Agricultural and Forest Meteorology 256–257 (2018) 315–333

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Comparing crop growth and carbon budgets simulated across AmeriFlux agricultural sites using the Community Land Model (CLM)

Ming Chen[a,b,*], Tim J. Griffis[a,*], John M. Baker[a,c], Jeffrey D. Wood[d], Tilden Meyers[e], Andrew Suyker[f]

[a] Department of Soil, Water and Climate, University of Minnesota, Saint Paul, MN, USA
[b] Department of Forest Resources, University of Minnesota, Saint Paul, MN, USA
[c] USDA-ARS, USA
[d] School of Natural Resources, University of Missouri, Columbia, MO, USA
[e] Atmospheric Turbulence and Diffusion Division, NOAA/ARL, Oak Ridge, TN, USA
[f] School of Natural Resources, University of Nebraska-Lincoln, Lincoln, NE, USA

ARTICLE INFO

Keywords:
Crop model
Crop phenology
AmeriFlux
Carbon budget
Model sensitivity test
Multi-site-year calibration

ABSTRACT

Improvement of process-based crop models is needed to achieve high fidelity forecasts of regional energy, water, and carbon exchanges. However, most state-of-the-art Land Surface Models (LSMs) assessed in the fifth phase of the Coupled Model Inter-comparison project (CMIP5) simulated crops as unmanaged C_3 or C_4 grasses. This study evaluated the crop-enabled version of one of the most widely used LSMs, the Community Land Model (CLM4-Crop), for simulating corn and soybean agro-ecosystems at relatively long-time scales (up to 11 years) using 54 site-years of data. We found that CLM4-Crop had a biased phenology during the early growing season and that carbon emissions from corn and soybean were underestimated. The model adopts universal physiological parameters for all crop types neglecting the fact that different crops have different specific leaf area, leaf nitrogen content and vcmax25, etc. As a result, model performance varied considerably according to crop type. Overall, the energy and carbon exchange of corn systems were better simulated than soybean systems. Long-term simulations at multiple sites showed that gross primary production (GPP) was consistently over-estimated at soybean sites leading to very large short and long-term biases. A modified model, CLM4-CropM', with optimized phenology and calibrated crop physiological parameters yielded significantly better simulations of gross primary production (GPP), ecosystem respiration (ER) and leaf area index (LAI) at both short (hourly) and long-term (annual to decadal) timescales for both soybean and corn.

1. Introduction

Land surface models (LSMs) serve as important tools for studying the interactions between the atmosphere and ecosystems, understanding biophysical feedback processes, and predicting future climate. Most state-of-the-art LSMs assessed in the fifth phase of the Coupled Model Inter-comparison project (CMIP5), however, did not include process-based crop models with comprehensive physiology and phenology. The first attempt to incorporate explicit crop simulations into a land surface model was made early this century by Tsvetsinskaya et al. (2001), in which the Biosphere-Atmosphere Transfer Scheme (BATS) was modified to include plant growth functions for corn. The modified model showed improved simulation of seasonal phenology and significant changes in surface sensible (20%–35%) and latent (30%–45%) heat fluxes, especially during dry years in central and eastern Nebraska and eastern Kansas. Since then, the development and evaluation of land surface models (LSMs) with prognostic agricultural schemes became an active and important topic of inquiry (Tsvetsinskaya et al., 2001; Kucharik, 2003; Bondeau et al., 2007; Osborne et al., 2007; Stehfest et al., 2007; Gervois et al., 2008; Lokupitiya et al., 2009; Levis et al., 2012; Song et al., 2013; Wu et al., 2015; Chen et al., 2015).

Kucharik and Brye (2003) incorporated phenology, carbon allocation, and a corn-specific parameterization into the Integrated Biosphere Simulator (IBIS). The model was further developed to include other crops such as soybean, winter and spring wheat and was renamed Agro-IBIS. Its ability to simulate crop yields, water and energy balances, and impacts of climate change on agro-ecosystems in the United States were investigated at multiple sites and scales (Donner and Kucharik, 2003; Kucharik and Twine, 2007; Kucharik, 2003; Twine et al., 2013, 2004; Twine and Kucharik, 2009; Webler et al., 2012; Xu et al., 2016).

* Corresponding authors at: Department of Soil, Water and Climate, University of Minnesota, Saint Paul, MN, USA.
E-mail addresses: chen1718@umn.edu (M. Chen), tgriffis@umn.edu (T.J. Griffis).

https://doi.org/10.1016/j.agrformet.2018.03.012
Received 18 June 2017; Received in revised form 7 March 2018; Accepted 16 March 2018
Available online 05 April 2018
0168-1923/ © 2018 Elsevier B.V. All rights reserved.

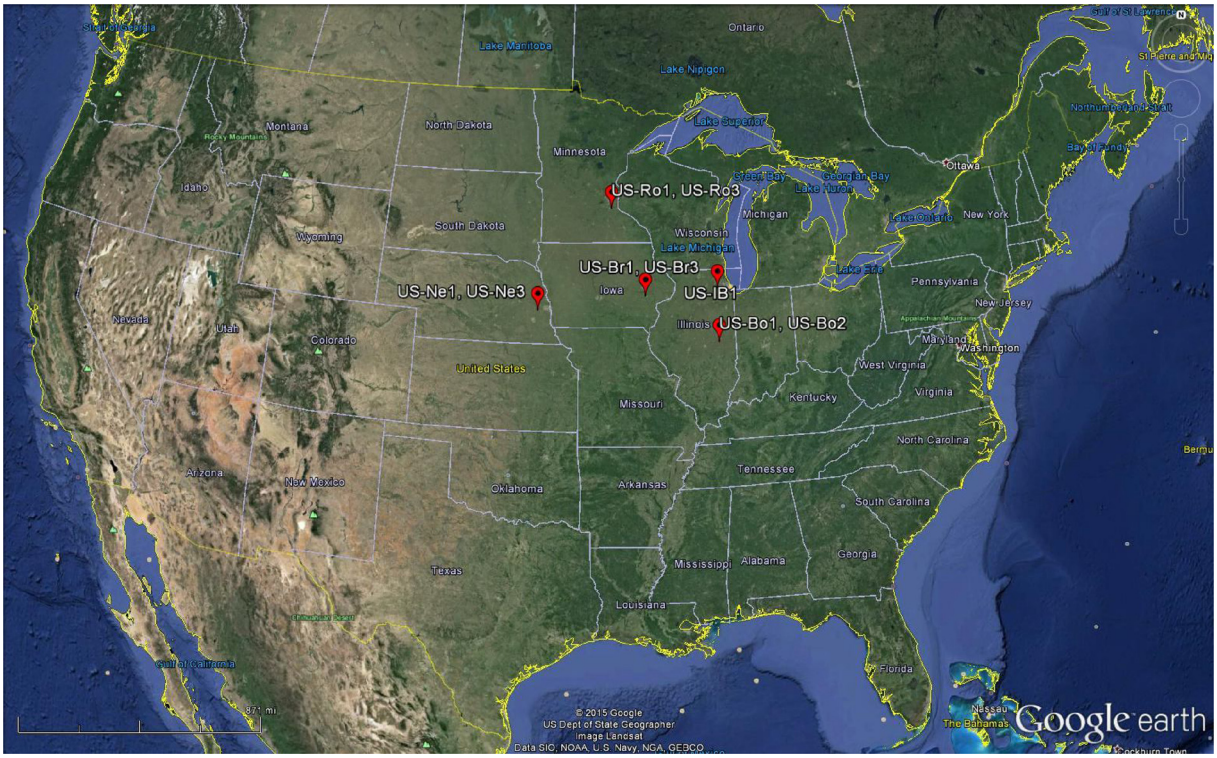

Fig. 1. Nine AmeriFlux crop sites within the US Corn Belt.

Recently, the Community Land Model (CLM), which is the land component of the widely used Community Earth System Model (CESM) (Oleson et al., 2010) adopted the prognostic crop scheme from Agro-IBIS into CLM (called CLM4-Crop). CLM4-Crop has the potential to provide better spatial and temporal information on climate-crop interactions and improved weather and climate prediction (Levis et al., 2012). For example, when CLM4-Crop was coupled to an atmospheric model (i.e. the Community Atmosphere Model, version 4 [CAM4.0] (Neale et al., 2010)), forecasted precipitation during the peak growing season was significantly improved over Midwestern North America. However, recent efforts have examined the phenology and the seasonality of net ecosystem CO_2 exchange (NEE) related to cropping systems over multiple years at three AmeriFlux sites and demonstrated high sensitivity of energy balance and carbon simulations to biases in phenology (Levis et al., 2012; Chen et al., 2015). Further improvements in CLM4-Crop to advance its carbon forecast capability are therefore urgently needed.

While LSMs that incorporate prognostic crop schemes are important for achieving realistic simulations of weather and climate and increasing our capacity to develop sound policies regarding the impacts of climate change on agricultural systems and the potential impacts of land management on climate (Levis et al., 2012), few studies have quantified the accuracy of the simulations at multiple sites and over relatively long time scales (> 10 years). Factors like plant biophysical properties, soil properties and climate vary considerably among sites (Fulton et al., 1996; Mzuku et al., 2005; Loescher et al., 2014; Mourtzinis et al., 2015). An important consideration, therefore, is the evaluation of these models across a broad range of sites to determine if the parameterizations are sufficiently general, while also providing acceptable performance across space and time. Evaluation of models over relatively long timescales can be used to help identify model deficiencies. For example, the ability of models to capture long-term variations in plant phenology and energy and carbon fluxes remains an important challenge (Piao et al., 2013; Richardson et al., 2007; Schwalm et al., 2010; Wang et al., 2012). Depending on the inter-annual variation of climate, it is suggested that 10–20 years of meteorological forcing data are generally necessary for reliable estimates of mean yield potential of crops and its inter-annual variability (Van Wart et al., 2015). Flux observations from networks such as AmeriFlux are just now providing long-term records approaching these important timescales and provide an opportunity for decadal-scale model assessment.

Here, we examine the performance of two versions of CLM4-Crop (CLM4-Crop and CLM4-CropM) at 9 agricultural sites with a focus on the ability of the model to capture seasonal and inter-annual variations in leaf area index (LAI), NEE, ER, and GPP. We use a total of 54 site-years of data to diagnose some of the key model deficiencies in CLM4-Crop and address the following questions:

1 Is the new phenology scheme able to simulate inter-annual variations in early growing season crop phenology at multiple sites across a climate gradient?
2 Are biases in simulated phenology across sites random, or are there systematic deficiencies that can be addressed with model calibration?
3 How well do the models perform under different climate conditions and does the new phenology scheme improve simulations for certain climate conditions?
4 Do the models adequately capture the long-term (> 10 years) dynamics of NEE for cropland?

2. Methods

2.1. Meteorology and biological data

The models were evaluated at nine AmeriFlux sites in the US Corn Belt located within latitude and longitude ranges of 40–45 °N and 88–97 °W, respectively (Fig. 1). These sites represent cropland systems in the northern (US-Ro1, US-Ro3), western (US-Ne1, US-Ne3), southern (US-Bo1, US-Bo2) and central (US-IB1, US-Br1, US-Br3) US Corn Belt. Climate and cropping information for these sites are provided in Table 1.

Table 1
AmeriFlux crop sites.

ID	Site Name	Site Description	State	Latitude	Longitude	Elevation	Mean Annual Temp (°C)	Mean Annual Precip. (mm)	Years	Corn phase	Soybean phase
1	US-Ro1	Rosemount	MN	44.7143	−93.0898	260 m	6.86 °C	806 mm	2007–2010	odd years	even years
2	US-Ro3	Rosemount alternative management	MN	44.7217	−93.0893	260 m	6.86 °C	806 mm	2007–2010	2007, 2008, 2010	2009
3	US-Bo1	Bondville	IL	40.0062	−88.2904	219 m	11.02 °C	991 mm	1997-2007	odd years	even years
4	US-Bo2	Bondville Companion Site	IL	40.0061	−88.2918	219 m	11.02 °C	991 mm	2005–2006	even years	odd years
5	US-IB1	Fermi Agricultural	IL	41.8593	−88.2227	225 m	9.18 °C	929 mm	2006–2008	even years	odd years
6	US-Br1	Brooks Field Site 10	IA	41.9749	−93.6914	275 m	8.95 °C	842 mm	2007–2010	odd years	even years
7	US-Br3	Brooks Field Site 11	IA	41.9747	−93.6936	314 m	8.95 °C	842 mm	2007–2010	even years	odd years
8	US-Ne1	Mead Irrigated	NE	41.1650	−96.4766	361 m	10.07 °C	784 mm	2002–2012	continuous	NA
9	US-Ne3	Mead Rainfed	NE	41.1797	−96.4396	363 m	10.11 °C	784 mm	2002–2012	odd years	even years

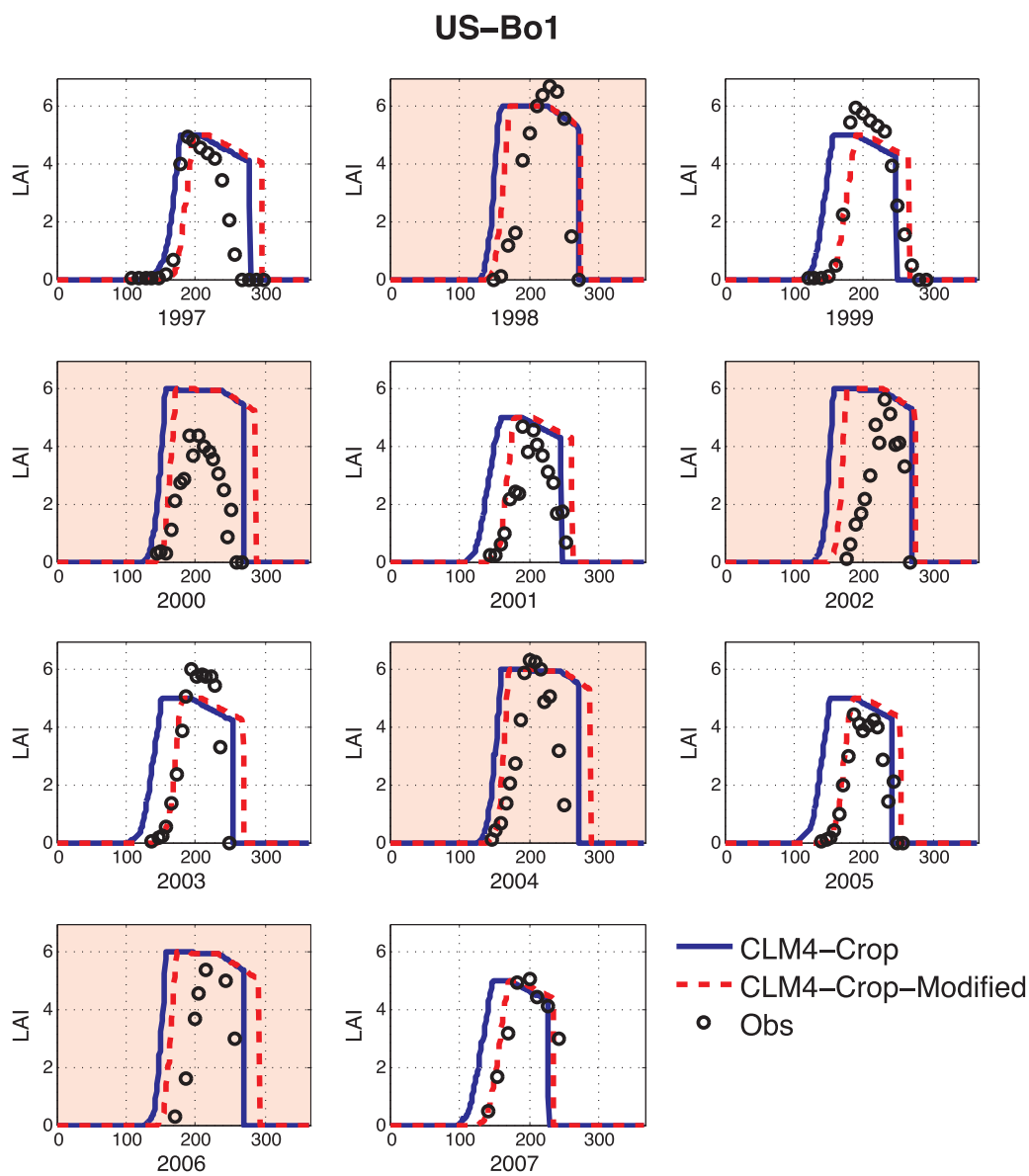

Fig. 2. Evaluation of the simulated phenology for corn (white background) and soybean (red background) at AmeriFlux site Bondville (US-Bo1). The blue line represents default CLM4-Crop simulation and the red dashed line represents CLM4-CropM with modified phenology according to Eq. (2) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Carbon and energy fluxes measured using the eddy covariance (EC) technique as well as meteorological data were obtained from AmeriFlux level 2 gap-filled data products for the longest available time series at each site. The biological data included LAI and harvested grain carbon (Level 1 AmeriFlux data). The US-Bo1, US-Ne1 and US-Ne3 sites have the longest available time series of LAI and were thus chosen for the LAI evaluation. The yield/harvest grain data were converted to gC m^{-2} y^{-1} to compare with model simulations, assuming grain carbon content is

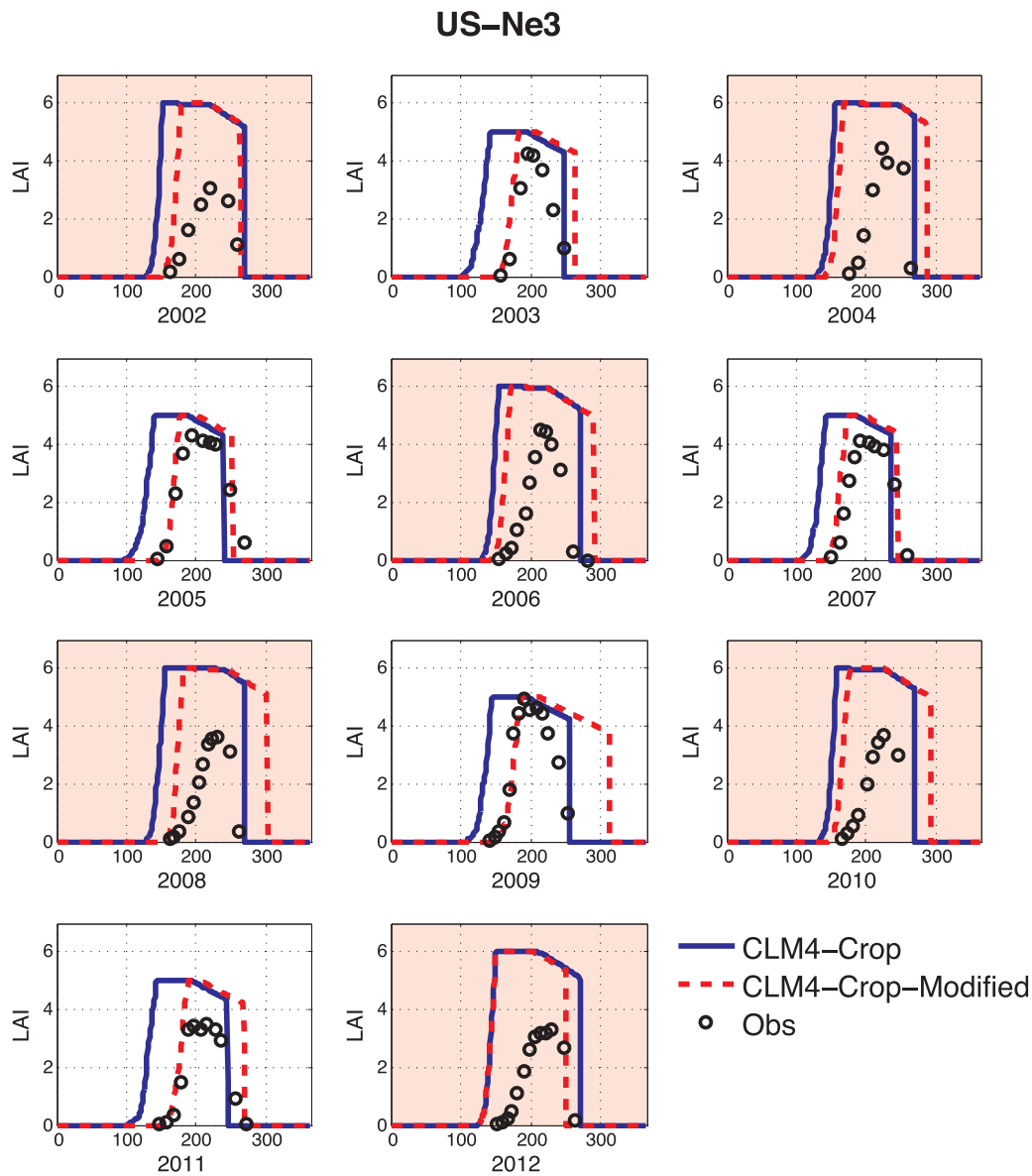

Fig. 3. Evaluation of the simulated phenology for corn (white background) and soybean (red background) at a rainfed AmeriFlux site Mead (US-Ne3). The blue line represents default CLM4-Crop simulation and the red dashed line represents CLM4-CropM with modified phenology according to Eq. (2) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

45% for corn and 54% for soybean based on laboratory measured mean values (Baker and Griffis, 2005). In order to close the annual carbon budget for cropping systems we have estimated the Net Biome Productivity (NBP) by adding the harvested grain carbon (a carbon loss) to the EC measured annual NEE:

$$\text{Carbon budget} = -\text{NBP} = \text{NEE} + C_{\text{grain}} \quad (1)$$

where C_{grain} is the harvested grain carbon calculated from the yield data at the sites (Baker and Griffis, 2005; Chen et al., 2015; Suyker and Verma, 2010). In Eq. (1), a positive carbon budget is defined as a net transfer of carbon from the land to the atmosphere.

2.2. CLM4-Crop and CLM4-CropM

CLM is the land surface component of CESM (Bonan and Oleson, 2002; Dai et al., 2003; Dickinson and Oleson, 2006; Oleson et al., 2010, 2008; Zeng et al., 2002). CLM simulates biophysical and biochemical processes between soil, plant, and the atmosphere. In this study, the fourth version of CLM with the crop scheme activated (CLM4-Crop) was used (Levis et al., 2012).

The crop algorithms in CLM4-Crop originated from the Agro-IBIS model (Levis et al., 2012). The crop types simulated in CLM4-crop include corn, soybean, and temperate cereals. Here we evaluate the simulation of corn and soybean since they represent the dominant crops in the United States Corn Belt. CLM4-Crop simulates three phenological phases: 1) planting to emergence, 2) leaf emergence to the beginning of grain fill, and 3) from the beginning of grain fill until harvest.

CLM4-Crop is known to overestimate LAI during early growing season (Chen et al., 2015; Levis et al., 2012). Planting occurs when three thresholds are met in CLM4-Crop: a 20-year averaged growing degree day (GDD; base temperature 8 °C) threshold, a threshold of 10-day running mean of air temperature and a threshold of daily minimum air temperature (Levis et al., 2012). Since the 20-year averaged GDD changes very slowly from year to year and the threshold is very low (50 ° days), the actual planting date is determined by the two temperature terms, which are also met very early in the year. Thus, when we compare the phenology simulated by CLM4-Crop to observation, the simulated growing season is always biased early. To improve the early

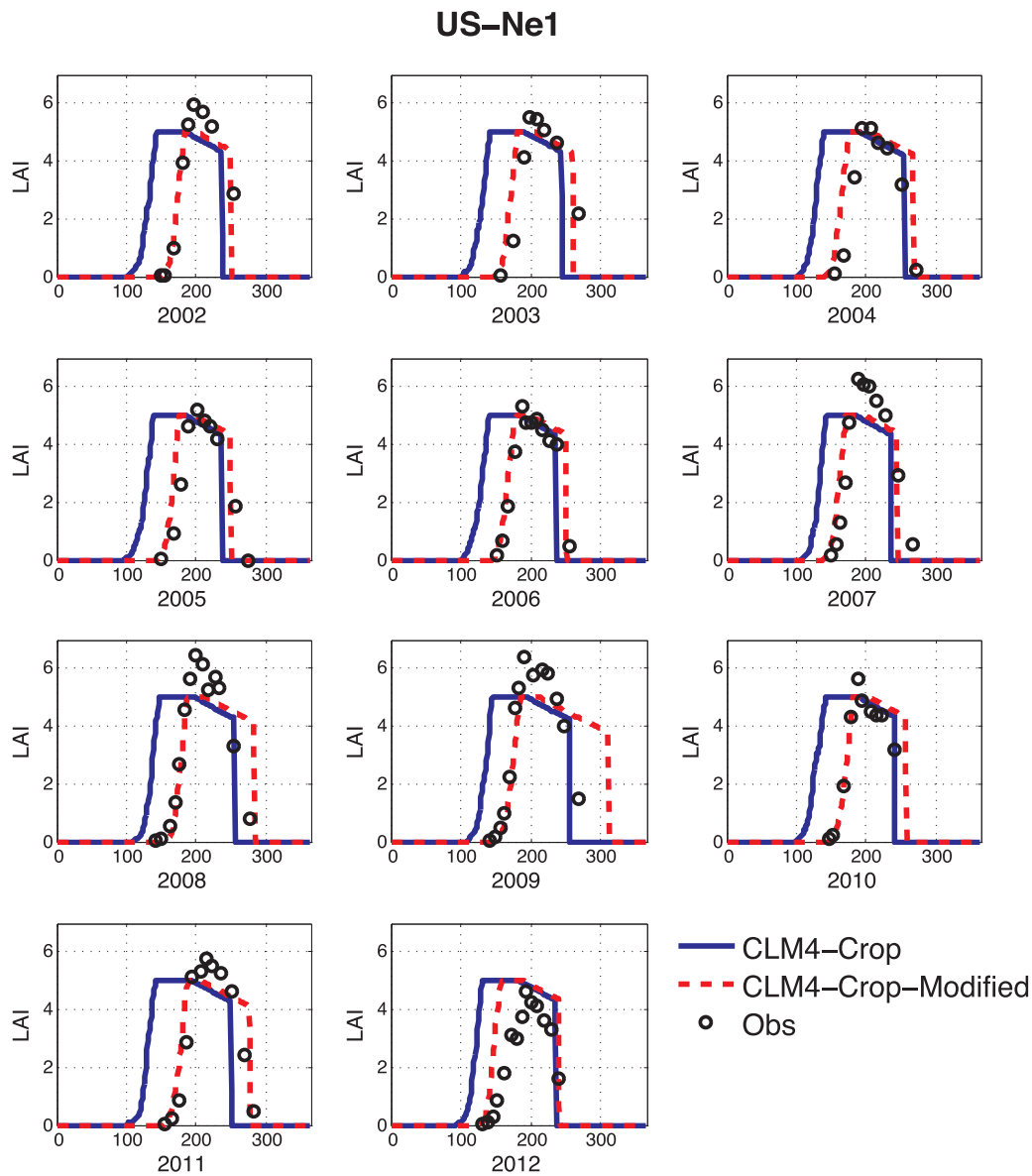

Fig. 4. Evaluation of the simulated phenology for corn at an irrigated AmeriFlux site Mead (US-Ne1). The blue line represents default CLM4-Crop simulation and the red dashed line represents CLM4-CropM with modified phenology according to Eq. (2) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

growing season phenology simulated in CLM4-Crop, we modified the model and simulated the planting date using growing degree time (GDT):

$$GDT = \int (T_{air} - T_{ref}) \cdot dt \ , \ \ when \ T_{air} > T_{ref} \tag{2}$$

where T_{air} is the air temperature at 2 m, T_{ref} is the reference temperature (8 °C), and dt is the model time step. The GDT thresholds are 450 for corn and 400 for soybean as specified in Chen et al. (2015). The modified model is called CLM4-CropM.

2.3. Model parameterization

There are more than 80 plant physiological parameters in CLM4. Here, only some of the key parameters that were determined from previous sensitivity analyses to control carbon sequestration of crop PFTs are discussed (Bilionis et al., 2015; Sargsyan et al., 2014). CLM4 uses a PFT-dependent water stress factor to describe the soil water constraint on the transpiration and photosynthetic rate. This water stress factor is calculated as:

$$\beta_t = \sum_i w_i r_i \tag{3}$$

where w_i is a plant wilting factor for layer i and r_i is the fraction of roots in layer i. The water stress factor β_t in CLM4-Crop was multiplied by 1.25 for increased drought tolerance of soybean (Levis et al., 2012).

To calculate the optical depth of the canopy as well as the fraction of sunlit and shaded leaves in the canopy, a leaf angle distribution factor (χl) is used in the model. Here $\chi l = -1$ indicates vertically distributed leaves; $\chi l = 1$ indicates horizontally distributed leaves and $\chi l = 0$ indicates a random distribution. This parameter ranges from −0.4 to 0.6 in CLM. The default leaf angle distribution factor for corn and soybean was −0.4. Theoretically, a more planophile canopy will reflect more short-wave radiation upward, and thus has less fraction of sunlit leaf. With the two-big-leaf approximation in CLM, the reduced fraction of sunlit leaf will reduce the canopy photosynthesis rate. Further sensitivity tests are discussed in Section 3.4.

The maximum carboxylation rate of rubisco at 25 °C (V_{cmax25}) was set to 100.7 μmol m^{-2} s^{-1} in CLM4-Crop for both corn and soybean (Bonan et al., 2012; Kattge et al., 2009; Levis et al., 2012). This value is

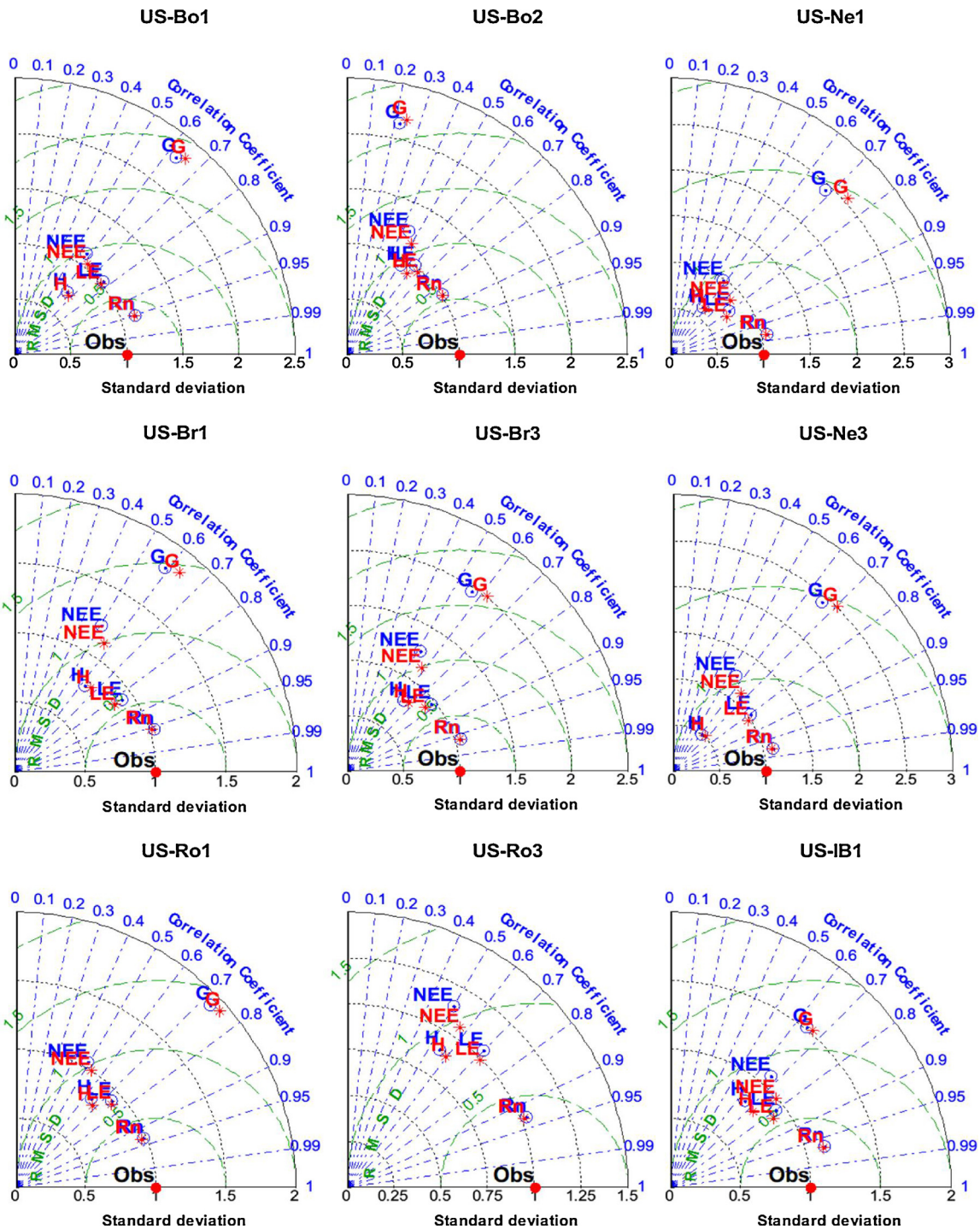

Fig. 5. Cross-site evaluation of the hourly energy fluxes and NEE simulation. Results are shown in Taylor diagrams. The standard deviations are normalized by the standard deviations of the observed values. The blue colored variables are the default CLM4-Crop simulations. The red colored variables are CLM4-CropM simulations (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

the mean V_{cmax25} derived for C_3 crops using a plant trait database (https://www.try-db.org/TryWeb/Home.php). Here we note that the V_{cmax25} value for C_3 crops has a relatively large standard deviation of 36.6 μmol m^{-2} s^{-1}.

2.4. Model simulation and evaluation

The models were forced by hourly meteorological data (solar radiation, air temperature, precipitation, air humidity, air pressure and wind speed) observed at each site (Table 1). Mineral soil texture data were extracted from a global data set (Oleson et al., 2010). To ensure that model soil carbon pools were at steady state, each model was spun up for 1000 years by re-cycling the available site meteorological data (Thornton and Rosenbloom, 2005). Afterwards, the models were forced by the meteorological data at the site for the available years and the hourly output of results was evaluated against measurements.

The model was run in single site mode so that there is one PFT that matches the land cover type in which the flux tower is located (Chen et al., 2015; Stöckli et al., 2008; Yuan and Liang, 2011). The metrics used for evaluating the simulated energy and carbon fluxes includes the

Table 2
Correlation coefficients of the simulated and observed hourly energy and NEE fluxes.

Model	R_n	H	LE	G	NEE
US-Bo1					
CLM4-Crop	0.95	0.64	0.77	0.63	0.58
CLM4-CropM	0.95	0.67	0.77	0.65	0.62
US-Bo2					
CLM4-Crop	0.84	0.51	0.60	0.22	0.45
CLM4-CropM	0.85	0.59	0.65	0.24	0.51
US-Br1					
CLM4-Crop	0.96	0.62	0.83	0.59	0.50
CLM4-CropM	0.95	0.67	0.82	0.63	0.56
US-Br3					
CLM4-Crop	0.96	0.62	0.78	0.56	0.52
CLM4-CropM	0.96	0.67	0.77	0.62	0.58
US-Ne1					
CLM4-Crop	0.98	0.57	0.80	0.68	0.58
CLM4-CropM	0.98	0.60	0.83	0.75	0.74
US-Ne3					
CLM4-Crop	0.98	0.60	0.83	0.75	0.74
CLM4-CropM	0.98	0.67	0.83	0.70	0.66
US-Ro1					
CLM4-Crop	0.93	0.64	0.74	0.72	0.50
CLM4-CropM	0.93	0.68	0.75	0.75	0.53
US-Ro3					
CLM4-Crop	0.93	0.55	0.70	0.45	0.50
CLM4-CropM	0.93	0.59	0.72	0.47	0.57
US-IB1					
CLM4-Crop	0.97	0.65	0.80	0.64	0.66
CLM4-CropM	0.97	0.73	0.83	0.67	0.76
Average					
CLM4-Crop	0.94	0.60	0.76	0.57	0.54
CLM4-CropM	0.94	0.65	0.77	0.61	0.62

All P values ≤ 0.01.

Pearson correlation coefficient, standard deviation and root mean square error. These three statistics were summarized and plotted as Taylor diagrams (Taylor, 2001).

We also used a refined index of agreement (Willmott et al., 2012) to quantify the degree to which observed LAI were captured by the models:

$$d_r = \begin{cases} 1 - \dfrac{\sum_{i=1}^{n} |P_i - O_i|}{c\sum_{i=1}^{n} |O_i - \overline{O}|}, & when\ \sum_{i=1}^{n} |P_i - O_i| \le c\sum_{i=1}^{n} |O_i - \overline{O}| \\ \dfrac{c\sum_{i=1}^{n} |O_i - \overline{O}|}{\sum_{i=1}^{n} |P_i - O_i|} - 1, & when\ \sum_{i=1}^{n} |P_i - O_i| > c\sum_{i=1}^{n} |O_i - \overline{O}| \end{cases} \quad (4)$$

where P_i stands for predictions and O_i is the pair-wise-matched observations. The range of refined index of agreement d_r is from −1 to 1. A d_r of 1 indicates perfect agreement between model and observation, and a d_r of −1 indicates either lack of agreement between the model and observation or insufficient variation in observations to test the model adequately.

3. Results and discussion

3.1. Simulation of phenology across sites

Our assessment of simulated crop phenology was restricted to comparisons with observations from sites with the longest and most complete LAI records (> 10 years): US-Bo1, US-Ne1 and US-Ne3 (Figs. 2–4), representing different geographic regions and climate in the Corn Belt (Table 1). The hydrologic cycle differs markedly among sites, and includes co-located fields where one is irrigated (US-Ne1) and the other rainfed (US-Ne3). Therefore, these sites constitute good boundaries on parameter space with respect to the coupled carbon cycle, hydrologic cycle and the interactions with climate.

The planting date simulated by CLM4-Crop was systematically biased early, with an average bias of −14 ± 3 days for corn and -18 ± 4 days for soybean. The emergence simulated by CLM4-Crop was also biased early, ranging from 30 to 50 days. CLM4-Crop early growing season LAI bias was similar at all sites, but somewhat larger at the two Mead sites (US-Ne1 and US-Ne3), with slightly colder climate condition than the Bondville site (US-Bo1). In a previous global scale study, CLM4-Crop simulated 20-year averaged monthly LAI at the grid cell where US-Bo1 was located (other PFTs coexisted in the grid cell) and it also showed an over-estimated LAI during the first half of the growing season and an over-estimated length of growing season (Levis et al., 2012). Our multi-site-year comparison confirmed that the agricultural system contributed to this bias. In contrast, the GDT method implemented in CLM4-CropM delayed the leaf emergence date and reduced the mean bias of emergence date across sites to −2 ± 3 days.

The simulated inter-annual variability in early growing season phenology based on CLM4-CropM also showed a notable improvement. For instance, for US-Bo1 in some years when CLM4-Crop simulated emergence times were close to the observations, CLM4-CropM did similarly well (e.g., 1997, 2000, 2004). In other years when CLM4-Crop estimated earlier emergence, CLM4-CropM gave better estimations (e.g., 1998–1999, 2001–2003, 2005–2007). This indicates that the CLM4-CropM phenology scheme is more sensitive to early spring temperature variation and has greater skill at capturing the observed inter-annual variability of the onset of the growing season compared to CLM4-Crop.

One point to note is that Agro-IBIS, which used the same mechanism to simulate crop phenology, did not show earlier onset of growing season or overestimated LAI during the early growing season (Twine et al., 2013; Webler et al., 2012). Twine et al. (2013) used a group of soybean specific physiological parameters other than the previous Agro-IBIS parameters which were collected firstly for corn and found

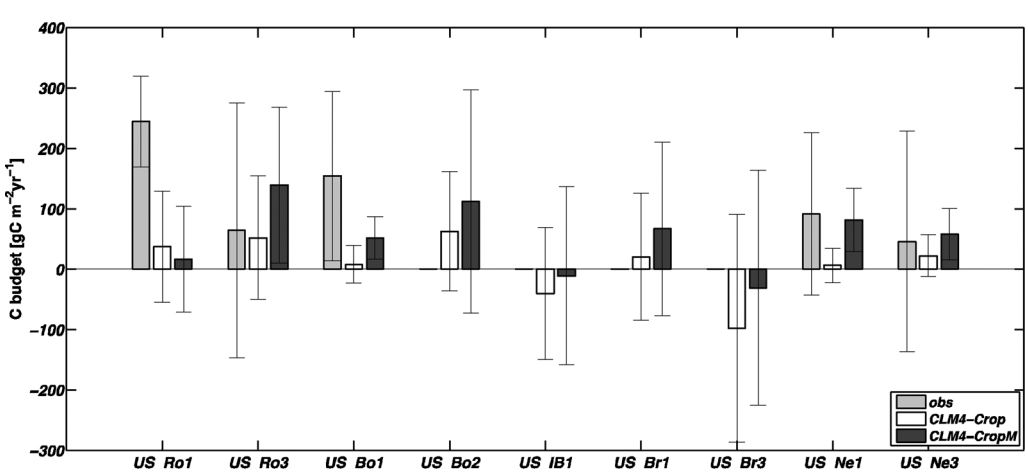

Fig. 6. Cross-site comparison of the GDD and GDT methods simulated annual carbon budgets (-NBP) at the nine AmeriFlux sites. The error bars represent the standard deviation of the annual mean NBP. At the sites US-Ro1, US-Ro3, US-Bo1, US-Ne1 and US-Ne3, yield data are available. Thus at those sites the measured annual carbon budgets were shown to be compred with the modeled values.

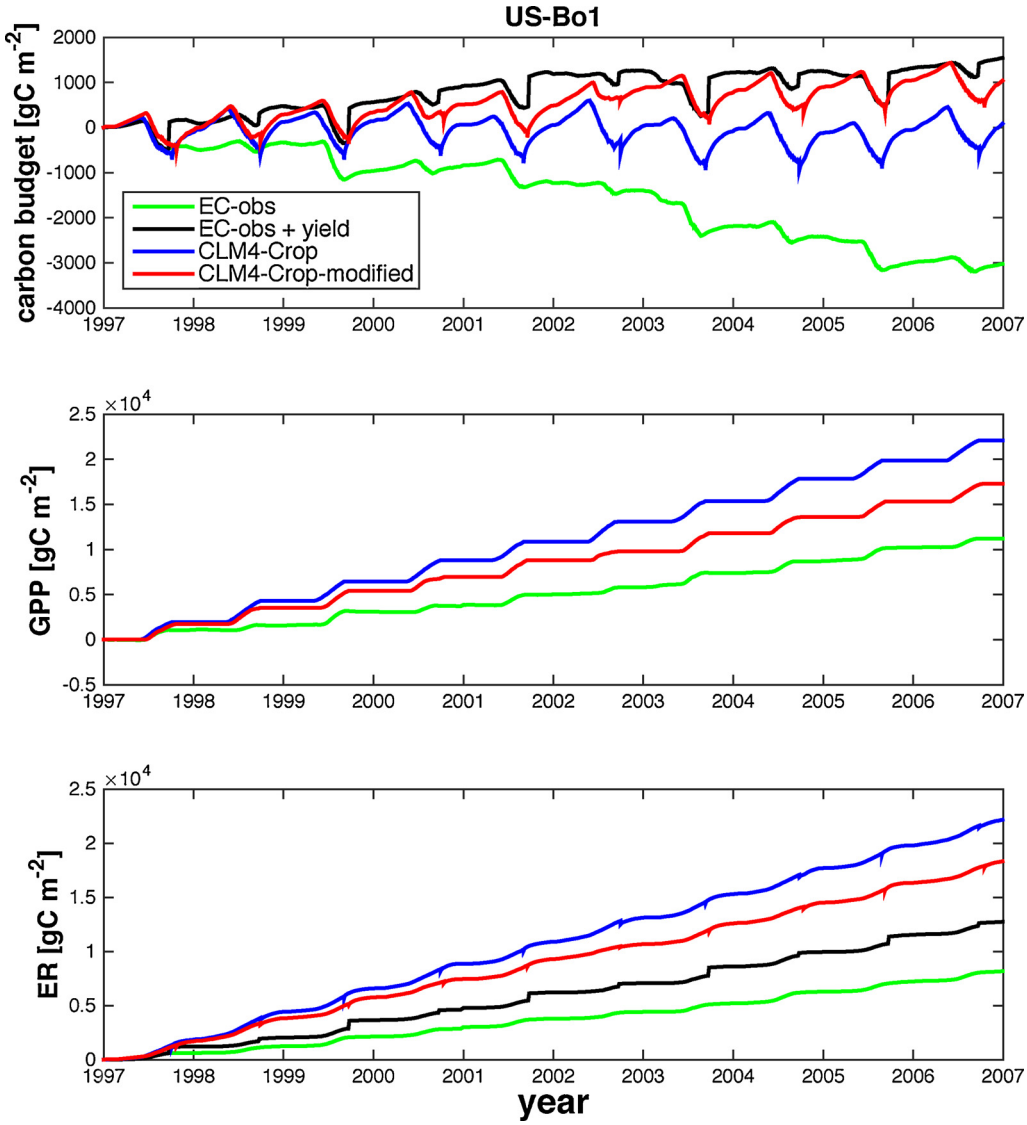

Fig. 7. Long-term carbon budget simulation at site US-Bo1. The green line represents carbon fluxes calculated using the eddy covariance measured carbon fluxes. The black line is the eddy covariance measured carbon fluxes plus annual yield grain carbon from the field. The blue line is the default CLM4-Crop simulation and the red line is CLM4-CropM simulation (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

improvements in phenology at a soybean free air concentration enrichment site (SoyFACE) with the new scheme. However, in their studies the planting dates were prescribed inputs. In future predictions with a changing climate, a prognostic crop scheme is desired.

Soybean displayed lower growth rates compared to corn between leaf emergence and grain filling. However, both models simulated similar growth rates for both crops, which was too high for soybean in the early growth phase, causing LAI to quickly reach the maximum constraint of 6 m^2 m^{-2}. This resulted in an overestimated LAI for soybean during the early growing season for both models, especially at the drier site US-Ne3. Multiple factors contribute to the model estimation of LAI including photosynthetic rate, canopy structure, and allocation of assimilated carbon to leaves. The sensitivity of LAI to the parameters used in these processes was explored further in Section 3.4.

The maximum LAI assumptions in CLM4 of 5 $m^2\,m^{-2}$ for corn and 6 $m^2\,m^{-2}$ for soybean worked well for seven years out of the eleven at US-Bo1 (1997, 2001–2002, 2004–2007). However for the other years, the maximum LAI threshold was either underestimated for corn (1999, 2003) and soybean (1998) or overestimated (2000). The maximum soybean LAI threshold of 6 $m^2\,m^{-2}$ was biased high by about 30% for the water-limited Mead site US-Ne3. At the irrigated Mead site US-Ne1, the maximum LAI constraint of 5 $m^2\,m^{-2}$ for corn in the models contributed to an underestimation of maximum LAI for seven years out of the eleven (e.g., 2002–2003, 2007–2009, 2010–2011) and damped the inter-annual LAI variability. Thus, there is still a need to improve the photosynthesis and carbon allocation processes in the model to eliminate the need for an upper LAI threshold, and to make the model fully prognostic so that it can capture the inter-annual variability in LAI.

The timing of grain fill and harvest time were well simulated by both models at all four sites (Figs. 2–4). Grainfill is usually initiated when the maximum LAI is reached, unless there is a disturbance event. This model behavior is consistent with a previous study at Rosemount, Minnesota (Chen et al., 2015). Although CLM4-Crop did not simulate the first half of the growing season well, the model captured the timing of grain fill at all sites.

Biases in the growing season dynamics of LAI simulations were closely coupled with water availability. During periods of probable water deficit and stress (i.e., annual average evaporative fraction (EF) < 0.5), both models showed large overestimation of LAI relative to observations. This was most prevalent at the rainfed US-Ne3 site. US-Ne3 is rainfed with corn and soybean rotation, and has the lowest annual precipitation (784 mm) among the nine sites (Table 1). The EF

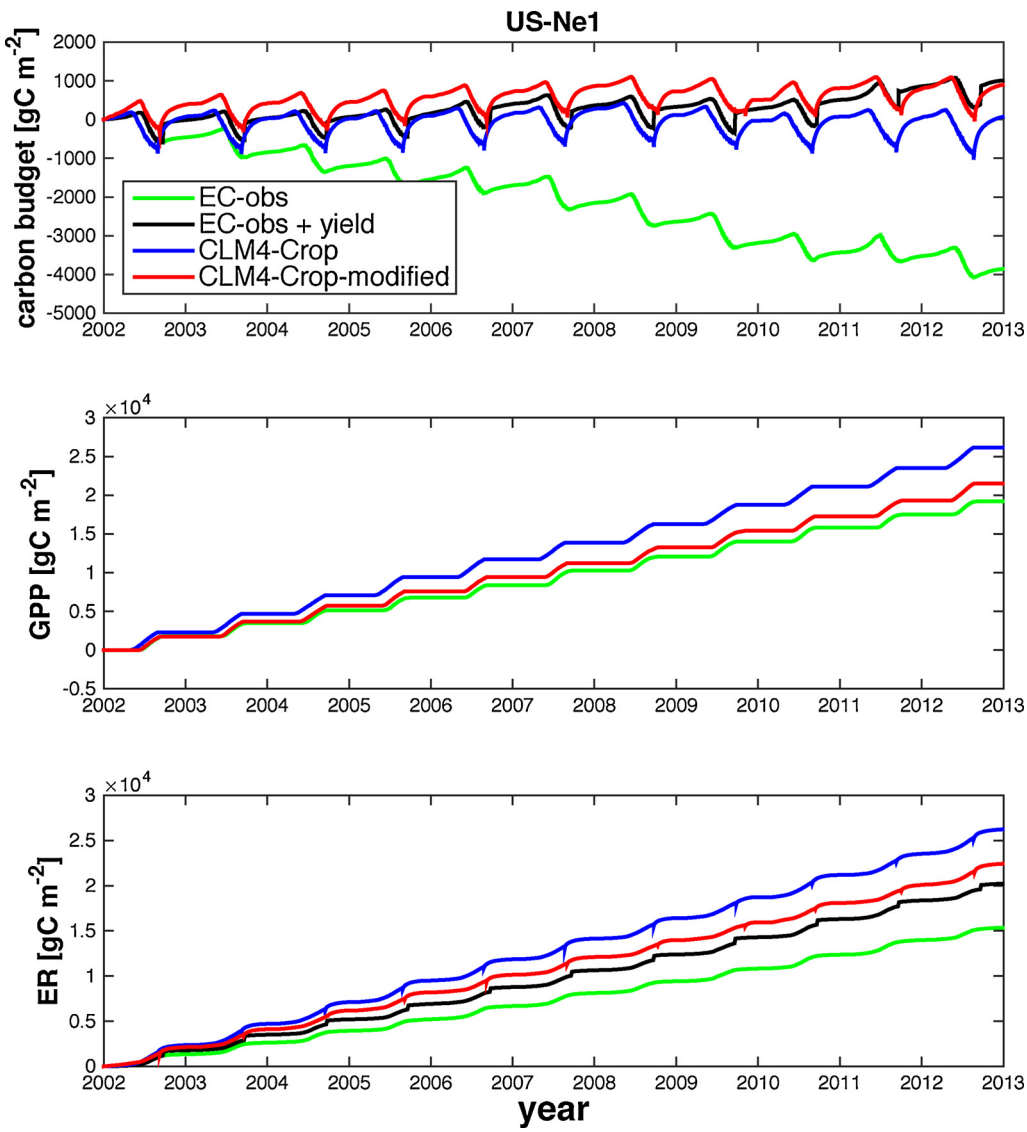

Fig. 8. Long-term carbon budget simulation at site US-Ne1. The green line represents carbon fluxes calculated using the eddy covariance measured carbon fluxes. The black line is the eddy covariance measured carbon fluxes plus annual yield grain carbon from the field. The blue line is the default CLM4-Crop simulation and the red line is CLM4-CropM simulation (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

revealed significant periods of potential water stress for soybean during the vegetative and reproductive growth phase in 2002. For corn, major dry periods occurred during silking and/or reproductive stages in 2003 and during vegetative/silking growth stages in 2005 (Suyker and Verma, 2012). In all cases, maximum simulated LAI was higher than observations, revealing an inadequacy in water stress limitations on phenological and/or physiological processes. Similar model behavior was rarely observed at US-Bo1, where rainfall was more plentiful, or at the irrigated US-Ne1 site.

While US-Ne1 did not display water deficit induced biases, there were years where both models underestimated seasonal LAI maxima. The US-Ne1 site is an irrigated cornfield. Irrigation at this site provided about 40–50% of the total water received (Suyker and Verma, 2010). Due to irrigation, the maximum LAI of corn at this site was higher than the rainfed US-Ne3 site, with some years reaching about $6\,m^2\,m^{-2}$ (Fig. 4). Since CLM4-Crop and CLM4-CropM do not simulate irrigation, we expected a lower simulated LAI (i.e. similar to the value of $5\,m^2\,m^{-2}$ at the rainfed site US-Ne3) in the model estimation. The biases associated with simulating phenology are expected to propagate into the energy and NEE flux simulations (Chen et al., 2015).

3.2. Evaluation of energy fluxes and NEE

The simulated hourly energy and NEE fluxes were evaluated against site observations and presented as Taylor diagrams (Fig. 5) with summary statistics provided in Table 2. Note that these evaluations were performed using the default generic corn and soybean physiological parameterization in CLM4-Crop. The only difference between CLM4-Crop and CLM4-CropM is the timing of planting date. The models were not tuned to any site-specific observations to optimize the results.

Among the energy fluxes, R_n was the best-simulated variable (Fig. 5). The averaged correlation coefficient was 0.94 for the nine sites (54 site-years) for both models. For sites US-Ro1 and US-Ro3, the correlation coefficients were comparatively lower (0.93 for both sites and both models). The correlation coefficients of R_n were lower at these two sites because during winter, a small bias in the precipitation forcing data followed by an underestimation of snow cover caused a large change in ground albedo, which affected the R_n simulation. In a previous study, R_n during growing season was evaluated at these two sites and the correlation coefficient was from 0.98 to 0.99 for the same years (Chen et al., 2015).

The simulated latent heat flux (*LE*) had a standard deviation that

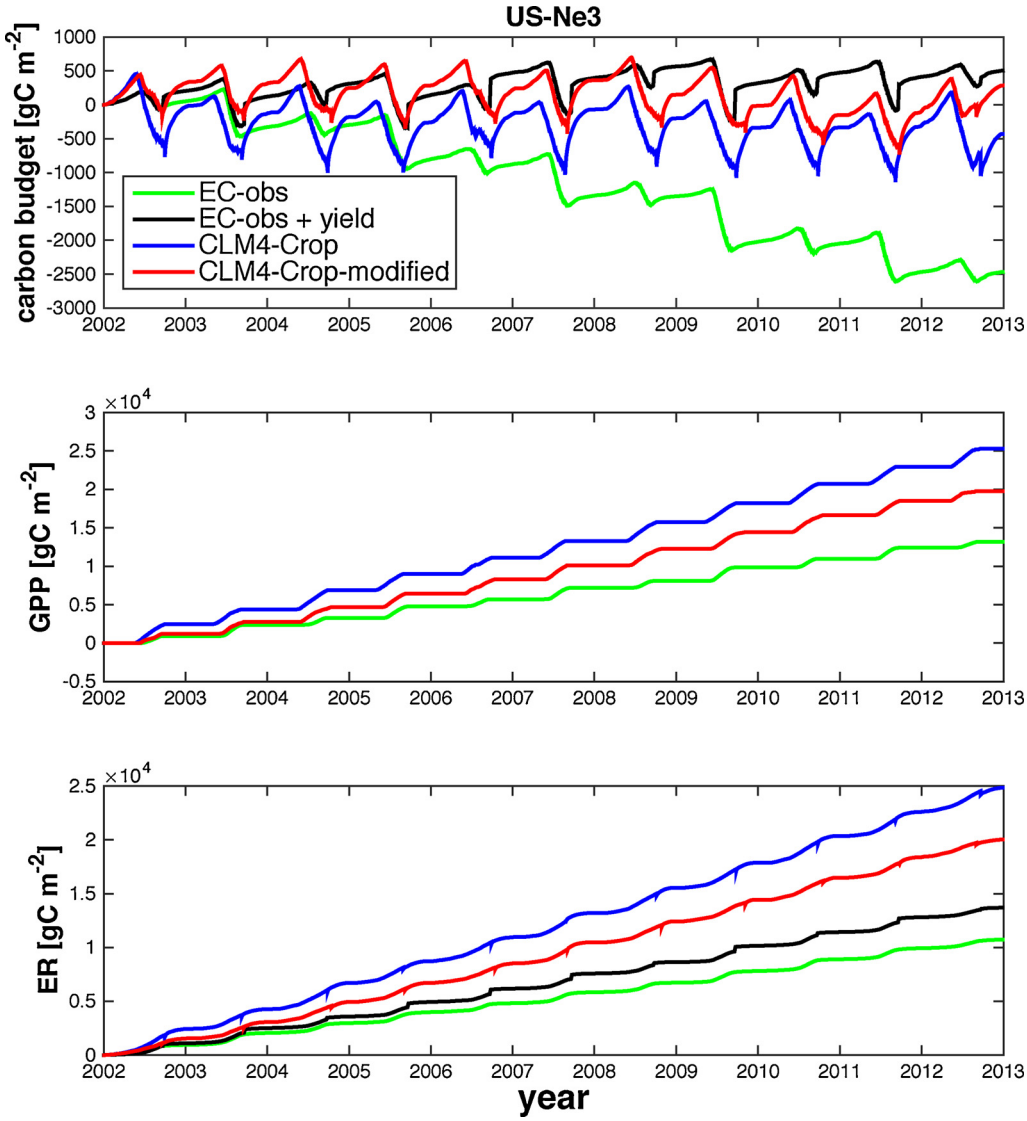

Fig. 9. Long-term carbon budget simulation at site US-Ne3. The green line represents carbon fluxes calculated using the eddy covariance measured carbon fluxes. The black line is the eddy covariance measured carbon fluxes plus annual yield grain carbon from the field. The blue line is the default CLM4-Crop simulation and the red line is CLM4-CropM simulation (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

was closer to the observations and a higher correlation coefficient (r = 0.76 for CLM4-Crop and r = 0.77 for CLM4-CropM) compared to the observed sensible heat flux (*H*) for all sites (r = 0.60 for CLM4-Crop and r = 0.65 for CLM4-CropM) (Fig. 5). This was also found in a previous study when CLM3.5-simulated daily mean energy fluxes were evaluated at 15 AmeriFlux sites (Yuan and Liang, 2011). In their study, simulated *LE* had a higher correlation with the observations compared to simulated *H* at 11 sites, including 2 crop sites. The simulated *H* generally had less variation (standard deviation) than the observations among all sites (Fig. 5). Ground heat flux (*G*), as an energy balance residual (Oleson et al., 2010), had a standard deviation 2 or more times greater than the standard deviation of the observed ground heat flux (Fig. 5). This was consistent among all sites. This bias has been reported for other process based models which calculate *G* as a residual (Webler et al., 2012). The skill of both models when simulating hourly energy fluxes were similar, however, CLM4-CropM generally had higher correlation coefficients and lower RMSDs with the observations when compared to CLM4-Crop. The modified phenological scheme improved the sensible heat flux simulations notably, because the improved early growing season phenology resulted in more realistic canopy temperature, which helped reduce the *H* bias during this period.

The model performance in simulating NEE was similar among sites (Fig. 5). Compared to the energy fluxes, the simulated hourly NEE was generally less correlated with the observations (r = 0.54 for CLM4-Crop and r = 0.62 for CLM4-CropM). There are two reasons for this phenomenon. First, compared to the simulation of the energy fluxes, more nonlinear processes are encompassed in the simulation of the carbon cycle. Second, the CLM4-Crop model assumes the harvested biomass is decomposed on site following harvest in order to close the carbon budget at crop sites (Chen et al., 2015; Levis et al., 2012). This results in an unrealistic carbon flux from the field into the atmosphere after harvest, which lowered the correlation coefficient of simulated and observed NEE. The correlation coefficients of the hourly NEE simulated by CLM4-Crop ranged from 0.45 (US-Bo2) to 0.66 (US-IB1) for the nine sites. The correlation coefficients of NEE simulated by CLM4-CropM ranged from 0.51 (US-Bo2) to 0.76 (US-IB1). The standard deviation of the simulated NEE is close to the observations, indicating the amplitude of NEE was well simulated by both models. CLM4-Crop generally had higher standard deviation than the observations (relevant SD from 1.04 at US-Ro1 to 1.22 at US-Br1), except for the site US-Ne1 (0.98). The standard deviation of the simulated NEE was lower for CLM4-CropM compared to CLM4-Crop across all sites. This lower standard deviation

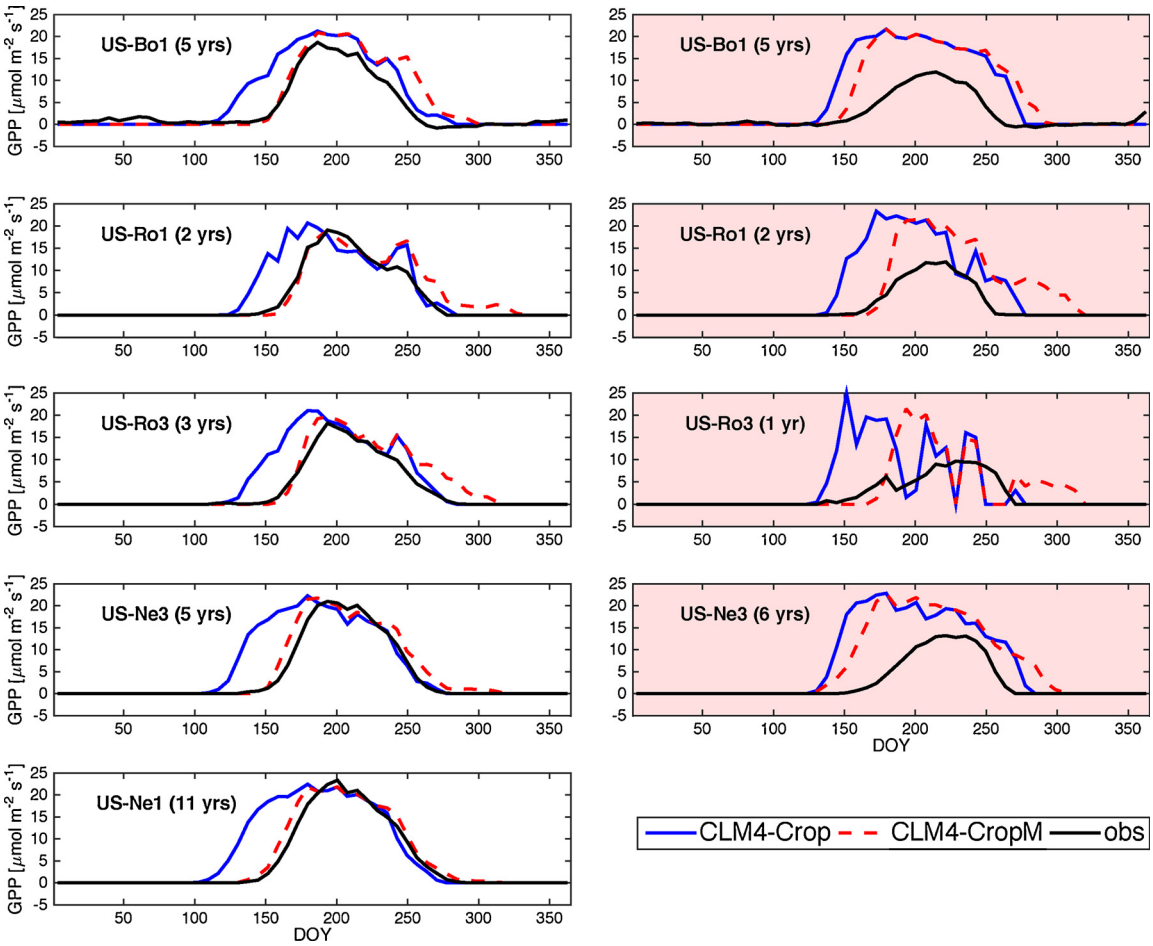

Fig. 10. Multi-year-averaged weekly GPP simulation for Corn and Soybean years respectively across sites. White background indicates corn years and red background indicates soybean years (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

simulated by CLM4-CropM was in better agreement with the observations (i.e. from 1.00 at US-Ro1 to 1.12 at US-Br1).

At the five sites where yield data were available (US-Ro1, US-Ro3, US-Bo1, US-Ne1 and US-Ne3), the mean annual carbon budgets (Eq. (1)) were positive, indicating those agro-ecosystems were carbon sources (Fig. 6). However, there was substantial inter-annual variation (shown in the figure as standard deviation) in both observed and simulated values of carbon budget. For seven of the nine sites, the carbon budget was positive (C source) in both models. On average, CLM4-CropM gave more positive annual carbon budget values. There are two sites out of the nine (US-IB1 and US-Br3) that were estimated as carbon sinks in both models.

In general, CLM4-CropM improved the NEE simulation, with increased correlation coefficient and standard deviation closer to observations. Since the photosynthetic parameterization is the same in both models, the improved NEE skill is directly linked to the simulations of phenology.

3.3. Long-term carbon budget

The long-term carbon budget was evaluated for over 10 years at site US-Bo1, and over 11 years at US-Ne1 and US-Ne3. The measured and modeled accumulated carbon budgets at these three sites are shown in Figs. 7–9, respectively.

3.4. US-Bo1

From 1997 to 2006, measured cumulative NEE was $-3026\,\mathrm{gC\,m^{-2}}$ (Fig. 7). After accounting for crop yields, the total carbon budget was $1543\,\mathrm{gC\,m^{-2}}$ (Eq. (1)), indicating a net carbon emission rate of $154.3\,\mathrm{gC\,m^{-2}\,yr^{-1}}$ over the 10-year period. Here we assumed that grain carbon is the only carbon taken from the field and that the above ground biomass remains as residue. At this site, CLM4-Crop estimated a cumulative carbon source of $86\,\mathrm{gC\,m^{-2}}$ with the carbon emission rate of $8.6\,\mathrm{gC\,m^{-2}\,yr^{-1}}$. CLM4-CropM predicted a carbon source of $516\,\mathrm{gC\,m^{-2}}$ - a carbon emission rate of $51.6\,\mathrm{gC\,m^{-2}\,yr^{-1}}$- still underestimated relative to the measured value, but substantially better than CLM4-Crop.

To better understand what contributes to the underestimated carbon emissions at US-Bo1, the 10-year accumulative GPP and ER[1] simulations were also examined (Fig. 7). GPP and ER were substantially overestimated by 10,902 and 9431 $\mathrm{gC\,m^{-2}}$, respectively, using CLM4-Crop. CLM4-CropM also overestimated both GPP and ER, but the bias was reduced for GPP and ER by 17% and 16%, respectively. The simulated GPP and ER for soybean were biased higher compared to corn. This was partly attributed to the fact that early growing season LAI of soybean was substantially overestimated compared to corn (Fig. 2). The GPP bias of the 5 soybean years contributed to 68% and 75% of the total GPP bias from CLM4-Crop and CLM4-CropM, respectively. ER bias for the 5 soybean years contributed to 58% and 64% of the total ER bias from CLM4-Crop and CLM4-CropM, respectively. Overall, due to the larger bias in simulated GPP compared to ER, the NEE values estimated by both models were more negative compared to the flux tower derived

[1] Here the yield carbon was added to the ecosystem respiration term to account for the total carbon released from the agro-ecosystem.

Table 3
Sensitivity tests for soybean simulations across sites. There are 5 years at US-Bo1, 6 years at US-Ne3, 2 years at US-Ro1 and 1 year at US-Ro3. The values of the parameters showed in the table are the default values. The left 4 columns and the right 4 columns showed the model responses after 20% increase and decrease of the parameter, respectively. Here xl is leaf angle distribution factor; slatop is specific leaf area at top of the canopy ($m^2\,gC^{-1}$); feafi is a dimensionless parameter for calculating initial carbon allocation to leaf; Vcmax25 is the maximum carboxylation rate of rubisco ($\mu mol\,m^{-2}\,s^{-1}$); GDDmat is the growing degree days required to reach plant maturity, i.e.- crop harvest.

	−20%				+20%			
	Δ NEE	Δ GPP	Δ ER	Δ LAI	Δ NEE	Δ GPP	Δ ER	Δ LAI
xl = 0								
US-Bo1	1.0%	−0.1%	0.0%	−0.3%	1.1%	−0.6%	−0.4%	0.1%
US-Ne3	−13.9%	1.1%	0.8%	0.3%	26.1%	−1.8%	−1.2%	−0.6%
US-Ro1	−50.1%	1.4%	0.8%	−0.1%	19.1%	−1.3%	−1.1%	−0.1%
US-Ro3	4.4%	0.6%	0.8%	0.1%	0.7%	−0.7%	−0.6%	−0.1%
slatop = 0.07								
US-Bo1	9.1%	−3.7%	−2.3%	−5.4%	−5.1%	2.1%	1.4%	3.2%
US-Ne3	47.6%	−4.2%	−3.1%	−8.5%	−26.7%	2.2%	1.6%	3.8%
US-Ro1	116.3%	−4.1%	−2.7%	−4.8%	−97.5%	2.7%	1.6%	2.9%
US-Ro3	11.2%	−2.6%	−1.8%	−5.6%	−4.0%	1.4%	1.0%	3.0%
feafi = 0.85								
US-Bo1	6.2%	−2.7%	−1.7%	−4.3%	−3.7%	1.7%	1.1%	2.6%
US-Ne3	28.7%	−3.0%	−2.4%	−6.5%	−19.3%	1.8%	1.3%	3.1%
US-Ro1	73.4%	−2.9%	−2.1%	−3.8%	−71.2%	2.0%	1.2%	2.3%
US-Ro3	8.1%	−2.0%	−1.4%	−4.3%	−3.3%	1.0%	0.8%	2.5%
V_{cmax25} = 100								
US-Bo1	13.0%	−8.0%	−5.7%	−2.8%	−4.7%	4.3%	3.3%	1.4%
US-Ne3	130.1%	−10.9%	−8.0%	−5.3%	−66.9%	5.6%	4.1%	2.2%
US-Ro1	279.6%	−12.4%	−9.1%	−3.2%	−208.9%	7.8%	5.3%	1.5%
US-Ro3	20.7%	−8.7%	−6.9%	−3.6%	−2.9%	5.0%	4.6%	0.9%
GDD_{mat} (dynamic)								
US-Bo1	−2.3%	−9.6%	−8.9%	−26.3%	−2.5%	2.0%	1.5%	7.6%
US-Ne3	343.5%	−20.0%	−12.4%	−29.7%	−52.4%	2.9%	1.8%	9.8%
US-Ro1	−43.6%	0.7%	0.2%	−0.9%	11.6%	−0.8%	−0.7%	−0.4%
US-Ro3	5.7%	0.9%	1.1%	−2.5%	5.7%	−0.8%	−0.4%	−0.6%

Table 4
Calibration of soybean simulation, part one. Numbers in the table are correlation coefficients and bias ($\mu mol\,m^{-2}\,s^{-1}$, in parenthesis) of the simulated weekly leaf area index (LAI), net ecosystem exchange (NEE), gross primary production (GPP), and ecosystem respiration (ER).

Model run	LAI	NEE	GPP	ER
Control run: gddmat = default vcmx = 100 slatop = 0.07 feafi = 0.85				
Control 1:				
US-Bo1(1998,2000)	0.85 (0.89)	0.51 (−98.70)	0.77 (1117.50)	0.55 (961.70)
US-Ne3(2002,2004,2006)	0.77 (1.40)	0.55 (−178.38)	0.72 (784.47)	0.67 (605.04)
US-Ro1(2008,2010)	0.62 (1.34)	0.79 (−196.21)	0.88 (822.78)	0.78 (625.72)
Control 1 average	0.75 (1.24)	0.61 (−160.71)	0.78 (890.57)	0.67 (712.85)
Control 2:				
US-Bo1(2002,2004,2006)	0.78 (1.18)	0.64 (53.11)	0.84 (577.65)	0.68 (629.94)
US-Ne3(2008,2010,2012)	0.72 (1.46)	0.47 (−125.23)	0.71 (936.9)	0.7 (809.69)
US-Ro3(2009)	0.67 (1.28)	0.52 (61.34)	0.73 (538.78)	0.49 (598.02)
Control 2 average	0.74 (1.32)	0.55 (−22.15)	0.77 (726.06)	0.66 (702.42)
Total site-year average	**0.75 (1.28)**	**0.58 (−91.43)**	**0.78 (808.32)**	**0.66 (707.63)**
*Calibration run: gddmat*0.8, jday*0.8 vcmx = 80 slatop = 0.06 feafi = 0.68*				
Calibration:				
US-Bo1(1998,2000)	0.97 (−0.08)	0.83 (−46.72)	0.86 (534.36)	0.67 (484.44)
US-Ne3(2002,2004,2006)	0.91 (0.10)	0.62 (123.45)	0.83 (-99.7)	0.85 (23.57)
US-Ro1(2008,2010)	0.69 (0.15)	0.81 (−38.34)	0.9 (238.25)	0.76 (199.87)
Calibration average	0.87 (0.06)	0.74 (28.61)	0.86 (178.02)	0.77 (205.62)
Validation:				
US-Bo1(2002,2004,2006)	0.95 (0.11)	0.93 (137.09)	0.97 (120.16)	0.86 (255.13)
US-Ne3(2008,2010,2012)	0.8 (0.10)	0.65 (87.62)	0.86 (-23.87)	0.85 (62.75)
US-Ro3(2009)	0.86 (0.04)	0.75 (113.65)	0.91 (197.80)	0.7 (311.04)
Validation average	0.87 (0.10)	0.78 (112.54)	0.91 (69.52)	0.83 (180.67)
Total site-year average	**0.87 (0.08)**	**0.76 (70.57)**	**0.89 (123.77)**	**0.8 (193.15)**

values.

3.5. US-Ne1

From 2002 to 2012, observed NEE at site US-Ne1 showed a carbon sink of $-3860\,gC\,m^{-2}$ (Fig. 8). However, after accounting for the harvested grain carbon, this site was actually a carbon source of $1009\,gC\,m^{-2}$ over the 11 years, an average carbon emission rate of $91.7\,gC\,m^{-2}$ per year. The CLM4-Crop simulated carbon budget for the 11 years was $69\,gC\,m^{-2}$, a carbon emission rate of $6.3\,gC\,m^{-2}\,yr^{-1}$. CLM4-CropM simulated carbon budget was $897\,gC\,m^{-2}$, or an annual carbon emission rate of $81.5\,gC\,m^{-2}\,yr^{-1}$, in closer agreement with the observations, largely because CLM4-CropM simulated a more realistic growing season length.

Table 5
Calibration of soybean simulation, part two. Numbers in the table are correlation coefficients and bias (μmol m^{-2} s^{-1}, in parenthesis) of the simulated weekly net radiation (*Rn*), sensible heat flux (*H*) and latent heat flux (*LE*).

Model run	*Rn*	*H*	*LE*
Control run: gddmat = default vcmx = 100 slatop = 0.07 fleafi = 0.85			
Control 1:			
US-Bo1(1998,2000)	0.98 (−5.92)	0.43 (−0.55)	0.92 (−4.71)
US-Ne3(2002,2004,2006)	0.97 (−5.92)	0.23 (4.19)	0.85 (−0.57)
US-Ro1(2008,2010)	0.94 (−34.71)	0.51 (−18.40)	0.9 (2.37)
Control 1 average	0.97 (−14.15)	0.36 (−3.62)	0.88 (−0.91)
Control 2:			
US-Bo1(2002,2004,2006)	0.9 (−2.83)	0.33 (5.45)	0.9 (−3.11)
US-Ne3(2008,2010,2012)	0.98 (−3.62)	0.2 (7.51)	0.88 (1.15)
US-Ro3(2009)	0.93 (−17.8)	0.63 (−13.97)	0.87 (1.28)
Control 2 average	0.94 (−5.31)	0.32 (3.56)	0.89 (−0.66)
Total site-year average	**0.95 (−9.73)**	**0.34 (−0.03)**	**0.89 (−0.78)**
*Calibration run: gddmat*0.8, jday*0.8 vcmx = 80 slatop = 0.06 fleafi = 0.68*			
Calibration:			
US-Bo1(1998,2000)	0.98 (−9.94)	0.49 (1.36)	0.91 (−10.44)
US-Ne3(2002,2004,2006)	0.98 (−13.87)	0.48 (7.35)	0.91 (−11.60)
US-Ro1(2008,2010)	0.95 (−39.19)	0.56 (−16.81)	0.88 (−4.06)
Calibration average	0.97 (−19.98)	0.5 (−1.26)	0.9 (−9.12)
Validation:			
US-Bo1(2002,2004,2006)	0.9 (−7.05)	0.5 (5.88)	0.95 (−7.59)
US-Ne3(2008,2010,2012)	0.98 (−11.35)	0.41 (10.33)	0.91 (−9.26)
US-Ro3(2009)	0.93 (−21.01)	0.72 (−14.72)	0.93 (−1.26)
Validation average	0.94 (−10.89)	0.49 (4.85)	0.93 (−7.40)
Total site-year average	**0.96 (−15.43)**	**0.5 (1.79)**	**0.91 (−8.26)**

Again, the 11-year accumulative GPP and ER were overestimated by both models (Fig. 9). However, the biases were smaller compared to US-Bo1 (Fig. 6). GPP and ER were overestimated by 6325 and 5470 gC m^{-2} (575 and 497 gC m^{-2} yr^{-1}) using CLM4-Crop. CLM4-CropM with better phenology simulation reduced the bias of GPP by 67% and reduced the bias of ER by 63%.

3.6. US-Ne3

The EC measured NEE showed that US-Ne3 was a carbon sink of −2468 gC m^{-2} in the 11 years analyzed (2002-2012). However, after accounting for the yield carbon, US-Ne3 was a moderate carbon source for the 11 years with a net carbon emission of 504 gC m^{-2}. The average carbon emission rate of 45.8 gC m^{-2} yr^{-1} was about half of the irrigated Mead site US-Ne1 and one third of the Bondville site US-Bo1. CLM4-Crop estimated the 11-year carbon budget to be 244 gC m^{-2}, with a carbon emission rate of 22.2 gC m^{-2} yr^{-1}, while CLM4-CropM predicted a mean carbon emission rate of 57.8 gC m^{-2} yr^{-1}.

Similar to the sites US-Bo1 and US-Ne1, both GPP and ER were overestimated at US-Ne3 (Fig. 8). GPP and ER were overestimated by 12,176 and 11,933 gC m^{-2} (1107 and 1085 gC m^{-2} yr^{-1}) using CLM4-Crop. CLM4-CropM reduced the overestimations of GPP by 51% and 13% for corn and soybean, respectively, while the overestimations of ER were reduced by 39% and 12% for corn and soybean, respectively.

The average carbon budget of the three sites under consideration equated to a carbon source of 97.3 ± 54.5 gC m^{-2} yr^{-1} (the reported uncertainty is the standard error). This value is comparable with a European study over seven crop sites from 2004 to 2007 (Kutsch et al., 2010), which estimated an annual carbon loss of 95 ± 87 gC m^{-2} yr^{-1} from the seven European sites. The CLM4-Crop mean value was 12.4 ± 8.6 and the CLM4-CropM value was 63.6 ± 15.8 gC m^{-2} yr^{-1}. The modified model CLM4-CropM improved the carbon budget estimation and the estimated value is within the range of the measured values.

The response of ecosystems to regionally heterogeneous stimuli such as historical land use, or rainfall and temperature anomalies and their effects on carbon fluxes is an important aspect for estimating large scale carbon sinks and sources (Schimel et al., 2001). In this case, comparing the three sites we can see large differences in the mean annual carbon budgets. It also emphasizes the importance of evaluating a model across different sites before applying it to the region.

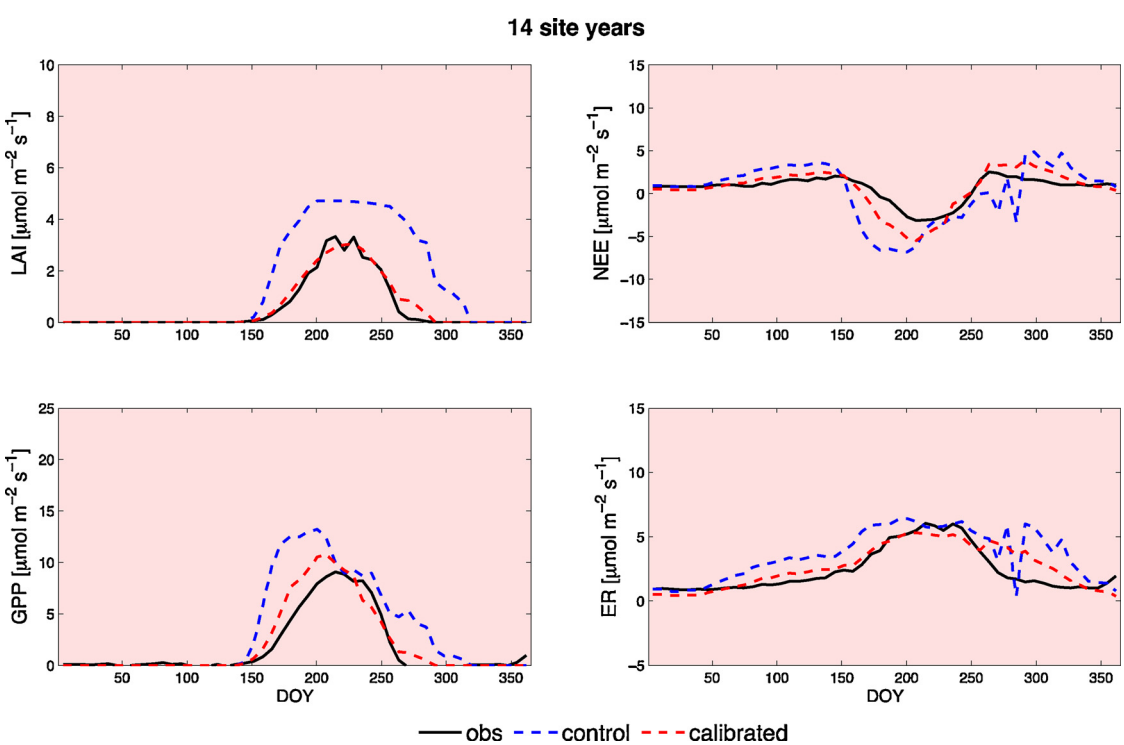

Fig. 11. Calibration of the model CLM4-CropM for soybean. Weekly averaged simulated soybean LAI, NEE, GPP and ER were compared with observations over 14 site years. The blue dotted line represents results for the uncalibrated CLM4-CropM and the red dotted line represents results for the calibrated model CLM4-CropM'. Parameters used in the control run and the calibration run can be found in table 3.5(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

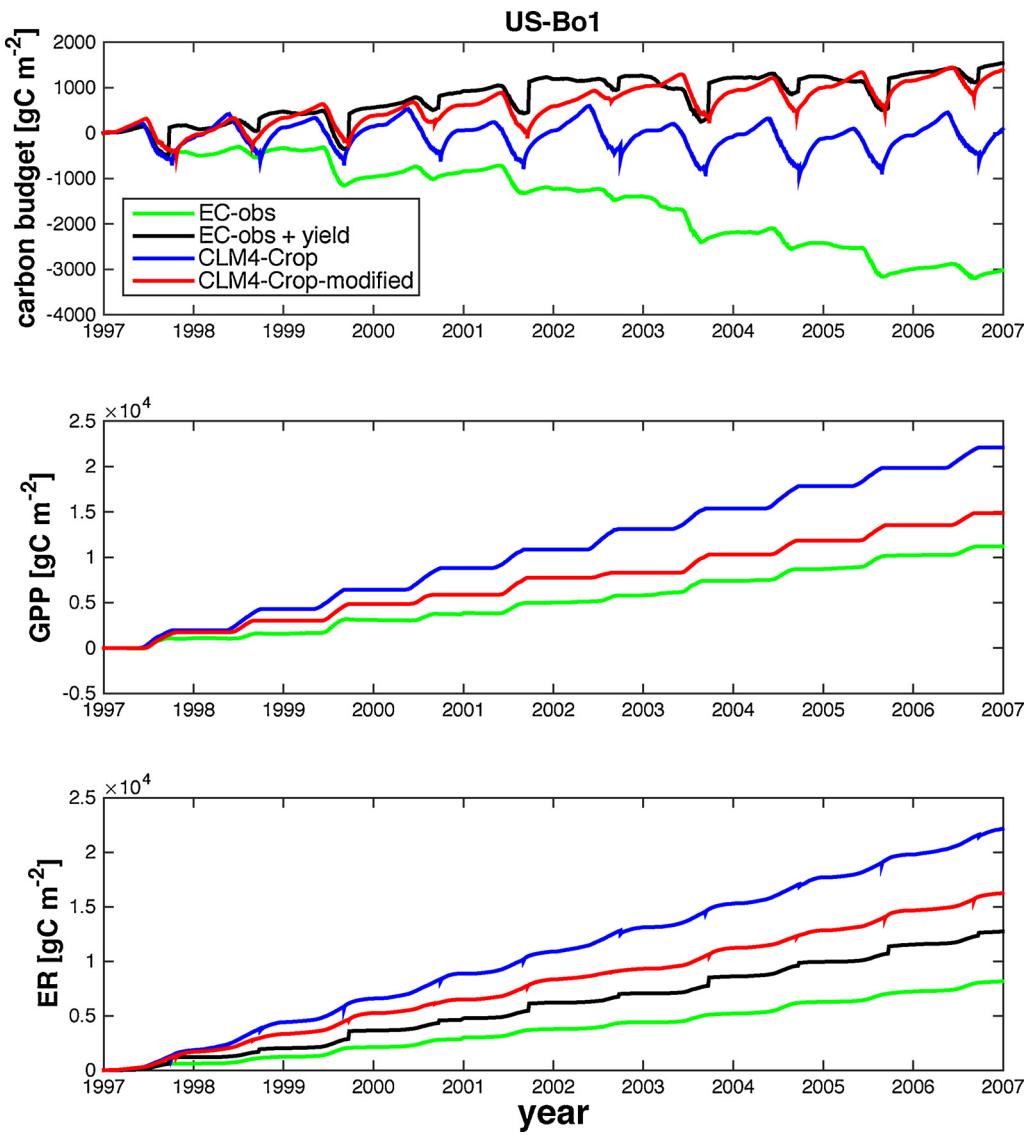

Fig. 12. Long-term carbon budget simulation at site US-Bo1 (after calibration). The red line is the calibrated CLM4-CropM' simulation. The green line represents carbon fluxes calculated using the eddy covariance measured carbon fluxes. The black line is the eddy covariance measured carbon fluxes plus annual yield grain carbon from the field. The blue line is the default CLM4-Crop simulation (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

3.7. Sensitivity tests and model calibration

The results above indicate that the over-estimated GPP is one of the major model deficiencies that needs to be resolved in order to perform reliable long-term simulations. The model simulated ER is intimately connected to the GPP simulation because plant maintenance respiration (MR) is simulated proportional to plant biomass. Growth respiration is currently simulated as 30% of the available GPP (GPP-MR), and heterotrophic respiration is proportional to litter and soil carbon pools, which are also related to plant biomass. Thus, our sensitivity tests focus on the simulated GPP. First, we compared the model simulated GPP for corn and soybean across sites (Fig. 10). The poor simulation of soybean GPP is consistent across sites and contributed to the overall bias in the cumulative integrals. The model-observation mismatch in soybean GPP can be separated into two categories. The first category is site/year specific errors. For example, we observed unrealistic variations in simulated GPP during the 2008 growing season at US-Ro1 and during the 2009 growing season at US-Ro3. This error was eliminated by reverting to the original calculation of β_t (Eq. (3), i.e. by eliminating the 1.25 soybean drought tolerance factor). This factor was recently introduced to increase soybean drought tolerance (Levis et al., 2012).

The second category is systematic bias across all sites. For example, the soybean leaf growth rate during the early growing season was overestimated across all sites. The estimated harvest time was also biased late, resulting in a longer growing season for soybean across all sites. Further, accumulative GPP and ER for soybean were consistently over-estimated at all sites. In this section, we aimed to improve the model by reducing the systematic bias across sites to provide more reliable carbon simulations for corn-soybean systems.

In order to diagnose the potential factors that contributed to the high bias of GPP simulation for soybean, as well as to diagnose the most sensitive parameters to carbon flux simulations across sites, a series of sensitivity tests were carried out across sites. Previously, global sensitivity analyses have been carried out for CLM in which 80 parameters were investigated with respect to their influence on LAI and carbon simulations (Sargsyan et al., 2014). However, for crop PFTs in CLM4, the nitrogen limitation has been turned off, therefore, previous sensitivity tests will differ from those carried out here.

Based on previous analyses of CLM model physiological parameters (Bilionis et al., 2015; Sargsyan et al., 2014), five parameters were

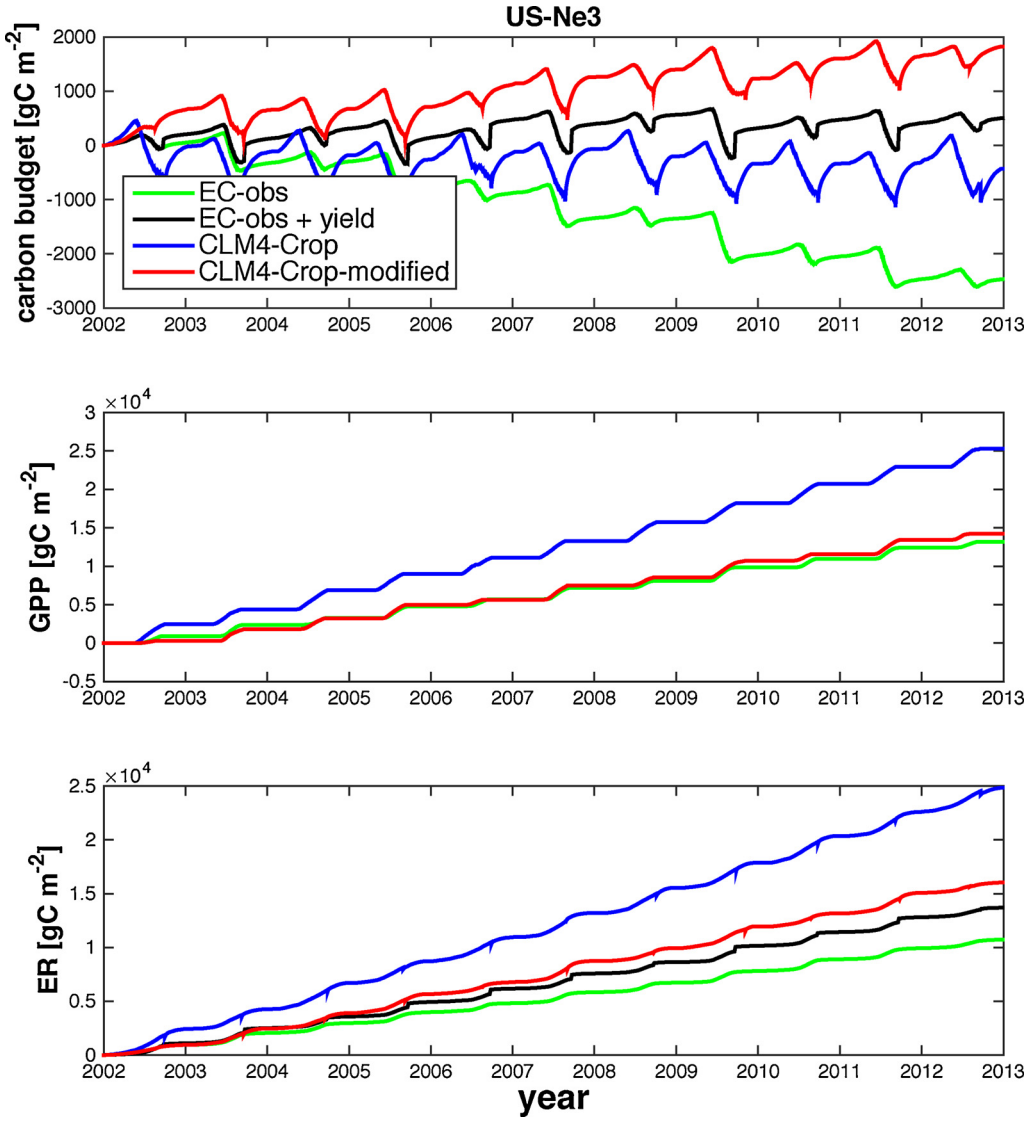

Fig. 13. Long-term carbon budget simulation at site US-Ne3 (after calibration). The red line is the calibrated CLM4-CropM' simulation. The green line represents carbon fluxes calculated using the eddy covariance measured carbon fluxes. The black line is the eddy covariance measured carbon fluxes plus annual yield grain carbon from the field. The blue line is the default CLM4-Crop simulation (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

chosen for the sensitivity tests because of their large impact on the simulation of LAI, GPP, and ER. Note here the leaf nitrogen related parameters were not tested because this version of the model assumes no nitrogen limitation for cropping systems. The parameters included: leaf angle distribution factor (χl), specific leaf area at the top of plant canopy (*slatop*), initial carbon allocation to leaf (*fleafi*), the maximum carboxylation rate of rubisco at 25 °C (V_{cmax25}), and growing degree days required to reach plant maturity (GDD_{mat}). Each parameter value was increased and decreased by 20% (except χl, which was 0 in the control run. It was increased and decreased by 0.2). The corresponding changes in LAI, and the carbon fluxes compared to the control run are shown in Table 3.

The sensitivity tests revealed that among the five parameters, V_{cmax25}, *slatop* and *fleafi* are the parameters to which LAI, GPP and ER are most sensitive. This is consistent among the four sites (US-Bo1, US-Ne3, US-Ro1 and US-Ro3). The sensitivity of NEE to those parameters varies greatly (Table 3) because NEE is the very small difference between the large opposing fluxes of GPP and ER. Thus, when describing model sensitivity to changing parameter values we refer to the gross fluxes.

The sensitivity of simulated soybean phenology to GDD_{mat} varies among sites. A 20% reduction of GDD_{mat} at US-Bo1 and US-Ne3 shortened the length of the growing season and therefore had a substantial influence on LAI, GPP and ER simulations. However, at the two Rosemount sites, a 20% reduction in GDD_{mat} had very little impact on LAI or the carbon fluxes. This is because at Rosemount, the model reached another important threshold (the number of days past planting) before GDD_{mat} was reached. This indicates that the current use of a universal GDD_{mat} is not sufficiently general across site locations. For example, at higher latitudes, soybean cultivars were selected to better adapt to cooler temperatures and thus require less GDD to reach plant maturity. Currently most of the crop models use static maximum growing degree days to estimate harvest time. However, a dynamic GDD_{mat} that changes with annual mean temperature could represent this cultivar selection at different latitudes and better predict the harvest time. Selecting crop cultivars for better adaptation to local environment and climate is a process similar to acclimation, but on a shorter time scale. When simulating crops at the global scale, it is impossible to parameterize each crop cultivar, thus a dynamic parameterization scheme that considers the cultivar selection will provide

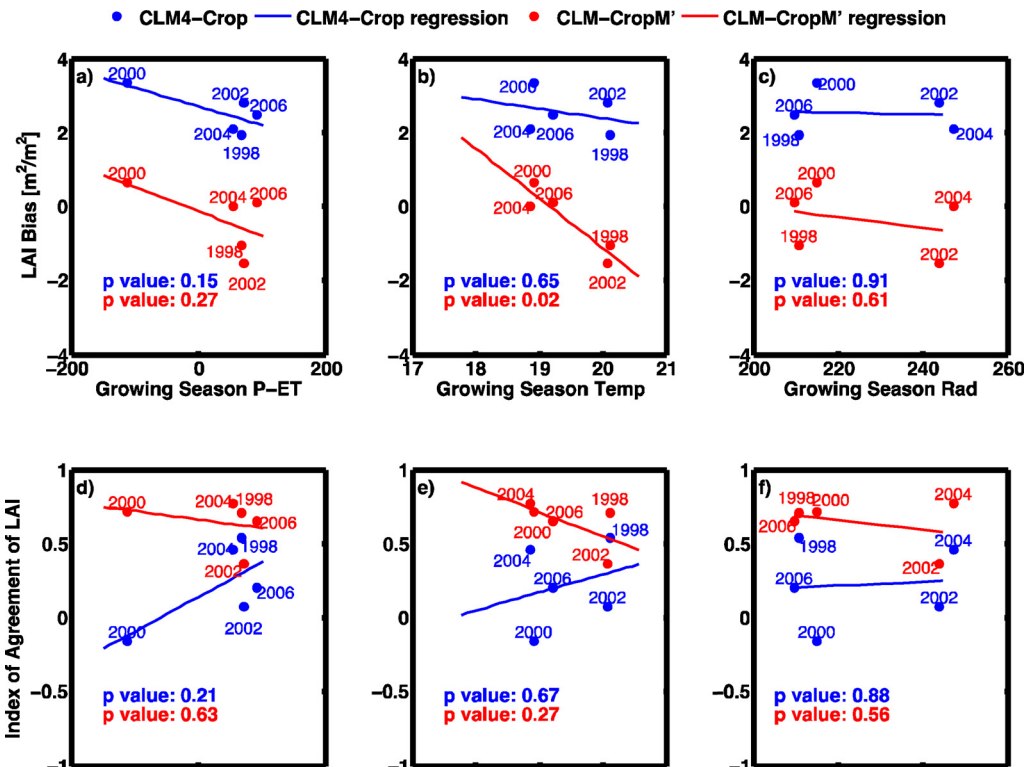

Fig. 14. Linear relationships of soybean LAI bias (model – observation, a–c) and climate factors and linear relationships of Index of agreement of LAI (d–f) and climate factors at site US-Bo1. The climate factors are difference of precipitation and evapotranspiration during growing season (mm, a,d), averaged air temperature during growing season (°C, b,e) and averaged surface solar radiation during growing season (W m^{-2}, c,f). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

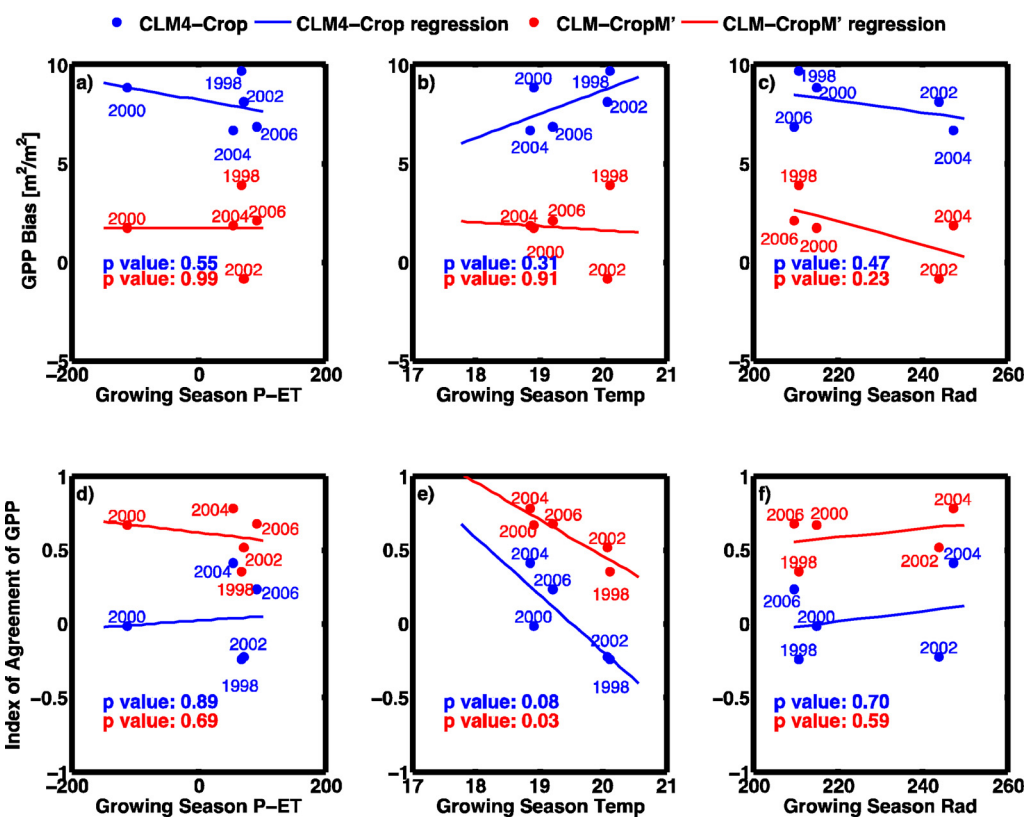

Fig. 15. Linear relationships of soybean GPP bias (model – observation, a, b, c) and climate factors and linear relationships of Index of agreement of GPP (d,e,f) and climate factors at site US-Bo1. The climate factors are difference of precipitation and evapotranspiration during growing season (mm, a,d), averaged air temperature during growing season (°C, b,e) and averaged surface solar radiation during growing season (W m^{-2}, c,f). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

more reasonable estimations of the crop phenology and regional carbon budget.

Following the sensitivity tests, the model was calibrated and validated using the site-years that have yield information available (for the purpose of evaluating ER. see Tables 4 and 5). Half of the site-years were used to calibrate the model and the other half were used to validate the model. Based on a previous CLM4-Crop model calibration study (Bilionis et al., 2015), the specific leaf area at top of the canopy

(*slatop*) was reduced from 0.07 to 0.06 $m^2\ gC^{-1}$. The initial fraction of carbon allocated to leaves (*fleafi*) was reduced from 0.85 to 0.68 (The range of this parameter in previous literature is from 0.85 (Levis et al., 2012) to 0.47 (Twine et al., 2013)). GDD_{mat} was reduced to 80% of the calculated value. The threshold of days past planting, which is a second constraint on harvest time, was also reduced by 80%, in order to improve the simulation at sites located at higher latitudes, like Rosemount.

Finally, the V_{cmax25} was reduced by one half of its observed standard deviation from 100.7 $\mu mol\ m^{-2}\ s^{-1}$ to 80 $\mu mol\ m^{-2}\ s^{-1}$(see Section 2.3). There are two reasons for this downward adjustment. First, directly implementing leaf trait V_{cmax25} values into big leaf models like CLM has been shown to cause overestimated canopy level GPP (Bonan and Oleson, 2002). Second, most estimates of V_{cmax25} ignore the importance of mesophyll resistance assume that the intercellular CO_2 concentration is equal to CO_2 concentration at the sites of carboxylation. This can lead to an overestimation of V_{cmax} by up to 20% (June, 2011; Long and Bernacchi, 2003). We refer to this calibrated version of the model as CLM4-CropM'.

The calibrated CLM4-CropM' simulations of LAI, NEE, GPP and ER over the 14 soybean years across sites were compared with the default CLM4-CropM results in Fig. 11. After calibration the simulated seasonal variation and magnitude of those variables was greatly improved. The correlation coefficients and biases of modeled and observed weekly LAI, NEE, GPP and ER values are shown in Table 4. Based on this calibration, the average correlation coefficient of LAI across 14 site-years for soybean increased from 0.75 to 0.87 while the bias of LAI reduced from 1.28 to 0.08 $m^2\ m^{-2}$. Further, the simulations of the gross carbon fluxes GPP and ER were also improved with increased correlation coefficients and reduced bias across site years (Table 4). The correlation coefficient of NEE increased from 0.58 to 0.76. The bias of NEE remained about the same magnitude before and after the calibration. Again, this is because NEE is a small difference between two much larger gross carbon fluxes acting in opposite directions. The CLM4-CropM' predicted carbon budget for US-Bo1 between 1997 and 2006 was 1391.8 $gC\ m^{-2}$, much closer to the observed 1543.2 $gC\ m^{-2}$ and the bias was reduced by 89.6% (Fig. 12). The site US-Ne3 from 2002 to 2012 was estimated by CLM4-CropM' to be a carbon source of 1825.2 $gC\ m^{-2}$, in the same direction but higher than the observed value of 504.9 $gC\ m^{-2}$ (Fig. 13).

Overall, with the phenology improvement and across-site calibration, CLM4-CropM' was able to significantly improve LAI, energy fluxes and carbon fluxes simulations across 9 sites.

3.8. Model performance and climate

After calibration, the two models' biases as well as the new index of agreement of simulated LAI and GPP were regressed with three climate factors: the difference between precipitation and evapotranspiration during growing season (calculated as April to September), mean growing season temperature, and the mean growing season net radiation (Figs. 14 and 15). The difference of modeled corn and soybean response to climate conditions are quite large, thus we analyzed the corn years and soybean years separately (for corn years see Figs. S1 and S2).

Overall, CLM4-CropM' has lower biases and higher indices of agreement with observations than CLM4-Crop for both soybean and corn for almost all climate conditions. CLM4-Crop overestimated both LAI and GPP at US-Bo1. For LAI, CLM4-Crop performs better in wetter (Fig. 14a, d) and warmer (Fig. 14c, e) years, while in the drier year 2000 (soybean) and 2001 (corn) when the growing season water budget is negative, CLM4-Crop largely overestimated LAI. CLM4-CropM' significantly improves LAI simulation in drier climate condition with lower bias and index of agreement closer to 1. Although the overall improvement was obvious, CLM4-CropM' underestimated LAI in two warmer years, 1998 and 2002, compared to overestimated LAI in the default model in those two years, indicating a higher model LAI simulation sensitivity to higher temperature. While both CLM4-Crop and CLM4-CropM' provides better simulation of GPP in colder years, CLM4-CropM' provides higher index of agreement and lower bias for all soybean years (Fig. 15).

4. Conclusions

This study used data from nine AmeriFlux sites (54 site-years) to examine the performance of two versions of CLM4-Crop (CLM4-Crop and CLM4-CropM) in capturing seasonal variations in leaf area index (LAI), NEE, ER, and GPP, as well as long term carbon budgets for agro-ecosystems. Some key parameters were analyzed and the model was calibrated across sites for soybean simulations. Our analyses indicate that:

1 Biases in CLM4-Crop simulated early growing season LAI were consistent across the 40 site years where LAI data were available. Evaluation of results showed that compared to the default crop phenology algorithm currently used in CLM4-Crop, the GDT approach improved the early growing season phenology, with better timing of leaf emergence and better representation of inter-annual variations across all sites. The high bias in early growing season LAI that was found in CLM4-Crop simulations has been largely reduced using the GDT approach and this is consistent across sites. However, in both models, soybean LAI was generally overestimated, especially in drier conditions.
2 *LE* was better simulated than *H* by both models (r = 0.76 and 0.77 for CLM4-Crop and CLM4-CropM, respectively) (r = 0.60 and 0.65 for CLM4-Crop and CLM4-CropM, respectively) at all sites. The GDT method notably improved sensible heat flux simulation.
3 Consistent overestimation of the carbon sink strength in CLM4-Crop accumulated in the long-term carbon analysis and resulted in a near-neutral carbon budget instead of a carbon source for most of the sites (except US-Ne3 where it estimated a carbon sink). After phenology improvement, NBP, GPP and ER were well estimated for corn years in CLM4-CropM. For soybean, the NBP budgets were reasonably estimated in CLM4-CropM but GPP and ER were overestimated. This was consistent across all sites.
4 Sensitivity tests show that V_{cmax25}, *slatop* and *fleafi* were the parameters that LAI, GPP and ER are mostly sensitive to. GDD_{mat} largely affects the length of the simulated growing season at lower latitude sites but did not affect the LAI simulation at the more northerly Rosemount sites. The calibrated model CLM4-CropM' generally had shorter growing season, reduced GPP and ER than CLM4-Crop, and also gave more positive carbon budgets, which were closer to the observations. The overestimation of annual GPP and ER for soybean was reduced by 84.7% and 72.7%, respectively, in the calibrated CLM4-CropM' model.
5 CLM4-Crop LAI simulation was better in wetter and warmer years, while the model largely overestimated LAI in drier years. For GPP simulation, CLM4-Crop did better in colder years, and overestimated LAI in warmer years. CLM4-CropM' improved LAI and GPP simulation in almost all climate conditions. However, in very warm growing seasons the model underestimated LAI.

In general, the calibrated CLM4-CropM' showed robust estimation of crop phenology, energy and carbon fluxes at all the sites. Although it still has some deficiencies (i.e., slight overestimation of the early growing season GPP), we expect it to give reasonable estimations over the US Corn Belt.

Acknowledgements

Funding for this research has been provided by the National Science Foundation, ATM-0546476 (TG) and by the Office of Science (BER) U.S. Department of Energy, DE-FG02-06ER64316 (TG and JB). We would

like to thank Jens Kattge at the Max Planck Institute for Biogeochemistry for his helpful discussions about vcmax of corn and soybean. JDW acknowledges support from the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Program, through Oak Ridge National Laboratory's Terrestrial Ecosystem Science (TES) Science Focus Area (SFA). ORNL is managed by UT-Battelle, LLC, for the U.S. DOE under contract DE-AC05-00OR22725.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.agrformet.2018.03.012.

References

Baker, J.M., Griffis, T.J., 2005. Examining strategies to improve the carbon balance of corn/soybean agriculture using eddy covariance and mass balance techniques. Agric. For. Meteorol. 128, 163–177. http://dx.doi.org/10.1016/j.agrformet.2004.11.005.

Bilionis, I., Drewniak, B.a., Constantinescu, E.M., 2015. Crop physiology calibration in the CLM. Geosci. Model Dev. 8, 1071–1083. http://dx.doi.org/10.5194/gmd-8-1071-2015.

Bonan, G., Oleson, K., 2002. The land surface climatology of the community land model coupled to the NCAR community climate model. J. Clim. 94, 3123–3150.

Bonan, G., Oleson, K., Fisher, R., Lasslop, G., Reichstein, M., 2012. Reconciling leaf physiological traits and canopy flux data: use of the TRY and FLUXNET databases in the Community Land Model version 4. J. Geophys. Res. 117, G02026. http://dx.doi.org/10.1029/2011JG001913.

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-campen, H., Müller, C., Reichstein, M., Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Glob. Change Biol. 13, 679–706. http://dx.doi.org/10.1111/j.1365-2486.2006.01305.x.

Chen, M., Griffis, T.J., Baker, J., Wood, J.D., Xiao, K., 2015. Simulating crop phenology in the Community Land Model and its impact on energy and carbon fluxes. J. Geophys. Res. Biogeosci. 120, 310–325. http://dx.doi.org/10.1002/2014JG002780.

Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., Denning, a.S., Dirmeyer, Pa., Houser, P.R., Niu, G., Oleson, K.W., Schlosser, C.A., Yang, Z.-L., 2003. The Common Land Model. Bull. Am. Meteorol. Soc. 84, 1013–1023. http://dx.doi.org/10.1175/BAMS-84-8-1013.

Dickinson, R., Oleson, K., 2006. The Community Land Model and its climate statistics as a component of the Community Climate System Model. J. Clim. 2302–2324.

Donner, S.D., Kucharik, C.J., 2003. Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin. Glob. Biogeochem. Cycl. 17. http://dx.doi.org/10.1029/2001GB001808.

Fulton, J.P., Wells, L.G., Barnhisel, R., 1996. Spatial variation of soil physical properties: a precursor to precision tillage. ASAE Paper No. 961002. Phoenix, Arizona.

Gervois, S., Ciais, P., de Noblet-Ducoudré, N., Brisson, N., Vuichard, N., Viovy, N., 2008. Carbon and water balance of European croplands throughout the 20th century. Glob. Biogeochem. Cycl. 22http://dx.doi.org/10.1029/2007GB003018. n/a-–n/a.

June, T., 2011. Analysis of CO2 fluxes: inclusion of Wall conductance (Gw) on the estimation of Rubisco activity, VCMAX of soybean leaves. HAYATI J. Biosci. 18, 43–50. http://dx.doi.org/10.4308/hjb.18.1.43.

Kattge, J., Knorr, W., Raddatz, T., Wirth, C., 2009. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. Glob. Change Biol. 15, 976–991. http://dx.doi.org/10.1111/j.1365-2486.2008.01744.x.

Kucharik, C., Twine, T., 2007. Residue, respiration, and residuals: evaluation of a dynamic agroecosystem model using eddy flux measurements and biometric data. Agric. For. Meteorol. 146, 134–158. http://dx.doi.org/10.1016/j.agrformet.2007.05.011.

Kucharik, C.J., 2003. Evaluation of a process-based agro-ecosystem model (Agro-IBIS) across the U.S. Corn belt: simulations of the interannual variability in maize yield. Earth Interact. 7, 1–33. http://dx.doi.org/10.1175/1087-3562(2003)007<0001:EOAPAM>2.0.CO;2.

Kucharik, C.J., Brye, K.R., 2003. Integrated biosphere simulator (IBIS) yield and nitrate loss predictions for Wisconsin maize receiving varied amounts of nitrogen fertilizer. J. Environ. Qual. 32, 247–268. http://dx.doi.org/10.2134/jeq2003.2470.

Kutsch, W.L., Aubinet, M., Buchmann, N., Smith, P., Osborne, B., Eugster, W., Wattenbach, M., Schrumpf, M., Schulze, E.D., Tomelleri, E., Ceschia, E., Bernhofer, C., Béziat, P., Carrara, a., Di Tommasi, P., Grunwald, T., Jones, M., Magliulo, V., Marloie, O., Moureaux, C., Olioso, a., Sanz, M.J., Saunders, M., Søgaard, H., Ziegler, W., 2010. The net biome production of full crop rotations in Europe. Agric. Ecosyst. Environ. 139, 336–345. http://dx.doi.org/10.1016/j.agee.2010.07.016.

Levis, S., Bonan, G.B., Kluzek, E., Thornton, P.E., Jones, A., Sacks, W.J., Kucharik, C.J., 2012. Interactive crop management in the Community Earth System Model (CESM1): seasonal influences on land–atmosphere fluxes. J. Clim. 25, 4839–4859. http://dx.doi.org/10.1175/JCLI-D-11-00446.1.

Loescher, H., Ayres, E., Duffy, P., Luo, H., Brunke, M., 2014. Spatial variation in soil properties among North American ecosystems and guidelines for sampling designs. PLoS One 9. http://dx.doi.org/10.1371/journal.pone.0083216.

Lokupitiya, E., Denning, S., Paustian, K., Baker, I., Schaefer, K., Verma, S., Meyers, T., Bernacchi, C.J., Suyker, A., Fischer, M., 2009. Incorporation of crop phenology in Simple Biosphere Model (SiBcrop) to improve land-atmosphere carbon exchanges from croplands. Biogeosciences 6, 969–986. http://dx.doi.org/10.5194/bg-6-969-2009.

Long, S.P., Bernacchi, C.J., 2003. Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. J. Exp. Bot 54, 2393–2401. http://dx.doi.org/10.1093/jxb/erg262.

Mourtzinis, S., Specht, J.E., Lindsey, L.E., Wiebold, W.J., Ross, J., Nafziger, E.D., Kandel, H.J., Mueller, N., Devillez, P.L., Arriaga, F.J., Conley, S.P., 2015. Climate-induced reduction in US-wide soybean yields underpinned by region- and in-season-specific responses. Nat. Plants 1, 8–11. http://dx.doi.org/10.1038/nplants.2014.26.

Mzuku, M., Khosla, R., Reich, R., Inman, D., Smith, F., MacDonald, L., 2005. Spatial variability of measured soil properties across site-specific management zones. Soil Sci. Soc. Am. J. 69, 1572. http://dx.doi.org/10.2136/sssaj2005.0062.

Neale, R.B., Richter, J.H., Conley, A.J., Park, S., Lauritzen, P.H., Gettelman, A., Williamson, D.L., Rasch, P.J., Vavrus, S.J., Taylor, M.A., Collins, W.D., Zhang, M., Lin, S.-J., 2010. Description of the NCAR Community Atmosphere Model (CAM 4.0). Ncar/Tn-485+Str.

Oleson, K.W., Lawrence, D.M., Gordon, B., Flanner, M.G., Kluzek, E., Lawrence, P., Levis, S., Swenson, S.C., Thornton, P.E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C.L., Hoffman, F., Lamarque, J.F., Mahowald, N., Niu, G., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z.L., Zeng, X.D., Zeng, X.B., 2010. Technical Description of Version 4.0 of the Community Land Model (CLM). National Center for Atmospheric Research, Boulder, Colorado.

Oleson, K.W., Niu, G.-Y., Yang, Z.-L., Lawrence, D.M., Thornton, P.E., Lawrence, P.J., Stöckli, R., Dickinson, R.E., Bonan, G.B., Levis, S., Dai, A., Qian, T., 2008. Improvements to the Community Land Model and their impact on the hydrological cycle. J. Geophys. Res. 113, G01021. http://dx.doi.org/10.1029/2007JG000563.

Osborne, T.M., Lawrence, D.M., Challinor, A.J., Slingo, J.M., Wheeler, T.R., 2007. Development and assessment of a coupled crop-climate model. Glob. Change Biol. 13, 169–183. http://dx.doi.org/10.1111/j.1365-2486.2006.01274.x.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J.G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P.E., Li, J., Lin, X., Lomas, M.R., Lu, M., Luo, Y., Ma, Y., Myneni, R.B., Poulter, B., Sun, Z., Wang, T., Viovy, N., Zaehle, S., Zeng, N., 2013. Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends. Glob. Change Biol. 19, 2117–2132. http://dx.doi.org/10.1111/gcb.12187.

Richardson, A.D., Hollinger, D.Y., Aber, J.D., Ollinger, S.V., Braswell, B.H., 2007. Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange. Glob. Change Biol. 13, 788–803. http://dx.doi.org/10.1111/j.1365-2486.2007.01330.x.

Sargsyan, K., Safta, C., Najm, H.N., Debusschere, B.J., Ricciuto, D., Thornton, P., 2014. Dimensionality reduction for complex models via Bayesian compressive sensing. Int. J. Uncertain. Quantif. 4, 63–93. http://dx.doi.org/10.1615/Int.J.UncertaintyQuantification.2013006821.

Schimel, D.S., House, J.I., Hibbard, K.a, Bousquet, P., Ciais, P., Peylin, P., Braswell, B.H., Apps, M.J., Baker, D., Bondeau, a, Canadell, J., Churkina, G., Cramer, W., Denning, a S., Field, C.B., Friedlingstein, P., Goodale, C., Heimann, M., Houghton, R.a, Melillo, J.M., Moore, B., Murdiyarso, D., Noble, I., Pacala, S.W., Prentice, I.C., Raupach, M.R., Rayner, P.J., Scholes, R.J., Steffen, W.L., Wirth, C., 2001. Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. Nature 414, 169–172. http://dx.doi.org/10.1038/35102500.

Schwalm, C.R., Williams, Ca., 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. 115, G00H. http://dx.doi.org/10.1029/2009JG001229.

Song, Y., Jain, A.K., McIsaac, G.F., 2013. Implementation of dynamic crop growth processes into a land surface model: evaluation of energy, water and carbon fluxes under corn and soybean rotation. Biogeosciences 10, 8039–8066. http://dx.doi.org/10.5194/bg-10-8039-2013.

Stehfest, E., Heistermann, M., Priess, J.A., Ojima, D.S., Alcamo, J., 2007. Simulation of global crop production with the ecosystem model DayCent. Ecol. Modell. 209, 203–219. http://dx.doi.org/10.1016/j.ecolmodel.2007.06.028.

Stöckli, R., Lawrence, D.M., Niu, G.-Y., Oleson, K.W., Thornton, P.E., Yang, Z.-L., Bonan, G.B., Denning, A.S., Running, S.W., 2008. Use of FLUXNET in the Community Land Model development. J. Geophys. Res. 113, 1–19. http://dx.doi.org/10.1029/2007JG000562.

Suyker, A.E., Verma, S.B., 2012. Gross primary production and ecosystem respiration of irrigated and rainfed maize-soybean cropping systems over 8 years. Agric. For. Meteorol. 165, 12–24. http://dx.doi.org/10.1016/j.agrformet.2012.05.021.

Suyker, A.E., Verma, S.B., 2010. Coupling of carbon dioxide and water vapor exchanges of irrigated and rainfed maize–soybean cropping systems and water productivity. Agric. For. Meteorol. 150, 553–563. http://dx.doi.org/10.1016/j.agrformet.2010.01.020.

Taylor, K., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106, 7183–7192.

Thornton, P.E., Rosenbloom, N.A., 2005. Ecosystem model spin-up: estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. Ecol.

Modell. 189, 25–48. http://dx.doi.org/10.1016/j.ecolmodel.2005.04.008.
Tsvetsinskaya, E.A., Mearns, L.O., Easterling, W.E., 2001. Investigating the effect of seasonal plant growth and development in three-dimensional atmospheric simulations. Part I: simulation of surface fluxes over the growing season. J. Clim. 14, 692–709. http://dx.doi.org/10.1175/1520-0442(2001)014<0692:ITEOSP>2.0.CO;2.
Twine, T.E., Bryant, J.J., T Richter, K., Bernacchi, C.J., McConnaughay, K.D., Morris, S.J., Leakey, A.D.B., 2013. Impacts of elevated CO2 concentration on the productivity and surface energy budget of the soybean and maize agroecosystem in the Midwest USA. Glob. Change Biol. 19, 2838–2852. http://dx.doi.org/10.1111/gcb.12270.
Twine, T.E., Kucharik, C.J., 2009. Climate impacts on net primary productivity trends in natural and managed ecosystems of the central and eastern United States. Agric. For. Meteorol. 149, 2143–2161. http://dx.doi.org/10.1016/j.agrformet.2009.05.012.
Twine, T.E., Kucharik, C.J., Foley, J.A., 2004. Effects of land cover change on the energy and water balance of the Mississippi River Basin. J. Hydrometeorol. 5, 640–655. http://dx.doi.org/10.1175/1525-7541(2004)005<0640:EOLCCO>2.0.CO;2.
Van Wart, J., Grassini, P., Yang, H., Claessens, L., Jarvis, A., Cassman, K.G., 2015. Creating long-term weather data from thin air for crop simulation modeling. Agric. For. Meteorol. 209–210, 49–58. http://dx.doi.org/10.1016/j.agrformet.2015.02.020.
Wang, T., Brender, P., Ciais, P., Piao, S., Mahecha, M.D., Chevallier, F., Reichstein, M., Ottlé, C., Maignan, F., Arain, A., Bohrer, G., Cescatti, A., Kiely, G., Law, B.E., Lutz, M., Montagnani, L., Moors, E., Osborne, B., Panferov, O., Papale, D., Vaccari, F.P., 2012. State-dependent errors in a land surface model across biomes inferred from eddy covariance observations on multiple timescales. Ecol. Modell. 246, 11–25. http://dx.doi.org/10.1016/j.ecolmodel.2012.07.017.
Webler, G., Roberti, D.R., Cuadra, S.V., Moreira, V.S., Costa, M.H., 2012. Evaluation of a dynamic agroecosystem model (Agro-IBIS) for soybean in Southern Brazil. Earth Interact. 16, 1–15. http://dx.doi.org/10.1175/2012EI000452.1.
Willmott, C.J., Robeson, S.M., Matsuura, K., 2012. A refined index of model performance. Int. J. Climatol. 32, 2088–2094. http://dx.doi.org/10.1002/joc.2419.
Wu, X., Vuichard, N., Ciais, P., Viovy, N., de Noblet-Ducoudré, N., Wang, X., Magliulo, V., Wattenbach, M., Vitale, L., Di Tommasi, P., Moors, E.J., Jans, W., Elbers, J., Ceschia, E., Tallec, T., Bernhofer, C., Grünwald, T., Moureaux, C., Manise, T., Ligne, A., Cellier, P., Loubet, B., Larmanou, E., Ripoche, D., 2015. ORCHIDEE-CROP (v0), a new process based Agro-Land Surface Model: model description and evaluation over Europe. Geosci. Model Dev. Discuss. 8, 4653–4696. http://dx.doi.org/10.5194/gmdd-8-4653-2015.
Xu, H., Twine, T.E., Girvetz, E., Iqbal, M., Goheer, M., Khan, A., Moriondo, M., Giannakopoulos, C., Bindi, M., Challinor, A., Simelton, E., Fraser, E., Hemming, D., Collins, M., Lobell, D., Hammer, G., McLean, G., Messina, C., Roberts, M., Schlenker, W., Harding, K., Snyder, P., Pryor, S., Barthelmie, R., Schoof, J., Sage, R., Kubien, D., Lobell, D., Schlenker, W., Costa-Roberts, J., Lobell, D., Field, C., Lobell, D., Asner, G., Twine, T., Kucharik, C., Ruiz-Vera, U., Siebers, M., Drag, D., Ort, D., Bernacchi, C., Kucharik, C., Serbin, S., Parry, M., Rosenzweig, C., Iglesias, A., Livermore, M., Fischer, G., Kucharik, C., Twine, T., Kucharik, C., Brye, K., Kucharik, C., Donner, S., Kucharik, C., Twine, T., Kucharik, C., Twine, T., Bryant, J., Richter, K.T., Bernacchi, C., Mcconnaughay, K., Morris, S., VanLoocke, A., Twine, T., Zeri, M., Bernacchi, C., Sacks, W., Kucharik, C., Taylor, K., Stouffer, R., Meehl, G.a, Maloney, E., Camargo, S., Chang, E., Colle, B., Fu, R., Geil, K., Vuuren, D., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Moss, R., Edmonds, J., Hibbard, K., Manning, M., Rose, S., Vuuren, D., van, Harris, I., Jones, P., Osborn, T., Lister, D., Gudmundsson, L., Bremnes, J., Haugen, J., Engen-Skaugen, T., Boé, J., Terray, L., Habets, F., Martin, E., Ummenhofer, C., Xu, H., Twine, T., Girvetz, E., McCarthy, H., Chhetri, N., Foley, J., Prentice, I., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., Kucharik, C., Foley, J., Delire, C., Fisher, V., Coe, M., Lenters, J., Collatz, G., Ball, J., Grivet, C., Berry, J., Farquhar, G., Sharkey, T., Miller, D., White, R., Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Duvick, D., Schlenker, W., Roberts, M., Dale, R., Shaw, R., Parry, M., Rosenzweig, C., Livermore, M., Arnell, N., Cannell, M., Hulme, M., Kovats, R., Mitchell, J., Nicholls, R., Challinor, A., Watson, J., Lobell, D., Howden, S., Smith, D., Chhetri, N., Gutmann, E., Pruitt, T., Clark, M., Brekke, L., Arnold, J., Raff, D., Challinor, A., Wheeler, T., Craufurd, P., Ferro, C., Stephenson, D., Manderscheid, R., Erbs, M., Weigel, H.-J., Markelz, R., Strellner, R., Leakey, A., 2016. Climate change and maize yield in Iowa. PLoS One 11, e0156083. http://dx.doi.org/10.1371/journal.pone.0156083.
Yuan, X., Liang, X.-Z., 2011. Evaluation of a conjunctive surface–subsurface process model (CSSP) over the contiguous United States at regional–local scales. J. Hydrometeorol. 12, 579–599. http://dx.doi.org/10.1175/2010JHM1302.1.
Zeng, X., Shaikh, M., Dai, Y., 2002. Coupling of the common land model to the NCAR community climate model. J. Clim. 15, 1832–1854.