https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

# Ensemble Agroecosystem Modeling Enhances Predictions of Crop Yields and Soil Carbon Across the United States

Sagar Gautam*[1,2], Chang Gyo Jung[1], Umakant Mishra[1,2], Rattan Lal[3], Klaus Lorenz[3], Jinyun Tang[4], DeAnn Ricks Presley[5], Alan J. Franzluebbers[6]

[1] Bioscience Division, Sandia National Laboratory, Livermore, CA 94550, United States
[2] Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA 94608, United States
[3]The Ohio State University, Rattan Lal Center for Carbon Management and Sequestration, Columbus, OH 43210, United States
[4]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States
[5]Department of Agronomy, Kansas State University, Manhattan, KS United States
[6]USDA-Agricultural Research Service, Raleigh, NC 27606, United States

**Corresponding author:** sgautam@sandia.gov

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

**Abstract.** Accurately estimating crop yields and soil organic carbon (SOC) dynamics is essential for agricultural planning, carbon accounting, and sustainable land management. However, process-based agroecosystem models often produce divergent estimates due to variations in model structure, parameterization, and underlying assumptions. In this study, we developed a multi-model ensemble framework that integrates three widely used process-based models-Daily Century (DAYCENT), DeNitrification DeComposition (DNDC), and Ecosystem model (ECOSYS)-to simulate crop yields and SOC stock changes (0-30 cm) across cultivated lands of the continental United States (CONUS) at 4 $km^2$ spatial resolution. Each model was parameterized using harmonized environmental, soil, and management datasets and evaluated using observed crop yields from the National Agricultural Statistics Service and measured SOC data from the Rapid Carbon Assessment. For the baseline period (2014-2023) under conventional corn-soybean rotation, the ensemble mean showed strong agreement with observations (corn: 7.7 vs. 8.5 Mg $ha^{-1}$, RMSE = 3.0; soybean: 2.5 vs. 3.0 Mg $ha^{-1}$, RMSE = 1.0), while simulated SOC stocks (5.5 vs. 4.8 kg C $m^{-2}$, RMSE = 2.5) closely matched measured data. Spatially, the ensemble model projected SOC gains in the Midwest and Southeastern regions and losses in the Great Plains and Western United States, underscoring the importance of region-specific management practices. Overall, the ensemble approach improved predictive accuracy and reduced uncertainty relative to individual models, providing a scalable pathway for robust, data driven assessments of soil carbon and crop productivity across U.S. agroecosystems.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

## 1 Introduction

Agroecosystems occupy a substantial portion of the global landscape, with agricultural land covering about 47.7 million km$^2$ (36.9%) of the global land area in 2022 (FAO, 2025). These system provide food, feed, fiber, and fuel, and also deliver ecosystem services (ESs) such as soil carbon (C) sequestration, water retention, and nutrient cycling (Robertson et al., 2014). As anthropogenic pressures intensify through consumption patterns, resource demand and weather variability, management of agroecosystems has become essential to maintain productivity and ecological balance. Enhancing soil organic carbon (SOC) sequestration is one of the key benefits of improved management, contributing to climate solutions and long-term soil health maintenance (Paustian et al., 2016).

Among agricultural practices, crop rotation particularly the inclusion of legumes such as soybean (*Glycine max* (L.) Merr.) is widely recognized for its positive effects on SOC accumulation (King & Blesh, 2018; Kumar et al., 2018). Legume-based rotations improve soil fertility through biological nitrogen fixation (BNF), reduce reliance on synthetic fertilizers, and enhancing carbon inputs from crop residues and root exudates (Drinkwater et al., 1998; Meena & Lal, 2018). They also stimulate microbial activity and stabilize SOC stocks compared to monocultures or less diverse crop rotations (Davis et al., 2012; Xiao et al., 2017). These benefits highlight the importance of understanding how site-specific management strategies and crop rotation influence SOC dynamics across diverse ecoregions and soil conditions.

Several process-based agroecosystem models have been developed to simulate the long-term effects of agricultural practices on SOC dynamics and crop productivity. Among the widely used models are the Daily Century (DAYCENT) model, DeNitrification-DeComposition (DNDC) model, and the Ecosystem Simulation Model (ECOSYS)  (Grant, 1998; C. Li et al., 1992; Parton et al., 1994). Each of these models provides unique structural features and strengths representing plant-soil-atmosphere interactions. DAYCENT a daily time-step version of the CENTURY model simulates C and N fluxes with-in plant-soil systems, incorporating processes such as plant growth, organic matter (OM) decomposition, and soil emissions. It is extensively used for evaluating SOC dynamics and nitrous oxide ($N_2O$) emissions under different land uses and climate conditions (SJ Del Grosso et al., 2005). DNDC, focuses on coupled C and N biogeochemistry,  simulate key emissions including methane ($CH_4$), carbon dioxide ($CO_2$), and $N_2O$ (C. Li et al., 1992). ECOSYS is a more mechanistic model that offers detailed representations of plant physiology, root growth, soil physics, microbial dynamics, and management practices, making it well-suited for capturing complex feedbacks in high-resolution agroecosystem analyses (Grant et al., 2001).

Given the strengths of individual models, there remains substantial variability in predicted crop yields and SOC dynamics due to differences in model structure, parameterization, and underlying assumptions (Smith et al., 1997; Tao et al., 2018). Consequently, recent studies increasingly employ multi-model ensemble approaches to represent uncertainty and enhance predictive reliability. Ensemble modeling leverages the complementary strengths of individual models to capture a range of plausible system responses under diverse climate and management conditions, thereby improving the robustness of land management and environmental assessments (Basso et al., 2025; Martre et al., 2015). The Agricultural Model Intercomparison and Improvement Project (AgMIP) exemplifies this progress, using ensemble to evaluate model sensitivities and parameter uncertainties in crop yield projections (Rosenzweig et al., 2014).

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

However, most model ensemble studies have focused primarily on crop productivity or provided limited evaluation of SOC dynamics, often constrained by small spatial scales (Basso et al., 2025; Rosenzweig et al., 2014). Furthermore, few studies have combined mechanistic models such as ECOSYS with semi-empirical models like DAYCENT and DNDC, despite their widespread use in soil carbon measurement, reporting, and verification (MRV) and carbon credit calculation. This gap limits comprehensive evaluations of how integrated management strategies influence both agronomic outcomes and ecosystem services (ESs).

In this work, we develop and benchmark an ensemble agroecosystem modeling framework integrating DAYCENT, DNDC and ECOSYS to predict crop yields (corn and soybean) and SOC dynamics across the continental United States (CONUS). This framework bridges agronomic, environmental, and computational disciplines, enabling improved carbon accounting and sustainability assessments at continental scales. By combining models with distinct process representations, we assess both consistency and divergence of predictions and identify areas of model convergence that inform robust management strategies. Unlike previous efforts, this study emphasizes the spatial variability in multi modal agreement across U.S. croplands, providing valuable insights into the design of resilient agroecosystems under weather variability. Furthermore, the results provide actionable information for policymakers, startups, and stakeholders engaged in carbon MRV and market-based sustainability initiatives.

The specific objectives of this study were to: (a) develop an ensemble agroecosystem modeling framework for CONUS croplands; (b) simulate crop yields and SOC stock changes under baseline conditions from 2014 to 2023; and (c) compare model outputs from DAYCENT, DNDC, and ECOSYS to characterize variability and uncertainty in predicted yields and SOC stocks. We hypothesized that: 1) the ensemble of DAYCENT, DNDC, and ECOSYS would capture the uncertainty of SOC stock change projections; and 2) ensemble mean corn and soybean yields would align more closely with observations while exhibiting spatially heterogeneous responses linked to geographic and environmental gradients.

## 2 Methods

### 2.1 Study area and model input

This study encompassed three primary land cover types (croplands, pastures, and grasslands) identified using National Land Cover Database (NLCD) (USGS, 2023). The daily precipitation and minimum and maximum air temperature data for the 30-year period (1994–2023) were obtained from Daily Surface Weather and Climatological Summaries (DAYMET) (Thornton et al., 2016). The key soil parameters for model initialization included multilayer soil texture, bulk density, soil moisture retention characteristics, hydraulic conductivity, soil organic carbon (SOC), and pH, derived from the Soil Survey Geographic Database (SSURGO) (Staff, 2023). The study simulated a corn-soybean rotation spanning 2014 to 2023 to evaluate crop management impacts on soil biogeochemical processes. Nitrogen (N) fertilizer was applied at a rate of 185 kg N/ha during the corn year following planting. Tillage operations included a spring chisel plow of 30 cm and a fall field cultivator to 20 cm, representing conventional management practices. In contrast, soybean cultivation followed a no-till (NT) approach with minimal soil disturbance during the growing season. This rotation setup reflects typical Midwestern U.S.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

agricultural practices and provides a consistent framework for evaluating long-term SOC dynamics and crop productivity across multiple models. The schematic overview of the modeling framework and workflow is presented in Figure. 1.

**2.2 Process model description and setup:**

DAYCENT is a daily time step, process-based agroecosystem model developed to simulate the biogeochemical cycles at a point scale (S Del Grosso et al., 2002). It is daily version of the CENTURY model (Parton et al., 1998). It can simulate exchange of carbon and soil nutrients [N, phosphorus (P), and sulfur (S)] between the atmosphere and terrestrial ecosystems, along with soil water and temperature dynamics (Schimel et al., 2001). In the DAYCENT model, SOC is predicted based on OM decomposition and C cycling among multiple soil pools, including surface litter (structural and metabolic),and fast, slow and passive C pools (S Del Grosso et al., 2002; Parton et al., 1998). Carbon inputs arise from plant residues and organic amendments, while decomposition is driven by soil temperature, moisture, and texture. DAYCENT simulates daily C fluxes between soil pools, accounting for mineralization, immobilization, and heterotrophic respiration losses (Parton et al., 1994). Additionally, DAYCENT also represent various management practices including tillage, crop rotation, residue retention, and fertilization. Crop yield is simulated through net primary production (NPP) as a function of sunlight, temperature, soil water, and nutrients. The assimilated carbon in partitioned among plant components including grain, roots and shoots based on their C:N ratios. The final yield is determined using the harvest index (HI), which defines the proportion of aboveground biomass converted into harvestable product, such as grain. This HI is crop specific and sensitive to environmental and stress conditions, enabling DAYCENT to represent realistic interactions among biophysical processes and management practices. The model was initialized with a long-term spin-up (1000 to 4,000 years) to achieve the steady-conditions for soil nutrient pools, representing pre-disturbance natural land conditions prior to human intervention. For the equilibrium simulations, ecoregion-specific native grass where grown, based on data from the US Environmental Protection Agency Inventory of US GHG emissions and sinks (USEPA, 2015). Historical land use and land cover data were compiled from historical datasets (Ogle et al., 2010). Historical management practices from initial tillage to the modern agricultural period was compiled at the Major Land Resource Area  (MLRA) level from multiple historical data sources (Ogle et al., 2010). Contemporary agricultural management was represented based on the multiyear analysis of the National Land Cover Database. Additionally, historical database was used to estimate the average annual nitrogen fertilizer rates for each agricultural crop across different ecoregion (USEPA, 2005).

DNDC is a process based soil biogeochemical model developed to simulate C and N cycling in agricultural ecosystems (C. Li et al., 1992; C. Li et al., 1994). It integrates key soil physical and chemical properties (e.g., soil texture, bulk density and soil pH) to represent soil biogeochemical processes. DNDC also includes a detailed plant growth module that simulates the entire crop life cycle, linking plant development with soil C and N dynamics. The model partitions the net assimilated C among leaves, stems, roots, and grain according to the crop's phenological stage. Final crop yield is determined at the end of the simulation as the total biomass allocated to the grain at physiological maturity. Due to its explicit representation of plant growth and decomposition processes, DNDC is widely used to evaluate the effects of agricultural

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

management practices on crop yields, SOC dynamics and soil emissions (Ding et al., 2023; Jiang et al., 2021).

ECOSYS is a process-based ecosystem model grounded in biophysical and biochemical principles that integrates water, energy, C and nutrient cycles (Grant et al., 2001). It explicitly represents interactions among soil, plant, and atmospheric components with high mechanistic details. Due to its comprehensive process representation, including plant photosynthesis and hydraulics, dynamic C and nutrient allocations in plants and microbes, soil biogeochemistry, and agricultural management practices, ECOSYS is a robust tool for simulating agroecosystem responses to diverse environmental and management conditions (Grant et al., 2006; Shekoofa et al., 2021; Zhou et al., 2021). The model incorporates detailed physical, chemical and biological processes to simulate plant-microbe-soil-atmosphere interactions across diverse soil types, climatic regimes, and land-use practices. It explicitly models nutrient exchange between microbes and plants, allowing nutrient-based evaluation of management effects on crop productivity and SOC dynamics. Crop yield is calculated based on dynamic plant C allocation, which balances photosynthetic C gain and respiratory losses, and partitions assimilated biomass. Grain filling is regulated by canopy temperature, available carbon and nutrient reserves, and physiological constraints, allowing ECOSYS to capture crop yield responses to environmental stress and management variability.

### 2.3 Model Evaluation

Each of the three models was validated using county level crop yield data from the National Agricultural Statistics Service (NASS) (NASS, 2025) and site-level SOC data from the Rapid Carbon Assessment (RaCA) dataset (Wills et al., 2014) across CONUS. Model parameterization was conducted to minimize the root mean square error (RMSE) between observed and predicted values, as defined by Equation (i):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \qquad \text{(Eq.i)}$$

where, $P_i$ and $O_i$ are the predicted and observed values, respectively, and n is the total number of observations. After model evaluation, each model was updated using a calibrated parameter set, and large-scale simulations were executed on the Sandia National Laboratories high performance computing (HPC) system in a parallel environment.

## 3 Results

### 3.1 Model Evaluation of crop yield and SOC

A comparison of simulated corn yields across the CONUS croplands from three agroecosystem models and their ensemble median against observed yields from the NASS dataset showed that the ensemble model provided the closest agreement with observations (Figure. 2a). The ensemble exhibited the lowest variability and achieved the smallest root mean square error (RMSE = 3.0 Mg ha$^{-1}$) (Figure. 2a). Its median corn yield of (7.7 Mg ha$^{-1}$) closely matched the observed median of 8.5 Mg ha$^{-1}$. Among individual models, DAYCENT and DNDC performed moderately well, producing RMSEs of 3.8 Mg ha$^{-1}$ and 5.0 Mg ha$^{-1}$, and median yields of 9.4 Mg ha$^{-1}$ and 8.7 Mg ha$^{-1}$, respectively. In contrast, ECOSYS exhibited the greatest deviation, with an RMSE of 4.1 Mg ha$^{-1}$ and a median yield of only 4.4 Mg. ha$^{-1}$, reflecting a consistent tendency to underpredict yields and higher variability across the CONUS region (Figure. 2a).

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

For soybean yields, the ensemble model again demonstrated the best performance, yielding the lowest root mean square error (1.0 Mg ha$^{-1}$) and most accurate representation of the observed yield distribution (Figure. 2b). The ensemble median soybean yield (2.5 Mg ha$^{-1}$) was in close agreement with the observed median (3.0 Mg ha$^{-1}$). Among the individual models, DAYCENT and DNDC yielded RMSEs of 1.3 Mg ha$^{-1}$ and 2.1 Mg ha$^{-1}$ with median yields of 3.2 Mg ha$^{-1}$ and 2.7 Mg ha$^{-1}$, respectively (Figure. 2b). ECOSYS again underpredicted soybean yields, with an RMSE and a median yield of 1.3 Mg ha$^{-1}$. Overall, the ensemble approach minimized prediction errors and more accurately represented the central tendency of observed yields compared with individual models. These results highlight the advantage of multi-model ensemble frameworks in enhancing prediction accuracy and reducing uncertainty in large-scale crop yield simulations.

A comparison of simulated SOC stocks (0-30 cm) from three agroecosystem models (DAYCENT, ECOSYS, and DNDC) and their ensemble against observations from the RaCA dataset revealed distinct differences in model performance (Figure. 3). The RaCA observations indicated a median SOC stock of 4.8 kg C m$^{-2}$ across CONUS for 0-30 cm layer. Among the individual models, DAYCENT performed most consistently with observations, producing a median SOC of 5.5 kg C m$^{-2}$ and an a RMSE of 2.7 kg C m$^{-2}$. ECOSYS underestimated SOC, yielding a median of 3.2 kg C m$^{-2}$ and an RMSE of 3.0 kg C m$^{-2}$, suggesting limitations in its representation of SOC accumulation processes. In contrast, DNDC consistently overestimated SOC, with median of 7.7 kg C m$^{-2}$ and the largest RMSE (4.9 kg C m$^{-2}$) among the three models (Figure. 3). The ensemble approach effectively reduced model-specific biases, producing a median SOC estimate of 5.5 kg C m$^{-2}$ with the lowest RMSE (2.5 kg C m$^{-2}$) across CONUS (Figure. 3). Collectively, these results demonstrate that the ensemble modeling enhances agreement with observed data and improves the reliability of SOC stock estimates at regional and national scales.

### 3.2 Model prediction of corn yield

During the 10-year simulation period (2014-2023), DAYCENT estimated annual corn yields ranging from 1.0 to15.3 Mg ha$^{-1}$, with a mean of 8.8 Mg ha$^{-1}$. DNDC projected yields between 1 and 15 Mg ha$^{-1}$ (mean = 8.7 Mg ha$^{-1}$), whereas ECOSYS computed a similar range but a substantially lower mean of 5.95 Mg ha$^{-1}$ (Figure. 4). All three models captured the broad spatial pattern of high corn productivity across the U.S. Midwest Corn Belt, particularly in central Iowa, Illinois, and Indiana. Regional difference among models reflected contrasts in structure representation and parameterizations. DNDC projected relatively higher yields in the Mississippi River Basin and the Southeast U.S., while DACYENT simulated more moderate yields in the same regions (Figure. 4). By contrast, ECOSYS estimated comparatively higher productivity across parts of the Northern Plains and Northwestern U.S. The ensemble yield map effectively moderated these model-specific extremes, generating a spatially coherent pattern that preserved agreement within high-yield zones while diminishing the influence of outlier projections. Overall, the ensemble framework produced robust regional yield estimates by integrating model variability, leveraging complementary strengths, and redacting individual model biases.

### 3.3 Model prediction of soybean yield

From 2014 to 2023, DAYCENT simulated annual soybean yields ranging from 0.1 to 5 Mg ha$^{-1}$ with a mean of 2.8 Mg ha$^{-1}$. DNDC projected yields within a similar range (0.1-5 Mg ha$^{-1}$) and a mean of 2.7 Mg ha$^{-1}$, while ECOSYS estimated a slightly broader range (0.01 to 6 Mg ha$^{-1}$) with

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

a mean of 2.5 Mg ha$^{-1}$ (Figure. 5). All three models reproduced the general spatial pattern of highest soybean productivity across the Midwestern and central U.S., particularly in major soybean producing regions such as Iowa, Illinois, and Indiana. Model-specific variability emerged in both the spatial extent and magnitude of yield predictions. DNDC projected high yields concentrated in central Iowa and surrounding areas, whereas DAYCENT and ECOSYS predicted more spatially uniform yield distributions across the Corn Belt (Figure. 5). Additionally, DAYCENT simulated comparatively higher yields across parts of the Southeast U.S.

The ensemble yield map integrated outputs from all three models, producing a harmonized spatial distribution characterized by a concentrated high-yield zone in the central Midwest and a gradual decline in productivity toward coastal and arid western regions. This ensemble approach provided a more coherent and robust estimate of soybean productivity across diverse agroclimatic zones, effectively smoothing local extremes and reducing model-specific outliers.

### 3.4 Model prediction of soil organic carbon change

Figure 6 illustrates the spatial distribution of projected annual SOC change (kg C m$^{-2}$ yr$^{-1}$) under a 10-year (2014-2023) conventional-till corn-soybean rotation across CONUS, based on simulations from the three agroecosystem models and their ensemble mean. Positive values represent SOC sequestration, whereas negative values indicate net C emissions. DAYCENT consistently projected widespread SOC sequestration across the Midwest and eastern U.S., with rates often exceeding 0.02 kg Cm$^2$ yr$^{-1}$ in many areas (Figure. 6). In contrast, DNDC and ECOSYS exhibited greater spatial heterogeneity, including notable zones of SOC decline in the Northern Great Plains and parts of the Midwest. DNDC simulated SOC losses surpassing 0.01 kg C m$^{-2}$ yr$^{-1}$, particularly within high-yielding zones (Figure. 6). The ensemble map integrated these model-specific responses, capturing the broad sequestration trends indicated by DAYCENT while incorporating SOC-loss zones highlighted by DNDC and ECOSYS. Overall, the ensemble projection suggested a modest net increase in SOC across much of the Corn Belt, underscoring the capacity of conventional corn-soybean systems to enhance carbon sequestration under region-specific environmental and management conditions.

## 4 Discussion

The ensemble modeling approach produced corn yield estimates across the CONUS that closely aligned with the observed data, underscoring the value of multi-model ensembles in reducing individual model biases and improving prediction accuracy (Asseng et al., 2013; Basso et al., 2025; Martre et al., 2015). The ensemble model benefited from complementary model structures, error compensation and statistical averaging, resulting in enhanced stability and performance across diverse environmental gradients. The moderate accuracy of DAYCENT and DNDC was consistent with prior evaluations demonstrating their strengths in simulating corn productivity under temperate conditions, particularly in the U.S. Midwest (Campbell et al., 2014; S Del Grosso et al., 2002; Ingraham & Salas, 2019). In contrast, ECOSYS substantially underpredicted corn yield, likely due to its complex coupling root and canopy processes, which are highly sensitive to input uncertainty and parameterization constraints in large-scale applications (Liu et al., 2024). Regional anomalies, such as DNDC's higher corn yield projections in the Southeast,

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

likely influenced by how it simulated water stress and nutrient availability under less favorable climatic conditions, as well as the linear response of crop yield to precipitation (Kaur et al., 2023). Collectively, these results supported the hypothesis that ensemble predictions provide more reliable and robust estimation of corn yield across diverse agroecological zones in CONUS.

Similarly, the ensemble model demonstrated superior performance in predicting soybean yield, achieving the lowest RMSE (1.0 Mg $ha^{-1}$) and reproducing the observed median yield (2.5 Mg $ha^{-1}$). This finding aligns with previous research highlighting the advantages of ensemble modeling in stabilizing projections across variable cropping systems (Battisti et al., 2017). The relatively strong performance of DAYCENT and DNDC reflected their ability to simulate soybean physiology and biological nitrogen fixation (BNF) processes in temperate regions. However, DNDC's spatial overestimation in Iowa is likely resulted from its limited representation of soil drainage and water balance (Helmers et al., 2012). ECOSYS underpredicted soybean yields, possibly due to its sensitivity to soil moisture dynamics and restricted calibration in drier regions. These model specific biases highlight the persistent challenges of representing complex interactions among soil texture, weather variability, and cultivar adaptability further reinforcing the utility of an ensemble approaches for yield estimation.

Ensemble projections for SOC stock changes (0-30 cm) across CONUS under a 10-year corn-soybean rotation indicated a modest net sequestration, consistent with previous finding identifying the Corn Belt as a region with high SOC sequestration potential (Grace & Robertson, 2021; Wu et al., 2025). Strong performance of DAYCENT in capturing SOC accumulation within the Midwest aligns with its established capacity to simulate long-term C cycling and residue turnover in U.S. cropping systems (Basso et al., 2025). In contrast, DNDC and ECOSYS underpredicted SOC stocks, likely due to model sensitivities to decomposition rates, soil moisture retention, and microbial activity, which are critical determinants of SOC turnover. The greater variability observed with DNDC was consistent with prior reports attributing its coarser parameterization of SOC pools and limited regional calibration, as its original design emphasis on $CH_4$ and $N_2O$ emissions rather than SOC dynamics (Zhang et al., 2014). The ensemble's intermediate median SOC and reduced RMSE illustrate its effectiveness in mitigating structural uncertainty and reproducing the SOC patterns captured in the RaCA database. These results underscore the promise of ensemble-based approaches for improving national scale carbon accounting and guiding sustainable land management strategies.

This study uses a uniform set of crop phenotypic and parameter values for corn and soybean across the entire spatial domain. While this standardization enabled large-scale simulations and computational efficiency, it inherently omitted genotype-environment interactions and region-specific cultivar adaptations. In ECOSYS, for example, crop phenology is highly sensitive to thermal time and maturity group parameters, both of which are vary significantly across U.S. regions. Using a single maturity group for corn and soybean may therefore have misrepresented region-specific growth dynamics and yield responses. Previous work have demonstrated that regional calibration of maturity group is essential for accurately reproducing phenological development (Z. Li et al., 2022). Although this approach supports continental scale consistency, it introduces limitations in representing localized cultivar adaptation and environmental responses, highlighting an important direction for future ensemble model refinement.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

## 5 Conclusions

This study advances multi-model integration in agroecosystem science by developing and benchmarking an ensemble framework that combines three process-based models, DAYCENT, DNDC, and ECOSYS, to improve large scale predictions of crop yields and soil carbon dynamics across the continental United States. The ensemble approach outperformed individual models, producing lower root mean square errors (RMSEs) and stronger agreement with observed corn and soybean yields and SOC stocks in terms of both mean values and variability. Ensemble results captured the central tendency of observed data more effectively, while generating spatially coherent yield patterns. For SOC dynamics, ensemble projections represented both sequestration potential and emission hotspots, capturing net carbon gains in the Midwest while reducing over and under estimation patterns from DNDC and ECOSYS, respectively. Overall, this study highlights the need of ensemble modeling framework for accurate regional to national scale assessment of crop productivity and soil carbon dynamics. By reducing structural uncertainty and bias, this approach strengthens the scientific basis for carbon measurement, reporting, and verification, carbon credit quantification, and regenerative agriculture initiatives. Ensemble agroecosystem modeling provides a critical foundation for advancing evidence-based land management and sustainable agricultural policy in a changing weather pattern.

## Data Availability

All data supporting the findings of this study are contained within the article.

## Acknowledgements

The Foundation for Food and Agriculture Research (Grant-ID: 22-000279; Title: Enhanced Soil Carbon Farming as a Climate Solution) provided financial support with matching contributions from Bayer Crop Science US, Corteva Agriscience, Herbert W. Hoover Foundation, Inter-American Institute for Cooperation on Agriculture (IICA), Microsoft, National Sorghum Producers, Ohio Corn & Wheat, Ohio Soybean Council, and The Ohio State University. In-kind support was provided by Agricultural Research Service of the US Department of Agriculture, Sandia National Laboratory of the US Department of Energy, Utah Department of Agriculture, and Western Geographic Science Center of the US Geological Survey. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the Foundation for Food & Agriculture Research. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. https://wp.sandia.gov/new-raa/guidance-for-publishing-a-journal-article/

**Author Contributions:** Conceptualization: S.G. and U.M.; methodology and investigation: S.G., and C.G.J.; formal analysis, visualization, and writing original draft: S.G., and C.G.J.; data

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

curation: S.G., and C.G.J.; resources: S.G., and C.G.J.; writing-review and editing: S.G., C.G.J., U.M., R.L., K.L., J.T., D.R.P., and A.J.F.; and funding acquisition: R.L.

**Competing Interest Statement:** The authors declare no competing interest.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

## References

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., et al. (2013). Uncertainty in simulating wheat yields under climate change. *Nature climate change, 3*(9), 827–832.

Basso, B., Tadiello, T., Millar, N., Robertson, G. P., Paustian, K., Maureira, F. S., et al. (2025). A multi model ensemble reveals net climate benefits from regenerative practices in US Midwest croplands. *Scientific Reports, 15*(1), 24881.

Battisti, R., Sentelhas, P. C., & Boote, K. J. (2017). Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. *Field Crops Research, 200*, 28–37.

Campbell, E. E., Johnson, J. M., Jin, V. L., Lehman, R. M., Osborne, S. L., Varvel, G. E., & Paustian, K. (2014). Assessing the soil carbon, biomass production, and nitrous oxide emission impact of corn stover management for bioenergy feedstock production using DAYCENT. *BioEnergy Research, 7*(2), 491–502.

Davis, A. S., Hill, J. D., Chase, C. A., Johanns, A. M., & Liebman, M. (2012). Increasing cropping system diversity balances productivity, profitability and environmental health.

Del Grosso, S., Ojima, D., Parton, W., Mosier, A., Peterson, G., & Schimel, D. (2002). Simulated effects of dryland cropping intensification on soil organic matter and greenhouse gas exchanges using the DAYCENT ecosystem model. *Environmental pollution, 116*, S75–S83.

Del Grosso, S., Parton, W., Mosier, A., Holland, E., Pendall, E., Schimel, D., & Ojima, D. (2005). Modeling soil CO2 emissions from ecosystems. *Biogeochemistry, 73*(1), 71–91.

Ding, W., Chang, N., Zhang, G., Kang, J., Yi, X., Zhang, J., et al. (2023). Soil organic carbon changes in China's croplands: A newly estimation based on DNDC model. *Science of the Total Environment, 905*, 167107.

Drinkwater, L. E., Wagoner, P., & Sarrantonio, M. (1998). Legume-based cropping systems have reduced carbon and nitrogen losses. *Nature, 396*(6708), 262–265.

FAO. (2025). *Agricultural land (% of land area)*. Retrieved from: https://ourworldindata.org/grapher/agricultural-land-percent-land-area

Grace, P., & Robertson, G. P. (2021). Soil carbon sequestration potential and the identification of hotspots in the eastern Corn Belt of the United States. *Soil Science Society of America Journal, 85*(5), 1410–1424.

Grant, R. (1998). Simulation of methanogenesis in the mathematical model ecosys. *Soil Biology and Biochemistry, 30*(7), 883–896.

Grant, R., Jarvis, P., Massheder, J., Hale, S., Moncrieff, J., Rayment, M., et al. (2001). Controls on carbon and energy exchange by a black spruce–moss ecosystem: testing the mathematical model ecosys with data from the BOREAS experiment. *Global biogeochemical cycles, 15*(1), 129–147.

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*(1), 235–248.

Helmers, M., Christianson, R., Brenneman, G., Lockett, D., & Pederson, C. (2012). Water table, drainage, and yield response to drainage water management in southeast Iowa. *Journal of soil and Water Conservation, 67*(6), 495–501.

Ingraham, P. A., & Salas, W. A. (2019). Assessing nitrous oxide and nitrate leaching mitigation potential in US corn crop systems using the DNDC model. *Agricultural Systems, 175*, 79–87.

Jiang, R., Yang, J., Drury, C. F., He, W., Smith, W., Grant, B., et al. (2021). Assessing the impacts of diversified crop rotation systems on yields and nitrous oxide emissions in Canada using the DNDC model. *Science of the Total Environment, 759*, 143433.

Kaur, N., Hui, D., Riccuito, D. M., Mayes, M. A., & Tian, H. (2023). Response patterns of simulated corn yield and soil nitrous oxide emission to precipitation change. *Ecological Processes, 12*(1), 17.

King, A. E., & Blesh, J. (2018). Crop rotations for increased soil carbon: perenniality as a guiding principle. *Ecological applications, 28*(1), 249–261.

Kumar, S., Meena, R. S., Lal, R., Singh Yadav, G., Mitran, T., Meena, B. L., et al. (2018). Role of legumes in soil carbon sequestration. In *Legumes for soil health and sustainable management* (pp. 109–138): Springer.

Li, C., Frolking, S., & Frolking, T. A. (1992). A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *Journal of Geophysical Research: Atmospheres, 97*(D9), 9759–9776.

Li, C., Frolking, S., & Harriss, R. (1994). Modeling carbon biogeochemistry in agricultural soils. *Global biogeochemical cycles, 8*(3), 237–254.

Li, Z., Guan, K., Zhou, W., Peng, B., Jin, Z., Tang, J., et al. (2022). Assessing the impacts of pre-growing-season weather conditions on soil nitrogen dynamics and corn productivity in the US Midwest. *Field Crops Research, 284*, 108563.

Liu, L., Zhou, W., Guan, K., Peng, B., Xu, S., Tang, J., et al. (2024). Knowledge-guided machine learning can improve carbon cycle quantification in agroecosystems. *Nature communications, 15*(1), 357.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J. W., Rötter, R. P., et al. (2015). Multimodel ensembles of wheat growth: many models are better than one. *Global change biology, 21*(2), 911–925.

Meena, R. S., & Lal, R. (2018). Legumes and sustainable use of soils. In *Legumes for soil health and sustainable management* (pp. 1–31): Springer.

NASS. (2025). *Quick Stats Database (Crop Yield Data)*. Retrieved from: https://quickstats.nass.usda.gov/

Ogle, S. M., Breidt, F. J., Easter, M., Williams, S., Killian, K., & Paustian, K. (2010). Scale and uncertainty in modeled soil organic carbon stock changes for US croplands using a process-based model. *Global change biology, 16*(2), 810–822.

Parton, W. J., Hartman, M., Ojima, D., & Schimel, D. (1998). DAYCENT and its land surface submodel: description and testing. *Global and planetary Change, 19*(1-4), 35–48.

Parton, W. J., Ojima, D. S., Cole, C. V., & Schimel, D. S. (1994). A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. *Quantitative modeling of soil forming processes, 39*, 147–167.

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G. P., & Smith, P. (2016). Climate-smart soils. *Nature, 532*(7597), 49–57.

Robertson, G. P., Gross, K. L., Hamilton, S. K., Landis, D. A., Schmidt, T. M., Snapp, S. S., & Swinton, S. M. (2014). Farming for ecosystem services: An ecological approach to production agriculture. *BioScience, 64*(5), 404–415.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., et al. (2014). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the national academy of sciences, 111*(9), 3268–3273.

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., et al. (2001). Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature, 414*(6860), 169–172.

Shekoofa, A., Safikhan, S., Snider, J. L., Raper, T. B., & Bourland, F. M. (2021). Variation in stomatal conductance responses of cotton cultivars to high vapour pressure deficit under controlled and rainfed environments. *Journal of Agronomy and Crop Science, 207*(2), 332–343.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

Smith, P., Smith, J., Powlson, D., McGill, W., Arah, J., Chertov, O., et al. (1997). A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. *Geoderma, 81*(1-2), 153–225.
Staff, S. S. (2023). *Soil Survey Geographic (SSURGO) Database* (Accessed July 16, 2025 ed.): USDA-NRCS.
Tao, F., Rötter, R. P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M. I., Semenov, M. A., et al. (2018). Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. *Global change biology, 24*(3), 1291–1307.
Thornton, P., Thornton, M., Mayer, B., Wei, Y., Devarakonda, R., Vose, R., & Cook, R. (2016). Daymet: daily surface weather data on a 1-km grid for North America, version 3. ORNL DAAC, Oak Ridge, Tennessee, USA. *USDA-NASS, 2019. 2017 Census of Agriculture, Summary and State Data, Geographic Area Series, Part 51, AC-17-A, 51*.
USEPA. (2005). *Inventory of U.S. greenhouse gas emissions and sinks: 1990–2003*.
USEPA. (2015). *Inventory of U.S. greenhouse gas emissions and sinks: 1990–2014*.
USGS. (2023). *National Land Cover Database (NLCD) 2021 land cover products*. Retrieved from: https://doi.org/10.5066/P9JZ7AO3
Wills, S., Loecke, T., Sequeira, C., Teachman, G., Grunwald, S., & West, L. T. (2014). Overview of the US rapid carbon assessment project: sampling design, initial summary and uncertainty estimates. In *Soil carbon* (pp. 95–104): Springer.
Wu, Y., Davis, E. C., & Sohngen, B. L. (2025). Crop rotation and the impact on soil carbon in the US Corn Belt. *Carbon Balance and Management, 20*(1), 6.
Xiao, X., Chen, W., Zong, L., Yang, J., Jiao, S., Lin, Y., et al. (2017). Two cultivated legume plants reveal the enrichment process of the microbiome in the rhizocompartments. *Molecular Ecology, 26*(6), 1641–1651.
Zhang, L., Yu, D., Shi, X., Xu, S., Xing, S., & Zhao, Y. (2014). Effects of soil data and simulation unit resolution on quantifying changes of soil organic carbon at regional scale with a biogeochemical process model. *PLoS One, 9*(2), e88622.
Zhou, W., Guan, K., Peng, B., Tang, J., Jin, Z., Jiang, C., et al. (2021). Quantifying carbon budget, crop yields and their responses to environmental variability using the ecosys model for US Midwestern agroecosystems. *Agricultural and Forest Meteorology, 307*, 108521.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

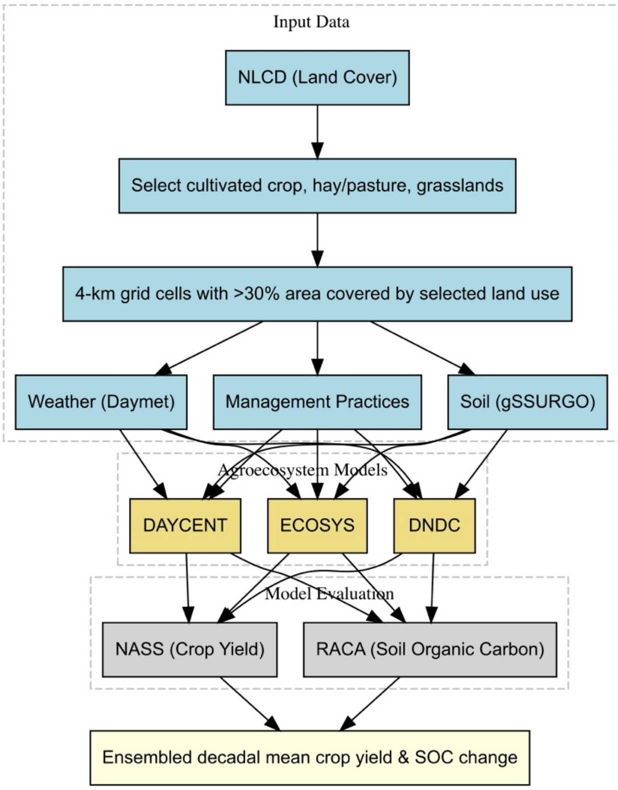


**Figure 1** Schematic representation of the workflow for large-scale application of the ensemble agroecosystem model (DAYCENT, DNDC, ECOSYS). Environmental and management inputs were derived from multiple datasets, including Daymet (daily surface weather data), SSURGO (Soil Survey Geographic Database), NASS (National Agricultural Statistics Service crop yield surveys), RaCA (Rapid Carbon Assessment soil carbon database), and NLCD (National Land Cover Database).

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

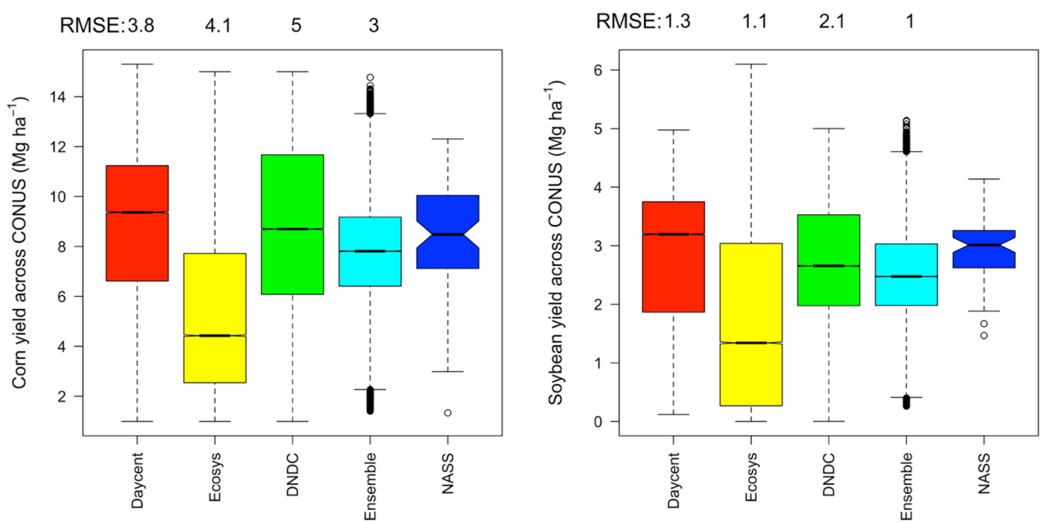


**Figure 2** Comparison of observed crop yields (Mg/ha; NASS, 2014-2023) with simulated yields and root mean square error (Mg/ha) from individual agroecosystem models and the ensemble approach for (a) corn and (b) soybean.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

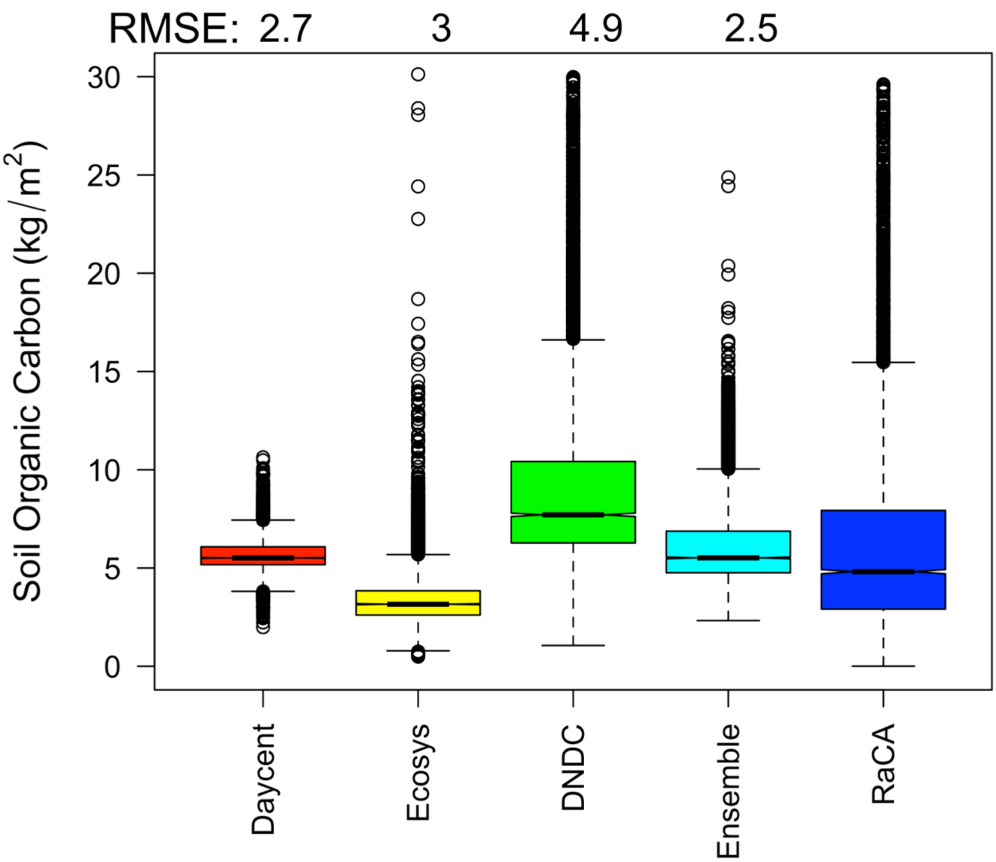


**Figure 3** Comparison of observed soil organic carbon (SOC) stocks (0-30 cm) (kg C $m^2$) from the RaCA dataset with SOC simulated and root mean square error (kg C $m^2$) by individual agroecosystem models and the ensemble approach.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

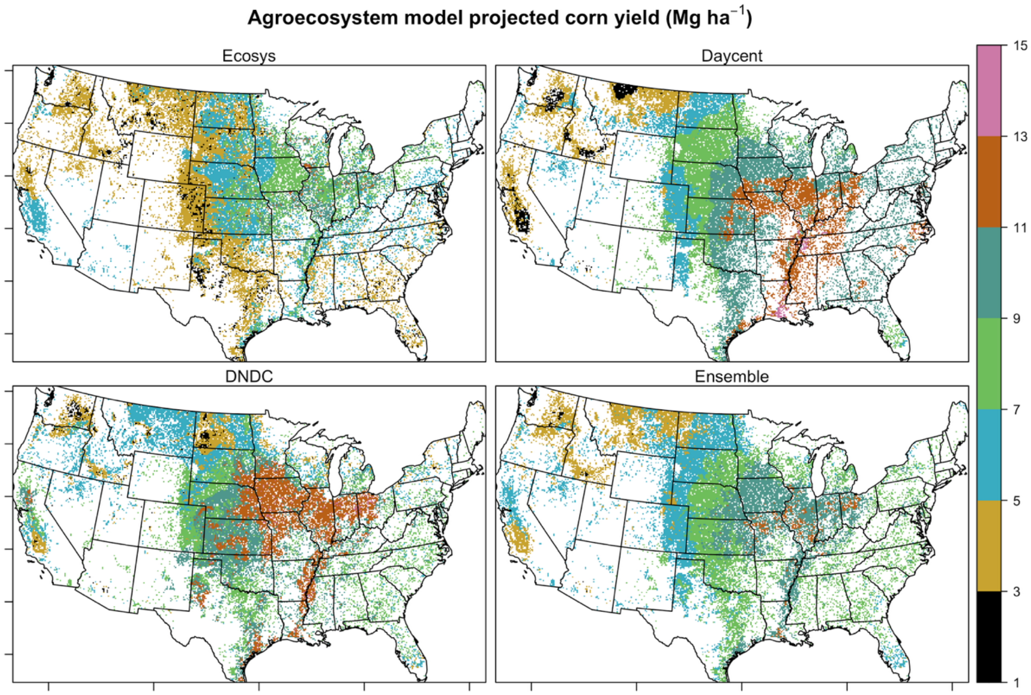


**Figure 4** Spatial distribution of annual corn yield (Mg ha$^{-1}$) across the CONUS simulated by the ensemble and individual agroecosystem models (DAYCENT, DNDC, and ECOSYS) for the 2014–2023 period.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

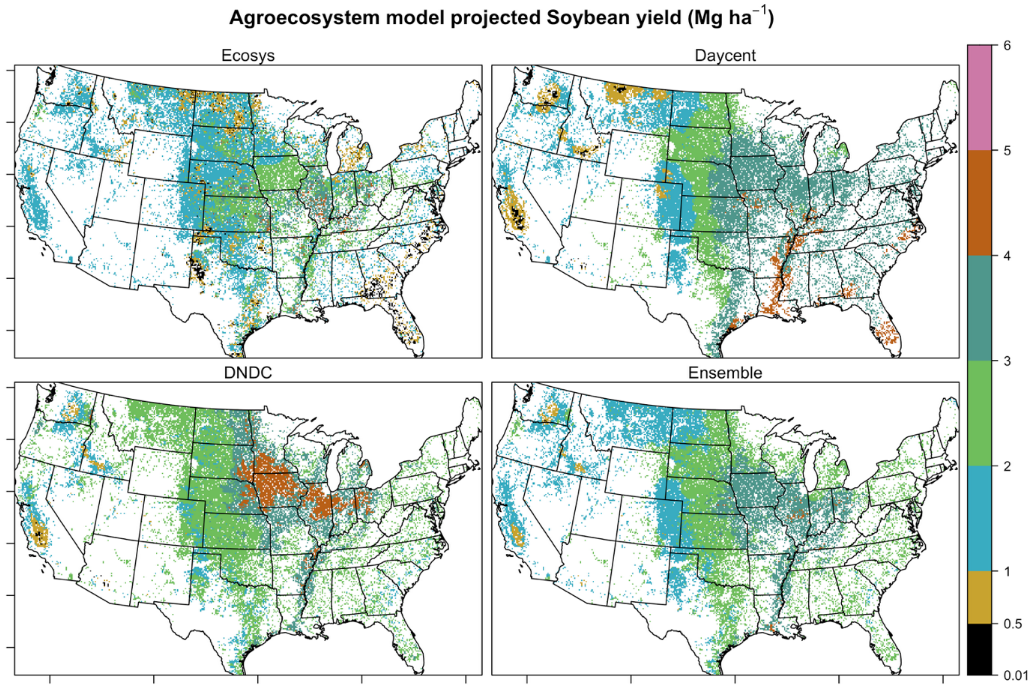


**Figure 5** Spatial distribution of annual soybean yield (Mg ha$^{-1}$) across the CONUS simulated by the ensemble and individual agroecosystem models (DAYCENT, DNDC, and ECOSYS) for the 2014–2023 period.

https://doi.org/10.5194/egusphere-2026-1094
Preprint. Discussion started: 9 March 2026
© Author(s) 2026. CC BY 4.0 License.

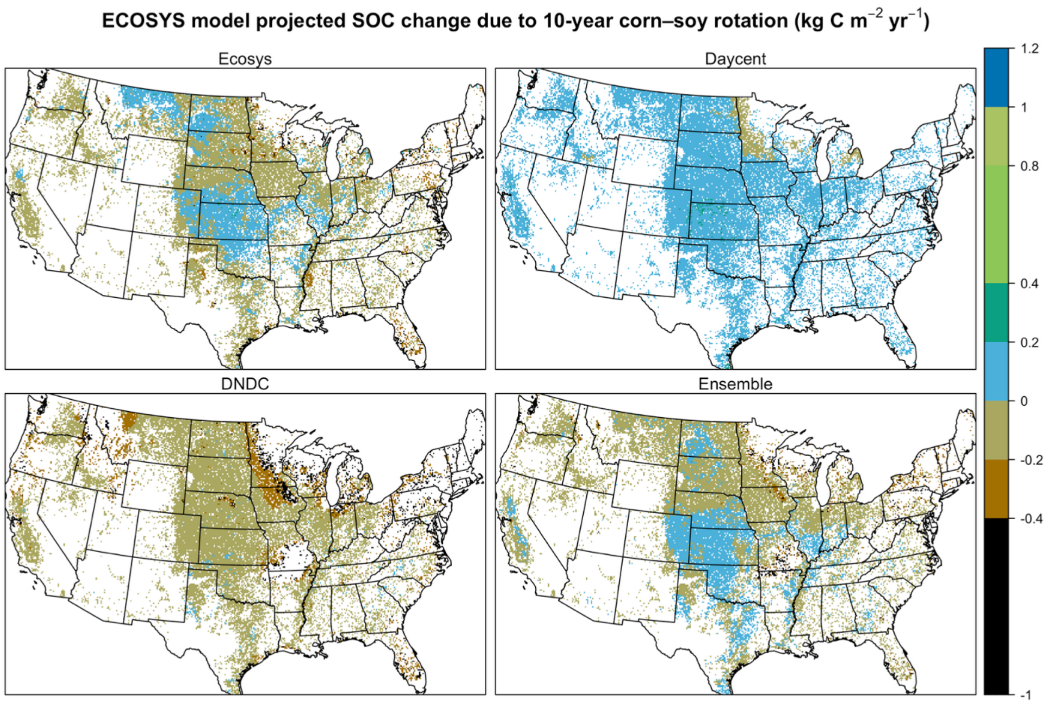


**Figure 6** Spatial distribution of simulated soil organic carbon (SOC) stock changes (kg C/m$^2$. yr) over a 10-year period (2014–2023) using individual agroecosystem models and the ensemble approach.