

Joint modeling of crop and irrigation in the Central United States using the Noah-MP land surface model

Zhe Zhang^{1,2}, Michael Barlage³, Fei Chen³, Yanping Li^{1,2}*, Warren Helgason^{1,4}, Xiaoyu Xu⁵, Xing Liu⁶, Zhenhua Li^{1,2}

1 Global Institute for Water Security, University of Saskatchewan, Saskatoon, SK, Canada

2 School of Environment and Sustainability, University of Saskatchewan, Saskatoon, SK, Canada

3 Research Application Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

4 College of Engineering, University of Saskatchewan, Saskatoon, SK, Canada

5 College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Jiangjun Road #29, Nanjing, 211106, China

6 College of Agriculture, Purdue University, Indiana, USA

Abstract

Representing climate-crop interactions is critical to earth system modeling. Despite recent progress in modeling dynamic crop growth and irrigation in land surface models (LSMs), transitioning these models from field to regional scales is still challenging. This study applies the Noah-MP LSM with dynamic crop-growth and irrigation schemes to jointly simulate the crop yield and irrigation amount for corn and soybean in the central U.S. The model performance of crop yield and irrigation amount are evaluated at county-level against the USDA reports and USGS water withdrawal data, respectively. The bulk simulation (with uniform planting/harvesting management and no irrigation) produces significant biases in crop yield estimates for all planting regions, with root-mean-square-errors (RMSEs) being 28.1% and 28.4% for corn and soybean, respectively. Without an irrigation scheme, the crop yields in the irrigated regions are reduced due to water stress with RMSEs of 48.7% and 20.5%. Applying a dynamic irrigation scheme effectively improves crop yields in irrigated regions and reduces RMSEs to 22.3% and 16.8%. In rainfed regions, the model overestimates crop yields. Applying spatially-varied planting and harvesting dates at state-level reduces crop yields and irrigation amount for both crops, especially in northern states. A “nitrogen-stressed” simulation is conducted and found that the improvement of irrigation on crop yields are limited when the crops are under nitrogen stress. Several uncertainties in modeling crop growth are identified, including yield-gap, planting date, rubisco capacity, and discrepancies between available datasets, pointing to future efforts to incorporating spatially-varying crop parameters to better constrain crop growing seasons.

Plain Language Summary

Modeling dynamic crop growth and irrigation processes in the earth system are critical to the understanding of climate-crop interaction and water availability for food security. While many of the existing models and parameters are developed at local sites, it is challenging to transition them to large regions. This study conducts a joint modeling effort of crop growth and irrigation in the central U.S. and focuses on transitioning model parameters. The results show that irrigation could significantly improve crop yields in the irrigated regions. By using spatially-varying planting and harvesting date, the model shows a better estimate for both crop yield and irrigation amount. A summary of model parameter uncertainties is provided. This urges future developments on detailed spatial crop data and further understanding of crop photosynthesis to have better constrain on model results.

Introduction

This study intends to extend the investigation of Xu et al. (2019), which focused on the transition of dynamic irrigation modeling from field to regional scales, by assessing the benefits and uncertainties in joint crop-growth and irrigation modeling in the context of capturing climate-crop-irrigation interactions in Earth System Models (ESMs). It has been recognized that climate change and variability play a major role in affecting crop production (Drewniak et al., 2013; Ray et al., 2015; Leng et al., 2016) from regional to global scales (Leng et al., 2016). Climate change has already impacted global agricultural production (Ray et al., 2019), and negative trends on crop yield per degree warming have been projected for major cultivars across the globe (National Research Council, 2011). In addition to mean climatic conditions, extreme climate events, such as drought and flooding, have also been emphasized as an important contributor to crop yield reduction (Hlavinka et al., 2009; Lobell et al., 2014).

Agriculture management activities such as irrigation and fertilization also play an essential role in increasing crop yields, especially in semi-arid climates and regions with strong seasonal variability of precipitation during crop reproductive stages (Grassini et al., 2009). Globally, ~20% of croplands are irrigated and contribute to ~40% of the world's food production (Siebert and Doll, 2010). Over the 55.8 million acres of irrigated U.S. farmland (as of 2012), 115 billion gallons of water was withdrawn for irrigation per day, accounting for more than one third of water-use nationwide in 2015 (Maupin et al., 2014; Dieter et al., 2018). Furthermore, agriculture is challenged to make efficient use of water to offset climate change impacts on freshwater availability and groundwater over-exploitation (Vorosmarty et al., 2000). Therefore, understanding the capability for freshwater to supply the world's major food production, such as in the U.S. Great Plain and Canadian Prairies, under climate change background, has become an overarching science goal in the Global Energy and Water Exchanges project (GEWEX, Grand Challenge on Water for the Food Baskets of the World: <https://www.gewex.org/about/science/wcrps-grand-challenges/water-for-the-food-baskets-of-the-world/>).

Agricultural management modifies surface water and energy balances, alters characteristics of land-atmosphere interactions, and hence impacts local and regional climate (Pielke et al., 2007). Furthermore, irrigation practices have been shown to increase humidity and decrease air temperature (Chen et al., 2018; Xu et al., 2019). This irrigation-cooling effect has shown to modify local environment, regional precipitation, and even reduce the chance of extreme heatwaves in the U.S. (Lu et al., 2015) and globally (Thiery et al., 2017).

To better understand the climate change, crop yield and freshwater nexus, as well as critical cropland-atmosphere interactions, it is important and necessary to improve the representation of dynamic crop growth and irrigation in ESMs. Recent efforts have been dedicated to implement crop growth dynamics and agricultural management into land surface models (LSM) within ESMs (Levis et al., 2012; Drewniak et al., 2013; Liu et al., 2016; Leng et al. 2016; McDermid et al., 2016). For instance, crop growth models were introduced into the Community Land Model version 4 with carbon-nitrogen cycle (CLM4CN) by Levis et al. (2012), which focused on the crop coverage in mid-latitude regions. The results showed improvement on simulating leaf area index (LAI), an index for crop growth, and summer precipitation, compared to the default setting of CLM4.5. This work also highlights the importance of accurate representation of the cropping

calendar, as a “late-planting” sensitivity test improved the simulated annual cycle of net ecosystem exchange (NEE) in midwestern North America. More recently, a dynamic crop growth model was incorporated into the Noah with multiple-physics (Noah-MP, Niu et al. 2011) model and tested for two field sites in Illinois and Nebraska for corn and soybean (Liu et al., 2016). In Noah-MP-Crop, crop growth stages are solely dependent on growing degree days (GDD). The Noah-MP-Crop model improved the simulation of surface energy balance and LAI and provided reasonable estimates of biomass. While these works demonstrated widespread potential for agriculture-climate interactions in some key agroecology regions, it is still challenging to accurately represent crop-climate-hydrology interactions in general and specifically the spatial variations of crop-model parameters across various scales.

Similarly, irrigation parameterizations have been incorporated into various LSMs using the “soil moisture deficit” approach. For example, Ozdonga et al. (2010) used the soil field capacity as a threshold, below which irrigation is triggered, and calculated the irrigation demand from subtracting current root-zone soil moisture from field capacity. Lawston et al. (2015) applied this soil moisture deficit approach in the coupled Weather Research and Forecast (WRF) model and found the regional climate is highly sensitive to the irrigation method chosen (drip, flood, and sprinkler). Xu et al. (2019) used a similar approach to mimic sprinkler irrigation at the county level in the central U.S. Instead of using a uniform value of field capacity, a spatially-varying soil moisture threshold parameter is determined through regional calibration against the USGS water withdrawal data, which enables transforming model parameters from field to regional scale.

The above-mentioned crop-focused and irrigation-focused modeling approaches are inadequate to comprehensively address climate-crop-water interactions. In crop-focused models, a significant amount of irrigation water as important input to the surface-water-budget equation is neglected in semi-arid croplands and will result in a warm/dry surface environment through land-atmosphere interactions, as well as loss in crop yield due to water stress. On the other hand, irrigation-only models fail to capture the feedback between irrigation water demand and crop growth stages. Therefore, regional irrigation modeling will benefit from the dynamic representation of crop heterogeneity, such as constraining simulated irrigation amount by crop planting/harvest date. Thus, it is necessary to perform joint crop-irrigation modeling in LSMs.

Leng et al. (2016) provided the first joint modeling effort with crop and irrigation on large-scale in the U.S., and optimized irrigation and fertilization practice in CLM4.5CN. The results showed that without optimization, the corn yield is much underestimated, due to the quick denitrification in CLM4.5CN previously reported by Oleson et al. (2013). The irrigation optimization increases yield only in the irrigated region and the fertilization optimization showed significant improvement in all regions. However, the improvement of irrigation scheme on crop yield under sufficient nutrition condition is not discussed. Moreover, uncertainties associated with crop model parameters, sparse agricultural datasets at both spatial and temporal scales, and even discrepancies between available datasets still remain unsolved.

Given the wide use of Noah-MP LSM in the community WRF model and in the operational National Water Model (NWM), it is important to understand and improve its capability in simulating concurrently crop growth and irrigation, because both processes affect surface heat and water-vapor fluxes (as lower boundary conditions in WRF) and streamflow. Therefore, the primary

141 objectives of this study are to: (1) assess the Noah-MP model's performance in joint crop and
142 irrigation modeling; (2) investigate methods of transforming irrigation and crop modeling from
143 field to regional scales; and (3) identify uncertainties and challenges in crop modeling in LSMs.
144 We focus on two crops (corn and soybean) in this study, since they are the two crops currently
145 represented in Noah-MP-Crop and are two major field crops in the central U.S. Section 2
146 introduces the data required for model input and evaluation, and the Noah-MP crop and irrigation
147 schemes. The model results for crop yield and irrigation amount are presented in Section 3. The
148 uncertainties in simulating crop yield are discussed in Section 4. We conclude our findings in
149 Section 5.

2. Description of input data, evaluation data, and models

2.1 Data Preparation

In this work, several agriculture management datasets are used to help constrain crop and irrigation models and to define the crop growing season, cultivated land fraction, and irrigated fractions. The planted area for corn and soybean are obtained from the 30-m CropScape data from the U.S. Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS)/George Mason University (GMU) (<https://nassgeodata.gmu.edu/CropScape/>). This is a geo-referenced, crop-specific land cover data layer created for the contiguous U.S. using satellite imagery and has been supported by extensive agricultural ground truthing. The CropScape dataset is originally derived from the planting frequency in 11 years (from 2008 to 2018) and used to calculate the fractional coverage of total cropland (relative to the grid cell's vegetated area; hereafter F_{crop}) and of each crop type (relative to the grid cell's total cropland area; F_{corn} and $F_{soybean}$). In this study, the planting areas are determined on two criteria: (1) the $F_{crop} > 0.5$; and (2) F_{corn} or $F_{soybean} > 0.3$, for corn and soybean, respectively. The planting area for these two crops and their planting fraction are shown in Figure 1.

Figure 1. Planted-area fractions for (a) corn and (b) soybean in the Central U.S. domain derived from the USDA-NASS CropScape dataset.

The 2010 USDA report on usual planting and harvesting dates is used to define the length of growing season for corn and soybean. This survey reports the most active period of usual planting and harvesting dates for each state. In our study, the middle dates of planting and harvest windows are selected for the states within our study domain (see Figure 2). Although the middle dates for each crop in each state may not reflect the complex decision of actual planting and harvesting, it represents to some degree the spatial variation of planting and harvesting at state-level. The impacts of uncertainties in planting/harvesting dates on simulated crop yield and irrigation amount are discussed in section 3.2. For details of the planting and harvesting dates in each state, please see Appendix A.

Figure 2. USDA-NASS state-level planting and harvest dates in Julian day for corn and soybean.

For each year, the USDA NASS reports the average yields for various crops at the county-level over the U.S (<https://quickstats.nass.usda.gov/>). These data are based on harvested yields, reported by a sample of farmers within each county, and verified with independent yield samples taken by USDA staff when the crop reaches maturity (FAO and DWFI, 2015). Therefore, the model simulated biomass (g/m^2) will need to be converted to standard yield (*bushel/acre, bu/ac*) to compare with the USDA county-level data, following the instruction: (see <http://www.ag.ndsu.edu/pubs/plantsci/crops/ae905w.htm>)

$$\text{corn yield [bu/ac]} = \text{biomass}[g/m^2] * (1 - 0.155) * 4.046[km^2/ac] / 25.4[kg/bu] \quad (1)$$

$$\text{soybean yield [bu/ac]} = \text{biomass}[g/m^2] * (1 - 0.13) * 4.046[km^2/ac] / 27.4[kg/bu] \quad (2)$$

In the Eq. (1) and (2), 0.155 and 0.13 are the standard moisture content (15.5% and 13%) for corn and soybean, respectively. Harvested corn usually contain an initial moisture content greater than 15.5% (15.5~32%). For transportation and storage purpose, mechanical drying method is typically applied to reduce the initial moisture to the standard moisture. Two sources of weight loss are associated with this process: 1) the weight of the moisture loss (also known as “water shrink”) and

2) the weight loss due to handling processes (Hicks and Cloud, 1992). The handling loss could range from 0.04% to 5.22%, depending on the initial moisture content and shrinkage loss. Therefore, the calculated dry mass losses tend to be variable among different growers. This uncertainty is worth noting when comparing the model simulated dry mass with standard yield in the USDA survey.

The irrigation locations are defined by the 500-m MODIS-based irrigation fraction map (Ozdogan and Gutman, 2008) and the critical irrigation threshold parameter, IRR_CRI, from Xu et al. (2019) is applied in this study (see Figure 3). IRR_CRI is a threshold parameter for the soil water content, below which the irrigation scheme will be activated and was calibrated at county-level in Xu et al. (2019). To evaluate the model irrigation amount, the five-year report from the U.S. Geological Survey (USGS) on fresh water withdrawals for irrigation (<http://water.usgs.gov/watuse/>) is used to constrain and calibrate the irrigation parameters in the irrigation module (for details of irrigation modeling, see section 2.3 and Xu et al., 2019).

Two Ameriflux sites with irrigated agriculture (Ne1 and Ne2 in Mead, NE; <https://ameriflux.lbl.gov/sites/>) are analyzed (Suyker, 2001). Ne1 is an irrigated continuous maize site and Ne2 is an irrigated maize-soybean rotation site. Data collected at the Ameriflux sites, including LAI, leaf mass per area (LMA), and harvested biomass, are used to evaluate the model output at these two locations with and without the irrigation scheme. Also, the measured leaf biomass per area (LMA; g/m^2) is equivalent to the Noah-MP-Crop parameter that converts biomass to LAI (BIO2LAI), which is assumed to be a constant.

Figure 3. (a) The irrigation fraction used in this study. (b) The critical irrigation threshold parameter used in this study, calibrated in Xu et al. (2019).

2.2 Noah-MP-Crop model

Noah-MP is a land component of the Weather Research and Forecast (WRF) model (Skamarock et al., 2008; Niu et al., 2011; Yang et al., 2011), which has been widely applied in numerical weather prediction (NWP), regional climate and hydrology studies (Liu et al., 2017; Barlage et al., 2015; Zhang et al., 2020). It has been also used to simulate the land surface processes for streamflow forecasts in the National Water Model (www.water.noaa.gov/about/nwm).

The Noah-MP-Crop crop module consists of three components: a photosynthesis (PSN)-stomata scheme, a carbon allocation scheme, and a dynamic crop growth scheme. The leaf-level PSN rate and stomatal conductance are calculated based on the model of Farquhar et al. (1980) and Collatz et al. (1992) for C3 and C4 plants, respectively. However, there is only one set of PSN parameters for a generic C3 crop in the default Noah-MP. This simplified treatment doesn't represent corn (C4), a major productive species in Central U.S. Therefore, in this study, a set of C4 PSN parameters are adapted from a synthesis of literature and model sensitivity tests (see Appendix B).

Following a similar approach used in traditional crop models (Hybrid-Maize for corn, Yang et al., 2004; DSSAT for soybean, the Decision Support System for Agrotechnology Transfer, Jones et al., 2003), the dynamic crop growth model in Noah-MP-Crop uses the accumulated growing degree days (GDD) to determine eight plant growth stages (PGS, Liu et al., 2016): before seeding, emergence, initial vegetative, normal vegetative, initial reproductive, to maturity, after maturity, and after harvesting. Also in Liu et al (2016), the dynamic crop growth parameters, such as planting/harvest dates and GDD-based thresholds to determine plant growth stages are calibrated at two Ameriflux sites in Bondville (Bo1), IL, for corn and Mead (Ne3), NE, for soybean.

Finally, the Noah-MP-Crop model allocates the assimilated carbohydrate to different parts of plant, depending on the growth stages. For each stage, the total carbohydrate from the PSN scheme is partitioned to the leaf, stem, root and grain according to stage-function fraction parameters (from 0 to 1). For example, during the vegetative stage, more carbon is allocated to leaf relative to stem and root; while in the reproductive stage, most of the assimilated carbon is allocated to grain. Then, the simulated leaf biomass is converted to LAI based on a model parameter, BIO2LAI (or specific leaf area, SLA), in the following equation:

$$LAI = Leaf_{mass} * BIO2LAI \quad (3)$$

The values of BIO2LAI are constants and are different for corn (0.015) and soybean (0.030), respectively (Liu et al., 2016).

2.3. Irrigation scheme

A dynamic irrigation scheme was integrated into Noah-MP and tested at field and regional scales without using the Noah-MP-Crop model (Xu et al. 2019). In this study, we adopt the same approach and couple it with dynamic crop growth, enabling two-way crop-irrigation interactions.

Plant photosynthesis and respiration processes are limited by water stress during droughts. Therefore, irrigation plays a critical role in both the water and carbon cycle through relieving water stress, especially for crops planted in arid and semi-arid regions. In Noah-MP, the water stress function is plant- and soil-dependent and is determined by the integrated soil moisture availability (SMA) in root zones. As in Xu et al. (2019), the root-zone SMA is also employed as a basic irrigation trigger. For the irrigated cropland, the root-zone SMA is defined as the ratio of the current root-zone available soil moisture (current $SM - SM_{wlt}$, wilting point) and non-stress soil moisture ($SM_{ref} - SM_{wlt}$):

$$SMA = (SM - SM_{wlt}) / (SM_{ref} - SM_{wlt}) \quad (5)$$

The irrigated cropland is defined as the fraction within a cultivated grid cell ($F_{irr-crop}$) and takes the smaller value of F_{irr} and $F_{crop} \cdot F_{veg}$ (cropland fraction relative to the model grid cell's total area) in Figure 3(a):

$$F_{irr-crop} = \min(F_{irr}, F_{crop} \cdot F_{veg}) \quad (6)$$

The irrigation triggering mechanism includes: (1) $F_{irr-crop} > IRR_FRC$ (an irrigation fraction threshold); (2) within the growing season, defined by the planting/harvesting date map above; (3) $SMA < IRR_CRI$ (soil moisture trigger, see Figure 3(b)); and (4) stop irrigation on rainy days. These criteria are checked daily, and if irrigation is triggered, the potential irrigation amount for the day (IWA) is computed to maintain SMA to a non-stress level (SM_{ref}): $IWA = \min(SM_{ref} - SM, IRR_LIM)$, where IRR_LIM is the daily maximum irrigation amount, which is limited by the capability of the irrigation system and water availability.

The above irrigation scheme would be executed for the crop type in each irrigated grid cell to obtain the irrigation water amount for corn (IWA_{corn}) and soybean ($IWA_{soybean}$), respectively.

2.4 Model setup

The model domain is identical to the central U.S. domain in Xu et al. (2019). The model domain is 600 grids (north-south) \times 700 grids (west-east) at 4-km resolution, covering the major part of the corn-belt in the Central U.S. The simulation period ranges from 1999-10-01 to 2004-12-31, covering five growing seasons. The atmospheric forcing data are from the North American Land Data Assimilation System (NLDAS, Cosgrove et al., 2003) forcing dataset at 0.125-degree and hourly resolutions. The precipitation forcing are generated by combining observations from field stations, Stage IV radar retrievals from Next Generation Weather Radar System and satellite. A 10-year spin-up period was used to ensure the soil moisture and temperature reach an equilibrium state. An elevation adjustment was applied to the surface pressure, longwave radiation, near-surface temperature and humidity fields to account for topography differences between the model and NLDAS grids.

Six experiments were performed to assess Noah-MP's performance in joint crop-irrigation modeling (see Table 1). The first experiment (BULK) is a simulation with dynamic crop but without irrigation, in which a uniform planting and harvest date is applied in the whole domain. It adopts the default planting/harvest date (day of year) initially calibrated for corn in Bondville, IL, and soybean in Mead, NE (for corn: Julian day 111/300; for soybean: Julian day 130/280). The second experiment (BULK_IRR) is the same as BULK but with the calibrated dynamic irrigation scheme activated (Xu et al., 2019). The third (STATE) and the fourth simulation (STATE_IRR) are the same as the BULK and BULK_IRR but used the state-level planting and harvest date as shown in Fig. 2. The BULK/BULK_IRR simulations were referred as the baseline simulations and the difference between BULK/BULK_IRR and STATE/STATE_IRR represents the impacts of spatially-varied planting/harvest date on crop yield and irrigation amount. The fifth (0.5N) and the sixth (0.5N_IRR) simulation are the same as STATE and STATE_IRR but reduce the nitrogen concentration by half. The difference between STATE/STATE_IRR and 0.5N/0.5N_IRR can be attributed to the impacts of nitrogen concentration. Furthermore, comparing the results between STATE_IRR and STATE with 0.5N_IRR and 0.5N will demonstrate the impacts of irrigation under N-sufficient and N-stressed conditions.

Table 1. Description of the Numerical Experiments.

3. Results

3.1. Model Performance

Figure 4 shows the county-level corn yields reported by USDA and results from the six experiments (five-year average from 2000 to 2004). Yield results from the BULK and STATE compare well with the USDA report in the magnitude and spatial pattern in the rainfed region but are underestimated in heavily irrigated regions such as Southeast Nebraska. Using the dynamic irrigation scheme in BULK_IRR and STATE_IRR reduces the yield bias in irrigated regions. The differences between the BULK and STATE will be further discussed in section 3.2. The 0.5N experiment significantly reduces yield for more than 60% of the domain due to nitrogen stress, which is similar to the CROP_DFLT scenario in Leng et al. (2016) for the fast denitrification in the default version of CLM4.5. In this case, using irrigation scheme (0.5N_IRR) has little improvement under nitrogen stress.

Figure 4. For Corn: Yield (bushel/acre) from USDA NASS county survey and six model simulations (five-year average from 2000-2004).

Figure 5. For Soybean: Yield (bushel/acre) from USDA NASS county survey and six model simulations (five-year average from 2000-2004).

As for soybean yields shown in Figure 5, BULK and STATE show good estimate of yield in the major soybean production areas in the U.S (MI, IL, IO, WI, MN, SD), but markedly underestimate the yield in the irrigated regions such as NE, AR and MS. In the 0.5N nitrogen-stressed condition, soybean yields are much under predicted for the entire domain. The dynamic irrigation scheme can help improve yield in the BULK_IRR and STATE_IRR simulation, but it doesn't show much impact under nitrogen stress condition in 0.5N_IRR. These results from corn and soybean suggest that the impacts of irrigation on yields in the irrigated regions are significant but only occur with sufficient fertilization supply.

3.2 Transition from field to regional scale crop modeling

The second objective of this study is to transition crop modeling from field to regional scale by first exploring the use of spatially-varying planting/harvesting dates for regional simulation. The impacts of spatially varying planting/harvest date on modeling crop yield and irrigation amount can be assessed by comparing the results from the BULK_IRR and STATE_IRR simulation, as shown in Figure 6. The bars are ranked by the yield from low to high in each of these states and the black lines represent the delayed days in planting date compared to the uniform planting date in BULK_IRR (111 for corn and 130 for soybean in Julian day). The delayed planting for each state implies a shorter growing season, which results in lower yields in STATE_IRR than in BULK_IRR for both corn and soybean. These reduced yields help improve the high bias of BULK_IRR in all states, except for South Dakota and Minnesota, where STATE_IRR underestimates in both corn and soybean yield.

Figure 6. Bar plot of the USDA and modeled yield for each state from the BULK_IRR and STATE_IRR simulation for (a) corn and (b) soybean (five-year average, 2000-2004). The delayed days in planting date in STATE_IRR (compared to the uniform date in BULK_IRR) are shown in black lines.

Figure 7 shows the impacts of delayed planting date on reduced yield (bu/ac/day) for corn and soybean. This impact of planting date on yield may be more complex than a linear relationship, but strong spatial variation exists across states on the sensitivity of modeled yield to delay in planting date. For both corn and soybean, a clear north-to-south gradient can be witnessed, as the impacts of planting date are strong in Northern states, such as Minnesota, Iowa, Wisconsin and Michigan. While for soybean, the planting region in lower Mississippi river valley shows a clear dependence on planting day as well. Moreover, this north-to-south gradient of yield dependence on planting date also exhibits in each particular state as well. This is most obvious in Minnesota, Iowa, Illinois, and Indiana, for both corn and soybean, that the modeled yield in northern part of the states are more sensitive to delay in planting date than in the south.

Figure 7. The impacts of delayed planting date on modeled yield (bu/ac/day) for (a) corn and (b) soybean.

In South Dakota the model shows very little sensitivity to the planting date, suggesting the modeled yield may be impacted by water stress (Figure S1 confirms this speculation that the underestimated yields in Eastern South Dakota and Western Minnesota are water-limited). However, the low irrigation fractions in these two regions (Figure 3a) suggested irrigation is not a significant water source for crop production. Therefore, we suspect that the perched shallow water table in the northern corn belt plays a role in supplying water for corn production (Rizzo et al., 2018). Note that the model applies a free drainage scheme for deep soil drainage and the complex two-way groundwater exchange processes are not considered in this study.

Transforming the planting date from uniform value at point scale to spatially-varied at state-level could also influence the modeled irrigation amount, as the irrigation period is constrained by the crop growing season. Figure 8 shows the spatial distribution of USGS water withdrawal report at county-level in 2000 and the modeled irrigation amount from the BULK_IRR and STATE_IRR. The BULK_IRR, with uniform planting/harvesting date, overestimates irrigation amount compared to the USGS reported data, especially in the Lower Mississippi River Basin (LMRB). The largest overestimation in irrigation amount is over 100 mm and occurs in Poinsett, Arkansas, with USGS reported 459.2 mm and the BULK_IRR simulated 561.3 mm. The overestimated

irrigation amount in the BULK_IRR has an intuitive explanation; the longer the growing season, the more water is needed to maintain soil moisture at the critical level. The scatter plot in Figure 9 for the irrigation amount from two simulations also confirms the overestimate of irrigation amount in the BULK_IRR, especially in the LMRB. After applying the spatially-varying planting/harvesting date, the performance in STATE_IRR is improved compared to the BULK_IRR (RMSEs improve from 29.67 to 26.24 mm, and coefficient of determination, R^2 , increases from 0.89 to 0.92) in LMRB. The STATE_IRR also reduces irrigation amount in Nebraska as well, but not as much as in LMRB. In fact, the USGS county-level report represents an upper bound of the total water withdrawal, but the water is not necessarily used all for irrigation. Therefore, the model simulated irrigation amount shouldn't exceed the USGS report. Hence, the STATE_IRR simulates less irrigation amount and provides better performance than the BULK_IRR.

Figure 8. Irrigation amount (mm) in 2000, from (a) USGS county-level water withdrawal report; (b) modeled irrigation amount from the BULK_IRR simulation; and (c) the STATE_IRR simulation.

Figure 9. Scatter plot of the model irrigation amount against the USGS water withdrawal data in two heavily irrigated region, Nebraska and Lower Mississippi River Basin (LMRB).

3.3 Impacts of irrigation on crop yield

Figure 10 shows the LAI and grain mass at the two Ameriflux sites (Ne1 and Ne2). STATE and STATE_IRR simulated LAI have good agreement in Ne1 for corn throughout the growing season, but underestimate LAI in Ne2 in 2002 for soybean. When it comes to the crop reproductive stage (grain production), the differences in yield between these two simulations are evident. The STATE simulation significantly underestimates corn yield at both sites, ranging from 31% to 80%, but using the irrigation scheme greatly improves corn yield at both sites.

As for the soybean yield, irrigation doesn't improve soybean yield as much as it did for corn yield, even with similar total irrigation amount. This is also noticed in Chen et al. (2018), as the increase in crop yield due to irrigation has a strong dependence on crop species. This may be attributed to the different biogeochemical characteristics between these two plants (corn is C4 and soybean is C3) in their water-use efficiency, including photosynthesis and respiration.

Figure 10. Timeseries of LAI and harvested grain in Ne1 and Ne2 sites from 2000 to 2005. Ne1 is irrigated continuous corn site and Ne2 is irrigated maize-soybean rotation; black boxes in Ne2 indicate soybean years.

Figure 11 shows the USDA yield data (five-year average) and the six simulations in this study, aggregated at state level. The comparison between BULK and BULK_IRR, and STATE and STATE_IRR in irrigated regions shows the improvement of yield with the irrigation scheme activated. The yield in BULK_IRR (156.5 bu/ac) is even double the amount than in BULK (74.61 bu/ac) for corn. The difference between BULK_IRR and STATE_IRR shows the impacts of prolonged growing season on overestimating modeled yield in BULK_IRR, due to the increase in modeled irrigation amount.

Figure 11. Bar plots of yield (five-year average) for (a) corn and (b) soybean from USDA survey and six simulations in this study. The red and blue bars represent the crop yields in the whole domain and in the irrigated region, respectively.

Moreover, the STATE_IRR and 0.5N_IRR represents the impacts of irrigation on crop yield under the conditions of sufficient and stressed nitrogen, respectively. The doubled irrigated yield in STATE_IRR (from 74.28 to 143.5 bu/ac) decreases under nitrogen stress condition (from 51.52 to 68.41 bu/ac) in 0.5N_IRR. This is similar to Leng et al. (2016) results, in which the irrigation scheme was applied to the default CLM4.5 run with fast denitrification rate. Thus, the irrigation impacts in such nitrogen-stressed conditions is limited. However, when the nitrogen concentration is unstressed, the impacts of irrigation manifest and improve crop yield.

Table 2 presents the statistics from all simulations, including RMSE (in both bu/ac and relative to USDA report) and the coefficient of determination (R^2). These statistics confirm that under sufficient nitrogen concentration and state-level planting/harvest management, the application of a dynamic irrigation scheme (STATE_IRR) improves the modeled yield performance for both corn and soybean, reducing RMSE from 47.8 to 22.3% for corn and from 18.9% to 16.8% for soybean.

Table 2. Summary of the model performance in simulating county-level corn and soybean yield from 2000-2004 (5 growing seasons) as compared to USDA report data for the whole domain and only irrigated regions (in parentheses).

4. Discussion

Several uncertainties can contribute to the differences between simulated crop yields and the USDA report, including those associated with discrepancies between available datasets, crop yield gaps, and crop/irrigation model parameters, which is the subject of discussion in this section.

4.1 Yield gaps between actual yield and modeled potential or water-limited yield

The yield potential (Y_p) is defined as the yield an adapted crop cultivar could achieve by alleviating all abiotic and biotic stresses through optimal crop and soil management (Lobwell et al., 2009). Thus, Y_p is achieved when management eliminates all limitations to crop growth and yield from nutrient deficiencies, water deficit or surplus, toxicities, salinity, weeds, insect pests, and pathogens. In our study, for irrigated corn and soybean, the model provides sufficient water and nitrogen, hence, the modeled yield should be close to Y_p . For rainfed crops, the modeled yield is less than the potential yield due to water limitation (Y_w , water-limited yield). The actual yield (Y_a) is collected from USDA NASS dataset. Therefore, the relative yield gap (Y_g) can be calculated in:

$$Y_g = (1 - Y_a/Y_p) * 100\%; \text{ for irrigated crop} \quad (7)$$

$$Y_g = (1 - Y_a/Y_w) * 100\%; \text{ for rainfed crop} \quad (8)$$

Quantifying the yield gaps for each crop cultivar in different growing regions is still a research topic in the food production community. The Global Yield Gap Atlas (GYGA, www.yieldgap.org) provides estimates of untapped crop production potential on existing farmland based on current climate and available soil and water resources. GYGA's estimated Y_g in US are 10~20% for irrigated corn and 20~30% for rainfed corn, respectively. In our study, Y_g are calculated between USDA county-level data and our model simulations and listed in Table 3, which are 13~25% for irrigated corn and 17~28% for rainfed corn. These numbers are comparable to the numbers given by GYGA. However, the yield gaps for soybean are 15~32% for irrigated and 14~39% for rainfed soybean, which are higher than other studies (e.g., 9~24% in Egli and Hatfield, 2014; 10~30% in Grassini et al., 2015), especially for the rainfed soybean, which agrees with the overestimation in IL, IN and OH.

4.2 Uncertainties in crop model parameters

The development of LSMs has expanded from its initial purpose to provide reliable lower boundary conditions for the coupled climate and weather models by including terrestrial biogeochemical processes, land use change, and dynamic vegetation growth (Bonan et al., 2011). Many LSMs adopt the Farquhar-Ball-Berry scheme to simulate the coupled leaf-level photosynthesis and stomatal conductance (Farquhar and von Caemmerer, 1982; Ball et al., 1987; Collatz et al., 1991; Collatz et al., 1992; Niu et al., 2011; Oleson et al., 2013). Those biophysiological models require a variety of plant-specific parameters, such as the minimum stomatal conductance, respiration rate, and rubisco capacity (V_{cmx25}), and they are usually measured under field experimental conditions. Bonan et al. (2011) reviewed the past literatures on PSN-stomata parameterization in LSMs and found that V_{cmx25} is the most critical parameter in modeling plant photosynthesis. This parameter characterizes the maximum carbon assimilation rate and is measured in laboratory conditions, given sufficient radiation upon leaf level and CO_2 concentration at 25 °C. Bonan et al. (2011) concluded that the leaf-level measured V_{cmx25} , when scaled up to LSM model grid cell, could lead to higher photosynthetic rates when nitrogen was non-limiting (such as for cropland systems). Furthermore, the V_{cmx25} parameter is little constrained and remains model dependent over LSMs.

Table S1 in Appendix B provides a synthesis of the parameters used in several studies. The wide range of V_{cmx25} values (from 30 to 101 $\mu mol\ m^{-2}\ s^{-1}$) and different treatments of product-limiting pathway in PSN calculation (K_p) demonstrate a significant uncertainty in specifying the model-dependent PSN parameters. Hence, calibration of the PSN parameters becomes critical, but has been usually conducted at field scales using measurements of moisture and carbon fluxes. The Noah-MP-Crop model (Liu et al. 2016) uses the generic crop PSN parameters, which don't distinguish C3 and C4 crops. To incorporate corn-specific PSN parameters into Noah-MP-Crop parameter table, we performed a calibration for C4 corn using the LAI and biomass data in the Ameriflux Bo1 site in Bondville, IL. The calibrated values are listed at the bottom row of Table S1, noted as "Adjust", meaning they are calibrated and subject to adjustment. The main result of the calibration is to reduce overestimated rain-fed corn yield by reducing V_{cmx25} from the default value (80 $\mu mol\ m^{-2}\ s^{-1}$) to a lower value (60 $\mu mol\ m^{-2}\ s^{-1}$). The calibration results are presented in Figure S2. As for soybean, the default crop parameters for C3 was used in this study.

He et al. (2019) provides a global rubisco capacity map from satellite-observed solar-induced chlorophyll fluorescence (SIF) record. Through data assimilation methods, the 11-year record of SIF shows both spatial and temporal variation of V_{cmx25} in world's major crop production regions. Future efforts of incorporating the spatial map of V_{cmx25} into ESMs and LSMs would be highly useful to address the wide range of this model parameter.

4.3 Crop Model parameter uncertainties – planting/harvesting management

Representing dynamic crop phenology in LSMs is critical for predicting the energy, water, and carbon budgets in croplands and may even influence the atmospheric boundary layer, especially in areas with large cropland coverage (Betts, 2005; Ma et al., 2012). In some LSMs, the determination of planting and harvesting, as well as plant growth stages are calibrated against field data. Therefore, these calibration efforts are local and there are few studies quantifying the impacts of planting on simulating crop phenology over a large region. For example, in the CLM4-Crop, the planting is activated by three temperature thresholds, a 20-year averaged GDD threshold, a threshold of 10-day running mean of air temperature, and a threshold of daily minimum temperature (Levis et al., 2012). Chen et al. (2018) evaluated the CLM4-Crop over multiple Ameriflux sites over the U.S. corn belt and found there is an early season overestimate of LAI, due to a too-early start of planting. A modified simulation with locally-accurate planting dates showed improvement in simulating energy and water fluxes, as well as the NEE.

In Noah-MP-Crop, the planting and harvesting date are prescribed parameters to reflect the spatial and year-to-year variation of planting/harvesting date for Bo1 and Ne3 sites in Liu et al. (2016). In this study, the BULK_IRR simulation with an early and spatially-invariant planting date overestimated the crop yield and irrigation amount for corn and soybean, consistent with the results of Chen et al. (2018). By contrast, the STATE_IRR simulation with spatially-varying and delayed planting dates effectively mitigated those overestimations (Figure 6). Figure 7 shows that the northern states in the corn belt are relatively more affected by delayed planting date than the southern states, and this north-to-south gradient is evident within each state as well.

Although the state-level planting/harvesting date applied in STATE_IRR represented to some degree of their spatial variations, uncertainties still exist. The USDA usual planting/harvesting date report gives the most active window for planting and harvesting through the survey of last 20 years. In the STATE and STATE_IRR simulation, the middle date of the window time is selected for each state. However, applying the single planting/harvesting date on state-level is still unrealistic. Figure 7 shows the spatial variations of the modeled crop yield sensitivities to delay in planting date and the range of these crop yield responses are calculated in Table 3.

To better constrain the crop growing seasons, it is necessary to incorporate the spatially detailed crop calendars. For example, the planting and harvesting windows can be dynamically modeled based on field workability, considering snow cover and rainfall, and crop biological requirement for heat and moisture (Iizumi et al., 2018). Dynamically modeling the crop calendar will likely to reduce the uncertainties of specifying crop growing seasons in future crop model development, especially in regions where agricultural management data are sparse.

4.4 Crop Model Parameter Uncertainties – convert leaf mass to LAI

Figure 12 shows the reciprocal of measured leaf mass per unit area (LMA, g/m^2) from Ne1 and Ne2 from 2001 to 2007, which demonstrates significant in-season variations for both corn and soybean. For corn, this reciprocal decreases from 0.03 at the early growing stage to 0.01 m^2/g at the end of the growing season. This characterizes a general corn leaf growth feature: extensive leaf growth (larger LAI) at the beginning of the growing season with small amount of mass, and later growing thicker (more mass) with slight increase in LAI. The inverse of LMA for soybean has less variability and the values are generally higher than for corn during the growing season (ranging from 0.018 to 0.029 m^2/g).

The ranges of LMA are listed in Table 3 as compared to the default constant value of BIO2LAI in Noah-MP-Crop that has the same physical meaning as the $1/LMA$ and is used to convert the prognosed leaf mass to diagnosed LAI. BIO2LAI is set as constants for corn (0.015) and soybean (0.030). Such a constant conversion coefficient is used in other LSMs too, e.g., the specific leaf area parameter (SLA) in CLM (Oleson et al., 2013). The substantial seasonal variations of $1/LMA$ in Figure 12 points to the challenges of using a constant BIO2LAI throughout the entire crop growing season, and a time-varying conversion coefficient is needed in future model development.

Figure 12. The reciprocal of the measured leaf mass per area (LMA) from two Ameriflux sites, US-Ne1 and US-Ne2. The inverse of LMA is the same as BIO2LAI parameter in the Noah-MP-Crop model. The black boxes in US-Ne2 indicates soybean years.

4.5 Summary of the uncertainties in validating crop modeling

Table 3 summarizes the aforementioned uncertainties and provides the default values in Noah-MP-Crop and the ranges of uncertainties of three parameters: yield gaps (between USDA-report actual yield and modeled yield), model parameters (V_{cmx25} , planting date, and BIO2LAI). The uncertainty associated with mechanical drying after harvest mentioned in Section 2 are also include in Table 3.

Table 3. Summary of the sources of uncertainties in conducting crop modeling and validating model outputs.

5. Conclusion

This study evaluated the performance of Noah-MP-Crop's joint modeling of crop and irrigation at in the Central U.S. By incorporating spatial datasets of high-resolution crop and irrigation fraction, and state-level planting/harvesting date, the crop model can be applied to regional scale. The impacts of irrigation on crop yield are assessed from field to regional scale as well as under nitrogen sufficient and stressed conditions. Also, several uncertainties including model parameters, yield gaps, and discrepancies between available datasets are assessed.

The results showed that in the U.S corn-belt the bulk simulation (with uniform planting/harvesting date and no irrigation) captured the magnitude and spatial variation of corn yield against the USDA county-level report (RMSE = 28.1% for the whole domain). But in the heavily irrigated region, for example in Nebraska, the yield was much underestimated (RMSE = 48.7% in the irrigated region). Adding irrigation modeling capability effectively improved yield simulation over irrigated region (RMSE=23.1%). The RMSEs for soybean over the whole domain and irrigated region are 28.4% and 20.5%, respectively. The irrigation improvements on soybean yield are relatively small compared to that for corn. Noticeable overestimation of yield for corn and soybean still exist in Northeast of the domain in Indiana and Ohio, which may be attributed to early planting biases and the yield gap between actual yield and modeled yield.

Homogeneous transitioning of the crop model parameters from field to regional scale, two simulations with state-level planting/harvesting date were conducted. These spatially-varied planting/harvesting dates were in general later than the uniform planting dates. The delayed planting dates across states resulted in reduction in modeled yield and irrigation amount, which improved the overestimated yield bias associated with early planting bias. A spatial analysis also showed that the modeled yield in northern states was more sensitive to delayed planting than in southern states for rainfed corn and soybean. This north-to-south gradient was evident within each northern state as well (IL, IN, IO, MN, WI). This indicates that using one single value for planting/harvesting date for each state is still an over-simplified assumption, which is inadequate to address the complex decision of agricultural management. Comprehensive datasets of cropping calendar at high-resolution are needed for future crop model development.

Dynamic modeling of crop growth and irrigation application is challenging and there are many uncertainties. Several sources of uncertainties were identified, including yield gaps, model parameters associated with photosynthetic rubisco capacity and planting date, and discrepancies between different observation data. The rubisco capacity (V_{cmx25}), is a significant source of uncertainty and we calibrated it according to single-point simulation in Bondville for corn (C4 corn).

Fertilization has been identified as a source of uncertainties in previous studies (Leng et al., 2016). In this study, it was assumed that the crops are not nitrogen-stressed. To investigate the impacts of irrigation on crop yield under nitrogen-stress, two sets of additional simulations are conducted which halved the nitrogen concentration. When nitrogen concentration is reduced to half, nitrogen stress could cut crop yield by 48.6% and 73.8% for corn and soybean, respectively (comparing 0.5N with STATE). The irrigation improvements on crop yields under nitrogen stress are restricted (comparing 0.5N and 0.5N_IRR), with 32% and 1% increase for corn and soybean. These numbers are much less than under sufficient nitrogen condition (comparing STATE and STATE_IRR, 93%

642 for corn and 27% for soybean). This concludes that the manifestation of irrigation improvement
643 on crop yield relies on sufficient nitrogen concentration.

644
645 The present study contributed to the knowledge of simulating crop yield and irrigation water
646 amount in one of the world's most productive agriculture regions and investigated the impacts of
647 irrigation on crop yields. The irrigation effects on crop yield under no nutrition-stress condition is
648 addressed in this study, which was often ignored in previous research. However, other sources of
649 uncertainties arise from crop model photosynthesis and phenology parameters, yield gap and unit
650 conversion. To mitigate these uncertainties, we demonstrated that calibrating the crop rubisco
651 capacity parameters and constraining the growing season with spatially-varying
652 planting/harvesting date can improve crop simulation results. Finally, future efforts should be
653 dedicated to incorporating spatially detailed rubisco capacity parameters and crop calendar to
654 better constrain the crop growth dynamics.
655

Tables and Figures

Table 1. Description of the Numerical Experiments.

#	Experiment	Dynamic Crop	Dynamic Irrigation	Planting/ Harvest	Nitrogen Concentration	Note
1	BULK	Yes	No	Uniform date	Sufficient	Baseline simulation
2	BULK_IRR	Yes	Yes (calibrated)	Uniform date	Sufficient	
3	STATE	Yes	No	State-level	Sufficient	To test the impacts planting/harvest date at state-level
4	STATE_IRR	Yes	Yes (calibrated)	State-level	Sufficient	
5	0.5N	Yes	No	State-level	Reduced by half	To assess the impacts of nitrogen-stress
6	0.5N_IRR	Yes	Yes (calibrated)	State-level	Reduced by half	

660 **Table 2.** Summary of the model performance in simulating county-level corn and soybean yield from 2000-2004 (5
661 growing seasons) as compared to USDA report data for the whole domain and only irrigated regions (in parentheses).

Experiment	Cultivar	RMSE	RMSE	R ²
		[bu/ac]	[% relative to USDA]	
BULK	Corn	38.3 (72.0)	28.1% (48.7%)	0.70 (0.23)
	Soybean	11.4 (8.9)	28.4% (20.5%)	0.84 (0.83)
BULK_IRR	Corn	32.2 (34.1)	23.6% (23.1%)	0.79 (0.72)
	Soybean	11.3 (8.5)	28.1% (19.61)	0.86 (0.91)
STATE	Corn	35.9 (70.6)	26.3% (47.8%)	0.71 (0.24)
	Soybean	10.9 (8.2)	27.1% (18.9%)	0.80 (0.83)
STATE_IRR	Corn	29.4 (33.0)	21.5% (22.3%)	0.80 (0.71)
	Soybean	10.6 (7.3)	26.4% (16.8%)	0.82 (0.90)
0.5N	Corn	65.4 (78.6)	47.9% (53.2%)	0.71 (0.51)
	Soybean	23 (17)	57.4% (46.2%)	0.50 (0.38)
0.5N_IRR	Corn	64.1 (68.8)	47.0% (46.7%)	0.74 (0.72)
	Soybean	22 (17)	57.0% (44.8%)	0.50 (0.37)

662
663

664 **Table 3.** Summary of the sources of uncertainties in conducting crop modeling and validating model outputs.

Uncertainty source	Default setting	Range	Unit
Yield gap	[-]	13~25% for irrigated corn 17~28% for rainfed corn 15~32% for irrigated soybean 14~39% for rainfed soybean	% relative to potential yield for irrigated corn and water-limited yield for rainfed corn.
Model parameter, V_{cmx25}	80 for generic crop parameter	30~101 for corn 80~101 for soybean	$\mu mol m^{-2} s^{-1}$
Model parameter, planting date*	111 for corn 130 for soybean	-0.04~-1.22 -0.06~-0.72	bu/ac/day delayed after the default date
Model parameter, BIO2LAI	0.015 for corn 0.030 for soybean	0.010~0.030 0.018~0.029	m^2/g
Handling loss in mechanical drying	[-]	0.04 ~ 5.22 % for corn	% relative to final standard yield at 15.5% moisture content

665

666

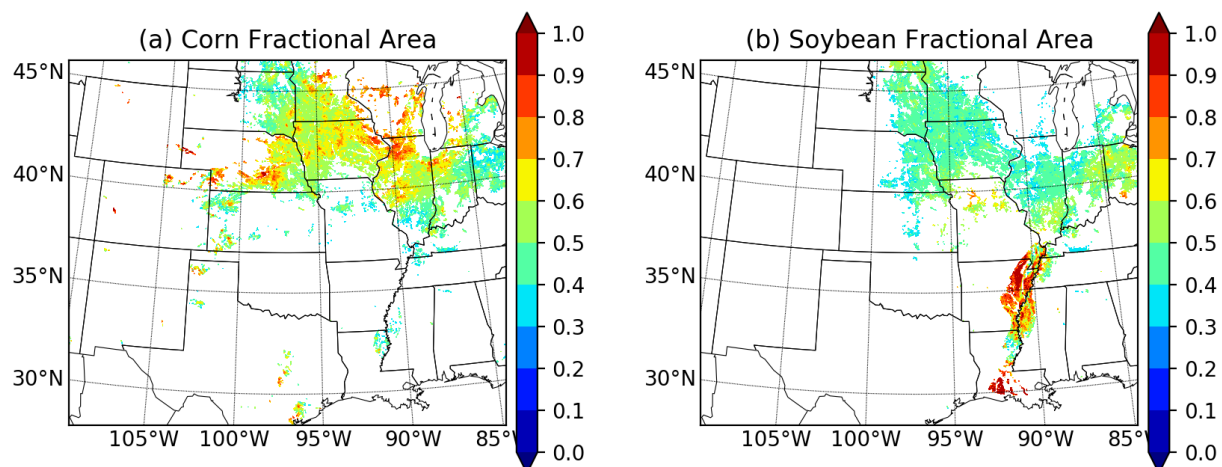


Figure 1. Planted-area fractions for (a) corn and (b) soybean in the Central U.S. domain derived from the USDA-NASS CropScape dataset.

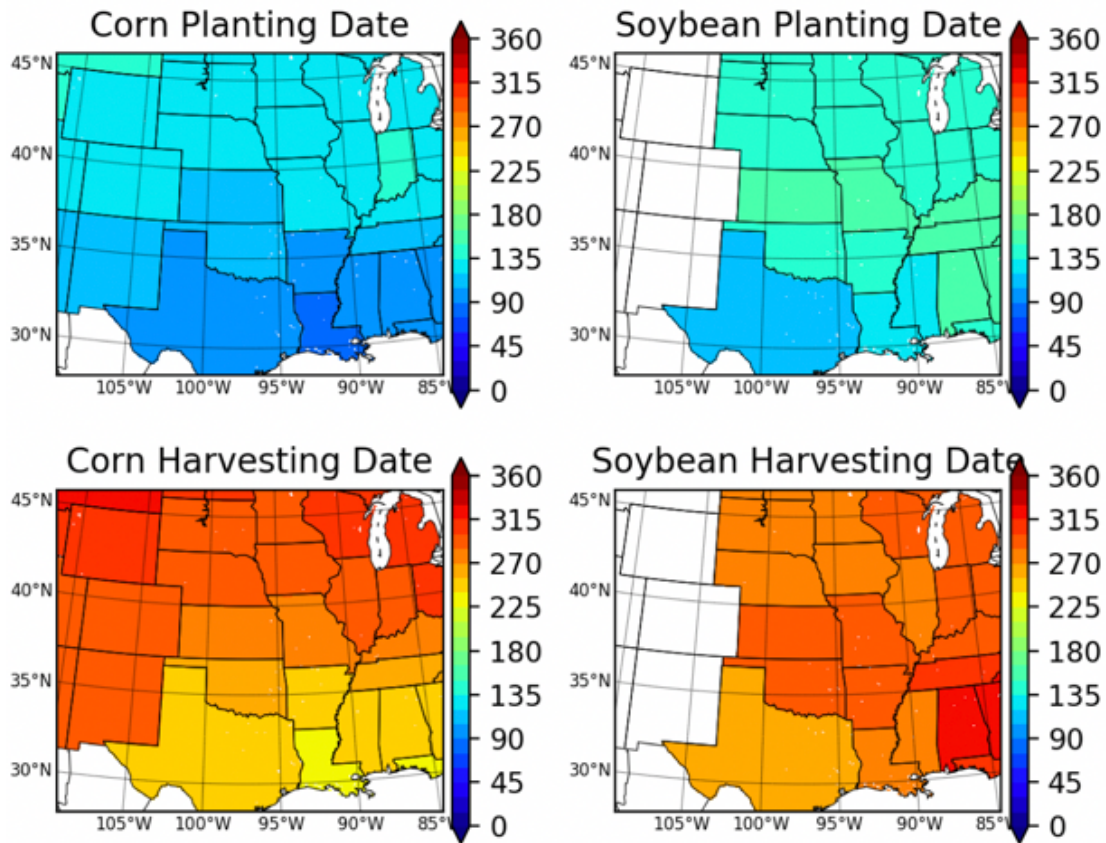


Figure 2. USDA-NASS state-level planting and harvest dates in Julian day for corn and soybean.

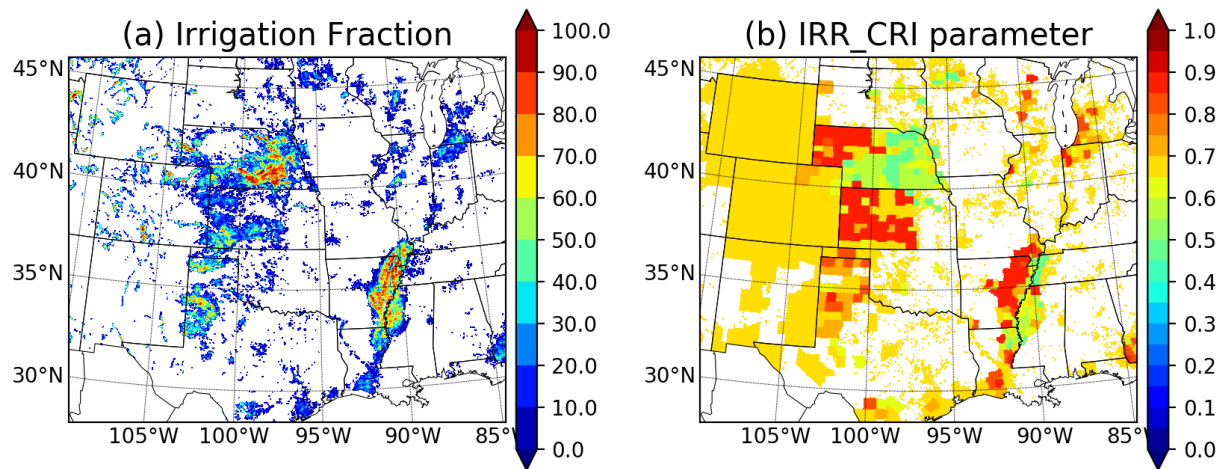
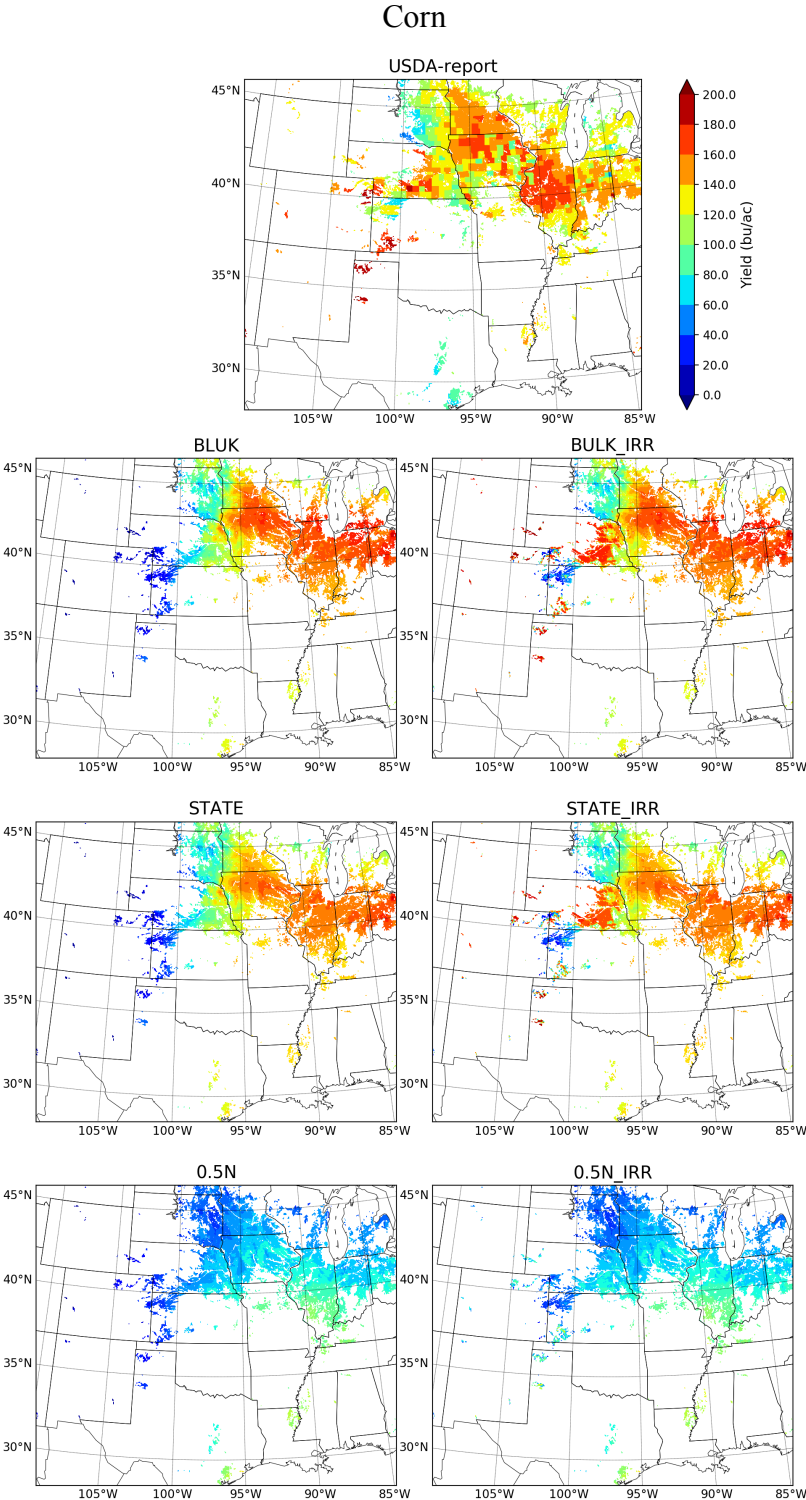


Figure 3. (a) The irrigation fraction used in this study. (b) The critical irrigation threshold parameter used in this study, calibrated in Xu et al. (2019).

678
679



680
681 **Figure 4.** For Corn: Yield (bushel/acre) from USDA NASS county survey and six model simulations (five-year
682 average from 2000-2004).
683
684

Soybean

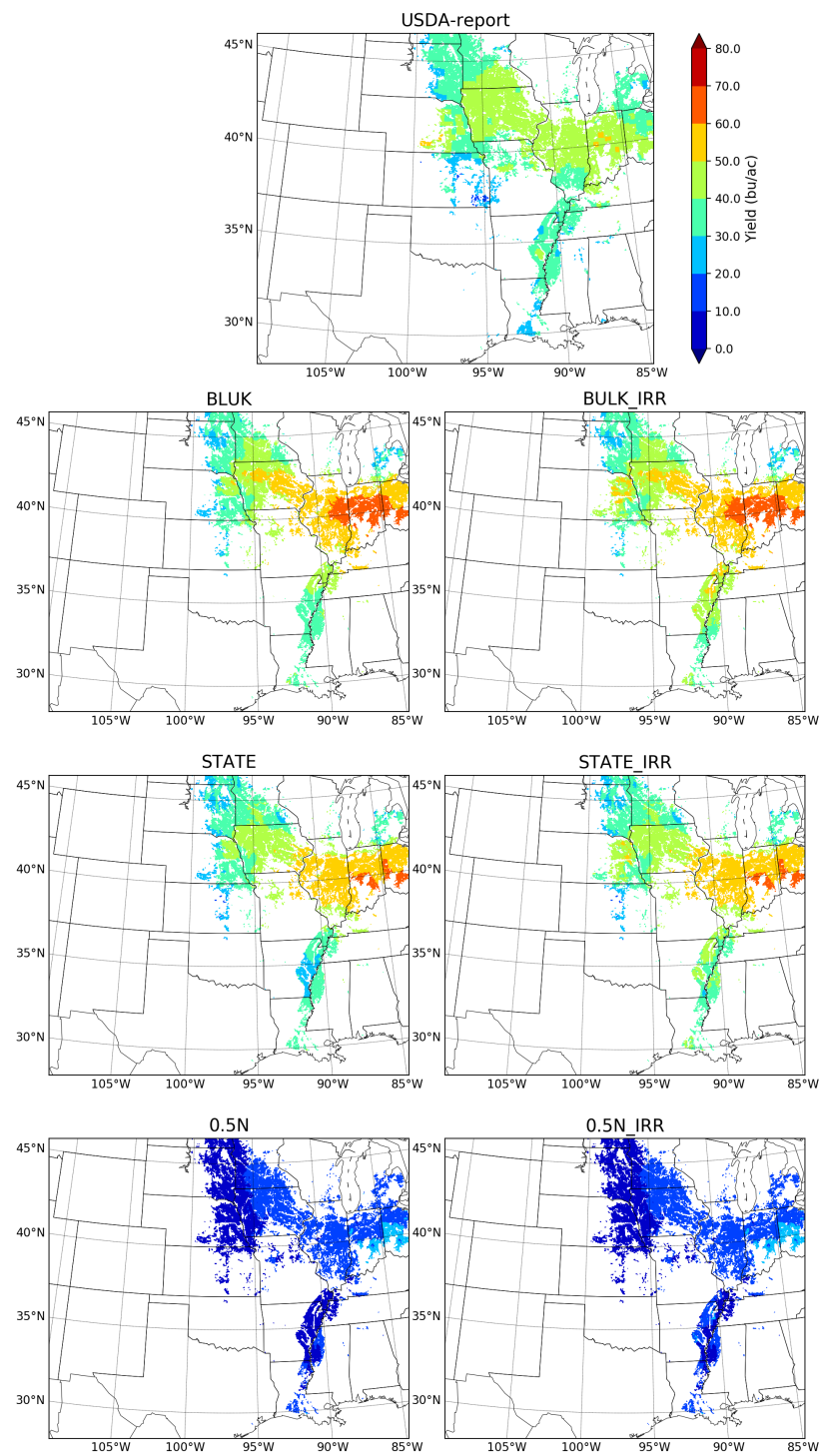


Figure 5. For Soybean: Yield (bushel/acre) from USDA NASS county survey and six model simulations (five-year average from 2000-2004).

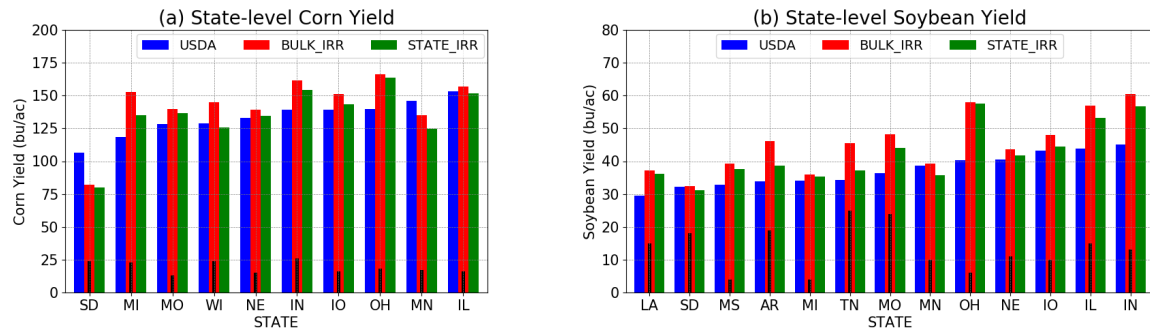


Figure 6. Bar plot of the USDA and modeled yield for each state from the BULK_IRR and STATE_IRR simulation for (a) corn and (b) soybean (five-year average, 2000-2004). The delayed days in planting date in STATE_IRR (compared to the uniform date in BULK_IRR) are shown in black lines.

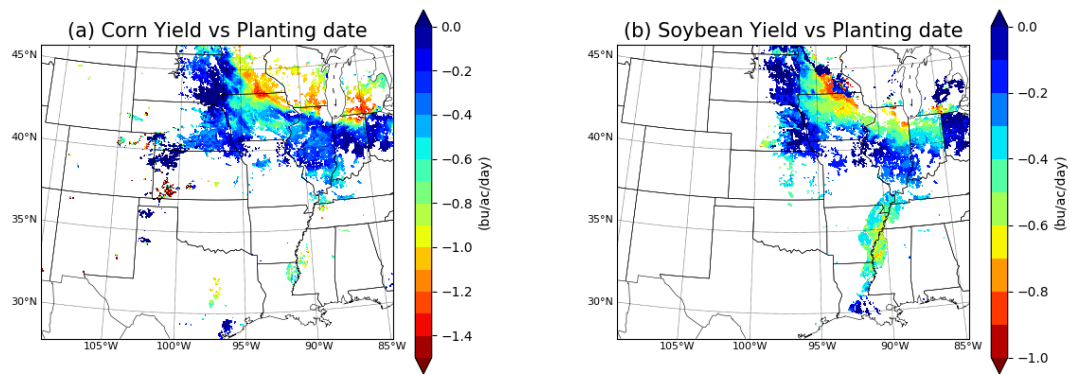


Figure 7. The impacts of delayed planting date on modeled yield (bu/ac/day) for (a) corn and (b) soybean.

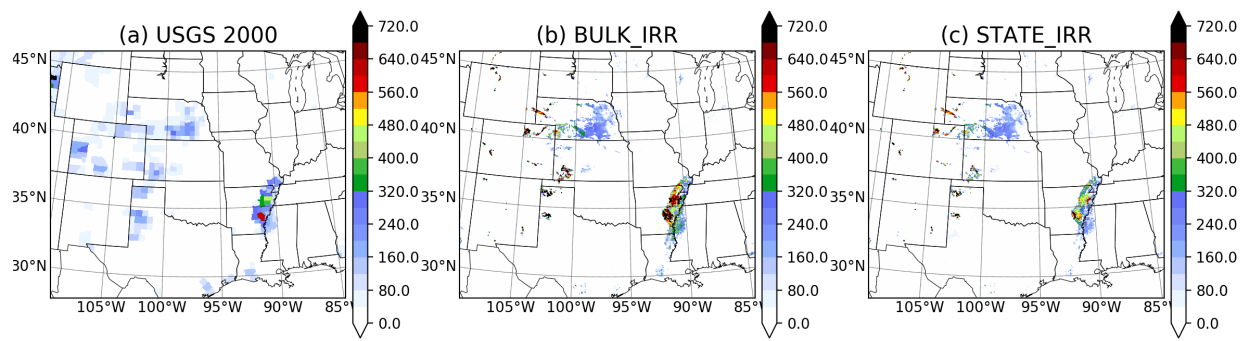


Figure 8. Irrigation amount (mm) in 2000, from (a) USGS county-level water withdrawal report; (b) modeled irrigation amount from the BULK_IRR simulation; and (c) the STATE_IRR simulation.

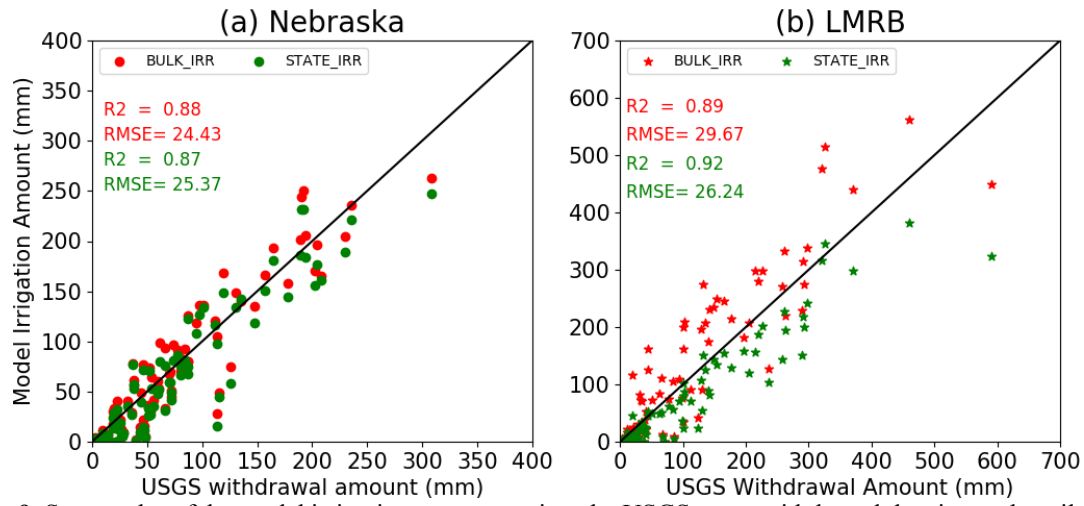


Figure 9. Scatter plot of the model irrigation amount against the USGS water withdrawal data in two heavily irrigated region, Nebraska and Lower Mississippi River Basin (LMRB).

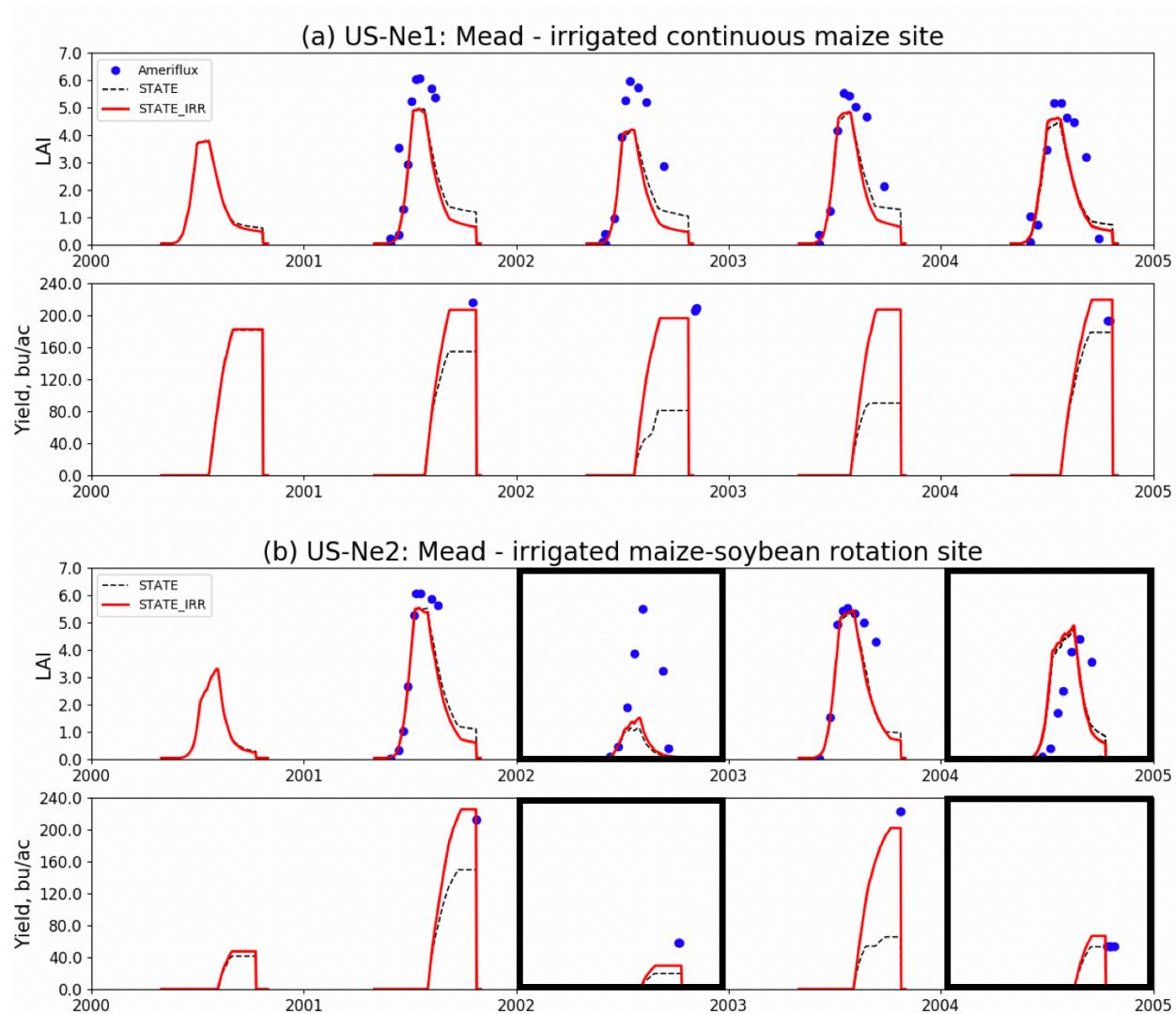


Figure 10. Timeseries of LAI and harvested grain in Ne1 and Ne2 sites from 2000 to 2005. Ne1 is irrigated continuous corn site and Ne2 is irrigated maize-soybean rotation; black boxes in Ne2 indicate soybean years.

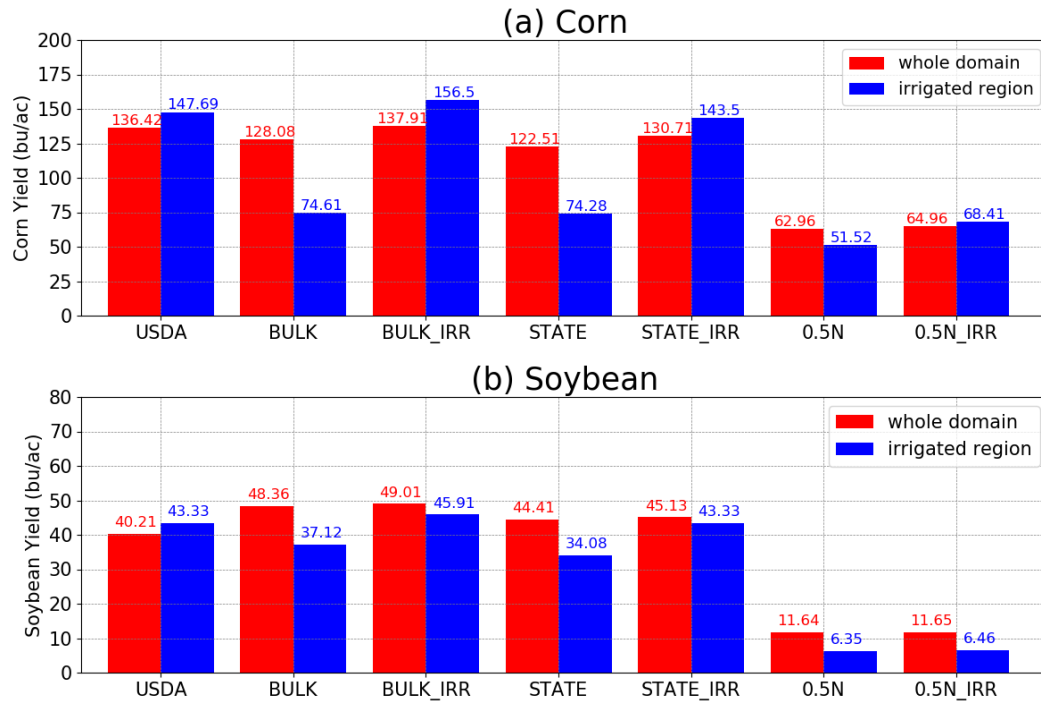


Figure 11. Bar plots of yield (five-year average) for (a) corn and (b) soybean from USDA survey and six simulations in this study. The red and blue bars represent the crop yields in the whole domain and in the irrigated region, respectively.

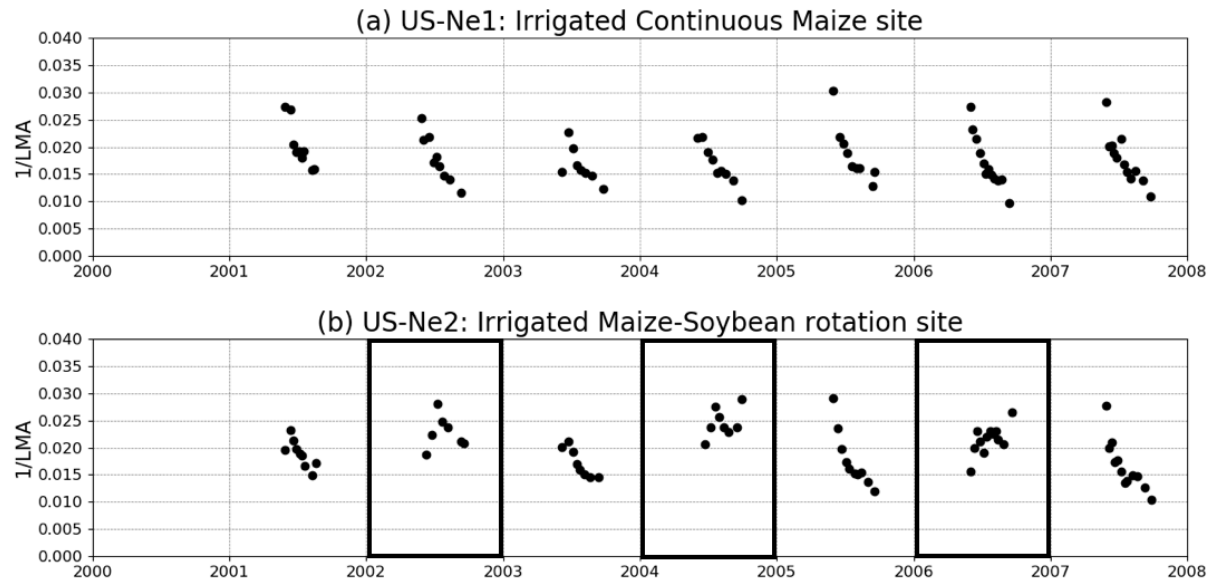


Figure 12. The reciprocal of the measured leaf mass per area (LMA) from two AmeriFlux sites, US-Ne1 and US-Ne2. The inverse of LMA is the same as BIO2LAI parameter in the Noah-MP-Crop model. The black boxes in US-Ne2 indicates soybean years.

722 Appendix A
 723 Field Crops Usual Planting and Harvesting Dates (October 2010)
 724 USDA, National Agricultural Statistics Service
 725 <https://usda.library.cornell.edu/concern/publications/vm40xr56k>
 726

727 **Corn for Grain Usual Planting and Harvesting Dates – States**

State	code	Usual planting dates				Usual harvesting date			
		Begin	Most active	End	Middle day	Begin	Most active	End	Middle day
Illinois	IL	Apr14	Apr21-May23	Jun 5	127	Sep 14	Sep 23-Nov 5	Nov 20	288
Indiana	IN	Apr20	May1-Jun1	Jun 10	137	Sep 15	Oct 1-Nov 10	Nov 25	294
Iowa	IO	Apr19	Apr25-May18	May26	127	Sep 21	Oct 5-Nov 9	Nov 21	296
Michigan	MI	Apr21	May1- May27	Jun 6	134	Sep 5	Oct10-Nov25	Dec 10	306
Minnesota	MN	Apr22	Apr26-May19	May29	128	Sep 27	Oct 8- Nov 8	Nov 23	297
Missouri	MO	Apr 3	Apr11- May27	Jun 12	124	Aug 29	Sep 8-Nov 3	Dec 22	279
Nebraska	NE	Apr19	Apr27-May15	May21	126	Sep 18	Oct 4 -Nov10	Nov 20	296
Ohio	OH	Apr18	Apr 24-May24	May30	129	Spe27	Oct11-Nov20	Dec 1	304
South Dakota	SD	Apr26	May2-May27	Jun 10	135	Sep24	Oct 6-Nov 16	Dec 3	300
Wisconsin	WI	Apr26	May -May27	Jun 6	135	Oct 2	Oct14-Nov17	Nov 28	304

728

729

730 **Soybean Usual Planting and Harvesting Dates - States**

State	code	Usual planting dates				Usual harvesting date			
		Begin	Most active	End	Middle day	Begin	Most active	End	Middle day
Arkansas	AR	Apr19	May5-Jun22	Jun5	149	Sep 10	Sep29-Nov13	Nov 26	294
Illinois	IL	May2	May8-Jun12	Jun24	145	Sep 19	Sep26-Oct26	Nov 7	284
Indiana	IN	May1	May5-Jun10	Jun25	143	Sep 20	Oct1-Nov1	Nov 10	289
Iowa	IO	May2	May8-Jun2	Jun16	140	Sep 21	Sep28-Oct20	Oct 31	282
Michigan	MI	May2	May11-Jun9	Jun18	145	Sep 25	Oct3-Nov3	Nov 13	291
Minnesota	MN	May2	May8-Jun2	Jun13	140	Sep 20	Sep27-Oct20	Oct 31	281
Missouri	MO	May2	May13-Jun24	Jul4	154	Sep25	Oct3-Nov8	Nov 23	294
Mississippi	MS	Apr19	Apr26-May31	Jun17	133	Sep10	Sep13-Oct31	Nov 9	280
Nebraska	NE	May5	May11-May31	Jun8	141	Sep23	Sep29-Oct24	Nov 2	284
Ohio	OH	Apr26	May3-May30	Jun10	136	Spe17	Sep24-Oct21	Nov 5	288
South Dakota	SD	May8	May15-Jun11	Jun21	148	Sep22	Sep28-Oct24	Nov 3	284
Tennessee	TN	May5	May15-Jun25	Jul5	155	Spe25	Oct5-Nov20	Nov 30	301

731

Appendix B
Parameters used in photosynthesis-stomata sub-model

Table S1 A synthesis of photosynthesis parameters used for C4 corn. In this study, we used the Adjust parameters for C4 corn parameters are the same as in the Noah-MP (2011).

Reference	K_p	V_{cmx25} ($\mu\text{molm}^{-2}\text{s}^{-1}$)	$QE25$ (α)	m	b ($\mu\text{molm}^{-2}\text{s}^{-1}$)	R_d ($\mu\text{molm}^{-2}\text{s}^{-1}$)
Noah-MP (2011)	$4000 * V_{cmx}$	80	0.06	9	2000	1.0 (carbon)
Collatz (1992)	0.7, 18000* V_{cmx}	39	0.04	3	80000	0.8, 0.021 * V_{cmx}
Bonan (1996)	$4000 * V_{cmx}$	33 (C4 grass)	0.04	5	2000	0.82 (C4 grass, carbon)
Sellers (1996)	20000* V_{cmx}	30 (C4 grass)	0.05	4	40000	0.025 * V_{cmx} (PSN)
CLM4	$4000 * V_{cmx}$	52 (C4 grass)	0.04	4	40000	-
Bonan (2011)	20000* V_{cmx}	52 (C4 grass; CLM4) 57 (crop; CLM4) 78 (C4 grass; Kattge2009) 101 (C3 crop; Kattge2009)	0.05	4	40000	0.025 * V_{cmx} (PSN)
CLM4.5	20000* V_{cmx}	52 (C4 grass) 101 (corn)	0.05	4	40000	0.025 * V_{cmx} (PSN)
Adjust	20000* V_{cmx}	60 (corn)	0.05	4	40000	0.8 (carbon)

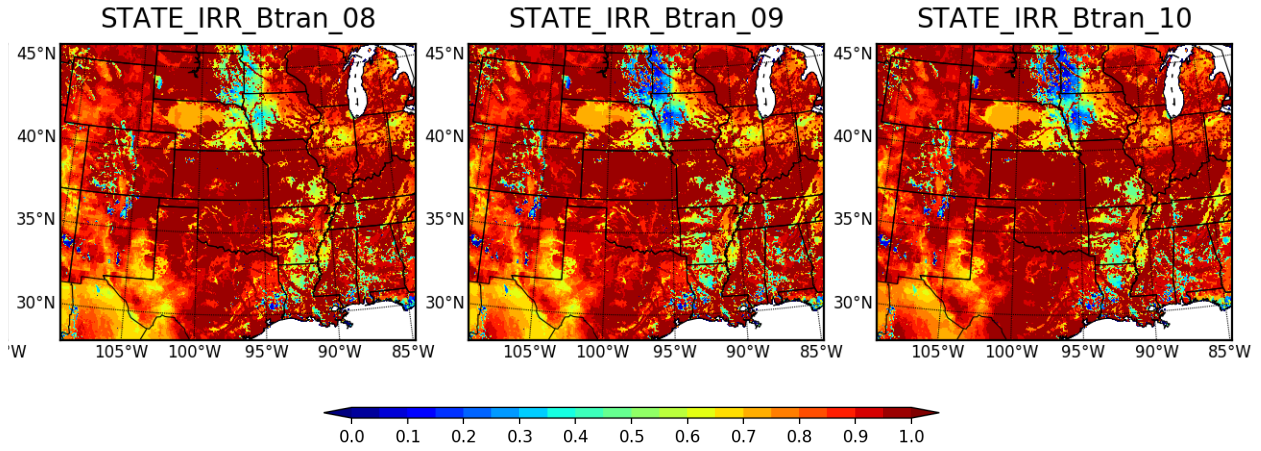


Figure S1. Monthly-averaged water stress factor, β_t , from STATE_IRR simulation from August to October. The blue regions show that the western Iowa, southwest Minnesota and eastern South Dakota are under water stress while the irrigation fraction (Figure 3a) in these regions are small. These suggest that while irrigation and rainfall are not significant water source, the water input from perched shallow water table might be the neglected component for the crop model.

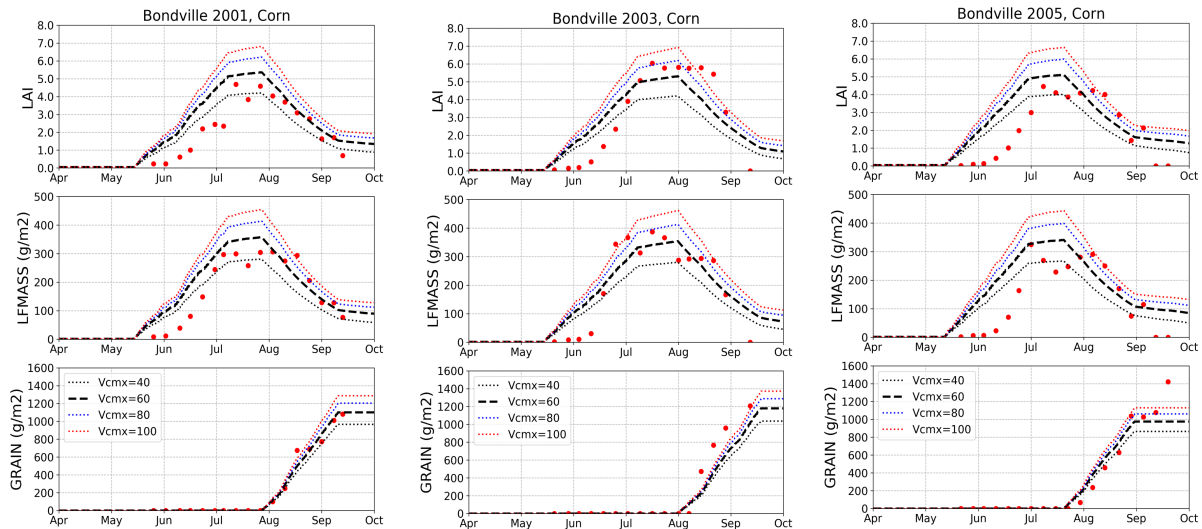


Figure S2. Calibration results for corn rubisco capacity V_{cmx25} , ranging from 40 to 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$, using the Ameriflux site Bo1 biomass data in 2001, 2003, and 2005. The $V_{cmx25} = 60 \mu\text{mol m}^{-2} \text{s}^{-1}$ (black dashed line) is the parameter value used in our regional simulations.

Acknowledgment The authors Zhe Zhang, Yanping Li, Warren Helgason, Zhenhua Li gratefully acknowledge the support from the Global Water Future project and Global Institute of Water Security at University of Saskatchewan. Yanping Li acknowledge the support from NSERC Discovery Grant. Michael Barlage, Fei Chen appreciate the support from by NCAR Water System, USDA NIFA Grants 2015 - 67003 - 23460, NSF INFEWS Grant #1739705, and NOAA OAR Grant NA18OAR4590381. NCAR is sponsored by the National Science Foundation. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Data availability statement

Two AmeriFlux sites data (US-Ne1 and US-Ne2) are available from the AmeriFlux website (<http://ameriflux.lbl.gov/>). The regional scale crop planting area data are available from the USDA/NASS website (<https://nassgeodata.gmu.edu/CropScape/>). The county-level crop yield data are available from USDA/NASS website (<https://quickstats.nass.usda.gov/>). The county-level water use data are from USGS website (<https://water.usgs.gov/watuse/data/2000/>).

Reference

- Ball, J. T., I. E. Woodrow, and J. A. Berry (1987), A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in Progress in Photosynthesis Research, vol. 4, edited by J. Biggins, pp. 221–224, Martinus Nijhoff, Dordrecht, Netherlands.
- Barlage, M., Tewari, M., Chen, F., Miguez-Macho, G., Yang, Z. L., & Niu, G. Y. (2015). The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP. Climatic Change, 129(3–4), 485–498. <https://doi.org/10.1007/s10584-014-1308-8>
- Betts, A. R. (2005). Integrated approaches to climate–crop modelling: needs and challenges. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1463), 2049–2065. <https://doi.org/10.1098/rstb.2005.1739>
- Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., et al. (2011). Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. Journal of Geophysical Research, 116(G2), 1–22. <https://doi.org/10.1029/2010jg001593>
- Chen, F., Liu, C., Dudhia, J., & Chen, M. (2014). A sensitivity study of high-resolution regional climate simulations to three land surface models over the western United States. Journal of Geophysical Research, 7271–7291. <https://doi.org/10.1002/2014JD021827>. Received
- Chen, F., Xu, X., Barlage, M., Rasmussen, R., Shen, S., Miao, S., & Zhou, G. (2018). Memory of irrigation effects on hydroclimate and its modeling challenge. Environmental Research Letters, 13(6). <https://doi.org/10.1088/1748-9326/aab9df>
- Chen, M., Griffis, T. J., Baker, J. M., Wood, J. D., Meyers, T., & Suyker, A. (2018). Comparing crop growth and carbon budgets simulated across AmeriFlux agricultural sites using the Community Land Model (CLM). Agricultural and Forest Meteorology, 256–257(March), 315–333. <https://doi.org/10.1016/j.agrformet.2018.03.012>
- Collatz, G. J., J. T. Ball, C. Grivet, and J. A. Berry (1991), Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: A model that includes a

laminar boundary layer, *Agric. For. Meteorol.*, 54, 107–136, doi:10.1016/0168-1923(91)90002-8.

Collatz, G. J., M. Ribas-Carbo, and J. A. Berry (1992), Coupled photosynthesis-stomatal conductance model for leaves of C4 plants, *Aust. J. Plant Physiol.*, 19, 519–538.

Cosgrove, B. A., et al. (2003), Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS), *J. Geo-phys. Res.*, 108(D22), 8845, doi:10.1029/2002JD00331603.

Dieter, C., Maupin, M., Caldwell, R., Harris, M., Ivahnenko, T., Lovelace, J., et al. (2015). Estimated use of water in the United States in 2015. Circular. <https://doi.org/10.3133/cir1441>

Drewniak, B., Song, J., Prell, J., Kotamarthi, V. R., & Jacob, R. (2013). Modeling agriculture in the Community Land Model. *Geoscientific Model Development*, 6(2), 495–515. <https://doi.org/10.5194/gmd-6-495-2013>

Egli, D. B., & Hatfield, J. L. (2014). Yield Gaps and Yield Relationships in Central U.S. Soybean Production Systems. *Agronomy Journal*, 106(2), 560–566. <https://doi.org/10.2134/agronj2013.0364>

Farquhar, G. D., S. von Caemmerer, and J. A. Berry (1980), A biochemical model of photosynthetic CO₂ assimilation in leaves of C3 species, *Planta*, 149, 78–90.

FAO and DWFI. 2015. Yield gap analysis of field crops – Methods and case studies, by Sadras, V.O., Cassman, K.G.G., Grassini, P., Hall, A.J., Bastiaanssen, W.G.M., Laborte, A.G., Milne, A.E., Sileshi, G., Steduto, P. FAO Water Reports No. 41, Rome, Italy.

Grassini, P., Yang, H., & Cassman, K. G. (2009). Limits to maize productivity in Western Corn-Belt: A simulation analysis for fully irrigated and rainfed conditions. *Agricultural and Forest Meteorology*, 149(8), 1254–1265. <https://doi.org/10.1016/j.agrformet.2009.02.012>

Hicks, D. R., and Cloud, H. A. (1992). Calculating grain weight shrinkage in corn due to mechanical drying. *National Corn Handbook-61*, Purdue Extension, Purdue University.

He, L., Chen, J. M., Liu, J., Zheng, T., Wang, R., Joiner, J., et al. (2019). Diverse photosynthetic capacity of global ecosystems mapped by satellite chlorophyll fluorescence measurements. *Remote Sensing of Environment*, 232, 111344. <https://doi.org/10.1016/j.rse.2019.111344>

Iizumi, T., Kim, W., & Nishimori, M. (2019). Modeling the Global Sowing and Harvesting Windows of Major Crops Around the Year 2000. *Journal of Advances in Modeling Earth Systems*, 11(1), 99–112. <https://doi.org/10.1029/2018MS001477>

Jones, J. W., G. Hoogenboom, C. H. Porter, K. J. Boote, W. D. Batchelor, L. A. Hunt, P. W. Wilkens, U. Singh, A. J. Gijsman, and J. T. Ritchie (2003), The DSSAT cropping system model, *Eur. J. Agron.*, 18, 235–265.

Leng, G., Huang, M., Tang, Q., Sacks, W. J., Lei, H., & Leung, L. R. (2013). Modeling the effects of irrigation on land surface fluxes and states over the conterminous United States: Sensitivity to input data and model parameters. *Journal of Geophysical Research Atmospheres*, 118(17), 9789–9803. <https://doi.org/10.1002/jgrd.50792>

Leng, G., Zhang, X., Huang, M., Asrar, G. R., & Leung, L. R. (2016). The Role of Climate Covariability on Crop Yields in the Conterminous United States. *Scientific Reports*, 6(September), 1–11. <https://doi.org/10.1038/srep33160>

Levis, S., B. B. Gordon, Erik Kluzek, P. E. Thornton, A. Jones, W. J. Sacks, and C. J. Kucharik (2012), Interactive crop management in the community earth system model (CESM1): Seasonal influences on land-atmosphere fluxes, *J. Clim.*, 25(14), 4839–4859, doi:10.1175/JCLI-D-11-00446.1.

- Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., et al. (2017). Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, 49(1–2), 71–95. <https://doi.org/10.1007/s00382-016-3327-9>
- Liu, X., Chen, F., Barlage, M., Zhou, G., & Niyogi, D. (2016). Noah-MP-Crop: Introducing dynamic crop growth in the Noah-MP land surface model. *Journal of Geophysical Research: Atmospheres*, 121(23), 13,953–13,972. <https://doi.org/10.1002/2016JD025597>
- Lobell, D. B., Cassman, K. G., & Field, C. B. (2009). Crop Yield Gaps: Their Importance, Magnitudes, and Causes. *Annual Review of Environment and Resources*, 34(1), 179–204. <https://doi.org/10.1146/annurev.enviro.041008.093740>
- Lobell, D. B., M. J. Roberts, W. Schlenker, N. Braun, B. B. Little, R. M. Rejesus, and G. L. Hammer (2014), Greater sensitivity to drought accompanies maize yield increase in the US Midwest, *Science*, 344, 516–519.
- Lu, Y., & Kueppers, L. (2015). Increased heat waves with loss of irrigation in the United States. *Environmental Research Letters*, 10(6), 064010+. <https://doi.org/10.1088/1748-9326/10/6/064010>
- Ma, S., Churkina, G., & Trusilova, K. (2012). Investigating the impact of climate change on crop phenological events in Europe with a phenology model. *International Journal of Biometeorology*, 56(4), 749–763. <https://doi.org/10.1007/s00484-011-0478-6>
- Maupin, M. A., Kenny, J. F., Hutson, S. S., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2014). Estimated use of water in the United States in 2010: U.S. Geological Survey Circular 1405, 56 p., <https://dx.doi.org/10.3133/cir1405>. Circular. <https://doi.org/10.3133/cir1405>
- McDermid, S. S., Mearns, L. O., & Ruane, A. C. (2017). Representing agriculture in Earth System Models: Approaches and priorities for development. *Journal of Advances in Modeling Earth Systems*, 9(5), 2230–2265. <https://doi.org/10.1002/2016MS000749>
- National Research Council 2011. *Climate Stabilization Targets: Emissions, Concentrations, and Impacts over Decades to Millennia*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12877>.
- Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research Atmospheres*, 116(12), 1–19. <https://doi.org/10.1029/2010JD015139>
- Oleson, K. W., et al. (2013), Technical Description of version 4.5 of the Community Land Model (CLM), NCAR Tech. Note NCAR/TN-503+STR, 422 pp., Natl. Cent. for Atmos. Res., Boulder, Colo., doi:10.5065/D6RR1W7M. [Available at http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf.]
- Ozdogan, M., & Gutman, G. (2008). A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US. *Remote Sensing of Environment*, 112(9), 3520–3537. <https://doi.org/10.1016/j.rse.2008.04.010>
- Pielke, R. A., Adegoke, J., Beltrán-Przekurat, A., Hiemstra, C. A., Lin, J., Nair, U. S., et al. (2007). An overview of regional land-use and land-cover impacts on rainfall. *Tellus B: Chemical and Physical Meteorology*, 59(3), 587–601. <https://doi.org/10.1111/j.1600-0889.2007.00251.x>
- Ray, D. K., Gerber, J. S., Macdonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, 6, 1–9. <https://doi.org/10.1038/ncomms6989>

- Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., & Chatterjee, S. (2019). Climate change already affects global food production, 1–18.
- Rizzo, G., Edreira, J. I. R., Archontoulis, S. V., Yang, H. S. and Grassini, P.: Do shallow water tables contribute to high and stable maize yields in the US Corn Belt?, *Glob. Food Sec.*, 18(April), 27–34, doi:10.1016/j.gfs.2018.07.002, 2018.
- Sellers, P. J., D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, D. A. Dazlich, C. Zhang, G. D. Collelo, and L. Bounoua (1996a), A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation, *J. Clim.*, 9, 676–705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2.
- Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz, and D. A. Randall (1996b), 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, doi:10.1175/1520-0442(1996)009<0706:ARLSPF>2.0. CO;2.
- Siebert, S., & Döll, P. (2010). Quantifying blue and green water uses and virtual water contents in global crop production as well as potential production losses without irrigation. *Journal of Hydrology*, 384(3-4), 198-217. <https://doi.org/10.1016/j.jhydrol.2009.07.031>
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., ... Powers, J. G. (2008). A Description of the Advanced Research WRF Version 3 (No. NCAR/TN-475+STR). University Corporation for Atmospheric Research. doi:10.5065/D68S4MVH
- Suyker, A., (2001-) AmeriFlux US-Ne1 Mead - irrigated continuous maize site, Dataset. <https://doi.org/10.17190/AMF/1246084>
- Suyker, A., (2001-) AmeriFlux US-Ne2 Mead - irrigated maize-soybean rotation site, Dataset. <https://doi.org/10.17190/AMF/1246085>
- Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., Hauser, M., & Seneviratne, S. I. (2017). Present-day irrigation mitigates heat extremes. *Journal of Geophysical Research: Atmospheres*, 122(3), 1403–1422. <https://doi.org/10.1002/2016JD025740>
- Vorosmarty, C. J. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science*, 289(5477), 284–288. <https://doi.org/10.1126/science.289.5477.284>
- Xu, X., F. Chen, M. Barlage, D. Gochis, S. Miao, and S. Shen, 2019: Lessons learned from modeling irrigation from field to regional scales. *J. Adv. Model. Earth Syst.*, DOI:10.1029/2018MS001595
- Yang, H. S., A. Dobermann, J. L. Lindquist, D. T. Walters, T. J. Arkebauer, and K. G. Cassman (2004), Hybrid-maize—A maize simulation model that combines two crop modeling approaches, *Field Crops Res.*, 87, 131– 154.
- Yang, Z. L., Niu, G. Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research Atmospheres*, 116(12). <https://doi.org/10.1029/2010JD015140>