GLOBAL CHANGE BIOLOGY
BIOENERGY

GCB Bioenergy (2017) **9**, 1543–1554, doi: 10.1111/gcbb.12473

# The greenhouse gas intensity and potential biofuel production capacity of maize stover harvest in the US Midwest

CURTIS D. JONES[1], XUESONG ZHANG[2], ASHWAN D. REDDY[1], G. PHILIP ROBERTSON[3,4,5] and ROBERTO CÉSAR IZAURRALDE[1,6]

[1]*Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA,* [2]*Joint Global Change Research Institute, Pacific Northwest National Laboratory and University of Maryland, College Park, MD 20740, USA,* [3]*Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48824, USA,* [4]*W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060, USA,* [5]*Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, USA,* [6]*Texas A&M AgriLife Research & Extension Center, Temple, TX 76502, USA*

## Abstract

Agricultural residues are important sources of feedstock for a cellulosic biofuels industry that is being developed to reduce greenhouse gas emissions and improve energy independence. While the US Midwest has been recognized as key to providing maize stover for meeting near-term cellulosic biofuel production goals, there is uncertainty that such feedstocks can produce biofuels that meet federal cellulosic standards. Here, we conducted extensive site-level calibration of the Environmental Policy Integrated Climate (EPIC) terrestrial ecosystems model and applied the model at high spatial resolution across the US Midwest to improve estimates of the maximum production potential and greenhouse gas emissions expected from continuous maize residue-derived biofuels. A comparison of methodologies for calculating the soil carbon impacts of residue harvesting demonstrates the large impact of study duration, depth of soil considered, and inclusion of litter carbon in soil carbon change calculations on the estimated greenhouse gas intensity of maize stover-derived biofuels. Using the most representative methodology for assessing long-term residue harvesting impacts, we estimate that only 5.3 billion liters per year (bly) of ethanol, or 8.7% of the near-term US cellulosic biofuel demand, could be met under common no-till farming practices. However, appreciably more feedstock becomes available at modestly higher emissions levels, with potential for 89.0 bly of ethanol production meeting US advanced biofuel standards. Adjustments to management practices, such as adding cover crops to no-till management, will be required to produce sufficient quantities of residue meeting the greenhouse gas emission reduction standard for cellulosic biofuels. Considering the rapid increase in residue availability with modest relaxations in GHG reduction level, it is expected that management practices with modest benefits to soil carbon would allow considerable expansion of potential cellulosic biofuel production.

*Keywords:* agricultural residues, bioenergy, biofuels, corn stover, crop modeling, Environmental Policy Integrated Climate, greenhouse gas intensity, soil organic carbon, sustainability, terrestrial ecosystem modeling

*Received 16 December 2016; revised version received 13 June 2017 and accepted 19 June 2017*

## Introduction

The 2007 US Energy Independence and Security Act (EISA), passed amid growing concern regarding increasing greenhouse gas (GHG) levels in the atmosphere and national dependence on imported fuels, mandated the production of 79.5 billion liters per year (bly) of advanced biofuel including 60.6 bly of cellulosic biofuel by 2022. An advanced biofuel is a biofuel derived from a nonmaize starch renewable source, while a cellulosic biofuel is a biofuel derived from renewable cellulose, hemicellulose, or lignin sources. Additionally, life cycle GHG emissions reductions of at least 50% and 60% relative to 2005 petroleum are required in order to qualify as advanced and cellulosic biofuels, respectively. The use of agricultural residues for cellulosic ethanol production has been identified as the most feasible near-term option to supply cellulosic feedstock to an emergent biofuels industry (Langholtz *et al.*, 2016) as other feedstock options require further time for development before large-scale availability.

Correspondence: Curtis D. Jones, tel. +1 202 813 9145, e-mail: cujo@umd.edu

Further, agricultural residue collection occurs on cultivated lands, avoiding both conversion of new lands to agriculture and displacement of existing feed and fiber producing lands (Solomon, 2010).

Despite the advantages of utilizing agricultural residues as a cellulosic feedstock, there remain concerns regarding the sustainability of residue removal from agricultural systems. Crop residues in cropping systems serve important functions related to nutrient cycling, soil erosion, soil structure, and soil water retention and provide habitat for ground-dwelling organisms. Excessive residue removal can accelerate soil erosion (Lal, 1976), deplete soil organic carbon (SOC) pools (Karlen *et al.*, 1994; Lal & Pimentel, 2007), and reduce the production potential of soils (Lal, 2006). Conversely, many other studies have demonstrated a capacity for sustainable crop residue harvesting, especially under no-till (NT) management where decomposition is slowed and near-surface SOC usually accumulates (Gollany *et al.*, 2011; Machado, 2011; Robertson *et al.*, 2011; Adler *et al.*, 2015).

It is evident that the relationship between soils and crop residues is strongly influenced by site-specific conditions as well as previous and current land management. As such, suitable rates of residue removal must be determined with due consideration for local topography, soil, land management, and climate. Thus, estimating the impacts of widespread residue removal requires detailed cropping systems models that can incorporate spatially explicit management scenarios to simulate locally relevant, scalable outcomes. For example, Muth *et al.* (2013) modeled US maize (*Zea mays* L.) production at 10–100 m resolution to estimate that 151 Tg $yr^{-1}$ of crop residues could be sustainably harvested in the United States based on simulated impacts on soil erosion, SOC, soil water, and soil temperature. Tan *et al.* (2012) estimated a more modest sustainable residue harvest of 31 Tg $yr^{-1}$ for the United States to avoid reducing current SOC levels. The 2016 Billion Ton Study (Langholtz *et al.*, 2016) estimates that 27–106 Tg $yr^{-1}$ of residue is available at a residue price of \$44–88 per Mg. Each of these studies identified the US Corn Belt as the dominant source of sustainably available residue, providing 60–85% of that available nationally (Tan *et al.*, 2012; Muth *et al.*, 2013; Langholtz *et al.*, 2016).

In contrast, Liska *et al.* (2014) estimated that harvesting 6 Mg residue $ha^{-1}$ $yr^{-1}$ from across the US Corn Belt over a 5- to 10-year period would result in an average C loss of 0.47–0.66 Mg C $ha^{-1}$ $yr^{-1}$ and life cycle emissions of 74–95 g $CO_2$-eq $MJ^{-1}$, similar to emissions from petroleum (94 g $CO_2$-eq $MJ^{-1}$). Such findings would eliminate maize residue as a viable feedstock source for meeting EISA cellulosic biofuel production standards. However, others (Robertson *et al.*, 2014; Sheehan *et al.*, 2014) noted that the model used was driven only by temperature, yield, and initial SOC, excluding important factors such as topography, soil texture, soil moisture, fertilization rates, and tillage, which are also known to affect SOC stocks. Additionally, the analysis focused on short-term responses, considered SOC changes only in the upper 30 cm of soil, and included litter C, which is derived from all unharvested crop biomass, along with soil humus C in SOC change calculations.

More biophysically based analyses have been conducted to assess the GHG impacts of residue-derived biofuels. The DayCent model, for example, is a physically based model that Campbell *et al.* (2014) demonstrated to simulate crop productivity, SOC responses and nitrous oxide fluxes reasonably well through an evaluation at five sites in the US Midwest. The DayCent model has been applied at county scale to assess the GHG intensities of corn stover-derived ethanol from three representative counties in the United States (Dwivedi *et al.*, 2015) as well as the entire United States (Hudiburg *et al.*, 2016), each considering SOC changes in the upper 30 cm of soil. Alternatively, LeDuc *et al.* (2017) applied the Environmental Policy Integrated Climate (EPIC) model at 30 m resolution and considering SOC changes in the whole soil profile to assess the productivity and SOC impacts of producing corn stover feedstocks on Conservation Reserve Program lands in Iowa.

While many different approaches have been used to assess the GHG intensity of corn stover-derived biofuels, here we intend to improve upon these estimates for the US Midwest in terms of a combination of modeling approach, modeling resolution, soil depth considered, and modeling region. The use of mechanistic models at fine resolution has been recommended for assessing large-scale cropland C budgets (Smith *et al.*, 2012). Zhang *et al.* (2014a) demonstrated improved EPIC estimates of county-level yield using finer-scale (1 : 24 000) SSURGO soil data rather than coarser resolution (1 : 250 000) State Soil Geographic soils data as well as considerable differences in net ecosystem production between the two approaches. Similarly, large differences in simulated SOC responses have been shown with soil datasets of differing spatial resolution using the DNDC model in China (Zhang *et al.*, 2014b). Using a number of different models, Izaurralde *et al.* (2001) demonstrated that finer-scale simulations produced more robust estimates of SOC sequestration rates, particularly for more heterogeneous regions. Zhang *et al.* (2015) used fine-scale EPIC simulations and demonstrated reasonable estimates of yield compared to county-level National Agricultural Statistics Service (NASS) yields, which also showed improved estimates compared to previous EPIC

yield simulations at coarser resolution (Thomson *et al.*, 2005), as well as reasonable simulation of C dynamics compared to county-level cropland C budgets. Additionally, consideration of SOC change in deeper soil layers is meaningful for accurate accounting of the SOC impacts of residue harvesting. While the concentration of SOC is generally much higher in upper soil layers, deeper soil layers have considerable capacity for SOC storage or loss (Lorenz & Lal, 2005). Analyses have indicated the importance of SOC changes at greater depths, demonstrating that the SOC changes following changes in tillage management can be offset when depths beyond 30–40 cm are considered (Angers & Eriksen-Hamel, 2008; Luo *et al.*, 2010; Powlson *et al.*, 2014). Similarly, Guo & Gifford (2002) demonstrated significant SOC changes at depths in some cases beyond 100 cm following land use conversions. Hence, inclusion of SOC change at greater depths is expected to provide more representative estimates of the SOC impacts of residue management options.

Here, we utilize an improved modeling approach and analysis to better estimate the life cycle emissions of maize residue-derived biofuels from the US Midwest. We use a spatially explicit process-based terrestrial ecosystems model, the EPIC model (Williams, 1995; http://epicapex.tamu.edu/epic/), which has been used previously for modeling potential biofuel producing landscapes in the US Midwest (Zhang *et al.*, 2010, 2015; Egbendewe-Mondzozo *et al.*, 2013; Gelfand *et al.*, 2013) and has been demonstrated to skillfully simulate observed yield ($R^2 > 0.69$) and soil C ($R^2 > 0.89$) responses of medium- and long-term experiments (Izaurralde *et al.*, 2006). The EPIC model simulates soil C cycling similarly to the Century model (Parton *et al.*, 1994), splitting soil C into slow, passive, and biomass pools stratified by layer and splitting litter into structural and metabolic pools (Izaurralde *et al.*, 2006). Transformations between pools are calculated on a daily basis and are regulated by soil temperature, nitrogen, water content, oxygen, and tillage (Izaurralde *et al.*, 2006). We conduct in-depth site-level model calibration and evaluation to ensure modeling accuracy and apply this modeling framework to assess the GHG impact of residue harvesting in the US Midwest considering field-scale differences in maize productivity, topography, soil texture, fertilizer rates, tillage, and SOC change. Our specific goals are to (i) provide more accurate and representative quantification of the GHG impacts of maize stover harvesting and the potential capacity for cellulosic biofuel production in the US Midwest; and (ii) assess the impact of removal rate, location, and estimation methodology for determining SOC change on the GHG impacts and availability of maize stover.

## Materials and methods

### *Field experiments, model calibration, and evaluation*

We identified experiments from the literature pertinent to biofuel production from continuous maize rotations in the US Midwest in order to create a suitable dataset for calibration and evaluation of the EPIC model. Experiments were considered pertinent if they included at least two rates of residue removal, included a rotation of continuously cropped maize, and included at least an initial and final SOC measurement taken to a depth of at least 15 cm. While several experiments reported bulk density measurements, missing bulk densities were estimated by EPIC with the initial bulk density estimated from the Soil Survey Geographic (SSURGO; websoilsurvey.nrcs.usda.gov) database according to the dominant soil series. Bulk density values and measured SOC fractions were then used to calculate SOC on a mass density basis.

Linear mixed-modeling was conducted using the R statistical environment version 3.2.1 (R Core Team, 2015) and the LME4 package (Bates *et al.*, 2015) to assess the effect of residue removal rate as well as additional implemented treatments on the rate of SOC change. This method was selected because it has been shown to be suitable for analyzing complex data structures (Suuster *et al.*, 2011; Philibert *et al.*, 2012). Based on the experimental designs within the dataset, experimental site was used as a random effect and rate of residue removal, average depth of sampled soil layer and nitrogen fertilization rate were used as quantitative fixed effects while tillage was used as an ordinal fixed effect and irrigation was used as a categorical fixed effect. Backwards selection was conducted to eliminate nonsignificant ($\alpha = 0.05$) effects from the model according to Satterthwaite's approximation. Histogram and QQ normal plots were utilized to assess normality of residuals. The rate of SOC change was calculated for each experimental treatment and sampling depth as the annualized difference in SOC between the first and last measurements. Sampling depths were aggregated to 0–15, 15–30, 30–60, 60–90, 90–120, and 120–150 cm depths to align the measurements with the most commonly sampled depth intervals in the dataset.

To calibrate and evaluate EPIC SOC and productivity simulations, model files were created to approximate the site- and treatment-specific conditions of the corresponding experiments. Soil information was derived from the SSURGO database according to soil series and was supplemented with available experiment-specific soil information. Initial SOC levels were aligned with initial measurements and as such were excluded from model performance calculations. Weather data were obtained from on-site or nearby weather stations when available and from the re-analysis North American Land Data Assimilation System 2 (NLDAS-2; ldas.gsfc.nasa.gov/nldas/) when unavailable. Management information was derived as much as possible from experimental records, publication descriptions or correspondence with experimental investigators, with gaps in management information filled with regionally appropriate practices.

To assess model performance, measurements were first split into calibration and evaluation datasets to allow the model to

be evaluated against independent data. Calibration treatments were randomly chosen with the limitation that no more than one-third of the treatments at a particular site be utilized. This approach led to the selection of nine treatments, which accounted for 20% of the SOC measurements and 14% of the yield measurements. Simulated SOC was aligned with measured values according to measurement period and soil horizon. We utilized the HYDROPSO package (Zambrano-Bigiarini & Rojas, 2013) to calibrate influential parameters related to root development (PRMT2), water stress response (PRMT35), and tillage-related decomposition (PRMT52). Parameter suitability was evaluated based on the goodness-of-fit of yield and SOC simulations according to the average Nash–Sutcliffe coefficient of efficiency (NSE) of each measurement type. Evaluation simulations were assessed according to the coefficient of determination ($R^2$), NSE, percent bias (bias), and root mean squared error (RMSE), allowing assessment of overall model skill as well as skill for various measurement depths, residue removal rates, tillage types, or experimental sites.

## *Regional simulations*

We applied a spatially explicit integrative modeling framework for EPIC, developed by Zhang *et al.* (2010, 2015), whereby we combined multiple data layers to define modeling units. Maps of the Cropland Data Layer (CDL; Johnson & Mueller, 2010), soils (SSURGO database), and county boundaries were discretized to raster format with a grid resolution of 56 m, which is consistent with the resolution of the CDL. These maps were further combined to define over 2 million homogeneous spatial modeling units (HSMUs) with a total area of approximately $68 \times 10^6$ ha (Fig. 1). Each HSMU includes a group of grids with a unique combination of land use type and soil within the boundary of a county. Here, only those HSMUs under cultivation were included in our simulations. For each modeling unit, we further derived elevation and climate information from the Shuttle Radar Topography Mission digital elevation model (Farr *et al.*, 2007) and NLDAS2, respectively. Additional information about the spatial data used is provided in Note S1. Planting and harvesting dates and heat units required to reach maturity are important for reliable simulation of crop growth and development. We compiled these data for each state using the Soil and Water Assessment Tool potential heat unit program (available at http://swat.tamu.edu/software/potential-heat-unit-program/) and typical planting and harvesting dates of major crops in the US Midwest provided by USDA-NASS (1997). Annual nitrogen and phosphorus fertilizer application rates were estimated based on the state-level statistics from USDA-ERS (2013). Initial SOC was allocated between active, passive, and biomass pools based on the duration under cultivation according to Izaurralde *et al.* (2012). We employed the Python-based parallel computing software of Zhang *et al.* (2013) to execute EPIC in parallel for the over 2 million modeling units using the Pacific Northwest National Laboratory's Institutional Computing cluster (http://pic.pnnl.gov/) and compiled spatially explicit modeling results into relational databases linked to the HSMU map for geospatial analysis and presentation.

We applied this modeling framework to assess the residue production and the SOC response to residue harvesting across the US Midwest. We simulated no-till continuous maize rotations across the region with 0, 33, and 66% of available residue removed annually. The continuous maize system modeled across the whole US Midwest in this study, which contrasts with the maize–soybean (*Glycine max* (L.) Merr.) system that is currently more common in the US Midwest and exceeds the realistic implementation of these systems, was selected to quantify the maximum potential residue production and GHG impacts from maize-based systems. Residue removal from maize–soybean systems would be expected to have lower GHG emission intensity than from continuous maize systems (Gelfand *et al.*, 2013), although residue availability would be lower due to biennial rather than annual stover harvesting. Erosion was not considered in the SOC change assessments because limitations in available data impaired representative simulation of fine-scale erosion losses (De Vente *et al.*, 2013; Panagos *et al.*, 2015). This is a reasonable simplification because large-scale impacts of erosion on C budgets tend to be neutral in the absence of river routing (Quinton *et al.*, 2010; Nadeu *et al.*, 2015), which was not considered in this modeling framework but could be included in landscape or watershed analyses. Moreover, no-till maize producing systems in the US Midwest tend to have limited soil erosion losses even at high residue removal rates (Wilhelm *et al.*, 2007; Gregg & Izaurralde, 2010).

All residue removal scenarios were initiated with equivalent settings that were derived from a spin-up run from 1991 to 2000 as described in Zhang *et al.* (2015), and scenarios were subsequently run for 50 years with historical climate data from 1991 to 2010 to generate model outputs. Simulation periods of 10, 30, and 50 years were selected to assess the effect of residue harvesting over varying periods of time. Changes in SOC were calculated within the upper 30 cm, upper 100 cm, and total soil profile to evaluate the contributions of different soil depths to SOC change. Finally, SOC changes were calculated considering soil humus C pools (Izaurralde *et al.*, 2006) as well as litter C to identify the importance of litter C for life cycle GHG calculations.

Postprocessing of EPIC results began with removal of HSMUs with no maize yield such that it would be excluded from further analysis. This was done to ensure residue removal treatments were properly applied to the analysis area, which was not possible on nonproductive HSMUs where there was no stover available for removal for any treatments. The SOC change was calculated for each remaining HSMU by depth of soil considered, inclusion of litter C as a component of SOC and length of simulation. Additional GHG emissions components were calculated according to Liska *et al.* (2014), with system boundaries including relevant processes from corn stover collection to biorefinery processing. Emissions due to replacement fertilizer, feedstock collection, transport, and conversion to ethanol were assumed to be fixed at 30 g $CO_2$-eq $MJ^{-1}$ based on Spatari & MacLean (2010). Nitrous oxide emissions were assumed to be reduced by 4.6 g $CO_2$-eq $MJ^{-1}$ under residue removal according to Liska *et al.* (2014). The EPIC model does contain process-based algorithms to simulate microbial denitrification and $N_2O$ fluxes (Izaurralde *et al.*, 2017). However,

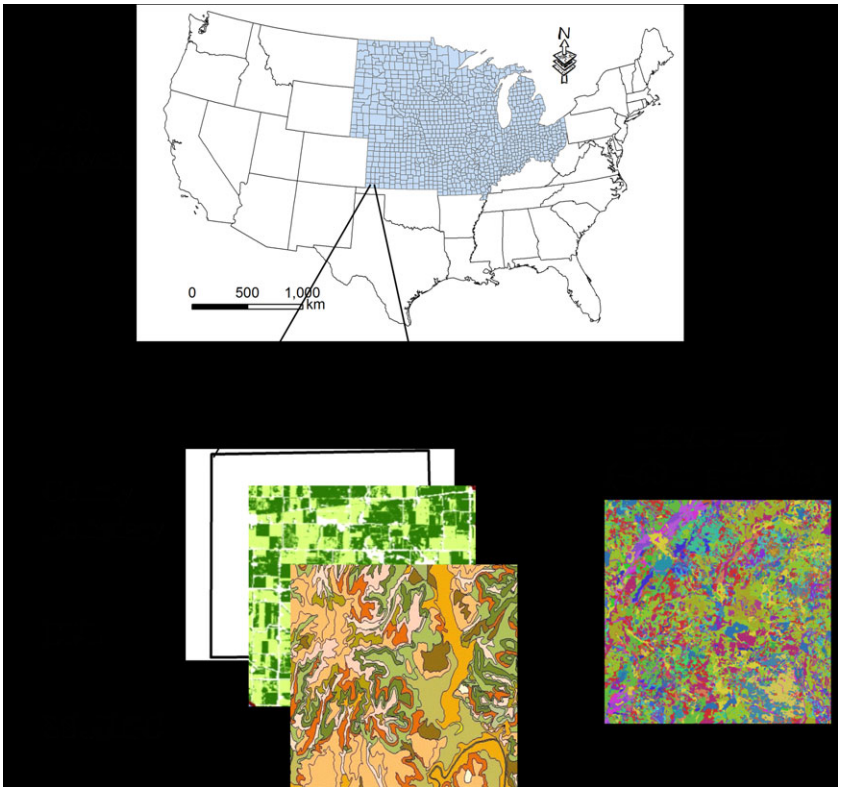



**Fig. 1** Spatial configuration of SEIMF (Adapted from Zhang *et al*. (2010)).

further efforts are needed to ensure proper model parameterization and suitability for regional applications. Thus, in this paper we use the simplified approach for estimating $N_2O$ emissions. Overall GHG intensities for the US Midwest were then calculated based on the area-weighted averages of SOC change under stover removal, SOC change without residue removal, maize stover harvested, and nitrous oxide benefit. Absolute SOC change and SOC change relative to zero stover removal were determined at the HSMU level and mapped. To determine the quantity of stover that could be collected while meeting a target GHG reduction threshold, we first ranked HSMUs in order from lowest to highest GHG emissions. We then calculated the amount of stover that could be collected while maintaining an aggregate GHG intensity with at least the target reduction. This process was conducted for GHG reduction targets of 50–60% representing biofuels that meet advanced to advanced cellulosic reduction standards.

## Results

### *Field experiments, model calibration, and evaluation*

The literature review identified ten suitable experiments in nine unique locations (Fig. 2) to provide 53 unique site treatments and 1737 SOC measurements from continuous maize rotations in the US Midwest that measured SOC responses to different rates of residue removal (Table S1). A linear mixed-effect model for rate of SOC change indicated significant effects of residue removal rate and soil depth and no significant effects for tillage type, nitrogen fertilization rate, or irrigation (Table S2). However, the dataset was primarily compiled for model calibration and evaluation. As such, experimental factors other than residue removal rate were sparsely replicated across sites, resulting in a dataset with limited power for identifying significance of factors.

We used nine treatments comprising 20% of the SOC measurements to calibrate EPIC, leaving the remaining data for model evaluation. The calibration process defined an optimal parameter set with PRMT2 set to 1.16 (Fig. S1a), PRMT35 set to 0.59 (Fig. S1b), and PRMT52 set to 5.44 (Fig. S1c), with near optimum ranges of roughly 1.15–1.20, 0.45–0.75, and 5.0–9.0, respectively. Evaluation of simulated SOC ($R^2$ = 0.83; NSE = 0.80, RMSE = 5.38 Mg C ha$^{-1}$, Bias = −1.90%, Fig. 3a) and yield ($R^2$ = 0.69; NSE = 0.65, RMSE = 2.17 Mg DM ha$^{-1}$, Bias = 7.50%, Fig. 3b) showed satisfactory agreement with measurements. This indicates good model capacity for capturing the SOC and yield dynamics of these continuous maize systems (Fig. 3). Model performance was largely consistent across soil layer depth (Fig. S2a), rate

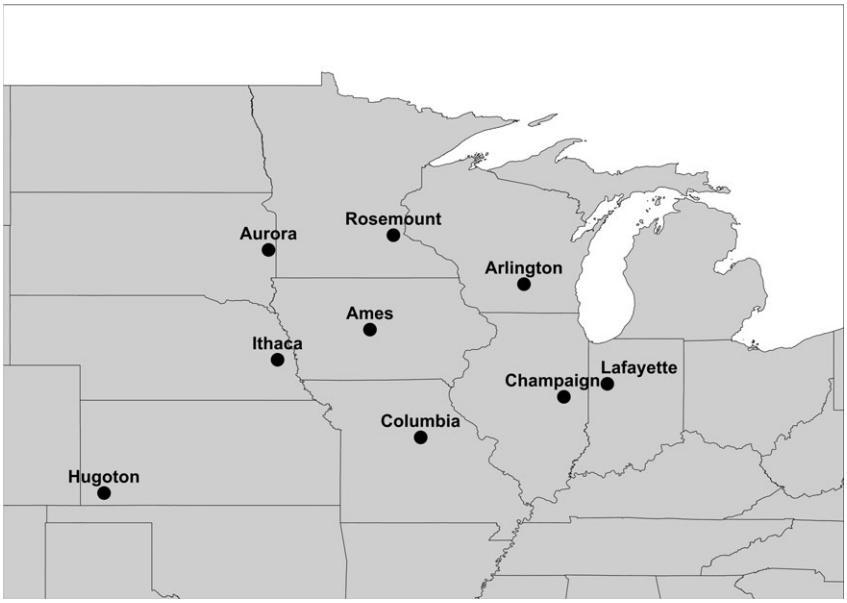



**Fig. 2** Residue removal experiment locations.

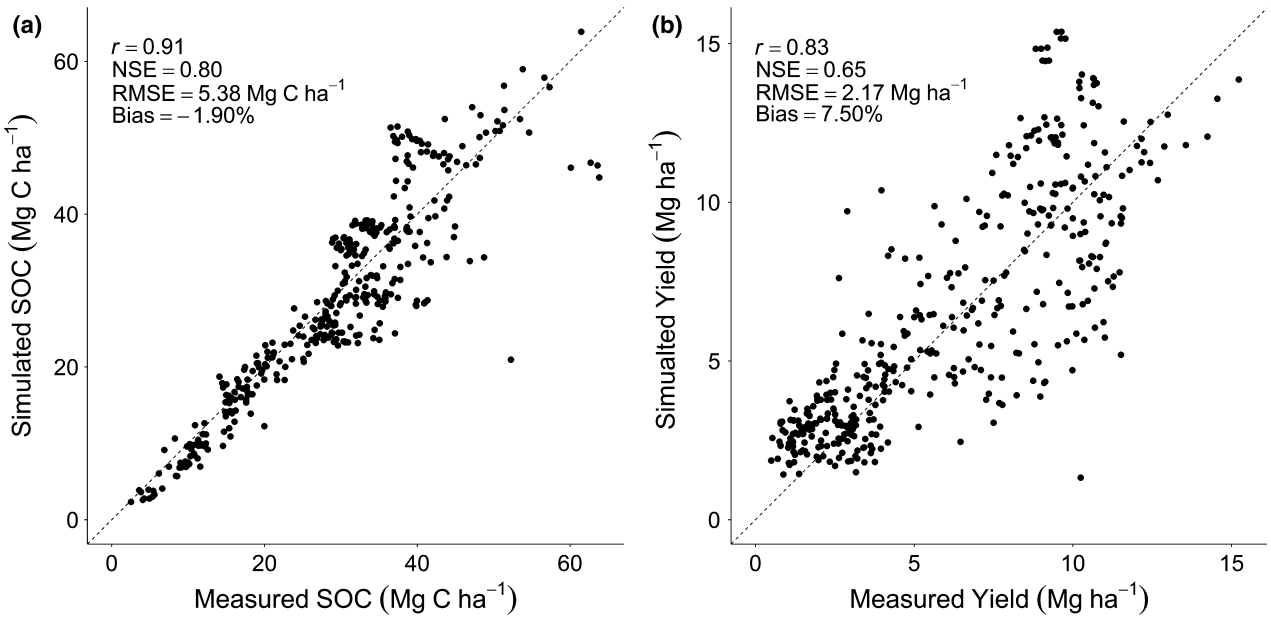



**Fig. 3** Simulated and measured SOC (a) and yield (b) from the 44 evaluation treatments from 10 experiments.

of residue removal (Fig. S2b), tillage intensity (Fig. S2c), and experimental site (Fig. S2d), although the number of measurements available at greater depths was limited. Model evaluation in terms of SOC change was poorer than in terms of absolute SOC ($R^2$ = 0.20; NSE = 0.20, RMSE = 4.20 Mg C ha$^{-1}$, Bias = −17.70%) but was comparable to similar model evaluations (Bhattacharyya *et al.*, 2013; Campbell *et al.*, 2014; Zhang *et al.*, 2015).

*Regional simulations*

The regional 50-year analysis of continuous maize that includes nonlitter SOC changes in the whole soil profile is the methodology that most representatively estimates the long-term impact of stover-based biofuel production on SOC stocks (see discussion). Under these conditions, average overall emission intensities were 46.4 and 47.0 g $CO_2$-eq MJ$^{-1}$ at residue removal rates of 33

(Fig. 4a) and 66% (Fig. 4b), respectively. As such, feedstocks derived from continuous maize systems across 68 × $10^6$ ha of the US Midwest could provide 153–310 Tg DM feedstock for 44.1–89.3 bly of ethanol production at 33–66% rates of residue removal. This range of residue harvesting would result in the average loss of

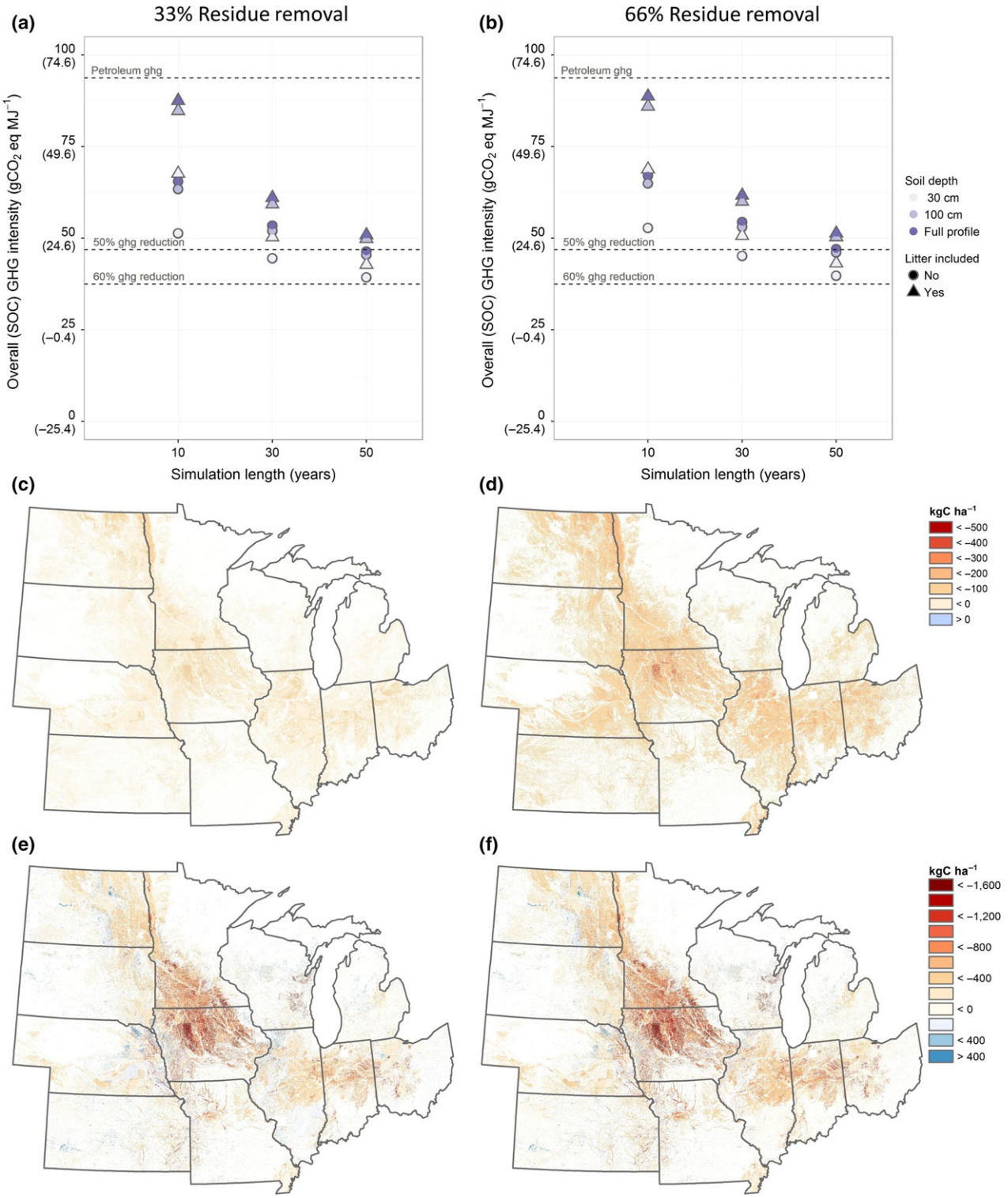



**Fig. 4** Impact of 33% (a, c, e) and 66% (b, d, f) residue removal on average GHG intensity (a, b), SOC change relative to no residue removal (c, d) and absolute SOC change (e, f). Plots c, d, e, and f are based on 50-year whole soil profile soil humus C changes.

0.078–0.163 Mg C $ha^{-1}$ $yr^{-1}$. The average aggregate GHG intensity of these feedstocks thus misses the 60% EISA reduction standards for cellulosic biofuels (37.5 g $CO_2$-eq $MJ^{-1}$) by 24 to 25%, but is close to the 50% reduction standard for advanced biofuels (46.8 g $CO_2$-eq $MJ^{-1}$), meeting the standard at 33% but not 66% residue removal. Due to site-specific variability in residue production and SOC response, some locations were capable of supplying feedstock meeting one or both of the EISA reduction standards while others were not.

While only 5.3 bly of ethanol (18.5 Tg DM residue) meeting cellulosic biofuel standards was found to be available, 89.0 bly of ethanol (308.9 Tg DM residue) was found to meet advanced biofuel standards (Fig. 5a). Thus, stover produced under these conditions could meet only 8.7% of the 2022 EISA target for cellulosic biofuels but 112% of the target for advanced biofuels, which is more than four times the target for noncellulosic advanced biofuels. This indicates a sizeable capacity to produce advanced biofuels on $67.6 \times 10^6$ ha of land but at the cost of losing SOC at an average rate of 0.163 Mg C $ha^{-1}$ $yr^{-1}$ (Fig. 4c, d). For context, an average SOC loss of 0.235 Mg C $ha^{-1}$ $yr^{-1}$ would occur with residue returned to the soil compared to 0.398 Mg C $ha^{-1}$ $yr^{-1}$ of SOC loss with residue harvested (Fig. 4e, f). The rapid increase in residue availability with small relaxations in GHG reduction level indicates management practices resulting in modest reductions in GHG emissions would considerably expand cellulosic biofuel feedstock availability. Major feedstock production areas suitable for advanced biofuels are located throughout most of Iowa and Illinois, in eastern Nebraska, South Dakota, North Dakota, and in southern and western Minnesota (Fig. 5b–e).

Detailed analysis of the regional simulations underscores the importance of soil profile depth, simulation duration, and consideration of litter C for estimating SOC responses to residue harvest. These methodological differences from earlier analyses produced average aggregate emissions ranging from 39.3 to 88.7 g $CO_2$-eq $MJ^{-1}$. Emissions were higher for shorter simulation periods, consideration of deeper portions of the soil profile, and inclusion of litter in estimates of SOC change. The impact of litter C on the SOC change attenuated over time because the immediate loss of litter C due to residue removal is ephemeral and represents a diminishing fraction of the overall C change as the change in humus C increasingly dominates the SOC response. Increasing the soil depth from 30 cm to 100 cm resulted in considerable increases in $CO_2$ emissions, while extending consideration beyond 100 cm resulted in only a slight increase in emissions, indicating analyses considering the upper 100 cm of the soil profile are sufficient for assessing residue removal impacts on SOC in comparable systems. Emission intensities were very similar at 33 and 66% residue removal rates, with only 0.8–2.8% greater intensities at the higher rate of removal owing to only slightly greater SOC loss relative to residue harvested.

## Discussion

In this work, extensive site-level experimental data from relevant US Midwest production systems were used to calibrate and evaluate the EPIC model. Overall model fit to data compared favorably with other modeling studies of yield (Izaurralde *et al.*, 2006; Wang *et al.*, 2012; Cheng *et al.*, 2014; Li *et al.*, 2015) and SOC (Cerri *et al.*, 2004; Izaurralde *et al.*, 2006; Miehle *et al.*, 2006; Lu *et al.*, 2008; Cheng *et al.*, 2014; Li *et al.*, 2015). As such, the model was parameterized and demonstrated to be suitable for assessing the productivity and SOC impacts of residue removal from these systems. This effort demonstrates the connection between pertinent experimental observations and model simulations, strengthening the soundness of the regional simulations. Through regional application of the model at fine resolution using publically available datasets, we are translating these experimental findings to policy-relevant scales using systems-level understanding in order to improve the accuracy of these estimates.

Our findings differ considerably from those of Liska *et al.* (2014) due to combined differences in modeling and analysis methodologies. To assess deviations derived from differences in modeling approach, we compared the GHG intensity estimates with 10-year simulation periods that include litter and humus SOC changes in the upper 30 cm of soil. This simulation scenario, duration, soil depth, and consideration of litter align closely with those implemented in Liska *et al.* (2014). Despite these similarities, estimates of average aggregate GHG emissions intensity were 67.6–68.8 g $CO_2$-eq $MJ^{-1}$, which is 7.3–8.6% lower than the 74.0–74.2 g $CO_2$-eq $MJ^{-1}$ estimated by Liska *et al.* (2014). Considering the impact of the depth of soil considered, we have shown that limiting the scope to the upper 30 cm severely reduces estimates of the SOC losses from these systems. This simulated response appears reasonable as experiments have demonstrated the importance of residue management to SOC stocks in deeper soil layers (Galdos *et al.*, 2009; Schmer *et al.*, 2014), although such experiments are sparse. Expanding the scope to the whole soil profile raises the GHG intensity to 87.5–88.7 g $CO_2$-eq $MJ^{-1}$. We excluded litter C in the SOC change because litter C change does not reflect the true soil humus C response to residue removal as a large difference in litter C is created from residue harvesting while a large portion of the unharvested litter will rapidly oxidize (Van Veen & Paul, 1981; Huggins *et al.*,

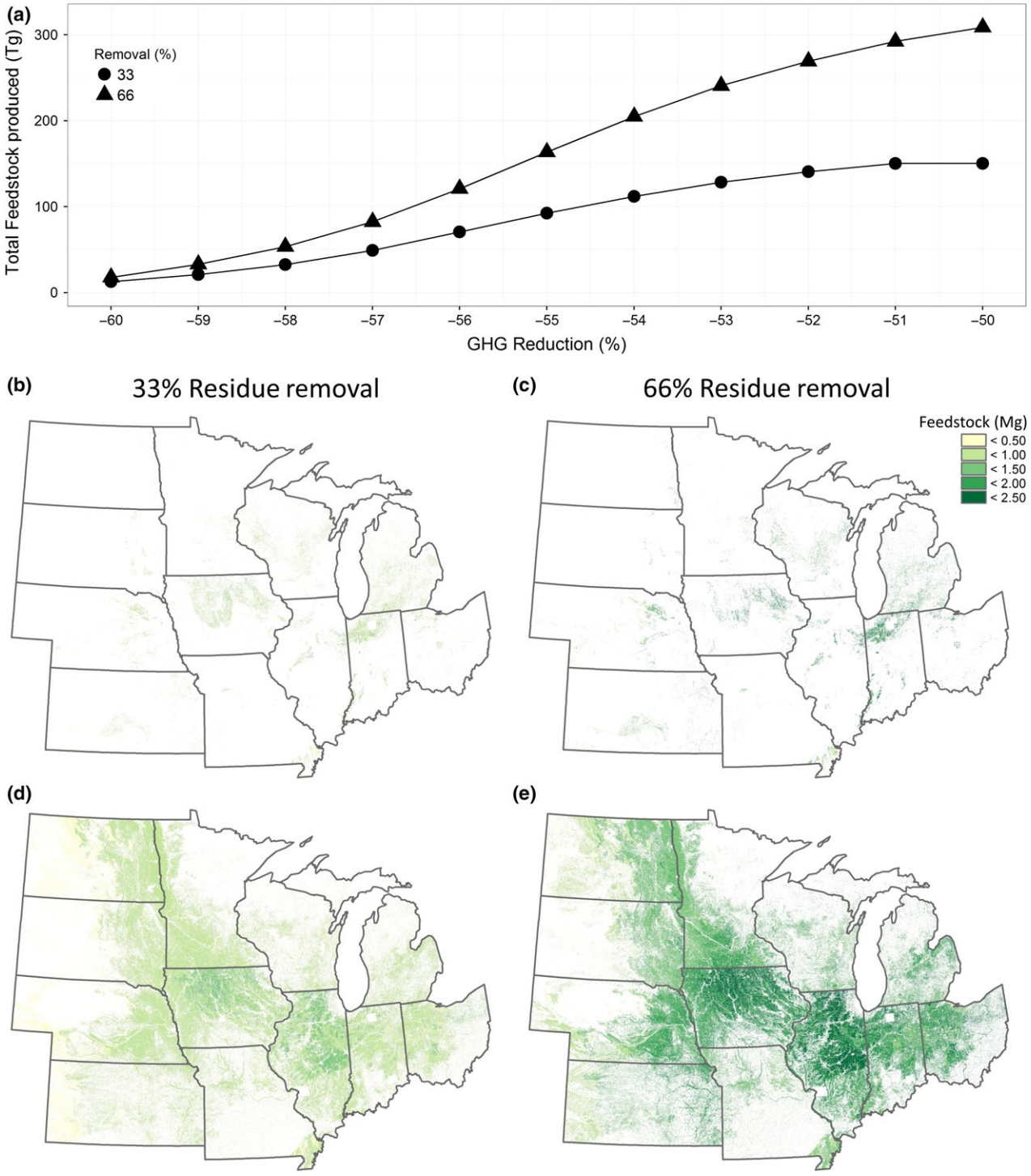



**Fig. 5** Available stover at 33% (a, b, d) and 66% (a, c, e) rate of residue removal for meeting 60% (b, c) and 50% (d, e) GHG reductions relative to 2005 petroleum.

2007). This is particularly impactful for shorter periods of study where the difference in litter C accounts for a greater proportion of the total C change, whereas over longer durations residue removal causes only small changes in litter C stocks (Gregg & Izaurralde, 2010). Removing litter from the calculations reduces the emissions to 65.2–67.0 g $CO_2$-eq $MJ^{-1}$, which is slightly less than the initial 10 year estimates for only the upper 30 cm of soil. Considering that feedstock production would likely persist for at least the 20–30+ year lifetime

of a biorefinery (Stephen *et al.*, 2010), it is reasonable to assess the impact of longer durations of residue removal on the GHG intensity of biofuel production. Considering soil humus C change in the whole soil profile, emission intensities for 30-year simulations were 53.5–54.4 g $CO_2$-eq $MJ^{-1}$ for 33–66% rates of removal, dropping intensities 17.9–18.8% from the comparable 10-year simulations. Further extending the simulation period to 50 years, which is the analysis combination presented here as the best estimate of long-term GHG impacts, the GHG emission intensity drops further to 46.4–47.0 g $CO_2$-eq $MJ^{-1}$, constituting a 13.3–13.6% reduction relative to the 30-year simulations and a 36.7–37.3% reduction compared to the 10-year estimates of Liska *et al.* (2014).

Overall, our analysis suggests that under common management practices only a modest amount of maize residue is available for cellulosic biofuel production in the US Midwest. However, with modest improvements in feedstock processing efficiency or crop management practices, an appreciable amount of feedstock could become available. To this end, increased use of cover crops represents a promising management practice as cover crops have been shown to increase SOC (Tonitto *et al.*, 2006) with measured increases of 0.10–1.0 Mg C $ha^{-1}$ $yr^{-1}$ under NT management relative to equivalent systems without cover crops under various climates and cropping systems (Blanco-Canqui, 2013). Mitigation practices of even modest effectiveness would greatly shift life cycle emissions of cellulosic feedstock production. For instance, an SOC increase of only 0.10 Mg C $ha^{-1}$ $yr^{-1}$ would shift the average aggregate emissions intensity to 33.8 Mg C $ha^{-1}$ $yr^{-1}$ under 66% residue removal. This represents a 64% GHG reduction relative to petroleum, allowing overall production in the US Midwest to meet the 60% GHG emission reduction standard for cellulosic biofuels. Residue harvesting from continuous maize systems under typical NT management will not produce appreciable amounts of feedstocks that meet the 60% cellulosic biofuel standard. Practical adjustments in production practices could considerably improve the GHG intensity of maize-derived biofuels from the US Midwest, dramatically increasing its potential for supplying cellulosic feedstock to meet near-term cellulosic biofuel production targets.

## Acknowledgements

We thank J.R. Williams (Texas A&M University) for EPIC model advice, R.E. Dunker and E.D. Nafziger (University of Illinois at Urbana-Champaign) for sharing data from the Morrow Plots, and K.A. Dolan (University of Maryland) for assistance producing maps and figures. Research was conducted with support from US DOE Office of Science (DE-FC02-07ER64494) to the Great Lakes Bioenergy Research Center. Additional support was provided from Texas AgriLife Research and Extension Center (Texas A&M University) to R.C.I. The high resolution EPIC simulations were conducted on the PNNL Institutional Computing (PIC) supercomputer (http://pic.pnnl.gov/). X. Zhang also received support from NASA Terrestrial Ecology Program (NNH12AU03I) for preparing geospatial input data to characterize agroecosystems in the Midwest U.S.

## References

Adler PR, Rau BM, Roth GW (2015) Sustainability of corn stover harvest strategies in Pennsylvania. *BioEnergy Research*, **8**, 1–11.

Angers DA, Eriksen-Hamel NS (2008) Full-inversion tillage and organic carbon distribution in soil profiles: a meta-analysis. *Soil Science Society of America Journal*, **72**, 1370–1374.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bhattacharyya T, Pal D, Ray S *et al.* (2013) Simulating change in soil organic carbon in two long term fertilizer experiments in India: with the RothC model. *Climate Change and Environmental Sustainability*, **1**, 104–117.

Blanco-Canqui H (2013) Crop residue removal for bioenergy reduces soil carbon pools: how can we offset carbon losses? *Bioenergy Research*, **6**, 358–371.

Campbell EE, Johnson JMF, Jin VL, Lehman RM, Osborne SL, Varvel GE, Paustian K (2014) Assessing the soil carbon, biomass production, and nitrous oxide emission impact of corn stover management for bioenergy feedstock production using DAYCENT. *BioEnergy Research*, **7**, 491–502.

Cerri CEP, Paustian K, Bernoux M, Victoria RL, Melillo JM, Cerri CC (2004) Modeling changes in soil organic matter in Amazon forest to pasture conversion with the Century model. *Global Change Biology*, **10**, 815–832.

Cheng K, Ogle SM, Parton WJ, Pan G (2014) Simulating greenhouse gas mitigation potentials for Chinese Croplands using the DAYCENT ecosystem model. *Global Change Biology*, **20**, 948–962.

De Vente J, Poesen J, Verstraeten G *et al.* (2013) Predicting soil erosion and sediment yield at regional scales: where do we stand? *Earth-Science Reviews*, **127**, 16–29.

Dwivedi P, Wang W, Hudiburg T *et al.* (2015) Cost of Abating Greenhouse Gas Emissions with Cellulosic Ethanol. *Environmental Science & Technology*, **49**, 2512–2522.

Egbendewe-Mondzozo A, Swinton SM, Izaurralde RC, Manowitz DH, Zhang X (2013) Maintaining environmental quality while expanding biomass production: sub-regional US policy simulations. *Energy Policy*, **57**, 518–531.

Farr TG, Rosen PA, Caro E *et al.* (2007) The shuttle radar topography mission. *Reviews of Geophysics*, **45**, RG2004.

Galdos M, Cerri C, Cerri C (2009) Soil carbon stocks under burned and unburned sugarcane in Brazil. *Geoderma*, **153**, 347–352.

Gelfand I, Sahajpal R, Zhang X, Izaurralde RC, Gross KL, Robertson GP (2013) Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*, **493**, 514–517.

Gollany H, Rickman R, Liang Y, Albrecht S, Machado S, Kang S (2011) Predicting agricultural management influence on long-term soil organic carbon dynamics: implications for biofuel production. *Agronomy Journal*, **103**, 234–246.

Gregg JS, Izaurralde RC (2010) Effect of crop residue harvest on long-term crop yield, soil erosion and nutrient balance: trade-offs for a sustainable bioenergy feedstock. *Biofuels*, **1**, 69–83.

Guo LB, Gifford RM (2002) Soil carbon stocks and land use change: a meta analysis. *Global Change Biology*, **8**, 345–360.

Hudiburg TW, Wang W, Khanna M *et al.* (2016) Impacts of a 32-billion-gallon bioenergy landscape on land and fossil fuel use in the US. *Nature Energy*, **1**, 15005.

Huggins D, Clapp C, Lamb J *et al.* (2007) Corn-soybean sequence and tillage effects on soil carbon dynamics and storage. *Soil Science Society of America Journal*, **71**, 145–154.

Izaurralde RC, Haugen-Kozyra KH, Jans DC, Mcgill WB, Grant RF, Hiley JC (2001) *Soil C Dynamics: Measurement, Simulation and Site-to-Region Scale-Up*. R. Lai *et al.* ; Lewis Publisher, Boca Raton, FL.

Izaurralde R, Williams JR, McGill WB, Rosenberg NJ, Jakas MQ (2006) Simulating soil C dynamics with EPIC: model description and testing against long-term data. *Ecological Modelling*, **192**, 362–384.

Izaurralde RC, Mcgill WB, Williams J (2012) *Development and Application of the EPIC Model for Carbon Cycle, Greenhouse-gas Mitigation, and Biofuel Studies*. Pacific Northwest National Laboratory (PNNL), Richland, WA.

Izaurralde RC, McGill WB, Williams JR *et al.* (2017) Simulating microbial denitrification with EPIC: model description and evaluation. *Ecological Modelling*, **359**, 349–362.

Johnson D, Mueller R (2010) The 2009 cropland data layer. *Photogrammetric Engineering & Remote Sensing*, **76**, 1201–1205.

Karlen D, Wollenhaupt NC, Erbach D, Berry E, Swan J, Eash N, Jordahl J (1994) Crop residue effects on soil quality following 10-years of no-till corn. *Soil and Tillage Research*, **31**, 149–167.

Lal R (1976) Soil erosion problems on an Alfisol in Western Nigeria and their control.

Lal R (2006) Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land Degradation & Development*, **17**, 197–209.

Lal R, Pimentel D (2007) Biofuels from crop residues. *Soil and Tillage Research*, **93**, 237–238.

Langholtz M, Stokes B, Eaton L (2016) *2016 Billion-Ton Report: Advancing Domestic Resources for a Thriving Bioeconomy, Volume 1: Economic Availability of Feedstocks*. U.S. Department of Energy, Oak Ridge National Laboratory, Oak Ridge, TN.

LeDuc SD, Zhang X, Clark CM, Izaurralde RC (2017) Cellulosic feedstock production on Conservation Reserve Program land: potential yields and environmental effects. *GCB Bioenergy*, **9**, 460–468.

Li ZT, Yang J, Drury C, Hoogenboom G (2015) Evaluation of the DSSAT-CSM for simulating yield and soil organic C and N of a long-term maize and wheat rotation experiment in the Loess Plateau of Northwestern China. *Agricultural Systems*, **135**, 90–104.

Liska AJ, Yang H, Milner M *et al*. (2014) Biofuels from crop residue can reduce soil carbon and increase $CO_2$ emissions. *Nature Climate Change*, **4**, 398–401.

Lorenz K, Lal R (2005) The depth distribution of soil organic carbon in relation to land use and management and the potential of carbon sequestration in subsoil horizons. *Advances in Agronomy*, **88**, 35–66.

Lu X, Cheng G, Xiao F, Huo C (2008) Simulating carbon sequestration and GHGs emissions in Abies fabric forest on the Gongga Mountains using a biogeochemical process model Forest-DNDC. *Journal of Mountain Science*, **5**, 249–256.

Luo Z, Wang E, Sun OJ (2010) Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. *Agriculture, Ecosystems & Environment*, **139**, 224–231.

Machado S (2011) Soil organic carbon dynamics in the Pendleton long-term experiments: implications for biofuel production in Pacific Northwest. *Agronomy Journal*, **103**, 253–260.

Miehle P, Livesley S, Feikema P, Li C, Arndt S (2006) Assessing productivity and carbon sequestration capacity of Eucalyptus globulus plantations using the process model Forest-DNDC: calibration and validation. *Ecological Modelling*, **192**, 83–94.

Muth DJ, Bryden KM, Nelson R (2013) Sustainable agricultural residue removal for bioenergy: a spatially comprehensive US national assessment. *Applied Energy*, **102**, 403–417.

Nadeu E, Gobin A, Fiener P, Wesemael B, Oost K (2015) Modelling the impact of agricultural management on soil carbon stocks at the regional scale: the role of lateral fluxes. *Global Change Biology*, **21**, 3181–3192.

Panagos P, Borrelli P, Meusburger K, van der Zanden EH, Poesen J, Alewell C (2015) Modelling the effect of support practices (P-factor) on the reduction of soil erosion by water at European Scale. *Environmental Science & Policy*, **51**, 23–34.

Parton WJ, Ojima DS, Cole CV, Schimel DS (1994) A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. In: *Quantitative Modeling of Soil Forming Processes* (eds Bryant RB, Arnold RW), pp. 147–167. Soil Science Society of America, Madison, WI.

Philibert A, Loyce C, Makowski D (2012) Assessment of the quality of meta-analysis in agronomy. *Agriculture, Ecosystems & Environment*, **148**, 72–82.

Powlson DS, Stirling CM, Jat ML, Gerard BG, Palm CA, Sanchez PA, Cassman KG (2014) Limited potential of no-till agriculture for climate change mitigation. *Nature Climate Change*, **4**, 678–683.

Quinton JN, Govers G, Van Oost K, Bardgett RD (2010) The impact of agricultural soil erosion on biogeochemical cycling. *Nature Geoscience*, **3**, 311–314.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robertson GP, Hamilton SK, Del Grosso SJ, Parton WJ (2011) The biogeochemistry of bioenergy landscapes: carbon, nitrogen, and water considerations. *Ecological Applications*, **21**, 1055–1067.

Robertson GP, Grace PR, Izaurralde RC, Parton WP, Zhang X (2014) $CO_2$ emissions from crop residue-derived biofuels. *Nature Climate Change*, **4**, 933–934.

Schmer M, Jin V, Wienhold B, Varvel G, Follett R (2014) Tillage and residue management effects on soil carbon and nitrogen under irrigated continuous corn. *Soil Science Society of America Journal*, **78**, 1987–1996.

Sheehan JJ, Adler PR, Del Grosso SJ, Easter M, Parton W, Paustian K, Williams S (2014) CO2 emissions from crop residue-derived biofuels. *Nature Climate Change*, **4**, 932–933.

Smith P, Davies CA, Ogle S *et al*. (2012) Towards an integrated global framework to assess the impacts of land use and management change on soil carbon: current capability and future vision. *Global Change Biology*, **18**, 2089–2101.

Solomon BD (2010) Biofuels and sustainability. *Annals of the New York Academy of Sciences*, **1185**, 119–134.

Spatari S, MacLean HL (2010) Characterizing model uncertainties in the life cycle of lignocellulose-based ethanol fuels. *Environmental Science & Technology*, **44**, 8773–8780.

Stephen J, Sokhansanj S, Bi X, Sowlati T, Kloeck T, Townley-Smith L, Stumborg M (2010) The impact of agricultural residue yield range on the delivered cost to a biorefinery in the Peace River region of Alberta, Canada. *Biosystems Engineering*, **105**, 298–305.

Suuster E, Ritz C, Roostalu H, Reintam E, Kõlli R, Astover A (2011) Soil bulk density pedotransfer functions of the humus horizon in arable soils. *Geoderma*, **163**, 74–82.

Tan Z, Liu S, Bliss N, Tieszen LL (2012) Current and potential sustainable corn stover feedstock for biofuel production in the United States. *Biomass and Bioenergy*, **47**, 372–386.

Thomson AM, Rosenberg NJ, Izaurralde RC, Brown RA (2005) Climate change impacts for the conterminous USA: an integrated assessment. Part 2: models and validation. *Climatic Change*, **69**, 27–41.

Tonitto C, David M, Drinkwater L (2006) Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: a meta-analysis of crop yield and N dynamics. *Agriculture, Ecosystems & Environment*, **112**, 58–72.

USDA-ERS (2013) Fertilizer use and price.

USDA-NASS (1997) *Usual planting and harvesting dates for U.S. field crops*.

Van Veen J, Paul E (1981) Organic carbon dynamics in grassland soils. 1. Background information and computer simulation. *Canadian Journal of Soil Science*, **61**, 185–201.

Wang X, Williams J, Gassman P, Baffaut C, Izaurralde R, Jeong J, Kiniry J (2012) EPIC and APEX: model use, calibration, and validation. *Transactions of the ASABE*, **55**, 1447–1462.

Wilhelm WW, Johnson JM, Karlen DL, Lightle DT (2007) Corn stover to sustain soil organic carbon further constrains biomass supply. *Agronomy Journal*, **99**, 1665–1667.

Williams JR (1995) The EPIC model. In: *Computer Models of Watershed Hydrology* (ed. Singh VP), pp. 909–1000. Water Resources Publications, Highlands Ranch, CO.

Zambrano-Bigiarini M, Rojas R (2013) A model-independent Particle Swarm Optimisation software for model calibration. *Environmental Modelling & Software*, **43**, 5–25.

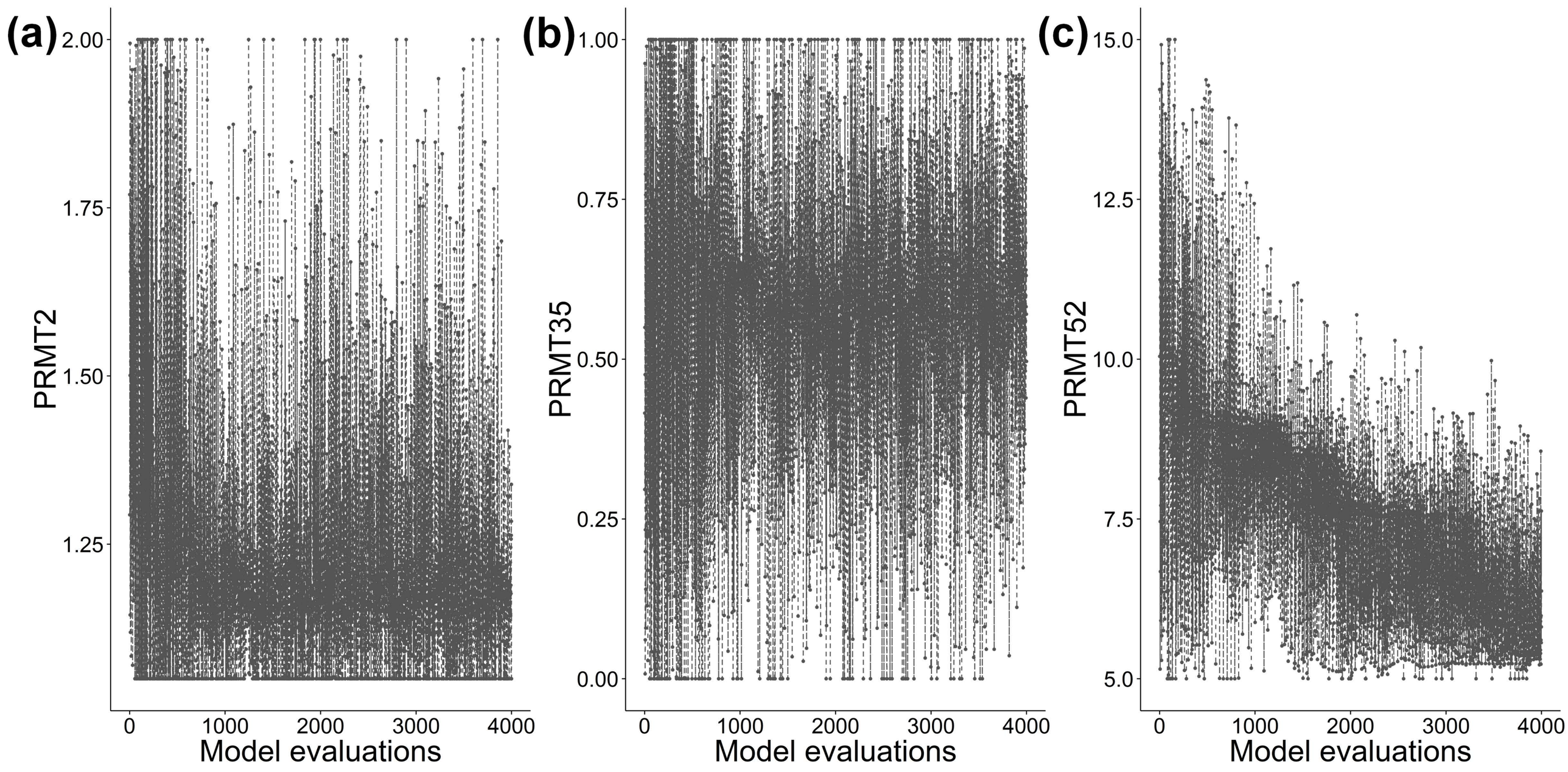
Zhang X, Izaurralde RC, Manowitz D *et al*. (2010) An integrative modeling framework to evaluate the productivity and sustainability of biofuel crop production systems. *GCB Bioenergy*, **2**, 258–277.

Zhang X, Beeson P, Link R *et al*. (2013) Efficient multi-objective calibration of a computationally intensive hydrologic model with parallel computing software in Python. *Environmental Modelling & Software*, **46**, 208–218.

Zhang X, Sahajpal R, Manowitz DH *et al*. (2014a) Multi-scale geospatial agroecosystem modeling: a case study on the influence of soil data resolution on carbon budget estimates. *Science of The Total Environment*, **479–480**, 138–150.

Zhang L, Yu D, Shi X, Xu S, Xing S, Zhao Y (2014b) Effects of soil data and simulation unit resolution on quantifying changes of soil organic carbon at regional scale with a biogeochemical process model. *PLoS ONE*, **9**, e88622.

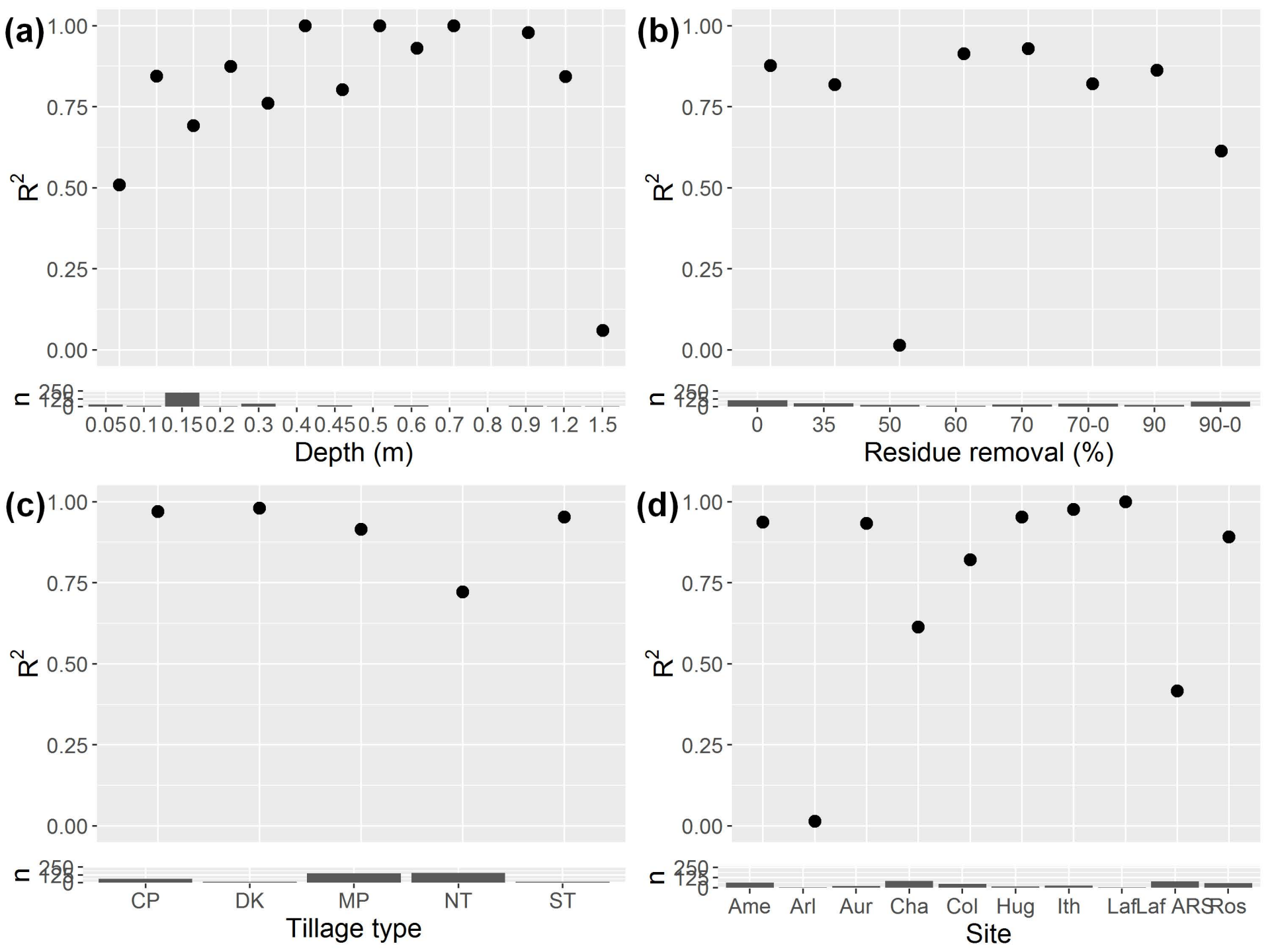
Zhang X, Izaurralde RC, Manowitz DH *et al*. (2015) Regional scale cropland carbon budgets: evaluating a geospatial agricultural modeling system using inventory data. *Environmental Modelling & Software*, **63**, 199–216.


(a)
PRMT2
2.00
1.75
1.50
1.25
0
1000
2000
3000
4000
Model evaluations
(b)
PRMT35
1.00
0.75
0.50
0.25
0.00
0
1000
2000
3000
4000
Model evaluations
(c)
PRMT52
15.0
12.5
10.0
7.5
5.0
0
1000
2000
3000
4000
Model evaluations

## Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Figure S1.** Evolution of PRMT2 (a), PRMT35 (b), and PRMT52 (c) parameter values during the calibration process. With greater iterations, parameters tend towards more optimal values.
**Figure S2**. Goodness-of-fit ($R^2$) and associated sample size (*n*) of simulated vs modelled SOC by depth (a), rate of residue removal (b), type of tillage (c) and experimental site (d).
**Note S1**. EPIC Simulations with a Spatially Explicit Integrated Modeling Framework (SEIMF).
**Table S1**. Description of residue removal experiments.
**Table S2**. Significance of factors in linear effects model for modeling measured rate of SOC change.


(a)
1.00
0.75
0.50
0.25
0.00
R2
n
250
125
0
0.05 0.1 0.15 0.2 0.3 0.4 0.45 0.5 0.6 0.7 0.8 0.9 1.2 1.5
Depth (m)
(b)
0 35 50 60 70 70-0 90 90-0
Residue removal (%)
(c)
CP DK MP NT ST
Tillage type
(d)
Ame Arl Aur Cha Col Hug Ith Laf Laf ARS Ros
Site

**Supporting information**

**Note S1**. EPIC Simulations with a Spatially Explicit Integrated Modeling Framework (SEIMF).

*Cropland*

Following Zhang *et al.* (2015) we overlaid four CDL maps from 2007-2010 to define the area of cropland that is cultivated under field crops in the U.S. Midwest. The total area of field crops is approximately 68 million ha.

*Soils*

The SSURGO data were used to derive a soil type distribution map and associated soil parameters for each map unit. We downloaded SSURGO vector maps at a scale of approximately 1:24,000 for each county in the US Midwest from the USDA Geospatial Data Gateway (http://datagateway.nrcs.usda.gov) and converted them into raster format with a resolution of 56 m. Soil properties processed for EPIC include the number of soil layers, layer depth, slope gradient, slope length, albedo, bulk density, pH, sand fraction, silt fraction, clay fraction, coarse fragments, organic C fraction, and total N.

**Topography.** To derive topographical information, we used data from the SRTM (Farr *et al.*, 2007), which produced a digital elevation model (DEM) for the region at a resolution of 30 m. Elevation was used by EPIC for atmospheric pressure, which is an input for evapotranspiration calculations.

**Climate data.** EPIC requires daily weather information, including maximum and minimum temperature, precipitation, solar radiation, wind speed, and relative humidity. The NLDAS2 database was used to derive daily weather files at approximately 12-km resolution.

**Defining spatially explicit modeling units**. The cropland map, SSURGO soil map, and US Midwest county map were overlaid to define HSMUs with unique combinations of these layers

(Figure 1). The HSMU map is of a spatial resolution of 56 m for the simulation domain. For each HSMU we derived average elevation based on the SRTM topography product. We prepared daily weather data for each HSMU by locating the closest NLDAS2 grid.

**Table S1**. Description of residue removal experiments.

| Site | Years[1]/ Treatments/ n[b] | Removal (%) | Tillage[c] | Fertilizer (kg N ha$^{-1}$) | Irrigation[d] | Depth[e] (m) | Reference(s) |
|---|---|---|---|---|---|---|---|
| Ames | 5.2/12/384 | 0, 35, 70 | CP, NT | 250, 225 | R | 0.15 | (Del Grosso *et al.*, 2013; USDA ARS, 2015) |
| Arlington | 5/2/45 | 0, 60 | NT | 160 | R | 0.25 | (GLBRC, 2015) |
| Aurora | 4.5/8/48 | 0, 60 | CP, NT | 168 | R, I | 0.6 | (Clay *et al.*, 2015) |
| Champaign | 104/3/199 | Yes-No | MP | 0, 77, 97 | R | 0.15 | (Nafziger & Dunker, 2011) |
| Columbia | 73/2/60 | Yes-No | MP | 0, 95 | I | 1.2 | (Miles & Brown, 2011) |
| Hugoton | 2/2/144 | 0, 70 | ST | 176 | I | 0.3 | (Ihde, 2011) |
| Ithaca | 9.5/6/432 | 0, 35, 70 | DK, NT | 202 | I | 1.5 | (Del Grosso *et al.*, 2013; Schmer *et al.*, 2014; USDA ARS, 2015) |
| Lafayette | 4.4/2/4 | 0, 60 | CP | 280 | R | 0.6 | (Barber, 1979) |
| Lafayette | 5.7/4/325 | 0, 35, 50, 70 | NT | 246 | R | 0.15 | (Del Grosso *et al.*, 2013; USDA ARS, 2015) |
| Rosemount* | 17/12/96 | 0, 90 | CP, MP, NT | 0, 200 | R | 0.45 | (Clapp *et al.*, 2000; Linden *et al.*, 2000) |

*Experiment was continuous corn from 1980-1992, continuous soybean from 1993-1998, and continuous corn after 1998

[1] Indicates the time between the first and last SOC measurements.

[b] Number of SOC measurements.

[c] CP – chisel plow, NT – no-till, MP – moldboard plow, ST – strip-till, DK – disk

[d] R – rainfed, I – irrigation.

[e] Maximum depth sampled.

**Table S2**. Significance of factors in linear effects model for modeling measured rate of SOC change.

| Factor | Sum of Squares | Mean Squares | Numerator df | Denominator df | F-value | P-value | Estimate | Standard Error |
|---|---|---|---|---|---|---|---|---|
| Tillage | 4.13E+06 | 1.03E+06 | 4 | 29.38 | 0.58 | 0.682 | | |
| N.rate | 2.91E+05 | 2.91E+05 | 1 | 285.23 | 0.16 | 0.6862 | | |
| Irrigation | 2.52E+06 | 2.52E+06 | 1 | 141.07 | 1.42 | 0.2352 | | |
| Residue | 2.59E+07 | 2.59E+07 | 1 | 355.42 | 14.66 | 0.0002 | -8.84 | 2.31 |
| Depth* | 1.91E+08 | 1.91E+08 | 1 | 357.78 | 107.82 | <1e-07 | -2065.25 | 198.89 |

* Average soil layer depth.

**Supporting references**


Barber SA (1979) Corn residue management and soil organic matter. *Agronomy Journal*, **71**, 625–627.

Clapp CE, Allmaras RR, Layese MF, Linden DR, Dowdy RH (2000) Soil organic carbon and 13 C abundance as related to tillage, crop residue, and nitrogen fertilization under continuous corn management in Minnesota. *Soil and Tillage Research*, **55**, 127–142.

Clay DE, Reicks G, Carlson CG, Moriles-Miller J, Stone JJ, Clay SA (2015) Tillage and Corn Residue Harvesting Impact Surface and Subsurface Carbon Sequestration. *Journal of Environmental Quality*, **44**, 803–809.

Del Grosso S, White J, Wilson G et al. (2013) Introducing the GRACEnet/REAP data contribution, discovery, and retrieval system. *Journal of environmental quality*, **42**, 1274–1280.

Farr TG, Rosen PA, Caro E et al. (2007) The shuttle radar topography mission. *Reviews of geophysics*, **45**.

GLBRC (2015) Great Lakes Bioenergy Research Center Sustainability Data Catalog.

Ihde NA (2011b) *Implications of residue removal on soil quality in southwest Kansas*. Kansas State University.

Linden DR, Clapp CE, Dowdy RH (2000) Long-term corn grain and stover yields as a function of tillage and residue removal in east central Minnesota. *Soil and Tillage Research*, **56**, 167–174.

Miles RJ, Brown JR (2011) The Sanborn Field experiment: Implications for long-term soil organic carbon levels. *Agronomy Journal*, **103**, 268–278.

Nafziger ED, Dunker RE (2011) Soil organic carbon trends over 100 years in the Morrow plots. *Agronomy journal*, **103**, 261–267.

Schmer M, Jin V, Wienhold B, Varvel G, Follett R (2014) Tillage and residue management effects on soil carbon and nitrogen under irrigated continuous corn. *Soil Science Society of America Journal*, **78**, 1987–1996.

USDA ARS (2015) US Department of Agriculture, Agricultural Research Service Data Portal.

Zhang X, Izaurralde RC, Manowitz DH et al. (2015) Regional scale cropland carbon budgets: evaluating a geospatial agricultural modeling system using inventory data. *Environmental Modelling & Software*, **63**, 199–216.

## Figure captions

**Figure S1.** Evolution of PRMT2 (a), PRMT35 (b), and PRMT52 (c) parameter values during the calibration process. With greater iterations, parameters tend towards more optimal values.

**Figure S2**. Goodness-of-fit ($R^2$) and associated sample size (n) of simulated vs modelled SOC by depth (a), rate of residue removal (b), type of tillage (c) and experimental site (d).