2016

Evaluating land cover influences on model uncertainties—A case study of cropland carbon dynamics in the Mid-Continent Intensive Campaign region

Zhengpeng Li
South Dakota State University, zli2807@jacks.sdstate.edu

Shuguang Liu
U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center

Xuesong Zhang
Pacific Northwest National Laboratory

Tristram O. West
U.S. Department of Energy

Stephen M. Ogle
Colorado State University

See last page for additional authors

Ecological Modelling 337 (2016) 176–187

Contents lists available at ScienceDirect

Ecological Modelling

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

Evaluating land cover influences on model uncertainties—A case study of cropland carbon dynamics in the Mid-Continent Intensive Campaign region

Zhengpeng Li[a,*], Shuguang Liu[b], Xuesong Zhang[c], Tristram O. West[d], Stephen M. Ogle[e], Naijun Zhou[f]

[a] Geospatial Sciences Center of Excellence, South Dakota State University, Brookings, SD 57007, United States
[b] U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, United States
[c] Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD 20740, United States
[d] U.S. Department of Energy, SC-23, Washington, DC 20585, United States
[e] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523, United States
[f] Department of Geographical Sciences, University of Maryland, College Park, 20742, United States

ARTICLE INFO

Article history:
Received 22 March 2016
Received in revised form 5 July 2016
Accepted 5 July 2016
Available online 18 July 2016

Keywords:
Biogeochemical model
Cropland carbon fluxes
Land cover
Uncertainty
Mid-Continent Intensive Campaign

ABSTRACT

Quantifying spatial and temporal patterns of carbon sources and sinks and their uncertainties across agriculture-dominated areas remains challenging for understanding regional carbon cycles. Characteristics of local land cover inputs could impact the regional carbon estimates but the effect has not been fully evaluated in the past. Within the North American Carbon Program Mid-Continent Intensive (MCI) Campaign, three models were developed to estimate carbon fluxes on croplands: an inventory-based model, the Environmental Policy Integrated Climate (EPIC) model, and the General Ensemble biogeochemical Modeling System (GEMS) model. They all provided estimates of three major carbon fluxes on cropland: net primary production (NPP), net ecosystem production (NEP), and soil organic carbon (SOC) change. Using data mining and spatial statistics, we studied the spatial distribution of the carbon fluxes uncertainties and the relationships between the uncertainties and the land cover characteristics. Results indicated that uncertainties for all three carbon fluxes were not randomly distributed, but instead formed multiple clusters within the MCI region. We investigated the impacts of three land cover characteristics on the fluxes uncertainties: cropland percentage, cropland richness and cropland diversity. The results indicated that cropland percentage significantly influenced the uncertainties of NPP and NEP, but not on the uncertainties of SOC change. Greater uncertainties of NPP and NEP were found in counties with small cropland percentage than the counties with large cropland percentage. Cropland species richness and diversity also showed negative correlations with the model uncertainties. Our study demonstrated that the land cover characteristics contributed to the uncertainties of regional carbon fluxes estimates. The approaches we used in this study can be applied to other ecosystem models to identify the areas with high uncertainties and where models can be improved to reduce overall uncertainties for regional carbon flux estimates.

1. Introduction

Understanding carbon sources and sinks is important for carbon management (Michalak et al., 2011). However, estimates of carbon dynamics in large regions still have large uncertainties among different methods (Ciais et al., 2010; Huntzinger et al., 2012; Ito, 2011). Intercomparisons between model estimates can help to identify the limitations of the models and suggest future research priorities. The North American Carbon Program (NACP) conducted a series of comparisons between model estimates and observations from local to continental scales (Huntzinger et al., 2012). For example, a comparison of 21 terrestrial biosphere models at multiple NACP tower sites showed that Net Ecosystem Exchange (NEE) simulation results were better in forest sites than in non-forest sites (Schwalm et al., 2010). Another study compared Gross Primary Production (GPP) between 26 terrestrial biosphere models and observations at flux tower sites (Schaefer et al., 2012).That study found that overall the model performance was poor in GPP estimates and

* Corresponding author.
E-mail address: zli2807@jacks.sdstate.edu (Z. Li).

http://dx.doi.org/10.1016/j.ecolmodel.2016.07.002
0304-3800/

was possibly caused by inadequate representation of observed light use efficiency. The results also suggested that model improvement should focus on improving leaf-to-canopy scaling and obtaining better estimates of the model parameters that control the light use efficiency. At the continental scale, a comparison of 19 terrestrial biosphere models found that ecosystem Net Ecosystem Productivity (NEP) for North America varied from −0.7 to + 2.2 PgC yr^{-1}, which was much narrower than estimates of GPP and respiration (Huntzinger et al., 2012). Another study on the North America carbon balance compared the NEE estimates between inventory-based estimates, atmospheric inversion models and terrestrial biosphere models (Hayes et al., 2012). The inventory-based estimate (−327 TgC yr^{-1}) was significantly different from the mean values of the atmospheric inversion models ((−931 TgC yr^{-1}) and the terrestrial biosphere models (−511 TgC yr^{-1}). For the terrestrial biosphere models, the estimated NEE values ranged from +29 TgC yr^{-1} to −3210 TgC yr^{-1}. Such large uncertainties in the model estimates could be driven by poorly simulated processes and input data (Hayes et al., 2012).

For regional simulations, land cover information usually is required as an important input to process-based models (Ahl et al., 2005). Different land cover inputs could bring different physical parameters to the biosphere model and create large differences in simulated outputs (Sellers et al., 1996). The comparison between multiple terrestrial biosphere models at flux tower sites found the biome classification was the most important factor controlling the model-data mismatch (Schwalm et al., 2010). Another comparison of global NPP estimates from multiple biosphere models also found that different vegetation inputs partially caused higher NPP differences at the borders of vegetation types (Cramer et al., 1999). However, the assessment of how land cover impacts the model uncertainties was informal, and so there is still a need for more research to better quantify the effects of land cover inputs on model uncertainty.

The Mid-Continent Intensive Campaign (MCI) was a project that focused on reducing the uncertainties in estimating carbon fluxes between the terrestrial surface and atmosphere (Ogle et al., 2006). Multiple methods have been applied in the MCI region to quantify ecosystem carbon fluxes (Li et al., 2014; Ogle et al., 2003; Schuh et al., 2013; West et al., 2010; Zhang et al., 2015). In these studies, land cover characteristics, such as the crop species and crop rotations, were found to impact the estimates of the carbon fluxes, such as NPP and soil organic carbon change (Li et al., 2014; Zhang et al., 2015). Based on these findings, it is possible to make further investigation into the influences of land cover on model uncertainties.

In this study, we investigated whether the observed patterns of the carbon fluxes uncertainties were related to the distribution of land cover. We made a null hypothesis: the spatial distribution of model uncertainties is random in the MCI region. This null hypothesis was tested on the uncertainties of three major carbon fluxes: net primary production (NPP), net ecosystem production (NEP) and change in soil organic carbon (SOC). The uncertainties of these carbon fluxes were calculated based on the estimates from three models: a crop inventory model; the Environmental Policy Integrated Climate (EPIC) model through the geospatial agricultural modeling system (GCAM) framework; and the General Ensemble biogeochemical Modeling System (GEMS). In situations where the null hypothesis was proved to be false, we further investigated the influences of three land cover characteristics with the uncertainties: cropland percentage, cropland richness and diversity.

2. Materials and methods

2.1. Study area

The research area is the Mid-Continent Intensive Campaign (MCI) region (Ogle et al., 2006). The MCI encompasses 678 counties from 11 states in the northern Great Plains and Western Corn Belt of the United States (Fig. 1). The land area in the MCI is about 124 million hectare (Mha) and more than 40% of the land area is used for agriculture. Corn, soybean, spring wheat, and winter wheat are the four major planted crops in the MCI region and occupy more than 90% of the planted area. The crop inventory data showed over 30 Mha of cropland area was used to plant corn and soybean, and about 10 Mha was planted with small grains and other crops in this region (West et al., 2008). The mean annual precipitation varies from 355 to 535 mm and the mean annual air temperature varies from 5 to 7 °C.

The spatial details of crop species in the MCI region are provided by the U.S. Department of Agriculture (USDA) crop land data layer (CDL) product (Boryan et al., 2011). The CDL program used remote sensing data from multiple satellite sensors and ancillary data to classify the crop types since the 1990s (Boryan et al., 2011). The two major satellite sensors used are the Advanced Wide Field Sensor (AWiFS) and Landsat Thematic Mapper (TM), both of which have high spatial resolution (56 m for AWiFS and 30 m for TM). The CDL map provided wall-to-wall mapping across the states with the spatial resolution at 30 m before 2005, and at 56 m between 2006 and 2010. The accuracies of the CDL products for major crop types are generally from 85% to 95% at the state level (Boryan et al., 2011). These high resolution crop maps have been widely used in biogeochemical models and with inventory data to estimate the carbon dynamics at regional and national scales (Li et al., 2014; West et al., 2010; West et al., 2008; Zhang et al., 2015). In the MCI region, CDL maps are available for all the states in 2007 and 2008.

2.2. Inventory

The inventory method estimates the carbon fluxes of crops based on county-scale crop yield data (NASS, 2013). The county-scale crop yield data include the reported crop planted and harvested area, crop production and crop yield estimates on an annual basis from 2001 to 2008. Yield data are reported for harvested crop commodities, therefore cover crops are not included. Generally the crop harvested area is about 1–3% smaller than the crop planted area at the state level, due to crop failures.

The inventory method calculated NPP for each crop from crop yield data using crop-specific parameters such as harvest indices, root:shoot ratio and estimated dry weight values (West et al., 2011; West et al., 2010). The SOC change is estimated by using empirical relationships between land management and soil carbon change based on crop species, land management, soil attributes and regional mean climate regimes (West et al., 2008). The annual estimates of NEE include the sum of net soil carbon change, uptake of crop carbon, and decomposition of above- and below-crop carbon. The spatial distribution of the NEE was calculated using weighted distribution and remote sensing land cover data (West et al., 2010). For this comparison, the NEP is estimated as the negative of NEE and the estimates are aggregated to county level.

2.3. EPIC

The Environmental Policy Integrated Climate (EPIC) model was originally developed based on site-level observations and has been extensively tested for many agricultural cropping systems landscapes (Causarano et al., 2008; Zhang et al., 2014, 2015). A recent development of the EPIC model used a geospatial agricultural modeling system (GAMS) to integrate the EPIC model with the spatially-explicit climate, land use, soil and management data for assessing regional carbon fluxes (Zhang et al., 2015, 2014).

Multi-year CDL maps (2007–2011) were processed by GAMS to provide crop rotation information for the regional simulation (Zhang et al., 2015). For each state, major crop rotations were

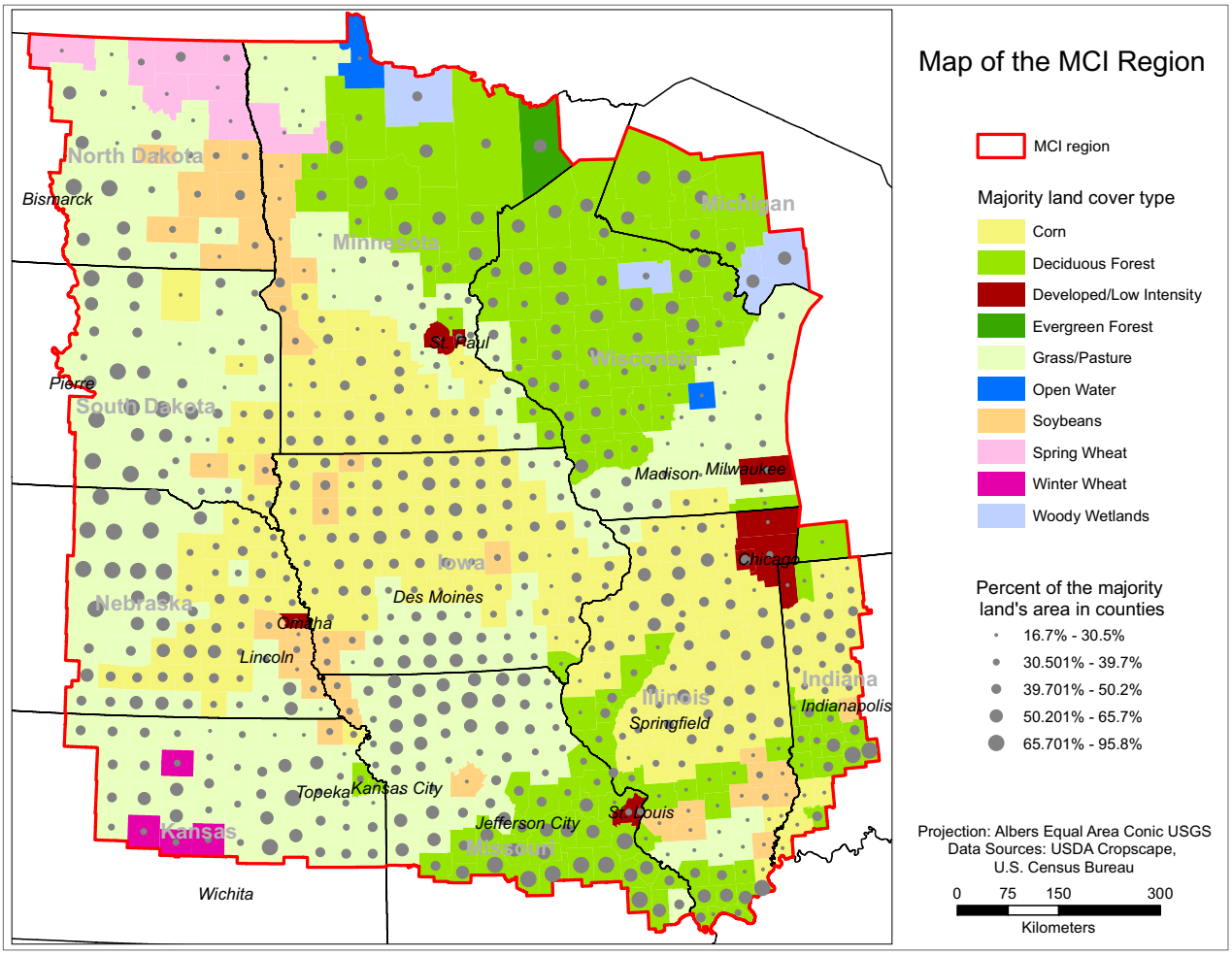

Fig. 1. The Mid-Continent Intensive Campaign (MCI) region boundary and land cover classes from the Cropland Data Layer land cover in 2008.

extracted from CDL maps and used to simulate land cover change in cropland areas. The soil data from the Natural Resources Conservation Service SSURGO were used for initializing soil carbon contents, and the climate inputs to the model were obtained from the North-American Land Data Assimilation System 2 (NASA, 2014). Crop management information such as tillage, conservation type, and fertilizer application rate was also used as inputs to the model. GAMS processed all the information into homogeneous spatial modeling units and performed EPIC simulations from 1991 to 2008 (Zhang et al., 2015).

In the EPIC model, NPP is computed a part of the plant canopy's interception of daily photosynthetically-active solar radiation. The NPP is affected by vapor pressure deficits, atmospheric CO_2 concentrations, nutrient availability, and other environmental controls and stresses. SOC dynamics is computed by considering many factors and processes, such as soil texture, pH, crop yields, atmospheric N input, fertilizer and manure, and tillage for the decomposition and transformation of soil C and nitrogen (N) from the model inputs. NEE was calculated as heterotrophic soil respiration minus the net C sequestration from the atmosphere into plant biomass (i.e. NPP) and is opposite in sign to NEP (Zhang et al., 2015). NEP is computed as the negative of NEE for the comparison.

2.4. GEMS

GEMS is a modeling framework developed to quantify the regional ecosystem carbon sequestration and its uncertainties (Liu, 2009; Liu et al., 2004). GEMS used an ensemble approach to apply land-cover/use data, along with information on soils, terrain, and management factors, to provide geospatially explicit inputs data to the ecosystem-level biogeochemical model. The uncertainty of model simulations can be quantified by a Monte Carlo based ensemble approach and multiple modeling runs in the region.

Spatial information about crop types was obtained from the CDL. The original crop types were regrouped into 6 representative crops (corn, soybean, spring wheat, winter wheat, other grains crops, other crops) for this study. The GEMS model was run for the MCI region using an equal distance (5 km) sampling approach and results were aggregated to the county level for comparison.

Meteorological inputs to the model were monthly minimum temperature, maximum temperature and precipitation from Oregon State University's Parameter-elevation Regressions on Independent Slopes Model (PRISM, 2004). The soil data were extracted from State Soil Geographic Data Base (STATSGO) (NRCS, 1994). The major crop growth parameters were calibrated using state-level crop yield data by GEMS internal subroutines (Li et al., 2014).

The biogeochemical model, EDCM, was used in GEMS to simulate carbon dynamics on agricultural land (Liu et al., 2003). EDCM is an ecosystem-level model that simulates soil carbon and nitrogen dynamics, vegetation primary productivity and water balance at monthly time steps. EDCM computes NPP based on vegetation potential production and environmental factors such as temperature, water and nitrogen. SOC dynamics are modeled as a

Table 1
Major inputs to the three models.

Method	Soil	LandCover	Crop rotation	Climate	Managements
Inventory	SSURGO	County reported harvest cropland area	Survey reported harvest data	Climate zone	Tillage
EPIC	SSURGO	NASS (56m)	NASS	NLDAS 2	Tillage, N fertilization
GEMS	STATSGO	NASS (5 km)	NASS	PRISM	Tillage, N fertilization

combination of soil movement (including the addition of manure) and decomposition. Decomposition of carbon is a function of soil carbon pool size and soil carbon decomposition rates, which are calculated based on the availability of temperature, water and nitrogen in each soil pool. The NEP on the cropland is calculated as the change of total ecosystem carbon plus the harvested carbon (grain and residue removal).

2.5. Model uncertainties and land cover characteristics

The major inputs for each model are presented in Table 1. To compare the model results and calculate the uncertainties, we aggregated the three carbon fluxes (NPP, NEP and SOC change) at the county level. For each county, the total cropland fluxes were calculated by adding the fluxes of all the crops. Then the mean value of the flux was calculated by dividing the total cropland fluxes by the total cropland area in the county.

The uncertainties of NPP and NEP fluxes were evaluated with the Coefficient of Variation (CV):

$$CV = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}}{\frac{1}{N}\sum_{i=1}^{N} x_i} \quad (1)$$

where N is the total number of estimated variables in each county (N = 3). x_i is the estimated variable from each method and $\bar{x}$ is the mean value of the three estimates (inventory, EPIC and GEMS). For SOC changes, which have a large portion of negative values, we used standard deviation (STDEV) instead of CV.

The cropland percentage in each county equals the planted cropland area divided by the total land area in the county. The planted cropland area is calculated using the number of cropland pixels in the annual CDL map multiplied by the pixel size. However, this planted cropland area is different from the cropland area inputs into the models. Each model uses its own approach to estimate the cropland area input: the inventory model used reported harvested cropland area; GEMS used 5 km sampling method and EPIC used 56 m CDL cropland area directly. Such input uncertainty can be propagated into the outputs. To evaluate the input uncertainty, we calculated the cropland area CVs using the cropland area inputs from three models:

$$CV = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}}{\frac{1}{N}\sum_{i=1}^{N} x_i} \quad (2)$$

where N is the number of cropland area used in each county (N = 3). x_i is the input cropland area from each method and $\bar{x}$ is the mean value of the three cropland area (inventory, EPIC and GEMS).

We used two indices to describe the land cover characteristics: land cover richness and land cover diversity. The land cover richness is defined as the number of unique land cover types inside each county. The land cover diversity is indicated by the Shannon equitability index. The Shannon equitability index is an index that is widely used in landscape ecology to describe the biodiversity. It is the Shannon diversity index divided by the maximum diversity and calculated as:

$$SI = \frac{-\sum_{i=1}^{M} p(i)\ln p(i)}{\ln(M)} \quad (3)$$

where, i is the land cover type in a county, $p(i)$ is the proportion of the value i to the total of the values, and M is the total number of values. For a well-sampled region, we can estimate this proportion as p(i) = area(i)/total_area, where area(i) is the area for each land cover within a county and total area is the area of all the land covers in the county. The Shannon equitability index takes values between 0 and 1, where lower values indicate more diversity and higher values indicate less diversity.

We designed a two-steps approach in this study: exploratory analysis and statistical analysis. The exploratory analysis methods include k-mean clustering and visual analyses, and the statistical analyses aim to detect spatial autocorrelation and to determine the linear relationship between variables.

The design of the statistic experiment is described in details in the Appendix. Both 2007 and 2008 data were used in the statistical analysis. The statistics and data mining method were implemented with R software and the spatial patterns were displayed using ArcGIS software.

3. Results

3.1. Model estimates on carbon fluxes

Fig. 2 shows the estimates of cropland area, NPP, NEP and SOC change in 2007 and 2008 at the county level for the MCI region. The total cropland area estimated from the three models was 53.0 ± 3.0 Mha in 2007 and 54.3 ± 3.1 Mha in 2008. The cropland area showed similar spatial distributions in both years (Fig. 2A, B). About 15% of the counties have cropland area smaller than 25,000 ha and about 30% of the counties have cropland area larger than 100,000 ha in the MCI region. Large cropland areas exist mainly in Illinois, Iowa, Nebraska, North Dakota, and South Dakota. The counties with small cropland area are in northern Minnesota, Missouri, Michigan and Wisconsin.

The total NPP estimated from the three models was 344.5 ± 5.8 TgC yr^{-1} in 2007 and 366.4 ± 38.4 TgC yr^{-1} in 2008. About 90% of the counties had NPP values between 250 and 850 gC m^{-2} yr^{-1} and 7% had NPP values greater than 850 gC m^{-2} yr^{-1} in 2007. In 2008, cropland NPP increased in most counties and 19% of the counties had NPP values higher than 850 gC m^{-2} yr^{-1}. These highest NPP values were mainly in Iowa and Illinois. Lower NPPs were in northern Minnesota, northern Wisconsin, and central Missouri (Fig. 2C, D).

The total NEP on croplands was 159.7 ± 7.7 TgC yr^{-1} in 2007 and 183.3± 47.8 TgC yr^{-1} in 2008 based on the three models. The county-level NEP had a smaller range than NPP. About 92% of the counties had NEP values between 250 and 450 gC m^{-2} yr^{-1} and 2% had NEP values greater than 450 gC m^{-2} yr^{-1} in 2007. In 2008, 78% of the counties had NEP between 250 and 450 gC m^{-2} yr^{-1},

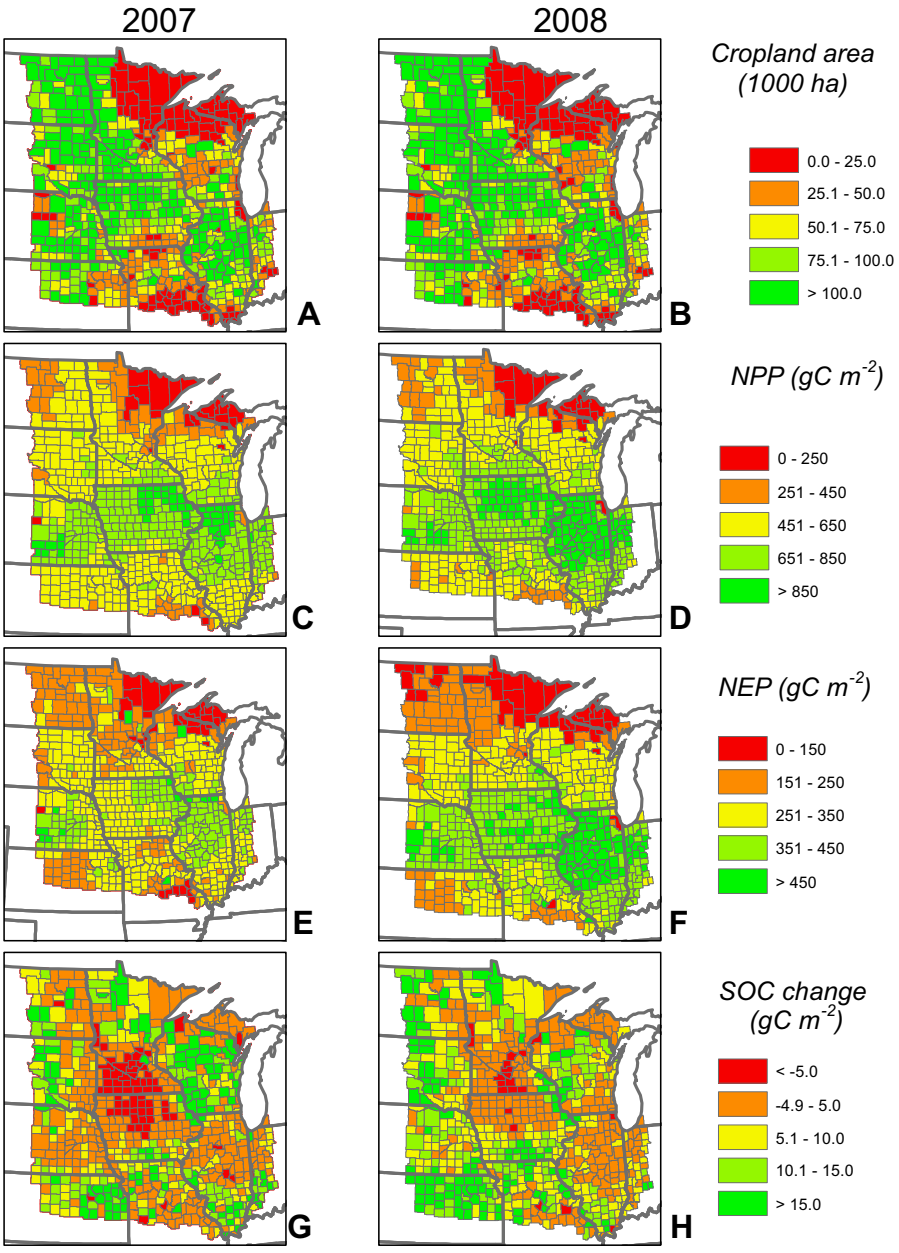

Fig. 2. Cropland area in 2007 (A) and 2008 (B); cropland mean Net Primary Production (NPP) in 2007 (C) and 2008 (D); cropland mean Net Ecosystem Production NEP in 2007 (E) and 2008 (F); cropland mean soil organic carbon (SOC) change in 2007 (G) and 2008 (H) in the Mid-Continent Intensive Campaign (MCI) region.

and about 16% of the counties had NEP values higher than 450 gC m^{-2} yr^{-1}. The spatial distributions of NEP showed similar patterns as NPP, with high values in Iowa and Illinois, and low values in northern Minnesota, northern Wisconsin, and central Missouri (Fig. 2E, F).

The total SOC change was 4.0 ± 4.9 TgC yr^{-1} in 2007 and 8.0 ± 10.5 TgC yr^{-1} in 2008. About 43% of the counties showed relatively small SOC changes (−4.9–5.0 gC m^{-2} yr^{-1}) in 2007. About 10% of the counties showed SOC change less than −5.0 gC m^{-2} yr^{-1} and these counties were located mainly in southern Minnesota and northern Iowa. In 2008, only 4% of the counties showed SOC change less than −5.0 gC m^{-2} yr^{-1} and about 60% of the counties showed SOC change higher than 5.0 gC m^{-2} yr^{-1}. The spatial distribution of SOC changes was quite different from the spatial distribution of NPP and NEP (Fig. 2G, H).

3.2. Model uncertainties

Fig. 3 shows the uncertainty of the estimates in cropland area, NPP, NEP and SOC change in 2007 and 2008. For cropland area, most counties had small CVs but some high CVs were found in northern Minnesota, northern Wisconsin, and central Missouri (Fig. 3A, B). The CVs of cropland area showed similar results in 2007 and 2008.

The three models agreed well on the NPP estimates in the MCI region. The CVs of NPP estimates showed that more counties had smaller CVs in 2007 than in 2008 (Fig. 3C, D). About 64% of the counties had CVs less than 0.2 in 2007 and only about 45% of the counties had CVs less than 0.2 in 2008. Higher CVs in 2008 were mainly located in Iowa and Illinois. It also seems that NPP CVs showed similar spatial patterns as the cropland area CVs. The highest NPP CVs tended to occur at counties with high cropland area CVs such as the northern Minnesota, northern Wisconsin and central Mis-

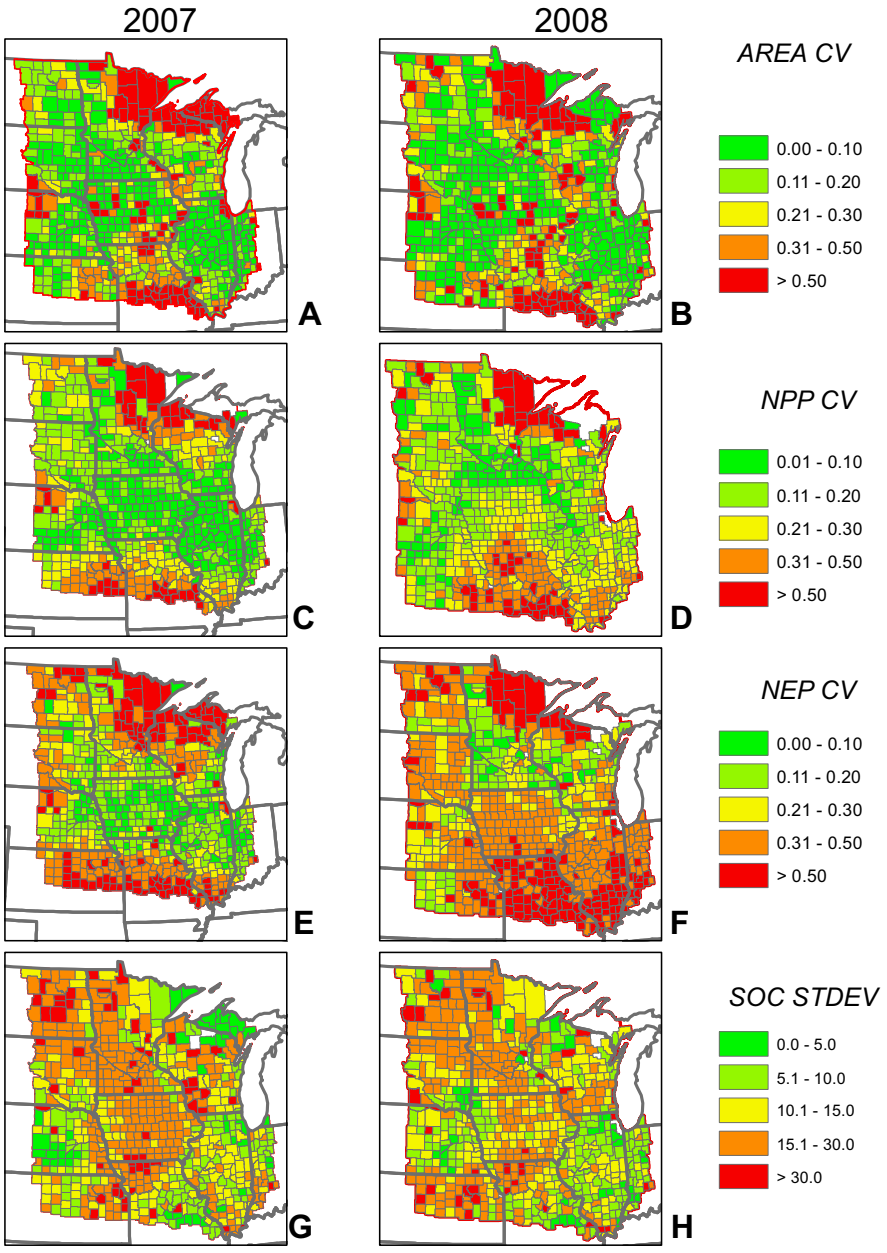

Fig. 3. Cropland area CVs in 2007 (A) and 2008 (B); cropland NPP CVs in 2007 (C) and 2008 (D); cropland NEP CVs in 2007 (E) and 2008 (F); cropland SOC change standard deviation in 2007 (G) and 2008 (H) in the Mid-Continent Intensive Campaign (MCI) region.

souri. The CVs of NEP showed similar spatial patterns as those for NPP but with higher values (Fig. 3E, F). Only 45% of the counties had NEP CVs less than 0.2 in 2007 and 15% of the counties had CVs less than 0.2 in 2008. This result indicates that NEP estimates from the three models had higher uncertainties than NPP. One noticeable difference between 2007 and 2008 was that NEP CVs were higher in Iowa and Illinois in 2008, similar as the CV changes in NPP.

The STDEVs of SOC changes showed different spatial patterns from the CV maps of NEP and NPP. High uncertainties indicated the model estimates diverged in the SOC changes in Iowa, Minnesota, and North Dakota. Low uncertainties were in Nebraska and Illinois (Fig. 3G, H). Based on these uncertainties, we are more confident that the cropland was a weak soil carbon sink in Nebraska and Illinois but less confident about the soil carbon loss in Iowa and south Minnesota where larger STDEVs were found.

We computed the correlation coefficients and p-values between the model uncertainties and the input land cover characteristics for all the counties (Table 2). For both 2007 and 2008, the CVs of cropland area showed significant positive correlations with the CVs of NPP and NEP. Meanwhile, there were significant negative correlations between the cropland percentage and the CVs of NPP and NEP. This indicated that in the counties with large cropland percentage, the cropland area CVs were small, as well as the CVs of the NPP and NEP. But in the counties with small cropland percentage, the CVs of cropland area, as well as the CVs of NPP and NEP, were large. In contrast, the STDEVs of SOC change did not show significant correlation with the CVs of cropland area, and less significant correlations with the cropland percentages than the CVs of NPP and NEP (Table 2).

Both cropland richness and Shannon equitability index showed negative correlations with the CVs of NPP and NEP (Table 2). That is, the uncertainties of the NPP and NEP were smaller in the county with higher richness or lower diversity. However, the p-values showed their correlations were less significant than cropland per-

Table 2
Correlation coefficient and p-value between the cropland area CVs, cropland percentage, richness and Shannon equitability index and model uncertainties in 2007 and 2008.

	NPP CV		NEP CV		SOC STDEV	
	Correlation coefficient	p-value	Correlation coefficient	p-value	Correlation coefficient	p-value
Cropland area	0.780	$<2.2\,e^{-16}$	0.700	$<2.2\,e^{-16}$	−0.087	0.0258
CVs	(0.812)[a]	($<2.2\,e^{-16}$)	(0.668)		(−0.038)	(0.324)
Cropland	−0.589	$<2.2\,e^{-16}$	−0.607	($<2.2\,e^{-16}$)	−0.098	0.0107
percentage	(−0.534)	($<2.2\,e^{-16}$)	(−0.404)		(−0.175)	($2.55\,e^{-10}$)
Cropland	−0.122	0.00145	−0.022	0.562	0.132	0.00059
richness	(−0.255)	($1.86\,e^{-15}$)	(−0.312)	($<2.2\,e^{-16}$)	(0.146)	(0.00134)
Shannon	−0.241	$2.14\,e^{-10}$	−0.182	$1.88\,e^{-6}$	−0.029	0.445
equitability	(−0.216)	($5.87\,e^{-13}$)	(−0.171)	($<1.03\,e^{-5}$)	(−0.069)	(0.105)

[a] The data in the parentheses are in 2008.

Table 3
Moran *I*'s analysis results in 2007 and 2008.

Variable	Moran's I index	z-score	p-value	Number of clusters[b]
NPP CV	0.457	26.420	0.000	6
	(0.475)[a]	(30.142)	(0.000)	(6)
NEP CV	0.374	26.118	0.000	7
	(0.373)	(24.887)	(0.000)	(7)
SOC change STDEV	0.193	13.224	0.000	6
	(0.198)	(12.914)	(0.000)	(6)

[a] The data in the parentheses are in 2008.
[b] From the k-means analysis.

centage. The STDEVs of the SOC changes did not show significant correlations with cropland richness and the Shannon equitability. These results indicated that the distribution of crop types had less impact on the uncertainties of SOC changes than the uncertainties of NPP and NEP.

3.3. Spatial patterns of model uncertainties

We performed the Moran's *I* analysis on the uncertainties and the results are listed in Table 3. Distributions of the model uncertainties exhibited statistically significant spatial patterns instead of being randomly distributed. With high Z-scores and low p-values all the results indicate that the model uncertainties (CVs and STDEVs) are positively spatially autocorrelated (i.e., similar CVs are clustered near one another). The uncertainties of NPP and NEP showed stronger spatial autocorrelation than the uncertainties of SOC in both years. Interestingly, the Moran's *I* values are similar for each type of uncertainty (NPP, NEP, SOC) between 2007 and 2008, indicating the spatial patterns of the model uncertainties are temporally stable.

The data mining method, k-means cluster analysis, was used to identify multiple clusters for the model uncertainties in both 2007 and 2008 (Fig. 4). The number of clusters for each fluxes uncertainties is given in Table 3. The clusters were not the same between the two years but showed some similarities. For example, a cluster with small NPP CVs was in Nebraska, Iowa and Illinois in 2007 and this cluster extended its range with larger CV values in 2008. This finding agrees with the NPP CV map of 2008, where larger CVs were shown in Iowa and Illinois. The cluster of NEP CVs also showed that the counties in Iowa were in one cluster in both 2007 and 2008. Generally for NPP and NEP, the clusters with small uncertainties are in cropland-dominated areas, such as Iowa and Illinois, and clusters with large uncertainties are in the counties with small cropland areas, such as northern Minnesota and northern Wisconsin. The clusters of SOC changes showed different spatial patterns than NPP and NEP. Clusters with high STDEV values were in Iowa, Minnesota, and North Dakota. Low uncertainties were in Nebraska and Illinois (Fig. 4E, F).

3.4. Hot spots and cold spots analysis

An analysis of hot/cold spots for cropland percentage, cropland cover richness and cropland Shannon equitability index within counties was conducted (Fig. 5). The hot spots of cropland percentage were located in corn and soybean dominated areas, such as central Iowa, southern Minnesota, eastern South Dakota, eastern Nebraska, and Illinois (Fig. 5A, B). The cold spots were mainly located in the northwestern MCI region (northern Minnesota, Wisconsin and Michigan) and northern Missouri, where cropland is not the major land cover type. The hot and cold spots of cropland richness showed different spatial patterns from the cropland percentage (Fig. 5C, D). The hot spots with a high number of crops planted in the county were in North Dakota, Minnesota and Wisconsin. The cold spots with a low number of crop types were mainly located in Iowa, eastern Nebraska and northern Missouri. The cropland richness hot/cold spots showed slightly different spatial patterns in 2007 and 2008 (Fig. 5C, D). Cold spots showed less coverage in 2008 than in 2007, while hot spots showed more coverage. The hot/cold spots of the Shannon equitability index showed more scattered results than cropland percentage and cropland richness (Fig. 5E, F). The hot spots were in North Dakota, central Minnesota, Wisconsin and southern Illinois in both 2007 and 2008. More hot spots were shown in southeastern Iowa and fewer hot spots were in North Dakota and Minnesota in 2008. The cold spots were in central Nebraska, northwestern Iowa, central Missouri and parts of Kansas.

The hot/cold spots of NPP CVs, NEP CVs and SOC change STDEVs are shown in Fig. 6. The NPP CVs showed similar patterns in both 2007 and 2008 (Fig. 6A, B). The hot spots were in northern Wisconsin, northern Minnesota and Missouri. The cold spots were in Iowa, parts of Nebraska and northern Illinois. The NEP CVs had a similar hot/cold spots pattern as the NPP CVs, except there were fewer cold spots in Nebraska and Kansas (Fig. 6C, D). The SOC change STDEVs showed more scattered results than NPP and NEP (Fig. 6E, F). The hot spots were in North Dakota, Kansas and along the border between Iowa and Missouri. The cold spots were in parts of Nebraska and Illinois.

The comparison between the cropland percentage hot/cold spots and the uncertainties hot/cold spots showed that in northern Wisconsin, western Michigan and Missouri, the cold spots of cropland percentages corresponded to the hot spots of NPP and NEP CVs, while the hot spots of cropland percentages corresponded to cold spots of CVs in Nebraska and southern Minnesota (Figs.Fig. 55A, B and Fig. 66A–D). Such correlations between cold and hot spots indicated that higher cropland percentages may lead to smaller difference in uncertainties for NPP and NEP. One interesting observation was that the counties in Iowa and Illinois had large cropland percentages but not low CVs.

The cropland richness and Shannon equitability index hot/cold spots showed quite different patterns from the hot/cold spots of the three carbon fluxes uncertainties. These differences may explain

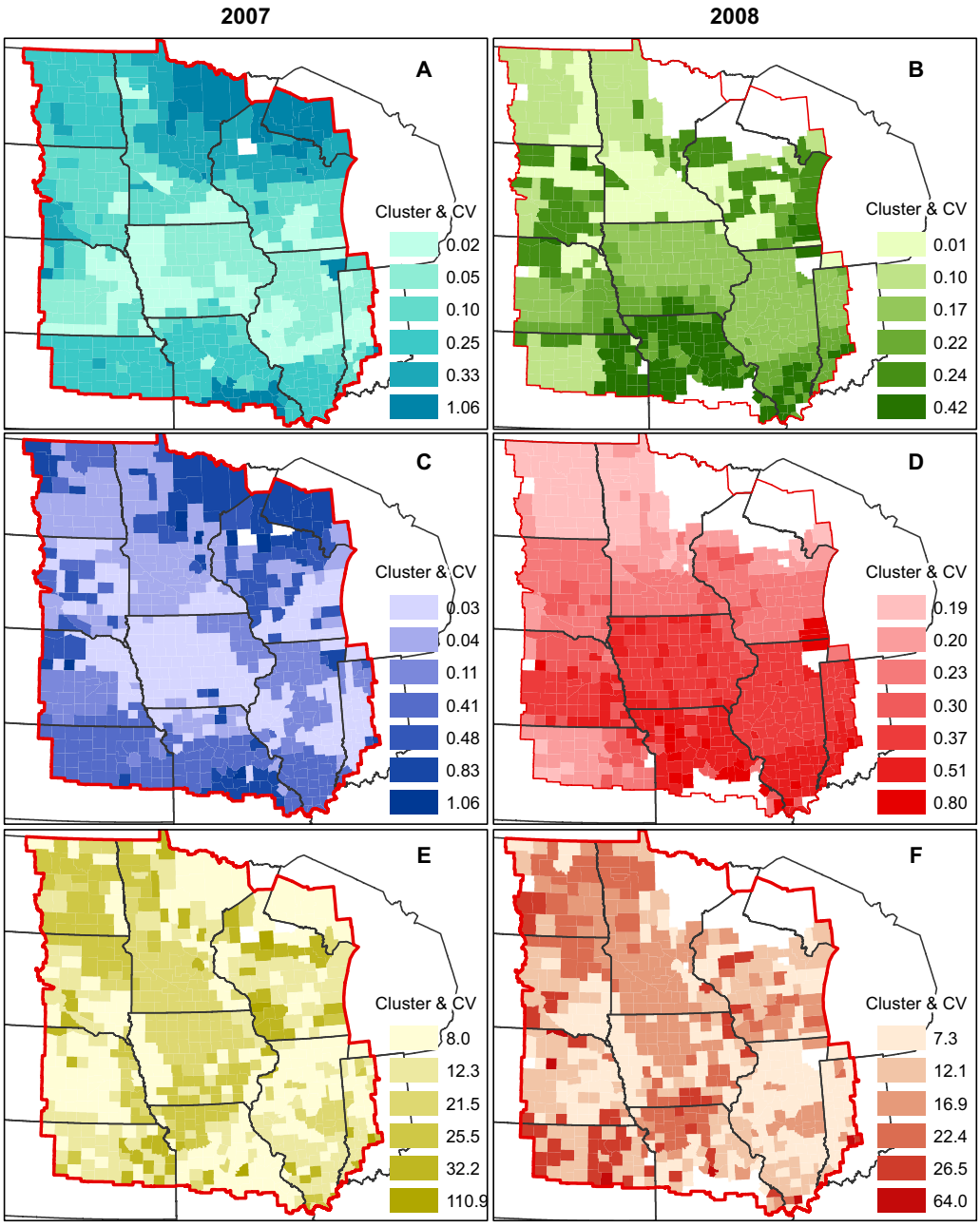

Fig. 4. k-means clustering analysis for NPP in 2007 (A) and 2008(B); NEP in 2007 (C) and 2008 (D); SOC change in 2007 (E) and 2008 (F).

the weak relationships between both characteristics and the uncertainties in the correlation analysis (Table 2).

4. Discussions

The evaluation of process-based models at the regional scale is necessary to assess the credibility of these models for large-scale carbon budget estimates (Zhang et al., 2015). In our study, we focused on analyzing the influence of land cover characteristics on the uncertainties of estimated cropland carbon fluxes.

The land cover characteristics impact the cropland inputs into the models. Each model used its own approach to estimate the cropland area input and resulting uncertainties could be propagated into the model results. The inventory method used the reported harvested cropland area to estimate the carbon fluxes and the harvest area usually is smaller than the planted cropland area. The EPIC model used the representative crop rotations instead of the observed CDL data (Sahajpal et al., 2014; Zhang et al., 2015). This approach reduced the redundancy and computation time but may have introduced some inaccuracies from year to year. For example, corn area in EPIC increased from 26.3 Mha in 2007–31.1 Mha in 2008, while in NASS the reported corn area decreased from 30.1 Mha in 2007–26.9 Mha in 2008. GEMS used a sampling method based on CDL data to simulate the annual crop rotations. This approach could result in large inaccuracies if the cropland area is smaller than the sampling size of the model. If the input data are not consistent between years at the pixel level, this sampling method may also bring uncertainties in the land cover inputs. Though our objective was not trying to evaluate the accuracy of CDL map, we did find some disagreement between years of CDL products, which may be caused by inconsistent classification algorithms applied among years. For example, the annual CDL map showed large amount of

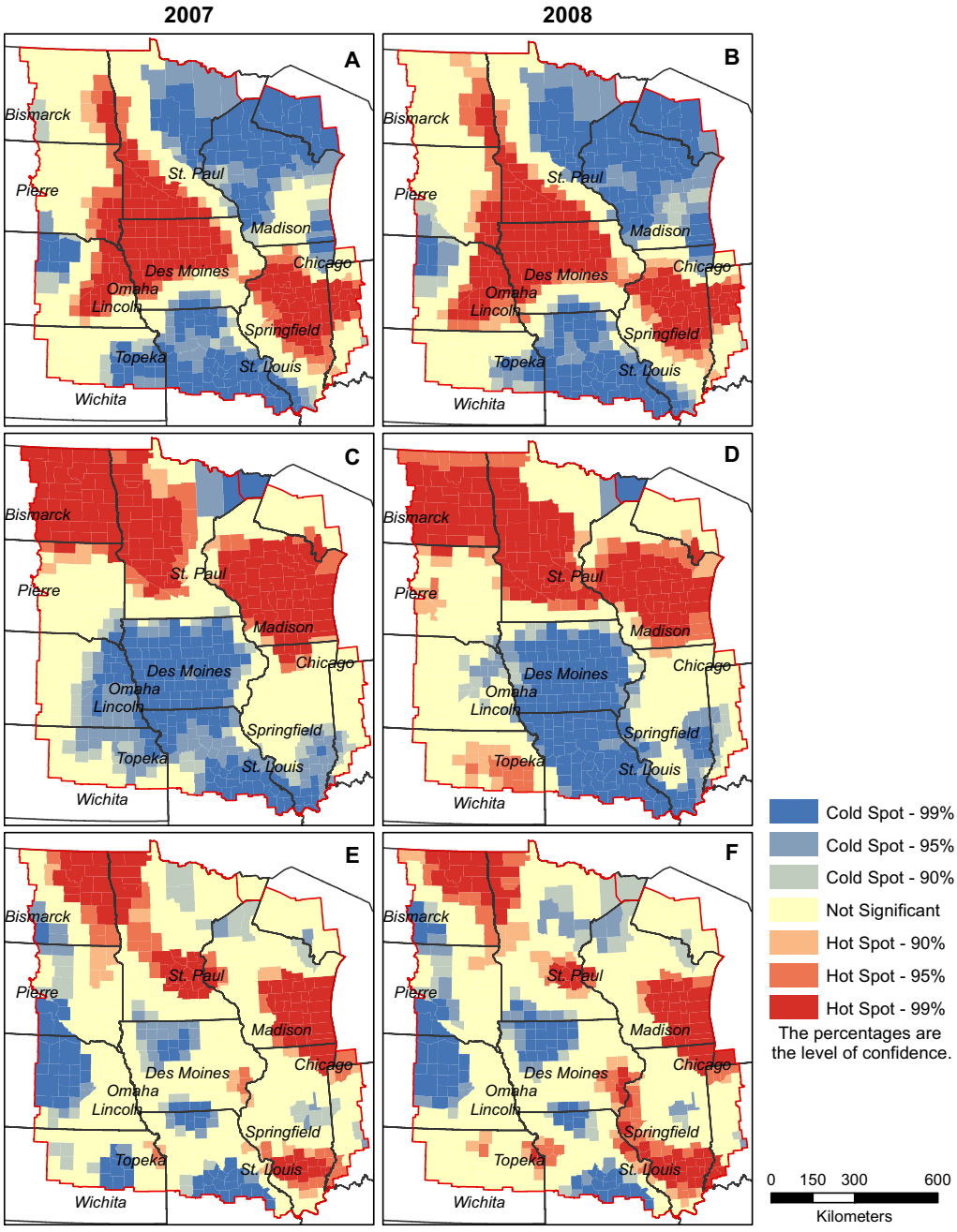

Fig. 5. Hot and cold spots analysis on cropland percentage in 2007 (A) and 2008(B); cropland richness in 2007 (C) and 2008 (D); and Shannon equitability in 2007 (E) and 2008 (F). Note: the percentages (99%, 95%, 90%) represent the areas with statistically significant clusters at alpha-levels of 0.01, 0.05, and 0.1.

grassland in 2006 transferred to forest land in 2007, and a large amount of forest land transferred back to grassland in 2008. In reality, this magnitude of change is unlikely within a single year. We did not find any support for this kind of transition in the literature so it is likely that the change is the result of classification error. Similar classification errors may occur in crop rotations. The land cover inputs differences is large in the county with small cropland percentage due to misclassification and representation of cropland. When cropland area is large, such differences will be small and have less impact on the model uncertainties.

The three models have different classification schemes for the crop types and this may bring uncertainties in model parameterization. The inventory method listed 19 crop types in the factor table to compute the NPP (West et al., 2011). EPIC used over 10 crop types and calibrated the model parameters for each crop using Fluxnet data (Zhang et al., 2015). GEMS used a more simplified approach and classified the crops into 6 categories (Li et al., 2014). These differences in representing the crop types may lead to greater uncertainties when there are more crop types in a single county. Both cropland richness and Shannon equitability index showed less significant correlations with the uncertainties of NPP and NEP in this study. Such differences may be caused by other cropland management practices in addition to crop types, such as cropland irrigation. Irrigation generally changes the water availability and plant growth in the cropland as well as the cropland carbon fluxes. Zhang et al. (2015) found that lack of spatial representation of irrigated cropland in CDL data could explain the discrepancies between the EPIC simulation and inventory estimates. Adding such information into the model inputs may reduce the uncertainties between the models.

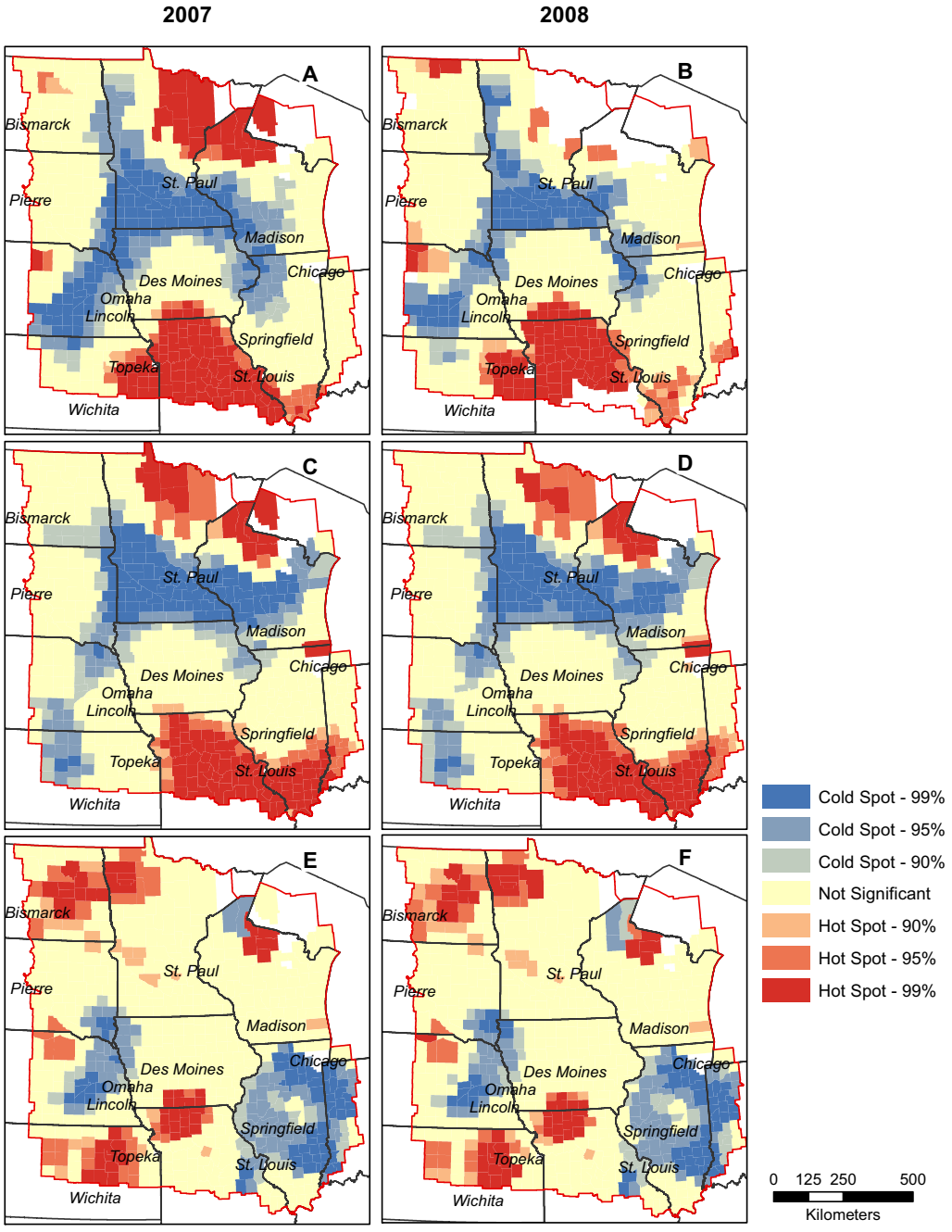

Fig. 6. Hot and cold spots analysis on model uncertainties for NPP CVs in 2007 (A) and 2008(B); NEP CVs in 2007 (C) and 2008 (D); SOC change STDEVs in 2007 (E) and 2008 (F). Note: the percentages (99%, 95%, 90%) represent the areas with statistically significant clusters at alpha-levels of 0.01, 0.05, and 0.1.

Another possible source of uncertainty related to crop types is from the model parameters. In the site-level intercomparison of the NACP models, Schwalm et al. (2010) pointed out that model parameter sets showed clear impact on model skill. The EPIC model used flux tower based measurements to calibrate the model parameters and then applied the same parameters in the MCI region. The GEMS model used the state-level crop inventory data to calibrate the crop growth parameters and used a different set of parameters in each state. When there are more cropland types in a county, the differences in the model parameters may bring higher uncertainties to the model results.

The NACP multi-scale synthesis and terrestrial model intercomparison project pointed to the need for evaluating model performances and better addressing the model differences (Huntzinger et al., 2013). Though our study compared only three model estimates, the data mining and spatial analysis techniques we used in this study could be easily applied to other model ensembles and their driving variables for different regions. Both Moran's *I* analysis and hot/cold spot statistics can help to find the areas with high uncertainties, which leads to identifying the sources of the uncertainties in both model inputs and structures. More research is needed to reduce the uncertainties and improve the model performance. Based on our study, we suggested that using high quality land cover inputs with crop species information is critical to reducing the uncertainties between the models. Integrating other cropland management information such as irrigation may also bring more accurate estimates for cropland fluxes estimates.

5. Conclusions

We used data mining and spatial statistics methods to study the relationships between land cover inputs and the uncertainty of carbon fluxes estimates in the MCI region. Our null hypothesis was proven to be false. The Moran's *I*'s analysis showed the uncertainties have significant positive autocorrelation in neighboring counties in the MCI region. The k-mean clustering analysis showed that the uncertainties in flux estimates are not distributed randomly but are instead formed into multiple clusters. For both NPP and NEP, the uncertainty of the estimates showed significant negative correlations with the cropland percentage in the county. But the uncertainty of the SOC change estimates showed no significant correlation with the cropland percentage. The cropland richness and Shannon equitability index showed a significant negative relationship with the uncertainties of NPP and NEP but not the uncertainties of SOC changes. Our results demonstrated that land cover inputs clearly had impacts on NPP and NEP estimates, but not on the SOC changes. Spatial analysis techniques are powerful tools for revealing the patterns and drivers of uncertainties in regional-scale carbon estimates.

Acknowledgments

We thank Dr. Michael C. Wimberly and two anonymous reviewers for their specific comments and helpful suggestions in improving the manuscript. Dr. Shuguang Liu, with support from the U.S. Geological Survey (USGS) Land Change Science Program, contributed to data analysis and writing of the paper. The fund for EPIC modeling and Dr. Xuesong Zhang is provided by the U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center and the National Aeronautics and Space Administration (NASA) (NNH12AU03I and NNH13ZDA001N). Contributions from Drs. Stephen Ogle and Tristram West are funded by a grant from NASA Terrestrial Ecology Program (NNX08AK08G).

Any use of trade, firm, or product names is for descriptive purpose only and does not imply endorsement by the U.S. Government.

Appendix. Design of the statistic experiment

In this study, we formulated a null hypothesis: the spatial distribution of the uncertainty is random. To test this hypothesis, we computed the spatial autocorrelation index, Moran's *I*, which measures the degree of association of uncertainty (e.g., the CVs of the three method results) between neighboring observations (Getis and Ord, 1992; Getis and Ord, 2010). Therefore, Moran's *I* can detect whether one or more spatial cluster of similar CV values exists in the study area. With a range of values between −1 and 1, Moran's *I* is positive when neighboring counties have more similar CVs, and Moran's *I* is 0 if the spatial distribution of CVs is random.

$$I = \frac{n \times \sum_{i=1}^{n}\sum_{j=1}^{n}[w_{i,j} \times (z_i - \overline{Z}) \times (z_j - \overline{Z})]}{(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}) \times (\sum_{i=1}^{n}(z_i - \overline{Z})^2} \quad (A.1)$$

In the formula, n is the number of counties, zi is the CV value in county i, z_j is the CV value in county j. $\overline{Z}$ is the mean CV of all the counties, and $w_{i,j}$ is the spatial weight. The spatial weight $w_{i,j}$ is computed as the inverse distance between county i and j.

We applied a data mining method (k-means clustering) to identify similar patterns of model estimates. For each county, all the estimates from the three models were treated as one vector, then all the counties were clustered into groups of vectors. The cluster size of each group was determined with the elbow method (Ketchen and Shook, 1996; Thorndike, 1953). First, we calculated the sum of square errors for different number of clusters using three estimates in the county:

$$Var = \sum_{n=1}^{k}\sum_{i=1}^{3}\sqrt{(x_i - c_i)^2} \quad (A.2)$$

In the formula, x_i is the flux value estimated by the method and c_i is the flux value of the center county of the cluster. K is the number of clusters. As the number of cluster increases, the variance decrease. The cluster size is chosen by finding a point when adding another cluster does not reduce much variance (Ketchen and Shook, 1996). The mean vector of each group is computed based on the cluster size. The cluster groups then were drawn to the maps (Fig. 4). The distribution of the clusters showed strong spatial pattern and we made further analysis to find out what land cover characteristics can contribute to such spatial patterns.

First we used linearly regression to compute the relationships between land cover characteristics and the uncertainties. The results were presented in. We also used the hot spot and cold spot statistics (Getis-Ord Gi* statistic) to analysis the spatial patterns of the uncertainties. For each feature i (county in this study), G_i^* will calculate the weighted sum of the variable (e.g., CV of cropland percentage) for the feature's local neighbors then compare the local sum with the global sum for the variable (Getis and Ord, 1992).

$$G_i* = \frac{\sum_{j=1}^{n} w_{i,j}z_j - \overline{Z}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{n\sum_{j=1}^{n} w_{i,j}^2 - (\sum_{j=1}^{n} w_{i,j})^2}{n-1}}} \quad (A.3)$$

where, n is the number of counties, z_j is the CV value of county j. $\overline{Z}$ is the mean of the CV values of all counties, and $w_{i,j}$ is the spatial weight calculated as the inverse distance between county i and j without row standardization. We used this analysis to test the hypothesis that the CVs of the estimates are impacted by the cropland percentage. For cropland percentage (with each county the total area of cropland divided by the total area), if a county is spatially surrounded by counties with high cropland percentage, the county is a hot spot of cropland percentage. Similarly, for the CVs of the three models, if a county is surrounded by counties with low CV values, the county is a cold spot of CVs. By comparing the hot and cold spots of the cropland percentage and the CVs, the spatial correlation between cropland percentage and the model CVs can be visually discovered.

References

Ahl, D.E., Gower, S.T., Mackay, D.S., Burrows, S.N., Norman, J.M., Diak, G.R., 2005. The effects of aggregated land cover data on estimating NPP in northern Wisconsin. Remote Sens. Environ. 97, 1–14.

Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. Geocarto Int. 26, 341–358 (http://dx.doi.org/10.1080/10106049.2011.562309).

Causarano, H.J., Doraiswamy, P.C., McCarty, G.W., Hatfi eld, J.L., Milak, S., Stern, A.J., 2008. EPIC modeling of soil organic carbon sequestration in croplands of Iowa. J. Environ. Qual. 37, 8.

Ciais, P., Wattenbach, M., Vuichard, N., Smith, P., Piao, S.L., Don, A., Luyssaert, S., Janssens, I.A., Bondeau, A., Dechow, R., Leip, A., Smith, P.C., Beer, C., Van Der Werf, G.R., Gervois, S., Van Oost, K., Tomelleri, E., Freibauer, A., Schulze, E.D., Carboeurope Synthesis, T., 2010. The European carbon balance: part 2: croplands. Global Change Biol. 16, 1409–1428, http://dx.doi.org/10.1111/j.1365-2486.2009.02055.x.

Cramer, W., Kicklighter, D.W., Bondeau, A., Iii, B.M., Churkina, G., Nemry, B., Ruimy, A., Schloss, A.L., intercomparison, T.p.o.t.P.N.m., 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results.

Global Change Biol. 5, 1–15, http://dx.doi.org/10.1046/j.1365-2486.1999.00009.x.
Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. Geogr. Anal. 24, 189–206.
Getis, A., Ord, J.K., 2010. The analysis of spatial association by use of distance statistics. In: Anselin, L., Rey, S.J. (Eds.), Perspectives on Spatial Data Analysis. Springer, Berlin Heidelberg, pp. 127–145.
Hayes, D.J., Turner, D.P., Stinson, G., McGuire, A.D., Wei, Y., West, T.O., Heath, L.S., de Jong, B., McConkey, B.G., Birdsey, R.A., Kurz, W.A., Jacobson, A.R., Huntzinger, D.N., Pan, Y., Post, W.M., Cook, R.B., 2012. Reconciling estimates of the contemporary North American carbon balance among terrestrial biosphere models, atmospheric inversions, and a new approach for estimating net ecosystem exchange from inventory-based data. Global Change Biol. 18, 1282–1299, http://dx.doi.org/10.1111/j.1365-2486.2011.02627.x.
Huntzinger, D.N., Post, W.M., Wei, Y., Michalak, A.M., West, T.O., Jacobson, A.R., Baker, I.T., Chen, J.M., Davis, K.J., Hayes, D.J., Hoffman, F.M., Jain, A.K., Liu, S., McGuire, A.D., Neilson, R.P., Potter, C., Poulter, B., Price, D., Raczka, B.M., Tian, H.Q., Thornton, P., Tomelleri, E., Viovy, N., Xiao, J., Yuan, W., Zeng, N., Zhao, M., Cook, R., 2012. North American carbon program (NACP) regional interim synthesis: terrestrial biospheric model intercomparison. Ecol. Modell. 232, 144–157, http://dx.doi.org/10.1016/j.ecolmodel.2012.02.004.
Huntzinger, D.N., Schwalm, C., Michalak, A.M., Schaefer, K., King, A.W., Wei, Y., Jacobson, A., Liu, S., Cook, R.B., Post, W.M., Berthier, G., Hayes, D., Huang, M., Ito, A., Lei, H., Lu, C., Mao, J., Peng, C.H., Peng, S., Poulter, B., Riccuito, D., Shi, X., Tian, H., Wang, W., Zeng, N., Zhao, F., Zhu, Q., 2013. The North American carbon program multi-Scale synthesis and terrestrial model intercomparison project –Part 1: overview and experimental design. Geosci. Model Dev. 6, 2121–2133.
Ito, A., 2011. A historical meta-analysis of global terrestrial net primary productivity: are estimates converging? Global Change Biol. 17, 3161–3175, http://dx.doi.org/10.1111/j.1365-2486.2011.02450.x.
Ketchen, D.J., Shook, C.L., 1996. The application of cluster analysis in strategic management research: an analysis and critique. Strateg. Manage. J. 17, 441–458.
Li, Z., Liu, S., Tan, Z., Bliss, N.B., Young, C.J., West, T.O., Ogle, S.M., 2014. Comparing cropland net primary production estimates from inventory, a satellite-based model, and a process-based model in the Midwest of the United States. Ecol. Modell. 277, 1–12.
Liu, S., Bliss, N., Sundquist, E., Huntington, T.G., 2003. Modeling carbon dynamics in vegetation and soil under the impact of soil erosion and deposition. Global Biogeochem. Cycles 17 (1074), http://dx.doi.org/10.1029/2002gb00201010.
Liu, S., Loveland, T.R., Kurtz, R., 2004. Contemporary carbon dynamics in terrestrial ecosystems in the southeastern plains of the United States. Environ. Manage. 33, S442–S456.
Liu, S., 2009. Quantifying the spatial details of carbon sequestration potential and performance. In: McPherson, B., Sundquist, E. (Eds.), Carbon Sequestration and Its Role in the Global Carbon Cycle. American Geophysical Union, Washington DC, pp. 117–128.
Michalak, A. M., R. B. Jackson, G., Marland, C. L., Sabine and the Carbon Cycle Science Working Group (2011). A U.S. carbon cycle science plan.
National Aeronautics and Space Administration (NASA), 2014. North American Land Data Assimilation System project phase 2 NASA GSFC Hydrological Sciences Laboratory (HSL) and Goddard Earth Sciences Data and Information Services Center (GES DISC) ldas.gsfc.nasa.gov/nldas.
National Agriculture Statistics Service (NASS), 2013. Agricultural Statistics Data Base. U.S. Department of Agriculture, Washington, D.C., USA http://www.nass.usda.gov.
Natural Resources Conservation Service (NRCS), 1994. U.S. General Soil Map (STATSGO). Soil Survey Staff, United States Department of Agriculture (Web Soil Survey, Available online at http://websoilsurvey.nrcs.usda.gov/).
Ogle, S.M., Jay Breidt, F., Eve, M.D., Paustian, K., 2003. Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997.
Ogle, S.M., Davis, K., Andrews, A., Gurney, K., West, T.O., Cooke, R.B., Parkin, T., Morisette, J., Verma, S., Wofsy, S., 2006. Science Plan: Mid-Continent Intensive Campaign of the North American Carbon Program. Greenbelt, MD.
Parameter-elevation Regressions on Independent Slopes Model (PRISM), 2004. PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu created Feb 2004.
Sahajpal, R., Zhang, X., Izaurralde, R.C., Gelfand, I., Hurtt, G.C., 2014. Identifying representative crop rotation patterns and grassland loss in the US Western Corn Belt. Comput. Electron. Agric. 108, 173–182.
Schaefer, K., Schwalm, C.R., Williams, C., Arain, M.A., Barr, A., Chen, J.M., Davis, K.J., Dimitrov, D., Hilton, T.W., Hollinger, D.Y., Humphreys, E., Poulter, B., Raczka, B.M., Richardson, A.D., Sahoo, A., Thornton, P., Vargas, R., Verbeeck, H., Anderson, R., Baker, I., Black, T.A., Bolstad, P., Chen, J., Curtis, P.S., Desai, A.R., Dietze, M., Dragoni, D., Gough, C., Grant, R.F., Gu, L., Jain, A., Kucharik, C., Law, B., Liu, S., Lokipitiya, E., Margolis, H.A., Matamala, R., McCaughey, J.H., Monson, R., Munger, J.W., Oechel, W., Peng, C., Price, D.T., Ricciuto, D., Riley, W.J., Roulet, N., Tian, H., Tonitto, C., Torn, M., Weng, E., Zhou, X., 2012. A model-data comparison of gross primary productivity: results from the North American Carbon Program site synthesis. J. Geophys. Res. Biogeosci. 117, G03010, http://dx.doi.org/10.1029/2012jg001960.
Schuh, A.E., Lauvaux, T., West, T.O., Denning, A.S., Davis, K.J., Miles, N., Richardson, S., Uliasz, M., Lokupitiya, E., Cooley, D., Andrews, A., Ogle, S., 2013. Evaluating atmospheric CO2 inversions at multiple scales over a highly inventoried agricultural landscape. Global Change Biol. 19, 1424–1439, http://dx.doi.org/10.1111/gcb.12141.
Schwalm, C.R., Williams, C.A., Schaefer, K., Anderson, R., Arain, M.A., Baker, I., Barr, A., Black, T.A., Chen, G., Chen, J.M., Ciais, P., Davis, K.J., Desai, A., Dietze, M., Dragoni, D., Fischer, M.L., Flanagan, L.B., Grant, R., Gu, L., Hollinger, D., Izaurralde, R.C., Kucharik, C., Lafleur, P., Law, B.E., Li, L., Li, Z., Liu, S., Lokupitiya, E., Luo, Y., Ma, S., Margolis, H., Matamala, R., McCaughey, H., Monson, R.K., Oechel, W.C., Peng, C., Poulter, B., Price, D.T., Riciutto, D.M., Riley, W., Sahoo, A.K., Sprintsin, M., Sun, J., Tian, H., Tonitto, C., Verbeeck, H., Verma, S.B., 2010. A model-data intercomparison of CO_2 exchange across North America: results from the north american carbon program site synthesis. J. Geophys. Res. Biogeosci. 115, G00H05, http://dx.doi.org/10.1029/2009JG0 01229.
Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A., Randall, D.A., 1996. A revised land surface parameterization (SiB2) for atmospheric GCMS. part II: the generation of global fields of terrestrial biophysical parameters from satellite data. J. Clim. 9, 706–737.
Thorndike, R.L., 1953. Who belongs in the family? Psychometrika 18, 267–276.
West, T.O., Brandt, C.C., Marland, G., De La Torre Ugarte, D.G., Larson, J., Hellwinckel, C.M., Wilson, B., Tyler, D.G., Nelson, R.G., 2008. Estimating regional changes in soil carbon with high spatial resolution. Soil Sci. Soc. Am. J. 72, 285–294.
West, T.O., Brandt, C.C., Baskaran, L.M., Hellwinckel, C.M., Mueller, R., Bernacchi, C.J., Bandaru, V., Yang, B., Wilson, B.S., Marland, G., Nelson, R.G., Ugarte, D.G.D.L.T., Post, W.M., 2010. Cropland carbon fluxes in the United States: increasing geospatial resolution of inventory-based carbon accounting. Ecol. Appl. 20, 1074–1086, http://dx.doi.org/10.1890/08-2352.1.
West, T.O., Bandaru, V., Brandt, C.C., Schuh, A.E., Ogle, S.M., 2011. Regional uptake and release of crop carbon in the United States. Biogeosciences 8, 2037–2046, http://dx.doi.org/10.5194/bg-8-2037-2011.
Zhang, X., Izaurralde, R.C., Manowitz, D.H., Sahajpal, R., West, T.O., Thomson, A.M., Xu, M., Zhao, K., LeDuc, S.D., Williams, J.R., 2015. Regional scale cropland carbon budgets: evaluating a geospatial agricultural modeling system using inventory data. Environ. Modell. Software 63, 199–216, http://dx.doi.org/10.1016/j.envsoft.2014.10.005.
Zhang, X., Sahajpal, R., Manowitz, D.H., Zhao, K., LeDuc, S.D., Xu, M., Xiong, W., Zhang, A., Izaurralde, R.C., Thomson, A.M., West, T.O., Post, W.M., 2014. Multi-scale geospatial agroecosystem modeling: a case study on the influence of soil data resolution on carbon budget estimates. Sci. Total Environ. 479–480, 138–150, http://dx.doi.org/10.1016/j.scitotenv.2014.01.099.