We'd like to thank Bill for his helpful review of our paper. We have carefully considered all the comments and suggestions and our point-by-point response to the review follows:

Comment: First, the authors only cite Levis et al. (2012) briefly, and in passing. Given that that study is quite similar to this one – also involving adding a crop model to CLM, and also using the Agro-IBIS formulation as a starting point for the crop module – the authors need to give more recognition to that study. I understand that the current study differs in some ways from Levis et al. (2012), but the authors need to be more explicit about this. I would like to see some mention of how the two studies are similar and how they differ.

Response: We agree with reviewer's suggestion on the need for a more detailed discussion of the Levis et al (2012). We have added text in our introduction to cite the work done by Levis et al. (2012) and have pointed out differences between that work and this work.

- The following text was added to the introduction: "Recently, a more sophisticated crop model was incorporated into the Community Land Model (Sacks et al., 2012). This addition adds a separate growth scheme for crops to simulate maize, soybean and cereals in the mid-latitudes, using algorithms from the Agro-IBIS model (Kucharik and Brye, 2003). Levis et al. (2012) used the new development to evaluate CO_2 fluxes from the modified leaf area index (LAI). The model showed promising improvements in annual net ecosystem exchange and the impact agriculture has on climate, such as reduced precipitation. However, this model lacked some important features of nitrogen cycling (i.e. nitrogen retranslocation, soybean nitrogen fixation) and management practices (i.e. fertilizer, residue harvest) that may have an important impact on the carbon fluxes."
- We added to the end of section 2.1.1 (Growth scheme): "The grain fill features of this model differ from the Levis et al. (2012) crop model through the maintenance of a separate pool for organ carbon and nitrogen to keep track of yield whereas Levis et al. (2012) allocate grain carbon into the stem pool."
- We added to the end of section 2.1.2 (Nitrogen and retranslocation): "The retranslocation scheme is included in the next release of CLM4.5."
- We added in section 2.1.3 (Fertilization): "The fertilizer scheme is included for the release of CLM4.5."
- We added to the end of section 2.1.4 (Soybean nitrogen fixation): "The soybean fixation scheme will also be added to the CLM4.5 crop model."
- We added in the first paragraph of section 2.1.5 (Crop root structure): "In Levis et al (2012), root density for all vegetation decreases exponentially with depth, but for crops did not vary with growth."
- We altered the text at the beginning of section 2.1.6 (Harvest management) to read "Crops are harvested as soon as maturity is reached, as done by Levis et al. (2012). However, in CLM-Crop, harvest is partitioned between the atmosphere and litter pools."

Comment: Second, there are a number of places where the authors state that the model agrees well with observations, yet their results – and even their own text – do not seem to support this conclusion. I appreciate the authors' careful assessments of the strengths and weaknesses of the model in this paper, and would like them to rephrase some of these introductory and concluding sentences to more accurately reflect the model's performance. I give specific instances of this in my specific comments below.

Response: We have gone through the manuscript and added more precise declarative statements in order to define 'well'. The text was adjusted to reflect more specific examples of agreement. Some examples of these changes are listed below (see specific comments 1, 10, 11, 12, 13, and 15).

Comment: Third, I am concerned that the authors may have mis-applied the Sacks et al. (2010) crop calendar dataset in some instances. This is most noticeable in Fig 4b, which indicates soybean planting at the end of April in Bondville, IL. The Sacks et al. dataset shows a soybean planting date of Day of Year 146.5 for Illinois, which is at the end of May (and even later than this in most of the surrounding states, so I don't think this is a problem of spatial averaging). I would like the authors to double-check this, and if necessary, rerun their simulations and analyses with correctly-applied planting dates.

Response: We have reviewed the application of the Sacks et al. (2010) crop calendar dataset and found that an error occurred during the re-gridding process which caused grid cells that had missing data to be included in the weighted average (with zero values). The impact only included grid cells which bordered bodies of water. We have corrected this error in the re-gridding tool so instead of spatial averaging, the data value used is the same as the grid with the most dominant weight. We re-ran all simulations with the revised dataset and have corrected our analysis to reflect the new results. Although the number of affected grid cells was small, we were able to expand the domain to include grid cells with small land area bordering the oceans. We include the revised figures at the end of our response.

Additional specific comments:
Comment: 1. P. 4138, L. 9-10: This is an instance where your broad "agrees well with measurements" statement is not entirely supported by your results; more nuances are needed here.

Response: Corrected, added "for soybean, but not as well for maize."

Comment: 2. Introduction, last paragraph: Although the authors apply their model globally, it is largely based on parameterizations for the United States, and most of their comparisons with observations are for the United States. This limitation should be mentioned somewhere; this would be a good place to do so.

Response: We added the following text to the paragraph: "We note that although CLM-Crop is designed to be implemented globally, a majority of the parameterizations are typical of crops grown in the U.S.; therefore we focus our analysis on this region with some limited discussion on the global results."

Comment: 3. P. 4143, L. 13: air temperature, I assume? should be explicit

Response: Air temperature is correct; text now defines T_{ave} as "two-meter air temperature"

Comment: 4. Section 2.1.3: For future work, the authors may want to consider Potter, P., N. Ramankutty, E. Bennett, and S. Donner., Characterizing the Spatial Patterns of Global Fertilizer Application and Manure Production, Earth Interactions, 14, 2010.

Response: We agree that this dataset is important for refining the fertilizer application in the model. We've added text in the discussion to note that spatial fertilizer data would help improve global representation of crops, specifically "Additionally, our parameter calibration is focused on crop species grown in the U.S.; expanding these parameters to more broadly capture other cultivars grown would improve the model's ability to capture global crop productivity. One example is to use fertilizer datasets to establish spatial fertilizer application by crop type, such as that developed by Potter et al. (2010)."

Comment: 5. Section 2.1.4: More details, including equation(s), would be helpful here

Response: We have expanded the text to include the equations governing the nitrogen fixation scheme and additional description of the model:

Soybean fixation is dependent on soil water, nitrogen availability and the growth stage of the crop, determined by

$$N_{fix} = N_{plant\_ndemand} * \min(1, fxw, fxn) * fxg \tag{3}$$

Where $N_{plant\_ndemand}$ is the balance of nitrogen needed to reach potential growth that cannot be supplied from the soil mineral nitrogen pool, *fxw* is the soil water factor, *fxn* is the soil nitrogen factor, and *fxg* is the growth stage factor calculated by

$$fxw = \frac{wf}{0.85} \tag{4}$$

$$fxn = \begin{cases} 0 & \text{for } sminn \leq 10 \\ 1.5 - 0.005 * (sminn * 10) & \text{for } 10 < sminn \geq 30 \\ 1 & \text{for } sminn > 30 \end{cases} \tag{5}$$

$$fxg = \begin{cases} 0 & \text{for } PHU \leq 0.15 \\ 6.67 * PHU - 1 & \text{for } 0.15 < PHU \geq 0.30 \\ 1 & \text{for } 0.30 < PHU \geq 0.55 \\ 3.75 - 5 * PHU & \text{for } 0.55 < PHU \geq 0.75 \\ 0 & \text{for } PHU \geq 0.75 \end{cases} \tag{6}$$

where *wf* is the soil water content as a fraction of the water holding capacity for the top 0.5 m, *sminn* is the total nitrogen in the soil pool (g/m2), and *PHU* is the fraction of growing degree days accumulated during the growth period. N_{fix} is added directly to the soil mineral nitrogen pool for use that time step. Nitrogen fixation does not occur in the early growth stage prior to the plant accumulating 15% of PHU or in the late growth stage after 75% of PHUs have accrued, shortly after grain fill begins. The soybean fixation scheme will also be added to the CLM4.5 crop model.

Comment: 6. Section 2.2.1 (or elsewhere): For flux site comparisons, do you just take the model results from the closest grid cell (as opposed to using site-specific data)?

Response: Since we were running the model at a coarse resolution, we chose the closest grid cell to the AmeriFlux coordinates. We have added the following paragraph to section 3.1.1 to clarify the center grid cell latitude and longitude: "For comparison purposes, since CLM-Crop is run at a global resolution of 2.8 degrees, we chose the grid cell closest to the site. In this case, the grid cell central coordinates are 40.46°N, 87.19°W for Bondville and 40.46°N, 95.625°W for Mead."

Comment: 7. Section 2.2.3: The Sacks et al. dataset only covers dominant growth areas of each crop, with very crude extrapolation outside of these areas. Please mention this limitation.

Response: We have addressed this comment by adding the following text: "The Crop Calendar Dataset accounts for generalized planting dates over large regions from the dominant crop cover, which may not capture small scale variability in planting both spatially and temporally (Sacks et al., 2012); however, as our resolution is course, we believe this database is appropriate for this application"

Comment: 8. P. 4149, L. 8-9: GPP is rarely observed directly at flux sites. How was this derived?

Response: GPP is derived using the Net Ecosystem Exchange and net ecosystem respiration. We added the following text to clarify: "Although GPP is not directly measured at AmeriFlux stations, GPP is calculated in the AmeriFlux Level 4 data as the difference between ecosystem respiration and NEE. Ecosystem respiration is estimated using Reichstein et al. (2005) and NEE data is gap filled using the Artificial Neural Network method (Papale et al., 2003)."

Comment: 9. Section 3.1.1, second paragraph: Please also give annual sums of GPP and NEE, either in the text or a figure

Response: Figure 2 shows the evolution of fluxes over the growth period for two cycles of each crop and the annual sums of GPP and NEE can be easily inferred. To summarize, the annual sum of GPP and NEE for maize during 2001 is 1197 gC/m2/yr and -322 gC/m2/yr, respectively at the Bondville site, and 1199 gC/m2/yr and -314 gC/m2/yr, respectively at the Mead site. The annual sum of GPP and NEE for soybean during 2002 is 1055 gC/m2/yr and -257 gC/m2/yr, respectively at the Bondville site, and 904 gC/m2/yr and -228 gC/m2/yr, respectively at the Mead site.

Comment10. P 4150, L 11-13: The RMSD is roughly equal to the mean; this doesn't seem like good agreement to me

Response: Text now reads: "Although the timing is not always synchronized with observations as a result of our use of fixed values for planting and growth period, the model does capture the general trend of NEE during the growth period at both the Bondville and Mead AmeriFlux sites, demonstrating the model's ability to simulate ecosystem respiration."

Comment: 11. P 4150, L 14-15: I agree with this statement for Mead, but not Bondville. You acknowledge the problems at Bondville in the next sentence; this first sentence needs revision to account for this

Response: Text now reads: "CROP-simulated GPP for soybean agrees well with observations at Bondville and Mead, but increases early and is too high at Bondville."

Comment: 12. P 4150, L 26-27: Here I disagree for Mead

Response: We assume you meant that you disagreed for Bondville. We edited text to read: "matching well with observations at Mead"

Comment: 13. P 4151, L 14-15: This is true for peak LAI, but simulated LAI seems to increase sooner than observations show.

Response: We've added the work "peak" before LAI to clarify our meaning.

Comment: 14. P 4151, L 15-17: For future work, the authors may want to consider the Agro-IBIS improvements documented here: Sacks WJ, Kucharik CJ (2011). Crop management and phenology trends in the U.S. Corn Belt: impacts on yields, evapotranspiration and energy balance. Agricultural and Forest meteorology, 7: 882-894.

Response: We are aware of this work and as we move forward with improvements in the model this will likely be a part of our effort.

Comment: 15. P 4152, L 19 and 27-28: Again, your conclusions about good agreement do not seem to be supported by the figures or your own text

Response: We deleted the "agree well with observations" on line 19 and we've altered the text in lines 26-27 to read: "The early general curve and timing of the carbon growth are simulated well for soybean and the peak carbon is simulated well in CROP"

Comment: 16. P 4153, L 3-5: Other than irrigation (which, in the US, should mainly be an issue in the Western US), it seems that your crop parameters should generally reflect presentday U.S. management, at least if they are taken from Agro-IBIS. So I'm not sure that I agree with this explanation

Response: Most of the crop parameters are taken from Agro-IBIS, however, some parameters do not exist in Agro-IBIS, such as the carbon-nitrogen ratio's that are used in the CLM model. We have calibrated these as best as we can, however, it was shown in Sacks and Kucharik (2011) that farmers have not only begun to plant early, but harvest seems to occur at the same time. This implies that farmers are using different hybrids which extend the growth period of the crop. This is something not included in our model, we can't account for changes in the plant type grown, which is considered technology.

Comment: 17. P 4153, L 18: Why does this value of 165 bu acre-1 differ from the value in Fig 8?

Response: The major reason for the difference is the USDA value of 165 bu acre^{-1} is derived from sample surveys from farmers for the year 2009, which incidentally was the peak of corn production in the US, and has since fallen to 123 bu acre^{-1}. The Monfreda et al. (2008) dataset is representative of the year 2000, and is derived from the USDA 2002 Census. The USDA's survey estimates yields for corn at 135 bu acre^{-1} for the year 2000.

Comment: 18. P 4154, L 10: Here and elsewhere: "trends" implies change over time, which isn't what you mean

Response: We've corrected this in several instances by deleting "trend(s)" on page 4148, line 18; page 4150, line 26; page 4153, line 4 17, and 19; and page 4154, line 10.

Comment: 19. P 4154, L 19-20: "Globally. . . observed yields": this agreement is not surprising given the huge range in observed yields

Response: We've added text to reflect this comment. Now reads: "Globally, the full range of CROP yields for maize and wheat falls within the range of observed yields, although the spread of yields is quite large. Regionally this result is not always true."

Comment: 20. P 4156, L 6 and 8, and elsewhere: Change "growing season" to "growth period" when referring to the time between planting and harvest (since "growing season length" generally refers to a climatic index, and a longer growing season is often associated with a shorter growth period)

Response: Thank you for pointing out the distinction. We have corrected all instances of "growing season" to "growth period".

Comment: 21. Section 3.2.1 and 3.2.2 and associated figures: it would help to show the slope of these relationships in addition to the correlation coefficients

Response: The slope of the relationships between the standardized anomalies of yield and temperature and yield and precipitation did not show the same significance as the correlation coefficients. The slope of the regression between yield and temperature was -0.26 for maize, -0.27 for spring wheat, and -0.63 for soybean. The slope of the regression between yield and precipitation was 0.19 for maize, 0.41 for spring wheat, and 0.38 for soybean.

Comment: 22. Section 4.1: Do you have any thoughts as to whether the problematic N cycle in CLM-CN leads to a too-large or too-small effect of these simulations?

Response: We believe the nitrogen deficiency leads to a large effect on these simulations, particularly for the low residue return. We added the following text to the end of the section: "We note however that the nitrogen deficiency in the model may exaggerate the results, especially for the low residue simulation where lack of nutrients affects future soil fertility."

Comment: 23. P 4159, L 6-8: How can earlier planting lead to later harvest in your model? Or is this just an artifact of spatially different responses?

Response: Yes, this was an artifact of spatially different responses. However, after re-running the model with the revised plant dates, we found that this no longer occurs. The revised text now indicates the new changes in planting dates and harvest which for maize, in the SH, planting is 25 days early and harvest is 17 days early.

Comment: 24. Section 4.2: It is interesting to me that both of the temperature-based planting rules generally lead to lower yields. I would like to see some expansion of this point, perhaps in the discussion. To me this indicates why the use of a simple rule-based approach to planting decisions generally doesn't work well in a global model.

Response: The yields were lower because the planting dates were earlier for both experiment compared to the CROP model. This was discussed in the text. We agree that using just temperature running means to bracket the planting dates without consideration to other factors, such as precipitation, may not be a sensitive parameter for global scale modeling. We find it more interesting that longer growing period and cooler growth period temperatures have not resulted in higher yields in these simulations for maize. We believe this is related to the nitrogen deficiency during the middle of the growth period. Preliminary tests indicate that without nitrogen stress, early planting does result in higher yields. However, a more complete parametric sensitivity study is needed to evaluate the model response. We have initiated this study as discussed in a companion paper (Zeng et al., in review, 2012).

Comment: 25. P 4160, L 16: To me this indicates a model deficiency: If farmers could achieve higher yields by planting later, they would do so.

Response: Yes, this is the result of the model's treatment of nitrogen. We've added a statement to this effect. We found in tests with unlimited nitrogen this is not the case, but the vegetation is over productive in those simulations so we did not include those analysis in our paper. Our added text is: "We believe again this is the result of the model treatment of nitrogen."

Comment: 26. P 4161, L 2-6: It isn't really true that this is the first version of CLM that fixes these issues, as I mention in my major points above.

Response: We've changed the text to reflect the work of Levis et al. (2012), now reads: "Cultivation has serious effects on the terrestrial carbon cycle, and the consequences of land management for carbon fluxes have not only recently (Levis et al., 2012) been included in earlier land surface modeling within the CLM framework. Previous versions of CLM had either a crude representation of crops and or omitted many traits that are important in simulating values (carbon fluxes, nutrient demands, soil carbon loss, etc.) that agree with observations such as fertilizer, soybean fixation, retranslocation, etc."

Comment: 27. P 4162, L 15-21: There is already a transient PFT dataset for CLM. How does your suggestion differ from that? Similarly, why did you choose to do your spinup using a steady-state scenario rather than using an existing transient scenario for spinup? (However, I'm more familiar with recent versions of the model than with the pre-CLM4 version the authors use, so it's possible that the transient scenarios I mention did not exist in that older version.)

Response: There are transient PFT datasets to be used in CLM, which include croplands, however, the current treatment of landuse change can only account for changes in the vegetated landunit of the gridcell. As we stated earlier, crops are grown on a separate landunit to avoid competition from natural vegetation with fertilizer inputs. The configuration of CLM does not allow expansion or contraction of landunits in a grid cell so modeling changes in landuse from cropland expansion or abandonment is not possible. In addition, running in dynamic/transient PFT mode causes shifts in vegetation each time step, whereas changes in croplands would more appropriately occur at a yearly interval. We have expanded the text in the paragraph to clarify our meaning: "The current CLM framework allows natural vegetation to change with time, however managed croplands as they are treated in this model cannot expand or contract. Using historical vegetation data to create a transient vegetation data set with appropriate deforestation/reforestation and grassland removal rates related to the growth or abandonment of cultivated land use could improve the performance of CLM-Crop."
We did not use a transient dataset for spinup since we felt it was unnecessary. Transitions would only occur in the natural vegetated landunit, and not on crop landunits. Using a transient dataset for spinup

would not add to our analysis, especially since we ran CLM-Crop with crops for 171 years after the spinup, to establish a new equilibrium.

Comment: 28. Table 1: Initial leaf and stem allocation seem to be missing from this table

Response: Initial leaf and stem allocation are variable depending on the fraction of phenological heat units accumulated. We have added two lines to the table for consistency.

Comment: 29. Fig 3: It's very hard to see the improvement here. Difference maps would probably make this easier to see

Response: Difference maps are not appropriate for the comparison between MODIS and CLM GPP, as a result of other features which dominate the difference plot. The percent change between the GRASS simulation and the CROP simulation of annual GPP is shown below. Although this figure doesn't indicate improvement from the crop development, it does show the change in GPP (usually a decrease) is largest in areas with intensive cultivation, such as the central U.S.
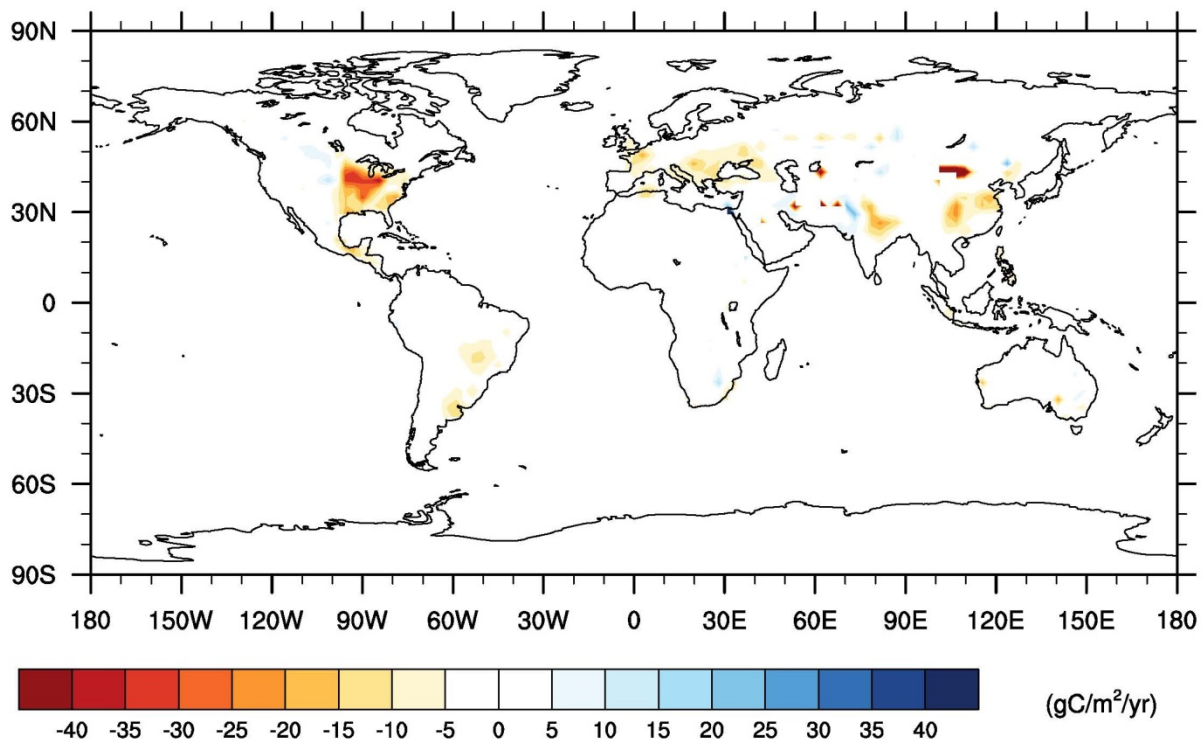
References:

Levis, S., Bonan, G. B., Kluzek, E., Thornton, P. E., Jones, A., Sacks, W. J., and Kucharik, C. J.: Interactive crop management in the Community Earth System Model (CESM1): Seasonal influences on Land-Atmosphere Fluxes, Journal of Climate, 25, 4839-4859, 2012.

Monfreda, C., Ramankutty, N., and Foley, J. A.: Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000, Global Biogeochem. Cycles, 22, GB1022, doi:10.1029/2007GB002947, 2008.

Sacks, W. J., Deryng, D., Foley, J. A., and Ramankutty, N.: Crop planting dates: an analysis of global patterns, Global Ecology and Biogeography, doi: 10.1111/j.1466-8238.2010.00551.x, 2010.

Sacks, W. J. and Kucharik, C. J.: Crop management and phenology trends in the U.S. Corn Belt: Impacts on yields, evapotranspiration, and energy balance, Agricultural and Forest Meteorology, 151, 882-894, 2011.

Zeng, X., Drewniak, B. A., and Constantinescu, E. M.: Calibration of the crop model in the Community Land Model, Geoscientific Model Development Discussion, 6, 379-398, 2013.
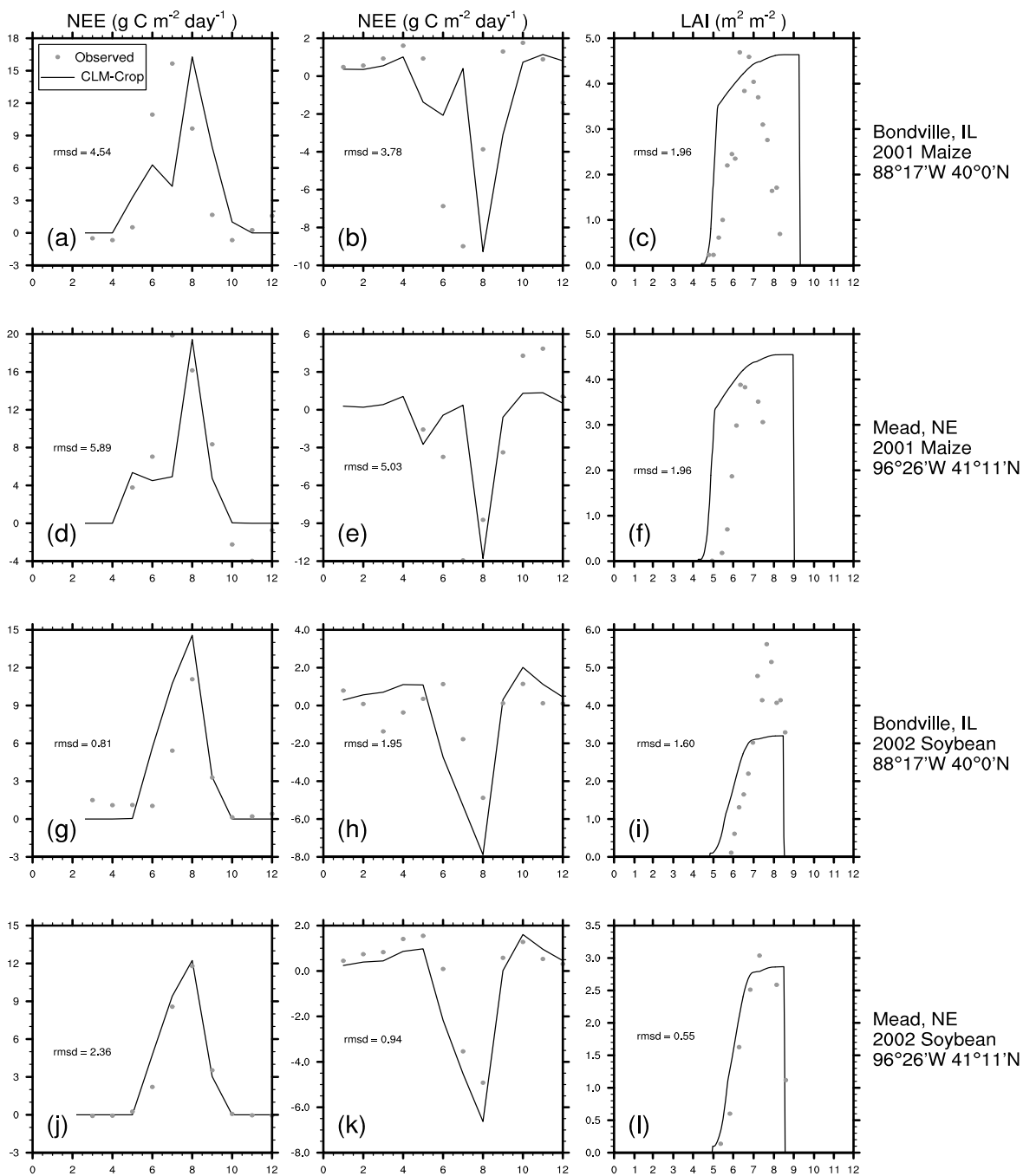
Figure 2. Simulated (lines) and observed (circles) monthly averaged gross primary productivity (GPP; g C m^{-2} day^{-1}), net ecosystem exchange (NEE; g C m^{-2} day^{-1}), and LAI (m^2 m^{-2}) during 2001 and 2002 for maize and soybean at two sites: Bondville, IL (40.01° N, 88.29° W), and Mead, NE (41.18° N, 96.43° W).
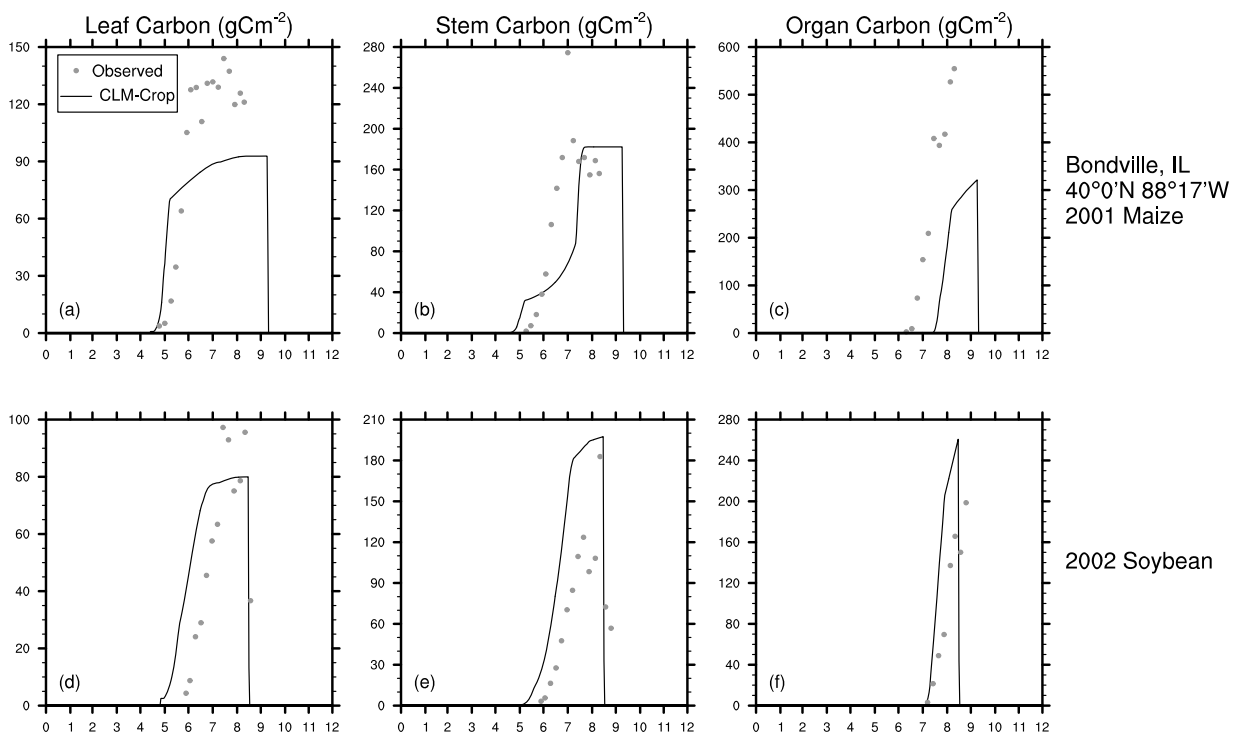
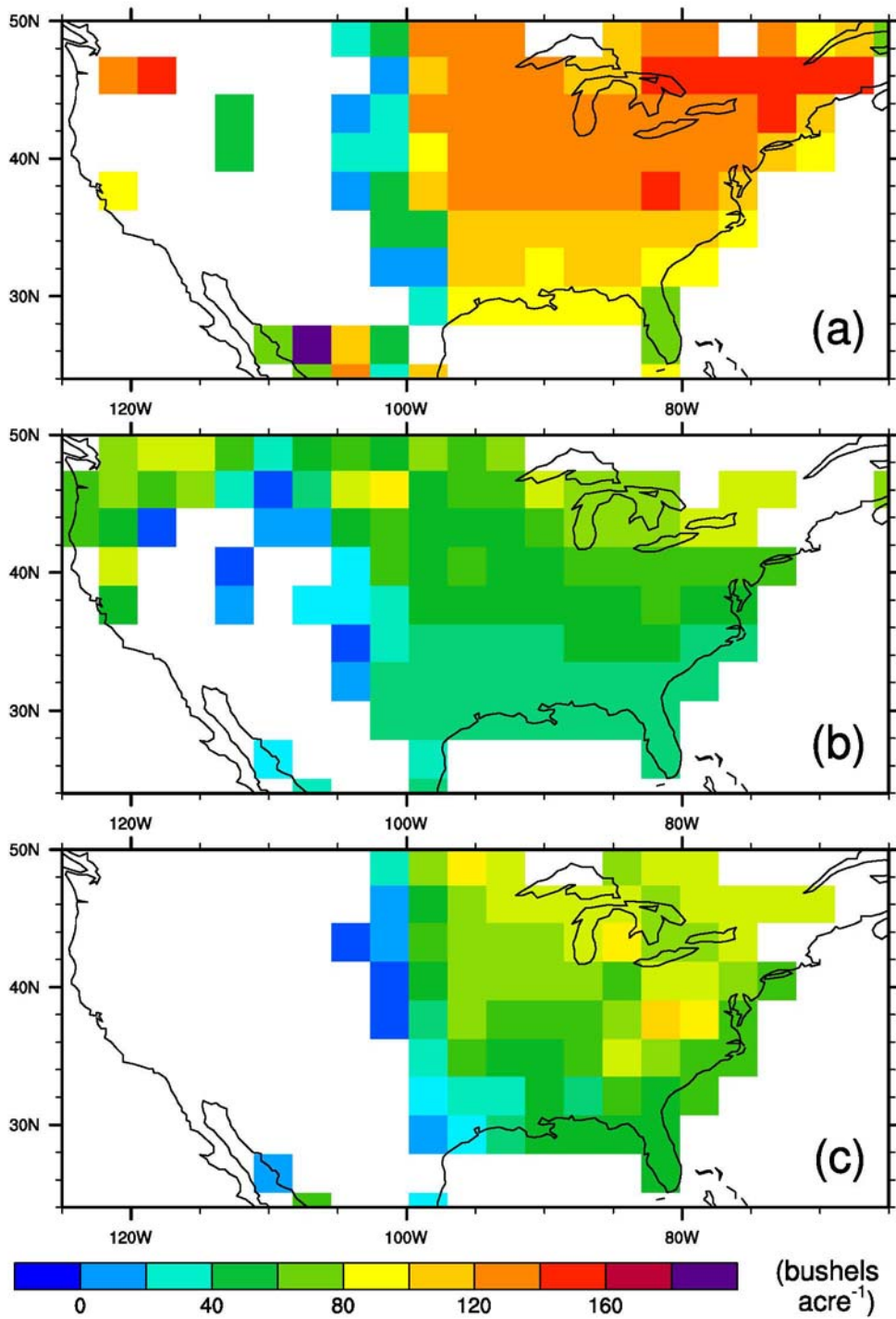Figure 4. Simulated (lines) and observed (circles) leaf, stem, and organ carbon (g C m^{-2}) during 2001 and 2002 for maize and soybean at Bondville, IL.

Figure 6. Simulated crop yields (bu acre^{-1}) for (a) maize, (b) wheat, and (c) soybean.
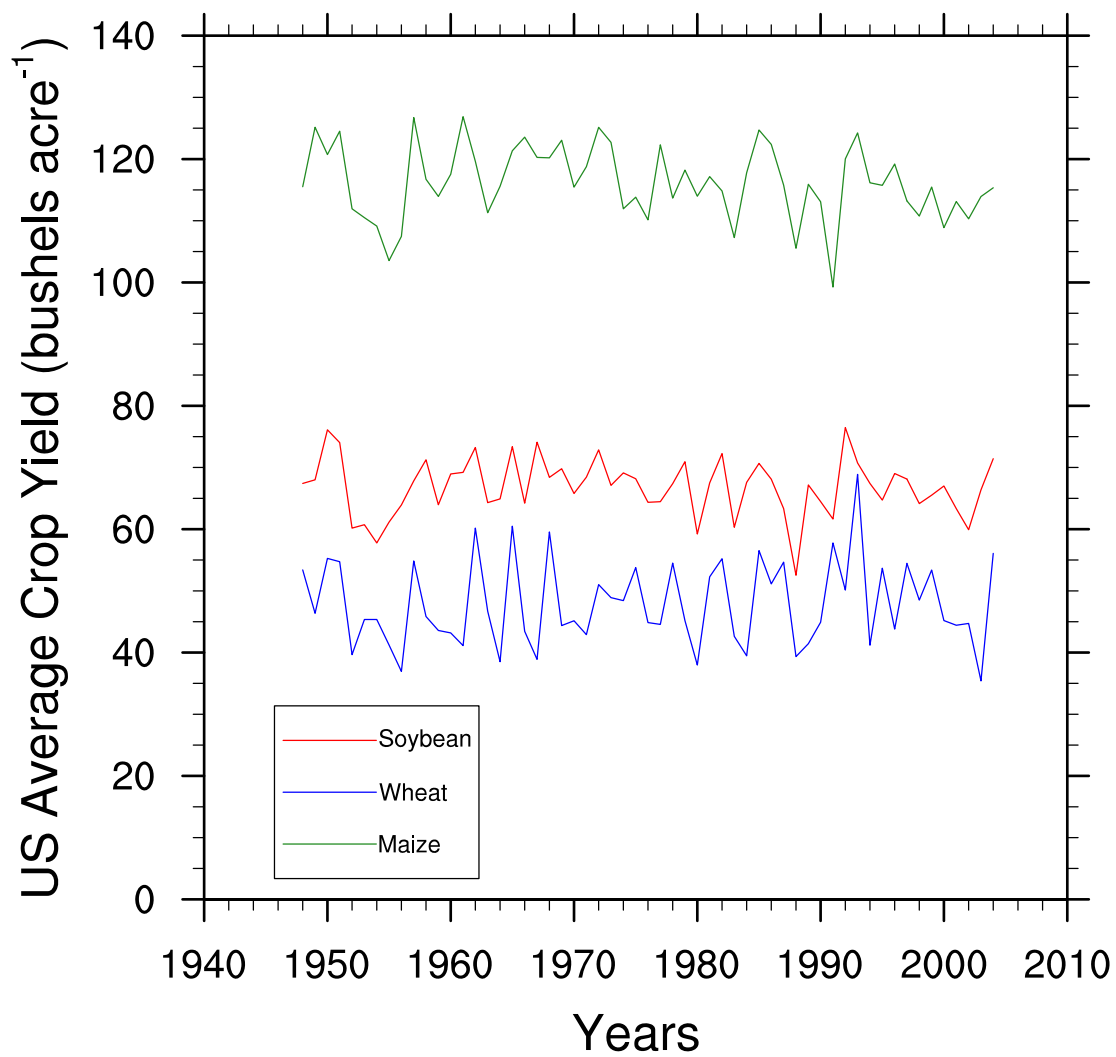
Figure 7. Simulated average U.S. crop yields (bu acre^{-1}) for maize, wheat, and soybean.
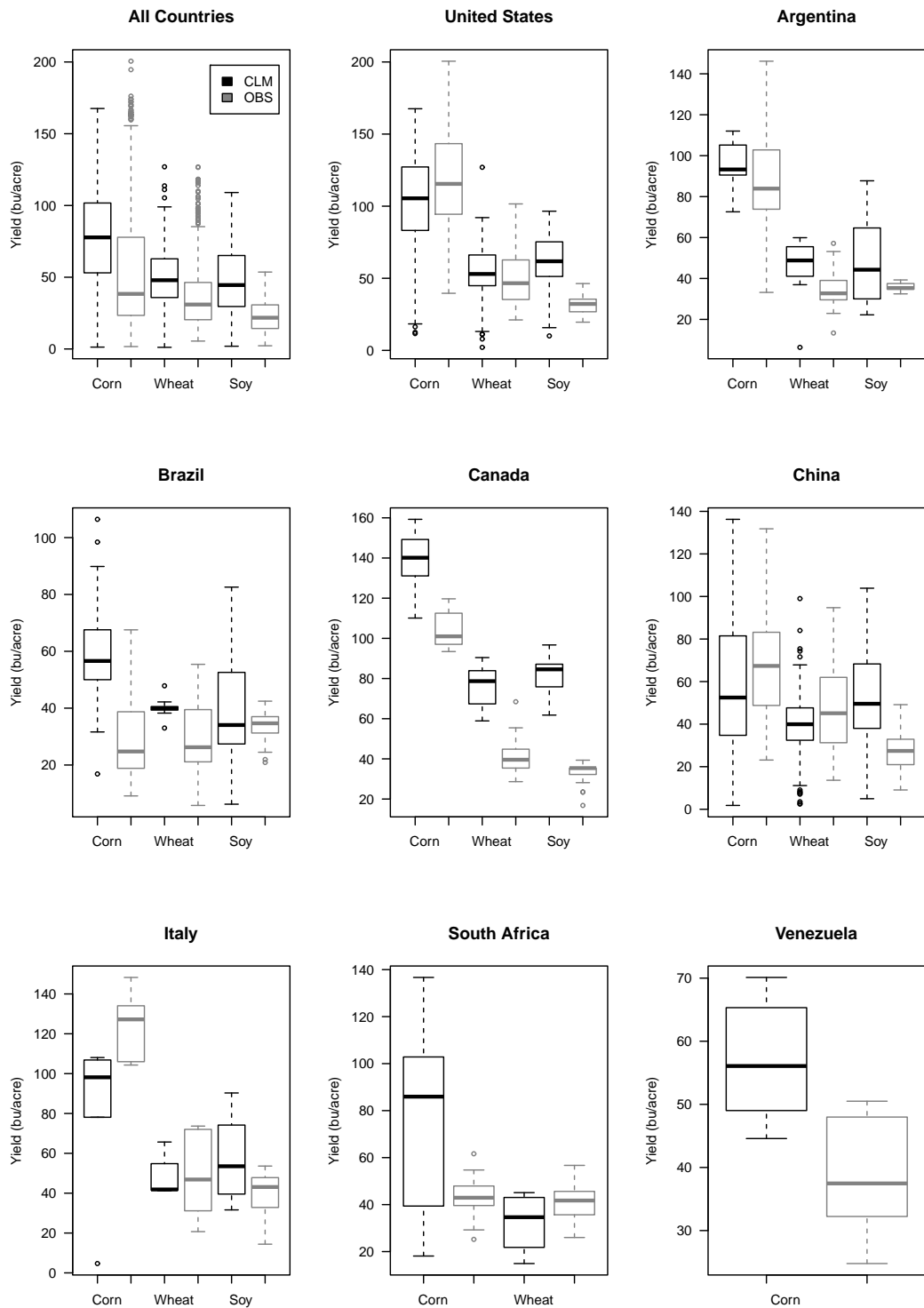
Figure 8. CLM-Crop-simulated (black) and observed (gray; from Monfreda et al., 2008) yield (bu acre^{-1}) for maize, wheat, and soybean of selected regions.
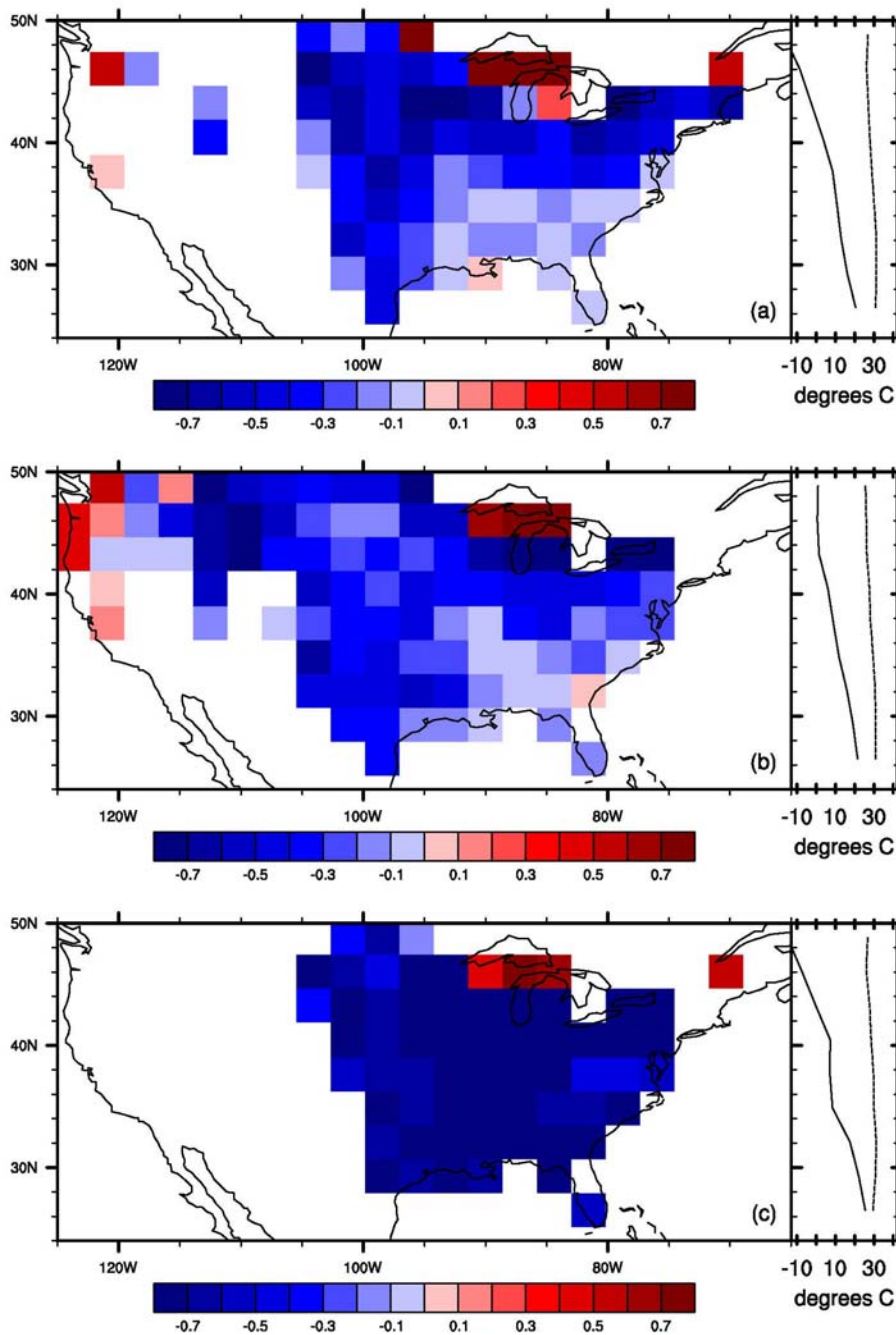
Figure 9. Correlation coefficient between temperature and yield for (a) maize, (b) spring wheat, and (c) soybean. The right half of each panel shows the latitudinal maximum and minimum temperatures (°C) during the growth period for each crop.
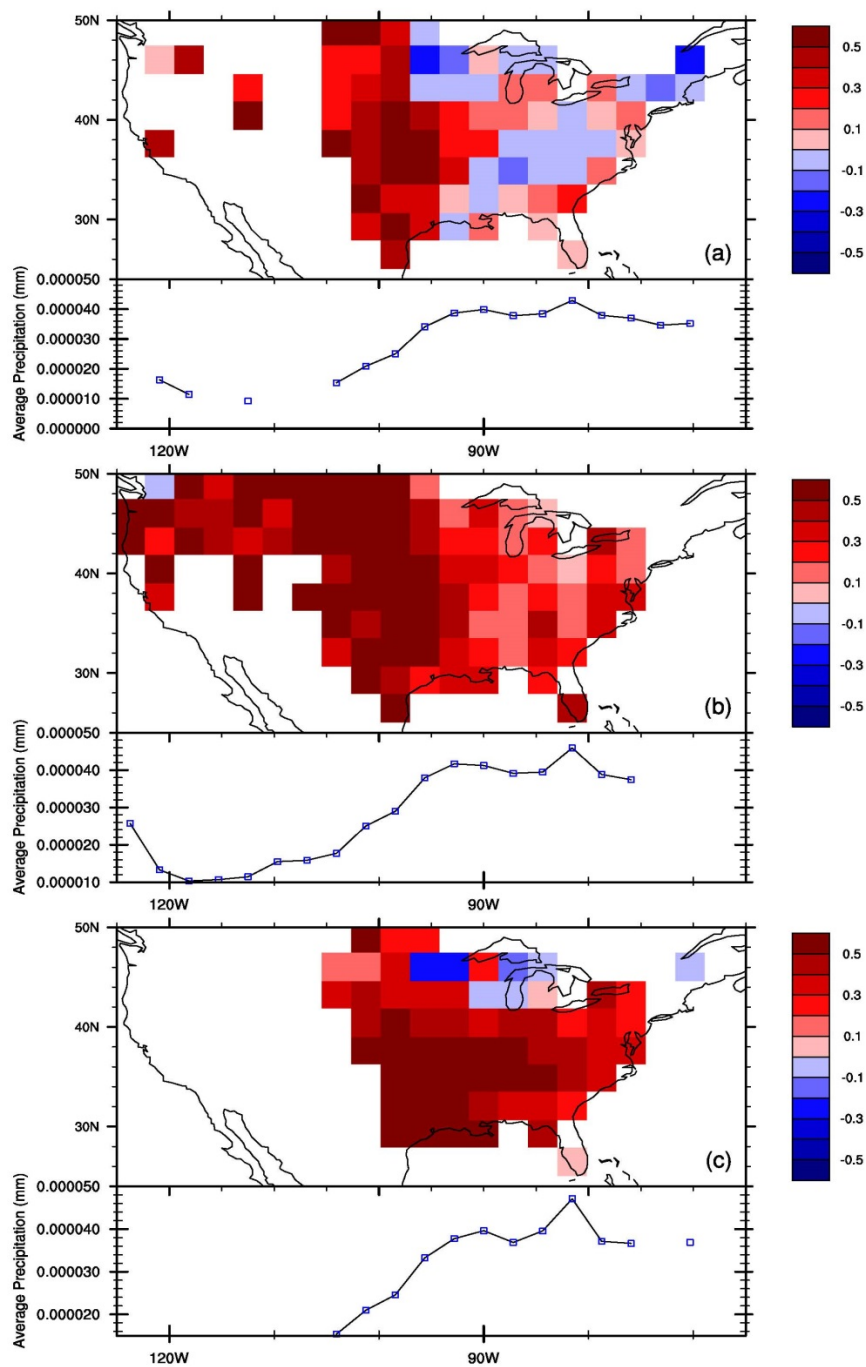
Figure 10. Correlation coefficient between precipitation and yield for (a) maize, (b) spring wheat, and (c) soybean. The bottom half of each panel shows the longitudinal average precipitation (mm) during the growth period for each crop.
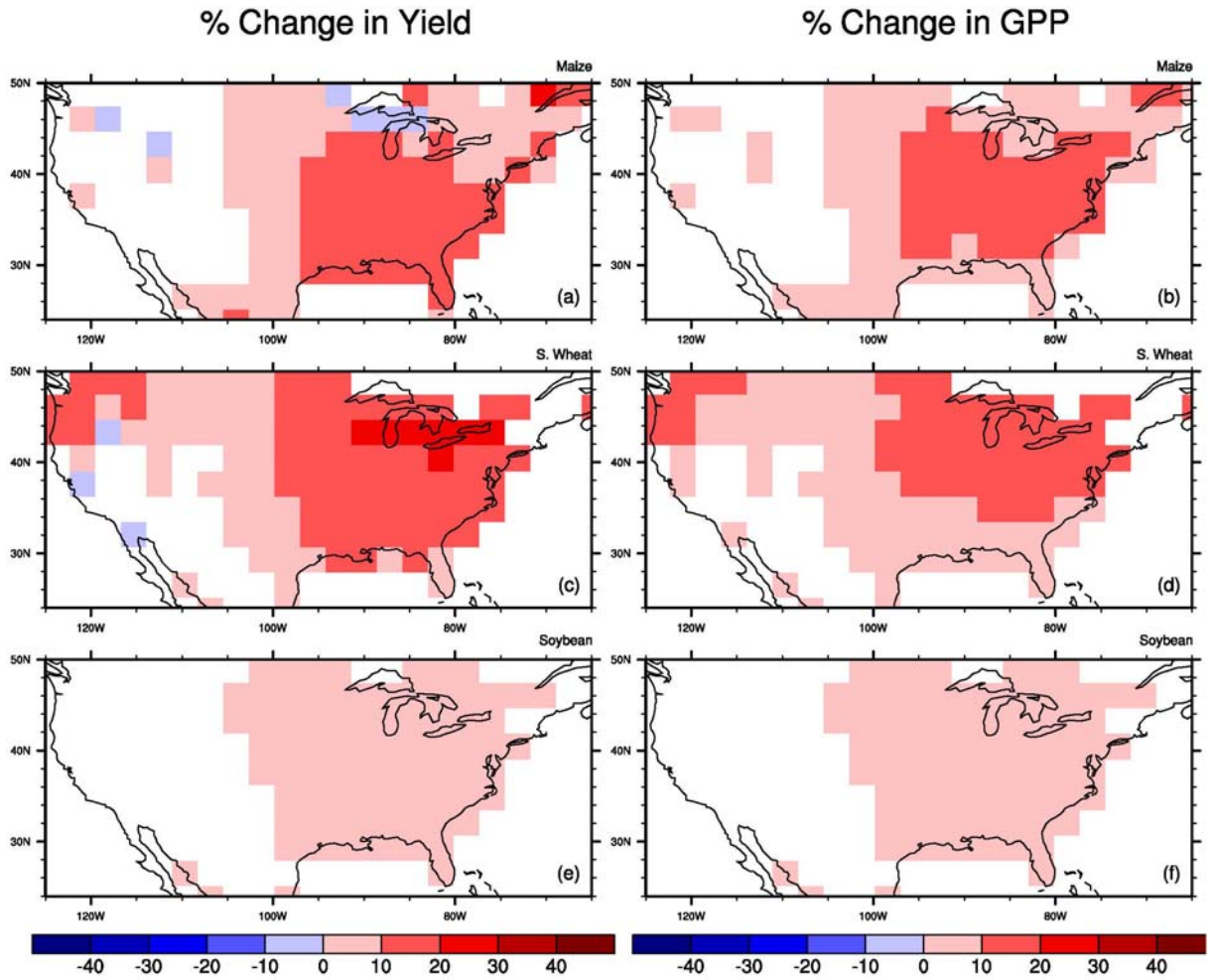
Figure 11. The percent change in yield (left column) and GPP (right column) for (a,b) maize, (c,d) spring wheat, and (e,f) soybean from a 70% residue return management practice (HIGHRES).
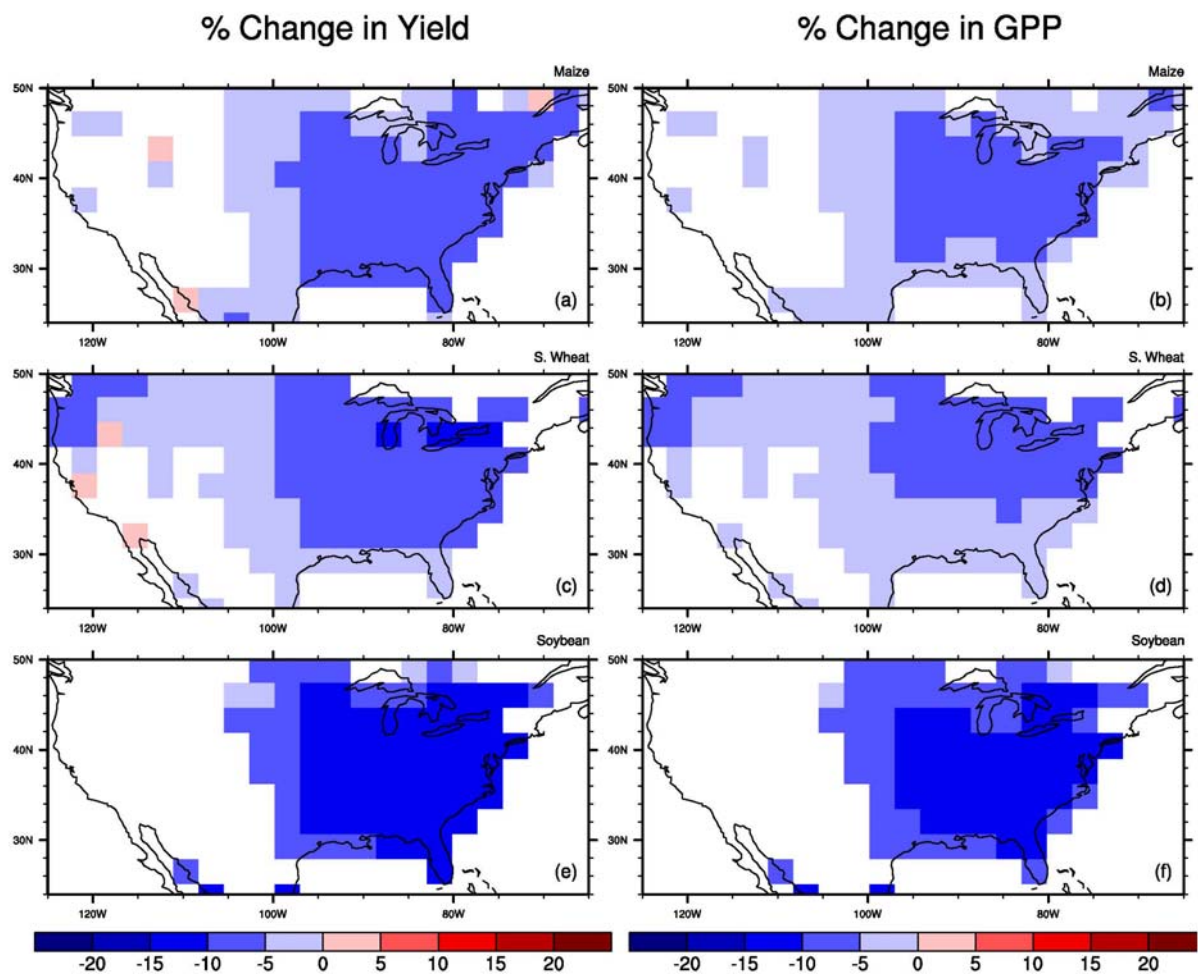
Figure 12. The percent change in yield (left column) and GPP (right column) for (a,b) maize, (c,d) spring wheat, and (e,f) soybean from a 10% residue return management practice (LOWRES).
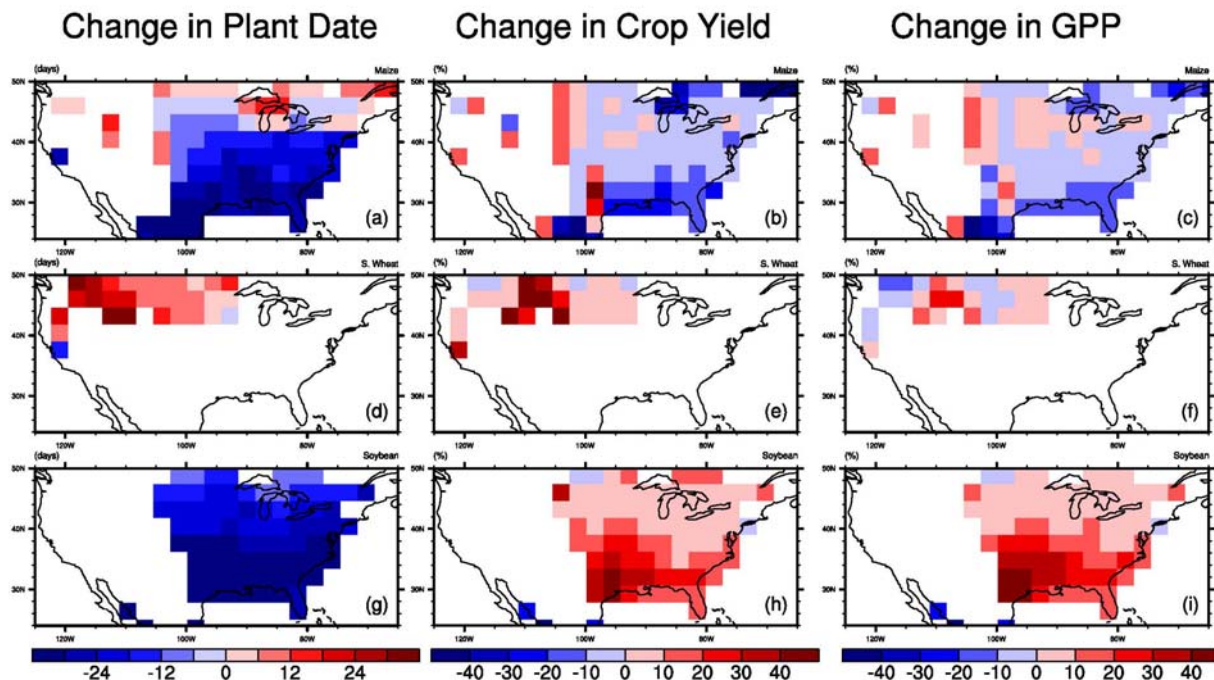
Figure 13. The left column is the change in planting date (days), represented by the difference between LowPTEMP and CROP for (a) maize, (d) spring wheat, and (g) soybean. The center and right columns are the percent change in crop yield for (b) maize, (e) spring wheat, and (h) soybean and the GPP for (c) maize, (f) spring wheat, and (i) soybean resulting from new planting dates.
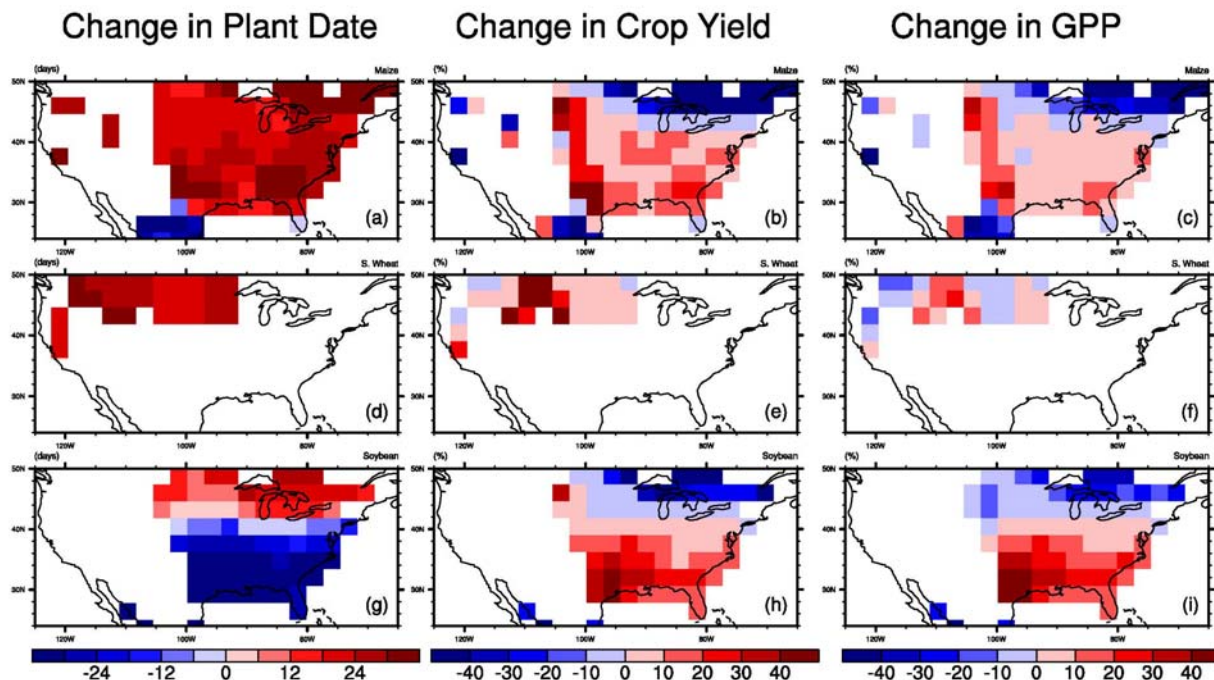
Figure 14. The left column is the change in planting date (days), represented by the difference between HighPTEMP and CROP for (a) maize, (d) spring wheat, and (g) soybean. The center and right columns are the percent change in crop yield for (b) maize, (e) spring wheat, and (h) soybean and the GPP for (c) maize, (f) spring wheat, and (i) soybean resulting from new planting dates.