Journal of Cleaner Production 429 (2023) 139448

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Digging deeper: Assessing the predictive power of common greenhouse gas accounting tools for soil carbon sequestration under organic amendment

K.R. Ball[a,b,*], I.C. Burke[c], D.P. Collins[a], C.E. Kruger[a], G.G. Yorgey[a]

[a] Center for Sustaining Agriculture and Natural Resources, Washington State University, Wenatchee and Puyallup, WA, United States
[b] School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom
[c] Department of Crop and Soil Sciences, Washington State University, Pullman, WA, United States

ARTICLE INFO

Handling Editor: Cecilia Maria Villas Bôas de Almeida

Keywords:
Climate smart agriculture (CSA)
COMET-Farm
Cool farm
DayCent
Emissions reduction
Compost

ABSTRACT

Effective project implementation and quantification of emissions reduction in climate-smart agriculture initiatives face challenges in measurement, monitoring, and verification. To address these challenges, predictive models are regularly used to estimate the emissions reduction potential of land management changes and to prioritize funding for projects. Despite their growing utility, few studies have evaluated the performance of publicly available model tools using site specific data. This study evaluated the performance and utility of four common model tools that represent the three Intergovernmental Panel on Climate Change model tiers to predict soil organic carbon storage and estimate greenhouse gas emissions on working lands under organic matter amendment. Field data from two long-term, compost application experiments in Washington State formed the basis for model simulations using DayCent, COMET-Farm, Cool Farm, and the Washington State Climate Smart Estimator (WaCSE). Soil carbon sequestration and emissions estimates varied among the evaluated models, which was expected given their differential data requirements and input capabilities. COMET-Farm, although easier to use, exhibited a higher level of bias compared to DayCent, which was expected as a mixed tier model. The DayCent model, the model engine for the COMET-Farm tool, demonstrated the ability to explain ~50% more of the variation in the observed values compared to COMET-Farm when initiated using the same parameters. Cool Farm was unsuitable for estimating SOC sequestration benefits from compost application primarily because it did not add carbon to the soil pool following amendment. The differences in emissions estimates derived from WaCSE compared with other tools could be attributed solely to its highly constrained input parameters and basis in tier 1 emissions factors. We conclude that online tools can provide rapid estimates of greenhouse gas emissions reduction potential over larger areas or groups of farms but should be used with caution for site-specific estimates. Hence, it is crucial to clarify the intended purpose of an assessment and the designed function of model tools when evaluating their suitability for prioritizing funding for climate-smart agriculture initiatives at the individual farm level.

1. Introduction

In recent years, countries worldwide have reaffirmed commitment to the Paris Climate Agreement: a critical global effort to combat climate change that sets rigorous targets for reducing greenhouse gas (GHG) emissions to limit global warming below 1.5 °C. A key aspect of this commitment is the adoption of 'climate-smart agriculture' (CSA); programs focused on GHG mitigation and enhancing soil organic carbon (SOC) storage on working lands (Legislature, 2020). Globally, governments and institutions recognize the significance of CSA and have made substantial investments to encourage their inception. For example, the United States Department of Agriculture has allocated over $3.2 billion to support sustainable commodity production at the national level. Additionally, various U.S. states, including Washington and California, have introduced initiatives like the Sustainable Farms and Fields and Healthy Soils programs, partially funded through emissions cap-and-trade auctions. These initiatives illustrate the momentum behind CSA, with varying levels of commitment observed across the global policy landscape (Australian Government Department of Agriculture and Forestry, 2022; Innovation, 2022; Johansson et al., 2022; McDonald et al., 2021).

Climate-smart agriculture continues to evolve globally, amid

* Corresponding author. Center for Sustaining Agriculture and Natural Resources, Washington State University, Wenatchee and Puyallup, WA, United States.
E-mail address: kirsten.ball1@abdn.ac.uk (K.R. Ball).

https://doi.org/10.1016/j.jclepro.2023.139448
Received 3 August 2023; Received in revised form 5 October 2023; Accepted 20 October 2023
Available online 31 October 2023

persistent scientific debate surrounding the extent to which techniques like reduced tillage, cover cropping, and organic amendments deliver significant and permanent SOC sequestration benefits. Despite scientific uncertainty, the development of CSA programs and initiatives remains robust, indicating a collective recognition of the need to explore and implement innovative agricultural practices to improve soil health and mitigate climate change. These discussions are crucial for achieving substantial reductions in GHG emissions, considering the prevailing intensive cropping practices employed in commercial agriculture (Anderson et al., 2020; Olen et al., 2021). In addition to public programs, there has been a notable increase in the number of private sector initiatives that offer financial incentives for CSA. These initiatives involve companies compensating farmers based on a predetermined price per potential ton of CO_2 equivalent (CO_2e) sequestered, with the companies usually retaining ownership of the resulting 'carbon credit.' While numerous carbon crediting programs have emerged over the last decade (Oldfield et al., 2022), several contentious aspects of carbon crediting CSA require resolution before these markets can be publicly legitimized (Chandra et al., 2017; de Freitas Netto et al., 2020; Paul et al., 2023).

Measurement, monitoring, and verification of practice implementation and quantification of emissions reduction pose significant challenges in the development of successful CSA initiatives. Direct measurement of SOC and GHGs on working lands is often time-consuming, operationally difficult, and subject to high variability (Stanley et al., 2023). There remains a lack of industry consensus regarding measurement requirements and protocols (Baumber et al., 2019; Paustian et al., 2016; van der Voort et al., 2023). Furthermore, changes to bulk SOC under differential management occur gradually and may not be detectable within short-term funding timeframes. Consequently, many programs, both public and private, rely on predictive models to estimate the potential outcomes of management changes and to establish funding priorities for CSA projects.

The Intergovernmental Panel on Climate Change (IPCC) developed a system of three methodological tiers that serve as a model framework for estimating emissions and removals. Each tier represents a different level of methodological complexity, with higher tiers requiring more comprehensive data. Tier 1 models incorporate basic data, commonly rely on IPCC-recommended default values at the country level and are not specific to individual sites. These models are typically employed at a national or regional scale, providing a broad overview of the potential climate impacts. Tier 2 models operate at an intermediate level of complexity and incorporate some site-level data. Tier 2 models are often used at the national or sub-national level, offering a more detailed assessment of the potential climate impacts within specific sectors, such as agriculture or energy. Tier 3 models, the most complex, demand extensive data and are best suited for providing site-specific estimates. Tier 3 models are typically applied at a local level and offer a highly detailed assessment of management change outcomes within specific ecosystems, such as a farm or river basin. Examples of tier 3 models include DayCent and DNDC (Del Grosso et al., 2012; Li et al., 1997).

In practice, models employ a combination of tiers to estimate emissions, as this approach considers both data availability and the need to simplify use. Particularly in the calculation of emissions from agricultural operations, a mixture of site-specific management data and Tier 1 & 2 emissions factors may be used to estimate emissions from diverse sources on-farm; an example of a widely used, mixed tier model is COMET Farm (Paustian et al., 2017; USDA Natural Resources Conservation Service, 2022). Initially, the tiers of model complexity were developed to help governments estimate their greenhouse gas emissions to satisfy reporting requirements under programs like the Kyoto Protocol and Paris Agreement (Anderson et al., 2020). More recently however, the structure is being used to establish guidelines and methodologies for CSA project development, and to assess the quality and credibility of emissions reductions projects - arguably, a purpose for which they were not designed. Specifically, the tiering structure for CSA schemes has been criticized for oversimplifying emissions estimates, having limited consideration of uncertainties, lack of consistency, inadequate accounting for indirect emissions, and potential for manipulation; all of which collectively undermine the accuracy, reliability, and integrity of carbon accounting (Oldfield et al., 2022).

Despite heavy reliance on model tools during both initialization and implementation phases of CSA projects, while there are many studies comparing different online model tools (Alex Thumba et al., 2022; Hillier et al., 2011; Moreno-García et al., 2022; Whittaker et al., 2013), few studies have evaluated the performance of publicly available greenhouse gas accounting tools using site specific data. Evaluating online GHG accounting tools with site-specific data is essential to ensure their accuracy in addressing a wide range of agricultural management decisions. This is especially crucial for data-deficient ecotypes like drylands, and for non-traditional practices such as broad-scale application of organic amendments, as it enables tailored solutions that address the specific challenges and opportunities in these environments and under these practices (Leger et al., 2022; Luján Soto et al., 2021; Parr et al., 1989). Only by accounting for the unique variables of each location can these tools effectively support sustainable practices and emissions reduction strategies in agriculture. Given this data gap, and acknowledging the continued growth of carbon farming programs, we asked the question: "to what extent can common model tools be relied upon to make estimates of soil carbon sequestration?" Further, considering ongoing debates and persistent questions surrounding carbon accounting and emissions reporting in United States CSA programs, we recognized the need to clarify model options. Concurrently, the Washington State Legislature sought to quantify the SOC sequestration capacity of organic amendment additions to croplands given the intention to incorporate this practice into CSA funding schemes. Therefore, our paper evaluates the effectiveness of four models representing the three tier types currently employed in the United States for predicting SOC storage and GHG emissions on working lands. For this analysis, field data was acquired from two long-term compost application studies carried out in Washington State, forming the foundation for model simulations. The models assessed in this study were DayCent, COMET-Farm, Cool Farm, and the Washington State Climate Smart Estimator (WaCSE); the latter being a Washington County-specific adaptation of the COMET-Planner tool. The models assessed represent all levels of complexity from those that can be applied with accuracy at the farm-scale using detailed site-specific data (DayCent), to those like the WaCSE tool which incorporates county-level climate, soil, and land-use data to estimate emissions following a land-use practice change and are likely more suitable for application at the regional level. While the chosen models are not an exhaustive list of available options, they represent the range of complexity exhibited in available model tools.

The focus of our work was to compare model estimates of the on-farm benefits of organic amendment applications. Therefore, the emissions predictions reported were calculated from the 'farm gate' perspective and do not encompass emissions associated with amendment production, transport, or application. Consequently, our analysis does not constitute a comprehensive life cycle assessment. Instead, it is an exploration and evaluation of the representative emissions assessment tools currently employed to estimate SOC sequestration and emissions reduction in croplands, in the specific context of organic amendments.

2. Methods

2.1. Study description

Using field data obtained from two organic amendment studies on agricultural lands in Washington, U.S.A — one in dryland wheat systems and one in a vegetable system with supplemental irrigation - predictions of SOC storage and GHG emissions (CO_2, CH_4, N_2O) were made using DayCent, COMET-Farm, Cool Farm, and a newly developed Washington

State Department of Agriculture tool, WaCSE. These models use Tiers 1, 2 and 3 emissions factors and management data (and a mixture of tiers). **1)** DayCent (Tier 3): Is an extensively published process-based, earth system model originally developed in 1998 and primarily in the field of academia, **2)** COMET-Farm (Tiers 1,2,3): is an online GHG estimator developed by the USDA and Colorado State University and released in 2005, that uses some parameters from the DayCent model, and a mixture of tier 1 and 2 emissions factors, **3)** Cool Farm (Tiers 1,2): An online greenhouse gas, water, and biodiversity calculator for farmers developed in the United Kingdom in 2010, and **4)** WaCSE (Tier 1): the Washington Climate Smart Estimator, an application-based online tool adapted in 2022 from the COMET-Planner tool (http://comet-planner.com/) to establish WA-county-specific estimates of changes in GHG emissions resulting from changes to land management practices in line with USDA-NRCS conservation practices. Model-specific features and input potentials are detailed in Table 1, and more detailed model descriptions and initialization parameters are provided below. For DayCent and COMET-Farm, model simulations of the experimental periods were carried out for 10 years following baseline and equilibrium scenarios owing to the constraints of the COMET-Farm model (section 3.2.2). Model performance was only statistically compared between the DayCent, and COMET-Farm simulations because, of the models employed, these were the only two for which sufficient model data is provided to validate model predictions against field (observed) data. To ensure accurate validation against real-world field data, it's imperative to furnish comprehensive backend information to online modeling tools. Fig. 1 outlines the project workflow.

2.2. Model descriptions (*Table 1*)

2.2.1. DayCent

DayCent is the daily time step version of the CENTURY model used to simulate C, N, and P dynamics in forests, grasslands, and croplands. The model has been widely applied to simulate agricultural management practices including the application of organic amendments to croplands in the USA. Its key sub-models include non-dynamic plant productivity and decomposition, soil water and temperature dynamics, soil organic C and N dynamics, and trace gas fluxes. Plant growth is primarily controlled by nutrient availability, water, and temperature. Soil organic carbon (SOC) is represented in three soil pools (active, slow, and passive/inert), and two surface organic matter pools (active and slow) each with a unique decomposition rate. The active SOC pool has a short turnover time of 1–5 years and consists primarily of microbial biomass and microbial products. The slow SOC pool (turnover time 10–50 years) constitutes up to 45–60% of total SOC and is made of resistant plant material and physically (mineral) protected SOC. Therefore, DayCent allocates faster turnover rate for coarse-textured (sandier) soils whereas a high silt and clay content will slow SOC turnover enhancing the stabilization of SOC. The passive (inert) pool is considered physically and chemically stabilized SOC, highly resistant to decomposition with a turnover time of hundreds to thousands of years, constituting 45–50% of total SOC. The SOC level in the model is a function of crop C input and organic matter additions, minus losses of C from turnover. The flow of C and nutrients between pools is controlled by the amount of each in the various pools, and the rate of decomposition of each pool varies based on the soil texture, water content, temperature, and crop residue N and residue lignin content. The decomposition rate constants used in this model for the active pool were DEC3 (2): 6, passive pool DEC4: 0.001 and the slow pool DEC5 (2): 0.20. Daily maximum/minimum temperature and precipitation, and the timing and description of management events (crop planting, fertilization/OA additions, tillage, and harvest), and soil texture data are required to initialize the model. DayCent is considered a higher tier (3) model and may be considered more accurate than lower tier models at the site level on the condition that adequate data is available to initialize and calibrate the model. An infinite number of years of management can be modelled, and there is no limit to the number of crops that can be entered into rotation. Only one crop can be grown at any given time.

Table 1
Input requirements and useability compared between the DayCent, COMET-Farm, Cool Farm and WaCSE models/tools..

	Site & Climate		Soil		Land Management			Cropping			Inputs				Other GHG Sources		Useability		
	Climate & Weather	Geographical Location	Texture / Water Holding Capacity / Bulk Density	Initial SOC / SOM Content	Historical Management	Conservation Practice Status	Tillage / Ground Operations	Crop Type	Crop Rotation	Planting & Harvest Dates	Fertilizer	Organic Amendments	Irrigation	Grazing	Fuel & Energy	Transport	Predictive Scope (years)	Data Requirement	Required Operational Skill Level
DAY-CENT	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Defined from external data	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Fully user-defined	Not defined	Not defined	100+ years	High	High
COMET Farm	Defined from external data	Defined from external data	Defined from external data	Defined from external data	Drop down	Drop down	Drop down	Drop down	Drop down	Fully user-defined	Fully user-defined	Drop down	Fully user-defined	Drop down	Defined from external data	Defined from external data	10 years	Low	Med.
COOL Farm	Not defined	Not defined	Drop down	Fully user-defined	Not defined	Drop down	Drop down	Drop down	Not defined	Not defined	Fully user-defined	Drop down	Fully user-defined	Defined from external data	Defined from external data	Defined from external data	1 year	Low	Low
WaCSE	Defined from external data	Drop down	Defined from external data	Defined from external data	Not defined	Drop down	Drop down	Not defined	Not defined	Not defined	Drop down	Drop down	Drop down	Drop down	Drop down	Not defined	1 year	Low	Low

Fully user-defined via manual entry or input flexibility
Defined from external data/other related tools
Defined through drop down (restricted choices)
Not defined

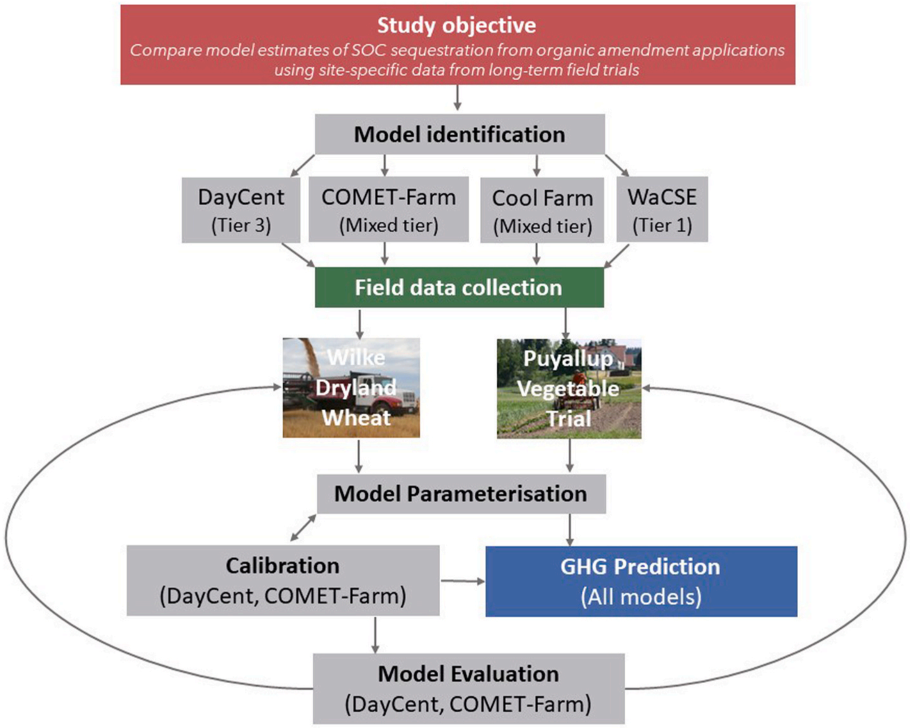

Fig. 1. Schema outlining the project workflow.

2.2.2. COMET-farm

The COMET-Farm (Paustian et al., 2017; USDA Natural Resources Conservation Service, 2022) system is a publicly available GHG accounting tool that incorporates state-of-the-art greenhouse gas quantification methods in an online accessible format. The tool was developed by United State Department of Agriculture (USDA) and Colorado State University for the purpose of providing a ranch and farm greenhouse gas accounting system for use by land managers. It can perform a full greenhouse gas assessment for CO_2, CH_4, and N_2O from all major on-farm emission sources (and CO_2 removal into biomass and soil sinks), and includes land management of annual and perennial crops, pasture, range and agroforestry systems, and emissions from livestock. It uses spatially explicit data on climate obtained from the PRISM climate database, and soil conditions from the Soil Survey Geographic Database (SSURGO). It allows the user to enter detailed information for field, crop, and livestock management, and can incorporate National Resource Conservation Service (NRCS) conservation practice standards (CPS) via a fully spatial mapping and menu-driven graphical user interface https://comet-farm.com/. The COMET-Farm model employs all three tiers of methodological complexity, requiring some user-defined, site-specific data, some data that is obtained from county-level measurements, and some IPCC-recommended country-level default values that are not site/county specific. Data can be entered manually and using drop-down menus and by specifying/manipulating some pre-defined values. The user specifies a period of historical management (pre-1980) followed by a baseline period of management. From the end of this baseline period, 10 years of management can be modelled, and users can specify up to 3 different crops in rotation (growing only one at a time) in any given year.

2.2.3. Cool Farm

Like COMET-Farm, Cool Farm is an online, menu-driven graphical interface tool that facilitates the calculation of GHG emissions and carbon sequestration under differential farm management and is intended for use by farmers and land managers https://coolfarmtool.org/. It can perform an assessment of CO_2e emissions from CO_2, N_2O and CH_4. Unlike some other tools, Cool Farm also allows the user to specify on-farm energy use for production and processing, and fuel use from transport. The Cool Farm tool incorporates tiers 1 and 2 of methodological complexity, requiring only a small amount of user-defined, site-specific data, but where most model inputs rely on IPCC-recommended country-level default values that are not site/county specific. Data can be entered manually and using drop-down menus and by specifying/manipulating some pre-defined values. Only one year of cropping/land management can be specified and modelled at any given time. If the user wishes to make multi-year predictions or specify crop rotations, the data must be modelled separately and summed.

2.2.4. The Washington State Climate Smart Estimator (WaCSE)

An adaptation of the online, user-friendly COMET-Planner tool (http://comet-planner.com/) developed by Colorado State University, WaCSE was advanced by the WSDA to support quantification of WA-county-specific estimates of changes in GHG emissions resulting from changes to land management practices in line with USDA-NRCS conservation practice benefits. It can provide county-rectified estimates of GHG emissions (CO_2e) savings from CO_2, N_2O and CH_4 sources following the implementation of NRCS conservation practices. WaCSE would be classified as a tier 1 model, because it requires a very low level of data input/complexity and uses non-site-specific data. It produces GHG estimates from IPCC-recommended country-level default values using WA county-rectified major land use area spatial units.

2.3. Field data

2.3.1. Site and study description (***Case study 1: Wilke Farm dryland wheat***)

A long-term compost study was established at the Washington State University Wilke Research Farm in Davenport, Washington in 2015 (47.6562, −118.09, Elevation = 2375 m). The no-tillage dryland wheat-based field site is situated on silt loam soil with an average of 337 mm of precipitation annually. The initial SOC concentration of the soil in 2015 before the initiation of the experiment was ~1.55%. Throughout the field sampling period (2016–2021), the region experienced several periods of drought (Fig S1 supplementary materials). In September 2016, municipal compost, obtained from Barr-Tech (Sprague, WA) was surface applied at one-time rates of 10,000, 25,000, and 50,000 kg ha-1 (8,921, 22,304, 44,608 lb. A-1) dry weight basis, compared against a conventionally fertilized control, fertilized annually (11.2 g N per m2 as 46-0-0-0, 3.36 g P per m2 as 16-20-0-13), and a control that received neither compost nor fertilizer. The compost, a mixture of yard and lawn trimmings, recyclable food materials and municipal biosolids, had a carbon to nitrogen ratio of 15, and added 270, 675 and 1350 g m2 of organic C to the system respectively. The compost treatments did not receive inorganic fertilizer. Two crop rotations were managed within each treatment: winter wheat-fallow (WF), and winter wheat – fallow – winter pea – fallow (WFPF) for a total of 4 replicates of each rotation and compost treatment. Crops were planted in September of each year and harvested in July of the following year. No irrigation water was applied to the experimental plots.

2.3.2. Sampling and analyses (**Case study 1: Wilke Farm dryland wheat**)

Soil samples were collected annually and analyzed for soil organic carbon, beginning in 2018 to a depth of 90 cm in four depth increments: 0–15, 15–30, 30–60, and 60–90 cm. Five soil cores (5 cm diam.) from each treatment plot were removed, composited in the field and subsampled. Soil was transported in a cooler to the lab, a subsample air-dried, and sieved to 2 mm and the remainder refrigerated. All visible roots and other particulate organic matter were removed during sieving. Air-dried soil was then stored for subsequent analysis. Soil organic matter content was determined by the Walkley-Black titration method using a spectrophotometer (Orion Aquamate 8000 UV–Vis Spectrophotometer, Thermo Scientific, Waltham, MA) (Gavlak et al., 2003). Soil organic carbon values were derived from soil organic matter by multiplying the SOM value by 0.58. Only the 0–15 cm and 0–30 cm depth values for SOC were used to calibrate and validate the models. Full experimental results from this study are in the process of being prepared for peer review publication.

2.3.3. Site and study description (**Case study 2: Puyallup vegetable trial**)

A long-term organic vegetable crop experiment was established in 2003 at the Washington State University (WSU) Puyallup Research and Extension Center in Puyallup, WA, USA (47° 11′24″ N, 122° 19′48″ W; elevation 13 m). The soil is classified as a Puyallup fine sandy loam (coarse-loamy over sandy, isotic over on-farm, mesic Fluventic Haploxerolls) with an average of 1040 mm of precipitation annually, and a mean annual temperature of 10.4 °C. Conventional row crops (predominantly silage corn (*Zea mays* L.)) were grown before the start of organic transition in 2001. The original experiment compared 12 combinations of organic management systems, consisting of three cover crop systems, two types of tillage, and two soil amendments arranged in a split-split plot design with four replicate blocks under drip irrigation. Cover crops were the main plots, tillage the first split, and soil amendment the second split. Main plots within each block measured 24.4 × 15.2 m. Details of the full experiment setup and full results are available (Cogger et al., 2016; Pritchett et al., 2011). For the current modeling study, one cover crop type, one tillage treatment, and two organic amendments were modelled. The primary (cash) crops included winter squash and broccoli (transplanted), and winter wheat and spinach (directly seeded). These crops were grown in rotation with a fall planted cover crop – a 50:50 mixture of cereal rye (*Secale cereale* L.) and hairy vetch (*Vicia villosa*) seeded at 134 kg ha-1 (119.5 lb. ac.). Vegetable crops were planted in April–May of each year and harvested in August–September and fall cover crops were seeded in September of each year, terminated, and residue incorporated before re-planting the vegetables. Between September 2003–2014, broiler litter and on-farm compost, both produced on-farm at WSU were surface applied annually to provide similar amounts of available nitrogen between each treatment (6000 kg ha-1 of broiler litter and 38,000 kg ha-1 of on-farm compost). The on-farm compost was made from locally available materials, and feedstocks included separated dairy solids, animal bedding from the Washington State Fair, yard debris, and small amounts of broiler litter and fish waste. The two applications resulted in average carbon inputs of 1760 kg ha-1 yr-1 (1570.2 lb. ac) of C from the broiler litter (low C input, C:N ~12, 3.8% N), and 6250 kg ha-1 yr-1 (5576 lb. ac) of C from the on-farm compost (high C input, C:N ~14, 1.7% N; **Table S1**). Organic amendments were applied with a manure spreader in the spring after mowing the cover crop and incorporated either on the day of, or the day after application. Spader-tillage (rotary spader, 1–2 passes at 1.3 km h-1 to 25 cm depth) was used prior to fall cover crop seeding in the fall-planted treatment and to incorporate cover crop residue in the spring.

2.3.4. Sampling and analyses (**Case study 2: Puyallup vegetable trial**)

Baseline samples (prior to experimental initiation) for soil carbon were obtained in 2002. Experimental soil samples were collected and analyzed annually from 2003 to 2012, and again in 2022 to a depth of 30 cm. 10 cores were collected in each plot with a hand probe (2-cm diameter). The cores were on-farm, sieved (<2 mm), air-dried at 30 °C, and ground prior to analysis. All cores were collected from the center two beds of each plot. Soil organic carbon was determined using standard combustion methods (Gavlak et al., 2003). Soil bulk density was determined using a hammer driven core sampler that collected a 6 cm deep by 5.4 cm diameter core (Grossman and Reinsch, 2002).

2.4. Model parameterization and calibration

Tables 2 and 3 detail simulation parameters for the DayCent, COMET-Farm, CoolFarm and WaCSE models for the Wilke Dryland Wheat and Puyallup Vegetable trials respectively. For the DayCent model, where model parameters like soil water holding capacity and crop inputs could be manually adjusted, this was performed iteratively until expected crop yields and measured SOC content were achieved in the simulation. These adjustments are detailed in Tables 2 and 3 (‘Soil parameters’ and ‘Crop parameters’). For an expanded parameterization method, please refer to ‘Model parameterization and calibration information’ in Supplementary Materials.

2.5. Model evaluation: case studies 1 & 2

Validation of model simulated values using observed SOC was only able to be performed for the DayCent model and COMET-Farm tool as they both provide detailed data outputs in. csv and. xml formats. For COMET Farm, the output. xml files were accessed from “file results” section of the report tab. From the “management information” files it was possible to access the background SSURGO data containing the soil parameters used to run the COMET-Farm model. Using these files, it was possible to determine the SSURGO imputed initial SOC content of the soil, soil texture information and bulk density, and to then quantify the predicted SOC content and GHG emissions changes for each modelled year. Information pertaining to the interpretation of DayCent and COMET output files were obtained from the individual model developers (Del Grosso et al., 2006; Easter, 2018).

2.5.1. Statistical analysis

To understand which experimental variables influenced the expression of modelled and measured values between models, linear on-farm effects models were fitted using the “ASReml-R” (Butler et al., 2009) and “asremlPlus” (Brien, 2019) packages in R (R Core Team, 2022). The response variables (either modelled or observed C) were square root transformed to improve the pattern of residuals. The models included fixed parameters for organic amendment treatment and rotation, and their interactions along with terms that allowed these trends to differ between treatments. A random term was included for year. Model

Table 2
DayCent, COMET-FARM, CoolFarm and WaCSE model parameters for **Case Study 1: Wilke Dryland Wheat**.

Model	DayCent	COMET-Farm	Cool Farm	WaCSE
Model Type	Earth System Model	Online whole farm and ranch carbon and greenhouse gas accounting system	Online carbon and environmental calculator for agriculture	Online calculator for generalized estimates of the greenhouse gas impacts of conservation practices for conservation planning purposes
Initialization	**Equilibrium scenario (4000 years):** Low productivity, cool season grassland. **Baseline scenario 1 (1881–1998):** winter wheat - fallow with "B" cultivation event (minimal tillage). **Baseline scenario 2 (1998–2015):** winter wheat - fallow, no tillage.	**Historic management scenario (Pre-1980):** non-irrigated grain-fallow under intensive tillage. **Baseline scenario (2000–2015):** non-irrigated winter wheat no-till.	No historical/baseline land-use history specified.	No historical/baseline land-use history specified.
Climate	Daily min-max temperature and precipitation from local weather station (1981–2021)	Defined by PRISM[d]	Not specified	Imputed using county-rectified **M**ajor **L**and **R**esource **A**rea[b] spatial units
Soil organic carbon	**Post equilibrium:** ~6170 g C m[b]. **Post baseline 1:** 6162.21 g C $m^{2.}$ **Post baseline 2:** 5519.14 g C m[b].	**SSURGO[a] designated:** SOC content (2015) 1.16%/4647.13 g C m[b]	**User-defined:** 1.15%	Imputed using county-rectified **M**ajor **L**and **R**esource **A**rea[b] spatial units
Soil Parameters	**User defined:** Sand = 43%, Clay = 18%, Bulk density = 1.31 g cm^3. Field capacity (column 4) = 0.27, wilting point (column 5) = 0.09, volumetric soil water content wilting point (wilting point-deltamin, column 11) = 0.08–0.00.	**SSURGO designated:** Clay = 11.5%, Bulk density = 1.33 g cm[c]	**User designated (constrained)[c]:** sandy (coarse) texture, average soil moisture dry, good soil drainage, pH 5.5–7.3	Imputed using county-rectified **M**ajor **L**and **R**esource **A**rea[b] spatial units
Management/ cropping scenarios	**2015–2025 (Wheat-fallow rotation):** Winter wheat (Planting: Yr. 1, Day 244. Harvest: Yr. 2, Day 210. Fallow: Yr. 2, Day 211 - Yr. 3 Da y 244). Organic matter additions occurred on Day 268. **2015–2025 (Wheat-pea-fallow rotation):** Winter wheat (Planting: Yr. 1, Day 244. Harvest: Yr. 2, Day 210. Fallow: Yr. 2, Day 211 - Yr. 3 Da y 244). Winter Pea (Planting: Yr. 3, Day 244, Harvest: Yr. 4, Day 210).	**2015–2025 (Wheat-fallow rotation):** Winter wheat (Planting: Yr. 1, Day 244. Harvest: Yr. 2, Day 210. Fallow: Yr. 2, Day 211 - Yr. 3 Da y 244). Organic matter additions occurred on Day 268. **2015–2025 (Wheat-pea-fallow rotation):** Winter wheat (Planting: Yr. 1, Day 244. Harvest: Yr. 2, Day 210. Fallow: Yr. 2, Day 211 - Yr. 3 Da y 244). Winter Pea (Planting: Yr. 3, Day 244, Harvest: Yr. 4, Day 210).	**2015–2016:** (Winter wheat rotation): Only one year of cropping was able to be simulated.	No crop rotation designated
Crop Parameters	**User defined:** 'PRDX (1)' value adjusted for winter wheat (W1) to 1.75. Crop.100 parameters for grass = GI3, winter wheat = W1, peas = AWP.	**User defined (constrained):** Predicted yield values for winter wheat (95 Bu/ac, 70% residue removal) and dry field pea (20 Bu/ac, 70% residue removal).	**User defined (constrained)**: Winter wheat (5700 lb./ac, 70% residue removal)	Not specified
Organic Matter/ Fertilizer Amendments	**User defined:** On-farm compost. C:N ratio 15. One-time rates of 10,000, 25,000, and 50,000 kg $ha^{-1.}$ Conventionally fertilized control applied annually (11.2 g N per m^2 as 46-0-0-0, 3.36 g P per m^2 as 16-20-0-13).	**User defined (constrained):** Compost or composted manure. One-time rates of 10,000, 25,000, and 50,000 kg ha^{-1}. Moisture content = 50%, C:N 15, 1.77% N. Conventionally fertilized control, applied annually (11.2 g N per m[b] as 46-0-0-0, 3.36 g P per m[b] as 16-20-0-13).	**User defined (constrained)**: Compost (zero emissions) 1% N. Applied at experimentally defined rates.	**User defined (constrained):** CPS 590[e], Replacing synthetic nitrogen fertilizer with compost (C:N ratio 15)

[a] The Soil Survey Geographic Database (SSURGO) is a comprehensive and detailed soil information system developed and maintained by the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS).
[b] A system used in the United States to classify and group areas with similar land use, physiography, climate, soils, and biological resources.
[c] In some cases, the online tools allow the user to select from pre-defined parameters using drop-down menus.
[d] PRISM (Parameter-elevation Regressions on Independent Slopes Model) weather data is a dataset that provides high-quality, fine-scale climate information for a specific geographic area.
[e] Specific conservation practices used in agriculture and natural resource management.

residuals were confirmed to be homogenous and normally distributed with different variances by treatment, using scatter and qqplots. Wald F-tests at $\alpha = 0.05$ were conducted for the fixed covariate terms, and non-significant terms were removed and are noted as "NA" in the pseudo-ANOVA tables.

Then, to validate the DayCent and COMET-Farm models, a series of quantitative metrics were produced (Smith et al., 1997). The primary metrics calculated were root mean squared error (RMSE; Eq. (1)) which details the mean error between modelled and observed values, and the normalized root mean square error (NRMSE; Eq. (2)): a standardized RMSE that allows the comparison of RMSE values of differing scales between models.

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} (M_i - O_i)^2} \Big/ n \tag{1}$$

Where M_i and O_i denote the modelled and observed values respectively and n is the number of observations.

$$\text{NRMSE} = RMSE * \frac{100}{\hat{\text{O}}} \tag{2}$$

Where $\hat{O}$ is the mean of the observed data.

Model efficiency (EF; Eq. (3)) is a metric with a value between 0 and 1 that shows how well the modelled values approximate the mean of the observed data (Karhu et al., 2012). Values of EF below 0 indicate that the modelled values less closely approximate the mean of the observed data, and positive values indicate that the modelled values describe the data better than the mean of the observed values, and values closer to 1 indicate a near perfect fit between modelled and observed values. Negative EF values can indicate lack of model fit (Nash and Sutcliffe, 1970). Relative error (RE) determines the bias in the total difference between modelled and observed values, with values closer to

Table 3
DayCent, COMET-FARM, CoolFarm and WaCSE model parameters for **Case Study 2: Puyallup vegetable trial**.

Model	DayCent	COMET-Farm	Cool Farm	WaCSE
Model Type	Earth System Model	Online whole farm and ranch carbon and greenhouse gas accounting system	Online carbon and environmental calculator for agriculture	Online calculator for generalized estimates of the greenhouse gas impacts of conservation practices for conservation planning purposes
Initialization	**Equilibrium scenario (4000 years):** Cool season grassland. **Baseline scenario (1980–2002):** Silage corn - fallow with "K" cultivation event (moldboard plough).	**Historic management scenario (Pre-1980):** non-irrigated grain-fallow under intensive tillage. **Baseline scenario 1 (2000–2002):** Irrigated silage corn with "K" cultivation. **Baseline scenario 2 (2002–2005):** Same as experimental scenario (2005–2015)	No historical/baseline land-use history specified.	No historical/baseline land-use history specified.
Climate	Daily min-max temperature and precipitation from local weather station (1981–2015)	Defined by PRISM[4]	Not specified	Imputed using county-rectified **M**ajor **L**and **R**esource **A**rea[2] spatial units
Soil organic carbon	**Post equilibrium:** ~6948 g C m^2. **Post baseline:** 6447.21 g C m^2.	**SSURGO[1] designated:** SOC content (2005) 5.22%/9755.20 g C m^2	**User-defined:** 3.11%	Imputed using county-rectified **M**ajor **L**and **R**esource **A**rea[2] spatial units
Soil Parameters	**User defined:** Sand = 45%, Clay = 8%, Bulk density = 1.17 g cm^3. No other soil parameters were altered.	**SSURGO designated:** Clay = 8%, Bulk density = 1.46 g cm^3	**User designated (constrained)[3]:** Silt (medium) texture, good soil drainage, pH 5.5–7.3	Imputed using county-rectified **M**ajor **L**and **R**esource **A**rea[2] spatial units
Management/ cropping scenarios	**2005–2015 (Vegetable – cover crop rotation):** Tomato (Planting: Yr. 1, Day 135. Harvest: Yr. 1, Day 227. Fall cover crop: Yr. 1, Day 278 - Yr. 2 Da y 100). Organic matter additions occurred on Day 110 of each year.	**2005–2015 (Vegetable – cover crop rotation):** Tomato (Planting: Yr. 1, Day 135. Harvest: Yr. 1, Day 227. Cover crop: Yr. 1, Day 278 - Yr. 2 Da y 100). Organic matter additions occurred on Day 110 of each year.	**2005–2006 (Tomato):** Only one year of cropping was able to be simulated. **2006–2007 (Grass-clover cover crop):** Only one year of cropping was able to be simulated. Organic matter additions occurred on Day 110 of each year. *Estimates for the two separate simulations were summed to achieve estimates for the entire rotation.*	No crop rotation designated
Crop Parameters	**User defined:** 'PRDX (1)' value adjusted for tomato (JTOM) = 2.0. Crop.100 parameters for grass = GI3, fall cover crop = G1CPD.	**User defined (constrained):** Predicted yield values for tomato (26,000 lb./ac, 70% residue removal).	**User defined (constrained):** Tomato (26,000 lb./ac, 70% residue removal). Grass-clover cover crop.	Not specified
Organic Matter/ Fertilizer Amendments	**User defined:** Broiler litter: C:N ratio 12, total N 3.8%. Annual rates of 6000 kg ha^{-1}. On-farm-compost: C:N ratio 15, total N 1.7%. Annual rates of 38,000 kg ha^{-1}	**User defined (constrained):** Chicken-broiler litter: C:N ratio 12, total N 3.8%. Annual rates of 6000 kg ha^{-1}. Compost/composted manure: C:N ratio 15, total N 1.7%. Annual rates of 38,000 kg ha^{-1}	**User defined (constrained):** Compost (zero emissions) 1% N (OFC), Poultry layer manure 1.9 % N (Broiler) Applied at experimentally defined rates.	**User defined (constrained):** CPS 590[5], Replacing synthetic nitrogen fertilizer with compost (C:N ratio 12 & 15)

0 suggesting a better model fit.

$$EF = \frac{\left[\sum_{i=1}^{n}(O_i - \hat{O})^2 - \sum_{i=1}^{n}(M_i - O_i)^2\right]}{\left[\sum_{i=1}^{n}(O_i - \hat{O})^2\right]} \tag{3}$$

$$RE = (M_i - O_i) / O_i \tag{4}$$

For each of the case studies, these qualitative assessments were performed for each level of experimental complexity. For Case study 1: Wilke Farm dryland wheat, this included generating values for treatment (organic amendment amount), rotation (Wheat-Fallow or Wheat-Fallow-Pea-Fallow) and the combination of treatment and rotation type. For Case study 2: Puyallup vegetable trial, this involved generating values between organic amendment treatments (on-farm compost vs broiler litter).

3. Results

3.1. *Model performance DayCent & COMET-farm* ***(Case study 1: Wilke Farm dryland wheat)***

For the DayCent model, treatment type (**$p < 0.01$;** Table S2) and crop rotation (**$p<0.05$**) were both significant predictors of the modelled SOC values. For the observed values, both treatment type (**$p<0.0001$;** Table S2) and rotation (**$p<0.0001$**) were the most significant predictors of SOC content. The extent to which the DayCent modelled values were consistent with the observed values was highly dependent on both treatment and crop rotation. Across the whole model, 39% of the variation in observed values could be explained by the modelled values, with an average RMSE of 580 g C m^2 and an NRMSE of 11%. Total model bias (RE) was 4%. Overall, the DayCent modelled values tended to describe the data better than the mean of the observed data (EF, Table 4). At the treatment level, the model explained between 1 and 87% of the variation in the observed SOC values (Table 4). For this level RMSE values were between 400 and 743 g C m^2, and NRMSE values were <13%. At the rotation level, the modelled values were only able to explain between 36% (Wheat-Fallow) and 44% (Wheat-Fallow-Pea-Fallow) of the variation in the observed SOC values, and model bias (RE) was only 3% for the Wheat-Fallow-Pea-Fallow rotation compared with 6% for the Wheat- Fallow rotation. Rotation by treatment modelled values could explain between 1 and 99% of the variation in the observed SOC values, with the model's predictive capacity being better for the wheat-fallow rotation and for both rotations, the model best explained the variation in the observed values for the fertilized control treatment (**$r2 = 0.88$,** Table 4). Model bias (RE) was between 7 and 12% depending on the rotation-treatment type; DayCent tended to over-estimate SOC content (Fig. 2).

For the COMET-Farm model, treatment type (**$p<0.0001$;** Table S2) and crop rotation (**$p<0.001$**) were both significant predictors of the

Table 4
Case study 1: Wilke Farm dryland wheat model performance metrics for each level of experimental complexity for the DayCent and COMET-Farm models. Includes root mean square error (RMSE, g C m^2), normalized root mean square error (NRMSE, % observed mean), model efficiency (EF) and relative error (RE).

	Model, Treatment, Rotation, Trt + Rotation	n	RMSE	NRMSE (%)	R^2	EF	RE
Model	**DAYCENT**	39	580	11	0.39	0.007	0.04
Treatment	No fertilizer	8	561	11.1	0.000	−0.979	0.077
	Fertilized Control	8	630	12.9	0.871	−3.01	0.117
	10,000 kg	8	531	10.3	0.167	−0.601	0.066
	25,000 kg	8	400	7.48	0.027	−0.366	0.037
	50,000 kg	7	743	12.3	0.506	−0.555	−0.067
Rotation	Wheat-Fallow	19	628	12.1	0.36	−0.11	0.06
	Wheat-Fallow-Pea-Fallow	20	530	9.9	0.44	−0.1	0.03
Rotation by treatment	Wheat-Fallow (No fertilizer)	4	610	12.1	0.020	−0.716	0.075
	Wheat-Fallow (Fertilized control)	4	661	13.7	0.874	−3.52	0.125
	Wheat-Fallow (10,000 kg)	4	625	12.6	0.098	−2.63	0.109
	Wheat-Fallow (25,000 kg)	4	415	7.93	0.036	−1.41	0.061
	Wheat-Fallow (50,000 kg)	3	818	13.3	0.992	−0.674	−0.077
	Wheat-Fallow-Pea-Fallow (No fertilizer)	4	508	10.1	0.074	−1.55	0.080
	Wheat-Fallow-Pea-Fallow (Fertilized control)	4	599	12.2	0.892	−2.61	0.109
	Wheat-Fallow-Pea-Fallow (10,000 kg)	4	418	7.83	0.332	−0.038	0.027
	Wheat-Fallow-Pea-Fallow (25,000 kg)	4	384	7.04	0.228	−0.088	0.015
	Wheat-Fallow-Pea-Fallow (50,000 kg)	4	682	11.3	0.303	−0.478	−0.059
Model	**COMET**	39	1196	22.7	0.27	−3.23	−0.127
Treatment	No fertilizer	8	562	11.2	0.246	−0.987	0.048
	Fertilized Control	8	354	7.27	0.000	−0.265	−0.008
	10,000 kg	8	620	12.1	0.127	−1.18	−0.077
	25,000 kg	8	1091	20.4	0.057	−9.16	−0.192
	50,000 kg	7	2381	39.3	0.269	−15	−0.383
Rotation	Wheat-Fallow	19	1251	24.1	0.301	−3.44	−3.17
	Wheat-Fallow-Pea-Fallow	20	1142	21.4	0.29	−3.17	−0.129
Rotation by treatment	Wheat-Fallow (No fertilizer)	4	632	12.5	0.143	−0.845	0.057
	Wheat-Fallow (Fertilized control)	4	337	6.98	0.008	−0.179	−0.018
	Wheat-Fallow (10,000 kg)	4	483	9.77	0.503	−1.17	−0.048
	Wheat-Fallow (25,000 kg)	4	1053	20.1	0.269	−14.5	−0.195
	Wheat-Fallow (50,000 kg)	3	2727	44.4	0.32	−17.6	−0.436
	Wheat-Fallow-Pea-Fallow (No fertilizer)	4	482	9.57	0.694	−1.29	0.038
	Wheat-Fallow-Pea-Fallow (Fertilized control)	4	370	7.55	0.082	−0.379	0.001
	Wheat-Fallow-Pea-Fallow (10,000 kg)	4	732	13.7	0.303	−2.19	−0.104
	Wheat-Fallow-Pea-Fallow (25,000 kg)	4	1128	20.6	0.151	−8.36	−0.189
	Wheat-Fallow-Pea-Fallow (50,000 kg)	4	2084	34.7	0.901	−12.8	−0.344

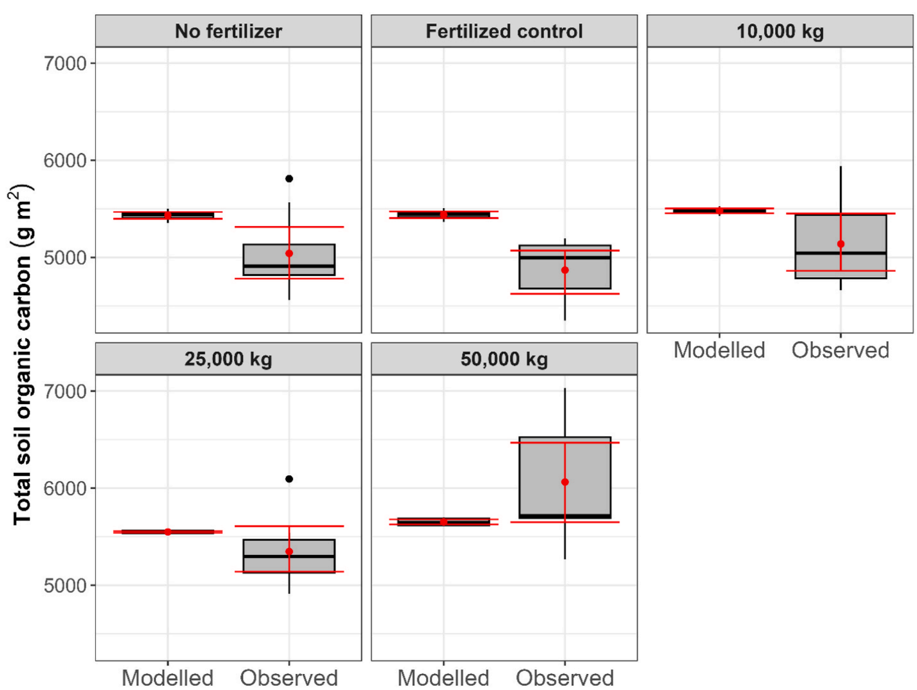

Fig. 2. *Case study 1: Wilke Farm dryland wheat* observed vs DayCent modelled SOC (g m^2) values at the treatment level. The red lines indicate the bootstrapped 95% confidence interval for the mean of each group. Black points indicate outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

modelled SOC values. Across the whole model, 27% of the variation in observed values could be explained by the modelled values, with an average RMSE of 1196 g C m^2 and an NRMSE of 22.7%. Total model bias (RE) was 12.7%. Across the whole model, the negative efficiency factor of −3.23 indicates a poor model fit (Table 4). The linear relationship between the modelled and observed values (R^2) did not appear to improve at the treatment level, and at the rotation level the amount of variation in the observed SOC values that could be explained by the modelled values was ~29%. Whether the COMET-Farm tool tended to over or underestimate the observed SOC values was highly dependent on treatment; the tendency being to underestimate SOC content for all organic matter additions (Table 4; Fig. 3).

3.2. DayCent model predictions (**Case study 1: Wilke Farm dryland wheat**)

3.2.1. SOC storage 2016–2025

In both rotations, total soil organic carbon was predicted to be higher with increasing compost amount. The WF systems were predicted to lose up to 6% and gain up to 3% SOC, depending on the amount of organic amendment added. The WFPF systems lost ~0.03% and gained up to 17% SOC between 2016 and 2025 (Fig. 4). Only the 50,000 kg treatment under WF, and the 25 and 50,000 kg treatments under WFPF were predicted to gain soil organic carbon through 2025 (Table 5).

3.2.2. Net greenhouse gas emissions

Since neither nitrous oxide nor methane emissions from soils were directly measured in the field or used to calibrate the model, net greenhouse gas emissions should be assessed with caution (Table 5). In Case study 1: Wilke Farm dryland wheat, all systems except the WFPF 50,000 kg treatment/rotation were predicted to be net-emitting between 2016 and 2025. Total emissions were lower under compost application compared with no fertilizer and fertilized controls for both rotations and were overall lower in the WFPF rotation. N_2O emissions were predicted to be lower than fertilizer application under compost in the WF system, but compost tended to increase N_2O emissions compared with fertilizer in the WFPF system. Overall, the emissions predictions were between −1.2 T ha^{-1} yr^{-1} (WFPF 50,000 kg) and +1.75 T ha^{-1} yr^{-1} (WF Unfertilized).

3.3. COMET-farm predictions (**Case study 1: Wilke Farm dryland wheat**)

3.3.1. Soil organic carbon storage 2016–2025

In both rotations, the unfertilized, fertilized control, and 10,000 kg compost treatments were predicted to gain soil organic carbon through 2025. Including the C added with the compost additions, the WF systems were predicted to increase SOC stocks by on average 9%, the greatest increase attributed to the unfertilized controls (Table 6). In the WF rotations under 25,000 kg and 50,000 kg compost, SOC loss averaged 89%; the loss being greatest under the 50,000 kg compost addition where the model predicted a loss of all added carbon, plus a further reduction of initial carbon stock. The average SOC increases in the WFPF system followed the same treatment pattern as the WF rotation but averaged 12%, while the predicted losses were smaller; with average SOC losses of ~48% between the 25,000 kg and 50,000 kg treatments. (Fig. 5). Unlike DayCent, COMET-Farm predicted higher SOC storage in unfertilized control, and less storage with increasing compost amount.

3.3.2. Net greenhouse gas emissions

For Case study 1: Wilke Farm dryland wheat, the no fertilizer, fertilized control, and 10,000 kg compost systems were predicted to be net-sequestration between 2015 and 2025 (Table 6) with overall higher net sequestration under the Wheat-Fallow-Pea-Fallow rotation. In the Wheat-Fallow rotation, total GHG emissions were generally higher under increasing compost application compared with the fertilized control except for the 10,000 kg application. In the Wheat-Fallow-Pea-Fallow rotation, total GHG emissions were generally higher under increasing compost application compared with the fertilized control. N_2O emissions were predicted to be lower than fertilizer application under 10 and 25,000 kg compost applications in both rotations, but the 50,000 kg treatment was predicted to have the highest N_2O emissions

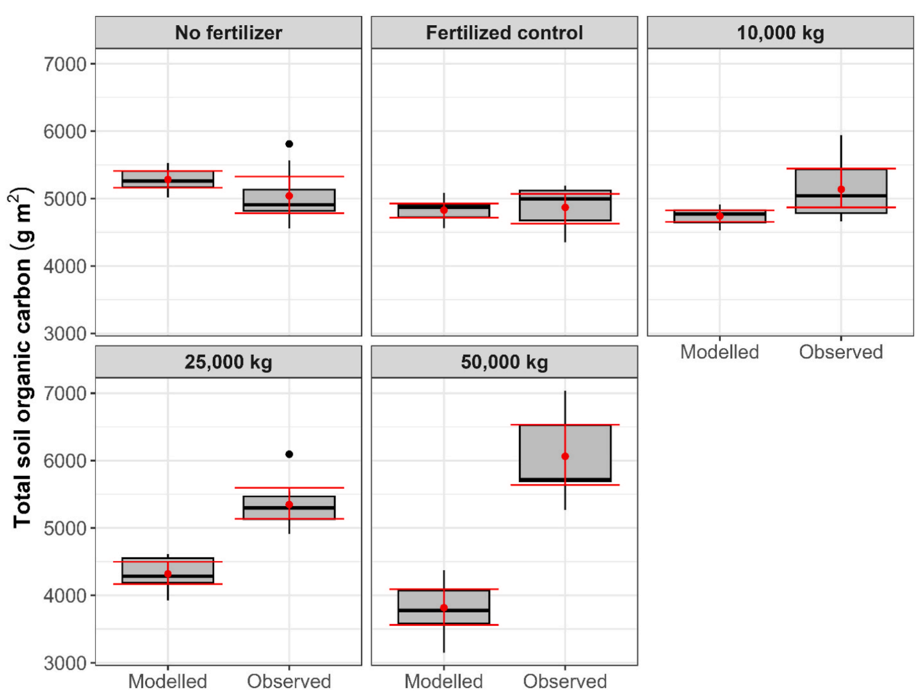

Fig. 3. *Case study 1: Wilke Farm dryland wheat* observed vs COMET-Farm modelled SOC (g m^2) values at the treatment level. The red lines indicate the bootstrapped 95% confidence interval for the mean of each group. Black points indicate outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

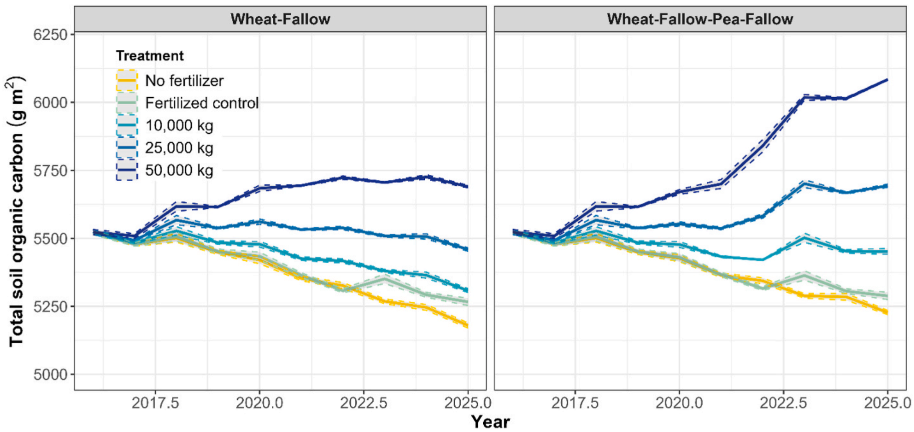

Fig. 4. DayCent simulated soil organic carbon (g m^2) for 2015–2025 between crop rotations and treatments (applied in kg ha^{-1}). WF = Wheat-fallow, WFPF = Wheat-Fallow-Pea-Fallow. Note that the initial compost amount is included in these predictions. Shaded areas represent the 95% confidence interval of the mean by treatment.

Table 5
Greenhouse gas (GHG) calculations as modelled by DayCent, reported as grams CO_2 equivalents (g CO_2e). **A positive value indicates a flux to the atmosphere while a negative value indicates a GHG sink**. Values shown are the cumulative annual averages for the period 2015–2025. Total CO_2e values include the added compost.

Treatment	N_2O CO_2e (g m^2)	CH^4 CO_2e (g m^2)	Soil CO_2e (g m^2) Compost included	Total CO_2e
WF 0	357.7	−26.5	1244.2	**1575.4**
WF Fert	601.0	−5.0	926.2	**1522.2**
WF 10,000 kg	398.2	−20.0	773.9	**1152.1**
WF 25,000 kg	487.5	−11.4	220.1	**696.2**
WF 50,000 kg	678.6	2.2	−628.9	**51.9**
WWPF 0	379.6	−17.7	1070.6	**1432.5**
WWPF Fert	527.1	−15.9	845.8	**1356.9**
WWPF 10,000 kg	450.2	−5.0	245.0	**690.2**
WWPF 25,000 kg	632.5	9.9	−641.1	**1.3**
WWPF 50,000 kg	970.8	29.2	−2072.5	**−1072.5**

Table 6
Greenhouse gas (GHG) calculations as modelled by COMET-Farm, reported as grams CO_2 equivalents (g CO_2e). **A positive value indicates a flux to the atmosphere while a negative value indicates a GHG sink**. Values shown are the cumulative annual averages for the period 2015–2025. Total CO_2e values include the added compost.

Treatment	N_2O CO_2e (g m^2)	CH^4 CO_2e (g m^2)	Soil CO_2e (g m^2) Compost included	Total CO_2e
WF 0	6.3	NA	−5391.48	**−5385.2**
WF Fert	18.0	NA	−58.07	**−40.0**
WF 10,000 kg	9.9	NA	−510.94	**−501.1**
WF 25,000 kg	17.3	NA	4368.98	**4386.3**
WF 50,000 kg	27.6	NA	10045.44	**10073.1**
WWPF 0	7.3	NA	−4538.91	**−4531.7**
WWPF Fert	12.7	NA	−2045.22	**−2032.5**
WWPF 10,000 kg	10.9	NA	−1072.15	**−1061.3**
WWPF 25,000 kg	18.7	NA	2413.15	**2431.8**
WWPF 50,000 kg	29.8	NA	7570.49	**7600.2**

across all treatments. Annually, net emissions were predicted to be between −5.98 T $ha^{-1}yr^{-1}$ in the WF unfertilized control, and +11.19 T $ha^{-1}yr^{-1}$ in the WF 50,000 kg treatment.

*3.4. Model performance DayCent & COMET-farm **(Case study 2: Puyallup vegetable trial)***

For the DayCent model, treatment type (**p = 0.01;** Table S3) was a significant predictor of the modelled SOC values. For the observed values, treatment type did not explain the variance in the measured values. The extent to which the modelled values were consistent with the observed values was highly dependent on treatment for both the DayCent and the COMET-Farm models. Across the whole DayCent model, 50% of the variation in observed values could be explained by the modelled values, with an average RMSE of 593 g C m^2 and an NRMSE of 9%. Total model bias (RE) was 0.02%. Overall, the DayCent modelled values tended to describe the data better than the mean of the observed data (EF, Table 7). With only two measured and modelled values at each treatment level, R^2 values are not considered a reliable metric for model evaluation. However, at the treatment level, RMSE values were 168 g C m^2 for the broiler litter treatment, and 822 g C m^2 for the on-farm compost treatment, with NRMSE values < 12%. The DayCent model tended to slightly overestimate the average SOC content for both treatments, however the observed values for the on-farm compost treatment demonstrated a far larger variance than for the broiler litter (Fig. 6).

For the COMET-Farm model, almost none of the variation in observed values could be explained by the modelled values, with an average RMSE of 1874 g C m^2 and an NRMSE of 27.4%. In interpreting these evaluations, it is important to note that the initial SOC value of 5.2% automatically imputed by the model from the SSURGO database likely significantly influenced the overall error in the COMET-Farm estimates. Total model bias (RE) was 23%. Across the whole model, the negative efficiency factor of −4.79 indicates a poor model fit (Table 7). The model was more accurate for the on-farm compost treatment (NRMSE 9.25%; Table 7) than the broiler litter (NRMSE 41%). Whether the COMET-Farm tool tended to over or underestimate the observed SOC values was highly dependent on treatment; the tendency being to overestimate SOC content for all organic matter additions (Table 6; Fig. 7).

*3.5. DayCent model predictions **(Case study 2: Puyallup vegetable trial)***

3.5.1. Soil organic carbon storage 2006–2015

The broiler litter compost was predicted to lose ~7%, and the on-farm compost to gain ~11% SOC from 2006 to 2015, mostly because of the total input of carbon being far greater in the on-farm compost treatment (Fig. 8). However, the DayCent model predicted 160% more

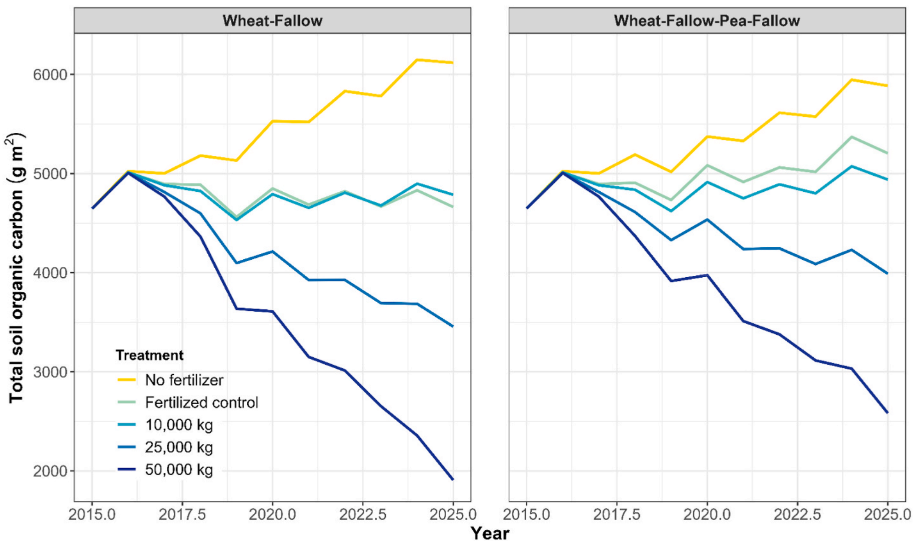

Fig. 5. *Case study 1: Wilke Farm dryland wheat* COMET-Farm simulated soil organic carbon (g m^2) for 2015–2025 between crop rotations and treatments (applied in kg ha^{-1}). WF = Wheat-fallow, WFPF = Wheat-Fallow-Pea-Fallow. **Note** that the initial compost amount has not been removed from these predictions. No 95% confidence interval was able to be calculated as the COMET-Farm output only provides a mean SOC change by year.

Table 7
Case study 2: Puyallup vegetable trial model performance metrics for each level of experimental complexity for the DayCent and COMET-Farm models. Includes root mean square error (RMSE, g C m^2), normalized root mean square error (NRMSE, % observed mean), model efficiency (EF) and relative error (RE).

	Model, Treatment	n	RMSE	NRMSE (%)	R^2	EF	RE
Model	**DAYCENT**	4	593	8.69	0.5	0.149	−0.02
Treatment	Broiler Litter	2	168	2.69	NA	0.529	0
	On Farm Compost	2	822	11.1	NA	−0.597	−0.05
Model	**COMET**	4	1874	27.4	0	−4.79	0.237
Treatment	Broiler Litter	2	2559	41.1	NA	−109	0.411
	On Farm Compost	2	688	9.25	NA	−0.11	0.092

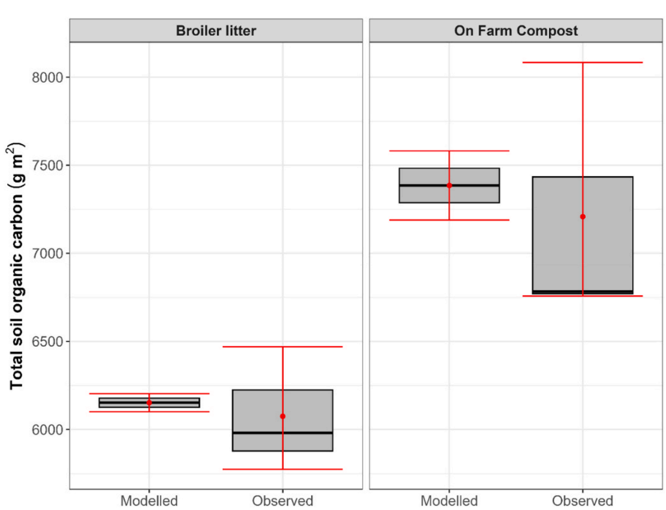

Fig. 6. *Case study 2: Puyallup vegetable trial* observed vs DayCent modelled SOC (g m^2) values at the treatment level. The red lines indicate the bootstrapped 95% confidence interval for the mean of each group. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

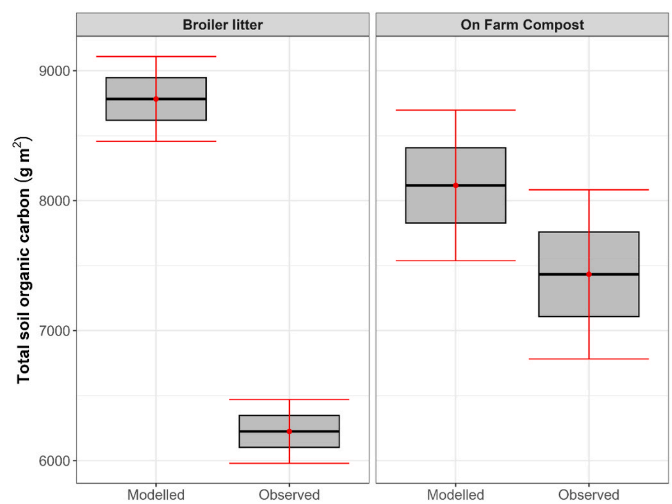

Fig. 7. *Case study 2: Puyallup vegetable trial* observed vs COMET-Farm modelled SOC (g m^2) values at the treatment level. The red lines indicate the bootstrapped 95% confidence interval for the mean of each group. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

soil organic carbon emissions from the broiler litter compared with the on-farm compost treatment (Table 8).

3.5.2. Net greenhouse gas emissions 2006–2015

The broiler litter treatment was predicted to be to net-emitting (0.61 T ha^{-1} yr^{-1}), while the on-farm compost was predicted to sequester 0.14 T ha^{-1} yr^{-1} (Table 8). While the on-farm compost had significantly higher N_2O emissions, the greater input of soil carbon from the amendment increased the net sequestration amount. Without a "business-as-usual" baseline, it is not possible to compare the climate impact of amendments in this system to conventional practices.

3.6. COMET-farm predictions (*Case study 2: Puyallup vegetable trial*)

3.6.1. Soil organic carbon storage 2006–2015

The broiler litter compost treatment was predicted to lose ~16.4%, and the on-farm compost ~26% SOC from 2006 to 2015 (Table 9;

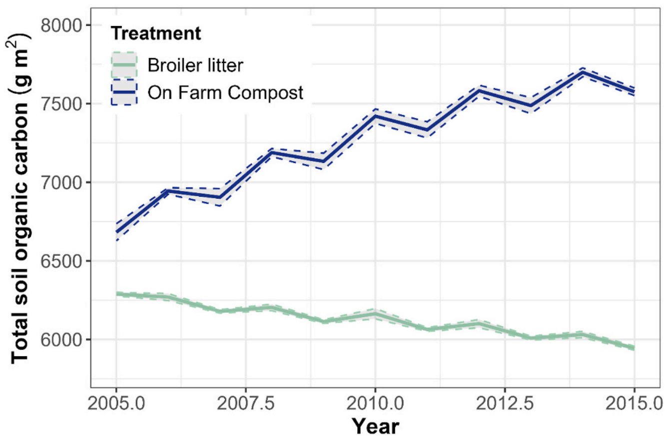

Fig. 8. DayCent simulated soil organic carbon (g m^2) for 2005–2015 between treatments (Trt, applied in kg ha^{-1}). **Note** that the initial compost amount has not been removed from these predictions. Shaded areas represent the 95% confidence interval of the mean by treatment.

Table 8
Greenhouse gas (GHG) calculations as modelled by *DayCent*, reported as grams CO_2 equivalents (g CO_2e). **A positive value indicates a flux to the atmosphere while a negative value indicates a GHG sink**. Values shown are the cumulative annual averages for the period 20*06*–20*15*. Total CO_2e values include the added compost.

Treatment	N_2O CO_2e (g m^2)	CH^4 CO_2e (g m^2)	Soil CO_2e (g m^2)	Total CO_2e
Broiler litter	1345.0	49.0	1564.2	**2958.1**
On-Farm compost	1933.7	88.5	−2592.1	**−569.9**

Table 9
Greenhouse gas (GHG) calculations as modelled by COMET-Farm, reported as grams CO_2 equivalents (g CO_2e). **A positive value indicates a flux to the atmosphere while a negative value indicates a GHG sink**. Values shown are the cumulative annual averages for the period 2005–2015. Total CO_2e values include the added compost.

Treatment	N_2O CO_2e (g m^2)	CH^4 CO_2e (g m^2)	Soil CO2e (g m^2)	Total CO_2e
Broiler litter	332.5	NA	5881.5	**6213.9**
On-Farm compost	992.8	NA	9499.4	**10492.1**

Fig. 9). Compared with the DayCent model, COMET-Farm predicted a steady decrease in soil carbon in both treatments despite annual amendment additions.

3.6.2. Net greenhouse gas emissions

COMET-Farm predicted that both treatments would be net-emitting (Table 9). The on-farm compost treatment was responsible for significantly higher N_2O and soil C emissions. The total emissions between the broiler litter and the on-farm compost were 2.58 T ha^{-1} yr^{-1} and 10.23 T ha^{-1} yr^{-1} respectively. Without a “business-as-usual” baseline, it is not possible to compare the climate impact of amendments in this system to conventional practices.

*3.7. All model predictions (inc. Cool Farm & WaCSE (**case studies 1 & 2;** Table 10)*

For case study 1: Wilke Farm dryland wheat, the DayCent (−0.29 T ha^{-1} yr^{-1}) and Cool-Farm (−0.03 T ha^{-1} yr^{-1}) tools predicted net SOC sequestration between ‘business as usual’ (synthetic fertilizer application) and application of 50,000 kg ha^{-1} of compost under the wheat-fallow rotation. *It is important to note that Cool Farm did not add carbon to the soil pool from amendments, nor increase crop growth in response to organic amendments, therefore the SOC sequestration benefit was static between conditions.* Under the practice change scenario CPS 590 (Nutrient Management), WaCSE predicted a 0.49 T ha^{-1} yr^{-1} loss of SOC, and COMET-Farm predicted 9.98 T ha^{-1} yr^{-1} loss of SOC. Increased N_2O emissions between BAU and high compost addition ranged 0.04 T ha^{-1} yr^{-1} for COMET-Farm, 0.07 T ha^{-1} yr^{-1} for DayCent, and 1.06 T ha^{-1} yr^{-1} for Cool Farm. WaCSE estimated a 0.12 T ha^{-1} yr^{-1} reduction in N_2O emissions under the practice change.

For Case study 2: Puyallup vegetable trial, only the DayCent model (−1.02 T ha^{-1} yr^{-1}) predicted net SOC sequestration between broiler litter and on-farm compost additions. SOC loss for COMET-Farm was predicted to be 0.36 T ha^{-1} yr^{-1}. Net C emissions of 0.15 T ha^{-1} yr^{-1} were recorded for the WaCSE simulated practice change. Cool Farm did not alter the amount of C added to the soil pool despite the large difference in carbon added with on-farm compost compared with the broiler litter. Increased N_2O emissions between broiler litter and on-farm compost addition ranged between 1.06 T ha^{-1} yr^{-1} for Cool Farm, 0.58 T ha^{-1} yr^{-1} for DayCent, and 0.009 T ha^{-1} yr^{-1} for COMET-Farm. WaCSE estimated a 0.12 T ha^{-1} yr^{-1} reduction in N_2O emissions under the broiler litter practice change, and 0.09 T ha^{-1} reduction under the on-farm compost.

4. Discussion

The evaluation of the first two tools, COMET-Farm and DayCent, revealed that COMET-Farm, although relatively easier to use, exhibited a higher level of bias compared to DayCent. In the Wilke Farm dryland wheat study, DayCent explained approximately 50% more of the variation in the observed values compared to COMET-Farm. The disparity in performance was even more pronounced in the second case study. The difference in performance can be attributed to the lower data requirements and limited capability of COMET-Farm to incorporate site-level metrics, such as initial soil organic carbon (SOC) content. Models that require minimal input data may trade off ease-of-use for increased bias and uncertainty. The Cool Farm and WaCSE tools, for example, exhibited different abilities to estimate emissions and sequestration due to variations in their input parameters and data requirements. Cool Farm was found to be unsuitable for estimating SOC sequestration benefits from organic amendments because it did not add carbon to the soil pool despite large differences in the amount of C added with each differing amendment. Online tools may be more suitable for rapid estimates of greenhouse gas (GHG) emissions reduction potential over larger areas or groups of farms in the short-term, but caution should be exercised when relying on these tools for site-specific estimates. When used in isolation, these tools may not be suitable for establishing priority funding for climate-smart agriculture (CSA) initiatives at the individual farm scale.

4.1. Using field measured data to initialize models may reduce prediction biases

In the first case study (dryland wheat), the inability to initialize COMET-Farm using measured SOC content did not significantly impact the bias in model predictions, as the initialized value derived from the Soil Survey Geographic Database (SSURGO) was close to the measured value. However, in the second case study (the Puyallup vegetable trial), the SOC initialization value was more than 4% higher than the observed value, resulting in a significant source of bias in all subsequent predictions. Because of this initialization error, and given the high sand content at this site, it is possible that the model calculated the SOC content to be at capacity: sandy soils having less capacity to store SOC (Georgiou et al., 2022). To mitigate such biases, the ability to initialize the COMET-Farm model using measured data, where available, could prove crucial. This approach may help address the challenges

Table 10
Average yearly emissions (CO_2 eq) by model for the periods. **a)** Case study 1: Wilke Farm dryland wheat 2016–2025 and **b)** Case study 2: Puyallup vegetable trial: 2006–2015. Only one year of emissions data was able to be calculated for the COOL-Farm and WaCSE tools. **Because the WaCSE tool does not use a baseline scenario for Case study 1, care should be taken in directly comparing these numbers with the other estimates.** *BAU = 'business as usual', synthetic fertilizer application. For the Wilke Farm dryland wheat trial, data from the Wheat-Fallow rotation is reported. No CO_2e emissions costs from compost or fertilizer production, transportation, or application were included in this analysis. **A positive value indicates a flux to the atmosphere** while a negative value indicates a GHG sink. Estimations from all reported tools are correct as calculated in October 2023.

Case study 1: Wilke Farm dryland wheat								
	CO_2		N_2O		CH_4		Total CO_2 eq	
	(g m^2 yr^{-1})		(g m^2 yr^{-1})		(g m^2 yr^{-1})		(T ha^{-1} yr^{-1})	
Scenario	*BAU	50,000 kg ha^{-1}	*BAU	50,000 kg ha^{-1}	*BAU	50,000 kg ha^{-1}	*BAU	50,000 kg ha^{-1}
DAYCENT	92.62	−62.89	60.10	67.86	−0.5	0.02	1.52	0.049
COMET	−5.80	1004.5	1.8	2.76	NA	NA	−0.04	10.07
Cool Farm	−19.05	−22.31	14.04	120.2	0.00	0.00	−0.05	97.89
WaCSE (practice change)	*NA*	*49.42*	*NA*	*−12.35*	*NA*	*0.00*	*NA*	*0.37*
Case study 2: Puyallup vegetable trial								
	CO_2 (g m^2 yr^{-1})		**N_2O (g m^2 yr^{-1})**		**CH_4 (g m^2 yr^{-1})**		**Total CO_2 eq (T ha^{-1} yr^{-1})**	
Scenario	**Broiler**	**On-Farm**	**Broiler**	**On-Farm**	**Broiler**	**On-Farm**	**Broiler**	**On-Farm**
DAYCENT	156.42	−259.2	134.5	193.37	4.9	9.9	2.95	−0.55
COMET	588.1	949.9	33.2	99.2	0.00	0.00	6.21	10.49
Cool Farm	38.79	38.79	235.03	242.28	0.00	0.00	2.73	2.81
WaCSE (practice change)	*19.76*	*34.59*	*−12.35*	*−9.88*	*0.00*	*0.00*	*0.07*	*0.25*

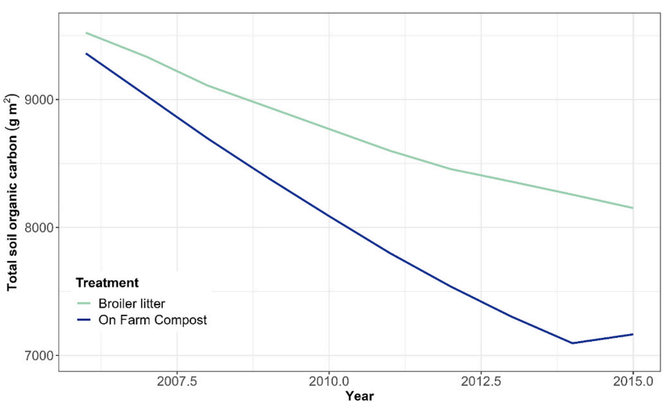

Fig. 9. *Case study 2: Puyallup vegetable trial* COMET-Farm simulated soil organic carbon (g m^2) for 2006–2015 treatments (Trt, applied in kg ha^{-1}). **Note** that the initial compost amount has not been removed from these predictions. No 95% confidence interval was able to be calculated as COMET-Farm only provides a mean SOC change by year.

encountered in reducing model bias. Consequently, caution should be exercised in accepting the predictions generated by COMET-Farm for the Puyallup vegetable trial.

4.2. SOC sequestration estimates likely differ under compost application owing to unquantified soil moisture and nutrient benefits, particularly in drylands

For case study 1 (Wilke Farm dryland wheat), the observed values were sensitive to rotation type and treatment, and the modelled values for both DayCent and COMET-Farm exhibited similar sensitivity. Both models tended to approximate the observed values more accurately for the fertilized control treatment compared to the organic amendments, and for the wheat-fallow-pea-fallow rotation compared to the wheat-fallow rotation. For both models, this may be in part because DayCent can be insensitive to changes in soil moisture content that result from surface residues (Wang et al., 2018), and in both case studies, increases in soil moisture were observed under the compost additions (Pritchett et al., 2011). Because soil moisture exerts significant pressure on soil organic matter decomposition (Serna-Chavez et al., 2013), not accounting for significant changes in the model would increase bias for this metric. Combinations of rotation and organic amendment treatments displayed a random pattern of bias for each model. Despite their similar sensitivity to crop rotation and treatment (though not in the same rotation-treatment combinations), the predictions of soil organic carbon (SOC) sequestration differed significantly between the models. DayCent predicted SOC sequestration with higher compost application in the wheat-fallow-pea-fallow rotations and reduced nitrous oxide (N_2O) emissions with compost application. On the other hand, COMET-Farm predicted more SOC under the fertilized control treatment and less SOC under compost, while also predicting reduced N_2O emissions with compost. It is worth noting that while this study did not specifically evaluate the management of soil nutrient inputs contributing to SOC storage and emissions by the models, DayCent and COMET-Farm appeared to handle these inputs differently. DayCent has been parameterized to simulate a wide range of cropping systems in the United States (Del Grosso et al., 2006; McClelland et al., 2021), although parameterizations have largely focused on temperate regions and may not fully account for dryland soil dynamics. Given that DayCent provides the background N_2O model for COMET-Farm, it was expected that their estimates would be more similar.

4.3. COMET-farm predicts more SOC sequestration under inorganic fertilizer than organic amendment

Existing literature supports the theory that organic amendments lead to increases in SOC content and reductions in N_2O emissions across various cropping systems, including drylands (De Rosa et al., 2018; Ding et al., 2013; Guangbin et al., 2021). That COMET-Farm predicted more SOC under inorganic fertilization and less SOC under compost, but lower N_2O emissions under compost compared to inorganic fertilizer, raises questions about whether COMET-Farm is well-equipped to accurately predict emissions reductions under organic amendments. It is important to recognize that COMET-Farm's N_2O model is under continuous development, and its high levels of uncertainty have been acknowledged (USDA Natural Resources Conservation Service, 2022). N_2O emissions predictions are among the most data-deficient areas of climate science, and better understanding of on-farm N_2O dynamics, and measurement of N_2O emissions are crucial for improving the predictive capabilities of numerous models (De Rosa et al., 2018; Del Grosso et al., 2006; Kostyanovsky et al., 2019). Additionally, on-farm N_2O emissions could not be verified using experimental data, and therefore the accuracy of the total CO_2e impact of both emissions and SOC sequestration for each dataset is assumed based on their ability to approximate measure SOC values. It is understood that this assessment does not provide a complete evaluation of model performance. Given that many climate-smart agriculture (CSA) programs recommend diversifying crop rotations, reducing fallow periods, and implementing soil amendments to optimize emissions reductions, it becomes increasingly important to assess the sensitivity of models to input parameters such as crop rotations and organic amendments.

4.4. Potential problems in relying on site-specific measurements of SOC to estimate benefits

Interestingly, the DayCent model demonstrated better explanation of the observed SOC data compared to the mean of the observed values. This finding indicates that there was a significant amount of natural variation in the observed SOC values, which is to be expected in highly heterogeneous landscapes. Unfortunately, most sampling regimes do not adequately account for this variability (Stanley et al., 2023) highlighting a potential problem in relying solely on localized soil sampling for benefits assessment (Prestele and Verburg, 2020). DayCent exhibited a higher ability to predict the variation in observed values for Case study 2: Puyallup vegetable trial (50%) compared to case study 1: Wilke Farm dryland wheat (39%), with the estimates improving significantly depending on the rotation and treatment. On the other hand, COMET-Farm could only explain 27% of the variation in the observed values for Case study 1: Wilke Farm dryland wheat and showed very limited explanatory power for Case study 2: Puyallup vegetable trial. The bias introduced by the SOC initialization value used by COMET-Farm in the Puyallup trial (case study 2) would have influenced all the emissions predictions. If the model assumes that the system is already at its maximum SOC capacity, any amendment would be interpreted as causing an increase in net emissions, resulting in a biased assessment of the benefits. All models, both simple and complex, inherently possess a level of uncertainty, with some aspects remaining unquantifiable (Rafique et al., 2015). To mitigate uncertainty, it is important to incorporate larger quantities of field data collected from diverse systems, considering both known and unknown spatial heterogeneity. This reinforces the need for establishing additional ongoing long-term research on organic amendment applications in diverse agricultural systems.

4.5. Important considerations when using model tools for SOC sequestration and emissions assessment

In addition to conducting a detailed assessment of readily available model tools, we wish to highlight some important considerations for using model tools: 1) The tiering structure, on which many platform-based tools are based, often relies on default values and simplified models, particularly in the lower tiers. This can lead to over-simplification and generalization of emissions estimates, which may not accurately capture the complexity and variability of emissions from different agricultural sectors or regions. This was made particularly clear in our assessment. And while lower tier models are not explicitly designed to make site-specific estimates, in the context of their use in CSA programs, there is a risk of underestimating or misrepresenting actual emissions, which can impact the integrity of individual programs. 2) Limited consideration of uncertainties: While the tiering structure acknowledges uncertainties associated with different tiers, it does not always provide a comprehensive treatment of uncertainties in emissions estimation. Unlike many other model tools, the COMET-Farm tool specifically attempts to estimate uncertainty in both N_2O and CO_2 emissions (USDA Natural Resources Conservation Service, 2022) and these components are undergoing continuous development. Uncertainties can arise from various sources, including data gaps, modeling assumptions, complex parameter interactions, and measurement errors. Failing to adequately address uncertainties can undermine the accuracy and reliability of emission estimates used for CSA programs, potentially leading to incorrect allocation of funding, or poor quality of the resulting carbon credits. It's essential to acknowledge that uncertainty estimation for mixed-tier online tools such as COMET-Farm and Cool Farm poses a significant challenge. Compared with higher tier models, which offer flexibility in altering singular parameters to quantify uncertainties, mixed-tier models feature a multitude of data sources, intricate parameter interactions, and dynamic variations in climate and agricultural practices, which are outside the user's control. 3) Particularly in lower tiers, the use of default values and simplified approaches creates opportunities for manipulation of emission estimates by project developers seeking to secure funding opportunities or maximize carbon credits. Inadequate attention to implementing verification processes may fail to detect deliberate or unintentional misrepresentation of emissions, undermining the environmental integrity of CSA schemes.

4.6. Limitations of current model approaches and future directions

Given the limitations of current approaches that our work highlights, alternative approaches should be explored. Specifically, multi-model ensemble frameworks may enhance decision-making for carbon sequestration and emissions reduction strategies by improving predictions and reducing uncertainties (Antle et al., 2018; Gupta et al., 2022; Semenov and Stratonovitch, 2010). These approaches involve leveraging diverse models with complementary strengths and weaknesses, using aggregation methods like averaging or weighting to combine their predictions. These approaches should also be adaptive, allowing for model updates over time. Furthermore, it is important to note that none of the models studied in this paper accounted for indirect emissions or environmental co-benefits. Because most models primarily focus on direct emissions sources, they may not fully consider the indirect emissions associated with complex supply chains or the indirect effects of mitigation activities, such as improvements in water or habitat quality and reductions in soil erosion (Ashton, 2022; Baumber et al., 2019). Neglecting to account for indirect emissions can result in an incomplete assessment of the overall carbon footprint. Similarly, failing to consider the environmental co-benefits of differential management may underestimate the full potential of these practices. Some US-based CSA programs are already displaying initiative by weighting funding criteria to include not only emissions reduction potential, but environmental co-benefits (Legislature, 2020). Table 11 details the key sources of bias encountered within our assessment and makes recommendations for improving model estimates.

4.7. Conclusions

Our findings emphasize the critical significance of using accurate initialization data to mitigate prediction biases. Further, we suggest that some of the model predictions exhibited variations due to unaccounted factors like soil moisture and nutrient benefits, underscoring the need to integrate a deeper understanding of on-farm dynamics when assessing controls on GHG emissions. It was clear that advanced process-based models offer the potential for more precise site-specific emissions simulations and predictions. However, striking a balance between model accuracy and data accessibility is difficult. For models that require low amounts of input data, increased bias, and therefore uncertainty at the site scale, may be a trade-off for ease-of-use. Platform-based tools are useful for rapid estimates of GHG emissions reduction potential in the short term and may be more relevant when applied over larger areas or groups of farms. Users should be cautious in relying on these tools for site-specific estimates, and they may not be suitable when used in isolation to establish priority funding for carbon farming initiatives. We suggest that model users will need to carefully consider several important questions when deciding whether there is a model suitable for calculating GHG emissions impacts from CSAs, and if so, which model(s) are most suitable. These questions include.

Table 11
Sources of model bias and recommendations for improvement.

Sources of model bias	Underlying mechanisms	Recommendations for improvement
Lack of accurate initialization data	Inaccurate initialization values lead to prediction biases that propagate throughout the model	Prioritize the collection of measured site-specific data for accurate model initialization. Use reliable soil databases and update them when available
Unaccounted factors in agricultural management	Models may not fully consider the impacts of agricultural management changes on other soil/ecosystems properties that influence model estimates (e.g.: increases to soil moisture and nutrient/carbon benefits from compost additions)	Develop and incorporate model parameters that account for changes in important soil metrics that indirectly influence SOC sequestration via their influence on crop growth and SOC decomposition
Uncertainties in N_2O emissions modeling	High levels of uncertainty in N_2O emissions estimates, especially for online tools	Continue the development of N_2O models in online tools to reduce uncertainty. Promote measurement and monitoring of on-farm N_2O emissions.
Limitations of localized soil sampling	Localized soil sampling may not adequately represent the variability in observed values	Conduct more extensive and diverse field data collection to reduce uncertainty. Explore the use of predictive mapping techniques to capture spatial heterogeneity
Oversimplification and generalization in lower tier models	Lower tier models often rely on default values and simplified model processes, leading to oversimplification	Develop tier-specific models with more accurate parameters, or that allow a greater flexibility in user-defined parameters. Provide clear guidelines for when and how to use lower tier models. Emphasize the need to exercise caution in applying lower-tier model estimates to individual farms

1) What is the desired outcome of the assessment (i.e.: Environmental impact assessment, regulatory compliance, farm planning and decision-making, carbon crediting and offsetting, research and education, benchmarking, and reporting, establishing baselines and targets for grants and funding applications)? Each of these outcomes require diverse data inputs and varying levels of precision for their distinct purposes. Where an absolute and accurate quantification of GHG emissions may be required for one purpose, an estimation of an expected change in emissions may be sufficient for another.
2) What data is available, and what resources are available for collecting additional field-level data?
3) What is the skill level of the user(s)?
4) Given that model tools have differing data requirements and capabilities which influence model accuracy, what level of uncertainty is acceptable? For example, this may be different for a field level assessment compared to a county level assessment that is primarily concerned with the overall directionality of the CO2e effect of a practice change.
5) If a high level of uncertainty is not acceptable, what resources are available to engage personnel capable of operating more complex models and to collect the needed data for parameterization?

Although model tools have limitations, it is crucial not to let these drawbacks become a barrier to action. Models, as tools designed to assist in decision-making, should not solely determine actions, funding decisions, or inaction in climate-smart agriculture programs. Moreover, despite the potential disruption and upfront costs associated with many climate change mitigation and adaptation actions, 'carbon farming' stands out as one of several feasible and low-cost options that is already available (IPCC., 2022). Simply,

"Waiting for better science to clarify choices can be rational, but only if the evidence accumulates faster than the situation deteriorates. Otherwise, the expected value of the new science is less than the cost of inaction" (Fischhoff, 2007).

CRediT authorship contribution statement

K.R. Ball: devised the study, conducted fieldwork, modeling and statistical analysis and wrote the manuscript. **I.C. Burke:** conducted fieldwork and provided data and contributed to manuscript development. **D.P. Collins:** conducted fieldwork and provided data and contributed to manuscript development. **C.E. Kruger:** devised the study and contributed to manuscript development. **G.G. Yorgey:** devised the study and contributed to manuscript development.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Acknowledgements and funding

The work that led to this publication was supported by the Washington State Legislature, the USDA-NIFA Organic Research and Extension Initiative Program (project nos. 2014-38421-22002 and 2019-51300-30476), and the CAHNRS Office of Research at Washington State University (Hatch project 1017286).

For the Puyallup field trials, we thank Emma Rast for her assistance in field and laboratory analysis and Liz Myhre for technical assistance, and Craig Cogger and Andy Bary for their efforts to initiate and operate a long-term organic farming research experiment at WSU Puyallup from 2003 to 2016. These trials were funded by the Western Sustainable Agriculture Research and Education Program (project no. SW-03-040), the USDA NIFA Integrated Organic Program (project no. WNP07725), and the Agricultural Research Center at Washington State University (Hatch Project 0722).

For the Davenport (Wilke Farm) field trials, we thank Rachel Zuger, Nicole Tautges, and Holly Lane.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclepro.2023.139448.

References

Alex Thumba, D., Lazarova-Molnar, S., Niloofar, P., 2022. Comparative evaluation of data requirements and level of decision support provided by decision support tools for reducing livestock-related greenhouse gas emissions. J. Clean. Prod. 373, 133886.
Anderson, K., Broderick, J.F., Stoddard, I., 2020. A factor of two: how the mitigation plans of 'climate progressive' nations fall far short of Paris-compliant pathways. Clim. Pol. 20 (10), 1290–1304.
Antle, J.M., Homann-KeeTui, S., Descheemaeker, K., Masikati, P., Valdivia, R.O., 2018. Using AgMIP regional integrated assessment methods to evaluate vulnerability, resilience and adaptive capacity for climate smart agricultural systems. Climate Smart Agriculture: Building Resilience to Climate Change 307–333.
Ashton, L., 2022. A framework for promoting natural climate solutions in the agriculture sector. Land Use Pol. 122, 106382.

Australian Government Department of Agriculture, Forestry, 2022. Budget Initiatives and Explanations of Appropriations Specified by Outcomes and Programs by Entity. Commonwealth of Australia, Canberra, ACT, Australia.

Baumber, A., Metternicht, G., Cross, R., Ruoso, L.-E., Cowie, A.L., Waters, C., 2019. Promoting co-benefits of carbon farming in Oceania: applying and adapting approaches and metrics from existing market-based schemes. Ecosyst. Serv. 39, 100982.

Brien, C.J., 2019. asremlPlus: augments ASReml-R in fitting mixed models and packages generally in exploring prediction differences. R package version 4, 1–28.

Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J., 2009. ASReml-R Reference Manual. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane.

Chandra, A., McNamara, K.E., Dargusch, P., 2017. The relevance of political ecology perspectives for smallholder Climate-Smart Agriculture: a review. Journal of political ecology 24 (1), 821–842.

Cogger, C.G., Bary, A.I., Myhre, E.A., Fortuna, A.M., Collins, D.P., 2016. Soil physical properties, nitrogen, and crop yield in organic vegetable production systems. Agron. J. 108 (3), 1142–1154.

de Freitas Netto, S.V., Sobral, M.F.F., Ribeiro, A.R.B., Soares, G.R.d.L., 2020. Concepts and forms of greenwashing: a systematic review. Environ. Sci. Eur. 32 (1), 1–12.

De Rosa, D., Rowlings, D.W., Biala, J., Scheer, C., Basso, B., Grace, P.R., 2018. N2O and CO2 emissions following repeated application of organic and mineral N fertiliser from a vegetable crop rotation. Sci. Total Environ. 637, 813–824.

Del Grosso, S.J., Parton, W.J., Adler, P.R., Davis, S.C., Keough, C., Marx, E., 2012. DayCent model simulations for estimating soil carbon dynamics and greenhouse gas fluxes from agricultural production systems. Managing agricultural greenhouse gases: coordinated agricultural research through GRACEnet to address our changing climate 241–250.

Del Grosso, S.J., Parton, W.J., Mosier, A.R., Walsh, M.K., Ojima, D.S., Thornton, P.E., 2006. DAYCENT national-scale simulations of nitrous oxide emissions from cropped soils in the United States. J. Environ. Qual. 35 (4), 1451–1460.

Ding, W., Luo, J., Li, J., Yu, H., Fan, J., Liu, D., 2013. Effect of long-term compost and inorganic fertilizer application on background N2O and fertilizer-induced N2O emissions from an intensively cultivated soil. Sci. Total Environ. 465, 115–124.

Easter, M., 2018. Interpreting DayCent Model Results from the COMET. Colorado State University. https://gitlab.com/comet-api/api-docs/-/blob/master/Interpreting%20DayCent%20model%20results%20from%20the%20COMET.docx.

Fischhoff, B., 2007. Nonpersuasive Communication about Matters of Greatest Urgency: Climate Change. ACS Publications.

Gavlak, R., Horneck, D., Miller, R.O., Kotuby-Amacher, J., 2003. Soil, Plant and Water Reference Methods for the Western Region. WCC-103 Publication, Fort Collins, CO.

Georgiou, K., Jackson, R.B., Vindušková, O., Abramoff, R.Z., Ahlström, A., Feng, W., Harden, J.W., Pellegrini, A.F.A., Polley, H.W., Soong, J.L., Riley, W.J., Torn, M.S., 2022. Global stocks and capacity of mineral-associated soil organic carbon. Nat. Commun. 13 (1), 3797.

Grossman, R.B., Reinsch, T.G., 2002. 2.1 Bulk density and linear extensibility. Methods of soil analysis: Part 4 physical methods 5, 201–228.

Guangbin, Z., Kaifu, S., Xi, M., Qiong, H., Jing, M., Hua, G., Yao, Z., Keith, P., Xiaoyuan, Y., Hua, X., 2021. Nitrous oxide emissions, ammonia volatilization, and grain-heavy metal levels during the wheat season: effect of partial organic substitution for chemical fertilizer. Agric. Ecosyst. Environ. 311, 107340.

Gupta, D., Gujre, N., Singha, S., Mitra, S., 2022. Role of existing and emerging technologies in advancing climate-smart agriculture through modeling: a review. Ecol. Inf. 71, 101805.

Hillier, J., Walter, C., Malin, D., Garcia-Suarez, T., Mila-i-Canals, L., Smith, P., 2011. A farm-focused calculator for emissions from crop and livestock production. Environ. Model. Software 26 (9), 1070–1078.

Innovation, U.R.a., 2022. Farming Futures R&D Fund.

IPCC, 2022. *Climate Change 2022: Impacts, Adaptation, and Vulnerability.* Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge, UK and New York, NY, USA).

Johansson, E.L., Brogaard, S., Brodin, L., 2022. Envisioning sustainable carbon sequestration in Swedish farmland. Environ. Sci. Pol. 135, 16–25.

Karhu, K., Gärdenäs, A.I., Heikkinen, J., Vanhala, P., Tuomi, M., Liski, J., 2012. Impacts of organic amendments on carbon stocks of an agricultural soil—comparison of model-simulations to measurements. Geoderma 189, 606–616.

Kostyanovsky, K.I., Huggins, D.R., Stockle, C.O., Morrow, J.G., Madsen, I.J., 2019. Emissions of N2O and CO2 following short-term water and n fertilization events in wheat-based cropping systems. Frontiers in Ecology and Evolution 7, 63.

Leger, A.M., Ball, K.R., Rathke, S.J., Blankinship, J.C., 2022. Mulch more so than compost improves soil health to reestablish vegetation in a semiarid rangeland. Restor. Ecol. 30 (6), e13698.

Legislature, W.S., 2020. Sustainable Farms and Fields Bill.

Li, C., Frolking, S., Crocker, G.J., Grace, P.R., Klír, J., Körchens, M., Poulton, P.R., 1997. Simulating trends in soil organic carbon in long-term experiments using the DNDC model. Geoderma 81 (1–2), 45–60.

Luján Soto, R., Martínez-Mena, M., Cuéllar Padilla, M., de Vente, J., 2021. Restoring soil quality of woody agroecosystems in Mediterranean drylands through regenerative agriculture. Agric. Ecosyst. Environ. 306, 107191.

McClelland, S.C., Paustian, K., Williams, S., Schipanski, M.E., 2021. Modeling cover crop biomass production and related emissions to improve farm-scale decision-support tools. Agric. Syst. 191, 103151.

McDonald, H., Frelih-Larsen, A., Lóránt, A., Duin, L., Pyndt Andersen, S., Costa, G., Bradley, H., 2021. Carbon Farming: Making Agriculture Fit for 2030. European Union, Luxembourg.

Moreno-García, B., Suvočarev, K., Reavis, C.W., Runkle, B.R.K., 2022. User-friendly Greehouse Gas calculators to assess water-saving practices in rice fields in Arkansas. Authorea Preprints. 1-2.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—a discussion of principles. J. Hydrol. 10 (3), 282–290.

Oldfield, E.E., Eagle, A.J., Rubin, R.L., Rudek, J., Sanderman, J., Gordon, D.R., 2022. Crediting agricultural soil carbon sequestration. Science 375 (6586), 1222–1225.

Olen, N.B., Roger, F., Brady, M.V., Larsson, C., Andersson, G.K.S., Ekroos, J., Caplat, P., Smith, H.G., Dänhardt, J., Clough, Y., 2021. Effects of farm type on food production, landscape openness, grassland biodiversity, and greenhouse gas emissions in mixed agricultural-forestry regions. Agric. Syst. 189, 103071.

Parr, J.F., Papendick, R.I., Hornick, S.B., Colacicco, D., 1989. Use of organic amendments for increasing the productivity of arid lands. Arid Soil Res. Rehabil. 3 (2), 149–170.

Paul, C., Bartkowski, B., Dönmez, C., Don, A., Mayer, S., Steffens, M., Weigl, S., Wiesmeier, M., Wolf, A., Helming, K., 2023. Carbon farming: are soil carbon certificates a suitable tool for climate change mitigation? J. Environ. Manag. 330, 117142.

Paustian, K., Easter, M., Brown, K., Chambers, A., Eve, M., Huber, A., Marx, E., Layer, M., Stermer, M., Sutton, B., 2017. Field-and farm-scale assessment of soil greenhouse gas mitigation using COMET-Farm. Precision conservation: geospatial techniques for agricultural and natural resources conservation 59, 341–359.

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

Prestele, R., Verburg, P.H., 2020. The overlooked spatial dimension of climate-smart agriculture. Global Change Biol. 26 (3), 1045–1054.

Pritchett, K., Kennedy, A.C., Cogger, C.G., 2011. Management effects on soil quality in organic vegetable systems in western Washington. Soil Sci. Soc. Am. J. 75 (2), 605–615.

Rafique, R., Kumar, S., Luo, Y., Kiely, G., Asrar, G., 2015. An algorithmic calibration approach to identify globally optimal parameters for constraining the DayCent model. Ecol. Model. 297, 196–200.

Semenov, M.A., Stratonovitch, P., 2010. Use of multi-model ensembles from global climate models for assessment of climate change impacts. Clim. Res. 41 (1), 1–14.

Serna-Chavez, H.M., Fierer, N., van Bodegom, P.M., 2013. Global drivers and patterns of microbial abundance in soil. Global Ecol. Biogeogr. 22 (10), 1162–1172.

Smith, P., Smith, J.U., Powlson, D.S., McGill, W.B., Arah, J.R.M., Chertov, O.G., Coleman, K., Franko, U., Frolking, S., Jenkinson, D.S.J.G., 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments, 81 (1–2), 153–225.

Stanley, P., Spertus, J., Chiartas, J., Stark, P.B., Bowles, T., 2023. Valid inferences about soil carbon in heterogeneous landscapes. Geoderma 430, 116323.

USDA Natural Resources Conservation Service, 2022. C.S.U. In: COMET-FARM. https://data.nal.usda.gov/dataset/comet-farm.

van der Voort, T.S., Verweij, S., Fujita, Y., Ros, G.H., 2023. Enabling soil carbon farming: presentation of a robust, affordable, and scalable method for soil carbon stock assessment. Agron. Sustain. Dev. 43 (1), 22.

Wang, Y., Dou, F., Del Grosso, S., Paustian, K., Anex, R., Hons, F., 2018. DAYCENT Assessment under Multiple Treatment-Measurement Scenarios in Bioenergy Sorghum Production.

Whittaker, C., McManus, M.C., Smith, P., 2013. A comparison of carbon accounting tools for arable crops in the United Kingdom. Environ. Model. Software 46, 228–239.