Validation Report of the *ecosys* Model Version 1.0

HABITERRE

Prepared for: Climate Action Reserve (CAR)

Model requirements version:

Requirements and Guidance for Model Calibration, Validation, Uncertainty, and

Verification For Soil Enrichment Projects, Version 1.1a

Prepared by: HabiTerre, Inc.

Contact: 60 Hazelwood Dr Ste. 224, Champaign, IL 61820

science@habiterre.com

Table of Contents

1. Report Type
   1.1 Model validation Report type
   1.2 SEP version
   1.3 SEP model requirements version
   1.4 Model version
   1.5 Changes from previous validation report
2. Introduction
   2.1 Overview of the *ecosys* model
3. Responsible parties
4. Model Calibration
5. Model simulations
   5.1 Model setup
   5.2 Documentation of model parameters
6. Project domain
   6.1 Practice categories
   6.2 Crop functional type
   6.3 Land Resource Regions and Climate types
   6.4 Soil texture classes
   6.5 Emission sources
7. Description of validation data collection process and final validation dataset
   7.1 Description of validation data collection process
   7.2 Procedures to handle missing data
   7.3 Description of validation dataset
8. Overall model performance
9. Model bias evaluation
   9.1 Calculation of bias and PMU
   9.2 An example of PMU calculation
   9.3 Comparison between bias and PMU
   9.4 Bias for each study across the PC, CFG, and ES combinations

10. Model prediction error

10.1 Model prediction error for practice change effects on SOC changes across all practice categories and all crops

10.2 Model prediction error for practice change effects on direct N_2O emissions across all practice categories and all crops

References

Appendix A: Carbon flux modeling with HabiTerre's advanced model-data fusion framework

Appendix B: Key parameters used in the model employed in this model version report

Appendix C: Methods of temporal aggregation to estimate annually or seasonally accumulated N_2O emission

Appendix D: Studies in the Validation Dataset

Appendix E: The sensitivity of SOC stock to SOC sampling depth

Appendix F: Deviation Request Email Correspondence with CAR

Appendix G: Practice and crop specific model prediction errors

G.1 Model prediction error for practice change effects on SOC changes under CROPPING practices

G.2 Model prediction error for practice change effects on SOC changes under DISTURBANCE practices

G.3 Model prediction error for practice change effects on SOC changes with corn under FERTILIZER practices

G.4 Model prediction error for practice change effects on SOC changes under ORGANIC practices

G5 Model prediction error for practice change effects on SOC changes in corn fields

G.6 Model prediction error for practice change effects on SOC changes in soybean fields

G.7 Model prediction error for practice change effects on direct N_2O emissions under CROPPING practices

G.8 Model prediction error for practice change effects on direct N_2O emissions under DISTURBANCE practices

G.9 Model prediction error for practice change effects on direct N_2O emissions under FERTILIZER practices

G.10 Model prediction error for practice change effects on direct N_2O emissions under ORGANIC practices

G.11 Model prediction error for practice change effects on direct N_2O emissions in corn fields
G.12 Model prediction error for practice change effects on direct N_2O emissions in soybean fields

1. Report Type

1.1 Model validation Report type

This model validation report is a Type 2 (Generalized) report.

1.2 SEP version

This report is following SEP Version 1.1a (dated May 31, 2022), accessed on 13 Feb, 2023.

1.3 SEP model requirements version

This report is following Requirements and Guidance for Model Calibration, Validation, Uncertainty, and Verification for Soil Enrichment Projects, Version 1.1a (dated April, 2022), accessed on 13 Feb 2023 (referred to hereafter as the "SEP Model Requirements").

1.4 Model version

The *ecosys* model is an hourly time-step model with multiple canopy and soil layers that include comprehensive biophysical and biogeochemical interactions that determine ecosystem structure and function as well as the agricultural outcomes from diverse management practices. The *ecosys* model was originally developed at University of Alberta (Grant, 2001). The ecosys model was publicly released under the BSD-3-Clause license in 2019. HabiTerre has adapted and maintained an internal *ecosys* version 1.0 (ecosys_HT, hereafter ecosys is still used as the name for simplicity), spun off from the *ecosys* model released in 2019, to support the carbon outcome quantification of sustainable agricultural practices.

1.5 Changes from previous validation report

NA

2. Introduction

This report describes the validation of the *ecosys* model in simulating Soil Organic Carbon (SOC) changes and direct nitrous oxide (N_2O) emissions from soil (hereafter N_2O emission means direct soil emission) under different conservation land management practices in annual row crop systems covered by the validation datasets.

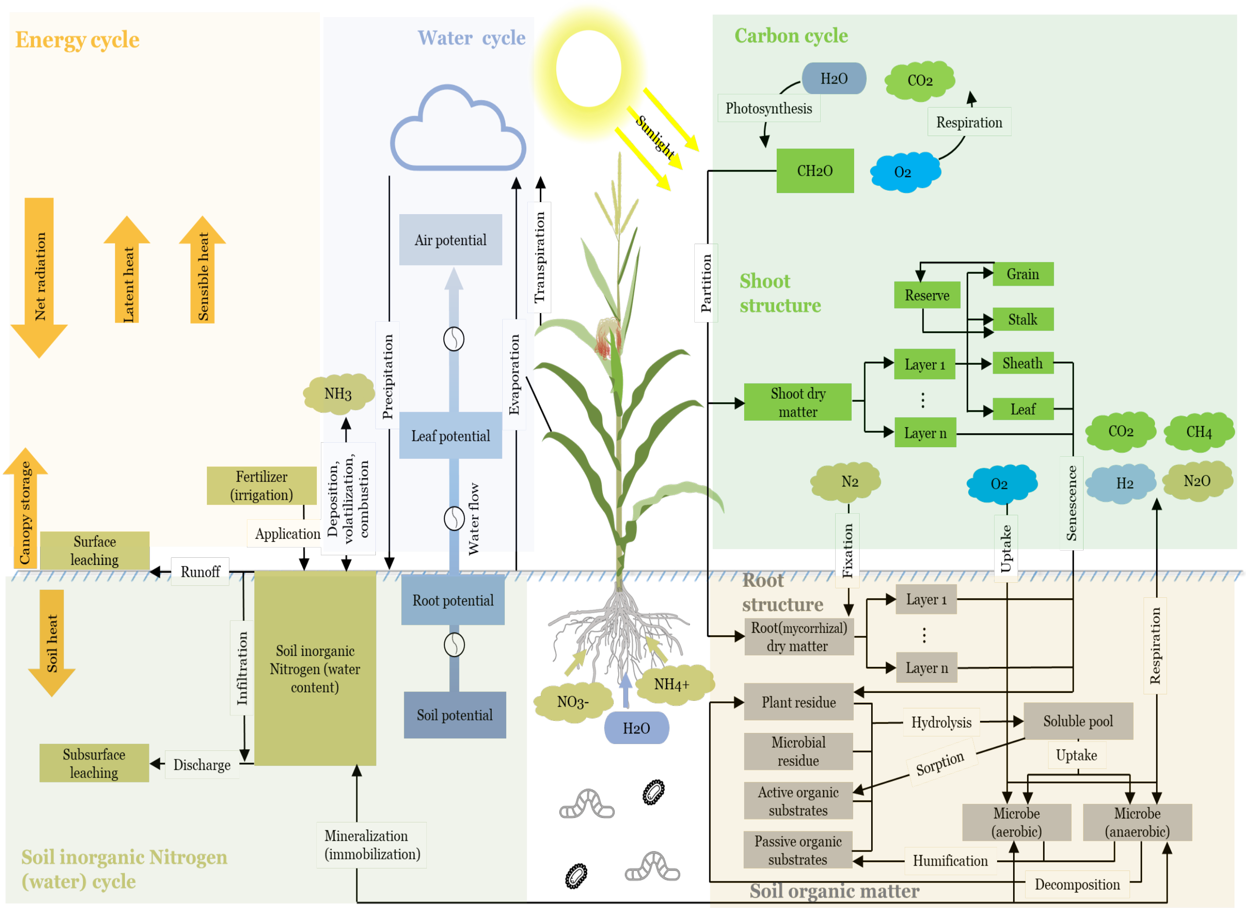

Figure 1. A conceptual structure of processes represented in the *ecosys* model (Qin et al., 2021).

2.1 Overview of the *ecosys* model

The *ecosys* model (Grant, 2001) is a process-based ecosystem model which simulates the holistic and coupled energy, water, carbon, and nutrient cycles from aboveground to belowground using a mechanistic way (Figure 1). The *ecosys* model employs complete physical and chemical theories in simulating plants and soil-related processes, and therefore is widely applicable to different soils, climates and land management conditions. The model explicitly includes microbes' competitive and symbiotic nutrient interactions with plants, enabling a nutrient-based analysis of how various management practices could affect crop productivity. Major farming practices that affect SOC dynamics and N_2O emission (i.e., crop rotation, tillage, cover crop, fertilizer and manure application, and irrigation) can be simulated by *ecosys*. Previous work using *ecosys* has fully demonstrated its capabilities in simulating the soil nitrogen cycle, N_2O and CH_4 emissions, long-term SOC trend, and impacts of different conservation land management practices. The *ecosys* version evaluated in this model validation report is a model version archived by the HabiTerre Science Team, which has been calibrated for the major row crops and cover crops across the U.S. Midwest in previous studies (Zhou et al., 2021; Qin et al., 2021; 2023; Li et al., 2022; Liu et al., 2021; Yang et al., 2022).

Unique strengths of the ecosys model include:

(1) The *ecosys* model simulates the whole carbon budget on farmland and the change of SOC at field level is holistically related to the farmland carbon budget in *ecosys* (see equation in Figure 2).
(2) The *ecosys* model simulates the SOC transformation rates using Michaelis-Menten equations, which are more mechanistic than the first-order decay.
(3) The *ecosys* model explicitly simulates the dynamics of multiple microbial groups and considers the microbial population and biophysiological activity in simulating SOC change, N_2O and CH_4 emissions.
(4) The *ecosys* model explicitly simulates the environmental drivers/regulators for microbial activities, like soil nutrient, water, temperature, oxygen, and pH conditions.
(5) The *ecosys* model explicitly simulates canopy energy balance for the aboveground processes, uses start-of-the-art representation of plant photosynthesis through linking stomatal conductance and biochemical photosynthesis models, and simulates plant water and nutrient uptake and transport from root to stem to leaf with advanced plant hydraulics, which are missing in many other models.
(6) The *ecosys* model is readily applicable to cropland, grassland, and all other natural ecosystems (like forest and wetland). For cropland and grassland, the impacts of agricultural management practices (e.g., crop rotation, nutrient management, manure application, cover cropping, conservation tillage, etc) are fully considered.
(7) With more detailed process representation and considerations of fundamental multi-process interactions, most of the responses in *ecosys* (such as the responses of SOC decomposition rate to soil water stress, and the responses of SOC change and GHG emissions to management changes) are simulated, rather than prescribed.

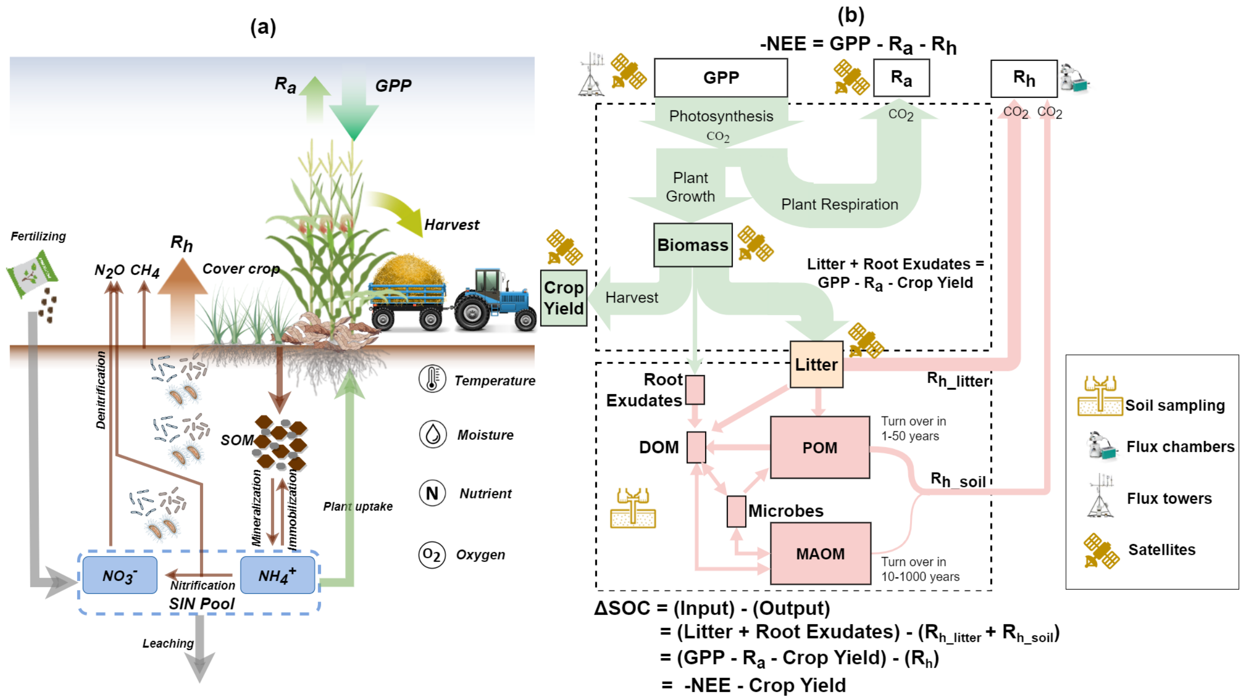

Figure 2. The *ecosys* model presents the holistic carbon and nitrogen balance and its linkage with greenhouse gas emissions throughout the agroecosystem (a) and uses a mass-balance-based approach to quantify the change of soil organic carbon (SOC) (b). GPP: gross primary productivity; Ra: autotrophic respiration; Rh: heterotrophic respiration; NEE: net ecosystem exchange; DOM: dissolved organic matter; POM: particulate organic matter; MAOM: mineral-associated organic matter. For more details, refer to Guan et al. (2023).

3. Responsible parties

Calibration, validation, and running of *ecosys* for this report were all performed by the Science Team of HabiTerre. As required in Section 5 of the SEP Model Requirements, HabiTerre has the requisite expertise to calibrate and validate *ecosys* for assessing model performance and uncertainty.

4. Model Calibration

The *ecosys* model has been well calibrated for the major row crops and cover crops across the U.S. Midwest in previous studies (Zhou et al., 2021; Qin et al., 2021; 2023; Li et al., 2022; Liu et al., 2021; Yang et al., 2022). It is noteworthy that the model calibration efforts of crop-related parameters as referenced above are totally independent from the model validation effort documented here, as the above publications did not use any observations from the sites referenced in this validation report. Those published studies also did not use any observations of SOC change or direct N_2O emissions as constraints in the model calibration. Instead, crop carbon uptake (photosynthesis or gross primary productivity, GPP), crop yield, and their responses to environmental or management factors (such as fertilizer application rates) were used in the model calibration. The above studies found that *ecosys* can reasonably reproduce the measured surface energy balance, carbon and water fluxes, crop yield and their responses to environmental and management factors (Zhou et al., 2021). In this model validation report, we employ the row crop and cover crop parameters reported in peer-reviewed, published *ecosys* articles in journals that were pre-approved by the SEP Model Requirements (section 3.6 of the SEP Model Requirements), which includes *Agricultural and Forest Meteorology* (Zhou et al., 2021; Yang et al., 2022), *Field Crops Research* (Qin et al., 2021; Li et al., 2021), and *Global Change Biology* (Qin et al., 2023). The model parameters and processes were not calibrated for the model validation runs. Accordingly, this report relies upon independent measurements (i.e., independent from measurements used in developing model parameters) of SOC changes and N_2O emissions to validate the performance of the *ecosys* model.

The HabiTerre team acknowledges that some of the crop-related parameters can be location- or site- specific, although all crop-related parameters were held constant here. HabiTerre has been using the SYMFONI (The "**Sy**stem of Systems" Solution for Co**m**mercial **F**ield-Level Quantification of Soil **O**rganic Carbon and **Ni**trous Oxide Emission for Scalable Applications) solution, which is an advanced model-data fusion (MDF) solution developed in the DOE ARPA-E SMARTFARM project, to constrain those location-specific crop-related parameters with remote sensing observations of GPP and crop yield, which can provide even better performance for SOC changes and N_2O emissions than what is reported here from a mass-balance and systems perspective

(Guan et al., 2023). A qualitative example on applying SYMFONI to field-level greenhouse gas emissions quantification is provided in Appendix A.

5. Model simulations

5.1 Model setup

The *ecosys* model was run for all relevant publicly available study sites and measurement treatments. Besides crop-related parameters, site-specific input files include site information, management inputs, soil inputs, and atmospheric forcing. The reported site information and management information in the study publications were used as inputs to the model. For soil inputs, we used soil data reported from the study publications. When a soil parameter was not reported, we used the information from gSSURGO, which is a gridded soil dataset from USDA. For the atmospheric forcing data, the North American Land Data Assimilation System (NLDAS) data was used. Historical mean atmospheric concentrations for oxygen (21%), nitrogen (78%), carbon dioxide (405 ppm), methane (1.8 ppm), nitrous oxide (0.3 ppm), and ammonia (0.005 ppm) were used, if not measured. The tillage date, type (e.g., conservation tillage and conventional tillage), and depth were based on the reported tillage practices, if applicable. The conversion of manure type into elemental (carbon and nitrogen) percentages was based on the carbon nitrogen ratios reported in Rynk et al. (1992). All the necessary model inputs, extracted using the above-mentioned methods, are documented in Table 1. All the simulations were performed after a 20-year spin up period for model initialization (prior to the start of the experimental period), using the soil and atmospheric forcing data described above.

Table 1. The list of necessary input data required to run the *ecosys* model.

Category	Input data name
Site location	Latitude
	Longitude
Weather conditions (Hourly)	Near-Surface Air Temperature
	Precipitation
	Downwelling Solar Radiation
	Near-Surface Relative Humidity
	Wind speed

Soil and hydrological conditions by soil depths	Initial soil organic carbon
	Initial soil bulk density
	Initial soil pH
	Sand content
	Silt content
	Initial water content at field capacity
	Initial water content at wilting point
	Initial saturated hydraulic conductivity
	Initial water table depth
Atmospheric conditions (Historical mean or field measurements)	Atmospheric oxygen concentration
	Atmospheric nitrogen concentration
	Atmospheric carbon dioxide concentration
	Atmospheric methane concentration
	Atmospheric nitrous oxide concentration
	Atmospheric ammonia concentration
Management practice	Crop type
	Planting/harvesting date
	Plant density
	Plant depth
	Cover crop type
	Cover crop planting/termination date
	Cover crop plant density
	Cover crop plant depth
	Tillage date
	Tillage depth
	Tillage type (categorical) or intensity (mixing ratio)
	Fertilizer type
	Fertilizer composition (NPK percentages)
	Fertilizer rate

	Fertilizer application date
	Manure type
	Manure composition (CNP percentages)
	Manure rate
	Manure application date

5.2 Documentation of model parameters

The *ecosys* model has been extensively tested against eddy covariance fluxes and related ecophysiological measurements with a wide range of sites and weather conditions in croplands (Grant, 1995; Grant, 1997; Grant and Heaney, 1997; Grant and Pattey, 2003; Grant et al., 2006; Grant and Pattey, 2008; Grant et al., 2020; Zhou et al., 2021; Qin et al., 2021; 2022; Li et al., 2022; Yang et al., 2022), grasslands (Grant and Flanagan, 2007; Grant et al., 2012a), forests (Grant et al., 2010), and wetlands (Mezbahuddin et al., 2014). A qualitative summary of the *ecosys* model structure is provided in Grant (2001) and Grant (2013). All *ecosys* model structures and key parameters are unchanged from those described in these earlier studies.

The HabiTerre team has documented key model parameters for the SOC and N_2O emissions simulations in Appendix B. The observational database used in this model validation report, processing scripts, and the model source code necessary to reproduce the results presented here have been archived in a version controlled Github repository.

6. Project domain

6.1 Practice categories

This model validation report intends to validate the modeled SOC changes and direct N_2O emissions under 4 practice categories (PCs) defined in the SEP Model Requirements: 1) Cropping practices, planting and harvesting (CROPPING), 2) Inorganic nitrogen fertilizer application (FERTILIZER), 3) Organic amendments application (ORGANIC), and 4) Soil disturbance and/or residue management (DISTURBANCE). The full descriptions of each practice category is provided in Table 2. We did not include stacked practices in this report, and there was at least one study that isolates the effect of the practice change being validated.

Table 2. Practice categories defined by the SEP Model Requirements and examined in this model validation report

Practice Category	**Description of practice effects**	**Abbreviation**
Cropping practices, planting and harvesting	Variety of crops grown, which includes the comparison of single-crop crop rotation and double-crop rotation and the effects of planting cover crops)	CROPPING
Inorganic nitrogen fertilizer application	Magnitude, form (including enhanced efficiency fertilizers), or timing for nitrogen fertilizer applied	FERTILIZER
Organic amendments application	Magnitude, timing, and variation in C:N ratio for the manure applied	ORGANIC
Soil disturbance	Soil disturbance driven by tillage activity	DISTURBANCE

6.2 Crop functional type

This model validation report intends to validate the modeled SOC changes and direct N_2O emissions under 2 crop functional types (CFTs): 1) Corn and 2) Soybean.

Table 3. Crop functional type (CFTs) covered in this model validation report

Crop functional types (CFTs)	**Crop name**
Annual, C4, herbaceous, non-N-fixing, non-flooded crops	Corn
Annual, C3, herbaceous, N-fixing, non-flooded crops	Soybean

6.3 Land Resource Regions and Climate types

This model validation report intends to validate the modeled SOC changes and direct N_2O emissions within 8 Land Resource Regions (LRRs) spanning across 3 climate types, as detailed in Table 4 and Table 5, respectively. The climate types appearing in this model validation report follow the IPCC climate reference regions (IPCC, 2006).

Table 4. Land Resource Regions (LRRs) covered in this model validation report

Land Resource Regions (LRRs)	LRR code
Northern Great Plains Spring Wheat Region	F
Western Great Plains Range and Irrigated Region	G
Central Great Plains Winter Wheat and Range Region	H
Northern Lake States Forest and Forage Region	K
Lake State Fruit, Truck Crop, and Dairy Region	L
Central Feed Grains and Livestock Region	M
East and Central Farming and Forest Region	N
South Atlantic and Gulf Slope Cash Crops, Forest, and Livestock Region	P

Table 5. IPCC climate types (CTs) covered in this model validation report

IPCC climate types	Abbreviation
Cool temperate moist	CTM
Warm temperate moist	WTM
Cool temperate dry	CTD

6.4 Soil texture classes

This model validation report intends to validate the modeled SOC changes and direct N_2O emissions under 8 soil texture classes: 1) Clay loam, 2) Loam, 3) Loamy sand, 4) Sandy clay loam, 5) Sandy loam, 6) Silt loam, 7) Silty clay, and 8) Silty clay loam (Table 6). The soil texture classes evaluated were defined per the USDA soil texture triangle.

Table 6. Soil texture classes included in this model validation report

Soil texture class	Clay content (%)
Clay loam	30-33
Loam	20-25
Loamy sand	4
Sandy clay loam	21-30
Sandy loam	11-17
Silt loam	19-26
Silty clay	48
Silty clay loam	27-37

6.5 Emission sources

This model validation report includes two emission sources (ES): SOC changes and direct N_2O emissions. Emissions of CH_4 and indirect N_2O emissions are not included in this model validation report. The practice category, crop functional type, land resource region, climate zone, and clay content embedded in each emission source is summarized in Table 7. The duration of direct N_2O emissions measurements at each site and year is reported in Table D2. For SOC changes, the validation is based on an equivalent mass basis, i.e., comparing the modeled and measured SOC changes at the measurement depth. The duration of SOC change measurements is shown in Figure E2, which ranges from 3 to 95 years with most of the sites reporting SOC changes over 12 years.

Table 7. The combination of practice category (PC), crop functional type (CFT), land resource region (LRR), climate zone, clay content, study sites, and emission sources evaluated in this model validation report. Detailed site information for each study site is provided in Appendix D.

SOC					
PC	**CFT**	**LRR**	**Climate zone**	**Clay content (%)**	**Study site (sampling depth)**
CROPPING	Corn	F, M, N, P, K	CTM, WTM	11-30	Central (15 cm), Dixon Springs (75 cm), East Lansing (5 cm), South (15 cm), SouthDakota (25 cm), Tenessa (15 cm), Urbana (46 cm), Wisconsin (50 cm)
CROPPING	Soybean	M, N, K	CTM, WTM	11-30	Central (15 cm), Crossville (20 cm), Dixon Springs (75 cm), South (15 cm), Urbana (46 cm)
DISTURBANCE	Corn	L, M, N	CTM, WTM	19-48	Ames (60 cm), Armstrong (60 cm), Crawfordsville (60 cm), Dixon Springs (75 cm), Kanawha (60 cm), Hoytville2 (15 cm), McNay (60 cm), Nashua (60 cm), Sutherland (60 cm)
DISTURBANCE	Soybean	L, M, N, H	CTM, WTM	19-48	Ames (60 cm), Armstrong (60 cm), Crawfordsville (60 cm), Dixon Springs (75 cm), Kanawha (60 cm), Kansas (30 cm), Hoytville2 (15 cm), McNay (60 cm), Nashua (60 cm), Sutherland (60 cm)
FERTILIZER	Corn	G, M, N	CTM	19-48	Central (15 cm), Crawfordsville (60 cm), Lexngton (30 cm), Mead (30 cm), McNay (60 cm), Northwest (15 cm), Rosemount, South (15 cm), Southeast (15 cm), Ithaca (30 cm)
FERTILIZER	Soybean	M, N	CTM	20-48	Central (15 cm), Crawfordsville (60 cm), Fort Collins (30 cm), Mead (30

					cm), McNay (60 cm), Northwest (15 cm), South (15 cm), Southeast (15 cm)
ORGANIC	Corn	M	CTM	22-30	Minnesota (30 cm), Urbana (15 cm)
		N_2O			
PC	CFT	LRR	Climate zone	Clay content (%)	Study site
CROPPING	Corn	F, H, K, M	CTM	20-31	KS1, IN1, IA1, SD1, WI1
CROPPING	Soybean	F, H, K, M, P	CTM	20-31	AL1, KS1, IN1, IA1, SD1, WI1
DISTURBANCE	Corn	G, M	CTM, CTD	20-25	CO1, IA1
DISTURBANCE	Soybean	M, L	CTM	4-20	IA1, MI1
FERTILIZER	Corn	G, M, N	CTM, CTD	17-25	CO1, CO3, CO4, KY1, IA2, MN2
ORGANIC	Corn	G, M, N	CTD	23	CO4, IN1, KY1

7. Description of validation data collection process and final validation dataset

7.1 Description of validation data collection process

Following SEP Model Requirements (Section 3.3 Requirement 1, p12), we compiled a database of measured long-term SOC changes (measured using dry combustion) and direct N_2O emissions (measured using chambers) through a comprehensive search of the published literature. For SOC, we selected studies using the following criteria:

- SOC measurements should be taken for at least two separate years spanning a total interval of at least five years.
- There is sufficient reported information related to site location, soil, agronomic management and experimental treatments in the study to ensure the model simulation can be properly set up.

For direct N_2O emissions, we selected studies using the following criteria:

- Annual or seasonal accumulated direct N_2O emissions are reported in the study or there are sufficient N_2O measurements reported in the study such that annual or seasonal accumulated N_2O values can be derived using a standard method (see Appendix C).
- The sampling period for direct N_2O emissions should cover at least 4 months in each growing season or year.
- There is sufficient reported information related to site location, soil, agronomic management and experimental treatments in the study to ensure the model simulation can be properly set up.

7.2 Procedures to handle missing data

The external water table depth (EWTD) is one of the key location-specific inputs for N_2O simulations conducted by the *ecosys* model. In this model validation effort, as the EWTD is not recorded or monitored at most experimental sites, we infer site-specific EWTD with multi-year soil moisture observations. Specifically, we use monthly soil moisture or seasonal mean soil moisture as constraints to estimate the EWTD as site-specific inputs. We then used this inferred site-specific EWTD for all the treatments over a particular site to quantify prediction errors driven by the model engine evaluated in this model validation report.

We prepared SOC and N_2O validation datasets based on publicly available SOC and N_2O measurements. We did not include stacked practices. The Pooled Measurement Uncertainty (PMU, see more details in section 9.1 of this report), a key factor determining the model validation outcomes, was calculated at sites where measurement errors were reported following the procedures defined in the SEP Model Requirements.

7.3 Description of validation dataset

Following the procedure and criteria described above, we evaluated the modeled practice-change effects on SOC changes and direct N_2O emissions based on the methods described in SEP Model Requirements. The analyses for SOC changes and direct N_2O emissions were presented in the unit of Ton $CO_{2,eq}$ per acre per year and Ton $CO_{2,eq}$ per acre, respectively, to reconcile the different measurement periods and lengths presented at different sites. A global warming potential over a 100-year time horizon of 273 (IPCC, 2013) was used for N_2O to convert the emissions into the CO_2 equivalent units in this model validation report, which is consistent with US EPA's Inventory of U.S. Greenhouse Gas Emissions and Sinks and complies with international GHG reporting standards under the United Nations Framework Convention on Climate Change (UNFCCC).

We evaluated field-level model performance by averaging SOC and N_2O measurements replicates collected from the same field under the same treatment. The SOC dataset includes 269 pairs of observed practice-change effects for the corn and soybean crops examined in this report (Figure 3). The soil sampling depths range from 5 to 75 cm across the SOC dataset. Treatment-pairs for SOC change were required to share a common soil sampling depth and length between sampling dates over each site. The N_2O dataset includes 175 pairs of observed practice-change effects for the corn and soybean crops examined in this report (Figure 4). Among the SOC and N_2O datasets, eight sites have measurement error reported for SOC and six sites have measurement error reported for N_2O for corn and soybean (Appendix D). Treatment-pairs for N_2O emissions were required to share a common sampling coverage length over each site.

We note that 70% of the SOC measurements used in the validation report are collected at 30 cm or deeper soil depth (62% are deeper than 60 cm). Though some SOC measurements collected at shallower soil depths (<30cm) are used to improve data availability and thereby extend the validity of this model validation report, those shallower SOC measurements do not appear to overestimate SOC changes as compared to SOC measurements collected at deeper soil depths (>30 cm) (Appendix E).

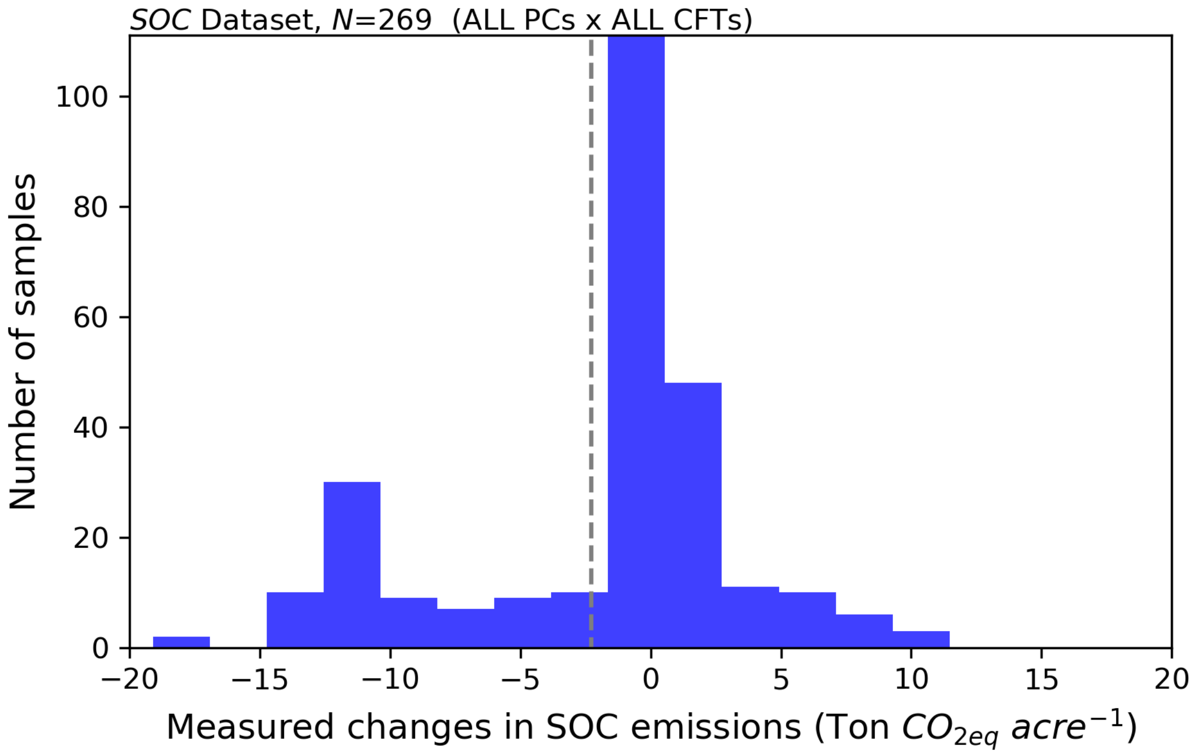

Figure 3. The histogram of measured changes in SOC emissions used for model validation, including all practice categories (PCs) and all crop functional types (CFTs) evaluated in this report. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

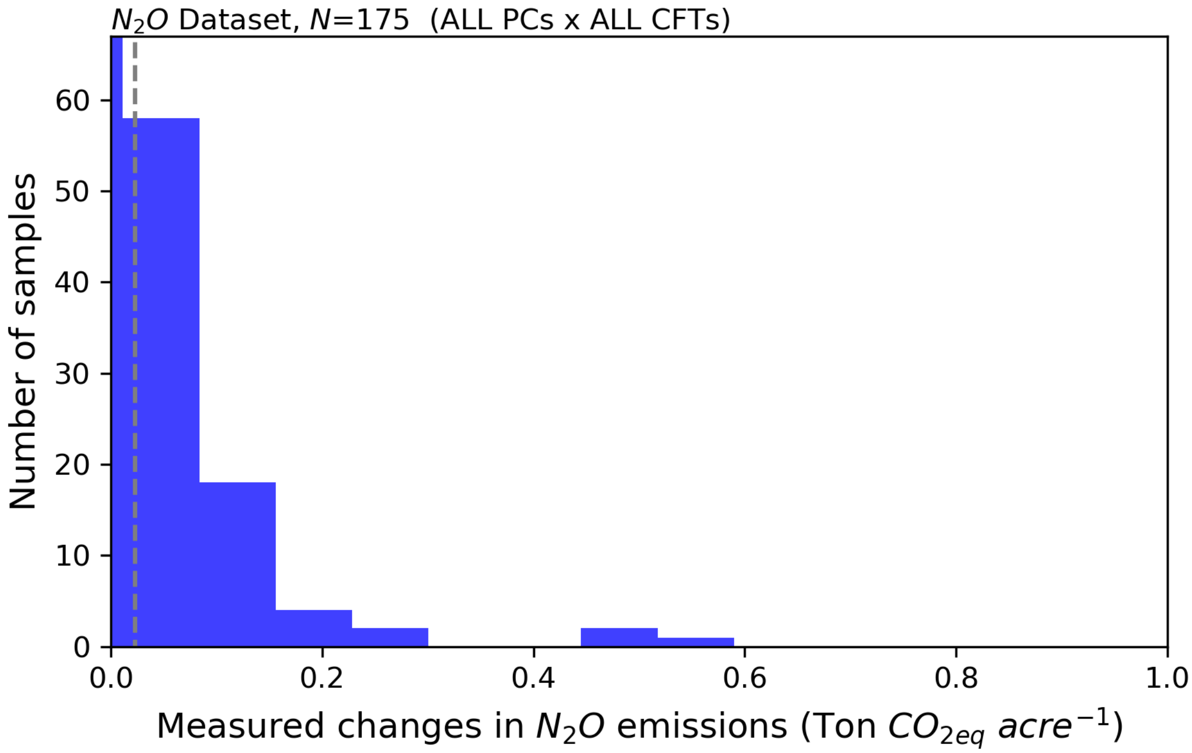

Figure 4. The histogram of measured changes in direct N_2O emissions used for model validation, including all practice categories (PCs) and all crop functional types (CFTs) evaluated in this report. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

8. Overall model performance

We evaluated the overall performance of the modeled practice change effects on SOC changes and direct N_2O emissions by comparing the simulations against observations collected from publicly available dataset (Figure 5 and Figure 6). The R^2 values were found to be 0.58 and 0.45 for the modeled practice change effects on changes in SOC and direct N_2O emissions, respectively.

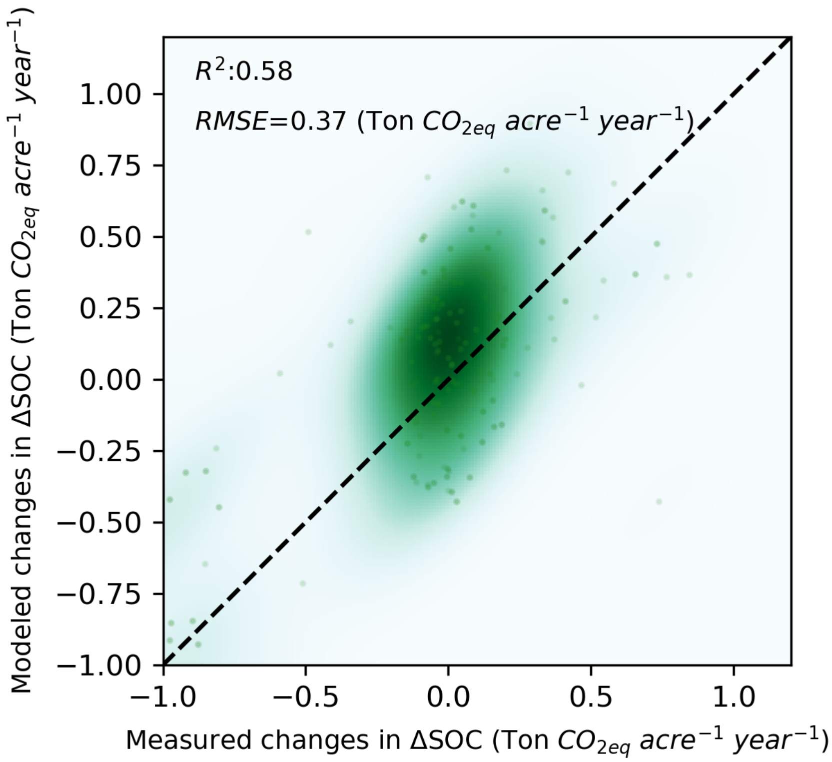

Figure 5. Overall performance of the modeled changes in SOC emissions in response to practice changes across the SOC Dataset evaluated in this report. Darker color represents denser data points. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

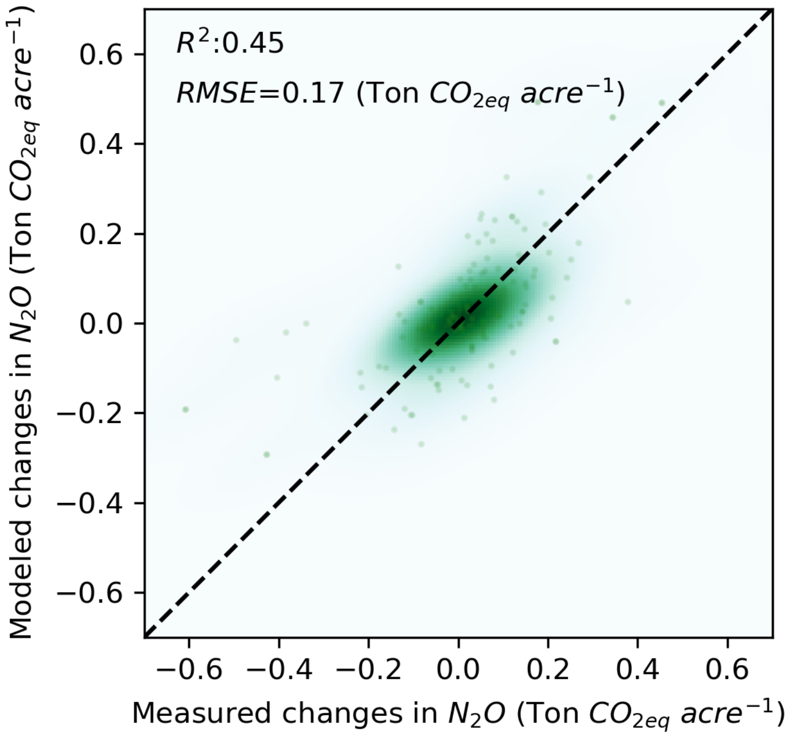

Figure 6. Overall performance of the modeled changes in direct N_2O emissions in response to practice changes across the N_2O Dataset evaluated in this report. Darker color represents denser data points. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

9. Model bias evaluation

9.1 Calculation of bias and PMU

We calculated bias and Pool Measurement Uncertainty (PMU) for each PC/CFG/ES combination. Specifically, the bias was calculated as the mean difference between modeled and observed practice-change effects:

$$bias = \frac{1}{n}\sum_{i=1}^{n} \quad (modeled_i - observed_i) \qquad (1)$$

Where $modeled_i$ and $observed_i$ represent the ecosys-simulated and measured practice-change effects on SOC changes and direct N_2O emissions, and n is the number of treatment pairs.

The PMU is calculated as:

$$PMU = \sqrt{\frac{\sum_{i=1}^{k} \quad \sigma_i^2(n_{i1}+n_{i2}-2)}{\sum_{i=1}^{k} \quad (n_{i1}+n_{i2}-2)}} \qquad (2)$$

Where k is the total number of treatment pairs with reported measurement uncertainty information; σ_i is the standard error of the i^{th} measurement of practice-change effects on SOC changes or direct N_2O emissions; n_{i1} and n_{i2} are the number of replicates in the two different treatments of a particular treatment pair; and $n_{i1} + n_{i2} - 2$ is the degrees of freedom of σ_i^2.

9.2 An example of PMU calculation

Table 8. An example of calculating the pooled measurement uncertainty (PMU) based on available estimates of measurement uncertainty from the FERTILIZER x Corn SOC validation dataset.

site	Number of samples in treatment1 (n1)	Number of samples in treatment2 (n2)	Standard error in treatment1 (σ_1)	Standard error in treatment2 (σ_2)	Degree of freedom (df)	Standard error in the measurement pair (σ_1^2+σ_2^2)	(σ_1^2+σ_2^2) * df
12	4	4	4.48	4.72	6	42.39	254.33
12	4	4	4.48	3.80	6	34.54	207.25
12	4	4	4.48	6.07	6	56.96	341.75

12	4	4	4.48	5.30	6	48.20	289.23
sum	NA	NA	NA	NA	24	182.09	1092.56
PMU = $\sqrt{\frac{\sum (\sigma_1^2+\sigma_2^2)*df}{\sum df}}$	NA	NA	NA	NA	NA	NA	6.00 (MgC/ha) =8.90 (Ton $CO_{2,\,eq}$ per acre)

9.3 Comparison between bias and PMU

We followed the protocol defined in the SEP Model Requirements to assess the performance of our simulated practice change effects on SOC changes and direct N_2O emissions, except for one deviation. The bias and PMU were calculated and reported against an emissions source specific domain (i.e., one for SOC and one for N2O emissions) comprising multiple land resource regions, crop functional types, soil texture classes and practice changes to improve the representativeness of validation data collected from public datasets. This deviation was necessary as the granular requirements for model validation combinations defined in the SEP Model Requirements are untenable for a generalized model validation report, which limits the spatial coverage and practice change can be validated with publicly available datasets. This deviation has been approved by CAR to ensure the best use of public data for model validation (Appendix F). Our evaluation below (Table 9) demonstrates that the mean bias is less than the estimated PMU for both SOC and N2O emissions, which fulfills the SEP Model Requirements.

Table 9. PMU and Bias for SOC and N_2O for different practice categories and crop functional types

SOC				
Practice categories (PCs)	**Crop functional types (CFTs)**	**PMU (Ton $CO_{2,\,eq}$ acre^{-1} year^{-1})**	**Bias (Ton $CO_{2,\,eq}$ acre^{-1} year^{-1})**	**Is Bias smaller than PMU?**
All PCs	All CFTs	0.60	0.09	Yes
CROPPING	All CFTs	-	0.20	Yes
DISTURBANCE	All CFTs	-	0.05	Yes
FERTILIZER	All CFTs	-	0.08	Yes

ORGANIC	All CFTs	-	-0.16	Yes
All PCs	Corn	-	0.08	Yes
All PCs	Soybean	-	0.08	Yes
N_2O				
Practice categories (PCs)	**Crop functional types (CFTs)**	**PMU (Ton $CO_{2, eq}$ acre^{-1})**	**Bias (Ton $CO_{2, eq}$ acre^{-1})**	**Is Bias smaller than PMU?**
All PCs	All CFTs	0.13	0.01	Yes
CROPPING	All CFTs	-	0.06	Yes
DISTURBANCE	All CFTs	-	-0.002	Yes
FERTILIZER	All CFTs	-	-0.01	Yes
ORGANIC	All CFTs	-	-0.24	No*
All PCs	Corn	-	0.01	Yes
All PCs	Soybean	-	-0.001	Yes
*Represent 4% of the data points evaluated in this model validation report.				

9.4 Bias for each study across the PC, CFG, and ES combinations

We followed the protocol defined in the SEP Model Requirements to report the bias of our simulations of practice change effects on SOC change (Table 10) and direct N_2O emissions (Table 11) at individual study sites. The study sites below are reported and ranked according to protocol requirements from highest to lowest bias values calculated from all practice categories and crop functional types.

Table 10. Study site specific bias for simulated practice change effects on SOC change from each study site across different practice categories and crop functional types.

Site	Measured change (Ton $CO_{2,eq}$ acre^{-1} year^{-1})	Modeled change (Ton $CO_{2,eq}$ acre^{-1} year^{-1})	Bias (Ton $CO_{2,eq}$ acre^{-1} year^{-1})
Wisconsin	-1.38	1.102	2.482
Ithaca	-0.417	0.361	0.778
Crossville	0.203	0.733	0.53
Tenessa	0.125	0.531	0.407

Southeast	-0.075	0.247	0.321
South	0.084	0.3	0.216
Nashua	-0.352	-0.204	0.148
Crawfordsville	-0.399	-0.266	0.133
Rosemount	-1.178	-1.054	0.125
Armstrong	-0.381	-0.303	0.078
McNay	-0.376	-0.335	0.04
Sutherland	-0.311	-0.286	0.025
Lexington	0.155	0.18	0.025
Central	0.353	0.345	-0.008
East Lansing	0.011	0.001	-0.01
Ames	-0.37	-0.382	-0.012
Northwest	0.12	0.099	-0.022
Kanawha	-0.313	-0.336	-0.023
Mead	0.14	0.115	-0.025
Dixon Springs	0.354	0.328	-0.026
Urbana2	-0.115	-0.148	-0.033
Urbana	-0.065	-0.106	-0.041
SouthDakota	-0.179	-0.357	-0.178
Kansas	0.24	0.019	-0.221
Hoytville2	0.185	-0.158	-0.344
Minnesota	0	-0.387	-0.387
Fort Collins	0.737	-0.427	-1.164

Table 11. The bias of simulated practice change effects on direct N_2O emissions calculated from each study site across different practice categories and crop functional types.

Site	Measured change (Ton $CO_{2,eq}$ acre^{-1})	Modeled change (Ton $CO_{2,eq}$ acre^{-1})	Bias (Ton $CO_{2,eq}$ acre^{-1})
MN2	-0.546	-0.11	0.436
WI1	-0.854	-0.583	0.271

KY1	0.555	0.633	0.079
CO1	0.14	0.206	0.066
SD1	-0.048	0.011	0.059
CO4	0.032	0.086	0.054
KS1	0.043	0.048	0.004
AL1	-0.006	-0.002	0.004
MI1	0.027	-0.034	-0.061
CO3	0.018	-0.063	-0.081
IA1	0.02	-0.074	-0.094
IN1	0.038	-0.151	-0.189
IA2	1.924	0.781	-1.142

10. Model prediction error

We followed the analytical error propagation (SEP Appendix D.1) method to quantify the model prediction error. Specifically, we calculated the mean squared error (MSE) of validation simulations for individual practice category and crop functional type combinations:

$$MSE = \frac{\sum_{i=1}^{k} (modeled_i - observed_i)^2}{k} \quad (3)$$

We then followed the protocol defined in the SEP Model Requirements (SEP Section 3.5) to compute the confidence coverage rate (CCR) of the 90% prediction intervals against the model validation datasets for both SOC and N2O (Section 7.3). The 90% prediction intervals were defined as the predicted effects plus and minus 1.64 times the standard deviation of the calculated bias (SEP Section 3.5). The protocol requires a minimum confidence coverage rate of 90% for 90% prediction intervals (i.e., the 90% prediction intervals should contain the measured value for at least 90% of the validation data); however, it is recognized that there may be circumstances where model uncertainty bounds are appropriately set even if 90 % conference coverage is not achieved. Besides MSE and CCR, we included the scatterplot of the model predictions versus measurements and histogram of residuals (the differences between predictions and measurements) in this report, as requested by the SEP Model Requirements. Besides the overall model prediction error below, practice change and crop functional type specific model prediction errors are provided in Appendix G.

10.1 Model prediction error for practice change effects on SOC changes across all practice categories and all crops

The mean square error is calculated as 0.140 (Ton $CO_{2,\,eq}$ $acre^{-1}$ $year^{-1})^2$ for the modeled practice change effects on SOC changes effects across all practice categories and all crops. The confidence coverage rate is 94.8%, which is within the 90% prediction intervals recommended in the SEP Model Requirements.

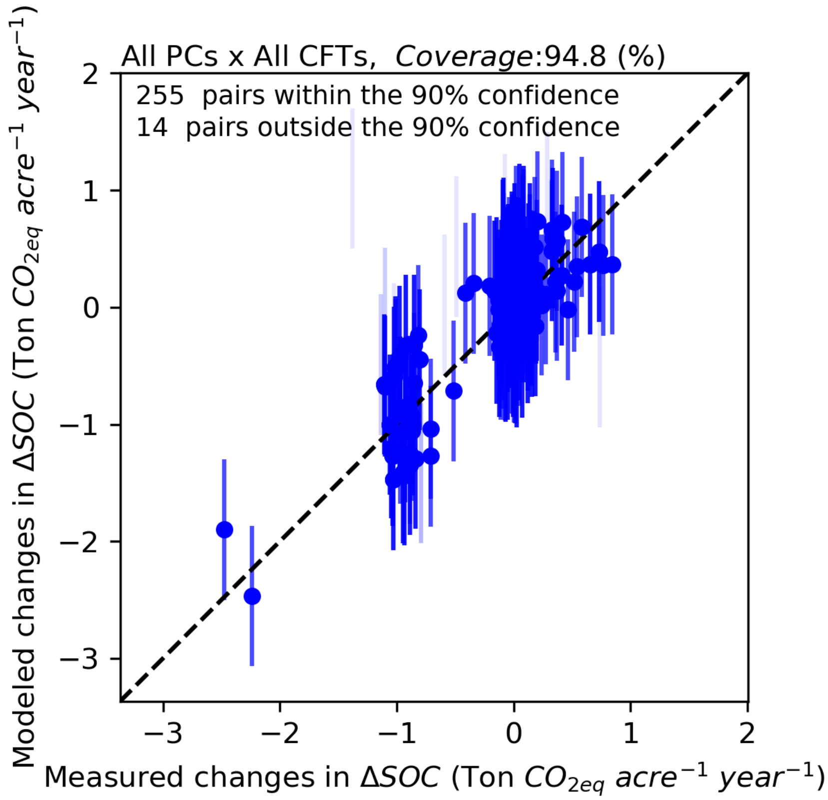

Figure 7. Modeled versus measured practice change effects on SOC changes in response to practice changes across all practice categories and all crop functional types. The treatment pairs were generated by averaging replicates within the same fields under the same treatment. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements.

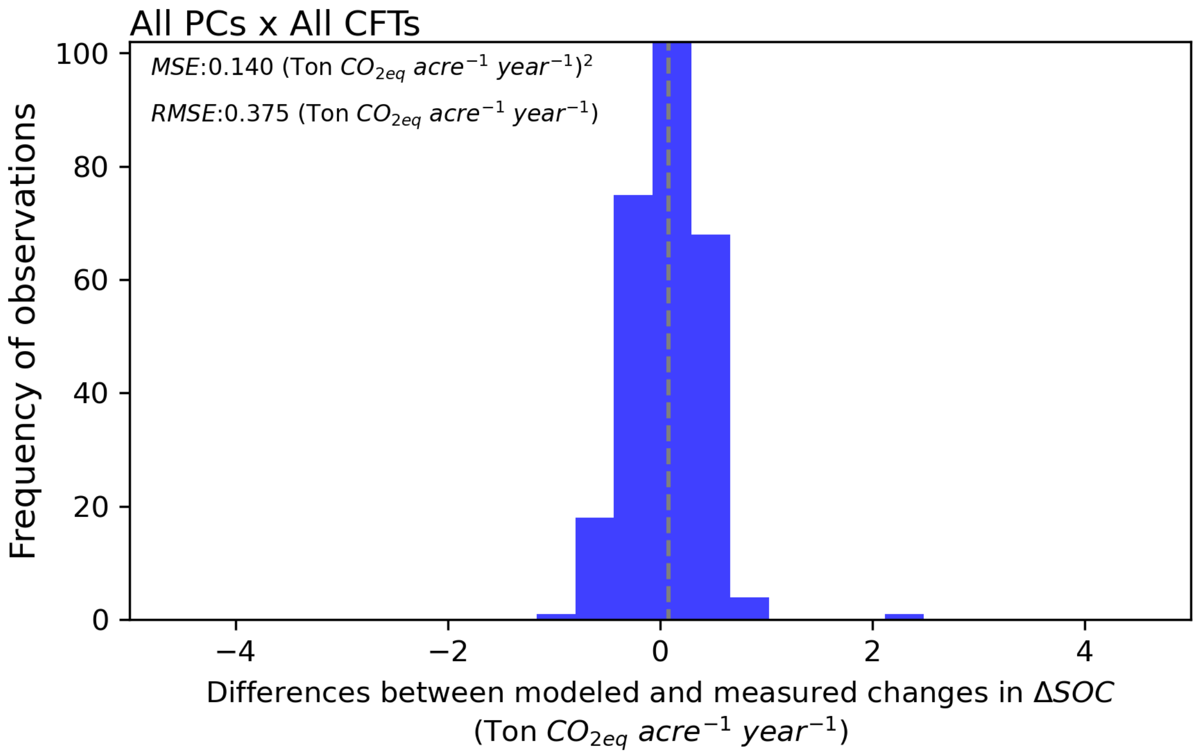

Figure 8. Histogram of the differences between modeled and measured practice change effects in SOC changes in response to practice change across all practice categories and all crop functional types. The treatment pairs were generated by averaging replicates within the same fields under the same treatment. The dashed line represents the mean value among the validation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively.

10.2 Model prediction error for practice change effects on direct N_2O emissions across all practice categories and all crops

The mean square error is calculated as 0.030 (Ton $CO_{2,\,eq}$ acre^{-1}) for the modeled practice change effects on direct N_2O emissions effects across all practice categories and all crops. The confidence coverage rate is 93.1%, which is within the 90% prediction intervals recommended in the SEP Model Requirements.

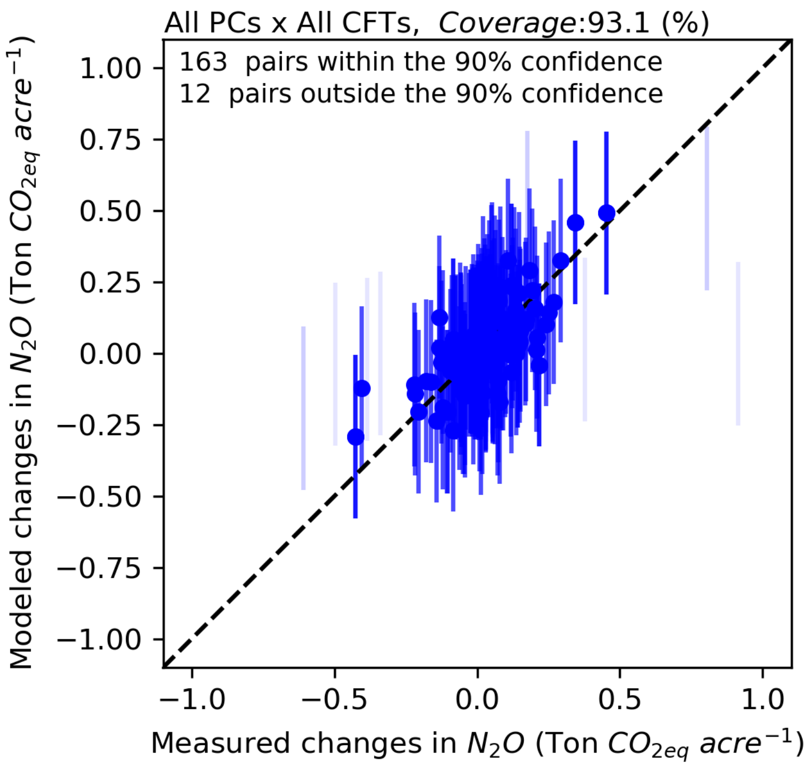

Figure 9. Modeled versus measured practice change effects on direct N_2O emissions in response to practice changes across all practice categories and all crop functional types. The treatment pairs were generated by averaging replicates within the same fields under the same treatment. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements.

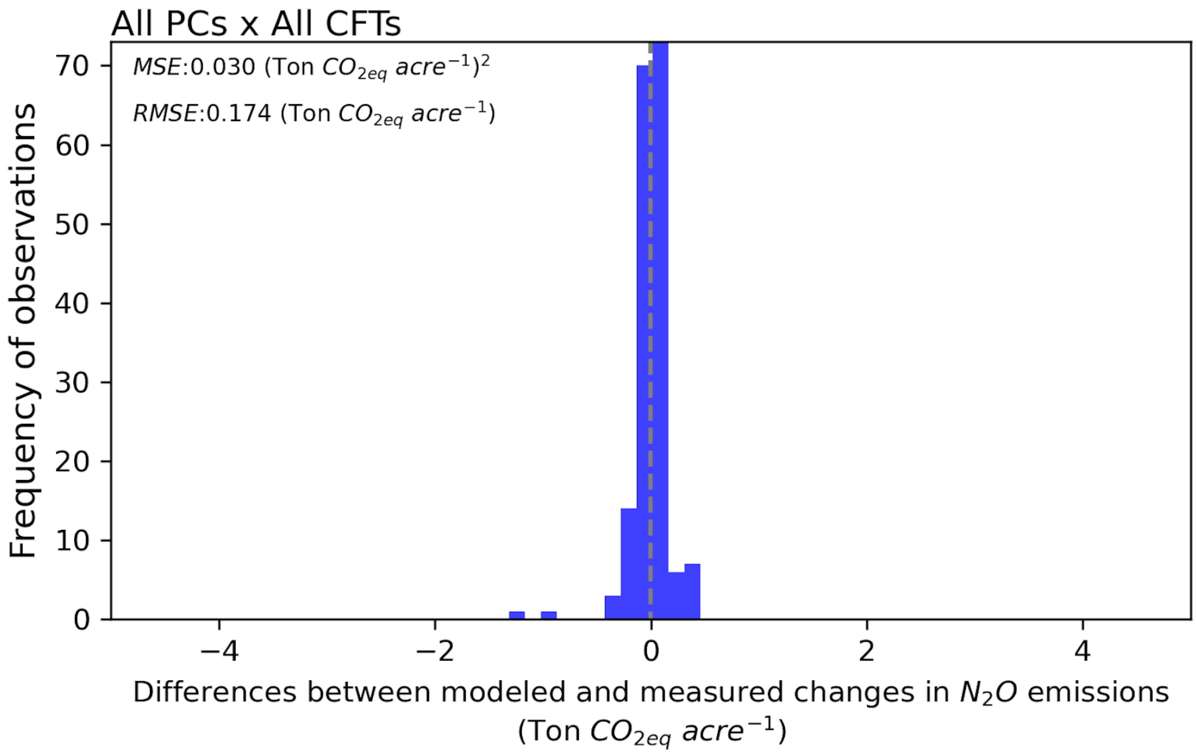

Figure 10. Histogram of the differences between modeled and measured practice change effects on direct N_2O emissions in response to practice changes across all practice categories and all crop functional types. The treatment pairs were generated by averaging replicates within the same fields under the same treatment. The dashed line represents the mean value among the validation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively.

References

Grant, R.F. (1995). Dynamics of energy, water, carbon and nitrogen in agricultural ecosystems: simulation and experimental validation. Ecological Modelling, 81, 169-181, https://doi.org/10.1016/0304-3800(94)00169-I

Grant, R.F. (1997). Changes in Soil Organic Matter under Different Tillage and Rotation: Mathematical Modeling in ecosys. Soil Science Society of America Journal, 61, 1159-1175, https://doi.org/10.2136/sssaj1997.03615995006100040023x

Grant, R. (2001). A Review of the Canadian Ecosystem Model-ecosys. Modeling Carbon and Nitrogen Dynamics for Soil Management: CRC Press

Grant, R. F. (2013). Modelling changes in nitrogen cycling to sustain increases in forest productivity under elevated atmospheric CO2 and contrasting site conditions, Biogeosciences, 10, 7703–7721, https://doi.org/10.5194/bg-10-7703-2013

Grant, R., Arkebauer, T., Dobermann, A., Hubbard, K., Schimelfenig, T., Suyker, A., Verma, S., & Walters, D. (2007). Net biome productivity of irrigated and rainfed maize–soybean rotations: modeling vs. measurements. Agronomy Journal, 99, 1404-1423, https://doi.org/10.2134/agronj2006.0308

Grant, R. F., Baldocchi, D. D., and Ma, S. (2012). Ecological controls on net ecosystem productivity of a seasonally dry annual grassland under current and future climates: Modelling with ecosys, Agricultural and Forest Meteorology, 152, 189-200， https://doi.org/10.1016/j.agrformet.2011.09.012

Grant, R. F., Barr, A. G., Black, T. A., Margolis, H. A., Mccaughey, J. H., and Trofymow, J. A. (2010). Net ecosystem productivity of temperate and boreal forests after clearcutting – a Fluxnet-Canada measurement and modelling synthesis, Tellus B: Chemical and Physical Meteorology， 62(5), 475-496, https://doi.org/10.1111/j.1600-0889.2010.00500.x

Grant, R., Dyck, M., & Puurveen, D. (2020). Nitrogen and phosphorus control carbon sequestration in agricultural ecosystems: modelling carbon, nitrogen, and phosphorus balances at the Breton Plots with ecosys under historical and future climates. Canadian Journal of Soil Science, 1-22, https://doi.org/10.1139/cjss-2019-0132

Grant R. F. and Flanagan L. B. (2007). Modeling stomatal and nonstomatal effects of water deficits on CO2 fixation in a semiarid grassland, J. Geophys. Res.. Grant, R.F., & Heaney, D.J. (1997). Inorganic Phosphorus Transformation and Transport in Soils:

Mathematical Modeling in ecosys. Soil Science Society of America Journal, 61, 752-764, https://doi.org/10.2136/sssaj1997.03615995006100030008x

Grant, R., & Pattey, E. (2003). Modelling variability in N2O emissions from fertilized agricultural fields. Soil Biology and Biochemistry, 35, 225-243, https://doi.org/10.1016/S0038-0717(02)00256-0

Grant, R., & Pattey, E. (2008). Temperature sensitivity of N2O emissions from fertilized agricultural soils: Mathematical modeling in ecosys. Global Biogeochemical Cycles, 22, GB4019, https://doi.org/10.1029/2008GB003273

Grant, R., Pattey, E., Goddard, T., Kryzanowski, L., & Puurveen, H. (2006). Modeling the effects of fertilizer application rate on nitrous oxide emissions. Soil Science Society of America Journal, 70, 235-248, https://doi.org/10.2136/sssaj2005.0104

Guan, K., Jin, Z., Peng, B., Tang, J., DeLucia, E., West, P., Jiang, C., Wang, S., Kim, T., Zhou, W., Griffis, T., Liu, L., Yang, W., Qin, Z., Yang, Q., Margenot, A., Stuchiner, E., Kumar, V., Bernacchi, C., Coppess, J., Novick, K., Gerber, J., Jahn, M., Khanna, M., Lee, D., Chen, Z., Yang, S. (2023). A scalable framework for quantifying field-level agricultural carbon outcomes. Earth-Science Reviews. Available online 29 May 2023, 104462, https://doi.org/10.1016/j.earscirev.2023.104462

IPCC (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds). Published: IGES, Japan.

IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Li, Z., Guan, K., Zhou, W., Peng, B., Jin, Z., Tang, J., Grant, R., Nafziger, E., Margenot, A., Gentry, L., DeLucia, E., Yang, W., Cai, Y., Qin, Z., Archontoulis, S., Fernández, F., Yu, Z., Lee, D.K., Yang, Y. (2022). Assessing the impacts of pre-growing-season weather conditions on soil nitrogen dynamics and corn productivity in the U.S. Midwest. Field Crops Research, 284, 108563, https://doi.org/10.1016/j.fcr.2022.108563

Liu, L., Xu, S., Jin, Z., Tang, J., Guan, K., Griffis, T., Erickson, M., Frie, A., Jia, X., Kim, T., Miller, L., Peng, B., Wu, S., Yang, Y., Zhou, W., Kumar, V. (2022). KGML-ag: A Modeling Framework of Knowledge-Guided Machine Learning to Simulate Agroecosystems: A Case Study of Estimating N2O Emission using Data from Mesocosm

Experiments. Geoscientific Model Development, 15, 2839–2858, https://doi.org/10.5194/gmd-15-2839-2022

Mezbahuddin, M., Grant, R. F., and Hirano, T. (2014). Modelling effects of seasonal variation in water table depth on net ecosystem CO2 exchange of a tropical peatland, Biogeosciences, 11, 577–599, https://doi.org/10.5194/bg-11-577-2014.

Qin, Z., Guan, K., Zhou, W., Peng, B., Villamil, M., Jin, Z., Tang, J., Grant, R., Gentry, L., Margenot, A., Bollero, G., Li, Z. (2021). Assessing the impacts of cover crops on maize and soybean yield in the U.S. Midwestern agroecosystems. Field Crops Research, 273, 108264, https://doi.org/10.1016/j.fcr.2021.108264

Qin, Z., Guan, K., Zhou, W., Peng, B., Tang, J., Jin, Z., Grant, R., Hu, T., Villamil, M.B., DeLucia, E., Margenot, A.J., Mishra, U., Chen, Z., & Coppess, J. (2023). Assessing long-term impacts of cover crops on soil organic carbon in the central U.S. Midwestern agroecosystems. Global Change Biology, 29(9), 2572-2590, https://doi.org/10.1111/gcb.16632

Rynk, R., M. van de Kamp, G.B. Willson, M.E. Singley, T.L. Richard, J.J. Kolega, F.R. Gouin, L. Laliberty Jr., D. Kay, D.W. Murphy, H.A.J. Hoitink, and W.F. Brinton. (1992). On-Farm composting handbook. (ed). R. Rynk. p.6-13, 106-113. Northeast Regional Agricultural Engineering Service, Ithaca, N.Y.

Smith, C., Z.R.J. Nicholls, K. Armour, W. Collins, P. Forster, M. Meinshausen, M.D. Palmer, and M. Watanabe (2021). The Earth’s Energy Budget, Climate Feedbacks, and Climate Sensitivity Supplementary Material. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Available from https://www.ipcc.ch/.

Yang, Y., Liu, L., Zhou, W., Guan, K., Tang, J., Kim, T., Grant, R., Peng, B., Zhu, P., Li, Z., Griffis, T., Jin, Z. (2022). Distinct driving mechanisms of non-growing season N2O emissions call for spatial-specific mitigation strategies in the US Midwest. Agricultural and Forest Meteorology, 324, 109108, https://doi.org/10.1016/j.agrformet.2022.109108

Zhou, W., Guan, K., Peng, B., Tang, J., Jin, Z., Jiang, C., Grant, R., Mezbahuddin, S. (2021). Quantifying carbon budget, crop yields and their responses to environmental variability using the ecosys model for U.S. Midwestern agroecosystems. Agricultural and Forest Meteorology, 307, 108521, https://doi.org/10.1016/j.agrformet.2021.108521

Appendix A: Carbon flux modeling with HabiTerre's advanced model-data fusion framework

*A.1 The "**Sy**stem of Systems" Solutions for Co**m**mercial **F**ield-Level Quantification of Soil **O**rganic Carbon and **Ni**trous Oxide Emission for Scalable Applications (SYMFONI)*

HabiTerre's Carbon Solution, i.e. SYMFONI, is an advanced model-data fusion (MDF) framework to quantify the carbon footprints and carbon outcomes of farming (Guan et al., 2023). This solution uses the novel "System-of-Systems" concept, meaning a single sensor or a model alone could not solve the complex agroecosystem problems, and only by effectively integrating various approaches (e.g. diverse observations, sensor/in-situ data, modeling) a viable solution is made possible. There are three major components in HabiTerre's carbon solution: model engine, observation suite for Crop (C), Management (M), and Environmental (E) conditions, and model-data fusion (MDF) module (Figure A1).

- Model engine: HabiTerre's carbon solution uses an internal version of an advanced process-based biogeochemistry model - *ecosys* (Grant, 2001) as its model engine. For more detailed description of the ecosys model, please refer to section 2.1.
- C-M-E observation suite: The C-M-E observation suite in HabiTerre's carbon solution provides capability to track the daily changes at each individual farmland parcel by using advanced sensing (ground, airborne hyperspectral imaging, and satellite remote sensing), including crop conditions (crop type, crop variety traits, phenology, maturity groups, photosynthetic capacity, crop carbon uptake, crop water use, and yield), management practices (planting/harvesting date, tillage practices, intercropping, crop rotation, cover cropping, fertilizer/pesticide applications), and environment conditions (weather information, soil temperature and moisture).
- Model-data fusion module: All the above observations will be digested and integrated with the model engine in the model-data fusion (MDF) module of HabiTerre's carbon solution (Figure A1). Some observations are directly used as model inputs (e.g. crop type, management practices, and weather information), while others are used as model constraints to reduce the model uncertainties. The MDF module uses advanced AI techniques to train deep neural network models as surrogates of Ecosys and GPU-based optimization to speed up the MDF such that it can be scaled up to millions of fields in a highly cost-efficient manner.

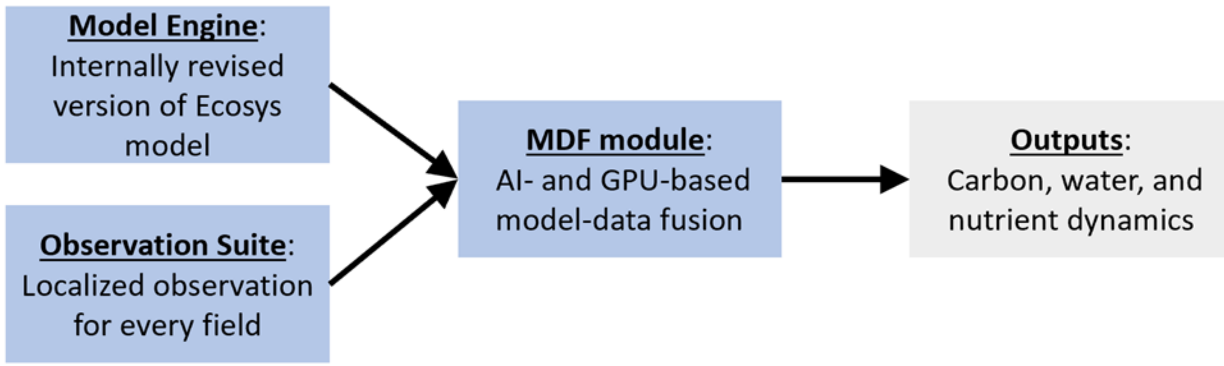

Figure A1. The overall model-data fusion (MDF) framework of HabiTerre's carbon solution.

In SYMFONI, changes in the carbon stock of a system are equal to the sum of the total carbon inflows and outflows of the system, as detailed above as a mass balance equation (Figure 2b). Therefore, validating the changes in carbon stock can be achieved by validating the carbon inflows and outflows. In a cropland system, gross primary productivity (GPP) and crop yield are the two most important carbon flows. GPP, i.e., photosynthesis, is the main source of carbon input into the system (~100% for systems with no addition of organic matter), while crop yield is a major source (~20%-30%) of carbon moving out of the system. The largest source of carbon moving out of the system is plant respiration (Ra in Figure 2b, ~50% of carbon efflux), which is closely related to GPP. Therefore, accurate measurement of crop yield and GPP at the field level enables more accurate estimation of SOC changes and N_2O emissions because:

1) Verifying GPP and crop yield using the mass balance approach validates approximately 100% of the carbon influx and 75% of carbon efflux in total;
2) The difference between GPP and crop yield upper bounds the potential changes in carbon stock;
3) Crop yield and GPP data are used as strong constraints to the carbon cycle simulation in HabiTerre's MDF solution, which helps to constrain the remaining 25% of the carbon efflux (i.e. Rh in Figure 2b) and thus improves the overall accuracy of the "measure + model approach" for estimation of SOC changes;
4) As the nitrogen cycle and carbon cycle are intrinsically linked together (Figure 2a), better constraining the carbon cycle through MDF will also lead to better estimation of the nitrogen cycle, including N_2O emissions.

According to SEP Model Requirements, model validation requires direct observations of SOC change and GHG emissions, which are usually difficult to obtain. In addition, for most of the literature, SOC changes and N_2O emissions were measured at small agronomic trial plots and thus cannot represent the field or landscape scale. We strongly suggest that validating a model through comparing the simulated carbon budget with observations (such as those from eddy-covariance towers at the landscape scale) should be added as one of the options in the SEP Model Requirements. In this version of HabiTerre's model validation report, advanced MDF-based calibration was not conducted at the validation sites because: (1) the current SEP Model Requirements recommend use of direct measurements of SOC changes and N_2O emissions from small trial plots, which greatly limits the feasibility of conducting MDF with remote sensing observations; and (2) the validation data periods for SOC changes and N_2O emissions largely do not overlap with periods with available remote-sensing-based GPP and crop yield estimations.

A.2 Field-level carbon outcome quantification with SYMFONI

We evaluated the performance of carbon flux modeling with and without implementing HabiTerre's advanced model-data fusion module at three agricultural sites in the U.S. Midwest using eddy covariance measurements from the AmeriFlux network (https://ameriflux.lbl.gov/). The three sites (US-Ne1, US-Ne2, and US-Ne3) have reported the necessary measurements to perform model-data fusion, and they cover the continuous corn and corn-soybean rotation systems examined in this model validation report.

The hourly gap-filled meteorological variables from AmeriFlux, soil information from the Gridded Soil Survey Geographic Database (gSSURGO), and management practices from site records were used to drive the *ecosys* model. We performed two sets of simulations at each site: (1) MDF run with HabiTerre's advanced model-data fusion module to optimize model parameters based on constraints inferred from gross primary productivity (GPP) observations, and (2) default run with default corn and soybean parameters. Both sets of simulations were evaluated by the daily gap-filled net ecosystem exchange (NEE) observations from 2001 to 2012, which presents the mass balance of carbon uptake and respiration at the field-level scale.

Our results show that using parameters optimized by HabiTerre's advanced model-data fusion framework improves the model performance of the corn and soybean systems simulated in the U.S. Midwest (Figure A1). The higher correlation coefficient (*r*) and slope values suggest that observational constraints, either from in situ measurements and/or remote sensing techniques, can be used to inform location specific parameters to reduce the modeling uncertainties in field-level greenhouse gas emissions quantification. Importantly, HabiTerre's SYMFONI solution (Guan et al., 2023) has integrated a range of observational constraints into a machine-learning based algorithm to optimize field-level greenhouse gas emissions quantification for agroecosystems, which has been validated by extensive ground truth data including eddy covariance flux measurements, long-term soil organic carbon observations, water quantity and quality observations. We note here that the parameter optimization process conducted by HabiTerre's model-data fusion module is an objective data driven algorithm without human intervention, providing auditability for field-level greenhouse gas emissions quantification.

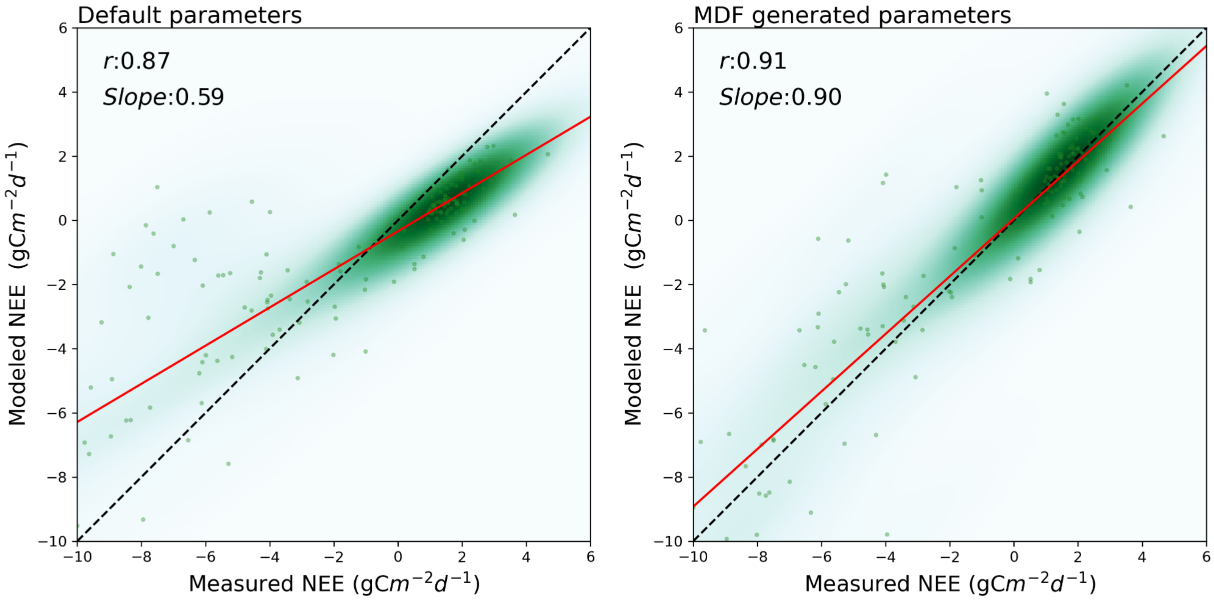

Figure A2. Overall performance of the net ecosystem exchange (NEE) modeled at the U.S Midwest corn and soybean systems. Red and black dashed represent linear best fit and one-to-one lines, respectively. Darker color represents denser data points. The simulations were conducted by two sets of parameters: (1) default model parameters; (2) MDF generated parameters based on parameter optimization performed by HabiTerre's advanced model-data fusion (MDF) technique.

Appendix B: Key parameters used in the model employed in this model version report

Category	Description	Value	Unit
Soil C, N and P Transformations	Michaelis–Menten constant for NH_4^+ uptake at microbial surfaces	0.4	gN/m^3
	Michaelis–Menten constant for NO_3^- uptake at microbial surfaces	0.35	gN/m^3
	Michaelis–Menten constant for reduction of O_2 by microbes, roots and mycorrhizae	0.064	gO_2/m^3

	specific heterotrophic respiration of microbial biomass under non-limiting DOC, O_2, nutrients, and water at 25°C	0.125	gC/m^2/h
	specific maintenance respiration at 25°C	0.0115	gC/m^2/h
	specific decomposition rate of labile component of microbial mass pool at 25 °C	1.6*10^{-3}	gC/m^2/h
	specific decomposition rate of resistant component of microbial mass pool at 25 °C	8*10^{-5}	gC/m^2/h
	energy requirement for growth of microbial biomass	25	kJ/gC
	energy requirement for non-symbiotic N_2 fixation by heterotrophic diazotrophs	5	gC/gN
	fraction of microbial growth allocated to labile component of microbial biomass	0.55	NA
	fraction of microbial growth allocated to labile component of resistant biomass	0.45	NA
Soil-Plant Water Relations	axial resistivity to water transport along root or mycorrhizal axes for deciduous plants	4*10^9	MpA h /m^4
	axial resistivity to water transport along root or mycorrhizal axes for deciduous plants	1*10^{10}	MpA h /m^4
	radial resistivity to water transport from surface to axis of roots or mycorrhizae	1*10^4	MpA h /m^2
	radius of secondary roots or mycorrhizae for trees	2*10^{-4}	m
	radius of secondary roots or mycorrhizae for bushes	1*10^{-4}	m
	radius of secondary roots or mycorrhizae for mycorrhizae	5*10^{-6}	m
Gross Primary Productivity, Autotrophic			

Respiration, Growth and Litterfall			
	energy cost of nutrient uptake	2.15	gC/gN or gC/gP
	fraction of leaf protein in chlorophyll	0.025	NA
	fraction of leaf protein in rubisco	0.125	NA
	rate constant for shoot-root N, P transfer	0.1	per hour
	rate constant for root-mycorrhizal C transfer	0.1	per hour
	rate constant for root-mycorrhizal N, P transfer	0.1	per hour
	energy of activation for electron transport	$43 * 10^3$	J/mol
	energy of activation for oxygenation	$60 * 10^3$	J/mol
	energy of activation for carboxylation	$65 * 10^3$	J/mol
	inhibition constant for remobilization of leaf or root N during senescence	0.1	gN/gC
	Michaelis-Menten constant for NH_4^+ uptake at root or mycorrhizal surfaces	0.4	gN/m^3
	Michaelis-Menten constant for NO_3^- uptake at root or mycorrhizal surfaces	0.35	gN/m^3
	Michaelis-Menten constant for root or mycorrhizal O_2 uptake	0.064	g/m^3
	N content of protein remobilized from leaf or root	0.4	gN/gC
	specific chlorophyll e^- transfer at 25 ^oC	450	$\mu mol\ g^{-1}$ chlorophyll s^{-1}
	specific rubisco oxygenation at 25 ^oC	9.5	$\mu mol\ g^{-1}$ rubisco s^{-1}
Symbiotic N_2 Fixation			

	direct energy cost of N_2 fixation	0.25	gN/gC
	Michaelis-Menten constant for nodule N_2 uptake	0.14	gN/m^3
	nodule growth yield	0.67	gC/gC
Inorganic N Transformations			
	Michaelis-Menten constant for oxidation of NH_3 by nitrifiers	0.01	gN/m^3
	Michaelis-Menten constant for microbial NH_4^+ uptake	0.35	gN/m^3
	Michaelis-Menten constant for reduction of NO_2^- by denitrifiers	3.5	gN/m^3
	Michaelis-Menten constant for reduction of NO_2^- by nitrifiers	3.5	gN/m^3
	Michaelis-Menten constant for oxidation of NO_2^- by nitrifiers	10	gN/m^3
	Michaelis-Menten constant for reduction of NO_3^- by denitrifiers	3.5	gN/m^3
	Michaelis-Menten constant for reduction of N_2O by denitrifiers	0.35	gN/m^3
	Michaelis-Menten constant for reduction of O_2 by heterotrophs	0.064	gO_2/m^3
	Michaelis-Menten constant for reduction of O_2 by NH_3 oxidizers	0.32	gO_2/m^3
	Michaelis-Menten constant for reduction of O_2 by NO_2^- oxidizers	0.32	gO_2/m^3
	Michaelis-Menten constant for oxidation of DOC by heterotrophs	12	gC/m^3

Appendix C: Methods of temporal aggregation to estimate annually or seasonally accumulated N_2O emission

For studies with reported aggregated values, we directly used the reported values for model validation. For sites lacking standard deviation/standard error of annually or seasonally aggregated N_2O emission, we derive them using the following steps: (1) For each daily observation with reported mean and corresponding standard deviation, sampling n pseudo-measurements within the range of [mean-3*std,mean+3*std] following uniform distribution where n is the number of measurement replicates, and mean and std represent the mean and standard deviation of daily observations reported in the experimental studies; (2) Using linear interpolation to get annually or seasonally aggregated values for the n sampled sequences; and (3) calculate the standard deviations of the n aggregated values.

Appendix D: Studies in the Validation Dataset

Table D1. The study sites analyzed in the model validation report. ES, LRR, CFTs, PCs, and N stand for Emissions sources, Land Resource Region, Crop Functional Types, Practice Categories, and Number of paired measurements respectively.

Study site	ES	LRR	CFTs	PCs	Fertilizer treatment	Soil Texture Classes	Clay content (%)	N	DOIs
Ames	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silt loam	25	8	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Ames	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silt loam	25	19	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Armstrong	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay loam	35	10	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Armstrong	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silty clay loam	35	11	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Central	SOC	M	Corn, Soybean	CROPPING	NA	Loam	21	4	10.1371/journal.pone.0172293
Central	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Loam	21	12	10.1371/journal.pone.0172293

Crawfordsville	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay loam	30	8	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Crawfordsville	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silty clay loam	30	21	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Crossville	SOC	N	Soybean	CROPPING	NA	Sandy loam	11	1	10.2136/sssaj1992.03615995005600050040x
Dixon Springs	SOC	N	Corn, Soybean	CROPPING	NA	Silt loam	19	4	10.1097/SS.0b013e3181cf7959
Dixon Springs	SOC	N	Corn, Soybean	DISTURBANCE	NA	Silt loam	19	2	10.1097/SS.0b013e3181cf7959
East Lansing	SOC	P	Corn	CROPPING	NA	Silty clay loam	30	2	10.1097/SS.0000000000000131
Fort Collins	SOC	G	Corn	FERTILIZER	Magnitude	Loam	25	1	10.2134/agronj2015.0402
Hoytville 2	SOC	L	Corn, Soybean	DISTURBANCE	Magnitude	Clay loam	33	2	10.1097/01.ss.0000162286.95137.70
Ithaca	SOC	M	Corn	FERTILIZER	Magnitude	Silty clay loam	33	2	10.2136/sssaj2015.02.0053
Kanawha	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silt loam	26	8	10.1016/j.agee.2004.08.002;

									10.1002/saj2.20003
Kanawha	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silt loam	26	14	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Kansas	SOC	H	Soybean	DISTURBANCE	NA	Silt loam	20	1	https://www.researchgate.net/profile/Charles-Rice-2/publication/228647002_Soil_Carbon_Sequestration_in_Kansas_Long-Term_Effect_of_Tillage_N_Fertilization_and_Crop_Rotation/links/00b49528d246d39222000000/Soil-Carbon-Sequestration-in-Kansas-Long-Term-Effect-of-Tillage-N-Fertilization-and-Crop-Rotation.pdf
Lexngton	SOC	N	Corn	FERTILIZER	Magnitude	Silt loam	20	6	10.2136/sssaj1994.03615995005800010028x
McNay	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silty clay	48	15	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003

McNay	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay	48	8	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Mead	SOC	M	Corn, Soybean	CROPPING	NA	Silty clay loam	30	2	10.2134/agronj1994.00021962008600020021x
Mead	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay loam	30	2	10.2134/agronj1994.00021962008600020021x
Minnesota	SOC	M	Corn	ORGANIC (manure)	NA	Clay loam	30	1	10.2136/sssaj2017.09.0344
Nashua	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silt loam	20	12	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Nashua	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silt loam	20	8	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Northwest	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay loam	37	18	0.137 1/journal.pone.0172293
Rosemount	SOC	M	Corn, Soybean	CROPPING	NA	Silt loam	23	2	10.1016/S0167-1987(00)00110-0
Rosemount	SOC	M	Corn	FERTILIZER	Magnitude	Silt loam	23	2	10.1016/S0167-1987(00)00110-0

South	SOC	M	Corn, Soybean	CROPPING	NA	Silty clay loam	27	2	0.1371/journal.pone.0172293
South	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay loam	27	18	0.1371/journal.pone.0172293
SouthDakota	SOC	F	Corn	CROPPING	NA	Sandy clay loam	30	2	10.2136/sssaj2008.0020
Southeast	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silt loam	21	17	0.1371/journal.pone.0172293
Sutherland	SOC	M	Corn, Soybean	DISTURBANCE	NA	Silty clay	48	11	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Sutherland	SOC	M	Corn, Soybean	FERTILIZER	Magnitude	Silty clay	48	8	10.1016/j.agee.2004.08.002; 10.1002/saj2.20003
Tenessa	SOC	P	Corn	CROPPING	NA	Silt loam	22	2	https://link.springer.com/chapter/10.1007/978-3-319-04084-4_28
Urbana	SOC	M	Corn	ORGANIC (manure)	NA	Silt loam	22	2	10.1016/S0167-1987(99)00051-3
Wisconsin	SOC	K	Corn, Soybean	CROPPING	NA	Silt loam	22	1	10.1016/j.still.2015.09.008

AL1	N2O	P	Soybean	CROPPING	NA	Sandy clay loam	21	9	10.2489/jswc.2023.00042
CO1	N2O	G	Corn	FERTILIZER	Magnitude	Loam	25	24	10.2134/jeq2007.0268
CO1	N2O	G	Corn	DISTURBANCE	NA	Loam	25	9	10.2134/jeq2007.0268
CO3	N2O	G	Corn	FERTILIZER	Form	Loam	23	42	10.2134/jeq2007.0268
CO4	N2O	G	Corn	FERTILIZER	Form	Loam	23	18	10.2134/jeq2015.08.0426
CO4	N2O	G	Corn	ORGANIC (manure)	NA	Loam	23	3	10.2134/jeq2015.08.0426
IA1	N2O	M	Corn, Soybean	CROPPING	NA	Loam	20	4	10.2134/agronj2018.03.0187
IA1	N2O	M	Corn, Soybean	DISTURBANCE	NA	Loam	20	4	10.2134/agronj2018.03.0187
IN1	N2O	M	Corn, Soybean	CROPPING	NA	Silt loam	24	4	10.2134/agronj2013.0184 Nitrification Kinetics and Nitrous Oxide Emissions when Nitrapyrin is Coapplied with Urea¨CAmmonium Nitrate

IN1	N2O	M	Corn	ORGANIC (manure)	NA	Silt loam	24	2	10.2134/agronj2013.0184 Nitrification Kinetics and Nitrous Oxide Emissions when Nitrapyrin is Coapplied with Urea¨CAmmonium Nitrate
IN2	N2O	M	Corn, Soybean	FERTILIZER	Form	Silt loam	20	2	10.2134/agronj2013.0184 Nitrification Kinetics and Nitrous Oxide Emissions when Nitrapyrin is Coapplied with Urea¨CAmmonium Nitrate
KS1	N2O	H	Corn, Soybean	CROPPING	NA	Clay loam	31	6	10.2134/agronj2018.03.0187
KY1	N2O	N	Corn	FERTILIZER	Magnitude	Silt loam	22	14	10.2134/jeq2011.0197 Atmospheric emissions of nitrous oxide, methane, and carbon dioxide from different nitrogen fertilizers
KY1	N2O	N	Corn	ORGANIC (manure)	NA	Silt loam	22	2	10.2134/jeq2011.0197 Atmospheric emissions of nitrous oxide, methane, and

									carbon dioxide from different nitrogen fertilizers
MI1	N2O	L	Soybean	DISTURBANCE	NA	Loamy sand	4	21	10.2134/jeq2005.0166
MN2	N2O	M	Corn	FERTILIZER	Form	Sandy loam	17	5	10.2136/sssaj2009.0078
SD1	N2O	F	Corn, Soybean	CROPPING	NA	Clay loam	30	8	10.2136/sssaj2016.01.0021
WI1	N2O	K	Corn, Soybean	CROPPING	NA	Silt loam	22	2	10.2134/jeq2014.02.0077

Table D2. The duration of direct N_2O emissions measurements analyzed in the model validation report.

Study site	Measurement start date	Measurement end date	Measurement ID	Measurement duration (days)
AL1	2012-05-01	2012-11-01	AL1_1	184
AL1	2013-05-01	2013-11-01	AL1_2	184
AL1	2014-05-01	2014-11-01	AL1_3	184
AL1	2015-05-01	2015-11-01	AL1_4	184
AL1	2016-05-01	2016-11-01	AL1_5	184
CO1	2005-05-02	2005-11-04	CO1_1	186
CO1	2006-04-28	2006-10-02	CO1_2	157
CO1	2005-05-02	2005-11-04	CO1_3	186
CO1	2006-04-28	2006-10-02	CO1_4	157
CO1	2005-05-02	2005-11-04	CO1_5	186
CO1	2006-04-28	2006-10-02	CO1_6	157
CO1	2005-05-02	2005-11-04	CO1_7	186

CO1	2006-04-28	2006-10-02	CO1_8	157
CO1	2005-05-02	2005-11-04	CO1_9	186
CO1	2006-04-28	2006-10-02	CO1_10	157
CO1	2005-05-02	2005-11-04	CO1_11	186
CO1	2006-04-28	2006-10-02	CO1_12	157
CO1	2005-05-02	2005-11-04	CO1_13	186
CO1	2006-04-28	2006-10-02	CO1_14	157
CO1	2005-05-02	2005-11-04	CO1_15	186
CO1	2006-04-28	2006-10-02	CO1_16	157
CO1	2005-05-02	2005-11-04	CO1_17	186
CO1	2006-04-28	2006-10-02	CO1_18	157
CO1	2005-05-02	2005-11-04	CO1_19	186
CO1	2006-04-28	2006-10-02	CO1_20	157
CO1	2005-05-02	2005-11-04	CO1_21	186

CO1	2006-04-28	2006-10-02	CO1_22	157
CO1	2005-05-02	2005-11-04	CO1_23	186
CO1	2006-04-28	2006-10-02	CO1_24	157
CO1	2005-05-02	2005-11-04	CO1_25	186
CO1	2006-04-28	2006-10-02	CO1_26	157
CO1	2005-05-02	2005-11-04	CO1_27	186
CO1	2006-04-28	2006-10-02	CO1_28	157
CO1	2005-05-02	2005-11-04	CO1_29	186
CO1	2006-04-28	2006-10-02	CO1_30	157
CO1	2005-05-02	2005-11-04	CO1_31	186
CO1	2006-04-28	2006-10-02	CO1_32	157
CO3	2009-05-05	2009-09-29	CO3_1	147
CO3	2010-05-06	2010-09-29	CO3_2	146
CO3	2009-05-05	2009-09-29	CO3_3	147

CO3	2010-05-06	2010-09-29	CO3_4	146
CO3	2009-05-05	2009-09-29	CO3_5	147
CO3	2010-05-06	2010-09-29	CO3_6	146
CO3	2009-05-05	2009-09-29	CO3_7	147
CO3	2010-05-06	2010-09-29	CO3_8	146
CO3	2009-05-05	2009-09-29	CO3_9	147
CO3	2010-05-06	2010-09-29	CO3_10	146
CO3	2009-05-05	2009-09-29	CO3_11	147
CO3	2010-05-06	2010-09-29	CO3_12	146
CO3	2009-05-05	2009-09-29	CO3_13	147
CO3	2010-05-06	2010-09-29	CO3_14	146
CO3	2009-05-05	2009-09-29	CO3_15	147
CO3	2010-05-06	2010-09-29	CO3_16	146
CO3	2009-05-05	2009-09-29	CO3_17	147

CO3	2010-05-06	2010-09-29	CO3_18	146
CO3	2009-05-05	2009-09-29	CO3_19	147
CO3	2010-05-06	2010-09-29	CO3_20	146
CO3	2009-05-05	2009-09-29	CO3_21	147
CO3	2010-05-06	2010-09-29	CO3_22	146
CO3	2009-05-05	2009-09-29	CO3_23	147
CO3	2010-05-06	2010-09-29	CO3_24	146
CO3	2009-05-05	2009-09-29	CO3_25	147
CO3	2010-05-06	2010-09-29	CO3_26	146
CO3	2009-05-05	2009-09-29	CO3_27	147
CO3	2010-05-06	2010-09-29	CO3_28	146
CO3	2009-05-05	2009-09-29	CO3_29	147
CO3	2010-05-06	2010-09-29	CO3_30	146
CO3	2009-05-05	2009-09-29	CO3_31	147

CO3	2010-05-06	2010-09-29	CO3_32	146
CO3	2009-05-05	2009-09-29	CO3_33	147
CO3	2010-05-06	2010-09-29	CO3_34	146
CO3	2009-05-05	2009-09-29	CO3_35	147
CO3	2010-05-06	2010-09-29	CO3_36	146
CO3	2009-05-05	2009-09-29	CO3_37	147
CO3	2010-05-06	2010-09-29	CO3_38	146
CO3	2009-05-05	2009-09-29	CO3_39	147
CO3	2010-05-06	2010-09-29	CO3_40	146
CO3	2009-05-05	2009-09-29	CO3_41	147
CO3	2010-05-06	2010-09-29	CO3_42	146
CO4	2012-03-20	2012-10-31	CO4_1	225
CO4	2012-11-01	2013-03-19	CO4_2	138
CO4	2013-03-20	2013-10-31	CO4_3	225

CO4	2013-11-01	2014-03-19	CO4_4	138
CO4	2014-03-20	2014-10-31	CO4_5	225
CO4	2014-11-01	2015-03-19	CO4_6	138
CO4	2012-03-20	2012-10-31	CO4_7	225
CO4	2012-11-01	2013-03-19	CO4_8	138
CO4	2013-03-20	2013-10-31	CO4_9	225
CO4	2013-11-01	2014-03-19	CO4_10	138
CO4	2014-03-20	2014-10-31	CO4_11	225
CO4	2014-11-01	2015-03-19	CO4_12	138
CO4	2012-03-20	2012-10-31	CO4_13	225
CO4	2012-11-01	2013-03-19	CO4_14	138
CO4	2013-03-20	2013-10-31	CO4_15	225
CO4	2013-11-01	2014-03-19	CO4_16	138
CO4	2014-03-20	2014-10-31	CO4_17	225

CO4	2014-11-01	2015-03-19	CO4_18	138
IA1	2003-04-12	2004-03-29	IA1_1	352
IA1	2004-04-05	2005-02-11	IA1_2	312
IA1	2003-04-12	2004-03-29	IA1_3	352
IA1	2003-04-12	2004-03-29	IA1_4	352
IA1	2004-04-05	2005-02-11	IA1_5	312
IA1	2003-04-12	2004-03-29	IA1_6	352
KS1	2011-01-01	2011-12-31	KS1_1	364
KS1	2012-01-01	2012-12-31	KS1_2	365
KS1	2013-01-01	2013-12-31	KS1_3	364
KS1	2011-01-01	2011-12-31	KS1_4	364
KS1	2012-01-01	2012-12-31	KS1_5	365
KS1	2013-01-01	2013-12-31	KS1_6	364
MI1	1991-01-01	1991-12-31	MI1_1	364

MI1	1992-01-01	1992-12-31	MI1_2	365
MI1	1993-01-01	1993-12-31	MI1_3	364
MI1	1994-01-01	1994-12-31	MI1_4	364
MI1	1995-01-01	1995-12-31	MI1_5	364
MI1	1996-01-01	1996-12-31	MI1_6	365
MI1	1997-01-01	1997-12-31	MI1_7	364
MI1	1998-01-01	1998-12-31	MI1_8	364
MI1	1999-01-01	1999-12-31	MI1_9	364
MI1	2000-01-01	2000-12-31	MI1_10	365
MI1	2001-01-01	2001-12-31	MI1_11	364
MI1	2002-01-01	2002-12-31	MI1_12	364
MI1	2003-01-01	2003-12-31	MI1_13	364
MI1	2004-01-01	2004-12-31	MI1_14	365
MI1	2005-01-01	2005-12-31	MI1_15	364

MI1	2006-01-01	2006-12-31	MI1_16	364
MI1	2007-01-01	2007-12-31	MI1_17	364
MI1	2008-01-01	2008-12-31	MI1_18	365
MI1	2009-01-01	2009-12-31	MI1_19	364
MI1	2010-01-01	2010-12-31	MI1_20	364
MI1	2011-01-01	2011-12-31	MI1_21	364
MN2	2006-04-20	2006-10-05	MN2_1	168
MN2	2007-04-01	2007-10-01	MN2_2	183
MN2	2005-04-15	2005-11-10	MN2_3	209
MN2	2007-04-01	2007-10-01	MN2_4	183
MN2	2006-04-20	2006-10-05	MN2_5	168
SD1	2009-04-01	2009-11-01	SD1_1	214
SD1	2010-04-01	2010-11-01	SD1_2	214
SD1	2011-04-01	2011-11-01	SD1_3	214

SD1	2012-04-01	2012-11-01	SD1_4	214
WI1	2010-04-01	2010-11-01	WI1_1	214
WI1	2011-04-01	2011-11-01	WI1_2	214

Appendix E: The sensitivity of SOC stock to SOC sampling depth

CAR SEP provided the following measurement guideline to quantify changes in soil organic carbon (SOC) stocks: "*Measured datasets of SOC stock change may be made at any depth, but the model must also predict SOC stock change at the corresponding depth. Thus, a fully compiled dataset for validating model performance and uncertainty may contain different depths for SOC stock change measurements as long as the model is predicting SOC stock change at each corresponding depth*".

We followed the guidelines provided in the SEP Model Requirements and evaluated the measured and modeled SOC stock changes at the same depth (based on the SOC measurement depth). We further examined the sensitivity of SOC stock to SOC sampling depth to evaluate the validity of aggregating SOC stock collected among different depths into the same SOC dataset. Our results showed that changes in SOC stock are comparable when aggregated into different measurement depth groups (Figure E1). Therefore, it is beneficial to include SOC stock measured at different depths to improve the data availability in the model validation dataset used in this model validation report.

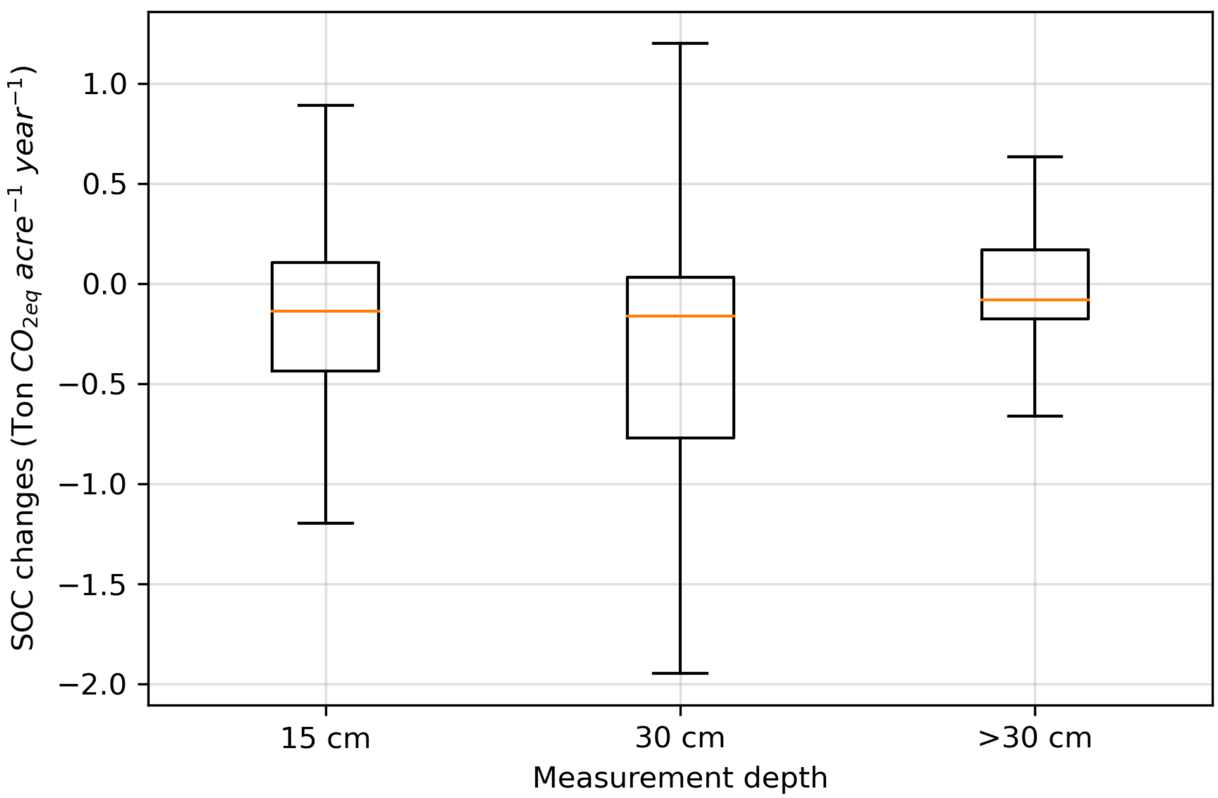

Figure E1. Boxplots showing the distribution of measured changes in SOC stock driven by different treatment pairs among study sites collected in this model validation report. Orange lines represent the median of the SOC stock change measured at the corresponding measurement depth. The bottom and top edges of each box indicate the 25th and 75th percentiles, respectively. The black whiskers extend to the most extreme data points not considered outliers.

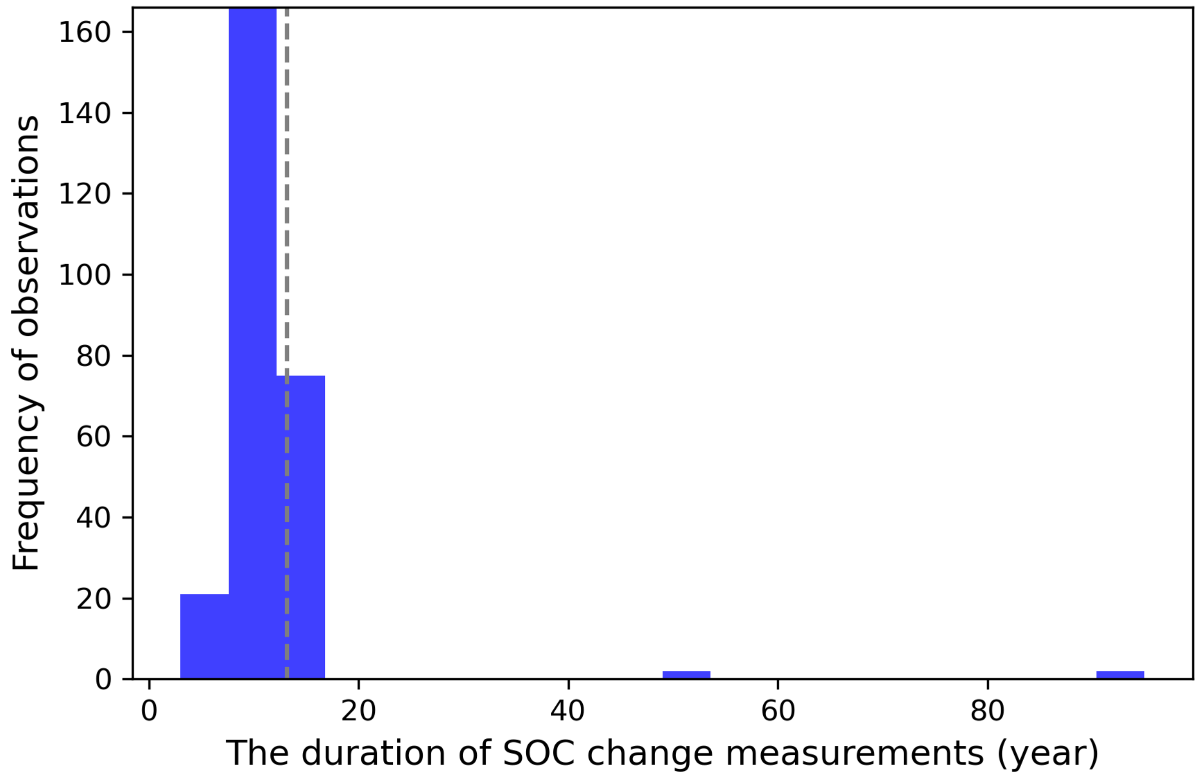

Figure E2. The duration of SOC change measurements used for model validation, including all practice categories (PCs) and all crop functional types (CFTs) evaluated in this report. The dashed line represents the mean duration of SOC change measurements (13 years).

Appendix F: Deviation Request Email Correspondence with CAR

Approval email from CAR:

From: McKenzie Smith <msmith@climateactionreserve.org> Fri, Oct 13, 2023 at 10:31 AM

To: Nick Reinke <nick.reinke@habiterre.com>

Cc: Brian McConkey <brianmcc.soils22@gmail.com>, Bin Peng <bin.peng@habiterre.com>, reserve <reserve@climateactionreserve.org>, Kaiyu Guan <kaiyu.guan@habiterre.com>, Kuangyu Chang <kuangyu.chang@habiterre.com>

Good morning,

Thank you again for reaching out with your question as it relates to the SEP Model Guidance and for your patience as we reviewed internally.

We can approve you moving forward with the described deviation from the SEP Model Guidance but will note that the approval of the model validation itself is dependent on our review of the model validation report. We will be designing and implementing a new template to summarize and communicate the applicability of various tool as it relates to these available data inputs and the reviewer's comments. More specifically, it is our goal to create a template that could communicate what areas the model has been validated for relative to the crop functional groups and practice change comparisons for emission sources over LRRs. It is our hope that this form will help summarize and communicate any model applicability specifics to the public and the verification body as it relates to these input parameters.

In terms of timelines, please feel free to move forward with your model review and will try to design this template as soon as possible to send back to you and Brian to complete as a part of your review. This document would be made public on our webpage along with the model validation report.

Please feel free to reach out with any questions or to discuss further.

Best,

McKenzie

McKenzie Smith, M.Sc.

Associate Director

msmith@climateactionreserve.org

Climate Action Reserve, the most trusted global offset registry.

(she/her) | California | office: (213) 542-0282 | mobile: 408-759-3125

CAR Notification of receiving the request:

On Fri, Sep 8, 2023 at 2:24 PM McKenzie Smith <msmith@climateactionreserve.org> wrote:

Good afternoon,

Thank you both for your time to develop this report and provide additional context as it relates to the geographical coverage.

We will conduct our review of the report internally and get back with any questions or concerns. We will do our best to complete this by the end of the month but it will depend on our review of this variance.

Thanks again and please don't hesitate to reach out in the meantime.

Best,

McKenzie

McKenzie Smith, M.Sc.

Associate Director

msmith@climateactionreserve.org

Climate Action Reserve, the most trusted global offset registry.

(she/her) | California | office: (213) 542-0282 | mobile: 408-759-3125

Comments from model reviewer:

From: Brian McConkey <brianmcc.soils22@gmail.com>

Sent: Wednesday, August 30, 2023 9:58 AM

To: Bin Peng <bin.peng@habiterre.com>

Cc: McKenzie Smith <msmith@climateactionreserve.org>; reserve <reserve@climateactionreserve.org>; Chloe Ney <cney@climateactionreserve.org>; Kaiyu Guan <kaiyu.guan@habiterre.com>; Kuangyu Chang <kuangyu.chang@habiterre.com>; Nick Reinke <nick.reinke@habiterre.com>

Subject: Re: HabiTerre model validation update

Dear McKenzie:

I am the formal reviewer of the model validation report of Habiterre. I believe that Habiterre has thoroughly searched for data suitable for validating the model Ecosys. Unfortunately, there is insufficient data available to provide an adequate geographical coverage, as indicated by having at least 3 LRR included, for validation based on CFGxPCxES. In practice, for application for a project, the model will be used across a wide range of sites (farms and fields). In addition, for model validation, most of the variation in model performance variation is by site, so the more sites involved in each validation provides the best evaluation of the model. Therefore, I recommend that they be allowed the variance from the Requirements and Guidance for Model Calibration, Validation, Uncertainty, and Verification For Soil Enrichment Projects, Version 1.1a to have the validation can be done across CFG and PC for each ES (SOC and N2O) as Habiterre has requested. This recommendation also relieves another problem of insufficient data for validation by CFGXPCxES of enabling calculation of a more representative PMU based on multiple studies (only a fraction of the studies provide the data required for PMU so more studies included in the PMU calculation is preferred).

To increase transparency of the validation and show that the validation across the CFG and PC does not hide weak performance for a CFG or PC, I will require the model validation report include an appendix with validation results by CFG across PC and by PC across CFG. This breakdown allows for finer resolution but provides a good number of locations, although not enough locations to always meet the LRR requirements as laid out in the Model Validation Guidance. This appendix would not be used for the final assessment of validity as that would be for the recommended validation across PC and CFG as described earlier. Although the current results do not indicate any weaknesses by CFG or PC, if there were obvious problems by CFG and/or PC shown in these validation results in the Appendix that cannot be resolved, these problems would affect my assessment regarding the valid scope for model application in projects.

Thanks for your consideration of this recommendation.

Best regards,

Brian McConkey, PhD

Validation Report Reviewer

Initial request from HabiTerre:

On Mon, Aug 28, 2023 at 12:09 PM Bin Peng <bin.peng@habiterre.com> wrote:

Dear McKenzie,

Hope you are doing very well. We are writing to provide you a quick progress update and inquire about a potential exception for our validation report with regard to data availability by the Land Resource Region (LRR). We have prepared the model validation report and have been making good progress with our independent reviewers toward a final report to be shared for review and approval in the near future.

One specific question we have for you is about the 3 Land Resource Regions (LRRs) requirements for validating a model for each combination of Crop Functional Group (CFT) and Practice Category (CFTxPC). We have conducted comprehensive literature surveys and collected all available studies with the data required to validate ecosys within the domain. However, we found most CFTxPC do not have 3 LRRs to meet the requirements for both soil organic carbon (SOC) and nitrous oxide(N2O). Given this, we validated ecosys and evaluated its bias and uncertainty for both SOC and N2O over a single validation domain comprising multiple LRRs, CFPs, and PCs, which is a decision we have to make because there is insufficient data at finer resolution.

Given that this exception has been granted for other validated models, we wish to inquire as to whether an exception could be granted here for the above approach. We look forward to guidance from CAR on this before we move forward to finalize our model validation report for your review.

Thanks a lot,

Bin.

Appendix G: Practice and crop specific model prediction errors

Table G1. Model prediction bias, root mean square error (RMSE), and confidence coverage rate (CCR) for practice change effects on SOC changes and direct N_2O emissions for different practice categories and crop functional types. LRR stands for Land Resource Regions.

SOC					
Practice category	**Bias (Ton $CO_{2,eq}$ acre^{-1} year^{-1})**	**RMSE (Ton $CO_{2,eq}$ acre^{-1} year^{-1})**	**CCR (%)**	**Clay content (%)**	**LRR**
CROPPING	0.298	0.59	90.9	11-30	F, K, M, N, P

FERTILIZER	0.084	0.34	96.3	19-48	G, M, N
DISTURBANCE	0.054	0.36	93.5	19-48	H, L, M, N
ORGANIC	-0.156	0.23	100	22-30	M
Soybean	0.102	0.41	97.1	11-48	F, K, M, N, P
Corn	0.078	0.40	93.4	11-48	H, K, L, M, N
N2O					
Practice category	**Bias (Ton $CO_{2,eq}$ acre^{-1})**	**RMSE (Ton $CO_{2,eq}$ acre^{-1})**	**CCR (%)**	**Clay content (%)**	**LRR**
CROPPING	0.057	0.13	93.6	20-31	F, H, K, M, P
FERTILIZER	-0.008	0.18	93.2	17-25	G, M, N
DISTURBANCE	-0.002	0.08	100.0	4-25	G, L, M
ORGANIC	-0.236	0.387	57.1	23	G, M, N
Soybean	-0.001	0.09	97.7	4-31	F, L, H, K, M, P
Corn	-0.006	0.19	91.6	17-31	F, G, H, K, M, N

G.1 Model prediction error for practice change effects on SOC changes under CROPPING practices

The mean square error is calculated as 0.344 (Ton $CO_{2,\ eq}$ acre^{-1} year^{-1})^2 and the confidence coverage rate is 90.9% for the modeled practice change effects on SOC changes under CROPPING practices.

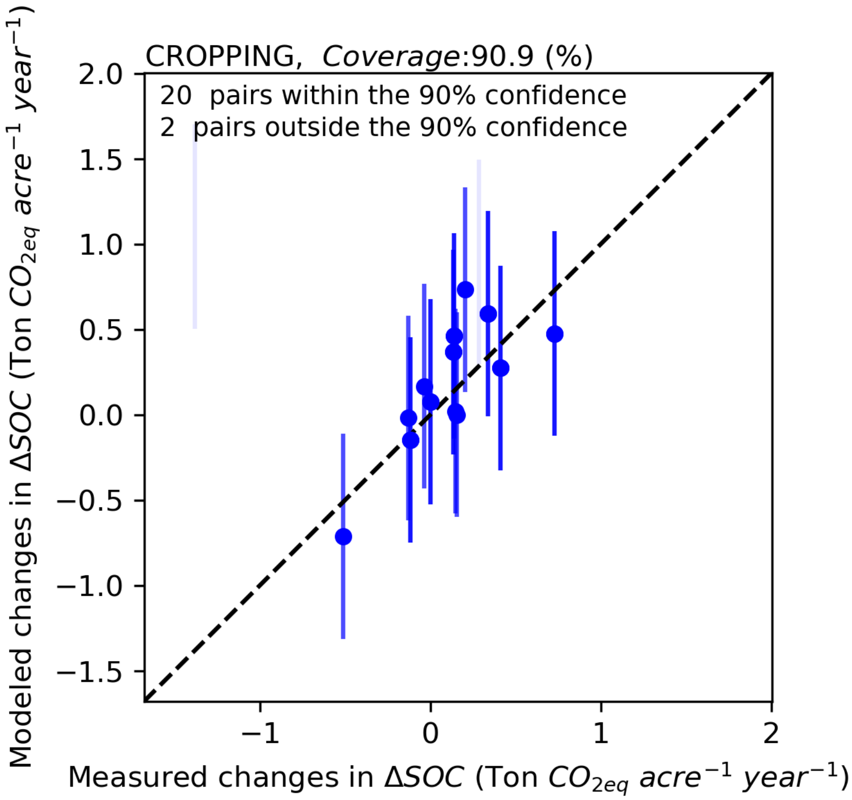

Figure G1. Modeled versus measured practice change effects on SOC changes in response to cropping practice changes for corn. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

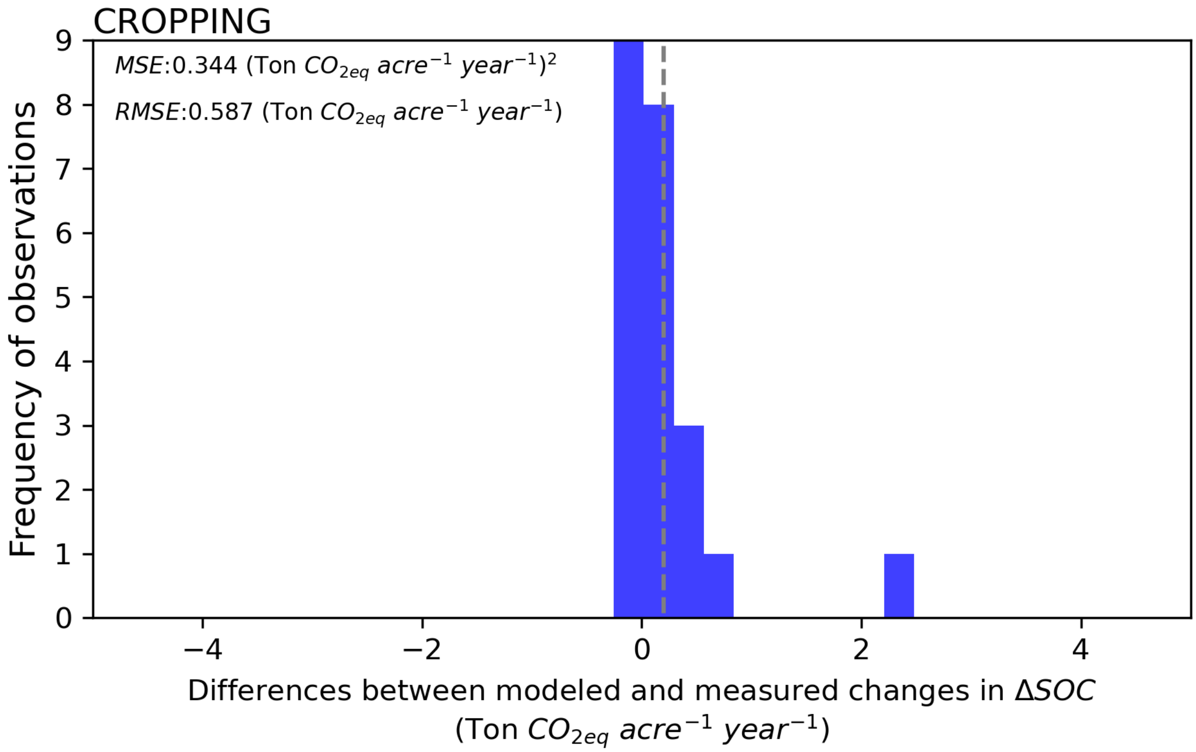

Figure G2. Histogram of the differences between modeled and measured practice change effects on SOC changes in response to cropping practice change for corn. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.2 Model prediction error for practice change effects on SOC changes under DISTURBANCE practices

The mean square error is calculated as 0.131 (Ton $CO_{2,\,eq}$ acre^{-1} year^{-1})^2 and the confidence coverage rate is 93.5% for the modeled practice change effects on SOC changes under DISTURBANCE practices.

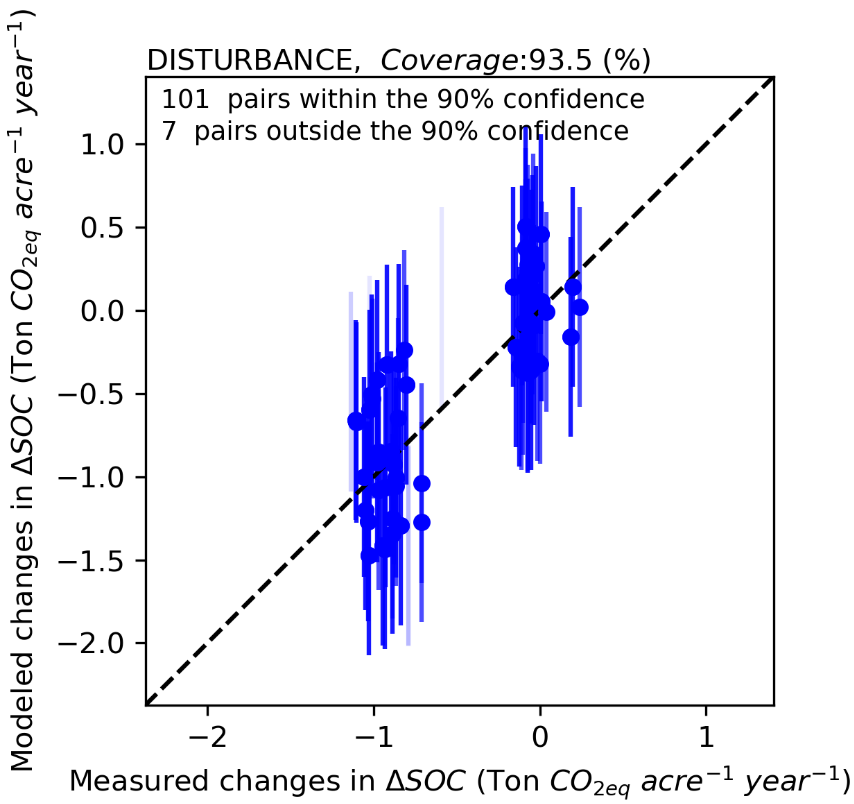

Figure G3. Modeled versus measured practice change effects on SOC changes in response to disturbance practice changes for corn. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

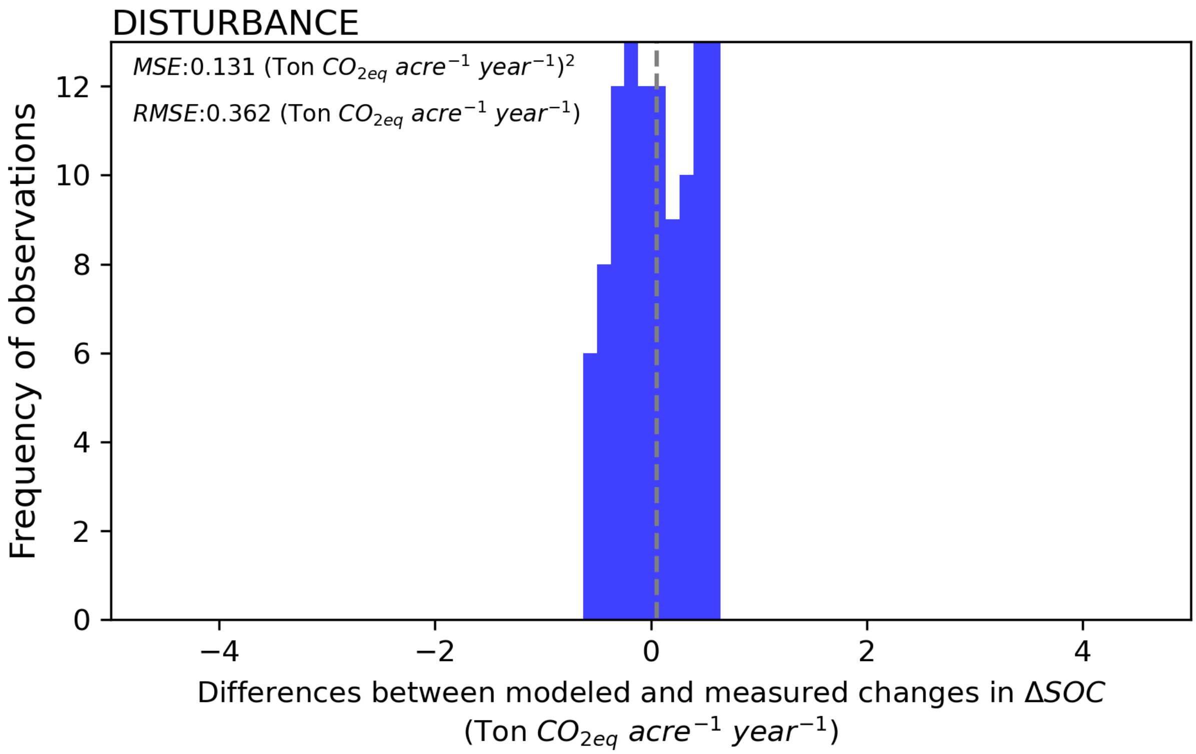

Figure G4. Histogram of the differences between modeled and measured practice change effects on SOC changes in response to disturbance practice change for corn. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.3 Model prediction error for practice change effects on SOC changes with corn under FERTILIZER practices

The mean square error is calculated as 0.117 (Ton $CO_{2,\,eq}$ acre^{-1} year^{-1})^2 and the confidence coverage rate is 96.3% for the modeled practice change effects on SOC changes with corn under FERTILIZER practices.

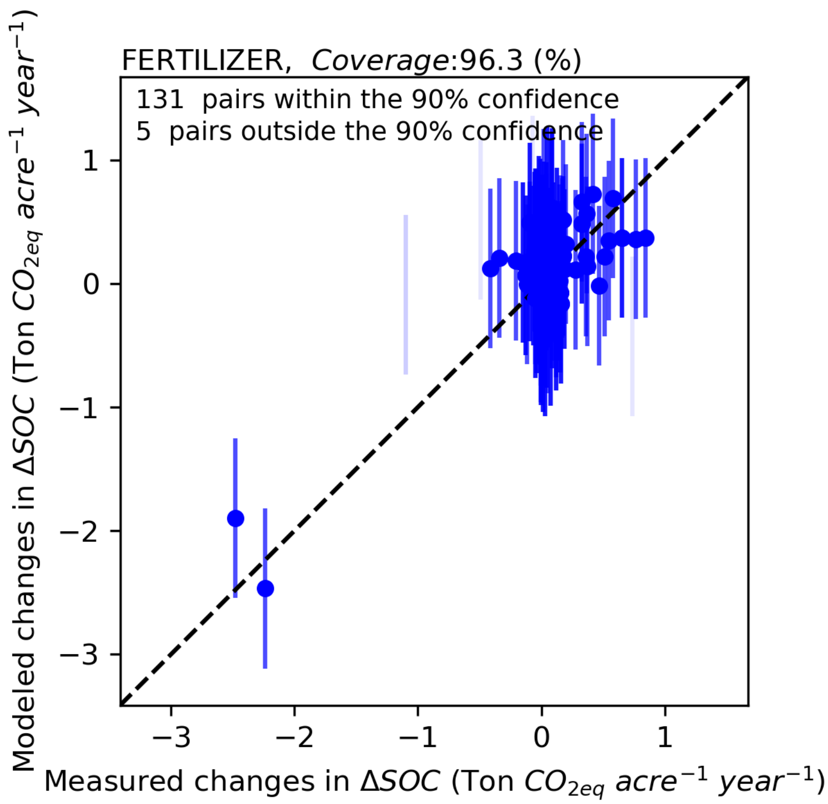

Figure G5. Modeled versus measured practice change effects on SOC changes in response to fertilizer practice changes for corn. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

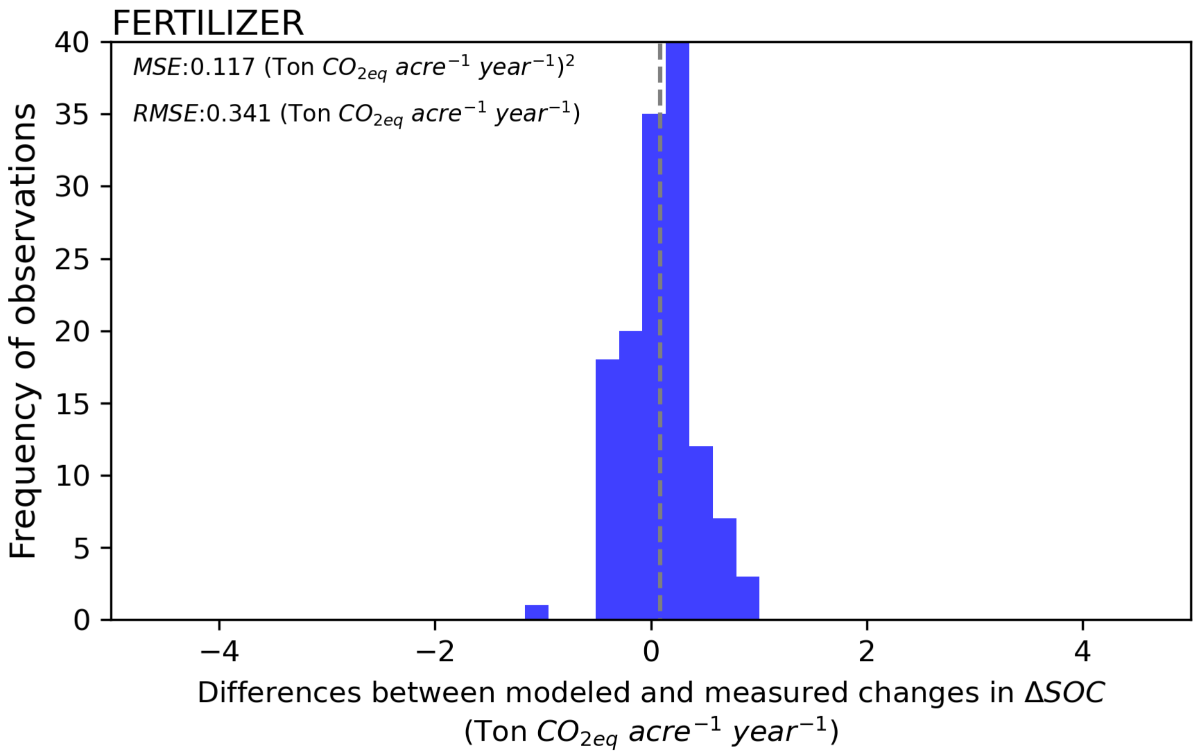

Figure G6. Histogram of the differences between modeled and measured practice change effects on SOC changes in response to fertilizer practice change for corn. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.4 Model prediction error for practice change effects on SOC changes under ORGANIC practices

The mean square error is calculated as 0.053 (Ton CO2, eq acre-1 year-1)^2 and the confidence coverage rate is 100% for the modeled practice change effects on SOC changes with corn under ORGANIC practices.

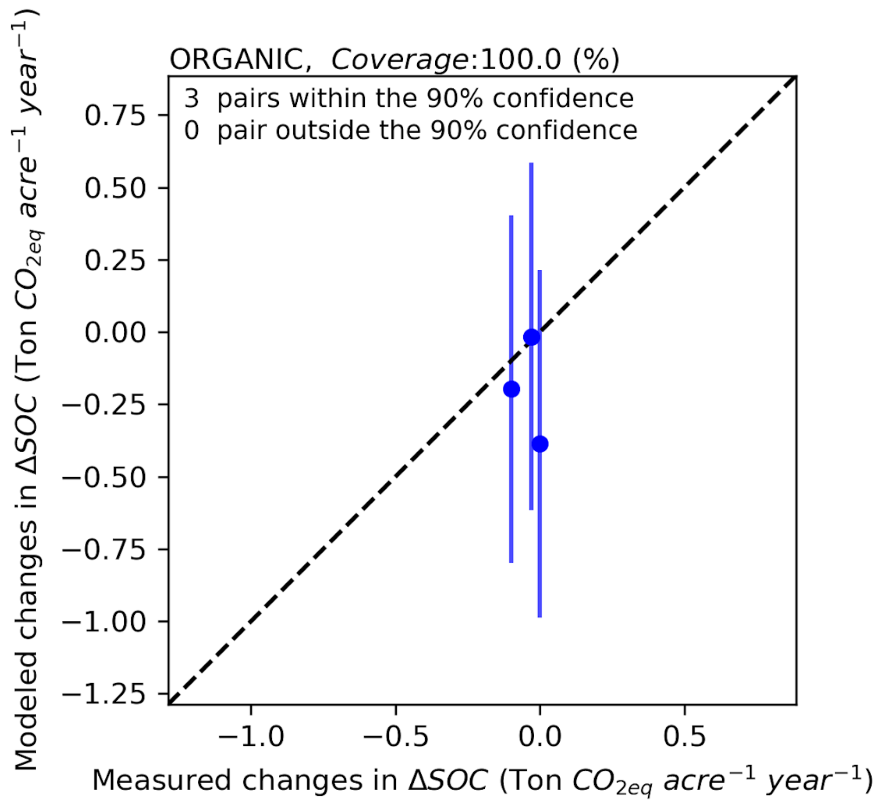

Figure G7. Modeled versus measured practice change effects on SOC changes in response to organic practice changes for corn. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

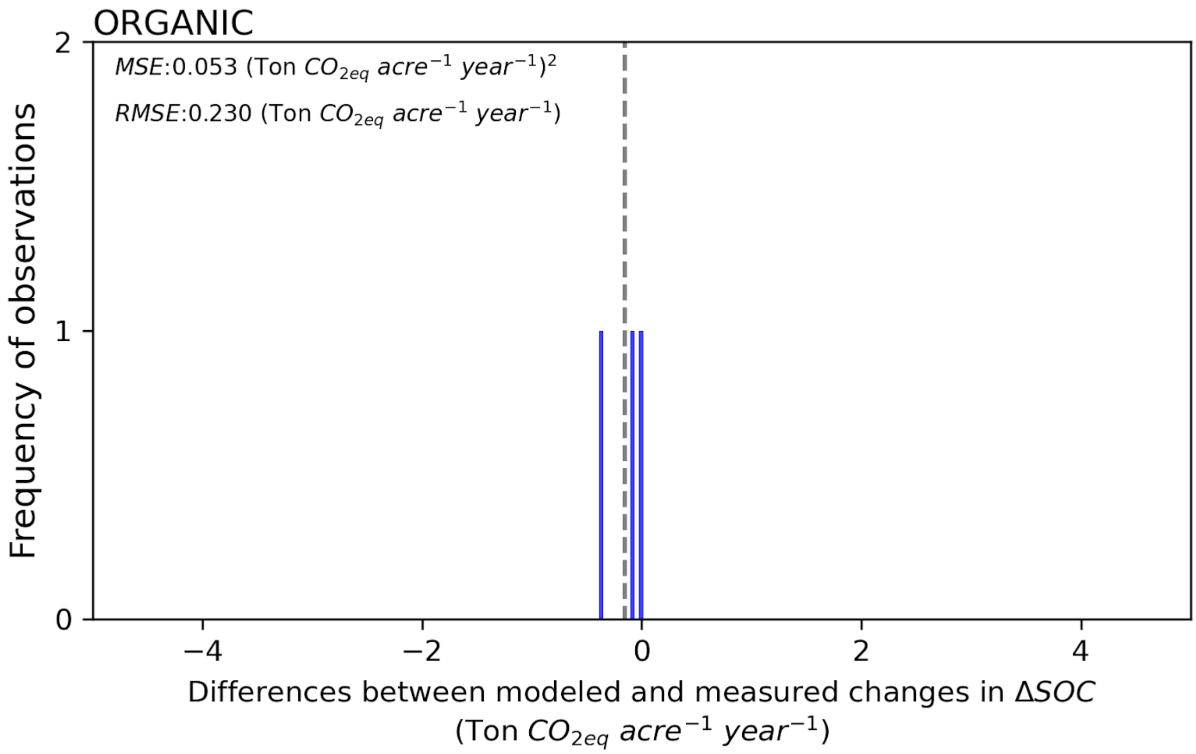

Figure G8. Histogram of the differences between modeled and measured practice change effects on SOC changes in response to organic practice change for corn. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G5 Model prediction error for practice change effects on SOC changes in corn fields

The mean square error is calculated as 0.161 (Ton $CO_{2,\ eq}$ acre^{-1} year^{-1})^2 and the confidence coverage rate is 93.4% for the modeled practice change effects on SOC changes in corn fields.

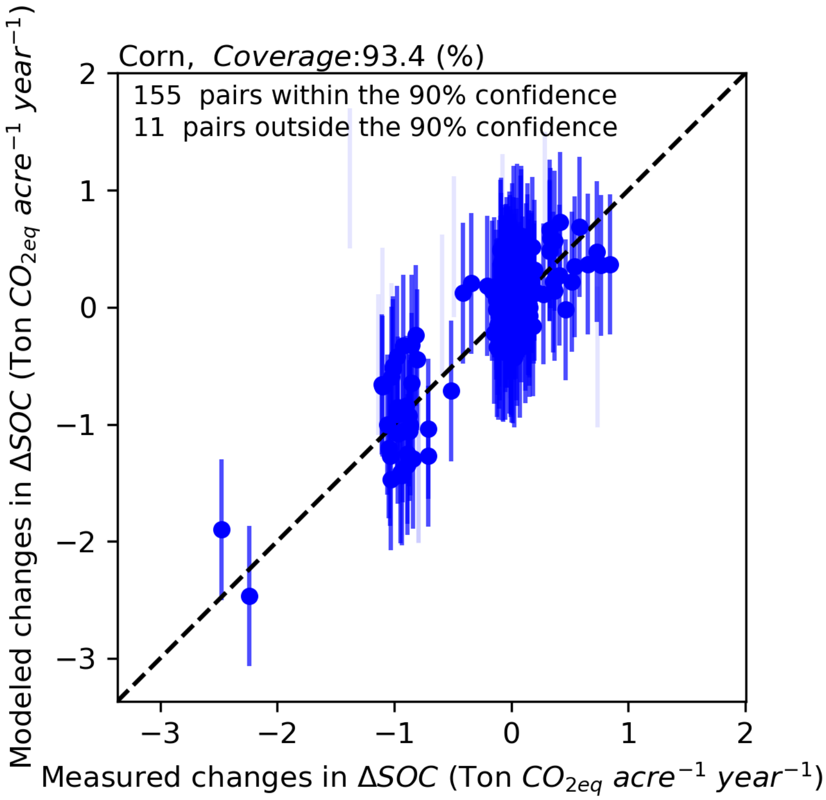

Figure G9. Modeled versus measured practice change effects on SOC changes in response to practice changes in corn fields. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

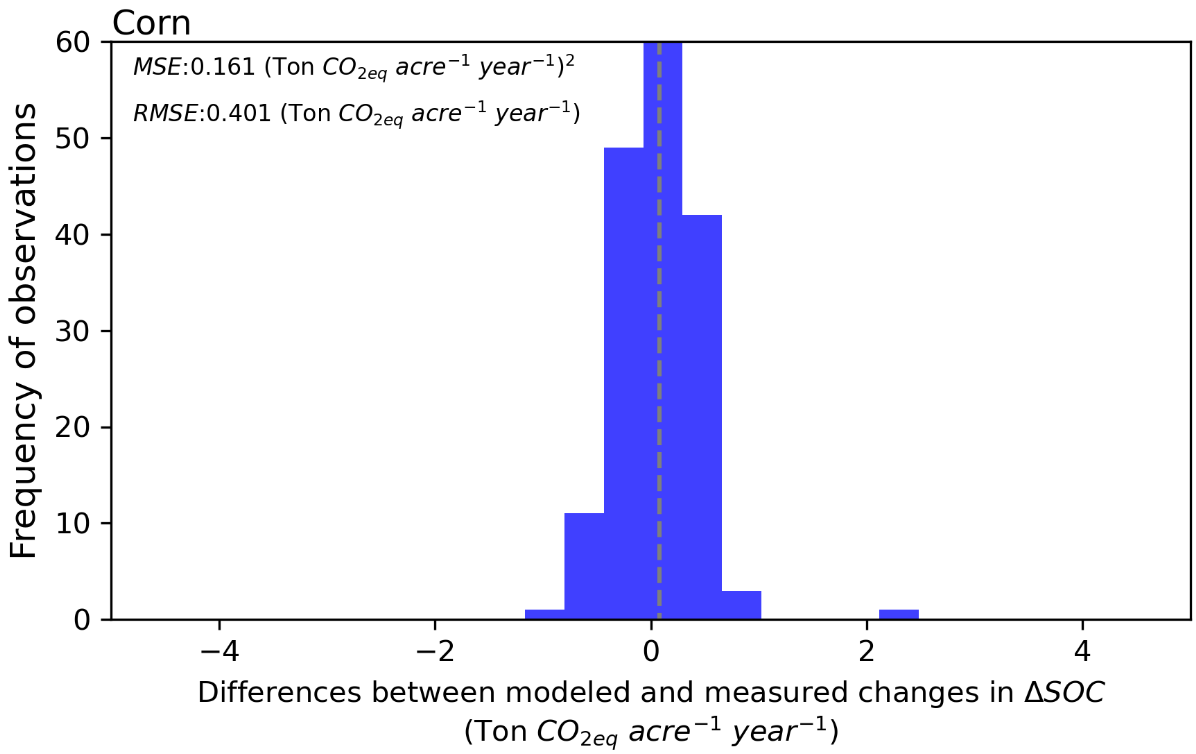

Figure G10. Histogram of the differences between modeled and measured practice change effects on SOC changes in response to practice changes in corn fields. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.6 Model prediction error for practice change effects on SOC changes in soybean fields

The mean square error is calculated as 0.107 (Ton $CO_{2,\,eq}$ $acre^{-1}$ $year^{-1}$)^2 and the confidence coverage rate is 97.1% for the modeled practice change effects on SOC changes in soybean fields.

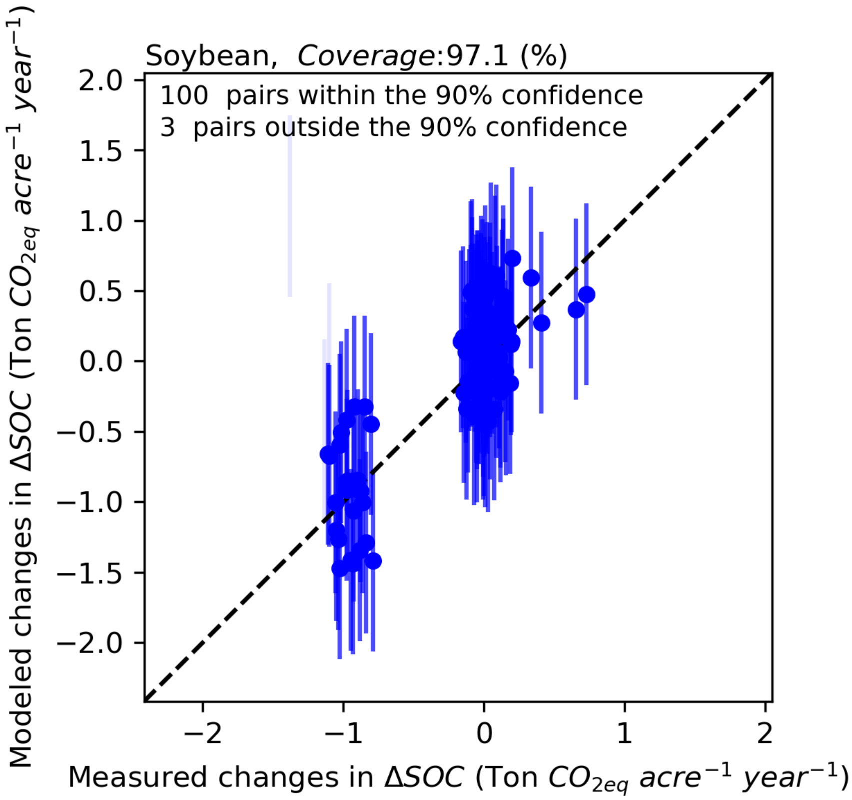

Figure G11. Modeled versus measured practice change effects on SOC changes in response to practice changes in soybean fields. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

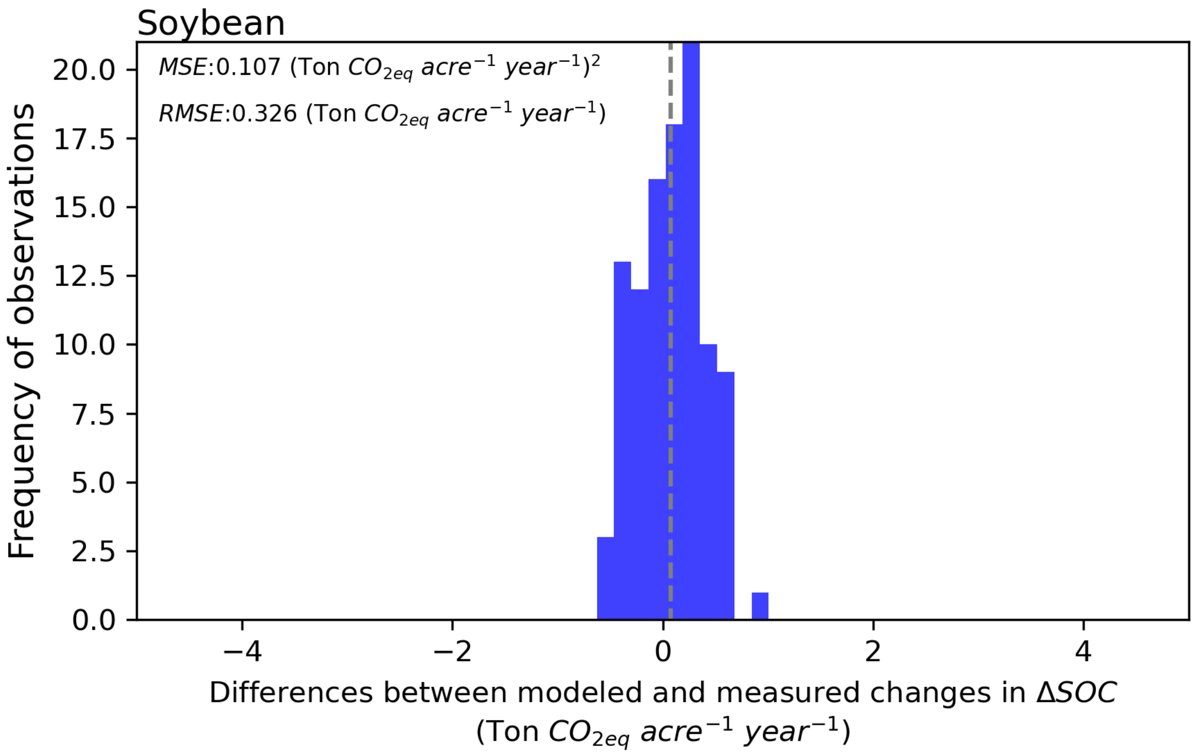

Figure G12. Histogram of the differences between modeled and measured practice change effects on SOC changes in response to practice changes in soybean fields. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.7 Model prediction error for practice change effects on direct N_2O emissions under CROPPING practices

The mean square error is calculated as 0.018 (Ton $CO_{2,eq}$ $acre^{-1})^2$ and the confidence coverage rate is 93.5% for the modeled practice change effects on direct N_2O emissions under CROPPING practices.

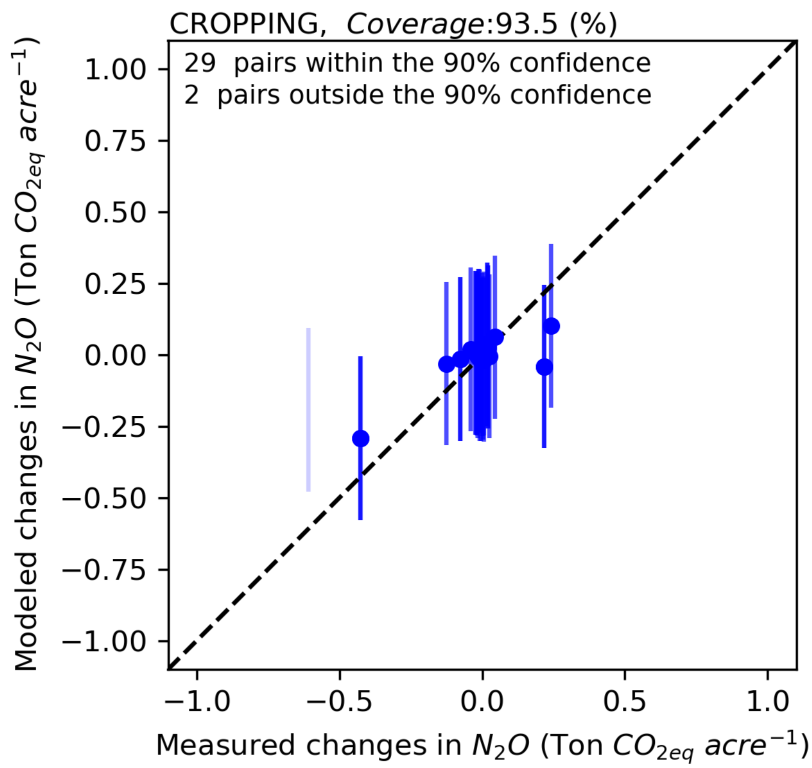

Figure G13. Modeled versus measured practice change effects on direct N_2O emissions in response to cropping practice changes. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

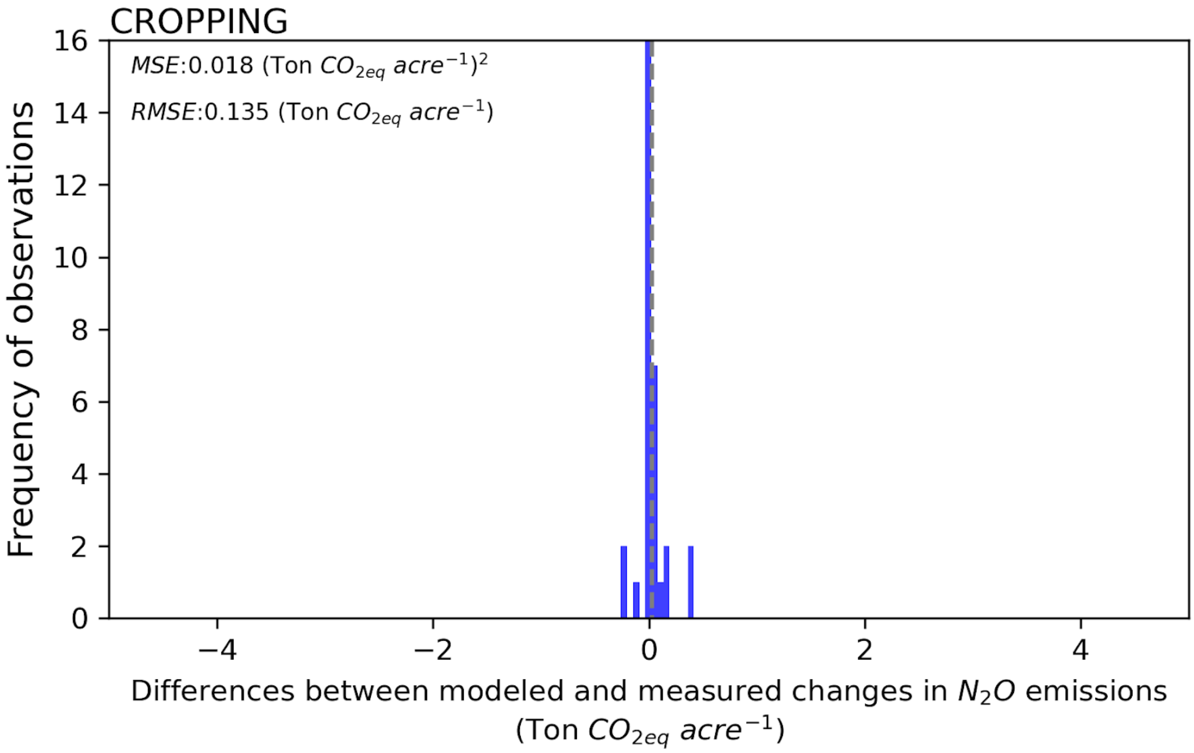

Figure G14. Histogram of the differences between modeled and measured practice change effects on direct N_2O emissions in response to cropping practice change. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.8 Model prediction error for practice change effects on direct N_2O emissions under DISTURBANCE practices

The mean square error is calculated as 0.007 (Ton $CO_{2,eq}$ acre^{-1})^2 and the confidence coverage rate is 100.0% for the modeled practice change effects on direct N_2O emissions under DISTURBANCE practices.

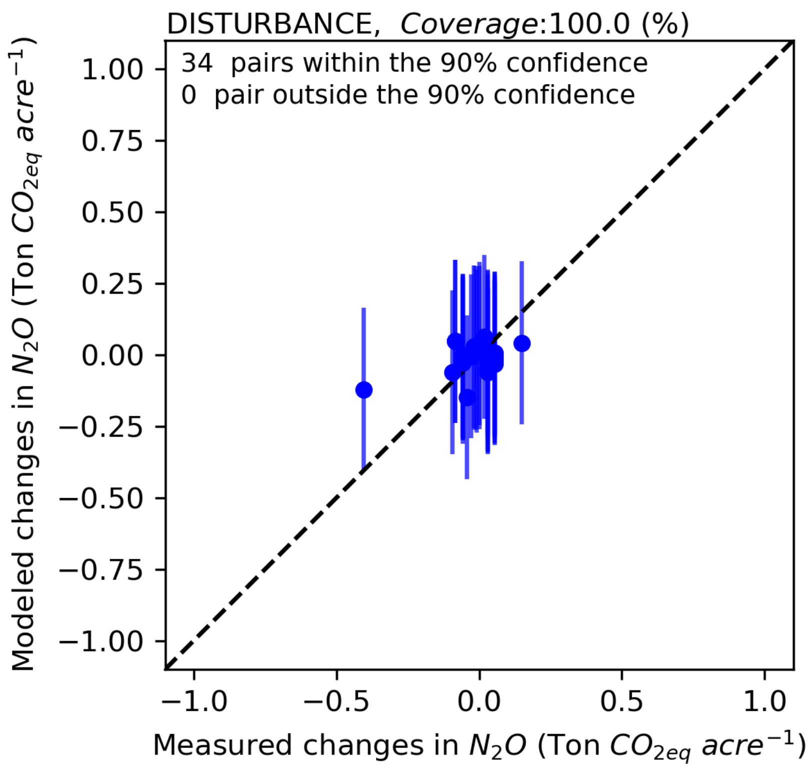

Figure G15. Modeled versus measured practice change effects on direct N2O emissions in response to disturbance practice changes. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

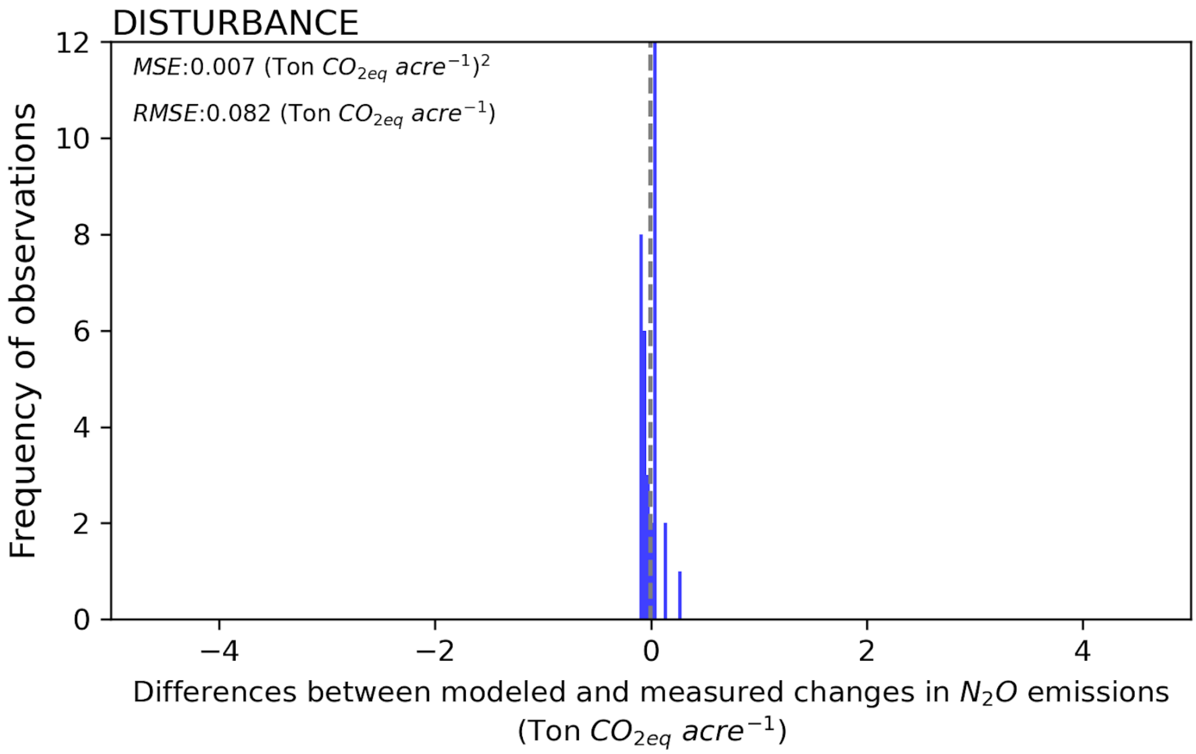

Figure G16. Histogram of the differences between modeled and measured practice change effects on direct N2O emissions in response to disturbance practice change. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.9 Model prediction error for practice change effects on direct N_2O emissions under FERTILIZER practices

The mean square error is calculated as 0.034 (Ton $CO_{2,eq}$ acre^{-1})^2 and the confidence coverage rate is 93.2% for the modeled practice change effects on direct N_2O emissions under FERTILIZER practices.

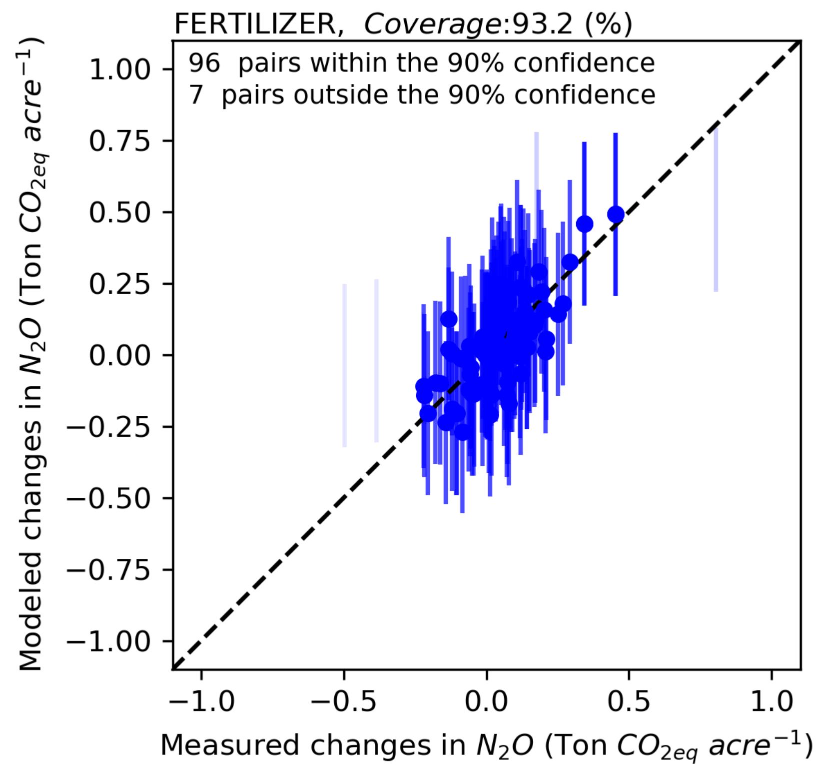

Figure G17. Modeled versus measured practice change effects on direct N_2O emissions in response to fertilizer practice changes. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

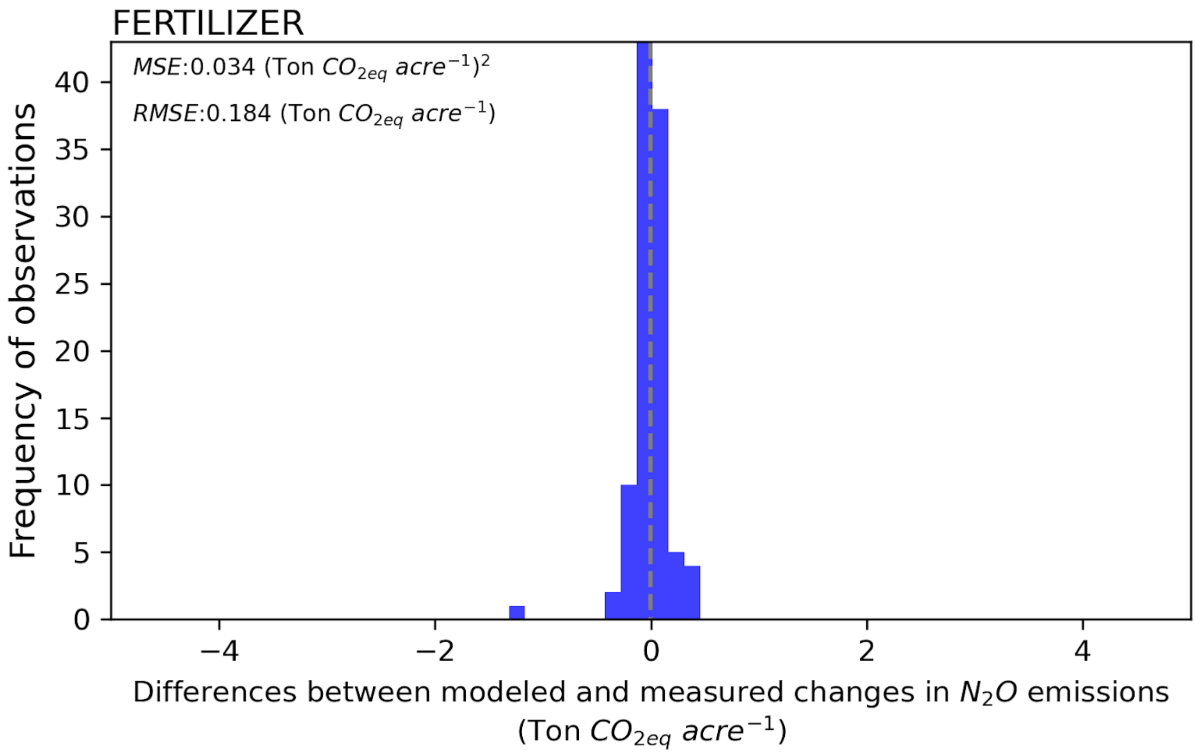

Figure G18. Histogram of the differences between modeled and measured practice change effects on direct N_2O emissions in response to fertilizer practice change. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.10 Model prediction error for practice change effects on direct N_2O emissions under ORGANIC practices

The mean square error is calculated as 0.147 (Ton $CO_{2,eq}$ $acre^{-1}$)^2 and the confidence coverage rate is 57.1% for the modeled practice change effects on direct N_2O emissions under ORGANIC practices.

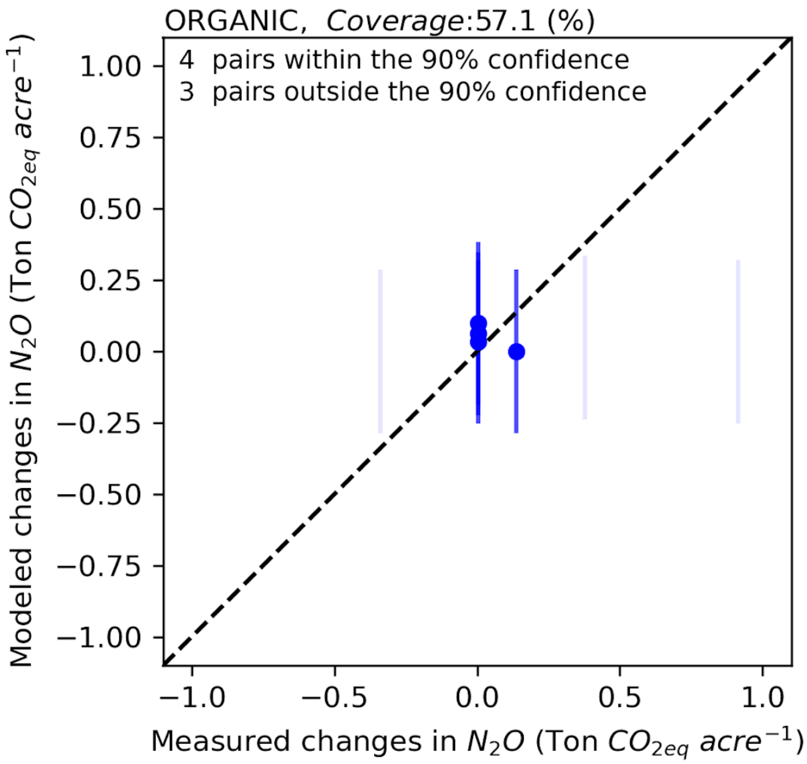

Figure G19. Modeled versus measured practice change effects on direct N_2O emissions in response to organic practice changes. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

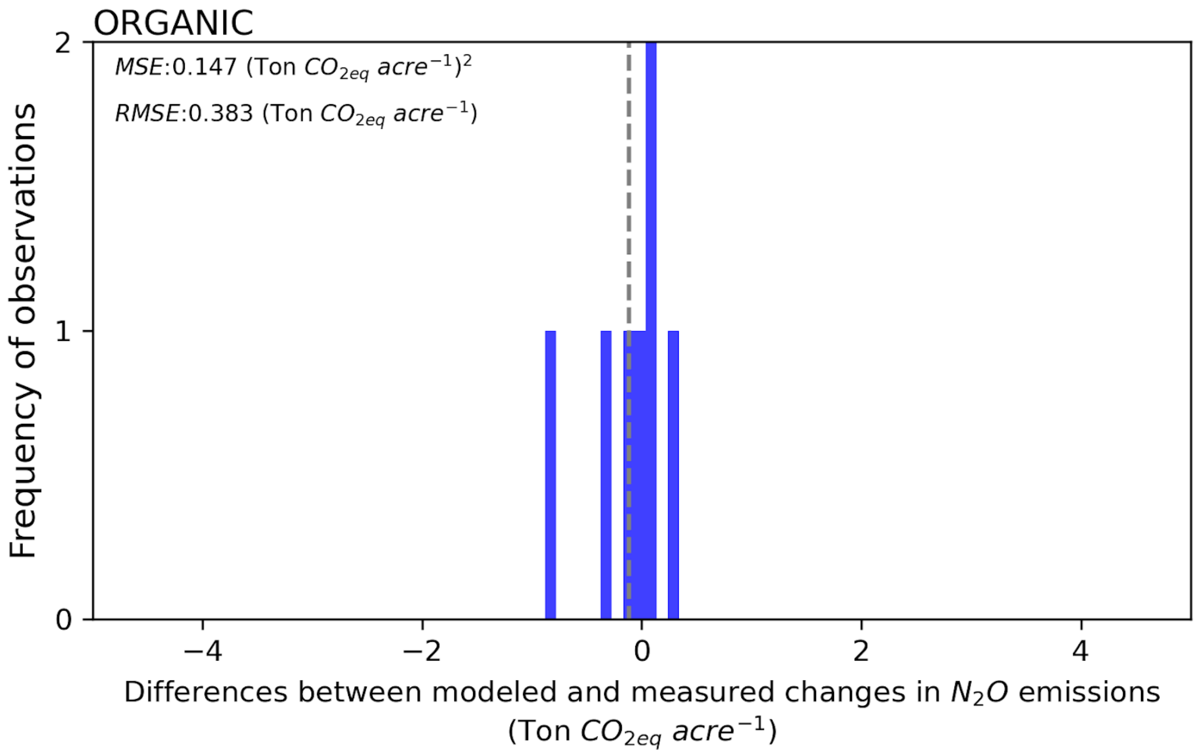

Figure G20. Histogram of the differences between modeled and measured practice change effects on direct N_2O emissions in response to organic practice change. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.11 Model prediction error for practice change effects on direct N_2O emissions in corn fields

The mean square error is calculated as 0.038 (Ton $CO_{2,eq}$ $acre^{-1})^2$ and the confidence coverage rate is 91.6% for the modeled practice change effects on direct N_2O emissions in corn fields.

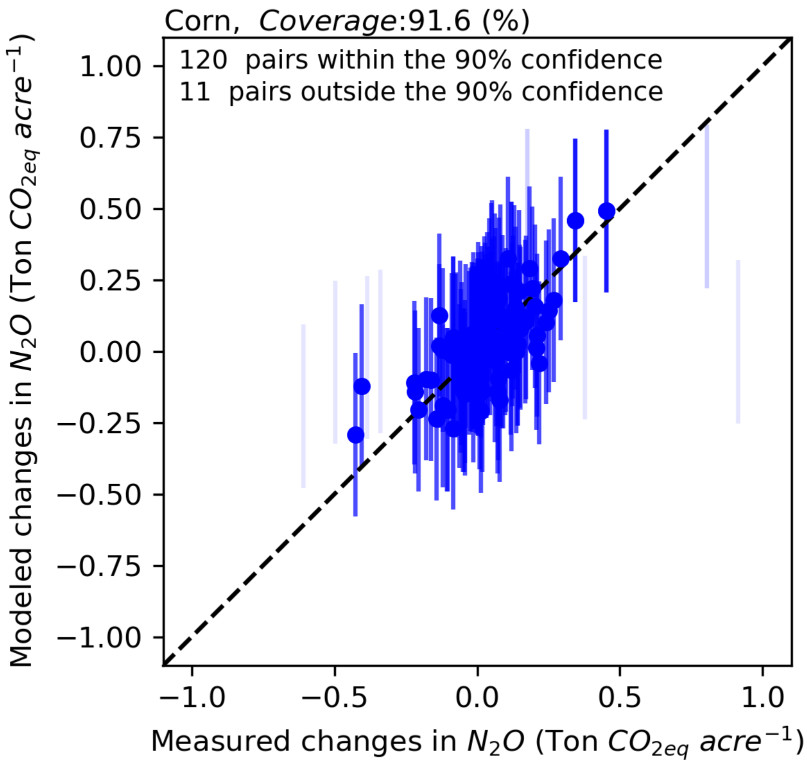

Figure G21. Modeled versus measured practice change effects on direct N_2O emissions in response to practice changes in corn fields. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

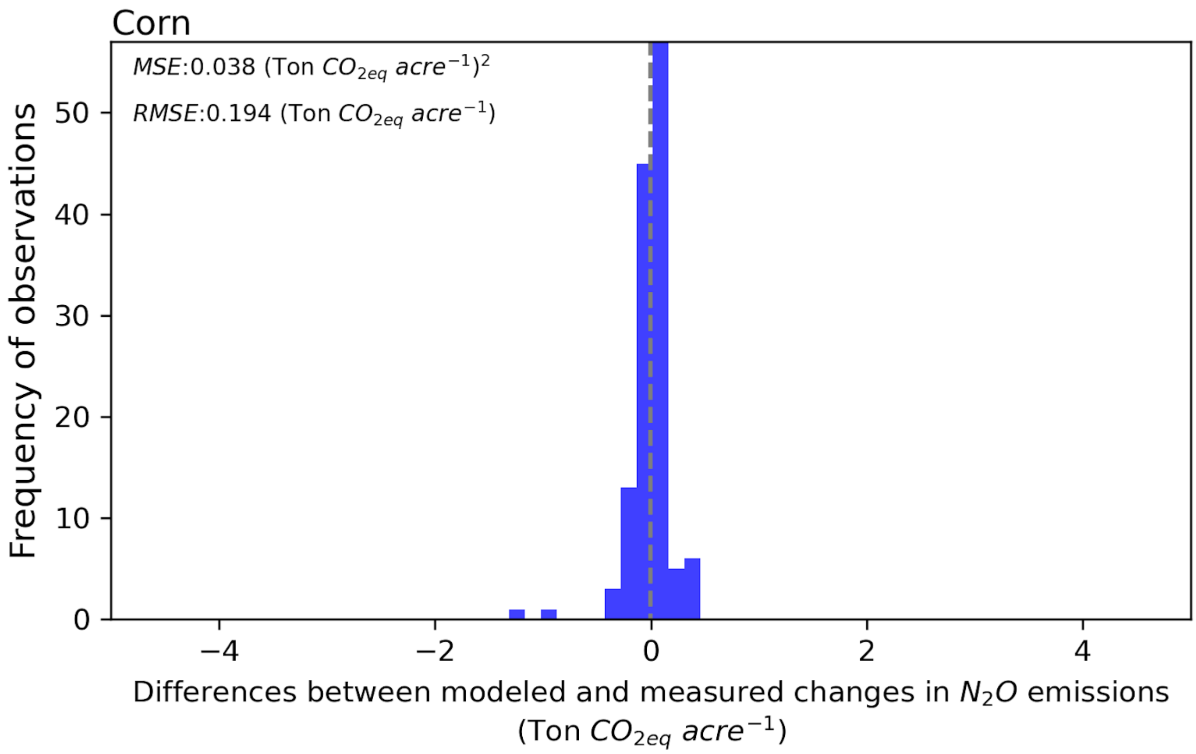

Figure G22. Histogram of the differences between modeled and measured practice change effects on direct N_2O emissions in response to practice changes in corn fields. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

G.12 Model prediction error for practice change effects on direct N_2O emissions in soybean fields

The mean square error is calculated as 0.009 (Ton $CO_{2,eq}$ $acre^{-1}$)^2 and the confidence coverage rate is 97.7% for the modeled practice change effects on direct N_2O emissions in soybean fields.

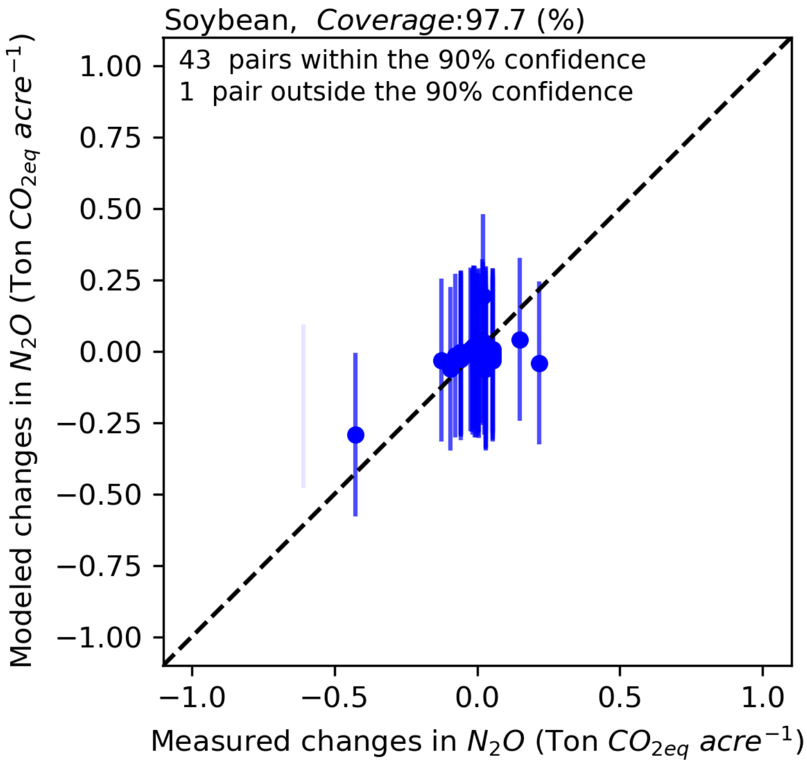

Figure G23. Modeled versus measured practice change effects on direct N_2O emissions in response to practice changes in soybean fields. The black dashed line represents the one-to-one line between measurements and predictions. Error bars show 90% prediction intervals, which overlaps with the one-to-one line when predictions meet the SEP Model Requirements. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.

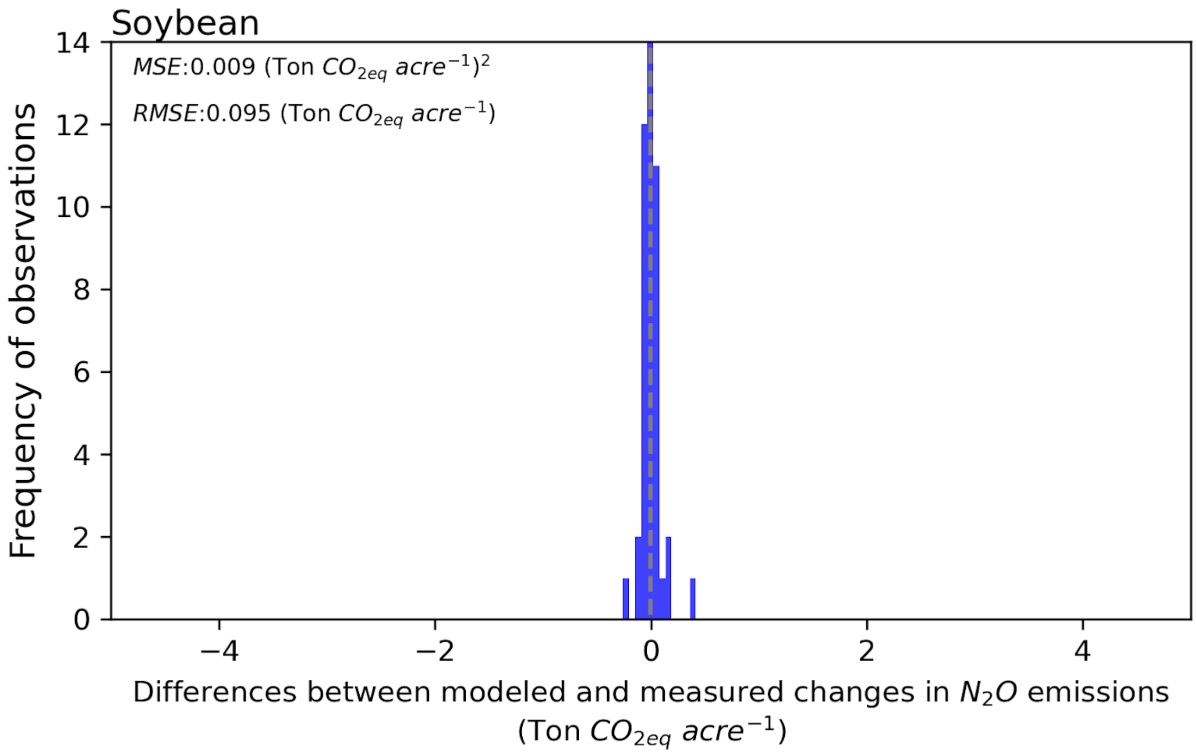

Figure G24. Histogram of the differences between modeled and measured practice change effects on direct N_2O emissions in response to practice changes in soybean fields. The dashed line represents the mean value among the observation dataset. MSE and RMSE stand for mean squared error and root mean squared error, respectively. The treatment pairs were generated by averaging replicates within the same fields under the same treatment.