNDC, and EPIC models. *Ecological Applications*, *28*, 694–708.

Gasche, R., & Papen, H. (2002). Spatial variability of NO and NO_2 flux rates from soil of spruce and beech forest ecosystems. *Plant and Soil*, *240*, 67–76.

Gelfand, I., Shcherbak, I., Millar, N., Kravchenko, A. N., & Robertson, G. P. (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.

Gilhespy, S. L., Anthony, S., Cardenas, L., Chadwick, D., Del Prado, A., Li, C., Misselbrook, T., Rees, R. M., Salas, W., Sanz-Cobena, A., Smith, P., Tilston, E L., Topp, C. F. E., Vetter, S., & Yeluripati, J. B. (2014). First 20 years of DNDC (denitrification decomposition): model evolution. *Ecological Modelling*, *292*, 51–62.

Giltrap, D. L., Li, C., & Saggar, S. (2010). DNDC: A process-based model of greenhouse gas fluxes from agricultural soils. *Agriculture, Ecosystems & Environment*, *136*, 292–300.

Glenn, A. J., Moulin, A. P., Roy, A. K., & Wilson, H. F. (2021). Soil nitrous oxide emissions from no-till canola production under variable rate nitrogen fertilizer management. *Geoderma*, *385*, 114857.

Grandy, A. S., Loecke, T. D., Parr, S., & Robertson, G. P. (2006). Long-term trends in nitrous oxide emissions, soil nitrogen, and crop yields of till and no-till cropping systems. *Journal of Environmental Quality*, *35*, 1487–1495.

Groffman, P. M., Butterbach-Bahl, K., Fulweiler, R. W., Gold, A. J., Morse, J. L., Stander, E. K., Tague, C., Tonitto, C., & Vidon, P. (2009). Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry*, *93*, 49–77.

Hansen, S., Berland Frøseth, R., Stenberg, M., Stalenga, J., Olesen, J. E., Krauss, M., Radzikowski, P., Doltra, J., Nadeem, S., Torp, T., Pappa, V., & Watson, C A. (2019). Reviews and syntheses: Review of causes and sources of N_2O emissions and NO_3 leaching from organic arable crop rotations. *Biogeosciences*, *16*, 2795–2819.

Hoben, J., Gehl, R., Millar, N., Grace, P., & Robertson, G. (2011). Nonlinear nitrous oxide (N_2O) response to nitrogen fertilizer in on-farm corn crops of the US Midwest. *Global Change Biology*, *17*, 1140–1152.

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*, 143–157.

Huang, Y., Lan, Y., Thomson, S. J., Fang, A., Hoffmann, W. C., & Lacey, R. E. (2010). Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture*, *71*, 107–127.

Jarecki, M. K., Parkin, T. B., Chan, A. S., Hatfield, J. L., & Jones, R. (2008). Comparison of DAYCENT-simulated and measured nitrous oxide emissions from a corn field. *Journal of Environmental Quality*, *37*, 1685–1690.

Joshi, D. R., Clay, D. E., Clay, S. A., Moriles-Miller, J., Daigh, A. L. M., Reicks, G., & Westhoff, S. (2024). Quantification and machine learning based N_2O–N and CO_2–C emissions predictions from a decomposing rye cover crop. *Agronomy Journal*, *116*, 795–809.

Kätterer, T., Reichstein, M., Andrén, O., & Lomander, A. (1998). Temperature dependence of organic matter decomposition: A critical review using literature data analyzed with different models. *Biology and Fertility of Soils*, *27*, 258–262.

Kaye, J. P., & Quemada, M. (2017). Using cover crops to mitigate and adapt to climate change. A review. *Agronomy for Sustainable Development*, *37*, Article 4.

Kim, D.-G., Hernandez-Ramirez, G., & Giltrap, D. (2013). Linear and nonlinear dependency of direct nitrous oxide emissions on fertilizer nitrogen input: A meta-analysis. *Agriculture*, *Ecosystems & Environment*, *168*, 53–65.

Kitzler, B., Zechmeister-Boltenstern, S., Holtermann, C., Skiba, U., & Butterbach-Bahl, K. (2006). Controls over N_2O, NOx and CO_2 fluxes in a calcareous mountain forest soil. *Biogeosciences*, *3*, 383–395.

Kravchenko, A. N., Toosi, E. R., Guber, A. K., Ostrom, N. E., Yu, J., Azeem, K., Rivers, M. L., & Robertson, G. P. (2017). Hotspots of soil N_2O emission enhanced through water absorption by plant residue. *Nature Geoscience*, *10*, 496–500.

Li, C. (2000). Modeling trace gas emissions from agricultural ecosystems. In R. Wassmann, R. S. Lantin, & H.-U. Neue (Eds.), *Methane emissions from major rice ecosystems in Asia* (pp. 259–276). Springer.

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

Liang, D., & Robertson, G. P. (2021). Nitrification is a minor source of nitrous oxide (N_2O) in an agricultural landscape and declines with increasing management intensity. *Global Change Biology*, *27*, 5599–5613.

Liao, J., Huang, Y., Li, Z., & Niu, S. (2023). Data-driven modeling on the global annual soil nitrous oxide emissions: Spatial pattern and attributes. *Science of the Total Environment*, *903*, 166472.

Linn, D. M., & Doran, J. W. (1984). Effect of water-filled pore space on carbon dioxide and nitrous oxide production in tilled and nontilled soils. *Soil Science Society of America Journal*, *48*, 1267–1272.

Luehmann, M. D., Peter, B. G., Connallon, C. B., Schaetzl, R. J., Smidt, S. J., Liu, W., Kincare, K. A., Walkowiak, T. A., Thorlund, E., & Holler, M. S. (2016). Loamy, two-storied soils on the outwash plains of southwestern lower Michigan: pedoturbation of loess with the underlying sand. *Annals of the American Association of Geographers*, *106*, 551–572.

Lussich, F., Dhaliwal, J. K., Faiia, A. M., Jagadamma, S., Schaeffer, S. M., & Saha, D. (2024). Cover crop residue decomposition triggered soil oxygen depletion and promoted nitrous oxide emissions. *Scientific Reports*, *14*, Article 8437.

Ma, B., Wu, T., Tremblay, N., Deen, W., Morrison, M., McLaughlin, N., Gregorich, E., & Stewart, G. (2010). Nitrous oxide fluxes from corn fields: on-farm assessment of the amount and timing of nitrogen fertilizer. *Global Change Biology*, *16*, 156–170.

Maharjan, B., & Venterea, R. T. (2013). Nitrite intensity explains N management effects on N_2O emissions in maize. *Soil Biology and Biochemistry*, *66*, 229–238.

Mei, K., Wang, Z., Huang, H., Zhang, C., Shang, X., Dahlgren, R. A., Zhang, M., & Xia, F. (2018). Stimulation of N2O emission by conservation tillage management in agricultural lands: A meta-analysis. *Soil and Tillage Research*, *182*, 86–93.

Millar, N., & Robertson, G. P. (2015). Nitrogen transfers and transformations in row-crop ecosystems. 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. 213–251). Oxford University Press.

Millar, N., Urrea, A., Kahmark, K., Shcherbak, I., Robertson, G. P., & Ortiz-Monasterio, I. (2018). Nitrous oxide (N_2O) flux responds exponentially to nitrogen fertilizer in irrigated wheat in the Yaqui Valley, Mexico. *Agriculture, Ecosystems & Environment*, *261*, 125–132.

Nouri, A., Lee, J., Yin, X., Tyler, D. D., & Saxton, A. M. (2019). Thirty-four years of no-tillage and cover crops improve soil quality and increase cotton yield in Alfisols, Southeastern USA. *Geoderma*, *337*, 998–1008.

Oertel, C., Matschullat, J., Zurba, K., Zimmermann, F., & Erasmi, S. (2016). Greenhouse gas emissions from soils—A review. *Geochemistry*, *76*, 327–352.

Panday, D., Saha, D., Lee, J., Jagadamma, S., Adotey, N., & Mengistu, A. (2022). Cover crop residue influence on soil N_2O and CO_2 emissions under wetting-drying intensities: An incubation study. *European Journal of Soil Science*, *73*, e13309.

Parton, W., Holland, E., Del Grosso, S., Hartman, M., Martin, R., Mosier, A., Ojima, D., & Schimel, D. (2001). Generalized model for NO_x and N_2O emissions from soils. *Journal of Geophysical Research: Atmospheres*, *106*, 17403–17419.

Peyrard, C., Mary, B., Perrin, P., Véricel, G., Gréhan, E., Justes, E., & Léonard, J. (2016). N2O emissions of low input cropping systems as affected by legume and cover crops use. *Agriculture, Ecosystems & Environment*, *224*, 145–156.

Philibert, A., Loyce, C., & Makowski, D. (2013). Prediction of N_2O emission from local information with random forest. *Environmental Pollution*, *177*, 156–163.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ramírez-Melgarejo, M., Reyes-Figueroa, A., Gassó-Domingo, S., & Güereca, L. P. (2020). Analysis of empirical methods for the quantification of N_2O emissions in wastewater treatment plants: Comparison of emission results obtained from the IPCC Tier 1 methodology and the methodologies that integrate operational data. *Science of the Total Environment*, *747*, 141288.

Rashti, M. R., Wang, W., Moody, P., Chen, C., & Ghadiri, H. (2015). Fertiliser-induced nitrous oxide emissions from vegetable production in the world and the regulating factors: A review. *Atmospheric Environment*, *112*, 225–233.

Richards, M., Metzel, R., Chirinda, N., Ly, P., Nyamadzawo, G., Duong Vu, Q., De Neergaard, A., Oelofse, M., Wollenberg, E., Keller, E., Malin, D., Olesen, J. E., Hillier, J., & Rosenstock, T. S. (2016). Limits of agricultural greenhouse gas calculators to predict soil N2O and CH4 fluxes in tropical agriculture. *Scientific Reports*, *6*, Article 26279.

Robertson, G. P. (2015). A sustainable agriculture? *Daedalus*, *144*, 76–89.

Robertson, G. P. (2023). Denitrification and the challenge of scaling microsite knowledge to the globe. *mLife*, *2*, 229–238.

Robertson, G. P., & Groffman, P. M. (2024). Nitrogen transformations. In E. A. Paul & S. D. Frey (Eds.), *Soil microbiology, ecology, and biochemistry* (5th ed., pp. 407–438). Elsevier.

Robertson, G. P., & Hamilton, S. K. (2015). Long-term ecological research at the Kellogg Biological Station LTER site. 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). Oxford University Press.

Rochette, P., & Janzen, H. H. (2005). Towards a revised coefficient for estimating N2O emissions from legumes. *Nutrient Cycling in Agroecosystems*, *73*, 171–179.

Rochette, P., Liang, C., Pelster, D., Bergeron, O., Lemke, R., Kroebel, R., MacDonald, D., Yan, W., & Flemming, C. (2018). Soil nitrous oxide emissions from agricultural soils in Canada: Exploring relationships with soil, crop and climatic variables. *Agriculture, Ecosystems & Environment*, *254*, 69–81.

Rowlings, D., Grace, P., Scheer, C., & Liu, S. (2015). Rainfall variability drives interannual variation in N_2O emissions from a humid, subtropical pasture. *Science of the Total Environment*, *512*, 8–18.

Saha, D., Basso, B., & Robertson, G. P. (2021). Machine learning improves predictions of agricultural nitrous oxide (N_2O) emissions from intensively managed cropping systems. *Environmental Research Letters*, *16*, 024004.

Saha, D., Kaye, J. P., Bhowmik, A., Bruns, M. A., Wallace, J. M., & Kemanian, A. R. (2021). Organic fertility inputs synergistically increase denitrification-derived nitrous oxide emissions in agroecosystems. *Ecological Applications*, *31*, e02403.

Saha, D., Kemanian, A. R., Montes, F., Gall, H., Adler, P. R., & Rau, B. M. (2018). Lorenz curve and Gini coefficient reveal hot spots and hot moments for nitrous oxide emissions. *Journal of Geophysical Research: Biogeosciences*, *123*, 193–206.

Saha, D., Kemanian, A. R., Rau, B. M., Adler, P. R., & Montes, F. (2017). Designing efficient nitrous oxide sampling strategies in agroecosystems using simulation models. *Atmospheric Environment*, *155*, 189–198.

Scheer, C., Rowlings, D. W., & Grace, P. R. (2016). Non-linear response of soil N_2O emissions to nitrogen fertiliser in a cotton–fallow rotation in sub-tropical Australia. *Soil Research*, *54*, 494–499.

Sexstone, A. J., Revsbech, N. P., Parkin, T. P., & Tiedje, J. M. (1985). Direct measurement of oxygen profiles and denitrification rates in soil aggregates. *Soil Science Society of America Journal*, *49*, 645–651.

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*, *111*, 9199–9204.

Six, J., Ogle, S. M., Jay Breidt, F., Conant, R. T., Mosier, A. R., & Paustian, K. (2004). The potential to mitigate global warming with no-tillage management is only realized when practised in the long term. *Global Change Biology*, *10*, 155–160.

Song, X., Ju, X., Topp, C. F., & Rees, R. M. (2019). Oxygen regulates nitrous oxide production directly in agricultural soils. *Environmental Science & Technology*, *53*, 12539–12547.

Syakila, A., & Kroeze, C. (2011). The global nitrous oxide budget revisited. *Greenhouse Gas Measurement and Management*, *1*, 17–26.

Syswerda, S., Basso, B., Hamilton, S., Tausig, J., & Robertson, G. (2012). Long-term nitrate loss along an agricultural intensity gradient in the upper Midwest USA. *Agriculture, Ecosystems & Environment*, *149*, 10–19.

Tian, L., Cai, Y., & Akiyama, H. (2019). A review of indirect N_2O emission factors from agricultural nitrogen leaching and runoff to update of the default IPCC values. *Environmental Pollution*, *245*, 300–306.

Ussiri, D. A., Lal, R., & Jarecki, M. K. (2009). Nitrous oxide and methane emissions from long-term tillage under a continuous corn cropping system in Ohio. *Soil and Tillage Research*, *104*, 247–255.

Van Kessel, C., Venterea, R., Six, J., Adviento-Borbe, M. A., Linquist, B., & Van Groenigen, K. J. (2013). Climate, duration, and N placement determine N_2O emissions in reduced tillage systems: A meta-analysis. *Global Change Biology*, *19*, 33–44.

Venterea, R. T., Halvorson, A. D., Kitchen, N., Liebig, M. A., Cavigelli, M. A., Grosso, S. J. D., Motavalli, P. P., Nelson, K. A., Spokas, K. A., Singh, B. P., Stewart, C. E., Ranaivoson, A., Strock, J., & Collins, H. (2012). Challenges and opportunities for mitigating nitrous oxide emissions from fertilized cropping systems. *Frontiers in Ecology and the Environment*, *10*, 562–570.

Wang, J., & Zou, J. (2020). No-till increases soil denitrification via its positive effects on the activity and abundance of the denitrifying community. *Soil Biology and Biochemistry*, *142*, 107706.

Wanyama, I., Pelster, D. E., Arias-Navarro, C., Butterbach-Bahl, K., Verchot, L. V., & Rufino, M. C. (2018). Management intensity controls soil N_2O fluxes in an Afromontane ecosystem. *Science of the Total Environment*, *624*, 769–780.

Warncke, D., Dahl, J., Jacobs, L., & Laboski, C. (2009). *Nutrient recommendations for field crops in Michigan*. Michigan State University Extension.

Yin, Y., Wang, Z., Tian, X., Wang, Y., Cong, J., & Cui, Z. (2022). Evaluation of variation in background nitrous oxide emissions: A new global synthesis integrating the impacts of climate, soil, and management conditions. *Global Change Biology*, *28*, 480–492.

Zhang, Y., Zhang, N., Yin, J., Yang, F., Zhao, Y., Jiang, Z., Tao, J., Yan, X., Qiu, Y., Guo, H., & Hu, S. (2020). Combination of warming and N inputs increases the temperature sensitivity of soil N_2O emission in a Tibetan alpine meadow. *Science of the Total Environment*, *704*, 135450.

SUPPORTING INFORMATION

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

How to cite this article: Dhaliwal, J. K., Panday, D., Robertson, G. P., & Saha, D. (2024). Machine learning reveals dynamic controls of soil nitrous oxide emissions from diverse long-term cropping systems. *Journal of Environmental Quality*, 1–15. https://doi.org/10.1002/jeq2.20637-