Simulating Agriculture in the Community Land Model Version 5

Danica L. Lombardozzi, Yaqiong Lu, Peter J. Lawrence, David M. Lawrence, Sean Swenson, Keith W. Oleson, William R. Wieder, and Elizabeth A. Ainsworth

Abstract Agricultural expansion and management have greatly increased global food production and altered Earth’s climate by changing physical and biogeochemical properties of terrestrial ecosystems. Few Earth system models represent agricultural management practices due to the complexity of the interactions between human decisions and biological processes on global scales. We describe the new capabilities of representing crop distributions and management in the Community Land Model (CLM) Version 5, which includes time-varying spatial distributions of major crop types and their management through fertilization and irrigation, and temperature-based phenological triggers. Including active crop management increases peak growing season gross primary productivity (GPP), increases the amplitude of Northern Hemisphere net ecosystem exchange, and changes seasonal and annual patterns of latent and sensible heat fluxes. The CLM5 crop model simulates the global observed historical trend of crop yields with relative fidelity from 1850 to 1990. Cropland expansion was important for increasing crop production, especially during the first century of the simulations, while fertilization and irrigation were important for increasing yields from 1950 onward. From 1990 to present day, observed crop production continued to increase while CLM5 production levels off, likely because intensification practices are not represented in the model. Specifically, CLM does not currently include increasing planting density, crop breeding and genetic modification, representations of tillage, or other management practices that may also affect crop-climate and crop-carbon cycle interactions and alter trends in yields. These results highlight the importance of including crop management in Earth system models, particularly as global data sets for parameterization and evaluation become more readily available.

Plain Language Summary While agricultural expansion and management practices are critical for increasing global food production, there is limited understanding of how they impact fluxes of carbon, water, and energy from the land surface to the atmosphere. Global land models are useful for understanding these possible climate impacts, yet few global land models explicitly represent crops or crop management given the complexity of interactions between human decisions, crop phenology, and land processes at global scales. Our analysis illustrates that representing specific crop types, as well as irrigation and fertilization, in the Community Land Model (CLM) increases the amount of carbon that plants draw out from the atmosphere and also changes patterns of evapotranspiration. Additionally, crop yield estimates from CLM compare well to observed crop yields until ~1990, when modeled crop yields level off. This occurs because CLM does not represent management practices associated with modern agricultural intensification. Overall, our results illustrate the impact that crop expansion and management may have on climate and highlight that global models should represent specific crop types and crop management to accurately capture carbon, water, and energy fluxes from the land surface.

1. Introduction

Agriculture has significantly changed the land surface, with approximately 12% of today’s land surface devoted to cultivated crops (Goldewijk et al., 2017; Leff et al., 2004). As human population and the demand for food increases (FAO, 2013), understanding the interactions between agricultural management practices and changes in Earth’s climate will help to inform future decisions about how to achieve the necessary increases in agricultural production while simultaneously meeting climate goals.
Crops can alter regional and global climate through both biogeochemical and biogeophysical pathways. Due to selective breeding and intensive management that minimizes resource constraints, crop productivity can exceed that of natural vegetation. Higher productivity in cropping systems compared to natural grasslands is observable from satellite estimates of Sun-induced chlorophyll fluorescence (SIF; Guanter et al., 2014) and contributes to the increased seasonal amplitude of atmospheric CO₂ concentrations (Gray et al., 2014; Zeng et al., 2014). Despite accelerated uptake of CO₂ by crops during the growing season, agricultural management practices, like tillage harvest and fertilization, can considerably increase net greenhouse gas emissions from agricultural lands. A recent analysis by Tubiello et al. (2015) estimates that agricultural (crop and livestock) greenhouse gas emissions are growing by 1% annually and now surpass the emissions from land use change.

Beyond these biogeochemical effects, crop management can also change biogeophysical properties of the land surface, which impact climate. For example, increased latent heat due to irrigation and intensification of crop productivity has contributed to regional cooling (Lobell et al., 2006; Mueller et al., 2015) and changed precipitation (Levis et al., 2012). Recent work highlights how management of crop albedo, which is greater than forest ecosystems but lower than perennial grasses, could potentially help to mitigate the climate impacts of crops by changing the available energy at the land surface (Bagley et al., 2015; Miller et al., 2015). Moreover, practices such as tillage, crop residue management, and cover cropping can also change biogeophysical properties of croplands by changing soil moisture retention, albedo, and regional temperatures (Bagley et al., 2015; Lobell et al., 2006; Lombardozzi et al., 2018).

Although crops are an important part of the land surface and have a rich history of consideration in crop and agroeconomic models (e.g., Elliott et al., 2015), the representation of crops in the land component of Earth system models (ESMs) that are used to make climate projections are less sophisticated. Many land models typically represent crops as grasses, even though crop phenology can be quite different from natural grasslands, or, if explicitly included, are represented simplistically (e.g., without management like irrigation or fertilizer application; McPherson et al., 2004). Due to the importance of agricultural practices on carbon, water, and energy fluxes, several land components of ESMs have recently developed capabilities to represent crops, finding that crops improve simulated carbon and water fluxes (McDermid et al., 2017). For example, the Simple Biosphere model (SiB), ORCHIDEE, and older versions of the Community Land Model (CLM: CLM4.0 and CLM4.5) have each incorporated crop modules to represent several varieties of temperate crops, which improved simulated leaf area indices and net ecosystem exchange (Drewniak et al., 2013; Levis et al., 2012, 2016; Lokupitiya et al., 2009; Wu et al., 2016). The Joint UK Land Environment Simulator (JULES) included a global representation of crops, finding improvements in simulated seasonality of leaf area index (LAI), gross primary productivity (GPP), and canopy height, though no improvement of surface energy fluxes (Osborne et al., 2015). Given biogeochemical and biogeophysical effects of agriculture and agricultural management on terrestrial processes, our understanding of the climate tradeoffs of crop management would benefit from representation of crops and crop management within ESMs (Betts, 2005). However, to date no crop module represents the changing distributions of crop types and management through time.

Here, we document improvements made to the CLM Version 5.0 (CLM5) crop model, which is the land component of the Community Earth System Model (CESM) being used in CMIP6 experiments, including specific crop types (corn, soy, wheat, rice, cotton, and sugarcane), as well as fertilization and irrigation management practices. We present simulations that are not coupled to the atmosphere in order to best evaluate the model, as the simulations are run with observationally derived atmospheric conditions and will simulate crop life cycle and yields with more fidelity than a coupled model. Additionally, land-only simulations allow identification of direct changes in carbon, water, and energy fluxes without atmospheric feedbacks. Using the updated dynamic crop representation in CLM5, we examine how global crop yields vary spatially and temporally over the historical period (1850–2010). Simulated crop yields are evaluated against gridded data sets derived from downscaled country-level information. Additional simulations are performed to examine the effect on crop yields, productivity, and surface energy fluxes due to (1) crop-specific plant types, (2) the relative impact of crop area expansion and management practices (defined hereafter as fertilizer application and irrigation), and (3) changes in atmospheric CO₂ concentrations and nitrogen availability. This provides a better understanding of the
sensitivity of crops to potential environmental changes and how crop management affects yields and, at the same time, potentially affects the climate system.

2. Methods

2.1. Model Description

The CLM Version 5.0 (CLM5) is the land component of the CESM Version 2.0 (CESM2) and is described by (Lawrence et al., 2019). The CLM5 has numerous new developments compared to CLM4.5 that impact plant productivity, including changes to plant carbon and nutrient dynamics (Fisher et al., 2019), a new stomatal conductance calculation (Franks et al., 2018), the addition of plant hydraulics (Kennedy et al., 2019), and changes to soil hydrology (Lawrence et al., 2019; Swenson & Lawrence, 2014, 2015). In addition to the more general developments that change plant productivity, the CLM5 includes a prognostic crop model based on the Agro-IBIS model (Kucharik, 2009) that can simulate crop yields (Levis et al., 2016). The crop model uses the same physiology as the natural vegetation, including the Farquhar model of photosynthesis with acclimation to temperature (Lombardozzi et al., 2015), though with different crop-specific parameter values, phenology and allocation parameterizations, and fertilizer and irrigation management. The CLM5 crop model includes several new developments compared to CLM4.5, with four primary changes including (1) additional crop functional types, (2) active management of all crop areas, (3) updated fertilization and irrigation application, and (4) the capability for simulation of changing distributions of crops and crop management due to the introduction of dynamic land units (carbon, nitrogen, water, and energy are conserved during all transitions). Soil biogeochemical processes are simulated using a vertically resolved decomposition cascade (Koven et al., 2013), similar to the approach used in the Century model (Parton et al., 1988), with variable soil depths (Pelletier et al., 2016). While additional information is provided below, a full description of the CLM5 crop model and the developments since CLM4.5, as well as information about aspects of CLM5 not described here, can be found in the CLM5.0 Technical Description, which is available as an appendix to Lawrence et al. (2019; also see online: https://escomp.github.io/ctsm-docs/doc/build/html/tech_note/index.html).

CLM5 includes infrastructure to represent unique managed crop types, each split by water management strategy into rain-fed or irrigated crops. The spatial distribution is based on combining the Land Use Harmonization Version 2 (LUH2; Hurtt et al., 2011) with data sets created by Portmann et al. (2010). The LUH2 data set (Hurtt et al., 2011) includes spatial distributions of broad crop categories (C3 and C4; annual and perennial; nitrogen fixing) from 850 to 2015, estimating changes in crop area through time. The LUH2 crop categories are allocated into 31 major crop types, with rain-fed and irrigated fractions for each, using Portmann et al. (2010) crop distributions for 2000. The distributions of each crop in 2000 are then proportionally scaled through time based on the LUH2 crop area and categories. Each of the crops is represented on independent soil columns (Levis et al., 2016), thereby eliminating plant competition for water and nutrients that is the default assumption for natural vegetation in CLM5. Presently, CLM5 includes parameter values for eight unique crop types based on the availability of corresponding algorithms in Agro-IBIS and as developed by Badger and Dirmeyer (2015), including temperate corn, tropical corn, temperate soybean, tropical soybean, sugarcane, spring wheat, cotton, and rice (Levis et al., 2016). The remaining 23 crop types do not have associated parameters required for active management and are mapped to their closest analog of the current eight active crops (e.g., barley is simulated as spring wheat) so that global crop areas and crop-land expansion capture the dynamics associated with agricultural management (additional information provided in the CLM5 technical documentation; see Lawrence et al., 2019). Soybean, wheat, cotton, and rice crop types use the C3 photosynthetic pathway, while corn and sugarcane crop types use the C4 photosynthetic pathway. Soybean is the only nitrogen-fixing crop. All analyses focus on these eight crop types.

Crop phenological phases in CLM5 are triggered primarily by growing-degree-day threshold values and include four distinct triggers: planting, leaf emergence, grain fill, and harvest. Planting requires that the crop-specific growing-degree-day threshold is met during a crop-specific planting window. If the growing-degree-day threshold is not reached during the planting window, the crop will be planted on the last day of the planting window. Seed carbon is planted and stored for available growth until the onset of the leaf emergence phase, at which point the seed carbon is transferred to the leaf carbon pool. Nitrogen allocation during the leaf emergence phase follows that of natural vegetation and is supplied by the soil mineral
### 2.2. Simulation Descriptions

Model simulations used here were land-only simulations run using CLM5.0 with active biogeochemistry (Lawrence et al., 2019) at 1° resolution using the GSWP3 atmospheric forcing data set (http://hydro.iis.u-tokyo.ac.jp/GSWP3/) and are described in Table 1. The carbon pools were equilibrated to 1850 climate conditions before running for the historical time period (1850 through 2010; Lawrence et al., 2019). The control simulation included transient climate, CO2, nitrogen deposition, land cover change, and crop management (including irrigation and fertilization). It should be noted that climate data are only available starting in 1900, so the simulated time from 1850 to 1900 recycles the 1901 to 1920 climate data. To determine the impact of explicitly representing crops and management options, the control simulation was compared to a simulation that used the same configuration and forcings but without active crops, that is, with crops represented as generic C3 grasses with no irrigation or fertilization.

To isolate the impacts of crop expansion and management, four additional simulations were conducted, each using the same configuration and forcings as the control simulation but with individual options turned off as per the protocol in the Land Use Model Intercomparison Project (LUMIP, Lawrence et al., 2016). These simulations included (1) no irrigation, (2) no industrial fertilizer application, (3) no irrigation and no nitrogen pool. During the grain fill phase, allocation of carbon and nitrogen shifts from leaf and stem pools to the grain pool. Additionally, during the grain fill phase nitrogen is remobilized from the leaf and stem pools for all crops and also from the roots for wheat and rice and is reallocated to the grain pool (i.e., retranslocated). This is based on the observation that crops tend to reuse N from vegetative parts of the plant during grain fill to maximize N use efficiency and minimize external N demand (Masclaux-Daubresse et al., 2008).

When the crop is harvested, the carbon and nitrogen in the grain pool is removed and put into a grain product pool, with 3 g C m⁻² transferred from the grain to seed carbon pool for planting in the next year. The grain product pool carbon is respired to the atmosphere with an assumed turnover time of 1 yr to account for human and animal consumption. The remaining carbon and nitrogen stocks in the crop residue are assumed to remain in the field and are transferred to litter pools. In reality, there is variability as to whether crop residue is harvested or left on the field, and the impacts of residue management is an active area of research (e.g., Turmel et al., 2015).

Additional management capabilities in CLM5 include fertilization and irrigation. Fertilizer, which includes manure and industrial fertilizer, is added directly to the soil mineral nitrogen pool during the leaf emergence phase of crop development and continues for 20 days. Manure-based fertilizer is applied to all crop columns for the entire time series at a rate of 2 g N m⁻² yr⁻¹, a value consistent with the global application of 21 Tg N yr⁻¹ in 1850 (Riddick et al., 2016). While aerial rates do not change through time, total manure application does increase over the historical period due to cropland expansion. Industrial fertilizer is prescribed by crop type, year, and country based on LUH2 fertilization rates (Hurtt et al., 2011). More realistic development of manure application, including transient application rates and N fluxes, is under development (Riddick et al., 2016; Vira et al., 2019) but not included in the released version of CLM5. The application of irrigation water is limited to only the irrigated crop columns. Irrigation amount responds dynamically to the soil moisture conditions simulated in the model. Irrigation is applied if the crop leaf area is >0 and the soil moisture at 6:00 a.m. local time is below a specified threshold that is based on calculated target and wiltting point volumetric soil moisture values. Irrigated water is applied directly to the ground surface. Irrigation water is removed from the river water when there is enough water to meet irrigation demand. When irrigation demand exceeds available river water, water is diffusively drawn from the ocean in coupled simulations. CLM5 also includes options to limit irrigation water to that available in the rivers, but that feature is not used for this study. Ongoing CLM developments are focusing on a more comprehensive representation of irrigation water sources (rivers, reservoirs, groundwater, and aquifers), which would allow realistic treatment of irrigation limitations due to water availability.

### Table 1

**Description of CLM5 Simulations Used in This Study**

| Simulation       | Active crop representation | Irrigation | Industrial fertilization | Cropland expansion |
|------------------|----------------------------|------------|--------------------------|--------------------|
| Control          | Yes                        | Yes        | Yes                      | Yes                |
| Generic crops    | No                         | No         | No                       | Yes                |
| No irrigation    | Yes                        | No         | Yes                      | Yes                |
| No fertilization | Yes                        | Yes        | No                       | Yes                |
| No irrigation or fertilization | Yes | No | No | Yes |
| No cropland expansion | Yes | Yes | No | No |
| + CO₂            | Yes                        | Yes        | Yes                      | Yes                |
| + N enrichment   | Yes                        | Yes        | Yes                      | Yes                |

*Note: All simulations used the GSWP3 atmospheric forcing.*
industrial fertilizer application (which isolates impact of crop area expansion), and (4) constant 1850 crop area (note that constant crop area also implies limited irrigation and industrial fertilizer application).

In addition, we assessed the sensitivity of crops to elevated CO2 and N enrichment to determine whether crop responses are similar to results from manipulative experiments using simulations described by Wieder et al. (2019). These sensitivity tests were initiated from the control simulation starting in 1990 and run through 2010. One simulation was run by increasing atmospheric CO2 by 200 ppm above ambient. A second simulation increased N by adding 5 g N m\(^{-2}\) yr\(^{-1}\) above ambient N through the deposition stream, which was distributed evenly throughout the year. These perturbation values were chosen based on values frequently used in field manipulations, as described by Wieder et al. (2019).

2.3. Description of Model Evaluation

To assess CLM5 simulated crop production and yield, estimates from CLM5 were compared to observational data sets available from the United States Department of Agriculture National Agricultural Statistics Service (USDA-NASS; https://www.nass.usda.gov/Data_and_Statistics/index.php) and from the United Nations Food and Agriculture Organization Statistics (FAOSTAT; FAO 2017). Simulated production from individual crops in CLM5 was calculated as the annual sum of the grain to food carbon flux, which is the flux of carbon into the grain pool. We integrate the flux because monthly average grain pool values archived from CLM5 underestimate the maximum grain yield by averaging over longer times, including when yields are 0 after harvest.

In calculating observed crop production (total amount of grain, tonnes), we assume that grain carbon is 45% of the total dry weight (Monfreda et al., 2008) and an 85% harvest efficiency (Kucharik & Brye, 2003) assuming that a portion of grain is lost due to weather, labor, and machine inefficiencies. Crop yields (grain per crop area, t ha\(^{-1}\)) are calculated by dividing the production by the total crop area.

The USDA-NASS data set included crop production, yield, and area harvested for each county within the United States from 1910 to 2016 for several crops, including corn, soybean, and wheat—the dominant crops in the United States that were used for comparison to CLM5. County-level USDA-NASS data were aggregated to the 1° resolution of the CLM5 simulations, and comparisons focused on U.S. yields from 1990 to 2010.

The FAOSTAT data set included area harvested, production, and yield globally for numerous crops by country from 1961 to 2016. To better understand the spatial distribution within each country, the FAOSTAT data were downscaled using observed crop distribution and yield data from EarthStat (http://www.earthstat.org/harvested-area-yield-175-crops/), a global data product that combines national, state, and county-level data to provide yield, production, and harvested area data at 10 km spatial resolution for 175 different crop types for the Year 2000 (Monfreda et al., 2008). Both the CLM5 and the downscaled FAOSTAT data were compared to the USDA-NASS data for the United States from 1990 to 2010 to evaluate the accuracy of downsampling using EarthStat data. To evaluate global temporal trends and spatial patterns of crop production and yield, CLM5 simulations were compared to the downscaled FAO-EarthStat data product.

3. Results and Discussion

3.1. Evaluation of CLM5 Crop Model

Global crop production estimated by CLM5 for the actively represented crop types increased nearly eightfold from ~300 Mt in 1850 to ~2,300 Mt in 2010 (Figure 1a), and global crop yield increased from ~1.1 t ha\(^{-1}\) in 1850 to 3 t ha\(^{-1}\) in 2010 (Figure 1b). Trends in global crop production account for changes in aerial expansion and management practices, whereas trends in global crop yield isolate the impact of management practices through standardizing by crop area. In CLM5 simulations, both production and yield gradually increased from 1850 until the early 1950s, at which time production and yield began to increase more rapidly in response to widespread use of industrial fertilizer. CLM5 global crop yields and production were similar to estimates from FAO-EarthStat for the same crop types from 1961, the first year available, to the mid-1990s. Production and yield estimates for these crop types from FAO-EarthStat continued to increase linearly through 2010, with production at nearly 3,250 Mt and yields of approximately 4.1 t ha\(^{-1}\), approximately 25% higher than CLM5 estimates. CLM5 diverged from these observed estimates in the mid-1990s when yields and production level off.
Observations suggest that yield increases primarily occur in high-yield environments due to intensification and specialization (Assefa et al., 2017; Grassini et al., 2013; Iizumi et al., 2014). The stagnation in CLM5 crop yields starting in the 1990s (Figure 1) suggests important recent advances in agricultural technology in high-yield environments, such as changes in planting density; advancement in sowing dates due to reduced tillage, glyphosate-resistant cultivars, and large equipment; and other advances in cultivar productivity—processes which CLM5 does not represent—dominate the trends in increased yield (e.g., Assefa et al., 2016, 2017; Fox et al., 2013; Grassini et al., 2013; Hammer et al., 2009). The 25% difference in global yields between CLM5 and UNFAO by 2010 (Figure 1) is similar in magnitude to the observationally estimated 24% global yield increase due to intensification (Rudel et al., 2009). Developing representations of intensification processes should therefore be a priority for future model development.

Spatial patterns of individual crop yields simulated by CLM5 are similar to downscaled FAO-EarthStat crop yields in many regions (Figures 2 and 3) and to USDA-NASS yields in the United States (Figures S4–S6). In general, CLM5 underestimates yields in temperate regions and overestimates yields in tropical regions (Figures 2 and 3; root-mean-square error [RMSE] and Bias in Figures S2 and S3). The net balance of high and low biases is evident in global trends, which illustrate that CLM5 underestimates wheat, rice, cotton, and sugarcane production and yield by 20–45% and overestimates global soybean and corn yields (~2 and ~3 ton ha\(^{-1}\), respectively) compared to FAO-EarthStat data (Figure S1). The static global parameterization used for each crop type in CLM5 likely oversimplifies the diversity of crop cultivars that are optimized for different environments. Future model development should broadly focus on refining crop model parameterizations, which could consider how to represent spatially explicit parameters that can account for cultivar differences and crop breeding.

Corn yields are generally overestimated in tropical regions and underestimated in temperate regions, whereas soybean yields are overestimated in most regions (Figures 2 and 3; RMSE and Bias in Figures S2 and S3). CLM currently does not account (or only simplistically accounts) for the impacts of environmental threats like extreme weather or crop pests and pathogens and is therefore unlikely to capture the variability and/or collapses in yields observed in several regions (Iizumi et al., 2014), potentially contributing to overestimated yields in tropical regions (Figures 2, 3, S2, and S3). CLM5 underestimates corn yields in most temperate regions compared to both FAO-EarthStat and USDA-NASS observational data products (Figures 3 and S4), due in part to not representing processes associated with intensification (Assefa et al., 2016, 2017). Both CLM5 and the downscaled FAO-EarthStat data product overestimate soybean yields in some parts of the central United States (e.g., Ohio and Missouri) compared to USDA-NASS (Figure S5), similar to an older
version of CLM-Crop (Drewniak et al., 2013). Uncertainties in all crop yields simulated by CLM5 are related to the representation of allocation and temperature-triggered phenology, as well as biological N fixation for soybeans. It should be noted that soybean and cotton RMSE and bias estimates are generally smaller than for other crops due to the smaller magnitude of maximum yields (Figures S2 and S3).

Globally, wheat yields are underestimated in many temperate regions, including in parts of India, Europe, and western North America. However, where county-level data are available from USDA-NASS, wheat yields (Figure S6) are overestimated in CLM5 as well as in the FAO-EarthStat data product, particularly...
in the eastern and central United States. Wheat can be troublesome to simulate in global crop models due to the fact that spring and winter wheat varieties are often grown within the same region (Müller et al., 2017), and data sets distinguishing where each variety is grown are not readily available. Although a winter wheat parameterization is in development, CLM5 currently only uses spring wheat phenology to simulate all wheat productivity, potentially contributing to inaccuracies in regions dominated by winter wheat, such as India.

Rice yields are overestimated throughout most tropical regions in CLM5 (e.g., high biases and RMSE in South America and Africa; Figures S2 and S3). However, they are too low in India and Southeast Asia compared to FAO-EarthStat. Yields of both cotton and sugarcane, new crop types within CLM5, are underestimated in many regions (Figure S3), though overestimated in parts of South America and Africa. There are some mismatches between observed and modeled spatial distribution of certain crops that may impact production, and biases are not reported for points where either observed or simulated yield was unavailable. Given that many globally gridded crop models have difficulty simulating rice yields and often do not include cotton or sugarcane (Müller et al., 2017), it is perhaps not surprising that yields of these crops are globally underestimated (Figure S1).

Inaccuracies in crop yields may be due to a variety of additional reasons that are likely different based on the region and crop type. For example, phenology, which is based on accumulated temperatures, is one potential source of error. The phenological parameters are globally fixed for each crop type, and the temperature thresholds to trigger planting have no dependence on water availability or technological advancements. Therefore, crops can be planted even if soil water is insufficient for the crop to establish, which is likely what causes such low yields in India (see Figure S7). Adding a soil moisture trigger or a spatially specific planting window may improve yields in India by changing the planting date and extending the growing season, allowing for crop growth when environmental conditions are more favorable. Accurately representing crop phenology is critical, not only for yields, but also because the specific crop phenology (relative to natural

**Figure 3.** Distributions of grain yields averaged from 1991 to 2010 estimated from United Nations Food and Agriculture Organization Statistics (FAOSTAT) and downscaled to 1° resolution using EarthSTAT data (see section 2 for more details; a, c) and CLM5 (b, d) for all C₄ crops that are actively managed in CLM5, including corn (a, b) and sugarcane (c, d).
grasslands) can affect the overall latent heat flux (Figure 7; also Chen et al., 2015; Sacks & Kucharik, 2011) and therefore temperature and precipitation when coupled to an atmosphere model (Levis et al., 2012). Additionally, the representation of carbon and nitrogen allocation can cause inaccuracies. If allocation to leaves is too high, for example, crop biomass requirements can lead to the crop rapidly using available soil N, resulting in nitrogen limitation during the grain fill phase of development, reducing crop yields. Peng et al. (2018) show that improving the representation of corn phenology and allocation in the CLM4.5 crop model improves yields, suggesting that both are important for accurate estimation of yields. The amount of fertilizer and irrigation applied to crops, which is homogenized across larger regions in global-scale models, can also affect water stress and plant nitrogen uptake and use, changing photosynthesis and therefore allocation.

3.2. Impact of Explicit Crop Representation

Including specific crop types and crop management in CLM increased annually averaged (1990–2010) GPP in several Northern Hemisphere regions relative to a representation of crops as grasses (Figure 4a). Maximum monthly averaged GPP increases even more strongly, with grid cell average increases of up to 2 g C m\(^{-2}\) day\(^{-1}\) (>25%) in heavily cropped regions like midwestern United States, Europe, and eastern China (Figure 4b). High annual maximum productivity in crop regions has been previously observed from satellite data products like solar-induced fluorescence (SIF; Guanter et al., 2014) and is often underestimated by ecosystem models, particularly when crops are represented as grasses, in these regions (Lokupitiya et al., 2016; Osborne et al., 2015).

Annual average and peak productivity decreased in many tropical regions, including sub-Saharan Africa and India, due to the timing of planting coinciding with low water availability (Figure 4). Consequently, the actively managed crops were not able to establish, and LAI remained low throughout the growing season, whereas the generic crops that do not rely on phenological planting triggers grew well during wet periods (see Figure S7). Future model development should focus on improving the phenological triggers for crop growth and account for water availability and regional variation, as this can have a large impact on the simulated land-atmosphere exchanges of water and carbon dioxide (Chen et al., 2015; Levis et al., 2012).

The peak productivity increases with explicit crop representation led to a larger terrestrial drawdown of carbon (e.g., negative net ecosystem exchange), increasing the amplitude of net ecosystem exchange north of 30°N, averaged from 1990 to 2010 (Figure 5), which will likely increase the amplitude of the atmospheric CO\(_2\) annual cycle when incorporated into an Earth system model like the CESM (Levis et al., 2012). This result is consistent with studies suggesting that increased agricultural productivity has contributed to the increased amplitude of the CO\(_2\) annual cycle observed in atmospheric CO\(_2\) measurements (Gray et al., 2014; Zeng et al., 2014). Currently, most ESMs do not capture this behavior (Graven et al., 2013), but the crop-driven amplification of annual CO\(_2\) exchange with the atmosphere (Figure 5) suggests that explicitly representing agriculture is an important feature in global carbon cycle simulations.

Historically, agricultural practices have reportedly decreased the amount of carbon stored in cropland soils over time (Sanderman et al., 2017). For most crop regions in CLM5, we see the opposite signal, with soil carbon increasing rather than decreasing (Figure 6) due in part to larger carbon inputs from higher productivity. The major reason for this discrepancy, however, is that CLM5 does not represent soil tillage, which enhances decomposition and leads to soil C losses (Haddaway et al., 2016; Levis et al., 2014; Lobell et al., 2006; Paustian et al., 2016). A recent estimate suggests that soil tillage contributed to accumulated losses of ∼40 Pg C over the past 10,000 yr (Sanderman et al., 2017). Including tillage into CLM is a high priority as it will also enable simulation of minimum- or no-tillage techniques and increased plant or litter additions that are promoted as climate-smart management practices because they can increase soil carbon storage (FAO, 2013; Jin et al., 2017; Paustian et al., 2016). A variety of other assumptions and simplifications in the model could also drive increased soil carbon in many crop regions. For example, actively managed crops in CLM increase productivity (Figure 4), concurrently increasing carbon fluxes from plant residue into the soil. Thus, overall increases in soil carbon (Figure 6) largely mirror the spatial patterns of annual average productivity change associated with explicit crop representation (Figure 4). Productivity biases, or not removing enough nongrain biomass with harvest, could drive unrealistic simulated increases in soil carbon. Future model developments should include a representation of residue management, varying the proportions of leaf and stem biomass pools transferred to the soil litter based on the residue management
strategy. There were some regions where CLM simulated decreased soil carbon, occurring where phenological triggers for planting were not aligned with water availability, similar to the response of productivity discussed above.

Crops can also impact climate through changing energy fluxes between the land surface and atmosphere (Figure 7). Explicitly representing crops in CLM5 increases the maximum monthly average latent cooling throughout most crop regions (Figure 7d; Levis et al., 2012), although it decreases latent heat fluxes in regions like India, where planting does not coincide with water availability. Increased latent heat fluxes in most regions are due to higher GPP and therefore transpiration, as well as the addition and evapotranspiration of irrigation water. Other studies have also shown that crop intensification led to cooler summer temperatures, especially in irrigated croplands (Mueller et al., 2015; Thiery et al., 2017), and has reduced summer temperature extremes over the last 50 yr (Mueller et al., 2017). Maximum changes in latent heat flux, however, were different from annual averaged changes in North and South America. The annual average latent heat flux decreases in these regions with explicit representation of crops (Figure 7c), possibly due to a depletion of soil water available for evapotranspiration. This is potentially caused by higher water use

Figure 4. The change in gross primary productivity (GPP) when crops are actively managed compared to crops simulated as grasses without any management. Panel (a) illustrates the change in annual averages, and panel (b) illustrates the change in the maximum monthly value each year. All data are averaged over 1990–2010, and all changes plotted are statistically significant at \( p < 0.05 \).
during the peak growing season leaving less available water for evapotranspiration during the remainder of the year or the fact that generic crops may have a longer growing season and therefore transpire more water over the course of the year (Levis et al., 2012). The decreased sensible heat fluxes in regions when crops are explicitly represented are likely caused by changes in energy partitioning (Figures 7a and 7b). While changes in roughness and albedo may also contribute to differences in sensible heat flux, these influences are likely small due to the fact that generic crops and explicit crops have similar roughness lengths and albedos.

3.3. Sensitivity to Areal Expansion and Management

Using CLM5, we are able to determine how changes in transient forcings, cropland expansion, fertilizer application, and irrigation contributed to the simulated increases in crop yields over the historical period (1850–2010; Figure 1). We were unable to determine the impact of other management strategies (e.g., herbicide use, crop genetics, planting density, and other aspects of intensification) on crop yields because CLM5 does not represent these processes. The changes in production between 1850 and 2010 in the simulation without cropland expansion (Figure 8, red line), which also limits fertilizer and irrigation to 1850 values, illustrates that transient forcings (e.g., CO₂, temperature, precipitation, irradiance, and N deposition) have a relatively small influence on global food production, highlighting the importance of expansion and management practices.

The expansion of crop area (Figure 8, gray line) has considerably contributed to increased crop production (Rudel et al., 2009), with CLM5 estimating a relatively linear increase of 712 Mt for active crop types over the last 160 yr.

The global effect of fertilizer is larger than crop expansion, increasing yields by 807 Mt (Figure 8, yellow line). Effects of fertilization began in the 1950s, when chemical fertilizer application became commonplace.

**Figure 5.** Net ecosystem exchange annual cycle averaged north of 30°N for a simulation where crops were actively managed (purple) and a simulation where crops were simulated as grasses without any management (generic crops; green). Data were averaged from 1990 to 2010, and shading represents ±1 standard deviation.

**Figure 6.** The change in annual average total soil organic matter carbon when crops are actively managed compared to crops simulated as grasses without any management. Data were averaged over 1990–2010, and all changes plotted are statistically significant at \( p < 0.05 \).
Increasing manure application (e.g., Riddick et al., 2016) is also expected to increase yields prior to 1950, although changes in the aerial application rate are not represented in CLM5. More robust manure application schemes are under development (Vira et al., 2019). Corn, wheat, and rice yields increased most with fertilizer application, whereas nitrogen-fixing soybean yields, as expected, did not change much with fertilizer application (Figure S8). Fertilizer application in CLM5 only supplies nitrogen because limitation by other nutrients, like phosphorus and potassium, is not represented. Fertilizer application in industrial regions is typically applied to maximize crop yields, and in many regions, such as China where fertilizer is applied beyond saturation, nitrogen is commonly lost to leaching and trace gas emissions (Ju et al., 2009). As with previous versions of the model, the representation of N losses from natural and managed ecosystems are poorly represented in CLM (Houlton et al., 2015; Thomas et al., 2013) and remains a priority for future model development.

The effect of irrigation is somewhat smaller than fertilization, increasing global production by 522 Mt by 2010. Since both fertilizer and irrigation are essential for crop growth, the effect of each is dampened by 170 Mt if one of these management activities is not included (Figure 8). Irrigation effects become more pronounced in the 1980s, when the difference between the simulations with and without irrigation is largest (Figure 8, blue line), and could possibly be due to a higher water demand associated with warmer temperatures (e.g., Levis et al., 2016). Irrigation has a similar and spatially homogenous impact on all crop types (Figure S9). The smaller overall effect of irrigation on global production is due to the fact that by 2010 only ~25% of global crop area is irrigated. Indeed, Müller et al. (2018) suggest that irrigation will be important in the future to increase global crop yields and also stabilize yields in regions with high yield variability, although it should be noted that the ability to expand irrigation in reality is severely constrained by water resource availability. Thus, developments in CLM are focused on including a more sophisticated representation of irrigation and water resource management.

**Figure 7.** The change in sensible (a, b) and latent heat fluxes (c, d) when crops are actively managed compared to crops simulated as grasses without any management. Panels (a) and (c) illustrate the change in annual averages, and panels (b) and (d) illustrate the change in the maximum monthly value each year. All data are averaged over 1990–2010, and all changes plotted are statistically significant at $p < 0.05$. 

Increasing manure application (e.g., Riddick et al., 2016) is also expected to increase yields prior to 1950, although changes in the aerial application rate are not represented in CLM5. More robust manure application schemes are under development (Vira et al., 2019). Corn, wheat, and rice yields increased most with fertilizer application, whereas nitrogen-fixing soybean yields, as expected, did not change much with fertilizer application (Figure S8). Fertilizer application in CLM5 only supplies nitrogen because limitation by other nutrients, like phosphorus and potassium, is not represented. Fertilizer application in industrial regions is typically applied to maximize crop yields, and in many regions, such as China where fertilizer is applied beyond saturation, nitrogen is commonly lost to leaching and trace gas emissions (Ju et al., 2009). As with previous versions of the model, the representation of N losses from natural and managed ecosystems are poorly represented in CLM (Houlton et al., 2015; Thomas et al., 2013) and remains a priority for future model development.

The effect of irrigation is somewhat smaller than fertilization, increasing global production by 522 Mt by 2010. Since both fertilizer and irrigation are essential for crop growth, the effect of each is dampened by 170 Mt if one of these management activities is not included (Figure 8). Irrigation effects become more pronounced in the 1980s, when the difference between the simulations with and without irrigation is largest (Figure 8, blue line), and could possibly be due to a higher water demand associated with warmer temperatures (e.g., Levis et al., 2016). Irrigation has a similar and spatially homogenous impact on all crop types (Figure S9). The smaller overall effect of irrigation on global production is due to the fact that by 2010 only ~25% of global crop area is irrigated. Indeed, Müller et al. (2018) suggest that irrigation will be important in the future to increase global crop yields and also stabilize yields in regions with high yield variability, although it should be noted that the ability to expand irrigation in reality is severely constrained by water resource availability. Thus, developments in CLM are focused on including a more sophisticated representation of irrigation and water resource management.
Effects of expansion and particular management strategies have local distinct effects on the water and energy budget of particular regions. For example, conversion of natural vegetation to cropland and the use of irrigation can increase regional cooling from latent heat fluxes associated with agriculture (section 3.2; Kueppers & Snyder, 2011; Zeng et al., 2017). In our simulations, latent heat flux declined in Jiangsu, China, when irrigation is excluded and when cropland expansion remains constant (Figure 9a), despite the relatively small crop expansion (Table 2). Expansion and management have an even larger impact on latent heat flux in Punjab, India, where peak latent heat flux decreases considerably without cropland expansion and declines even more when irrigation is excluded (Figure 9b). The changes caused by transient atmospheric forcing alone (including changes in CO₂, temperature, and other forcings) is calculated as the difference in 2010 from 1850 for the simulation without changes in land use or any management options (red line).

Effects of expansion and particular management strategies have local distinct effects on the water and energy budget of particular regions. For example, conversion of natural vegetation to cropland and the use of irrigation can increase regional cooling from latent heat fluxes associated with agriculture (section 3.2; Kueppers & Snyder, 2011; Zeng et al., 2017). In our simulations, latent heat flux declined in Jiangsu, China, when irrigation is excluded and when cropland expansion remains constant (Figure 9a), despite the relatively small crop expansion (Table 2). Expansion and management have an even larger impact on latent heat flux in Punjab, India, where peak latent heat flux decreases considerably without cropland expansion and declines even more when irrigation is excluded (Figure 9b). The changes caused by transient atmospheric forcing alone (including changes in CO₂, temperature, and other forcings) is calculated as the difference in 2010 from 1850 for the simulation without changes in land use or any management options (red line).

3.4. Sensitivity to Perturbations of Atmospheric CO₂ and Nitrogen Enrichment

We conducted CO₂ and N enrichment simulations to evaluate model sensitivities with observations from similar experimental manipulations. The simulation that increased CO₂ by 200 ppm, similar to the increases...
applied in Free Air CO2 enrichment (FACE) studies, globally increases crop yields by 27% for rain-fed crops and by 36% for irrigated crops (Figure 10). The increased yield is similar in magnitude to the increases in crop GPP (32% for rain-fed, 41% for irrigated; data not shown). The magnitude of the yield response is nearly double observations from FACE experiments, which estimate that crop yields increase by 17–19% in response to elevated CO2 (Ainsworth & Long, 2004; Bishop et al., 2014). Further, the larger increase found in irrigated crops is the opposite of observations, which show that the effect of CO2 decreases with water availability (Bishop et al., 2014). Several studies suggest that C4 crops, such as corn and sugarcane, are less sensitive to elevated CO2 than other crops (Kim et al., 2006; Leakey, 2009; Ruiz-Vera et al., 2015). While CLM5 simulations show that sugarcane, as expected, is largely insensitive to elevated CO2, the global change in corn yield (1.41 and 1.43 for rain-fed and irrigated, respectively; Figure 10) is much larger than observations.

Table 2
Changes in Cropland Expansion (Increase in % Crop Area) and Fertilizer Application Between 1850 and 2010

|                | Expansion (%) | Fertilizer (g N m⁻² yr⁻¹) |
|----------------|---------------|----------------------------|
| Jiangsu, China | 5             | 50                         |
| Punjab, India  | 31            | 47                         |
| Central Valley, CA | 13       | 25                         |
| Midwest United States | 31         | 1                          |

Figure 9. Average latent heat flux annual cycle for a 1° grid cell point in Jiangsu, China (a), Punjab, India (b), the Midwestern United States (c), and California’s Central Valley (d) for the CLM5 with all land use and management options active (purple line), without fertilization (orange line), without irrigation (blue line), and without cropland expansion (gray line). Data were averaged from 1990 to 2010.

10.1029/2019JG005529
suggest (1.10; Bishop et al., 2014). The simulated effect of CO2 on rain-fed wheat, soy, cotton, and rice yields ranged from 1.27 to 1.36, while irrigation had a relatively larger effect on rice (1.78) and soy (1.43) yields. These changes are larger than the average 12–23% increased yields typically observed in FACE experiments for soy, rice, and wheat, and fall near the high end of the observed confidence intervals (1.27 for rice and 1.35 for wheat; Ainsworth & Long, 2004; Ainsworth et al., 2008; Bishop et al., 2014; Cai et al., 2015; Long et al., 2006).

Although the simulated response to CO2 enrichment in CLM5 seems more appropriate than previous versions of CLM (Wieder et al., 2019), particular model assumptions likely need to be revised. For example, observations suggest that elevated CO2 changes leaf carbon allocation, such that leaf mass per area (LMA) changes more than LAI (Medlyn et al., 2015). That is, leaves under elevated CO2 tend to get thicker more than whole plant canopies get leafier, but land models—including CLM—struggle to capture this behavior (Kovenock & Swann, 2018; Zaehle et al., 2014). Allowing for changes in leaf morphology should be a priority for future model development. Similar to the response of crops found here, temperate C3 grasses also show a strong sensitivity to elevated CO2 (Wieder et al., 2019), suggesting that changes in leaf carbon allocation may be critical to accurately represent plant physiological response to elevated CO2. These dynamics may be especially important in the CLM crop model, which assigns a maximum value on crop LAI. Under elevated CO2 plants likely hit this peak value more quickly in the growing season so that more carbon and nitrogen are allocated to the grain fill phenological phase, possibly contributing to the larger than observed increases in crop yields simulated here. However, given the similar responses of crop GPP and yield to elevated CO2, it is more likely that assimilation rates are causing increased yields.

Nitrogen enrichment increased rain-fed and irrigated crop yields by 33% and 39%, respectively, suggesting that crop yields in CLM5 are N limited even with fertilizer application (Figure 10b). The magnitude of yield increase under N enrichment was smaller than that of CO2, consistent with global results from CLM5 (Wieder et al., 2019), but with additional N, yields of all crops, even N-fixing soy, increased globally, which is unrealistic. The effect of additional N was largest for sugarcane and rice (1.52 and 1.92, respectively, for irrigated varieties). The increased yields in response to additional N are likely due to the method of fertilizer application in CLM, which applies fertilizer over the first 20 days after crops are planted. If the crops use most of the N during leaf emergence, N might be limiting during the allocation to grain, which happens later in the growing season. The N enrichment applied here was applied consistently throughout the year (through the N deposition stream), suggesting that adding N during other crop phenological phases may boost yields.
4. Conclusions

The growing demand for food production from Earth's human population presents a challenge to society due to limited resources such as arable land, water, energy, and the need to reduce greenhouse gas emissions. To help meet this demand, a greater understanding of the interactions between agricultural management and the climate system is needed. By representing agricultural management in models like CLM5, changes in crop productivity due to expansion and specific management practices under the conditions of changing climate, and feedbacks to the climate system, can be quantified. Here, we document the explicit representation of crops in CLM5 using land-only simulations to show that this representation can feasibly simulate crop yields. This analysis illustrates that crops and crop management practices directly change spatial and temporal patterns of productivity and latent and sensible heat fluxes, highlighting the importance of including dynamic crop representation in Earth system models for capturing changes in carbon, water, and energy exchanges that will impact climate. Additionally, cropland expansion, industrial fertilizer, and irrigation application have been essential for improving crop yields since 1850, with yield increases from 1850 to 1950 driven primarily by cropland expansion and fertilizer and irrigation emerging as important drivers of yield increases since 1950.

Part of the challenge in explicitly representing crops in Earth system models like CESM stems from limited data availability as well as the ability to globally represent phenological and management processes. New advances in data availability and the development of parameterizations (e.g., Erb et al., 2017; Pongratz et al., 2018) will continue to allow for improved agricultural representation in ESMs. Near-term developments within CLM should target improving crop phenological representation by explicitly representing more crop types, connecting and expanding phenological triggers to respond to water availability and heat stress, and representing spatially varying planting windows. Additionally, the representation of agricultural management can be improved in CLM by developing representations for residue management and soil tillage, allowing fertilizer and manure application to be more flexible, and including more realistic limits on irrigation and more comprehensive irrigation methods and timing. Longer-term developments can focus on processes where data are not as readily available, including representing major technological advances associated with recent intensification, such as higher planting densities, earlier sowing dates, and genetic gain.

Although there is still room for improvement, we demonstrate that the crop model embedded within CLM5 exhibits a significant advance by including temporally changing spatial distributions of the world's major crops, irrigation and fertilization, and crop-specific allocation and phenological triggers for the six most common crops types, which are capabilities that are not commonly included in Earth system models. The inclusion of crops within models like CLM5 is not intended to replace the capabilities of agronomic models that help farmers make management decisions and provide accurate estimates of crop yields. Instead, these models aid in our understanding of how agricultural management impacts, and potentially mitigates, climate change by including the feedbacks between climate change, atmospheric CO2 concentrations, and crop behavior. Combining capabilities from both classes of models will allow for a better understanding of how managed ecosystems have historically and will continue to impact Earth’s climate as humanity strives to produce enough food to sustain a growing population.

Data Availability Statement

Data are archived and made publicly available at the UCAR/NCAR Climate Data Gateway (doi:10.5065/d6154fwh). CLM5.0 is publicly available through the Community Terrestrial System Model (CTSM) repository (https://zenodo.org/record/3779821#.XsRzrhNKjjB).

References

Ainsworth, E. A., & Long, S. P. (2004). What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. *New Phytologist, 162*(2), 351–372. https://doi.org/10.1111/j.1469-8137.2004.01224.x

Ainsworth, E. A., Rogers, A., & Leakey, A. D. B. (2008). Targets for crop biotechnology in a future high-CO2 and high-O3 world. *Plant Physiology, 147*(1), 13–19. https://doi.org/10.1104/pp.108.117101

Assefa, Y., Prasad, P. V. V., Carter, P., Hinde, M., Bhalla, G., Schon, R., et al. (2017). A new insight into corn yield: Trends from 1987 through 2015. *Crop Science, 57*(3), 2799–2813. https://doi.org/10.2135/cropsci2017.01.0066
