Science of the Total Environment 479–480 (2014) 138–150

Contents lists available at ScienceDirect

Science of the Total Environment

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

Multi-scale geospatial agroecosystem modeling: A case study on the influence of soil data resolution on carbon budget estimates

Xuesong Zhang [a,*], Ritvik Sahajpal [b], David H. Manowitz [a], Kaiguang Zhao [c], Stephen D. LeDuc [d], Min Xu [e], Wei Xiong [f], Aiping Zhang [f], Roberto C. Izaurralde [a,b], Allison M. Thomson [a], Tristram O. West [a], Wilfred M. Post [g]

[a] Joint Global Change Research Institute, Pacific Northwest National Laboratory and University of Maryland, College Park, MD 20740, USA
[b] Department of Geographical Sciences, University of Maryland, College Park, MD 20740, USA
[c] School of Environment and Natural Resources, The Ohio Agricultural Research and Development Center, Ohio State University, Wooster, OH 44691, USA
[d] U.S. Environmental Protection Agency, National Center for Environmental Assessment, Arlington, VA 22202, USA
[e] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA
[f] Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[g] Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

HIGHLIGHTS

- Spatially-explicit modeling of agroecosystems at a spatial resolution of 56 m for Iowa, USA.
- SSURGO- and STATSGO-derived cropland carbon fluxes substantially differ from each other at the county and grid scales.
- Pronounced difference in amount and distribution between marginal lands identified with different soil data

ARTICLE INFO

Article history:
Received 1 November 2013
Received in revised form 24 January 2014
Accepted 25 January 2014
Available online 19 February 2014

Keywords:
Climate change
Net Ecosystem Production
EPIC
Parallel computing
Spatial resolution
SSURGO
STATSGO

ABSTRACT

The development of effective measures to stabilize atmospheric CO_2 concentration and mitigate negative impacts of climate change requires accurate quantification of the spatial variation and magnitude of the terrestrial carbon (C) flux. However, the spatial pattern and strength of terrestrial C sinks and sources remain uncertain. In this study, we designed a spatially-explicit agroecosystem modeling system by integrating the Environmental Policy Integrated Climate (EPIC) model with multiple sources of geospatial and surveyed datasets (including crop type map, elevation, climate forcing, fertilizer application, tillage type and distribution, and crop planting and harvesting date), and applied it to examine the sensitivity of cropland C flux simulations to two widely used soil databases (i.e. State Soil Geographic—STATSGO of a scale of 1:250,000 and Soil Survey Geographic—SSURGO of a scale of 1:24,000) in Iowa, USA. To efficiently execute numerous EPIC runs resulting from the use of high resolution spatial data (56 m), we developed a parallelized version of EPIC. Both STATSGO and SSURGO led to similar simulations of crop yields and Net Ecosystem Production (NEP) estimates at the State level. However, substantial differences were observed at the county and sub-county (grid) levels. In general, the fine resolution SSURGO data outperformed the coarse resolution STATSGO data for county-scale crop-yield simulation, and within STATSGO, the area-weighted approach provided more accurate results. Further analysis showed that spatial distribution and magnitude of simulated NEP were more sensitive to the resolution difference between SSURGO and STATSGO at the county or grid scale. For over 60% of the cropland areas in Iowa, the deviations between STATSGO- and SSURGO-derived NEP were larger than 1 Mg C ha^{-1} yr^{-1}, or about half of the average cropland NEP, highlighting the significant uncertainty in spatial distribution and magnitude of simulated C fluxes resulting from differences in soil data resolution.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Stabilizing concentrations of carbon dioxide (CO_2) and other greenhouse gases (GHGs) in the atmosphere is critical for mitigating potential negative impacts of future climate change (IPCC, 2007). Agricultural lands not only provide essential ecosystem goods (such as food, livestock, and fiber) for humankind but also have a significant potential to sequester carbon (C) and attenuate the increase rate of atmospheric CO_2 (Lal and Bruce, 1999; Allmaras et al., 2000; West and Post, 2002; Smith et al., 2007). The biological dynamics of managed landscapes affect the fluctuations of atmospheric CO_2 levels (Moureaux et al., 2008) and need to be considered when understanding, quantifying, and

* Corresponding author at: 5825 University Research Court, Suite 1200, Joint Global Change Research Institute, Pacific Northwest National Laboratory and University of Maryland, College Park, MD 20740, USA.
E-mail address: xuesong.zhang@pnnl.gov (X. Zhang).

0048-9697/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.scitotenv.2014.01.099

regulating the global C cycle (Sus et al., 2010; West et al., 2013). The magnitude, spatial and temporal patterns of terrestrial C sinks and sources remain uncertain and require further examination (Denman et al., 2007). Refining the scale of sources and sinks is important for credibly determining their variation and strength in different regions or under varying land management systems (Dolman et al., 2009).

Due to the lack of extensive monitoring networks with sufficient numbers of soil organic carbon (SOC) sampling sites and flux towers, simulation models are regularly used by researchers and decision makers to estimate terrestrial C flux (Saby et al., 2008; Ogle et al., 2010; West et al., 2010). As one of the most important drivers for terrestrial ecosystem models, soil properties (e.g. SOC, albedo, layer depth, texture, and bulk density) can vary substantially across spatial scales among different soil databases. In the U.S., the two most widely used soil databases are the State Soil Geographic (STATSGO, soils.usda.gov/survey/geography/statsgo) and Soil Survey Geographic (SSURGO, soils.usda.gov/survey/geography/ssurgo) which are collected, stored, and maintained by the U.S. Department of Agriculture's Natural Resources Conservation Service (NRCS) (Soil Survey Staff, 2012). STATSGO maps are compiled by generalizing more detailed (SSURGO) soil survey maps (USDA-NRCS National Soil Survey Center, 1995). SSURGO allows much finer scale soil unit identification (1:24,000) than STATSGO (1:250,000). Although both databases were compiled by USDA-NRCS, pronounced differences have been reported when using SSURGO and STATSGO to derive SOC stocks at local and regional scales. Davidson and Lefebvre (1993) compared SOC stock estimates derived from SSURGO and STATSGO in the Old Orchard Beach Quadrangle in York and Cumberland Counties, Maine, and found that SSURGO estimated 13% less SOC than STATSGO. Buell and Markewich (2004) demonstrated that spatial patterns of SOC derived from SSURGO and STATSGO can vary substantially within a county in North Carolina. Wu et al. (2001) and Rasmussen (2006) also found noticeable deviations between SSURGO and STATSGO estimates of SOC at a county scale. Recently, Zhong and Xu (2011) showed that soil organic matter (SOM) from SSURGO has a high correlation ($R^2 = 0.625$; Legates and McCabe, 1999) with field observed SOM across 30 sites in Louisiana, while SOM derived from STATSAGO did not ($R^2 = 0.013$). They also showed that total SOC stocks in Louisiana were 1.9 vs. 1.4 Pg C derived from SSURGO and STATSGO, respectively. Mednick (2010) reviewed 18 publications on comparing the use of STATSGO and SSURGO data in water quality modeling, and concluded that higher resolution soil data more often provide greater accuracy for modeling hydrologic and water quality parameters. Thus, the difference between STATSGO and SSURGO has significant implications for hydrologic, nutrients, and water quality modeling, and may represent an important source of uncertainty for estimating terrestrial C budgets.

Although STATSGO and SSURGO both have been used in estimating and simulating terrestrial C budgets (Paustian et al., 2002; West et al., 2010; West et al., 2013; Causarano et al., 2008; Zhang et al., 2010; Zhang et al., 2010), they involve more or less aggregation of STATSGO or SSURGO from their original resolution to a coarser level. For example, Paustian et al. (2002) reported results at the state level, West et al. (2013) at the county scale, and Causarano et al. (2008) at a 1600 m resolution. However, the difference in C estimates arising from using STATSGO or SSURGO has not been examined in a spatially-explicit manner. As soil is the medium for plant growth, the pronounced variation of soil properties from changes in soil resolution is expected to lead to considerable deviations in simulations of the amount of C photosynthesized into vegetation and the dead plant material input into soil, thus introducing significant uncertainties into terrestrial C flux estimates. Therefore, the major objective of this research was to examine the effects of using STATSGO (both the dominant soil type and area-weighted approach) and SSURGO, and their accompanying differences in soil data resolution, on simulating C fluxes on croplands.

To examine this, we used the agroecosystem model Environmental Policy Integrated Climate (EPIC) (Williams, 1995; Izaurralde et al., 2006), which has been extensively tested for simulating SOC dynamics and land–atmosphere C flux on cropland (e.g. Izaurralde et al., 2006; He et al., 2006; Wang et al., 2005; Causarano et al., 2007, 2008; Apezteguía et al., 2009; Schwalm et al., 2010). In addition, EPIC is among the models used to assess conservation effects of the Conservation Reserve Program (USDA-FSA, 2008; CRP, USDA-FSA, 2010). We chose Iowa as the study area (Fig. 1), as it is one of most intensively cultivated states in the U.S. and is the leading hotspot experiencing agricultural expansion in response to human needs of food and energy (Secchi et al., 2011). Soils in Iowa are dominated by fertile Mollisols and Afisols (over 90% of the lands) that contain high soil organic matter (on average > 200 Mg SOC ha^{-1}) and are mainly comprised of fine particles classified as fine, fine-loamy, and fine-silty.

In this study, built on a spatially-explicit integrative modeling framework (SEIMF, Zhang et al., 2010; Nichols et al., 2011), we developed a parallel computing modeling system to efficiently run EPIC at a high spatial resolution across large scales, thus allowing us to examine the implications of soil map resolution on regional scale C balance analyses. The modeling system contains the following major functions: (1) it combines multiple spatial layers (e.g. crop data layer (CDL, Johnson and Mueller, 2010), county boundary, and soil maps from SSURGO and STATSGO) to derive spatially-explicit modeling units; (2) georeferences climate and crop management data to each unit and prepare location specific input files for EPIC; (3) performs parallel implementation of EPIC runs on all modeling units by simultaneously using hundreds of processors; (4) automatically processes and compiles the outputs of each EPIC run into an relational database; and (5) links the EPIC derived ecosystem variables to spatially-explicit units for visualization. We ran EPIC on each unit for a period from 1991 to 2008, with years from 1991 to 1999 to initialize conditions, to derive simulation results on crop yields and carbon results for analysis. We used both statistical methods and spatial maps to compare and illustrate the difference between simulations with SSURGO and STATSGO. As biofuels are promising for reducing C emissions and marginal lands could offer potential opportunities for large scale deployment of biofuel feedstocks (Tilman et al., 2006; Gelfand et al., 2013), we also employed a marginal land identification example to illustrate differences between SSURGO and STATSGO for analyses that require detailed spatial information. We discuss the implications of these results for understanding the sensitivity of model simulated C flux to changes in soil data resolution.

2. Materials and methods

2.1. The EPIC model

EPIC is a comprehensive terrestrial ecosystem model capable of simulating key biophysical and biogeochemical processes, such as plant growth and development, water balance, C and nutrient cycling, soil erosion, and greenhouse-gas emissions; and how these processes are influenced by climate conditions, landscape configurations, soil properties, and management practices (Williams, 1995; Surendran Nair et al., 2012). The plant growth sub-model of EPIC is a revised version of Crop Environment REsource Synthesis (CERES) (Williams et al., 1989; Jones et al., 1991), which employs the concept of radiation-use efficiency by which a fraction of daily photosynthetically-active solar radiation is intercepted by the plant canopy and converted into plant biomass. Daily gains in plant biomass are affected by vapor pressure deficits, atmospheric CO_2 concentrations, and other environmental controls and stresses. Currently, EPIC is parameterized for approximately 120 plant species including food crops, native grasses, and trees. EPIC represents all salient terrestrial water cycling processes including snowmelt, surface runoff, infiltration, soil water content, percolation, lateral flow, water table dynamics, and evapotranspiration (Williams, 1995). A modified version of the CENTURY C model (Parton et al., 1994) is used to simulate soil C and nitrogen (N) decomposition and transformation (Izaurralde et al., 2006). Soil C and N dynamics are influenced by

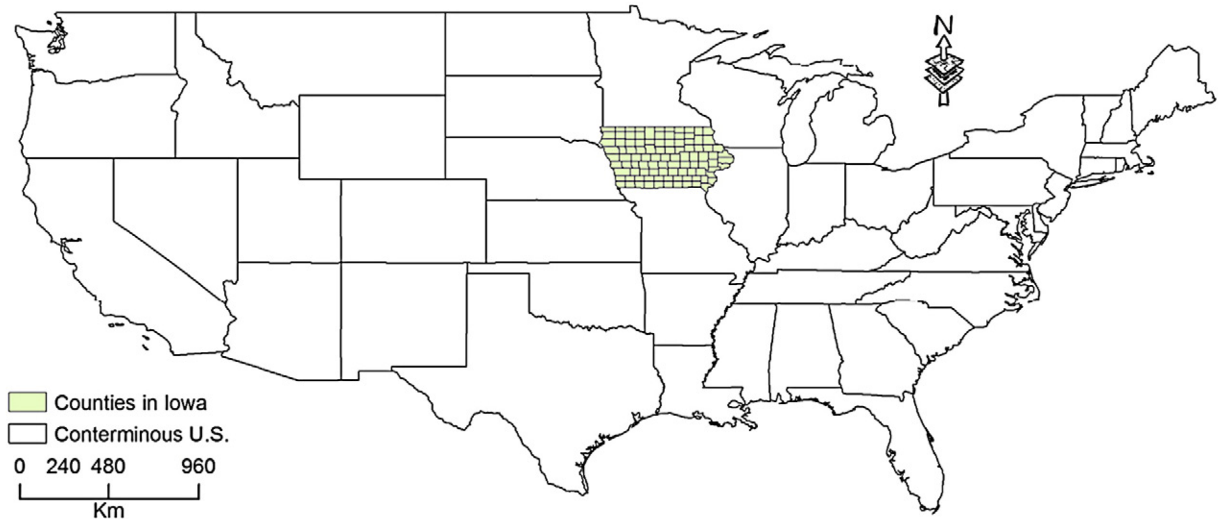

Fig. 1. Location of the study region.

many factors and processes, such as soil texture, pH, crop yields, atmospheric N input, fertilizer and manure, and tillage, among others. The EPIC model has been extensively tested for crop yield across numerous sites around the world (Gassman et al., 2007). Its capability to predict SOC dynamics under different crop rotation and management practices has also been evaluated against numerous field observations (e.g. Izaurralde et al., 2006, 2007; He et al., 2006; Causarano et al., 2008). Recent studies (Schwalm et al., 2010; Zhang et al., 2013b) showed that the C algorithm in EPIC simulated well Net Ecosystem Production (NEP) of diverse agroecosystems, where NEP was calculated as the net C sequestration from the atmosphere into plant biomass (or Net Primary Productivity, NPP) minus the C emission through soil heterotrophic respiration. When applying EPIC for simulating cropland C flux in Iowa, we borrowed the SOC cycling relevant parameters that have been tested in the aforementioned studies and further constrained and verified the model behavior with respect to NPP using county scale crop yield data from USDA-NASS' Quick Stats (quickstats.nass.usda.gov).

2.2. Geospatial database

For spatially-explicit assessment of cropland C flux across Iowa, we processed and compiled a series of geospatial databases that contain climate, land use/land cover, soil, and topography data to define spatial modeling units and provide relevant parameters to drive the EPIC model.

2.2.1. Soils

The SSURGO and STATSGO data were respectively used to define spatially-explicit modeling units and derive soil type distribution map and associated soil parameters for each map unit. For SSURGO, we downloaded the vector SSURGO maps for each county in Iowa from the US Department of Agriculture (USDA) Geospatial Data Gateway (datagateway.nrcs.usda.gov) and merged and converted them into a raster format with a resolution of 56 m that is consistent with the land use resolution. The STATSGO data compiled for the national scale Hydrologic Unit Model of the United States (HUMUS) in support of USDA analyses required for the 1997 Resources Conservation Act (Srinivasan et al., 1998, 2010; Arnold et al., 1999) were used here and also converted to a 56 m resolution. Soil properties processed for EPIC included the number of soil layers; layer depth; slope gradient and length; albedo, bulk density; pH; percent sand, silt, clay and coarse fragments; and percent organic C and total N.

2.2.2. Topography

To derive topographical information, we used data from the Shuttle Radar Topography Mission (SRTM), which produced a digital elevation model (DEM) for the region at a resolution of 30 m (Farr et al., 2007). Elevation was used by EPIC for atmospheric pressure calculation.

2.2.3. Climate data

EPIC requires daily weather information including daily temperature (maximum and minimum), precipitation, solar radiation, wind speed, and relative humidity. The hourly North-American Land Data Assimilation System 2 (NLDAS2, ldas.gsfc.nasa.gov/nldas/) climate forcing data were employed to derive the daily weather input at an 8-km resolution for EPIC.

2.2.4. Land use and land cover

The US National Agricultural Statistical Service' Cropland Data Layer (CDL, Johnson and Mueller, 2010) was used to develop a spatially-explicit map of crop rotations at a resolution of 56 m. We developed an algorithm to estimate crop rotation by combining and analyzing the CDL data for years 2007 to 2010. Unique combinations of annual crop rotations were aggregated into three dominant rotation classes (i.e. continuous corn, corn–soybean, and soybean–corn) that account for over 90% of the spatial and temporal crop patterns in Iowa. This simplification helped to substantially reduce the burden of preparing management files for EPIC for numerous modeling units, without leading to significant loss of land use information.

2.3. Spatially-explicit representation of cropping systems

The spatially-explicit integrative modeling framework (SEIMF, Zhang et al., 2010; Fig. 2) was employed to facilitate the processing of large quantities of spatial input data, define homogenous spatial modeling units (HSMUs) with the high-resolution data aforementioned, geo-referencing climate and management data to each unit, and preparing input files for executing EPIC. The CDL derived crop rotation map, SSURGO and STATSGO soil maps, 10-digit hydrologic unit, and county and state boundary maps were overlaid to define HSMUs with unique combinations of the above five spatial layers. Similar to CDL, the HSMU map was gridded to a resolution of 56 m consistency with the CDL maps. For each HSMU, we derived the average elevation based on the SRTM. Each HSMU possesses a unique series of properties such as soils, crop rotation, terrain, and county. We

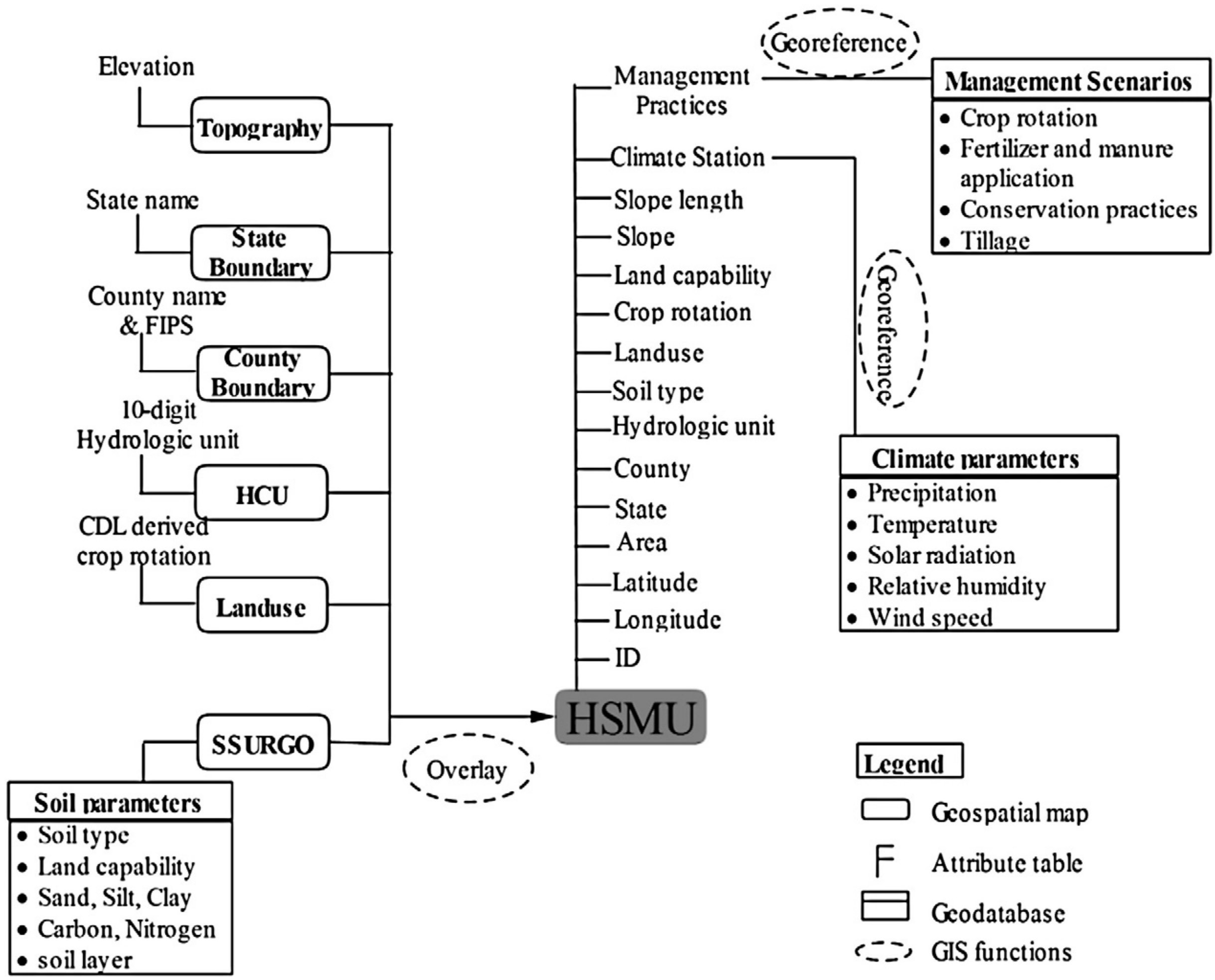

Fig. 2. Geospatial definition of HSMU and associated major attributes.
Adapted from Zhang et al., 2010.

prepared climate parameters for each HSMU by locating the closest NLDAS grid.

As there were no readily available spatially-explicit fertilizer, tillage, crop planting and harvesting date data, we further developed several procedures to assign these key crop management practices to each HSMU based on surveyed data at the county or state level. State level annual N and phosphorus application rates over 1990 to 2008 were obtained from USDA (USDA-ERS, 2013). Data gaps were filled with the value from the closest year. County scale data on fractions of tillage practices were collected and compiled by the Conservation Technology Information Center (CTIC, 2007), which were re-processed into three categories: conventional tillage, conservation tillage, and no-till and gap-filled by West et al. (2010) for 2000–2008. In order to cover the entire simulation period (1990–2008), we assumed that the spatial pattern of tillage practices for the initialization period (1990–1999) was similar to that averaged over 2000–2004. We also assumed that farmers apply no-till and conservation tillage to steep soils in order to preserve soil productivity and protect the environment. Therefore, we first assigned no-till to the steepest modeling units, followed by conservation tillage to modeling units with lower gradient slopes, and finally conventional tillage to those remaining tabular soils. Planting and harvesting dates and heat units required by corn (*Zea mays* L.) and soybean (*Glycine max* [L.] Merr.) to reach maturity are important for accurate crop growth and development prediction. We derived these data by using the potential heat unit program available at swat.tamu.edu/software/potential-heat-unit-program/ and typical planting and harvesting dates of major crops in the U.S. provided by USDA-NASS (1997). After processing the geospatial maps and deriving climate and management data, each HSMU possesses multi-dimensional information as illustrated in Fig. 2. The spatially-explicit scheme employed here allows us to preserve the spatial details of land use and soil patterns, while making it flexible for georeferencing climate data and crop management practices.

2.4. Deriving soil properties for spatially explicit modeling units

By alternatively using SSURGO and STATSGO maps, we derived two sets of modeling units for EPIC runs that allow us to examine the sensitivity of EPIC simulations to these two soil databases. As illustrated in Fig. 3, each map unit of STATSGO and SSURGO contains a series of soil components. However, there is no information available about the exact location of each component, making it ambiguous in selecting one or multiple soil components to represent soil properties of a HSMU overlapping with a soil map unit. The average number of soil components within one map unit of STATSGO is about ~18. This compares to ~2–3 for SSURGO. On average, the dominant soil component within each soil map unit is accounting for about 20% and 90% of the total area of the corresponding soil map unit for STATSGO and SSURGO, respectively. Much information is omitted if the dominant soil type (the one with the largest area coverage) is used to represent the soil properties of a STATSGO map unit. In contrast, the dominant soil type in a SSURGO map unit captures the majority of the soil

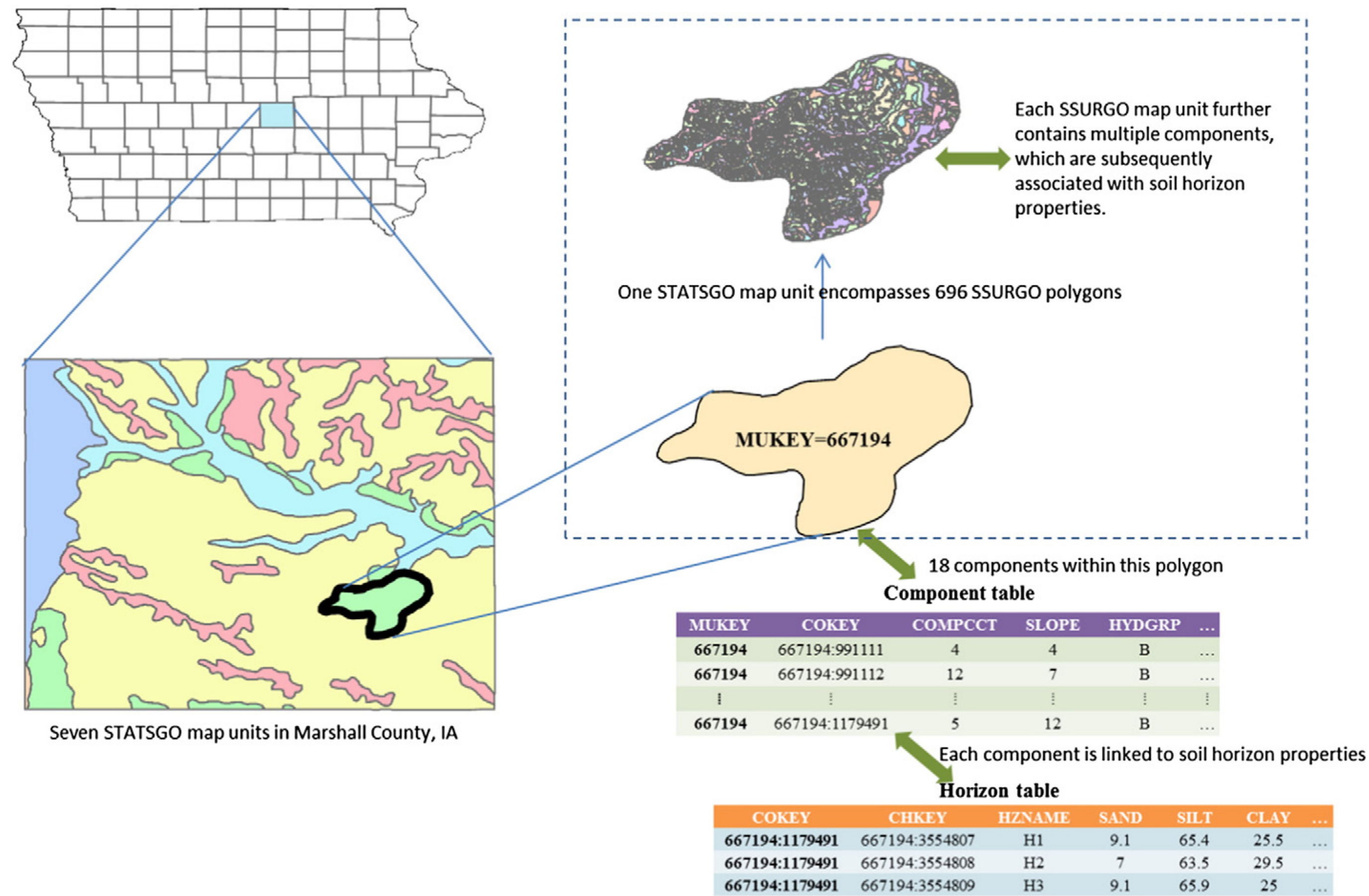

Fig. 3. Schematic illustration of the association between spatial soil map unit, soil components, and soil horizon properties and the difference between STATSGO and SSURGO in spatial resolution. (MUKEY is the id of each unique soil map unit; COKEY identifies the soil components in each map unit; COMPCCT is the area percentage; HYDGRP represents soil hydrologic group; Each COKEY is linked to multiple CHKEYs that contain soil horizon properties.).

properties of that map unit. Therefore, for the application of STATSGO, one question that immediately emerges concerns what are the implications of using the dominant soil type vs. area-weighting simulations using all soil components within a STATSGO map unit for agroecosystem modeling?

To address this question, we used two schemes to process the simulations based on STATSGO: (1) STATSGO-dominant—a HSMU is assumed to overlap with only the dominant soil type of that HSMU' soil map unit; and (2) STATSGO-weighted—each HSMU is assumed to overlap with all soil components within that HSMU's soil map unit and is further divided into n subunits with areas proportional to the areas of the n soil types within the corresponding soil map unit. SOC and simulated crop yields and NEP across the n subunits are area-weighted to derive aggregated estimates for a HSMU. The STATSGO-dominant scheme is the default in applying the Soil and Water Assessment Tool (SWAT) model that has been widely used in assessing effects of agricultural management practices on crop productivity, soil quality, and water quality in the U.S. (swat.tamu.edu), while the STATSGO-weighted scheme has rarely been evaluated.

For SSURGO, we only assessed the SSURGO-dominant scheme because (1) the dominant soil components covered over 90% of the soil map units, and (2) missing values in soil properties in the SSURGO database make it difficult to construct a consistent framework for performing an area-weighted analysis. In some SSURGO soil map units, the first dominant soil type had missing values, and we therefore selected the second or even the third dominant soil type to represent that soil map unit. Overall, the selected SSURGO dominant soil types accounted for approximately 91% of the total area of the combined map units.

2.5. Parallel implementation of EPIC

It is computationally expensive to execute EPIC for a large number of modeling units. Each EPIC run in Iowa took about 10 s. Were the runs implemented serially, the time required for completing the entire EPIC runs would have been 486 h when using SSURGO (175,184 runs) and 517 h when using STATSGO. As each EPIC run generates at least three output files, the subsequent processing and analysis of EPIC outputs will additionally cost significant amount of computational resources. Therefore, we alleviated the computational burden by adapting the parallel computing software developed by Zhang et al. (2013a) for parallel EPIC implementation, which is referred to as Parallel-EPIC hereafter. Parallel-EPIC combines Python (python.org), mpi4py (Dalcin et al., 2011), OpenMPI (www.open-mpi.org), and PostgreSQL (www.postgresql.org) to simultaneously employ multiple processors to execute EPIC runs and derive and upload simulation variables into a relational PostgresQL database for further data analysis.

Combining Python with OpenMPI through mpi4py enables Parallel-EPIC to simultaneously submit jobs to processors distributed across a computing cluster in a Master–slave configuration (Zhang et al., 2013a; Fig. 4). Parallel-EPIC first identifies a Master processor among the processors allocated to the submitted job, which splits the entire EPIC runs into M folders. Based on our experiments, Parallel-EPIC is much more efficient when M is large, as file related operations are more efficient in folders with smaller numbers of existing files. Next, the Master processor sends commands to the m slaves to execute EPIC and process its outputs in parallel for the first n folders. This procedure is repeated until all M folders are treated. We used the Evergreen

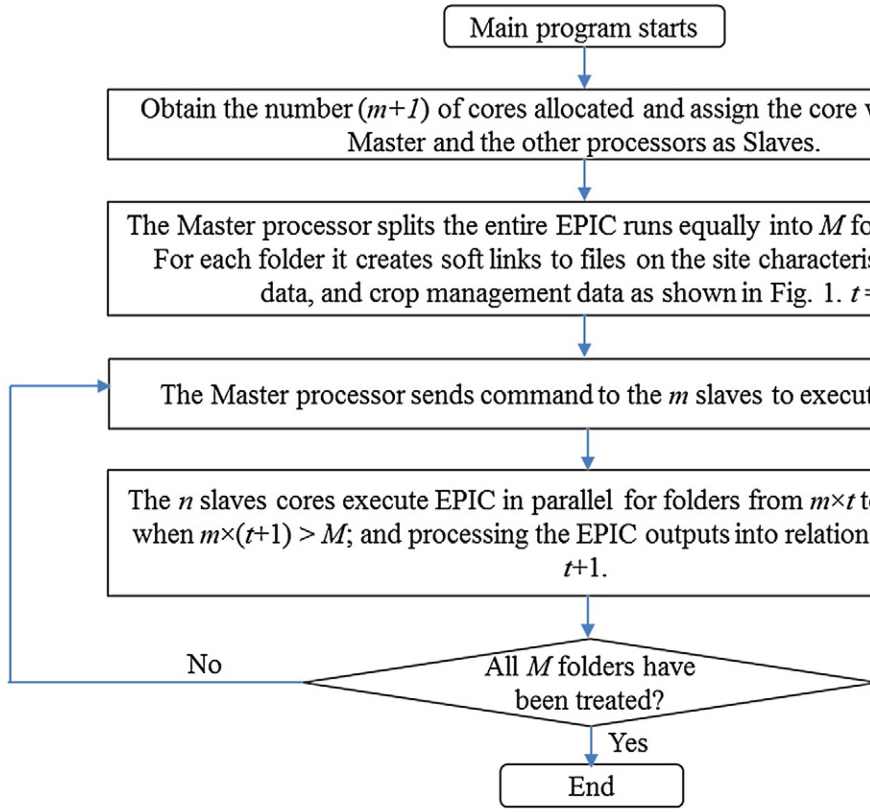

Fig. 4. Diagram of parallel execution of EPIC with a Master–slave configuration.

computing cluster hosted by the Pacific Northwest National Laboratory and University of Maryland (evergreen.umd.edu) as the platform for our model simulations. Its configuration and capacity have been detailed in Zhang et al. (2013a). Although Evergreen has over 2000 processors, we only employed 200 because of an input/output (I/O) bottleneck resulting from intensified resource competition and increased communication cost between computing nodes when larger number of processors is used. We estimated that Parallel-EPIC led to a 19-fold improvement in the efficiency of implementing EPIC when compared to a serial implementation.

3. Results and discussion

3.1. SOC stocks estimated using STATSGO and SSURGO

We derived SOC stocks on croplands in Iowa using both STATSGO and SSURGO databases. The calculated average entire soil profile SOC stocks were 245.7, 231.6, and 229.5 Mg C ha^{-1} for SSURGO, STATSGO-dominant, and STATSGO-weighted, respectively. As STATSGO maps were compiled by generalizing more detailed (SSURGO) soil survey maps, here we used SSURGO as the benchmark for comparison. Compared with SSURGO, the STATSGO-dominant and STATSGO-weighted schemes slightly underestimate SOC stock on the 8,799,949 ha cropland in Iowa by 5.6 (or 124 Tg C) and 6.6% (or 142 Tg C), respectively.

With different databases and schemes, we also obtained spatial SOC stock maps with notable difference in spatial patterns (Fig. 5). For example, the spatial map derived with STATSGO-dominant shows the area with the highest SOC intensity in north-east Iowa, while STATSGO-weighted and SSURGO identified the area with the highest SOC intensity in north central Iowa. As compared with the spatial distribution of SSURGO SOC, STATSGO-weighted and STATSGO-dominant had R^2 values of 0.8 and 0.27, respectively. In general, the spatial pattern of SOC derived from STATSGO-weighted is close to that of SSURGO; while there is substantial spatial discrepancy between STATSGO-dominant and SSURGO. In addition, the STATSGO-dominant and STATSGO-weighted schemes resulted in SOC stock maps exhibiting apparent differences in spatial pattern, as evidenced by a R^2 value of 0.5 that indicates 50% disagreement in SOC spatial variation between these two spatial maps.

These differences in SOC distributions are well illustrated at the local scale in Franklin County, Iowa (Fig. 5). Clearly, SSURGO preserves more heterogeneous details than the two STATSGO based schemes. Contrasting spatial patterns of SOC were also observed. SSURGO estimated higher SOC intensity in the west of the county, while STATSGO-dominant estimated higher SOC intensity in the northeast part. In addition, STATSGO-weighted lost much spatial variability in comparison to both SSURGO and STATSGO-dominant. Although the three schemes derived similar SOC estimates at the state level, the local scale analysis highlights the substantial divergence between their spatial SOC patterns, which represents an important source of uncertainty for carbon balance simulation.

3.2. Sensitivity of carbon flux simulations to different soil data

EPIC simulated crop yield of corn and soybean averaged over 2000–2008 at the county scale was compared with national inventory data from USDA-NASS' Quick Stats (quickstats.nass.usda.gov; Table 1). State level average crop yield matched well with NASS observations, with relative errors <5% for all scenarios. Simulated corn yield of all three scenarios had a relative error <20% for most counties and a relative error less than 10% for ~70% of the counties. EPIC's performance for soybean was slightly lower, with 75–81% counties with a relative error <20% and 32–48% counties with a relative error <10%. The small biases in simulated corn and soybean yields are indicative of the satisfactory performance of EPIC for simulating biomass growth on cropland in Iowa. As crop yield is directly determined by the total biomass

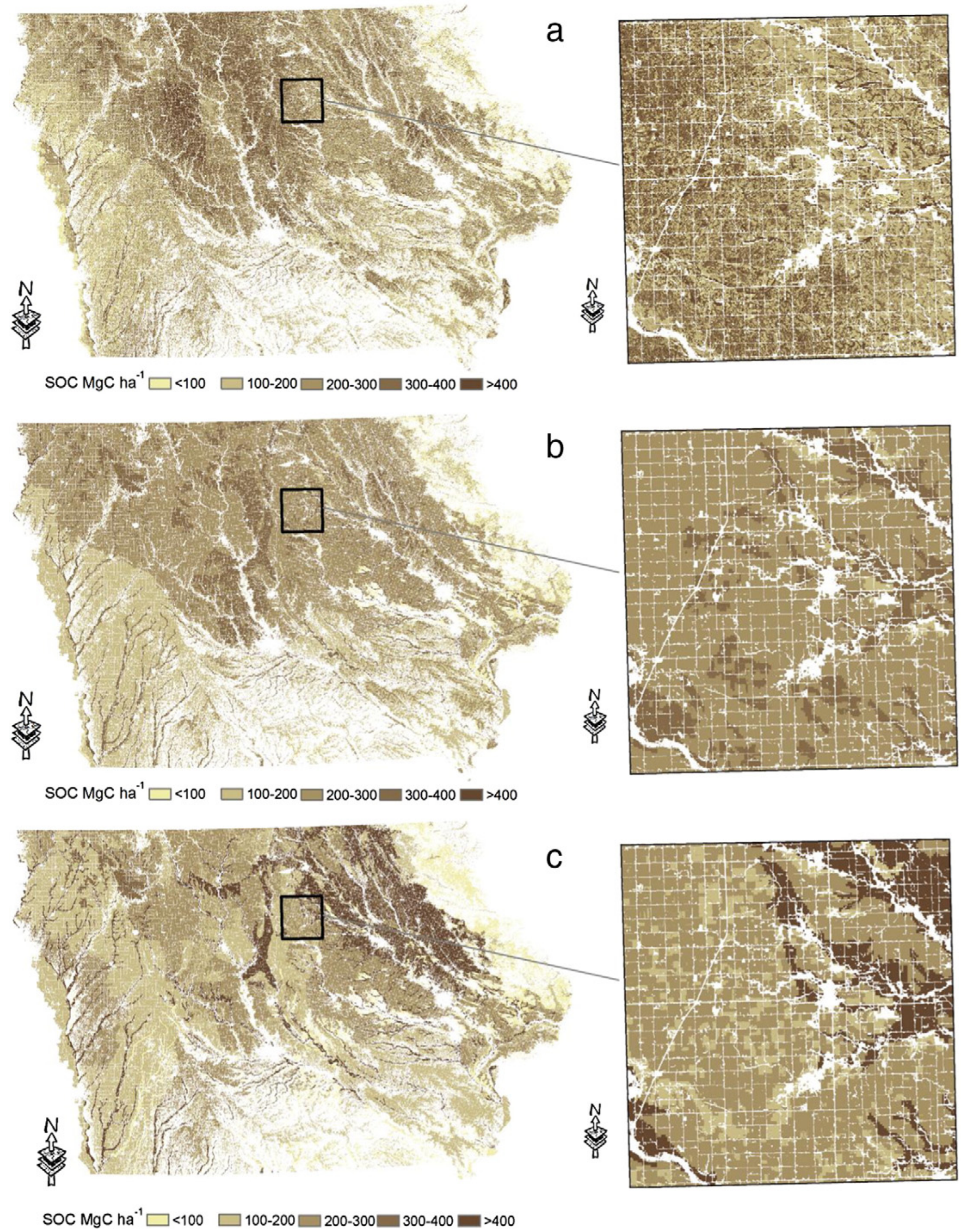

Fig. 5. Spatial distribution of SOC intensity over croplands in Iowa derived using (a) SSURGO, (b) STATSGO-weighted, and (c) STATSGO-dominant. These three maps for Franklin County, Iowa, are used to illustrate the difference between spatial SOC distributions derived from the three schemes at the local scale.

production (or NPP), it is reasonable to ascertain that EPIC capably simulated NPP thereby indicating that EPIC can serve as a useful tool for providing insights into the sensitivity of C budget modeling to soil data resolution.

Although all three soil datasets resulted in similar simulations of crop yield at the state level, we observed varying performances for capturing spatial variance at the county scale. SSURGO explained the highest amount of variation of observed spatial patterns in corn and soybean yields, followed by STATSGO-weighted and finally STATSGO-dominant, explaining the least. In addition, SSURGO simulated crop yield has similar or higher percentage of counties with relative errors less than 10 or 20% in comparison to STATSGO-dominant and STATSGO-weighted. Overall, SSURGO slightly outperformed STATSGO-weighted and both of them were much better than STATSGO-dominant in terms of capturing spatial variability of observed crop yields.

As SOC content is closely related to many important soil properties (such as texture), the remarkable difference in SOC content and its spatial distribution between SSURGO and STATSGO was expected to have significant implications for agroecosystem modeling with respect to crop productivity and C balance. On average, estimated state level NEP was −2383, −2437, and −2441 kg C ha^{-1} yr^{-1} for SSURGO, STATSGO-dominant, and STATSGO-weighted, respectively. These values are close to the range of annual cropland NEP (−2912 to −4076 kg C ha^{-1} yr^{-1}) in Iowa estimated with an inventory method

Table 1
Comparison between crop yields simulated by EPIC and reported by USDA-NASS.

	Corn				Soybean			
	Average yield[a]	R^{2b}	P10[c] (%)	P20[d] (%)	Average yield	R^2	P10 (%)	P20 (%)
USDA-NASS	9.21	–			2.99	–		
SSURGO	9.53	0.38	77	98	3.12	0.16	48	81
STATSGO-dominant	9.58	0.29	69	98	3.11	0.12	32	75
STATSGO-weighted	9.48	0.38	71	100	3.08	0.15	34	77

[a] Area weighted average of multi-year average crop yield across all counties in Iowa (Mg dry matter ha^{-1} yr^{-1}).
[b] Spatial correlation between county scale EPIC simulated and NASS observed crop yield.
[c] Percentage of counties with relative error less than 10%.
[d] Percentage of counties with relative error less than 20%.

(West et al., 2010), corroborating the credibility of the high-resolution agroecosystem modeling system for quantifying regional-scale C balance. Similar to the SOC analysis, the state level NEP simulated with different soil data did not much differ from each other.

We further examined the NEP maps at the county and grid scales, and found pronounced differences in EPIC-simulated NEP using the three soil datasets (Fig. 6). In general, the spatial patterns of NEP correspond well to those of SOC intensity (Figs. 5 and 6). Thus our spatial analyses of NEP and SOC stocks led to similar results. With the SSURGO estimated NEP as a baseline, STATSGO-dominant and STATSGO-weighted schemes resulted in R^2 values of 0.21 and 0.8, respectively, confirming that the STATSGO-dominant scheme has led to a significant loss of spatial details of soil data and in turn alters modeling outcomes at the local scale. The two STATSGO based schemes also exhibited significant difference in spatial NEP estimation, as indicated by the low spatial correlation (a R^2 value of 0.44 between them). The three maps for Franklin County illustrate the variation in the spatial patterns of NEP estimates derived from the different soil data.

To quantitatively measure the difference between the NEP maps, we calculated three spatial maps by subtracting NEP estimated with one scheme from that with another scheme (Fig. 7). As compared with SSURGO, for most cropland areas in Iowa (>60%), both STATSGO-weighted and STATSGO-dominant underestimated or overestimated NEP by more than 1 Mg C ha^{-1} yr^{-1}, which is about half of the average cropland NEP across Iowa. In Franklin County, the percentage of area with deviations larger than 1 Mg C ha^{-1} yr^{-1} was greater than 70%. The NEP difference between the two STATSGO based schemes is not as prominent as that between SSRURGO and STATSGO. Only about 12.5% of the total cropland had deviations larger than 1 Mg ha^{-1} yr^{-1}. However, we still identified over 30% area with deviations larger than 0.5 Mg ha^{-1} yr^{-1}. Overall, the local scale analysis emphasizes the prominent deviations between NEP estimated with different soil data.

3.3. An example of using SSURGO and STATSGO to identify marginal lands

Here, we used both SSURGO and STATSGO to identify marginal agricultural lands, a critical first step in modeling biomass production for cellulosic biofuels. Marginal lands have relatively low crop productivity due to diverse limiting factors, such as edaphic and climatic limitations. On the other hand, these lands are often suitable for grasses, short-rotation tree crops or other perennial vegetation with persistent roots that are better adapted to low-nutrient, erodible or droughty soils (Gelfand et al., 2013), thereby well suited for producing biomass feedstocks for advanced cellulosic biofuel. We followed Gelfand et al. (2013) to define marginal lands as rural lands falling into Land Capability Classes (LCCs) V–VII with slope gradients of <20% under nonforested vegetation. LCC describes land classes on the basis of use limitation such as erosion risk, soil depth, wetness and slope (Klingebiel and Montgomery, 1961). There are eight LCCs, ranging from class I (lands without any limitations for agricultural use) to class VIII (lands with severe limitations). Fig. 8 illustrates the distributions of marginal croplands in Iowa identified with SSURGO and STATSGO. SSURGO, STATSGO-weighted, and STATSGO-dominant identified 129,677, 221,963, and 152,179 ha marginal cropland, respectively. STATSGO-weighted almost doubled the area of marginal croplands identified with SSURGO. Although STATSGO-dominant estimated similar amount of marginal cropland to SSURGO, its pattern matched poorly with that estimated by SSURGO. The two STATSGO based marginal cropland maps exhibited marked differences in area and distribution. Clearly, using STATSGO to identify marginal lands could result in significantly biased estimates of the availability of marginal lands and where they are located, leading to unrealistic assessments of the capacity of these lands for biofuel production and the potential environmental outcomes. Much care should be taken when using STATSGO for spatial analyses that require detailed spatial information.

3.4. Discussion

As the rate of C flux between land and atmosphere is influenced by many interacting factors (such as terrain characteristics, soil properties, land use history, and management practices) varying at a local scale of tens to hundreds of m^2 (Houghton et al., 2009), spatially-explicit modeling approaches have been increasingly embraced by the community to reduce the scale of C sink/source identification and quantification, and understand the vulnerability of C stocks to future conditions. Both STATSGO and SSURGO are widely used in ecological modeling in the U.S. For C budget assessment, STATSGO has been employed in previous studies (e.g. West et al., 2010), while SSURGO has rarely been applied across large scales in a spatially explicit way. In this investigation, we examined the implications of using STATSGO and SSURGO for estimating spatially explicit C balances.

For SOC stocks on cropland, STATSGO and SSURGO provided very similar estimates at the state level with small differences around 5–6%. EPIC simulated NEP based on STATSGO and SSURGO were also close to each other at the state level. These results indicate that, for large scale C budget assessments, STATSGO can provide credible results, even when only the dominant soil type within each soil map unit was used. This is because STATSGO maps are compiled by generalizing more detailed (SSURGO) soil survey maps (USDA-NRCS National Soil Survey Center, 1995). When aggregated to the state level, high consistency is reached between STATSGO and SSURGO.

However, at grid and county scales, substantial deviations between SSURGO and STATSGO were observed, especially when only the dominant STATSGO soil component was selected. At the county scale, SOC and NEP estimated with STATSGO-weighted explained over 80% of the variation with SSURGO, while STATSGO-dominant only account for about 20% of the SSURGO spatial variability. The two schemes of using STATSGO also exhibited substantial differences in estimating C budgets, as indicated by the low R^2 of only about 50%. At the local scale (56 m), the two STATSGO based schemes either underestimate or overestimate NEP by at least 1 Mg ha^{-1} yr^{-1} for over 60% of the croplands in Iowa. Overall, spatial analyses of C budgets derived from different soil data emphasize that using STATSGO might result in significant biases in spatial patterns and magnitude of C budgets as compared with those

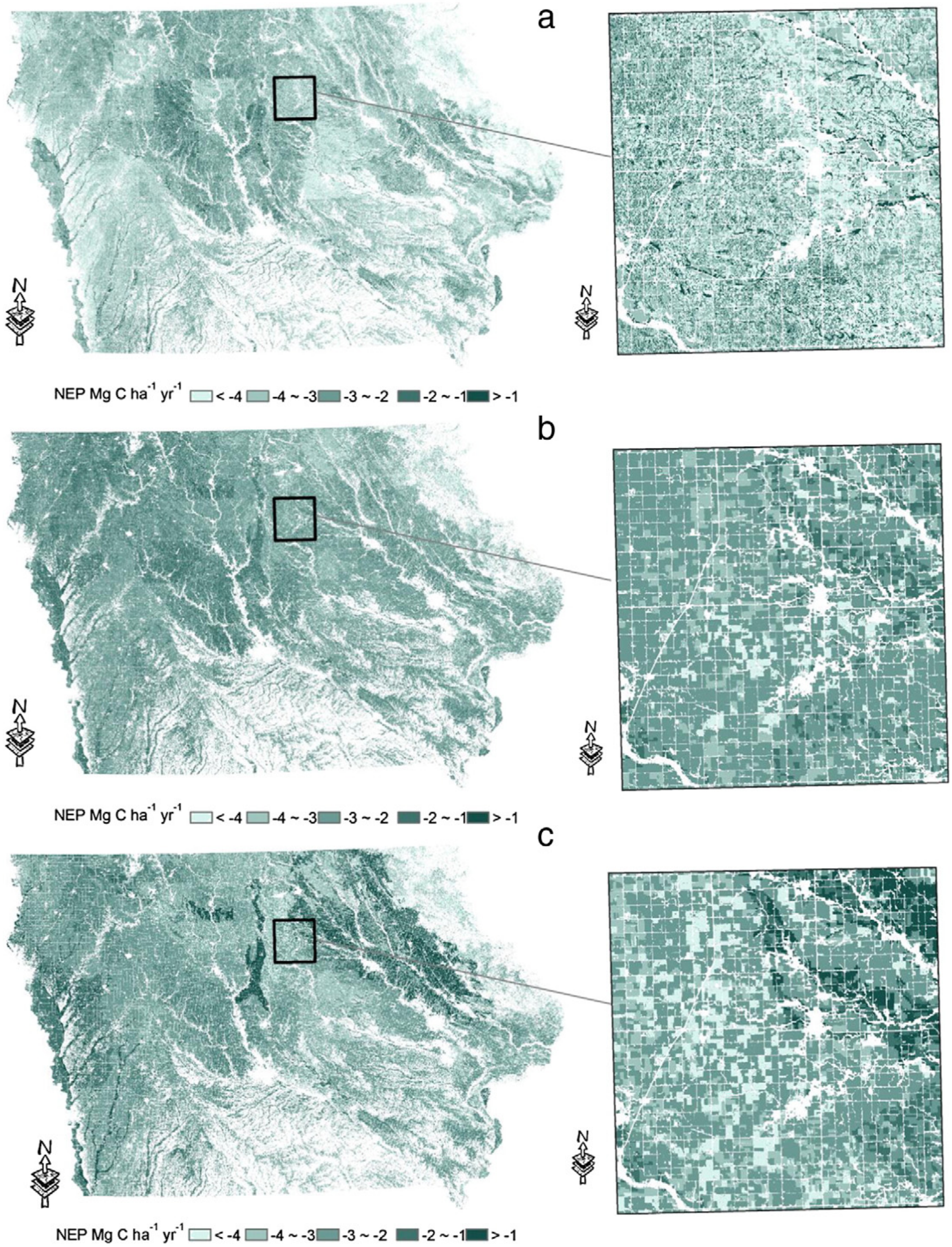

Fig. 6. Spatial distribution of NEP over croplands in Iowa derived using (a) SSURGO, (b) STATSGO-weighted, and (c) STATSGO-dominant. These three maps for Franklin County, Iowa, are used to illustrate the difference between spatial NEP distributions derived from the three schemes at the local scale.

predicted by SSURGO. In addition, the use of coarse resolution STATSGO data with the significant loss of spatial details of soil properties, precludes accurate spatial analyses at farm levels, such as identification of marginal lands for biofuel production (Gelfand et al., 2013). Collectively, SSURGO is more preferable to STATSGO for county and local scale assessment of C balance. We also note that only choosing the dominant soil type in each STATSGO polygon will cause further loss of spatial soil information, thus resulting in further reduction in accuracy.

The STATSGO-dominant scheme is popular in ecological modeling. For example, most of the publications of applying SWAT in the U.S. employed this scheme (www.card.iastate.edu/swat_articles). Our comparison shows that STATSGO-weighted achieved C simulations much closer to those based on SSURGO, STATSGO-weighted is preferred to STATSGO-dominant. The transition from STATSGO-dominant to STATSGO-weighted does not demand extra data preparation, but would greatly increase the computing and post-processing work load. On average each STATSGO polygon contains about 18 soil components, leading to the increase of the number of EPIC runs from 11,309 to 186,293 in the case of this study. It is important to note that the computational burden of running the STATSGO-weighted scheme is even higher than that of using SSURGO, which required 167,817 EPIC runs. However, although a gridded version of SSURGO was recently released (Soil Survey Staff, 2013), it does not provide a complete set of parameters that are necessary for driving process based agroecosystem models

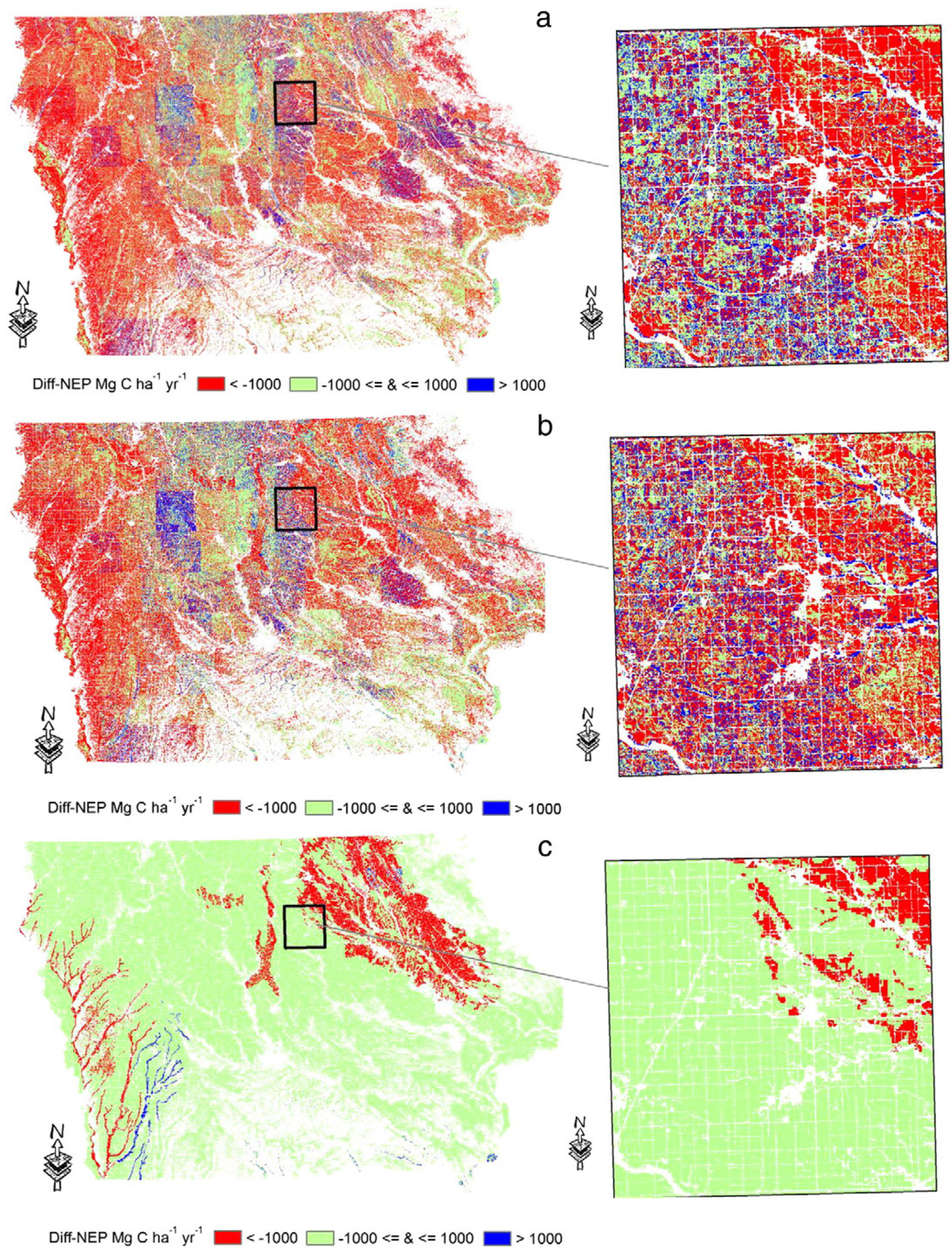

Fig. 7. Spatial distributions of the difference between estimated NEP over croplands in Iowa. (a as SSURGO minus STATSGO-weighted; b as SSURGO minus STATSGO-dominant; c as STATSGO-weighted minus STATSGO-dominant; The three maps for Franklin County in Iowa are used to illustrate the difference at the local scale).

like EPIC. Deriving soil parameters from the SSURGO database demands extra data processing and compilation time, especially for large scale simulations. Therefore, we suggest carefully balancing tradeoffs of using SSURGO and the two STATSGO based schemes and selecting the one that meets the accuracy objective with affordable cost of time and computation resources. For example, if the objective is to estimate C budgets at the state level, using STATSGO-dominant does not lead to much accuracy reduction, while demanding less than 10% of the computation cost of using SSURGO or STATSGO-weighted.

As noted above, a major concern regarding applying SSURGO and the STATSGO-weighted scheme is the amount of time and computational resources they require. For the Iowa simulations, implementation of EPIC would take more than two weeks with sequential computing. With a simple linear extrapolation to the U.S. Midwest, running EPIC with SSURGO and STATSGO-weighted might demand almost half a year. Spending so much time is often not allowed in order to meet project timelines, let alone the significant additional work load of preparing input data and processing and analyzing model outputs. When multiple scenarios are considered, the computational burden could be further aggravated. The Parallel-EPIC developed here helped cut the computational burden by a factor of 19, making the EPIC implementation affordable. In order to fully take advantage of the benefits of using

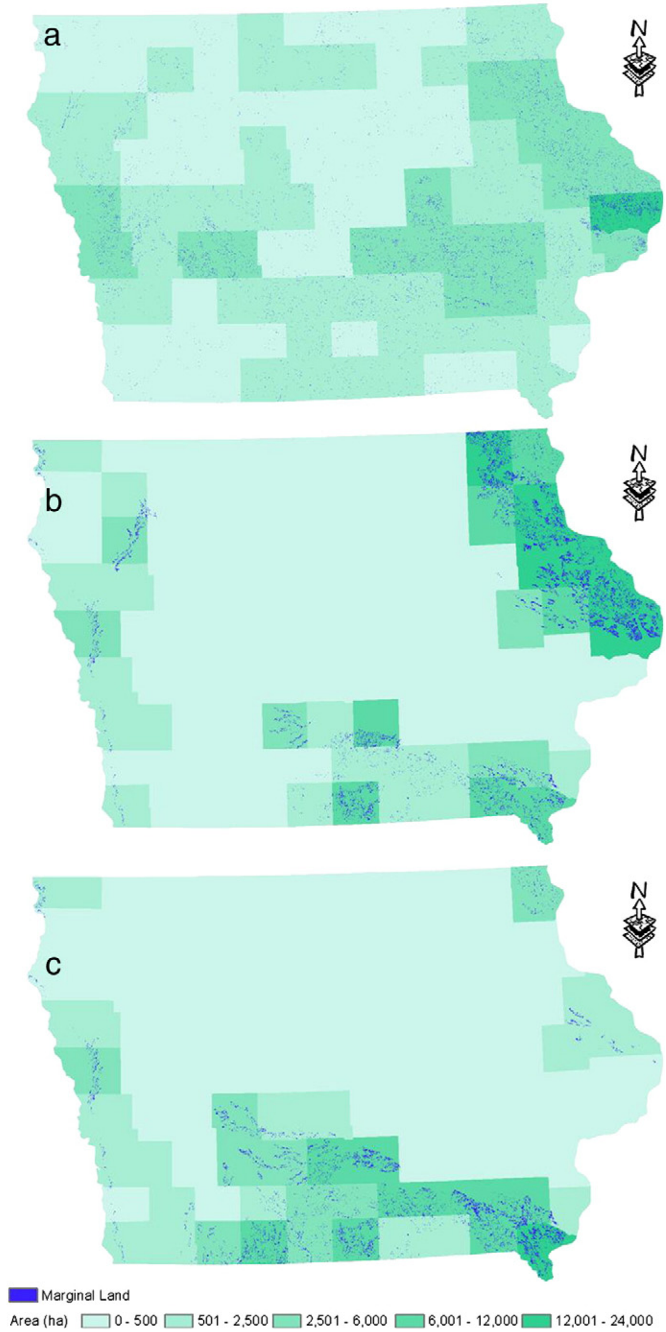

Fig. 8. Marginal lands identified with different soil data (a) SSURGO, (b) STATSGO-weighted, and (c) STATSGO-dominant.

high resolution data for C budgets assessment at the local scale, we envision that the traditional way of sequentially running agroecosystem models will transition to advanced parallel computing systems that can employ the strength of supercomputers to alleviate the computational burden.

We performed the sensitivity tests using the EPIC model only because of the intensive time requirements and the lack of readily available tools to build similar high-resolution systems for other models. As agroecosystem models represent biophysical and biogeochemical processes with diverse mechanisms and equations, this sensitivity to different soil data may vary substantially and deserves further research. In addition, our results were derived with only one combination of climate, land use, and crop management practice datasets. If other sources of data are acquired and embedded into our modeling system to replace the datasets employed here, the uncertainty of the estimated responses of C flux to soil resolution could be further amplified. Notably, a wide range of future climate change scenarios and our incomplete understanding of how terrestrial C cycling will respond to new water availability, temperature, and photosynthesis conditions may further widen the uncertainty of estimating C flux under various management practices. A systematic analysis of the resolution sensitivity of C flux simulations driven with diverse climate forcing, land use maps, and management details deserves further exploration in the future, in order to provide more credible assessment of the effectiveness of alternative C sequestration practices as climate change mitigation is becoming more imminent.

4. Conclusions

Accurate quantification of terrestrial C cycling is critical for designing effective policies and management practices to stabilize atmospheric CO_2 concentrations. Using high resolution spatial data has been promoted for reducing the uncertainty of the magnitude, spatial and temporal patterns of terrestrial C sinks and sources. In this study, we examined the sensitivity of EPIC, a widely used agroecosystem model, to different soil data of varying spatial resolutions. STATSGO (weighted- and dominant-approaches) and SSURGO were employed as input data to simulate crop yield and NEP. The computational burden of using STATSGO-weighted or SSURGO can be eased by adopting a parallel processing system as demonstrated here.

By alternatively implementing EPIC with SSURGO and STATSGO and analyzing simulation results, we found that EPIC simulated crop yield and NEP were not sensitive to soil data resolution at the state level, but exhibited substantial deviations from each other at the county and local scales. In general, SSURGO performed slightly better for crop yield simulation than STATSGO-weighted. STATSGO-dominant achieved much poorer county scale crop yield simulation, because the dominant soil components on average cover only about 20% of area of the corresponding STATSGO polygons. Using SSURGO as the benchmark, we found that NEP simulated with the two STATSGO schemes obtained close NEP at the state level, but deviated substantially at the county and local scales. At the grid (56 m) scale, for most cropland areas (over 60%), the deviations were larger than 1 Mg C ha^{-1} yr^{-1} or about half of the average cropland NEP across Iowa. Further comparison in the Franklin County clearly demonstrated the mismatch of NEP patterns derived with different soil data at the local scale. In addition, marked difference was observed between the marginal lands identified using SSURGO and STATSGO.

Our results show that using coarse soil resolution data could lead to significant uncertainty of C budgets assessments at these local scales. STATSGO, particularly the STATSGO-dominant soil approach, is more suitable for state-level applications requiring less computing cost, while employing the STATSGO-weighted soil approach or SSURGO is necessary when local-scale accuracy is emphasized (e.g. assessment of variation and strength in C sink/source across spatial cropping systems). Overall, choosing between different soil datasets should be based on a careful assessment of tradeoffs between accuracy requirements, application scales, and available computational resources. The modeling tool and analyses presented here represent an advancement in exploring the integration of mechanistic model, spatially-resolved data, surveyed management data, and super computing resources for understanding and addressing uncertainty of cropland C budgets as influenced by soil resolution, thereby contributing to effective C management.

Conflict of interest

The authors identify no conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, the submitted work.

Acknowledgments

We sincerely appreciate the valuable comments provided by the anonymous reviewers, which greatly improved the quality of this paper. This work was partially funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494, DOE BER Office of Science KP1601050, DOE EERE OBP 20469-19145), NASA (NNH08ZDA001N and NNH12AU03I), and USDA (CSREES-2009-34263-19774 (G-1449-1) and NIFA-2010-34263-21075 (G-1470-3)). The views expressed here are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

References

Allmaras RR, Schomberg HH, Douglas Jr CL, Dao TH. Soil organic carbon sequestration potential of adopting conservation tillage in U.S. croplands. J Soil Water Conserv 2000;55:365–73.

Apezteguía HP, Izaurralde RC, Sereno R. Simulation study of soil organic matter dynamics as affected by land use and agricultural practices in semiarid Córdoba, Argentina. Soil Tillage Res 2009;102:101–8.

Arnold JG, Srinivasan R, Muttiah RS, Allen PM. Continental scale simulation of the hydrologic balance1. JAWRA J Am Water Res Assoc 1999;35(5):1037–51.

Buell GR, Markewich HW. Data compilation, synthesis, and calculations used for organic-carbon storage and inventory estimates for mineral soils of the Mississippi River Basin. USGS Professional Paper 1686-A. Madison, WI: U.S. Geological Survey; 2004.

Causarano HJ, Shaw JN, Franzluebbers AJ, Reeves DW, Raper RL, Balkcom KS, et al. Simulating field-scale soil organic carbon dynamics using EPIC. Soil Sci Soc Am J 2007;71:1174–85.

Causarano HJ, Doraiswamy PC, McCarty GW, Hatfield JL, Milak S, Stern AJ. EPIC modeling of soil organic carbon sequestration in croplands of Iowa. J Environ Qual 2008;37:1345–53.

CTIC (Conservation Technology Information Center). Crop residue management survey. West Lafayette, Indiana, USA: Conservation Technology Information Center; 2007.

Dalcin LD, Paz RR, Kler PA, Cosimo A. Parallel distributed computing using python. Adv Water Resour 2011;34:1124–39.

Davidson EA, Lefebvre PA. Estimating regional carbon stocks and spatially covarying edaphic factors using soil maps at three scales. Biogeochemistry 1993;22:107–31.

Denman KL, Brasseur G, Chidthaisong A, Ciais P, Cox PM, Dickinson RE, et al. Coupling between changes in the climate system and biogeochemistry. Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press; 2007.

Dolman AJ, Gerbig C, Noilhan J, Sarrat C, Miglietta F. Detecting regional variability in sources and sinks of carbon dioxide: a synthesis. Biogeosciences 2009;6:1015–26.

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, et al. The shuttle radar topography mission. Rev Geophys 2007;45(2):RG2004.

Gassman PW, Reyes MR, Green CH, Arnold JG. The soil and water assessment tool: historical development, applications, and future research directions. Trans ASABE 2007;50(4):1211–50.

Gelfand I, Sahajpal R, Zhang X, Izaurralde RC, Robertson GP. Sustainable bioenergy production from marginal lands in the US Midwest. Nature 2013;493:514–7.

He X, Izaurralde RC, Vanotti MB, Williams JR, Thomson AM. Simulating long-term crop productivity and soil organic carbon dynamics with the EPIC model using data from Arlington, WI. J Environ Qual 2006;35:1608–19.

Houghton RA, Hall F, Goetz SJ. Importance of biomass in the global carbon cycle. J Geophys Res 2009;114:G00E03. http://dx.doi.org/10.1029/2009JG000935.

IPCC (Intergovernmental Panel on Climate Change). Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. In: Solomon S, Qin D, Manning M, Marquis M, Averyt K, Tignor MMB, Miller Jr HL, Chen Z, editors. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press; 2007.

Izaurralde RC, Williams JR, McGill WB, Rosenberg NJ, Quiroga Jakas MC. Simulating soil C dynamics with EPIC: model description and testing against long-term data. Ecol Model 2006;192(3–4):362–84.

Izaurralde RC, Williams JR, Post WM, Thomson AM, McGill WB, Owens L, et al. Long-term modeling of soil C erosion and sequestration at the small watershed scale. Clim Change 2007;80:73–90. http://dx.doi.org/10.1007/s10584-006-9167-6.

Johnson DM, Mueller R. The 2009 crop data layer. Photogramm. Eng. Remote Sens. 2010;76:1201–5.

Jones CA, Dyke PT, Williams JR, Kiniry JR, Benson VW, Griggs RH. EPIC: an operational model for evaluation of agricultural sustainability. Agr Syst 1991;37:341–50.

Klingebiel AA, Montgomery PH. Land-capability classification. Issue 210 of agriculture handbook, soil conservation service. U.S. Department of Agriculture; 196121.

Lal R, Bruce JP. The potential of world cropland soils to sequester C and mitigate the greenhouse effect. Environ Sci Pol 1999;2:177–85.

Legates DR, Gregory JM. Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation. Water Res Res 1999;35(1):233–41.

Mednick AC. Does soil data resolution matter? State soil geographic database versus soil survey geographic database in rainfall–runoff modeling across Wisconsin. J Soil Water Conserv 2010;65(3):190–9.

Moureaux C, Debacq A, Hoyaux J, Suleau M, Tourneur D, Vancutsem F, et al. Carbon balance assessment of a Belgian winter wheat crop (*Triticum aestivum* L.). Glob. Chang. Biol. 2008;14:1353–66.

Nichols J, Kang S, Post WM, Wang D, Bandaru P, Manowitz D, et al. HPC-EPIC for high resolution simulations of environmental and sustainability assessment. Comput Electron Agric 2011;79(2):112–5.

Ogle SM, Breidt FJ, Easter M, Williams S, Killian K, Paustian K. Scale and uncertainty in modeled soil organic carbon stock changes for US croplands using a process-based model. Glob Chang Biol 2010;16:810–22.

Parton WJ, Ojima DS, Cole CV, Schimel DS. A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. In: Bryant RB, Arnold RW, editors. Quantitative modeling of soil forming processes. Madison, WI: Soil Science Society of America; 1994. p. 147–67. [Special Publication 39].

Paustian K, Brenner J, Killian K, Cipra J, Williams S, Elliott ET, et al. State-level analyses of C sequestration in agricultural soils. In: Kimble JM, et al, editors. Agricultural practices and policies for carbon sequestration in soil. Boca Raton, FL: CRC Press; 2002. p. 193–204.

Rasmussen C. Distribution of soil organic and inorganic carbon pools by biome and soil taxa in Arizona. Soil Sci Soc Am J 2006;70(1):256–65.

Saby NPA, Bellamy PH, Morvan X, Arrouays D, Jones RJA, Verheijen FGA, et al. Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? Glob Chang Biol 2008;14:2432–42.

Schwalm CR, Williams CA, Schaefer K, Anderson R, Altaf Arain M, Baker I, et al. A model-data intercomparison of CO_2 exchange during a large scale drought event: results from the NACP site synthesis. J Geophys Res Biogeosci 2010;115:G00H05. http://dx.doi.org/10.1029/2009JG001229.

Secchi S, Kurkalova L, Gassman PW, Hart C. Land use change in a biofuels hotspot: the case of Iowa, USA. Biomass Bioenergy 2011;35(6):2391–400.

Smith P, Martino D, Cai Z, Gwary D, Janzen H, Kumar P, et al. Agriculture. In climate change 2007: mitigation. Contribution of working group III to the fourth assessment report of the intergovernmental panel on climate change. In: Metz B, Davidson OR, Bosch PR, Dave R, Meyer LA, editors. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press; 2007.

Soil Survey Staff. Soil survey geographic (SSURGO) database for [IOWA]. Available online at http://soildatamart.nrcs.usda.gov, 2012[verified on August 20, 2013].

Srinivasan R, Arnold JG, Jones CA. Hydrologic modeling of the united states with the soil and water assessment tool. Water Resour Dev 1998;14(3):315–25.

Srinivasan R, Zhang X, Arnold JG. SWAT ungauged: hydrological budget and crop yield predictions in the Upper Mississippi River Basin. Trans ASABE 2010;53(5):1533–46.

Surendran Nair S, Kang S, Zhang X, Miguez FE, Izaurralde RC, Post WM, et al. Bioenergy crop models: descriptions, data requirements, and future challenges. GCB Bioenergy 2012;4(6):620–33.

Soil Survey Staff. Gridded Soil Survey Geographic (gSSURGO) Database for Iowa. United States Department of Agriculture, Natural Resources Conservation Service; 2013 [Available online at http://datagateway.nrcs.usda.gov/, verified on September 24, 2013].

Sus O, Williams M, Bernhofer C, Beziat P, Buchmann N, Ceschia E, et al. A linked carbon cycle and crop developmental model: description and evaluation against measurements of carbon fluxes and carbon stocks at several European agricultural sites. Agric Ecosyst Environ 2010;139:402–18.

Tilman D, Hill J, Lehman C. Carbon-negative biofuels from low-input high-diversity grassland biomass. Science 2006;314:1598–600.

USDA-ERS (Economic Research Service). Fertilizer use and price. Available at http://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx#.UmluAvmkqpA, 2013[verified on October 24, 2013].

USDA-FSA. Fact sheet: Conservation Reserve Program (CRP) benefits: water quality, soil productivity and wildlife estimates. Washington, DC: U.S. Department of Agriculture, Farm Service Agency; 2008.

USDA-FSA (United Department of Agriculture - Farm Service Agency). Conservation Reserve Program Annual Summary and Enrollment Statistics FY 2010. Available at http://www.fsa.usda.gov/Internet/FSA_File/annual2010summary.pdf, 2010[accessed on July 6, 2012].

USDA-NASS (United States Department of Agriculture National Agricultural Statistics Service). Usual planting and harvesting dates for U.S. field crops. Available at http://www.nass.usda.gov/Publications/Usual_Planting_and_Harvesting_Dates/uph97.pdf, 1997[accessed on Oct. 18, 2011].

USDA-NRCS (United Department of Agriculture – National Resources Conservation Service) National Soil Survey Center. Soil survey geographic (SSURGO) data base data use information. Available at ftp://ftp.igsb.uiowa.edu/gis_library/Support_Data/Soils/SSURGO.PDF, 1995[verified on September 23, 2013].

Wang X, He X, Williams JR, Izaurralde RC, Atwood JD. Sensitivity and uncertainty analyses of crop productivity and soil organic carbon simulated with EPIC. Trans ASAE 2005;48:1041–54.

West TO, Post WM. Soil organic carbon sequestration by tillage and crop rotation: a global data analysis. Soil Sci Soc Am J 2002;66:1930–46.

West TO, Brandt CC, Baskaran LM, Hellwinckel CM, Mueller R, Bernacchi CJ, et al. Cropland carbon fluxes in the United States: increasing geospatial resolution of inventory-based carbon accounting. Ecol Appl 2010;20(4):1074–86.

West TO, Brown ME, Duran RM, Ogle S, Moss RH. Definition, capabilities, and components of a terrestrial carbon monitoring system. Carbon Manag. 2013;4(4):413–22.

Williams JR. The EPIC model. In: Singh VP, editor. Computer models of watershed hydrology. Highlands Ranch, CO: Water Resources Publications; 1995. p. 909–1000.

Williams JR, Jones CA, Kiniry JR, Spanel DA. The EPIC crop growth model. Trans ASAE 1989;32:497–511.
Wu J, Ransom MD, Kluitenberg GJ, Nellis MD, Seyler HL. Land-use management using a soil survey geographic database for Finney County, Kansas. Soil Sci Soc Am J 2001;65(1):169–77.
Zhang X, Izaurralde RC, Manowitz DH, West TO, Post MA, Thomson AM, et al. An integrated modeling framework to evaluate the productivity and sustainability of biofuel crop production. Glob Chang Biol Bioenergy 2010;2(5):258–77.
Zhang X, Beeson P, Link R, Manowitz D, Izaurralde RC, Sadeghi A, et al. Efficient multi-objective calibration of a computationally intensive hydrologic model with parallel computing software in python. Environ Model Software 2013a;46:208–18.
Zhang X, Izaurralde RC, Arnold JG, Williams JR. Modifying the SWAT model to simulate cropland carbon flux: model development and initial evaluation. Sci Total Environ 2013b;463–464:810–22.
Zhong B, Xu YJ. Scale effects of geographical soil datasets on soil carbon estimation in Louisiana, USA: a comparison of STATSGO and SSURGO. Pedosphere 2011;21(4):491–501.