Lawrence David M. (Orcid ID: 0000-0002-2968-3023)
Fisher Rosie A. (Orcid ID: 0000-0003-3260-9227)
Koven Charles D. (Orcid ID: 0000-0002-3367-0065)
Oleson Keith W. (Orcid ID: 0000-0002-0057-9900)
Swenson Sean C. (Orcid ID: 0000-0002-2923-1203)
Bonan Gordon (Orcid ID: 0000-0003-0830-6437)
Collier Nathan (Orcid ID: 0000-0002-7367-3981)
van Kampenhout Leo (Orcid ID: 0000-0001-9047-5031)
Kennedy Daniel (Orcid ID: 0000-0001-9494-3509)
Lawrence Peter J. (Orcid ID: 0000-0002-4843-4903)
Li Fang (Orcid ID: 0000-0002-3686-2257)
Lombardozzi Danica (Orcid ID: 0000-0003-3557-7929)
Riley William J. (Orcid ID: 0000-0002-4615-2304)
Shi Mingjie (Orcid ID: 0000-0002-2469-4831)
Vertenstein Mariana (Orcid ID: 0000-0001-9486-9994)
Wieder William R. (Orcid ID: 0000-0001-7116-1985)
Ali Ashehad A. (Orcid ID: 0000-0002-7328-306X)
Badger Andrew, M. (Orcid ID: 0000-0003-4537-9993)
Bisht Gautam (Orcid ID: 0000-0001-6641-7595)
van den Broeke Michiel (Orcid ID: 0000-0003-4662-7565)
Brunke Michael A. (Orcid ID: 0000-0002-2677-7446)
Buzan Jonathan (Orcid ID: 0000-0002-8199-3264)
Clark Martyn (Orcid ID: 0000-0002-2186-2625)
Drewniak Beth (Orcid ID: 0000-0002-8193-6091)
Fisher Joshua B. (Orcid ID: 0000-0003-4734-9085)
Flanner Mark (Orcid ID: 0000-0003-4012-174X)
Fox Andrew M. (Orcid ID: 0000-0002-4027-4538)
Gentine Pierre (Orcid ID: 0000-0002-0845-8345)
Hoffman Forrest (Orcid ID: 0000-0001-5802-4134)
Keppel-Aleks Gretchen (Orcid ID: 0000-0003-2213-0044)
Kumar Sanjiv (Orcid ID: 0000-0002-0472-6074)
Lenaerts Jan (Orcid ID: 0000-0003-4309-4011)
Leung L. Ruby (Orcid ID: 0000-0002-3221-9467)

The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty

David M. Lawrence[1], Rosie A. Fisher[1], Charles D. Koven[2], Keith W. Oleson[1], Sean C. Swenson[1]

Gordon Bonan[1], Nathan Collier[3], Bardan Ghimire[2], Leo van Kampenhout[4], Daniel Kennedy[5], Erik Kluzek[1], Peter J. Lawrence[1], Fang Li[6], Hongyi Li[7], Danica Lombardozzi[1], William J. Riley[2], William J. Sacks[1], Mingjie Shi[8,9], Mariana Vertenstein[1], William R. Wieder[1,18], Chonggang Xu[10]

Ashehad A. Ali[11], Andrew M. Badger[12], Gautam Bisht[2], Michiel van den Broeke[4], Michael A. Brunke[13], Sean P. Burns[14,35], Jonathan Buzan[15], Martyn Clark[1], Anthony Craig[1], Kyla Dahlin[16], Beth Drewniak[17], Joshua B. Fisher[8,9], Mark Flanner[19], Andrew M. Fox[20], Pierre Gentine[5], Forrest Hoffman[3], Gretchen

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2018MS001583

Keppel-Aleks[21], Ryan Knox[2], Sanjiv Kumar[22], Jan Lenaerts[23], L. Ruby Leung[24], William H. Lipscomb[1], Yaqiong Lu[25], Ashutosh Pandey[22], Jon D. Pelletier[26], Justin Perket[1,27], James T. Randerson[28], Daniel M. Ricciuto[29], Benjamin M. Sanderson[30], Andrew Slater[31], Zachary M. Subin[32], Jinyun Tang[2], R. Quinn Thomas[33], Maria Val Martin[34], Xubin Zeng[13]

[1] Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA
[2] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3] Computer Science and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[4] Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, The Netherlands
[5] Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA
[6] International Center for Climate and Environmental Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
[7] University of Houston, Houston, TX, USA
[8] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[9] Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, Los Angeles, CA, USA
[10] Los Alamos National Laboratory, Los Alamos, NM, USA
[11] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
[12] Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA
[13] Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA
[14] Department of Geography, University of Colorado, Boulder, CO, USA
[15] Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, Indiana, USA
[16] Department of Geography, Environment, and Spatial Sciences, Michigan State University, East Lansing, MI, USA
[17] Environmental Science Division, Argonne National Laboratory, Argonne, IL, USA
[18] Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA
[19] Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA
[20] School of Natural Resources and the Environment, University of Arizona, Tucson, USA
[21] Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, CA, USA
[22] School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA
[23] Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, CO, USA
[24] Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA
[25] Chinese Academy of Sciences, Chengdu, China
[26] Department of Geosciences, University of Arizona, Tucson, AZ, USA
[27] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[28] Department of Earth System Science, University of California, Irvine, CA
[29] Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN
[30] CERFACS, Toulouse, France
[31] National Snow and Ice Data Center, Boulder, CO, USA
[32] Energy & Environmental Economics, San Francisco, CA, USA
[33] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA
[34] Leverhulme Centre for Climate Change Mitigation, Animal and Plant Sciences Department, University of Sheffield, Sheffield, UK
[35] Mesoscale and Microscale Meteorology Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

Corresponding author: David Lawrence (dlawren@ucar.edu)

Key Points (up to 3):

- Updated Community Land Model has more hydrological and ecological process fidelity and more comprehensive representation of land management
- Model systematically evaluated using the International Land Model Benchmarking system, showing marked improvement over prior versions.

Keywords: Global land model, Earth System Modeling, carbon and nitrogen cycling, hydrology, benchmarking

Submitted to Journal of Advances in Modeling Earth Systems

December 3, 2018

Lawrence David M. (Orcid ID: 0000-0002-2968-3023)
Fisher Rosie A. (Orcid ID: 0000-0003-3260-9227)
Koven Charles D. (Orcid ID: 0000-0002-3367-0065)
Oleson Keith W. (Orcid ID: 0000-0002-0057-9900)
Swenson Sean C. (Orcid ID: 0000-0002-2923-1203)
Bonan Gordon (Orcid ID: 0000-0003-0830-6437)
Collier Nathan (Orcid ID: 0000-0002-7367-3981)
van Kampenhout Leo (Orcid ID: 0000-0001-9047-5031)
Kennedy Daniel (Orcid ID: 0000-0001-9494-3509)
Lawrence Peter J. (Orcid ID: 0000-0002-4843-4903)
Li Fang (Orcid ID: 0000-0002-3686-2257)
Lombardozzi Danica (Orcid ID: 0000-0003-3557-7929)
Riley William J. (Orcid ID: 0000-0002-4615-2304)
Shi Mingjie (Orcid ID: 0000-0002-2469-4831)
Vertenstein Mariana (Orcid ID: 0000-0001-9486-9994)
Wieder William R. (Orcid ID: 0000-0001-7116-1985)
Ali Ashehad A. (Orcid ID: 0000-0002-7328-306X)
Badger Andrew, M. (Orcid ID: 0000-0003-4537-9993)
Bisht Gautam (Orcid ID: 0000-0001-6641-7595)
van den Broeke Michiel (Orcid ID: 0000-0003-4662-7565)
Brunke Michael A. (Orcid ID: 0000-0002-2677-7446)
Buzan Jonathan (Orcid ID: 0000-0002-8199-3264)
Clark Martyn (Orcid ID: 0000-0002-2186-2625)
Drewniak Beth (Orcid ID: 0000-0002-8193-6091)
Fisher Joshua B. (Orcid ID: 0000-0003-4734-9085)
Flanner Mark (Orcid ID: 0000-0003-4012-174X)
Fox Andrew M. (Orcid ID: 0000-0002-4027-4538)
Gentine Pierre (Orcid ID: 0000-0002-0845-8345)
Hoffman Forrest (Orcid ID: 0000-0001-5802-4134)
Keppel-Aleks Gretchen (Orcid ID: 0000-0003-2213-0044)
Kumar Sanjiv (Orcid ID: 0000-0002-0472-6074)
Lenaerts Jan (Orcid ID: 0000-0003-4309-4011)
Leung L. Ruby (Orcid ID: 0000-0002-3221-9467)

Abstract

The Community Land Model (CLM) is the land component of the Community Earth System Model (CESM) and is used in several global and regional modeling systems. In this paper, we introduce model developments included in CLM version 5 (CLM5), which is the default land component for CESM2. We assess an ensemble of simulations, including prescribed and prognostic vegetation state, multiple forcing datasets, and CLM4, CLM4.5, and CLM5, against a range of metrics including from the International Land Model Benchmarking (ILAMBv2) package. CLM5 includes new and updated processes and parameterizations: (1) dynamic land units, (2) updated parameterizations and structure for hydrology and snow (spatially explicit soil depth, dry surface layer, revised groundwater scheme, revised canopy interception and canopy snow processes, updated fresh snow density, simple firn model, and Model for Scale Adaptive River Transport), (3) plant hydraulics and hydraulic redistribution, (4) revised nitrogen

cycling (flexible leaf stoichiometry, leaf N optimization for photosynthesis, carbon costs for plant nitrogen uptake), (5) global crop model with six crop types and time-evolving irrigated areas and fertilization rates, (6) updated urban building energy, (7) carbon isotopes, and (8) updated stomatal physiology. New optional features include demographically structured dynamic vegetation model (Functionally Assembled Terrestrial Ecosystem Simulator, FATES), ozone damage to plants, and fire trace gas emissions coupling to the atmosphere. Conclusive establishment of improvement or degradation of individual variables or metrics is challenged by forcing uncertainty, parametric uncertainty, and model structural complexity, but the multi-variate metrics presented here suggest a general broad improvement from CLM4 to CLM5.

Plain Language Summary

The Community Land Model (CLM) is the land component of the widely-used Community Earth System Model (CESM). Here, we introduce model developments included in CLM version 5 (CLM5), the default land component for CESM2 which will be used for the Coupled Model Intercomparison Project (CMIP6). CLM5 includes many new and updated processes including: (1) hydrology and snow features such as spatially explicit soil depth, canopy snow processes, a simple firn model, and a more mechanistic river model, (2) plant hydraulics and hydraulic redistribution, (3) revised nitrogen cycling with flexible leaf stoichiometry, leaf N optimization for photosynthesis, and carbon costs for plant nitrogen uptake, (4) expansion to six crop types (global) and time-evolving irrigated areas and fertilization rates, (5) improved urban building energy model, and (6) carbon isotopes. New optional features include a demographically-structured dynamic vegetation model, ozone damage to plants, and fire trace gas emissions coupling to the atmosphere. Model performance is generally improved for most assessed variables and metrics, though clear establishment of improvement or degradation is challenging due to model complexity as well as observational data limitations. Nonetheless, CLM5 is increasingly suited for research into a broad range of societally-relevant scientific questions related to the terrestrial system.

1. Introduction

Land models are classically used as tools to integrate terrestrial contributions and responses to weather, climate variability, and climate change. In addition, modern land models are increasingly expected to provide insight into weather and climate impacts of societally relevant quantities such as water availability, crop and timber yields, wildfire risk, human heat stress, and other ecosystem services (Bonan & Doney, 2018). The Community Land Model (CLM), which is the land component of the Community Earth System Model (CESM), has been developed and expanded over the last decade to provide an increasingly comprehensive platform that researchers can use to address these types of questions. More explicitly, CLM has been developed in accordance with two central themes: 1) terrestrial ecosystems, through their cycling of energy, water, momentum, carbon, nitrogen, and other trace gases, are important determinants of weather and climate, and 2) the land is a critical interface through which climate variability and climate change influence humans and ecosystems *and* through which humans and ecosystems can affect global environmental change.

Here, we introduce the Community Land Model version 5 (CLM5, http://www.cesm.ucar.edu/models/cesm2.0/land/), which builds on progress made in CLM4 (Lawrence et al., 2011) and CLM4.5 (Oleson et al., 2013). CLM is a community-developed model with CLM5 representing the outcome of model development and analysis efforts by a diverse group of scientists and software engineers from many institutions. Priorities for model development are set collectively by the CLM research and development community and are broadly focused on the enhancement of the capacity of the model to be applied to emerging questions that lie at the intersection of weather and climate with terrestrial processes. Examples of scientific topics that have driven CLM5 development include the following:

• Improved understanding of carbon and nitrogen cycle interactions and their influence on the long-term trajectory of the terrestrial carbon sink;

• Assessment of the response and vulnerability of ecosystems to climate change and disturbances (human and natural) and the possibility for ecosystem management to mitigate climate change;

• Quantification of the role of terrestrial processes in diurnal to interannual weather and climate variability including influence on droughts, floods, and extremes;

• Establishment of the availability of water resources under climate variability and climate change;

• Quantification of key land feedbacks to climate change including the permafrost climate-carbon feedback, and snow- and vegetation-albedo feedbacks;

• Representation and quantification of impacts of anthropogenic land-cover and land-use change on climate and the carbon cycle;

• Assessment of how land surface heterogeneity affects land-atmosphere interactions and carbon cycling;

• Examination of the impact of model structural and parameter uncertainty and exploration of parameter optimization techniques

The overarching development philosophy also rests on the notion that terrestrial systems are highly coupled and that development in one set of model processes can modify, and often improve, the simulation of other model processes (for example, improvements in the representation of soil hydrology is likely to improve carbon cycle simulations and vice versa) and can also expose problems in other parts of the model. Core biogeophysical and biogeochemical parameterization development has been complemented with expansions to model functionality (e.g., introduction of a global interactive crop model with fertilization and irrigation, introduction of an embedded ice sheet model) and model structural updates (e.g., increased soil vertical resolution and spatially variable soil depth). Many of the improvements adopted for CLM5 were independently developed by separate research groups for a range of reasons and applications. Therefore, a principal goal of this manuscript is to catalog and describe the full set of CLM5 model developments so that model users are aware of the new features of the model, including known strengths and limitations (Section 2). The model simulations and meteorological forcing datasets employed are described in Section 3. We include a high-level assessment of the integrated impact of these developments on the overall performance of the model, utilizing the International Land Model Benchmarking package (ILAMB, Collier et al., 2018), ecosystem experiment data, and other metrics (Section 4). A summary and discussion are provided in Section 5. A full Technical Description of the model is available here: http://www.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf.

2. **Model description**

2.1 CLM4

CLM4 was released in June, 2010 along with the Community Climate System Model version 4 (CCSM4). CLM4 has been used in CCSM4 (Gent et al., 2011) and CESM1 (Hurrell et al., 2013). CLM4 is described in Lawrence et al. (2011) and a full technical description is available in Oleson et al. (2010) (http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM4_Tech_Note.pdf). Briefly, CLM4 included more sophisticated representations of soil hydrology and snow processes than its predecessor, CLM3.5 (Oleson et al., 2008b). In particular, new treatments of soil column-groundwater interactions, soil evaporation, aerodynamic parameters for sparse/dense canopies, vertical burial of vegetation by snow (Wang & Zeng, 2009), snow cover fraction (Niu & Yang, 2007), and aging, black carbon and dust deposition, and vertical distribution of solar energy for snow were implemented (Flanner et al., 2007). CLM4 was the first version in the CLM series to include a prognostic above- and below-ground carbon-nitrogen cycle (CLM4CN, (Thornton et al., 2007) as well as the ability to represent transient land-cover change (Lawrence et al., 2012). CLM4 added a representation of organic soil and deep ground into the existing mineral soil treatment (Lawrence & Slater, 2008; Lawrence et al., 2008) to enable more realistic modeling of permafrost and active layer dynamics. An urban canyon model, to contrast rural and urban energy balance and climate, was also introduced (Oleson et al., 2008a).

2.2 CLM4.5

CLM4.5 was released in June, 2013 along with the Community Earth System Model version 1.2 (CESM1.2). A full technical description is available here: http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf (Oleson et al., 2013). The main modifications in CLM4.5 included updates to canopy processes including a revised canopy radiation scheme and canopy scaling of leaf processes, co-limitations on photosynthesis (Bonan et al., 2011; Bonan et al., 2012), and temperature acclimation of photosynthesis (Lombardozzi et al., 2015b). Hydrology updates included modifications such that hydraulic properties of frozen soils are determined by liquid water content only rather than total water content, introduction of an ice impedance function and allows for a perched water table above icy permafrost ground (Swenson et al., 2012). The snow cover fraction parameterization was revised to reflect hysteresis in fractional snow cover, for a given snow depth,

between accumulation and melt phases (Swenson & Lawrence, 2012). The lake model was thoroughly revised (Subin et al., 2012). A surface water store was introduced, replacing the wetland land unit. The surface energy flux calculation was modified to separately simulate snow-covered, water-covered, and snow/water-free portions of vegetated and crop land units, and snow-covered and snow-free portions of glacier land units (Swenson & Lawrence, 2012). Globally constant river flow velocity was replaced with variable flow velocity based on mean grid cell slope. A vertically resolved soil biogeochemistry scheme was introduced with base decomposition rates varying with depth and modified by soil temperature, water, and oxygen limitation and also including vertical mixing of soil carbon and nitrogen due to bioturbation, cryoturbation, and diffusion (Koven et al., 2013). Litter and soil carbon and nitrogen pool structure as well as nitrification and denitrification were modified to reflect the Century model (Koven et al., 2013). The fire model was replaced with a model that includes representations of natural and anthropogenic ignition sources and suppression as well as agricultural, deforestation, and peat fires (Li et al., 2013a; Li et al., 2012). The biogenic volatile organic compounds model was updated to MEGAN2.1 (Guenther et al., 2012).

Further additions to CLM4.5 included a methane production, oxidation, and emissions model (Riley et al., 2011) and an extension of the crop model to include interactive fertilization, organ pools (Drewniak et al., 2013), and irrigation (Sacks et al., 2009). Multiple urban density classes, rather than the single dominant urban density class used in CLM4, were modeled in the urban land unit. Carbon ^{13}C and ^{14}C isotopes for natural vegetation were introduced (Koven et al., 2013). A summary of the changes included in CLM4.5 relative to CLM4 are listed in Table 1.

2.3 CLM5

CLM5 is the default land model for CESM2 (http://www.cesm.ucar.edu/models/cesm2/). Developments for CLM5 build on the progress made in CLM4.5. Most major components of the model have been updated with notable changes made to soil and plant hydrology, snow density, river modeling, carbon and nitrogen cycling and coupling, crop modeling as well as new surface characterization and transient land use datasets and increased flexibility to represent landscape dynamics through specified or prognostic transitions in land unit weights. Much of the development reflects a push towards more mechanistic

treatment of key hydrologic and ecological processes and more comprehensive and explicit representation of anthropogenic land management.

Prior versions of CLM mainly included a single option for most parameterizations. With our new CLM codebase management philosophy, where new parameterizations or model structural decisions were defined for CLM5, we also maintained the CLM4.5 parameterization or configuration, thereby allowing users to switch back and forth between alternative parameterizations via namelist control. In this section, we briefly describe the full set of model developments. Except where explicitly noted, all described new parameterizations or features are active by default in CLM5. For full details of new and old CLM5 parameterizations, including equations and parameter values, we refer the reader to the cited papers and to the full technical description of CLM5 (http://www.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf). Additional documentation including information about how to access the code, tutorials about how to run the model, developer's guides, and model output diagnostics can be found here: http://www.cesm.ucar.edu/models/cesm2/land/. A schematic representation of the primary processes and functionality represented in CLM5 is shown in Figure 1. A summary of the changes in CLM5 relative to CLM4.5 are listed in Table 1 for reference.

2.3.1 Dynamic land unit weights and Plant Functional Type distribution

CLM5 includes a new capacity update land unit weights during a simulation either through a dataset or prognostically, a technical feature that was previously not possible which prevented representation of important specified or dynamic transitions. Spatial land surface heterogeneity in CLM is represented as a nested sub grid hierarchy in which grid cells are composed of multiple land units, columns, and patches (Figure 2). Each grid cell can have a different number of land units, each land unit can have a different number of columns, and each column can have multiple patches each with a specific plant functional type (PFT) or crop functional type (CFT). The first sub grid level, the land unit, is intended to capture the broadest spatial patterns of sub grid heterogeneity. The CLM5 land units are vegetated, lake, urban, glacier, and crop. New within CLM5 is the capacity to adjust the fractional area of each land unit throughout the course of a simulation either as specified through a land use dataset (e.g., deforestation for agriculture, transition of a fraction of vegetated land unit to crop land unit) or through prognosed

initiation or loss of glacier area (e.g., initiation of glacier area, transition of fraction of vegetated land unit to glacier land unit; only possible when two-way ice sheet interactions are activated). For natural vegetation, CLM operates under the assumption that all PFTs on the natural vegetation land unit compete for water and nitrogen and that all PFTs share the same soil column state (temperature, moisture, carbon, nitrogen). Note that prior research has shown that for some applications, particularly for studies of land cover change impacts on climate, it may be preferable for each PFT to operate on its own soil column to avoid implicit energy transfer from one PFT to another (Chen et al., 2018; Meier et al., 2018; Schultz et al., 2016). On the crop land unit, each CFT (irrigated and unirrigated) resides on its own soil column and therefore operates based on its own soil moisture and nitrogen conditions. For CLM5, transitions have only been enabled between natural vegetation, crop, and glacier land units; full transition capability (i.e., including ability for transitions to and from urban and lake land units) is under development. The default configuration of CLM5 (and the other CLM versions assessed here) does not include dynamic vegetation biogeography (i.e., CLM is not a Dynamic Global Vegetation Model). Instead vegetation distributions (natural and cropland PFTs and CFTs) are specified through time through a land use timeseries file (see below), but vegetation state (LAI, canopy height) is prognostic. Urban and lake land unit areas do not and cannot change during a simulation. Total grid cell water, energy, carbon and nitrogen are conserved for all transitions.

New land cover and land use data have been generated for CLM5 that combine updated versions of current day satellite land cover descriptions with the Land Use Model Intercomparison Project (LUMIP, (Lawrence et al., 2016)) past and future transient land use time series (Land Use Harmonization 2, LUH2, luh.umd.edu/data.shtml). The new CLM5 land surface input and timeseries data describe the distribution of plant functional types (PFTs) and crop functional types (CFTs), soil texture and, wood harvest, industrial N fertilizer application amounts, irrigation-equipped area, shifting cultivation (repeated clearing and abandonment of agricultural land), monthly PFT leaf area index (LAI) and canopy height (for simulations using prescribed satellite vegetation phenology - SP in CLM infrastructure nomenclature), and several additional required or optional input datasets (see Technical Description for more detail on input datasets).

The CLM5 surface datasets are created as in CLM4 and CLM4.5, but with updated methodology as described here. Present-day global land cover descriptions are generated at 1km resolution using updated

versions of the data and methods used for CLM4 and CLM4.5 (Lawrence & Chase, 2007). The basis for the land cover description comes from MODIS land cover (MCD12Q1 v5.1), vegetation continuous fields (MOD44B v5.1), leaf area index (MCD15A2 v5), and albedo (MCD43B3 v5) products for the years 2001 – 2015 (https://lpdaac.usgs.gov/dataset_discovery/modis). Additional information for tree leaf type and longevity are provided by the AVHRR continuous fields tree cover product (Defries et al., 2000). Global crop distributions are provided by the monthly irrigated and rainfed crop areas around the year 2000 (MIRCA2000) data set of Portmann et al. (2010). Canopy height data for tree PFTs are provided by the Geoscience Laser Altimeter System (GLAS) on the ICESat satellite as processed by (Simard et al., 2011). The LUH2 historical and future scenario data provide annual land use descriptions at 0.25^{o} resolution from the year 850 to 2100. The LUH2 data are generated from the History Database of the Global Environment (HYDE version 3.2, (Goldewijk et al., 2017)) for 850-2014, and from Integrated Assessment Model teams for multiple alternative scenarios of the future for 2015-2100. The LUH2 time series describes annual changes in primary and secondary forest and non-forest land units, along with five crop groups, managed pasture, rangeland and urban areas. The LUH2 data also includes information on wood harvest, both in terms of the mass of carbon extracted and the total harvest areal fraction (CLM5 uses carbon mass). Annual crop management is specified by crop type through industrial fertilizer application and the fraction of each crop irrigated. Finally, the CLM surface datasets and transient land use datasets are produced with the CLM Land Use Data Tool (http://www.cgd.ucar.edu/iam/projects/thesis/thesis-clm-landuse-tool.html). This tool takes the present-day land cover distribution and merges it with historical or future LUH2 transitions and management information and translates them into CLM PFT and CFT distributions and management information.

2.3.2 Soil hydrology

CLM5 includes several structural and parameterization improvements that increase the realism of the soil hydrology representation in the model. To resolve a deficiency in the seasonality of soil evaporation and soil water storage in semi-arid regions, the empirical soil evaporation resistance parameterization is replaced with a mechanistically-based parameterization where soil evaporation is controlled by the rate of diffusion of water vapor through a dry surface layer (Swenson & Lawrence, 2014). To account for spatial variation in soil thickness and columnar water holding capacity, CLM is updated so that different soil

thicknesses (by default ranging from 0.4 to 8.5m depth) can be applied for each soil column (Brunke et al., 2016; Swenson & Lawrence, 2015). The default spatially-explicit soil depths are derived from a spatially explicit soil thickness data product (Pelletier et al., 2016). The explicit treatment of soil thickness with underlying bedrock (currently assumed to be impermeable, i.e. zero flux bottom boundary condition) means that the soil saturated and unsaturated zones and associated water table depth are modeled explicitly. This allows for the deprecation of the unconfined aquifer parameterization ((Niu et al., 2007), which was used as part of the groundwater representation in CLM4 and CLM4.5. Note that an added benefit of the explicit representation of spatially-varying soil thickness underlain by impermeable bedrock is that it removes a logical inconsistency between the treatment of soil hydrologic and soil thermal calculations that existed in CLM4 and CLM4.5. The default model soil layer resolution is increased, especially within the top 3m, in part to more accurately simulate active layer thickness within the permafrost zone. The default configuration includes a total of 25 ground layers that extend to a depth of ~50m. The first 5 (0.4m thick soils) up to 20 (8.5m thick soils) layers in each column are considered soil and are hydrologically and biogeochemically active. The number of soil layers is specified independently for each column based on the imposed soil thickness dataset. The remaining ground layers in each column are considered to be dry bedrock. Note that since the number of active soil layers varies from grid cell to grid cell, users need to be careful when doing spatial averaging of soil moisture or carbon/nitrogen quantities since bedrock layers have very small prescribed constant soil moisture and carbon/nitrogen values.

An adaptive time-stepping solution to the Richard's equation is introduced (Clark & Kavetski, 2010; Kavetski et al., 2001). This improves the accuracy and stability of the numerical soil water solution by allowing for multiple substeps within the standard 30-minute model timestep when required. In test simulations, all instances of numerical instability in the Richards equation solution (i.e. negative soil moisture updates) were eliminated at a cost of an increase of less than 3% in model runtime. Sub-stepping is invoked (i.e., instabilities occur in Richard's equation solution) most frequently when and where the number of soil layers is small, which can be due to frozen soils or shallow bedrock. The process of subtracting the hydrostatic equilibrium soil moisture distribution from the vertical soil moisture profile before solving Richards' equation, proposed in Zeng and Decker (2009) and included in CLM4

and CLM4.5, has been deprecated because it is inconsistent with standard approaches used in soil hydrology (De Rooij, 2010).

2.3.3 Atmospheric surface layer stability

In Monin-Obukhov stability theory (Foken, 2006) atmospheric stability is characterized by a length scale *L*, called the Obukhov length, which is used to nondimensionalize the distance to the surface using variable *zeta = (z-d)/L*, where z is the reference height and *d* is the displacement height. In CLM4.5, the stability variable zeta is constrained to be less than or equal to 2. Using temperature and friction velocity measurements from a subalpine forest flux tower, Burns et al. (2018) showed that CLM4.5 exhibited a large and persistent nighttime low bias of canopy temperature and friction velocity. In that study, they alleviated this bias by implementing the Handorf et al. (1999) stability function in very stable conditions. For CLM5, we approximate the Handorf et al. (1999) stability function for very stable conditions by setting the maximum zeta value to 0.5. Ongoing development work since CLM5 was finalized indicates that this need for a maximum zeta value can be eliminated when a vegetation biomass heat storage capacity is explicitly modeled (Swenson et al., 2019). Stability corrections and the applicability of Monin-Obukhov similarity remains active research topics, which has recently leveraged high-resolution turbulent simulations such as Direct Numerical Simulations (DNS) (Li et al., 2018).

CLM4.5 includes an additional modification to under-canopy stability designed to increase aerodynamic resistance between the canopy and the ground in stable conditions (Sakaguchi & Zeng, 2009). Due to biases in surface to lowest atmosphere layer temperature profiles, also noted by Burns et al. (2018), it was found that the under-canopy stability parameterization did not perform as intended. Consequently, this under-canopy stability parameterization is inactive in CLM5. Within- and under- canopy stability remains an active area of research (e.g., Bonan et al., 2018).

2.3.4 Snow, glaciers, and ice sheets

Several changes are included that are mainly targeted at improving the simulation of surface mass balance, the difference between annual accumulation and ablation, over ice sheets. New parameterizations

for fresh snow density (updated temperature effects and wind effects), destructive metamorphism (the change in snow crystals from six-sided shapes to rounded, bonded ice grains due to disturbance, molecular motion, and pressure), and compaction by overburden pressure and drifting snow are included (van Kampenhout et al., 2017). For reference, fresh snow density as a function of temperature and wind speed is shown in Figure 1 of van van Kampenhout et al. (2017). The maximum number of snow layers and snow amount is increased from 5 layers and 1 m snow water equivalent to 12 layers and 10 m snow water equivalent, to allow for the formation of firn in regions of persistent snow-cover (e.g., glaciers and ice sheets) (van Kampenhout et al., 2017). The snow capping routine, which sets a limit on the maximum amount of accumulated snow, has been fixed to correctly allow surface snow density and grain size to refresh when new snow falls. The grain size of freshly fallen snow has been made a function of air temperature to address unrealistically high albedos over ice sheets. Instead of applying a fresh snow grain size of 54 μm at all temperatures, fresh snow grain size is set to 54 μm below -30^{o}C and to 204.5 μm above 0^{o} C, with a linear ramp applied between these temperatures. The parameters for snow grain ageing are maintained.

Multiple elevation classes (10 elevation classes by default) are specified on the glacier land unit to account for the strong topographic elevation gradients present over many glaciers and ice sheets (Lipscomb et al., 2013). Atmospheric surface temperature, potential temperature, specific humidity, density, and pressure are downscaled from the mean grid cell elevation to each glacier column elevation using a specified lapse rate (6.0^{o} km^{-1}) and an assumption of uniform relative humidity. Longwave radiation is downscaled by assuming a linear decrease in downwelling longwave radiation with increasing elevation (0.032 W m^{-2} m^{-1}, bounded to 0.5 to 1.5 times the grid cell mean value, then normalized to conserve grid cell total energy) (Tricht et al., 2016). This downscaling allows lower-elevation columns within a glacier land unit to undergo surface melting while columns at higher elevations remain frozen.

In typical configurations (e.g. by default in CESM2 and CLM5 land-only simulations), CLM5 computes ice sheet surface mass balance, but ice sheets do not evolve. CLM5 can also be coupled bidirectionally to CISM2.1 (Lipscomb et al., 2019) and thereby simulate an evolving Greenland ice sheet. The introduction of the capability to adjust land unit weights during a simulation (Section 2.3.1) means that a glacier can incept, grow, shrink, or disappear during a simulation when two-way coupling between the land and ice

sheet model is active. By default, two-way coupling is not active in CESM2 or CLM5 land-only simulations, including the simulations assessed here.

Vegetation canopy precipitation interception is updated to track liquid and solid water phases separately, with intercepted snow subject to unloading events due to wind or above-freezing temperatures similar to Roesch et al. (2001). Interception snow mass compares favorably with in-situ measurements from Storck et al. (2002). Additionally, the snow-covered fraction of the canopy, which is calculated based on the canopy snow mass and LAI, is used within the canopy radiation and surface albedo calculations.

Finally, CLM5 partitions total precipitation into rain and snow according to a linear temperature ramp. This partitioning occurs irrespective of what phase precipitation is calculated by the atmosphere model. For most land units, this ramp generates all snow below 0^{o}C, all rain above 2^{o}C, and a mix of rain and snow for intermediate temperatures. For glaciers, the end points of the ramp are -2^{o}C and 0^{o}C, respectively. To ensure energy conservation, a sensible heat flux correction term is applied when the phase of precipitation coming from the atmosphere is changed.

2.3.5 Rivers

The River Transport Model (RTM) used in CLM4.5 is replaced with the physically more realistic Model for Scale Adaptive River Transport (MOSART, Li et al., 2013b). Note that the river model is treated as a separate coupled component in CESM and therefore is not technically part of CLM, but we include it in this manuscript because of the clear relationship with and dependence on CLM, i.e., MOSART receives surface and subsurface runoff from CLM. MOSART represents an upgrade over RTM in several ways. RTM utilizes a simple linear reservoir method to calculate streamflow, whilst MOSART is based on the more physically-based kinematic wave method. MOSART also provides more information on river conditions; i.e., RTM only simulates streamflow whereas MOSART additionally simulates time-varying channel velocities, channel water depth, and channel surface water variations. In MOSART, surface runoff is routed across hillslopes and then discharged along with subsurface runoff into a tributary subnetwork before entering the main channel. MOSART assumes that all the tributaries within a spatial unit (either regular lat/lon grid or watershed) can be treated as a single hypothetical sub-network channel with a transport capacity equivalent to all the tributaries combined. Correspondingly, three routing

processes are represented in MOSART: 1) hillslope routing: surface runoff is routed as overland flow into the sub-network channel, while subsurface runoff directly enters the sub-network channel; 2) sub-network channel routing: the sub-network channel receives water from the hillslopes, routes water through the channel and discharges it into the main channel; 3) main channel routing: the main channel receives water from the sub-network channel and/or inflow, if any, from upstream, and discharges the water downstream or to the ocean. The capability to simulate flooding (water transfer from rivers back onto land under flood stages) that was implemented into RTM for CLM4.5 is retained for MOSART, but is not active by default. The representation of wetlands is unchanged from CLM4.5 wherein wetlands are no longer their own prescribed land unit, but instead are captured through a prognostic surface water storage that accounts for fine spatial scale variations in surface elevation (see Technical Description for details).

2.3.6 Vegetation physiology

A plant hydraulic stress (PHS) routine is introduced which explicitly models water transport through the vegetation according to a simple hydraulic framework (Kennedy et al., 2019). The plant hydraulics routine solves for vegetation water potentials (root, xylem and leaf) according to an electric circuit analogy, in which the flow (current) is the soil-to-leaf water supply (sap) which is set to meet the transpiration flux (demand) at every time step, i.e., no storage is assumed. Explicit prognosis of plant tissue water status improves the physical basis for many processes represented in CLM, such as the dynamics of root water uptake profiles, and the attenuation of photosynthesis and transpiration with drought, which was exaggerated in previous model versions (e.g., Powell et al., 2013). In PHS, 'unstressed' (atmospheric demand-driven) stomatal conductance is modulated for drought stress using a function of leaf water potential, requiring vegetation to regulate stomatal conductance to avoid excessively negative leaf water potential and thus plant desiccation and embolism in the xylem. This more mechanistic representation of vegetation water stress replaces the soil moisture stress (SMS) parameterization in prior versions of CLM in which water stress was calculated through a plant wilting factor that was based on soil water matric potential relative to PFT-dependent parameters for fully closed and fully open stomata, weighted by layer root fractions. An emergent feature of the plant hydraulics scheme (wherein water moves along water potential gradients within the soil-root-stem-leaf system) is a plant-mediated vertical hydraulic redistribution of soil water from wet to dry soil layers, which thus leads

to important nighttime and seasonal hydraulic redistribution, physically constrained by the plant hydraulic parameterization (Kennedy et al., 2019). To prevent unrealistically high soil evaporative losses of soil water due to continuous hydraulic redistribution, root water uptake and hydraulic redistribution is not allowed to occur in the 2 cm thick surface soil layer.

PHS advances the physical and empirical basis of the CLM vegetation hydrodynamics scheme. Previously used soil moisture stress functions (as in SMS) tend to lack either a strong physical or empirical justification and are a major source of uncertainty in land models (Trugman et al., 2018). PHS, in adopting a plant hydraulic framework, incorporates more physical root water uptake, following Darcy's Law, and a stress formulation based on avoiding excessive xylem tension. Likewise, PHS opens avenues for better empirical constraints on vegetation water use. The model parameters have physical meaning, and new prognostic vegetation water potential can be validated with field observations and, potentially, satellite remote sensing products (e.g., Anderegg et al., 2018; Konings et al., 2017; Li et al., 2017).

In CLM5, maximum stomatal conductance is obtained from the Medlyn "empirical-optimal" conductance model (Medlyn et al., 2011), rather than the Ball-Berry stomatal conductance model that was utilized in CLM4.5 and prior versions of the model. The Ball-Berry implementation used a single slope parameter for all C3 plants. In a recent study, Lin et al. (2015) estimated PFT-dependent slope parameters for the Medlyn model, which have been successfully used in CABLE (De Kauwe et al., 2015). The slope parameters used in CLM5 are from CABLE. Note that the slope parameter value is indicative of the plant's water use strategy – PFTs with a high slope parameter have high stomatal conductance per unit photosynthesis, and therefore a low water use efficiency. As discussed by Franks et al. (2017) and Franks et al. (2018), the primary difference between the two stomatal models, after accounting for different slope parameters, relates to the effects of extreme low and high vapor pressure deficit on stomatal conductance.

Two other relatively minor changes are included in CLM5. (1) The trigger for stress deciduous PFT phenology is augmented with an antecedent precipitation requirement (Dahlin et al., 2015). This additional trigger was implemented to reduce the occurrence of anomalous green-up during the dry season in many semi-arid regions that was being driven by upwards water movement from wet to dry soil layers and thereby triggering unrealistic leaf-out even in circumstances when there was not any recent rainfall. More recent work has demonstrated a broad array of stress deciduous phenology strategies that are not possible to resolve in the current CLM PFT scheme (Adole et al., 2018; Dahlin et al., 2017), but this

complexity could potentially be represented in FATES (see Section 2.3.12). (2) The rooting profiles, which were inconsistent for water and carbon in CLM4.5 were updated to be consistent in CLM5. The Jackson et al. (1996) rooting profile is preferred over the Zeng (2001) profile as it produces more realistic vertical soil C profiles, though the Zeng (2001) profile is retained as an option.

Lastly, ozone damage to vegetation is included as an optional feature in CLM5. The ozone damage parameterization is the same as implemented by Lombardozzi et al. (2015a) based on ozone damage response data compiled by Lombardozzi et al. (2013). Ozone damage to vegetation is applied directly and independently to photosynthesis and stomatal conductance for three broad PFT classes (broadleaf trees and shrubs, needleleaf trees and shrubs, and crops and grasses) based on the cumulative uptake of ozone. Cumulative uptake of ozone is calculated as the ozone concentration multiplied by stomatal conductance, integrated through time, to account for the fact that ozone primarily damages vegetation once it enters the leaf and total damage is dependent on the time period of exposure. The damage decays over the growing season to account for the fact that plants acquire new, undamaged leaves throughout the growing season, and also decays over the leaf lifespan for evergreen plant types.

2.3.7 Carbon dynamics

CLM5 applies a fixed C allocation scheme for woody vegetation where allocation to above and below ground biomass is held constant. The decision not to use the dynamic allocation scheme based on NPP, as was used in CLM4 and CLM4.5, was driven by the fact that observations indicate that plant biomass saturates with increasing productivity, which is inconsistent with the behavior in CLM4 and CLM4.5 where biomass perpetually increases with increasing productivity (Negrón-Juárez et al., 2015). Because the prior allocation rules implicitly led to a saturation of leaf carbon allocation, this change does lead to a possible tradeoff between accuracy of biomass and accuracy of leaf area, and remains a large uncertainty and an area of active research. Soil carbon decomposition processes are unchanged from CLM4.5 to CLM5, but assessment with a new metric for the temperature sensitivity of apparent soil carbon turnover times (Koven et al., 2017) pointed to the need to adjust the parameter that controls intrinsic depth limitation on soil carbon turnover towards a weaker depth limitation (rather than the strong depth limitation in CLM4.5) and to adjust the parameter that controls soil moisture limitation on soil carbon

turnover rates in dry soils to a wetter soil moisture level than that used in CLM4.5. Note that vertical C and N processes are only calculated for hydrologically active soil layers (see Section 2.3.2), which vary in space.

2.3.8 Nutrient dynamics

Plant nutrient dynamics are substantially updated in CLM5 to resolve several deficiencies with the representation of nutrient cycling in previous versions of the model. The Fixation and Uptake of Nitrogen (FUN) model, based on the work of Fisher et al. (2010), Brzostek et al. (2014), and Shi et al. (2016), is incorporated. The concept of FUN assumes that N uptake requires the expenditure of energy in the form of C (in CLM4.5 there was no C expenditure for N uptake), often a significantly large portion of NPP (Doughty et al., 2018; Marschner, 1995) and further, that there are numerous potential sources of N in the environment which a plant may exchange for C: symbiotic biological N fixation, arbuscular-mycorrhizal and ecto-mycorrhizal (two types of root fungus) uptake, direct root uptake, and leaf N retranslocation. The ratio of C expended to N acquired is therefore the C cost, or exchange rate, of N acquisition. This C is assumed to respire as it is used for N acquisition. As FUN calculates the rate of symbiotic N fixation, this N is passed straight to the plant, as opposed to passing through the soil mineral N pool. CLM5 now separately calculates rates of free-living N fixation as a function of evapotranspiration (modified from Cleveland et al., 1999), which is added to the soil inorganic ammonium (NH_4^+) pool. Previous versions of CLM added the N fixation flux, which was calculated as function of net primary productivity (without an associated C cost; Cleveland et al., 1999; Thornton et al., 2007; Wieder et al., 2015), to the soil mineral N pool.

The static plant carbon:nitrogen (C:N) ratios utilized in CLM4 and CLM4.5 are replaced with variable plant C:N ratios, as in Zaehle and Friend (2010), which allows plants to adjust their C:N ratio, and therefore their leaf N content, with the cost of N uptake (Ghimire et al., 2016). The implementation of a flexible C:N ratio means that the model no longer relies on instantaneous down-regulation of potential photosynthesis rates based on soil mineral N availability to represent nutrient limitation. Furthermore, stomatal conductance in CLM5 is based on the N-limited photosynthesis rate rather than on potential N-

unlimited photosynthesis as in CLM4 and CLM4.5, thereby allowing for more realistic coupling between plant C and water cycles (Medlyn et al., 2016).

Finally, the Leaf Use of Nitrogen for Assimilation (LUNA; Ali et al., 2016; Xu et al., 2012) model is incorporated. The model allocates N to maximize daily net photosynthetic carbon gain under the following two key assumptions: (1) N allocated for light capture, electron transport and carboxylation are co-limiting; and (2) respiratory nitrogen is allocated to maintain dark respiration determined by V_{cmax25}. Compared to traditional photosynthetic capacity models, a key advantage of LUNA is that it is able to predict potential acclimation of photosynthetic capacities for different environmental conditions as determined by temperature, radiation, CO_2 concentrations, day length, and humidity. Importantly, the inclusion of LUNA means that V_{cmax25}, the maximum rate of carboxylation, is a prognostic model quantity, dependent on leaf N per unit area and environmental conditions, whereas it was fixed for each PFT in CLM4 and CLM4.5.

2.3.9 Land management processes

Representation of human management of the land (agriculture and wood harvest) is augmented in several ways. Critically, the introduction of the capability to dynamically adjust land unit weights during a simulation means that the crop model can be run coincidentally with prescribed land-use change, which significantly expands the capabilities of the model. The CLM4.5 crop model is extended to operate globally through the addition of rice and sugarcane as well as tropical varieties of corn and soybean (Badger & Dirmeyer, 2015; Levis et al., 2018). These crop types are added to the existing temperate corn, temperature soybean, spring wheat, and cotton crop types. Industrial N fertilization amounts and irrigation-equipped area are updated annually based on crop type and geographic region through the land use time series dataset. The irrigation trigger is updated to remove the dependence on the CLM4.5 plant soil moisture stress calculation (replaced in CLM5 with PHS, Section 2.3.6) and instead uses a target soil moisture level, which was tuned to get reasonable irrigation amounts. Additional minor changes to crop model include: (1) crop phenological triggers vary by latitude for selected crop types, which is a temporary solution that generates more realistic global crop planting dates outside of the temperate regions for which the growing degree day-based crop planting window was originally parameterized

(though serious crop planting window errors still occur), and (2) grain C and N is transferred during crop harvest into a 1-year product pool with the C needed to seed the next season's crops removed from grain C while the rest of the crop vegetation residue is transferred to litter C and N pools. To better match wood harvest inventories specified in the LUH2 dataset, mass-based, rather than area-based, wood harvest is applied. Shifting cultivation is represented by calculating unrepresented gross transitions in the LUH2 time series and then removing above-ground C to account for the conversion of the gross forest PFTs to crop or pasture PFTs not included by the net transitions. Shifting cultivation is an optional feature of CLM5 and is off by default and in all simulations considered in this paper.

Changes to urban modeling capabilities include the introduction of several human heat stress indices for both urban and rural areas that are calculated and output by default (Buzan et al., 2015). A more sophisticated and realistic building space heating and air conditioning submodel that prognoses interior building air temperature and includes more realistic space heating and air conditioning waste heat factors is incorporated (Oleson & Feddema, 2019).

2.3.10 Fire

The fire parameterization in CLM5 simulates four types of fire: agricultural fires in cropland, deforestation fires in tropical closed forests, peat fires, and non-peat fires outside cropland and tropical closed forests (see Li & Lawrence, 2017 for details; Li et al., 2013a; Li et al., 2012). Burned area is affected by climate and weather conditions, vegetation composition and structure, and human activity. Once burned area is determined, the impact of the fire is calculated, including biomass and peat C losses, fire-induced vegetation mortality, adjustment of the vegetation C:N pools, and fire C and other trace gas emissions. The fire model is mainly unchanged from CLM4.5 except with a modified scheme for the dependence of fire occurrence and spread on fuel wetness for non-peat fires outside cropland and tropical closed forests and with the dependence of agricultural fires on fuel load removed.

The CLM5 fire model, when coupled to the Community Atmosphere Model, can simulate and transfer emissions of total C, aerosols (e.g., black C, organic C, DMS), greenhouse gases (e.g., CO_2, N_2O, CH_4) and other trace gases (e.g., CO, NO, NO_2, NH_3, HONO, SO_2 and over 15 non-methane hydrocarbon species) to the atmosphere (Ford et al., 2018). Fire emissions are estimated at the PFT level from total

fire C emissions, a conversion factor from C to dry matter (DM) (0.5 g C/g DM) and emission factors (g species/g DM) that convert dry matter burned into emissions. The emission factors for each species used in CLM5 are derived from up-to-date inventories compiled from field and laboratories studies ((Andreae & Merlet, 2001, updated to 2016) (Akagi et al., 2011, updated to 2014) and references therein). The vertical distribution of fire emissions is derived from PFT-dependent maximum injection heights (4.3 km for needleleaf trees, 3 km for other boreal and temperate trees, 2.5 km for tropical trees, 2 km for shrublands, and 1 km for grass and croplands). These injection heights are compiled from satellite-based observations of fire smoke plumes (Val Martin et al., 2018; Val Martin et al., 2010). The fire emissions module is not active by default in CESM2, but is available as a research option.

2.3.11 Parameters

Parameters of CLM5 were defined where possible from literature values and meta analyses, with some adjustments made to reduce large model biases, while accounting for errors in observational datasets and in the globally applied model structure. Default parameter values for all model parameters can be found in the CLM5 Technical Description. A brief description of the rationale for the values used for selected parameters is included here. Note that during the process of finalizing the CLM5 parameter set, we found several instances where parameter value tradeoffs needed to be made related to joint goals of relatively small biases for quantities such as GPP and LAI and reasonably high PFT survivability rates (see Section 4.2). Fisher et al. (2019) provide a more detailed assessment of CLM5 C and N cycle sensitivity to parametric uncertainty as well as additional discussion of parameter definition for CLM5. Note that ILAMB was not used during the parameter adjustment process.

Plant Hydraulics Parameters

The plant hydraulics scheme introduces four new parameters for each PFT (Kennedy et al., 2019), including the water potential at which half of the hydraulic conductivity of each plant element (root, stem, shaded leaf, sunlit leaf) is lost (*p50*), the conductivity of the soil-root interface (*krmax*), the conductivities at the interfaces between each of the plant elements (*kmax*), and the cavitation vulnerability curve shape-fitting parameter (*ck*). The code is structured so that in future investigations, parameter values for each plant element can be adjusted individually, but in the released version all plant elements use the same

value. Estimates of *p50* across PFTs are obtained from analysis of the dataset presented by Choat et al. (2010). Large datasets on comparable plant tissue conductivities (*kmax* and *krmax*) are not widely available. Further, because the resistances of the plant and roots act in series, the minimum conductivity among the plant elements largely controls the overall plant conductivity. Plant conductivities are therefore calibrated as follows: *kmax* values are set uniformly high and *krmax* is considered a free tuning parameter. The introduction of PHS represents the first instance where a plant hydrodynamic model has been applied globally across all biomes in CLM, or indeed, in any ESM of which we are aware. Consequently, the plant hydraulics parameter values included in the released CLM5, which were defined in a generally ad-hoc manner, should be considered an initial estimate of reasonable parameter valuables, that can and should be refined as required.

Vegetation Parameters

Several vegetation parameters were updated relative to those used in CLM4 or CLM4.5. (1) PFT-specific values for the slope of the Medlyn stomatal conductance (*medlynslope*) were adapted from Medlyn et al. (2011) as documented in Franks et al. (2017). (2) PFT-specific values of the respiration model intercept (*lmr_intercept_atkin*) were derived from Atkin et al. (2015). (3) Leaf longevity (*leaf_long*), target leaf CN ratio (*leafcn*) and specific leaf area (*slatop)* were all derived from the mean PFT-specific values identified in the TRY database (Kattge et al., 2011). With our final set of default CLM5 parameters, the productivity for boreal and temperate needleleaf evergreen trees is too high, particularly when the LUNA model is active. To calibrate model performance, *leafcn* was increased to one standard deviation above the mean reported value for these PFTs.

The parameters for carbon allocation (ratio of new coarse root to new stem allocation, *croot_stem;* ratio of new fine root to new leaf allocation, *froot_leaf;* ration of new stem to new leaf allocation, *stem_leaf*. The ratios of tissue biomass are the basis for the fixed carbon allocation scheme used in CLM5, which is an over-simplification of real allometric ratios that vary as plants age. Thus, it is difficult to directly connect the parametric allocation ratios used in CLM5 to those obtained from databases. The CLM5 allocation parameters (ratio of new coarse root to new stem, *croot_stem;* ratio of new fine root to new leaf, *froot_leaf;* ratio of new stem to new leaf, *stem_leaf*; and ratio of new live wood to new total wood) were initially derived from an analysis by Ghimire et al. (2016), but were further adjusted to reduce large biases in LAI in deciduous PFTs. CLM4.5 down-regulated leaf allocation with high NPP, whereas CLM5

adopts a fixed allocation scheme to rectify issues with woody biomass accumulation in tropical forests identified by Negrón-Juárez et al. (2015). For CLM5, allocation to stems and roots was increased for many PFTs, potentially compensating for the removal of a variable allocation parameterization, and potentially also contributing to low growth and survival in more marginal climate areas. This set of parametric tradeoffs reflects the need for a whole-plant based (as opposed to big leaf, tissue-based) allocation scheme, as is envisaged for future generations of the model (Fisher et al., 2018; Fisher et al., 2015).

Nitrogen Model Parameters

The introduction of the FUN model to CLM5 adds numerous parameters describing the costs of N acquisition from the environment and control on the flexibility of the tissue C:N ratios. Many of these parameter values are constrained by data, but still include some uncertainty since they represent processes (N uptake, fixation, allocation) that are sparsely documented in the literature. Nitrogen cycle models in general have large structural and parametric uncertainty. The maximum fraction of net carbon assimilation that can be spent (at a PFT level) on fixation is a proxy for the fraction of N fixers (*FUN_fracfixers*) in an ecosystem. *FUN_fracfixers* is set at 0.25 for each PFT, and 0 for all CFTs except temperate and tropical soy where it equals 1. Note that although *FUN_fracfixers* allows fixation, this does not necessarily mean it occurs if there are cheaper C costs for N acquisition from other pathways. Parameters for fixation cost (*a_fix, b_fix, c_fix, s_fix)* were derived from Houlton et al. (2008). The relative values of the six parameters of the active cost of N uptake (*akc_active, akn_active, ekc_active, ekn_active, kc_nonmyc, kn_nonmyc*) were taken from Brzostek et al. (2014). These parameters shape the C cost curves for the mycorrhizal and direct root uptake pathways. Note that N uptake costs of some PFTs were adjusted from Brzostek et al. (2014) values to reduce biases in GPP, especially broadleaf tropical deciduous trees and C4 grass, which Brzostek et al. (2014) did not provide. The parameters that adjust C expenditure on N uptake with changing environmental cost and existing tissue ratios (*fun_cn_flex_a, fun_cn_flex_b* and *fun_cn_flex_c)* were determined via an offline calibration exercise to achieve variations in tissue C:N ratios for the typical modeled N-cost range to be consistent with the range of observations. These parameters allowed FUN, which was originally parameterized for models with fixed plant C:N ratios, to work with the variable plant C:N ratios in CLM5. The fraction of ectomycorrhizal fungi (*per_ecm*) were derived from Shi et al. (2016).

2.3.12 FATES

Included as an option with CLM5 is the Functionally-Assembled Terrestrial Ecosystem Simulator (FATES, Fisher et al., 2015). FATES is a cohort model of vegetation competition and co-existence, allowing a representation of the biosphere which accounts for the division of the vegetated land into successional stages, and for competition for light between height structured cohorts of representative trees of various plant functional types. FATES allows the prediction of biome boundaries directly from plant physiological traits via their competitive interactions and includes the SPITFIRE model of Thonicke et al. (2010), modular allometry and allocation schemes, flexible trait-based PFT definition, interactive logging and plant hydrodynamics based on Christoffersen et al. (2016). FATES fast-timescale physiological processes are based on CLM, but resolved for a height-structured and multi-PFT canopy. FATES is not active by default in CLM5 and is not active within any simulations assessed in this manuscript. Open-source development and application of the codebase is ongoing at https://github.com/NGEET/fates.

2.3.13 Data assimilation capabilities

The capabilities for conducting data assimilation with CLM5 using the Data Assimilation Research Testbed (DART, Anderson et al., 2009) continue to improve, particularly with respect to computational efficiency. The CLM-DART system relies heavily on the CESM multi-instance capability and other workflows. The latest distribution of DART includes full support for CLM5 both in terms of the initial setup scripts provided to create a multi-instance case suitable for DA and the assimilation scripts called by CESM, and for the DART executables themselves. CLM-DART has the ability to assimilate many land observation types using the general DART framework, including in-situ and remote sensing measurements of soil moisture and temperature, eddy covariance flux tower measurements of carbon and water fluxes, and most recently leaf area index and aboveground biomass (Fox et al., 2018). Previous work with CLM-DART has concentrated on hydrometeorology, and describe capabilities to assimilate snow cover fraction (Zhang et al., 2014), AMSR-E brightness temperature for snow depth (Kwon et al., 2016), soil moisture (Zhao et al., 2016), and GRACE total water storage (Zhao & Yang, 2018). Work is

underway to add capability to assimilate solar induced fluorescence (SIF) and the latest generation of spaceborne soil moisture observations.

3 Simulations and assessment

3.1 Simulations

Table 2 lists the CLM4, CLM4.5, and CLM5 simulations that have been performed. This set of experiments provide a comprehensive assessment of CLM across model generations and across common CLM configurations, as well as the basis to assess the sensitivity to forcing datasets. The assessment of three model versions allows readers to understand the progression of model performance and provides context for CESM1 versus CESM2 These include simulations that apply LAI prescribed from satellite phenology (SP) and simulations with prognostic vegetation state and active biogeochemistry (BGC). Note that only CLM5 has the capability to dynamically simulate crop management and crop management change through time so this simulation is defined as CLM5 BGC-crop. All simulations were completed at a resolution of 0.9^{o} latitude by 1.25^{o} longitude and except where indicated include all required historical or future CLM forcings (as applicable for each configuration) including time series of CO_2, aerosol deposition, N deposition, and land use change. The projection period (2015-2300) simulations, which used the 'anomaly forcing' method (Lawrence et al., 2015), and the no land use change simulations are not assessed here but are available to the community via the data portal for use. The +N and $+CO_2$ simulations are 20-year long simulations starting in year 1995 that replicate the CLM4, CLM4.5, and CLM5 BGC simulations but with a step increase of: 1) nitrogen deposition (5 g N m^{-2} y^{-1} above ambient evenly distributed over the year) and 2) atmosphere CO_2 concentration (200 ppm over ambient) (see Wieder et al., 2019 for further detail). Results from the $+CO_2$ and +N experiments are described in Section 4.7. Note that we restrict our analysis to land-only simulations in this manuscript. However, for reference we include assessment of land quantities in CESM1 versus CESM2 with ILAMB and the CLM diagnostics package (see Section 3.3). The performance of CLM5 within CESM2 will be assessed in a separate manuscript.

The standard CLM spin-up protocol is used to achieve carbon, water and energy equilibrium at the start of the simulation. The year 1850 equilibrium conditions are calculated by integrating over a repeating 20-

year period of an atmospheric reanalysis dataset (i.e., years 1901 to 1920 from the forcing datasets described below) along with fixed atmospheric CO_2, N deposition, aerosol deposition, and land use (note that wood harvest is set to zero during spin-up). As with earlier versions of CLM, it is prohibitively expensive to run the full model for the period of time required to achieve a quasi steady-state. Thus, the spin-up procedure involves a new "accelerated decomposition" methodology, updated from that introduced in Thornton and Rosenbloom (2005) and Koven et al. (2013), with modifications for CLM5 to both add a geographic term to the acceleration and also to accelerate the stem and coarse root C turnover. During the accelerated decomposition phase, the decomposition of the slow C pools (e.g., the long turnover time soil C and coarse woody debris pools) are artificially increased to allow faster convergence on the equilibrium state (see Section 21.8 of CLM5 Technical Description for details). The CLM historical simulations assessed here were initialized from spinup simulations that consisted of ~400 years in accelerated mode, followed by an additional 400-800 years in 'normal mode'. Though the length of time for spin-up varies across configurations, by the end of the spin-up, the global total ecosystem C is drifting by less than 0.02 PgC yr^{-1} and fewer than 5% of grid cells are out of C balance by more than 1 gC m^2 yr^{-1}. For CLM5, initial/cold start (prior to spin-up) soil C and N stocks are increased substantially over earlier model versions, which was done to permit vegetation establishment in harsh environments (where the need for plants to pay for N uptake can inhibit growth under marginal conditions). In some high latitude grid cells, however, vegetation does not survive and soil C turnover is slow due to cold climate conditions. In these locations, the high initial soil C stocks do not deplete during the accelerated spin-up, which leads to unrealistically high equilibrium soil C stocks in those grid cells. To circumvent this undesirable feature, the C stocks of the slow C pools are set to zero where vegetation C is < 0.1 gC m^{-2} by the end of the accelerated spin-up phase.

3.2 Meteorological forcing datasets

For comparison, we utilize three historical meteorology/climate forcing datasets (1901-2014) which are drawn from standard forcing datasets that will be used within LS3MIP (Van den Hurk et al., 2016).

GSWP3v1 – The Global Soil Wetness Project forcing dataset (GSWP3) is the default forcing dataset for LS3MIP (Van den Hurk et al., 2016) and LUMIP (Lawrence et al., 2016)land-only simulations. It is a 3-

hourly 0.5° global forcing product (1901-2014) that was developed for the third phase of GSWP3 (http://hydro.iis.u-tokyo.ac.jp/GSWP3/). It is based on the 20th Century Reanalysis version 2 performed with the NCEP model (Compo et al., 2011). The reanalysis was dynamically downscaled to T248 (0.5°) resolution using the Global Spectral Model using a spectral nudging technique (Yoshimura & Kanamitsu, 2008). Bias correction for temperature, precipitation, and longwave radiation, and shortwave radiation were made using CRU TS v3.21 (Climate Research Unit, Jones & Harris, 2013), GPCCv7 (Global Precipiation Climatology Centre, Schneider et al., 2014), and SRB (Surface Radiation Budget) datasets, respectively. A wind-induced undercatch correction was applied.

CRUNCEPv7 – CRUNCEP is the default forcing dataset used in the Global Carbon Project TRENDY simulations (Le Quéré et al., 2018) and MsTMIP simulations (Huntzinger et al., 2013). It is also a secondary forcing dataset for LS3MIP land-only simulations. It is a 6-hourly 0.5° global forcing product (1901-2015) which is a combination of the CRU TS v3.24 monthly climate dataset (Jones & Harris, 2013) and NCEP reanalysis (Kalnay et al., 1996). The reanalysis is only used to generate diurnal and daily anomalies added to CRU TS monthly means. Precipitation, temperature, cloudiness, and relative humidity are all based on CRU while longwave radiation, pressure, and windspeed are taken directly from NCEP.

WATCH/WFDEI – WATCH is a 3-hourly or 6-hourly, 0.5° global forcing product (1901-2001). It uses the CRU TS2.1 (Mitchell & Jones, 2005) and GPCCv6 datasets to provide the mean climate and the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA-40) product to distribute the mean monthly climate to daily and hourly estimates. Years 1958-2001 are based directly on ECMWF Reanalysis (ERA-40) whereas years 1901–1957 are based on reordered ERA-40 data. Corrections have been applied for seasonal- and decadal-scale variations in the effects of tropospheric and stratospheric aerosol loading on solar radiation, thereby accounting for the effects of global “dimming” and “brightening.” Additional detail about the WATCH dataset is available in Weedon et al. (2011). Note that simulations with WATCH forcing only run through year 2001. We also utilize the WFDEI product, which utilizes the WATCH methodology to the ERA-Interim reanalysis dataset (Weedon et al., 2014). This WFDEI dataset covers the period 1979-2012. Due to the short record, we only use WFDEI dataset for SP simulations.

3.3 ILAMB

The International Land Model Benchmarking (ILAMBv2.1, Collier et al., 2018) package is used to assess the models. ILAMB is an open source land model evaluation system that operates on global-, regional-, and site-level data and provides a hierarchical scoring system to indicate model fidelity. The ILAMBv2 version used here integrates analysis for 28 variables utilizing more than 60 datasets and data products. For each variable, ILAMB produces statistics, maps, time series, and metrics for annual mean, bias, relative bias, RMSE, seasonal cycle phase, spatial distribution, interannual variability, as well as variable-to-variable assessments. Both global and regional assessments are included.

To address a range of questions related to the impacts of model configuration (e.g., prescribed satellite vegetation phenology (SP in CLM infrastructure terminology) versus prognostic vegetation and biogeochemistry (BGC in CLM infrastructure terminology), model structural evolution across CLM generations (CLM4 versus CLM4.5 versus CLM5), and forcing datasets (GSWP3v1 versus CRUNCEPv7 versus WATCH), we ran ILAMB for several different sets of the model simulations listed in Table 2 (http://www.cesm.ucar.edu/experiments/cesm2.0/land/diagnostics/clm_diag_ILAMB.html):

- Structural evolution BGC mode: (CLM4BGC, CLM4.5BGC, CLM5BGC; GSWP3v1 and CRUNCEPv7 forcing):
- Structural evolution prescribed vegetation mode (CLM4SP, CLM4.5SP, CLM5SP; GSWP3 forcing):
- Forcing uncertainty (CLM5BGC and CLM5SP with GSWP3, CRUNCEP, and WATCH/WFDEI)
- Structural and forcing uncertainty (CLM4, CLM4.5, CLM5 with GSWP3 and CRUNCEP, SP and BGC):

The CLM diagnostics package provides a vast set of additional plots and tables, including plots for many variables that are not included in ILAMB as well as seasonal comparisons against selected observed datasets. CLM diagnostic package results are available here for reference: http://www.cesm.ucar.edu/experiments/cesm2.0/land/diagnostics/clm_diag_PCKG.html.

4 Results

In this section, we present a representative sample of analyses that are selected to emphasize strengths and weaknesses of CLM5, relative to CLM4 and CLM4.5, as well as to highlight new features of the model. Due to the breadth of model improvements and the scope of the model output, the assessment presented here is necessarily incomplete. Companion manuscripts focused on CLM5 for the CESM2 Special Issue provide more in-depth assessment of specific aspects of the model (CO_2 and N-additions response, Wieder et al., 2019; plant hydraulics, Kennedy et al., 2019; C-N interactions and parameter uncertainty, Fisher et al., 2019; urban datasets, Oleson and Feddema, 2019; and terrestrial carbon cycle uncertainty, Bonan et al., 2019).

4.1 Assessment with ILAMB

Encouragingly, there is a general progression in the quality of the simulations across model generations. CLM5 outperforms CLM4 for the majority of assessed variables (Figure 3, see also http://www.cesm.ucar.edu/experiments/cesm2.0/land/diagnostics/clm_diag_ILAMB.html). We refer the reader to ILAMB output where vast amounts of additional figures and statistics are available. The improvements from CLM4.5 to CLM5 are comparatively subtle with several variables showing improvement (biomass, burned area, LAI, net ecosystem carbon balance, latent heat, terrestrial water storage, albedo, net ecosystem exchange, ecosystem respiration), but others showing degradation (soil carbon, runoff, surface net radiation, CO_2). The broad improvements across model generations is an emergent feature of the comprehensive model development activities described in Section 2. Definitive identification of the source of particular improvements (or degradation) is beyond the scope of this paper, but some insight is provided in the analyses below. Note that ILAMB results should be interpreted carefully. The summary scores shown in Figure 3 reflect integrated scores across multiple metrics (RMSE, bias, interannual variability, spatial pattern, etc) and for some variables also multiple observational datasets. An overall improved or degraded score for a particular variable can be a result of a mix of scores for individual metrics. For runoff, for example, the overall score is degraded in CLM5 which, when one drills down into ILAMB output, comes from a combination of degraded interannual variability, improved spatial distribution and a slightly greater mean bias (shifting from a low bias in CLM4 to a high bias of similar magnitude in CLM5 when forced with GSWP3v1; when forced with

CRUNCEPv7, all model versions show a large low bias in runoff). Consequently, the overall reduced score for runoff should be considered within this more nuanced perspective.

ILAMB scores indicate a degradation in the simulations of soil carbon stocks from CLM4.5 to CLM5, but the observed estimates are known to be highly uncertain. An alternative soil carbon metric that evaluates the models against apparent soil carbon turnover time shows an improvement from CLM4.5 to CLM5 (Section 4.6). This apparent disagreement between two metrics of soil carbon highlights one of the challenges of benchmarking. When there is disagreement across metrics, we argue that the metric that emphasizes a model process is more meaningful than one that simply evaluates a stock or flux. Consequently, in this instance, our interpretation (based partly on expert judgement) is that the representation of soil carbon is actually slightly improved in CLM5, even though the ILAMB assessment indicates otherwise. We refer the readers to Collier et al. (2018) for more information on how observed dataset uncertainty is accounted for in ILAMB and note that improved treatment of observational data uncertainty is ongoing within the ILAMB project.

ILAMB also assesses functional relationships between two variables (e.g., precipitation versus GPP or LAI). CLM5 performs better than CLM4 or CLM4.5 for the majority of the functional relationships assessed (Figure 4), suggesting improved process-representation in CLM5. In particular, the relationships between GPP and climate variables such as solar radiation and precipitation are improved, though there is a slight degradation (CLM4.5 to CLM5) of the relationship between GPP and surface air temperature. Relationships between burned area and climate are also improved (see ILAMB plots), with burned area correctly peaking at average annual precipitation rates of 2.5 to 5 mm d^{-1}, an ecoclimatic regime that is dry enough for fire, but productive enough to establish fuel loads.

The ILAMB system was designed to probe model performance across both timescales and spatial scales. At the global scale, the seasonal cycle of atmospheric CO_2 deduced from CLM carbon fluxes improved substantially from CLM4 to CLM5, especially in the mid-to-high northern latitudes. However, the magnitude of interannual variability has degraded, especially in the tropics. For all CLM model versions, the Northern Hemisphere interannual variability is at most one-third of that observed at NOAA marine boundary layer sites.

Utilizing ILAMB, we can also identify a significant sensitivity of simulation output to the forcing dataset (Figure 5). While all of the forcing datasets used in this study are observationally-derived, each one employs different methodology for downscaling and bias correction and can therefore potentially be assessed with ILAMB. GSWP3-forced simulations score best for most of the forcing variables (assessed forcing variables are surface air temperature, precipitation, surface relative humidity, surface downward shortwave and longwave radiation) with relative humidity being the exception. Generally, CLM5 scores best for simulations forced with the GSWP3 forcing dataset. The fact that model output variables score better with the best (according to ILAMB) forcing dataset suggests, not surprisingly, that land models are likely to perform better with more accurate forcing, particularly when functional relationships are represented reasonably by the model.

As noted in Section 3.1, it is beyond the scope and aim of this paper to provide an assessment of the performance of CLM5 within CESM2. However, we direct interested readers to the ILAMB results for CESM1/CLM4 versus CESM2/CLM5 that we provide on the ILAMB webpage associated with this paper. In those results, we see that the land climate forcing variables (e.g., surface air temperature, downwelling shortwave and longwave radiation, surface relative humidity, etc.) are generally marginally improved in CESM2 (with the exception of precipitation which shows slight degradation). The assessed land carbon, water, and energy variables show similar improvements in the coupled simulations (i.e., from CESM1 to CESM2) as they do in land-only simulations (CLM4 to CLM5). The modest improvement in coupled model land forcing quantities combined with the consistent relatively strong improvements in land-only and coupled simulations implies that the source of improvement in land surface variables derives from developments in CLM, rather than due to improvements in other components of CESM.

4.2 Plant Functional Type-level assessment

Biases in the annual monthly maximum Leaf Area Index (LAI) for selected PFTs are shown in Figure 6 and for all PFTs in Figure S2. CLM5 shows reduced root mean square error compared to MODIS LAI (Table 3) for 9 out of 14 PFTs compared to CLM4.5. Broadleaf evergreen tropical trees, broadleaf deciduous temperate trees and C4 grasses showed the biggest improvement.

During the course of the development of CLM5, we tested the model with parameter sets that resulted in considerable areal fractions of the vegetation not surviving for one or more PFTs. This result lead us to routinely track survival percentage throughout the model development process. Survival percentage for each PFT is reported in Table 3. In general, survival percentage is slightly higher in CLM5. Survival fraction plots in Whitaker space are shown in Figure S3. We can see, unsurprisingly, that for most PFTs survival fractions are low in dry and warm climates or in very cold climates. CLM PFTs have the same parameters across their entire geographical range, thus not accounting for geographical trait variations which could nonetheless regulate surface fluxes (Giardina et al., 2018; Konings & Gentine, 2017). Land models where PFTs or their parameters are more disaggregated, for example into those adapted for more and less productive environments (e.g., CLM-FATES), should in principle be able to circumvent this issue. It is important to note that in CLM, once a PFT dies (i.e., vegetation C goes to zero) in a particular grid cell, that PFT cannot grow back during the course of the simulation, even if climate conditions become more amenable for survival.

Maximum carboxylation rate at 25℃, V_{cmax25}, values (representing leaf canopy average) for each PFT and each model version are shown in Table 3, and are compared to the synthesized Kattge et al. (2009) observational estimates. In CLM4 and CLM4.5, the V_{cmax25} values are prescribed with the values in CLM4.5 specifically calibrated to reflect data in Kattge et al. (2009), except for broadleaf evergreen tropical trees which were adjusted upwards so as to produce a viable tropical forest photosynthesis levels. In CLM5, V_{cmax25} is a prognostic quantity (see Section 2.3.7), and the values shown in the table represent a spatially-weighted average monthly maximum V_{cmax25} for each PFT. With the model's current parameterization, CLM5 predicts V_{cmax25} values that are lower than the observational estimates for most PFTs, especially C3 grasses (Table 3, Figure S4). The discrepancy may be partially related to the fact that observed values of V_{cmax25} may not represent the environmental conditions (e.g., shading) as experienced by the plants in CLM, in addition to challenges associated with the limited spatial representativeness of the observed values. The ability of the model to represent photosynthesis and leaf area index reasonably well even with such low V_{cmax25} values is potentially indicative of a structural problem in the leaf-level versus canopy-scaled value (as discussed in Rogers et al., 2017) which will be investigated further using offline tools such as those presented by Walker et al. (2018). The prognostic V_{cmax25} values produced in CLM5 should be perceived as an initial effort to incorporate parameterizations that can simulate changes

in leaf N allocation and photosynthetic capacity under environmental change. Further investigation is needed to improve the model representation of photosynthetic capacity.

Simulated canopy height and canopy height biases with respect to those derived from ICEsat (Simard et al., 2011) are shown for all tree PFTs for CLM5BGC in Figure S5. On average, boreal needleleaf evergreen trees are too tall by 5-10 m while tropical broadleaf evergreen trees and temperate and boreal deciduous trees are too short by 5-10 m. These biases are related to simulated plant biomass as well as uncertainties in the specified allometric relationships between biomass and height. Biases in canopy height will affect the land surface roughness length and therefore turbulent heat flux exchange between the land and the atmosphere.

4.3 Hydrology

The main changes to soil hydrology (see Section 2.3.2) are: (1) introduction of spatially variable soil depth (depth to bedrock), (2) replacement of the unconfined aquifer that existed below the soil column with a no flux bottom boundary condition (Swenson & Lawrence, 2015) and (3) a revised soil evaporation parameterization that accounts for the rate of diffusion of water vapor through a dry surface layer (Swenson & Lawrence, 2014). Figure 7 illustrates the impacts of these new features for two example grid cells in southwest and southeast US. At the southwest US grid cell, one can see that ET is too variable compared to the observations for CLM4 and CLM4.5. With the dry surface layer in CLM5, soil evaporative water losses are restricted, resulting in improved ET seasonality. Water from snow melt and spring rains then infiltrates deeper into the soil column (which is 8.5 m deep at this location), providing a source of moisture for evaporation into the summer months. At the eastern US grid cell, we can observe a different feature of the new model. The shallow 1 m thick soil prescribed at this location in CLM5 cannot store much water. Consequently, we can see strong drying throughout the soil column in the low precipitation year of 1993, which then restricts ET from summer into fall, in agreement with observations. In CLM4 and CLM4.5, ET is unrealistically supported through this period by soil water that is stored deeper in the standard 3.5 m thick soils.

ILAMB and CLM diagnostics package results indicate only relatively small changes in the quality of annual streamflow for the top 50 biggest rivers. In particular, mean flow for the Amazon and Congo

rivers is increased and shows better agreement with observed flows, with the improvement mainly due to reduction of the excessively high tropical forest ET that was seen in CLM4. The mean bias in global annual mean river flow is slightly degraded, with CLM5 showing a high bias in global river discharge in both SP and BGC configurations (bias is larger in BGC mode). On the other hand, the global annual mean bias and bias/RMSE scores for ET show nominal improvement in CLM5. We also note that differences in simulated runoff and ET between forcing datasets is larger than the differences across model versions.

Assessment of the impact of hydrology changes on simulated land-atmosphere interactions is beyond the scope of this manuscript. However, we can infer that the relationship is likely to differ by examining the simulated soil moisture residence time (SMRT) across models. SMRT is the e-folding decay timescale of soil moisture due to evapotranspiration, is an integrative measure of soil-plant-atmosphere dynamics. We calculate SMRT for the root zone (0-0.5m) from daily soil moisture curves during post-rain periods using a procedure similar to the estimation of a base flow recession constant (Vogel & Kroll, 1996). This residence time metric is reflective of the evapotranspiration dry-down response time-scale (Teuling et al., 2006). In Figure 8, SMRT as simulated by CLM5 is shown for the continental United States for the May to October warm season and is compared to observationally-derived estimates from the North American Soil Moisture Database (Quiring et al., 2016). In general, the SMRT as simulated by CLM5 compares well with observations except for the western United States where observations show a wide range of residence times from less than 60 days to greater than 90 days whereas CLM5 shows uniformly longer residence time (120 days or more). At least some of the western US discrepancy could be attributed to the poorly resolved topographic gradients at the nominal 1^{o} resolution of these simulations. Figures 11b and 11c compare the soil moisture residence time in CLM5 with that in CLM4 and CLM4.5. Overall, the soil moisture residence time in CLM5 has increased across much of the eastern US and decreased in parts the western US compared to both CLM4 and CLM4.5. Identification of the source of the changes in residence time is beyond the scope of this paper, but the spatially-explicit soil depths, the introduction of the dry surface layer parameterization for soil evaporation, and soil moisture dynamics associated with the plant hydraulic stress routine are all likely to be factors. Averaged across the continental United States domain, soil moisture residence time is higher by 15% compared to CLM4.5, and 1.5% compared

to CLM4. Dirmeyer et al. (2016) concluded that the soil moisture residence time in CLM4 was 18% too low, so the lengthened residence time in CLM5 may represent a change in the desired direction.

The residence time metric suggests improvements in CLM5 compared to CLM4, and CLM4.5 with CLM5 shows a generally higher soil moisture residence time across majority of the soil moisture observing network, as one would expect with generally deeper soils and stronger soil evaporation limitations associated with the dry surface layer parameterization. In many regions this moves the model further from observed estimates (Supplementary Table S1) though caution is warranted when comparing CLM SMRT with observationally-derived SMRT due to uncertainties from a number of sources including uncertainties in observationally-derived SMRT due to different type of sensors and measurement techniques at each site, the substantial spatial-scale mismatch between grid cells and observational sites, as well as uncertainties in model parameterizations (Dirmeyer et al., 2016). We repeated our calculations using the soil moisture memory metric employed in Dirmeyer et al. (2016) and found a similar change in CLM5 compared to CLM4 and CLM4.5 (not shown). Changes in SMRT are likely to impact a range of land-atmosphere interaction phenomenon including land-driven climate predictability.

4.4 Plant Hydraulic Stress and ET partitioning

The plant hydraulic stress (PHS) configuration implements new parameterizations for root water uptake and water stress for CLM5. For comparison, we also ran CLM5 with PHS replaced with the soil moisture stress (SMS) parameterization included in prior CLM versions (see Section 2.3.6). One of the broadest impacts of PHS is a decrease in the coefficient of variation of GPP (CV_{GPP}) and transpiration (CV_{ET}) (Figure 9d,h). The global distributions of CV_{GPP} and CV_{ET} both shift towards lower values with PHS (Figure 9c,g), corresponding to global reductions in CV of 8.0 and 12.5% for GPP and ET, respectively, relative to SMS. Decreases in CV_{GPP} tend to occur in water-limited ecosystems with seasonal rainfall, such as the Sahel region of Africa and northern Australia (Figure 9d). PHS incorporates more flexible root water uptake (Kennedy et al., 2019), which can utilize more of the soil column to buffer shortfalls in precipitation, acting to reduce variability imposed by precipitation variations. CV_{ET} decreases follow roughly the same patterns, reflecting the coupling of transpiration and photosynthesis through stomatal conductance. (Figure 9h). With PHS, vegetation water stress is sensitive to atmospheric demand for transpiration and tends to narrow the range of transpiration values, which results in relatively larger

reductions in CV_{ET} as compared to CV_{GPP}. In some regions, variability increases with PHS, primarily at high latitudes (e.g., eastern Siberia) and in arid regions. Such increases in CV_{GPP} and CV_{ET} are generally associated with increases in the mean fluxes of GPP and ET in these regions with PHS. Other mechanisms unrepresented in CLM, including adaptive responses of V_{cmax25} to dry conditions and biochemical responses to stress (Keenan et al., 2009; Niinemets & Keenan, 2014) could in principle increase interannual variability of these fluxes, thus the decrease in variability seen here is not necessarily indicative of a structural degradation or inappropriate PHS parameters.

The partitioning of evapotranspiration into transpiration, canopy evaporation, and soil evaporation is a key emergent process simulated by land models, essential to assess ecosystem water use efficiency (Lawrence et al., 2007). In Figure 10, we show the transpiration fraction from each model compared to estimates of transpiration fraction from the Water, Energy, and Carbon with Artificial Neural Networks dataset (WECANN, Alemohammad et al., 2017 , available at https://gentinelab.eee.columbia.edu/content/datasets). In prescribed vegetation configurations, CLM5SP shows better agreement with WECANN transpiration fraction than either CLM4SP or CLM4.5SP, especially in the Tropics. Globally, the contribution of soil evaporation to ET is diminished in CLM5 relative to CLM4 and CLM4.5, resulting in a higher percentage of ET coming from transpiration (CLM4SP and CLM4.5SP 53%; CLM5SP 60%, Table 4), in line with recent isotopic data estimates of 61 ± 15% (Jasechko et al., 2013; Schlesinger & Jasechko, 2014). However, in prognostic vegetation mode, biases in simulated LAI lead to poorer agreement with WECANN for ET partitioning for all model versions. In particular, low LAI biases for tropical deciduous trees (Figure S2), especially in the Sahel and southern Africa, appear to correlate with low biases in transpiration fraction, though errors in the observations. Table 4 shows global percentages for transpiration, soil evaporation, and canopy evaporation for the CLM versions. Note that simulations forced with CRUNCEPv7 show a higher proportion of ET coming from canopy evaporation than GSWP3v1-forced simulations. This difference is likely due to the temporal frequency of the forcing precipitation (6-hourly for CRUNCEPv7, 3-hourly for GSWP3v1), which can have a strong impact on canopy evaporation.

4.5 Permafrost and snow density

Permafrost is a key feature of the earth system and uncertainty regarding the strength of the permafrost climate-carbon feedback is considerable (McGuire et al., 2018; Schuur et al., 2015). The permafrost climate-carbon feedback is a challenging research problem that depends on many features of a land modeling system. A known deficiency in prior versions of CLM was an unrealistically low fresh snow density, which led to excessive snow insulation of the ground, particularly at low snow depths (Slater et al., 2017). Several changes to fresh snow density and snow densification were introduced in CLM5 (van Kampenhout et al., 2017) resulting generally in denser snow for both seasonal and perennial snowpacks. The denser snow over Greenland and Antarctica is an improvement and along with the deeper snowpack allows the model to more realistically represent firn and the transition from snow to ice. The denser surface snowpack also largely eliminates excessive sub snow-surface melt that occasionally occurred in CLM4 and CLM4.5 in very cold climates where the simulated near-surface thermal conductivity was unrealistically low.

The changes to modeled snow density also have beneficial impacts on permafrost distribution and active layer thickness (ALT, the depth to which permafrost soils thaw each summer). In Figure 11, maps of ALT and February snow density are shown for CLM4.5 and CLM5 with GSWP3v1 and CRUNCEPv7. These maps reveal that there are strong relationships between the forcing dataset, the snow density formulation, and simulated ALT. Snow is denser across the permafrost domain in CLM5 (225 to 275 kg m^{-3}) compared to CLM4.5 (< 200 to 225 kg m^{-3}). This denser snow in CLM5 is more consistent with the values of 230 to 330 kg m^{-3} reported for northwest Alaska (Sturm et al., 2010). The denser snow reduces snow insulation and results in colder soils and shallower ALT in CLM5 compared to CLM4.5.

It is also relevant to note the impact of forcing dataset on snow density and ALT simulations. Snow tends to be less dense with CRUNCEPv7 forcing than with GSWP3v1 forcing. Taken in isolation, this should lead to shallower ALT with GSWP3v1 forcing, but instead ALT is generally deeper which appears to be due largely to greater downwelling longwave and shortwave radiation in GSWP3v1 forcing data. The large differences in simulated permafrost distribution and ALT between the two forcing datasets reveals an important aspect of uncertainty in permafrost modeling (which propagates to uncertainty in modeled soil carbon stocks, as discussed below). ILAMB output indicates that downwelling longwave radiation, downwelling solar radiation, and humidity variables all score significantly higher across the Arctic land domain with GSWP3v1 (other forcing quantities are roughly equivalent across these two forcing datasets)

which suggests that for permafrost studies, GSWP3v1 forcing may be more appropriate. If we consider just the GSWP3v1-forced simulations, we see that CLM4.5, with its low-density snow, exhibits ALT that is unrealistically deep (ALT > 1m deep across nearly the entire permafrost domain) while CLM5, with its denser snow, is more realistic. These results are an indirect indication that the CLM5 snow density parameterizations may represent an improvement.

4.6 Carbon and nitrogen fluxes and stocks

Table 5 lists the simulated global total carbon stocks and annual mean fluxes for the different model versions compared to available data products. Global GPP agrees best with available data products for CLM5 (119 PgC yr^{-1} in CLM5BGC, 134 PgC yr^{-1} in CLM4BGC, 118 PgC yr^{-1} for FLUXNET-MTE observed GPP estimate; values are for area of land intersection between model and observations, i.e., grid cells where model and observations agree there is land). Net primary productivity (NPP) increased in successive versions of the model, reflecting concurrent declines in autotrophic respiration (AR) fluxes. As such, annual mean carbon use efficiency (CUE, calculated as the quotient of grid cell NPP and GPP) increased from CLM4 to later versions of the model, although all three model versions show high spatial heterogeneity in CUE (Wieder et al., 2019). The latitudinal variation of CUE simulated by CLM5 seems plausible, based on published estimates (Campioli et al., 2015; DeLucia et al., 2007; Malhi et al., 2011; Vicca et al., 2012), but deserves further investigation. All three model versions reasonably replicate the global totals for vegetation carbon stocks, but the spatial distribution differs across models. ILAMB results show that CLM4BGC placed too much carbon into tropical rainforests and too little carbon into boreal forests, especially across Europe and Siberia. To first order, the biases are reversed in CLM5BGC with too little carbon in the tropical rainforests and too much carbon across the boreal forests, largely reflecting the spatial pattern of GPP biases but likely also related to changes in C allocation in CLM5.

Soil C stock patterns are more realistic in CLM4.5BGC and CLM5BGC than in CLM4BGC because of the introduction of vertically-resolved soil biogeochemistry in CLM4.5 (Koven et al., 2013), which allows the model to generate large C stocks across the northern high-latitude permafrost domain, as observed. The relationship between apparent soil C turnover times (defined as the ratio of mean soil C stocks over climatological annual mean NPP) and mean air temperature is more realistic in CLM4.5 and

CLM5 (Figure 12, metric reproduced as in Koven et al., 2017), with both of these model versions at least partially capturing the transition to longer apparent soil C turnover times in cold climates. This metric suggests that CLM5 apparent soil C turnover times are slightly improved over CLM4.5 with a steeper increase in turnover times at cold temperatures as well as a broader spread of turnover times in warm climates associated with soil wetness (short turnover times in warm-wet climates and long turnover times in warm-dry climates). Because of the greater permafrost extent and colder permafrost soil temperatures in CLM5 when forced by CRUNCEPv7 than by GSWP3v1, the stocks of soil C to 3m depth are a factor of 2 larger when forced by CRUNCEPv7 (4000 Pg C) than when forced by GSWP3v1 (1925 Pg C), demonstrating the extreme sensitivity of simulated permafrost soil C stocks to simulated permafrost conditions.

The spatial distribution and global sums of terrestrial N inputs and losses remain poorly constrained with data and highly variable among versions of CLM. Table 6 shows published estimates of global terrestrial N fluxes and corresponding estimates from the GSWP3-forced BGC simulations. Within CLM, N inputs come from N deposition and N fixation. Inputs from N deposition are consistent among model versions, with forcings coming from Lamarque et al. (2010), and show broad agreement with observationally derived estimates (Fowler et al., 2013). Estimates of global N fixation show greater spread among models. The empirical approach applied in CLM4 and CLM4.5 estimated biological N fixation rates as function of NPP (Cleveland et al., 1999). CLM5 calculates both symbiotic and free-living N fixation. Total N fixation in CLM5 is lower than in previous versions of the model and lies within the range of estimates of N fixation rates (Vitousek et al., 2013). Finally, with the ability to simulate a global interactive crop model, CLM5 provides opportunities to estimate anthropogenic changes to the terrestrial N cycle through planting N fixing crops and fertilizer application. The N fixation rates simulated by soy in the model are well below upscaled estimates of agricultural N fixation (Herridge et al., 2008), but simulated fertilization rates appear to be on target with observational estimates (Fowler et al., 2013). CLM simulates N losses through leaching, gaseous emissions, and biomass removal. Successive model versions show increasing hydrological N losses, though these have not been evaluated against data. Houlton et al. (2015) pointed out that gaseous N losses were too high in CLM4. The same is likely true with CLM5, which still suffers from poorly implemented representation of soil N dynamics resulting in a high bias in gaseous (as opposed to hydrologic) N losses. With intensification of land use and land

management, CLM5 also shows anthropogenically-driven N losses associated with wood harvest, crop harvest, and land use change. These N loss fluxes, as well as gaseous N emissions (including NOx emissions due to fire and soil N_2O fluxes) remain poorly constrained and an area for future model evaluation and development.

4.7 CO_2 and N-addition response

Over the course of model development, CLM (BGC configurations) transitioned from a model that exhibited strong N limitation of the terrestrial carbon cycle (CLM4) to a model that showed greater responsiveness to elevated concentrations of CO_2 in the atmosphere (CLM5) (Wieder et al., 2019), consistent with recent observations that suggest that there has been only weak N limitations on CO_2 fertilization (Campbell et al., 2017). Specifically, the carbon cycle simulated by CLM4 showed an unrealistically strong nitrogen limitation (Bonan & Levis, 2010; Thomas et al., 2013a; Thomas et al., 2013b) and a lower than observed response to CO_2 enrichment (Figure 13; Hoffman et al., 2014; Walker et al., 2014; Zaehle et al., 2014). With revisions to the photosynthesis parameterization and soil biogeochemical model (Bonan et al., 2011; Koven et al., 2013), CLM4.5 showed a lower sensitivity to N enrichment than its predecessor that was more in line with observations (LeBauer & Treseder, 2008), but it still exhibited lower sensitivity to CO_2 enrichment than observations from Free Air CO_2 Enrichment (FACE) sites (Ainsworth & Long, 2004). CLM5 includes a suite of model developments focused on improving the representation of vegetation C-N dynamics (outlined in section 2.3.6). The globally integrated response of terrestrial ecosystems to N and CO_2 enrichment suggest that CLM5 shows improved agreement with observed ecosystem response to these environmental manipulations (Figure 13; Ainsworth & Long, 2004; LeBauer & Treseder, 2008), though the globally-integrated improved agreement with these syntheses should not be overinterpreted . Besides capturing the appropriate magnitude of terrestrial C pools and fluxes to N enrichment, simulations with CLM5 also show increases in foliar N content and ecosystem C use efficiency that are consistent with observations (Campioli et al., 2015; Vicca et al., 2012; Wieder et al., 2019). Similarly, foliar N content and V_{cmax} decline under elevated CO_2, again consistent with observations (Ainsworth & Long, 2004). Together, these results suggest that CLM5 better captures terrestrial ecosystem responses to global change drivers (N and CO_2 enrichment) than previous versions of the model, and that CLM5 captures these responses for at least

some of the right reasons, though in a parameter uncertainty assessment, Fisher et al. (2018) demonstrate a strong sensitivity of CO_2 and N fertilization to several factors including to what extent plant communities can increase their fixation of nitrogen, stoichiometric flexibility of plant tissues, nitrogen uptake costs, and model initial state. The strong sensitivity to parameters as well as large differences in response across PFTs (see Wieder et al. (2019) for further figures and discussion) implies that the large-scale agreement with observations should not be overinterpreted as indicative that the model is necessarily reproducing the observed response for the right reasons.

Confronting land models with perturbations that are similar to experimental manipulations also exposes shortcomings in the model's structural assumptions and parameterizations. For example, although the bulk C cycle response to N enrichment simulated by CLM5 appears more appropriate than CLM4 or CLM4.5, the model still fails to capture observed shifts in plant C allocation towards greater above ground productivity or decreases in heterotrophic respiration that are commonly seen in nutrient addition experiments (Janssens et al., 2010; Liu & Greaver, 2010). Similarly, terrestrial sensitivities to elevated CO_2 simulated by CLM5 seem more in line with observed responses, but the model achieves higher productivity by increasing LAI and nitrogen fixation rates beyond what is likely to occur in natural ecosystems (Ainsworth & Long, 2004; Hungate et al., 2004; Medlyn et al., 2015; Terrer et al., 2018). Indeed, results from experimental manipulations emphasize that acclimation as well as changes to plant allocation (which are not represented in CLM5) and stoichiometry are important aspects of terrestrial ecosystem responses to global change drivers (Liu & Greaver, 2010; Luo et al., 2006; Reich et al., 2006). Despite its improvements, CLM5 still has limited capacity to capture these responses, highlighting priority areas that should be addressed in future model developments. Specifically, understanding and modeling appropriate changes in above and below ground C and N allocation remains uncertain, especially in response to global change (Giardina et al., 2005; Terrer et al., 2018). This is an outstanding challenge to be addressed in land models and evaluated with observations from experimental manipulations. Despite these limitations, the overall transition towards the use of optimality theories in N-cycle representation in CLM5 and in integrating N-processes directly into plant physiology, rather than the post-hoc reconciliation of N-unlimited and N-limited rates of GPP in CLM4, appears to broadly move the model in the right direction, though there is much work still to do (e.g., resolve limitations in

representation of soil nutrient competition between plants, microbes, and mineral surfaces; Zhu et al., 2016).

4.8 Land carbon accumulation over historical period

The global land C accumulation trends exhibit clear differences across model versions (Figure 14). As noted above, CLM4 produces an unrealistically strong nutrient limitation on photosynthesis, which limits that model's capacity for C uptake even as atmospheric CO_2 increases. Consequently, in CLM4 land-use and land-cover change (LULCC) C loss fluxes dominate over the CO_2 fertilization response resulting in an accumulated land C loss of ~60 Pg C over the period 1850 to 2014, which is outside the observational estimates of -8 Pg C (range +32 to -52 PgC, 1850-2010; Hoffman et al., 2014) CLM4.5, on the other hand shows C uptake and accumulation in response to CO_2 fertilization that is perhaps too strong, especially under the GSWP3v1-forced simulation. The CLM5 land C accumulation curve lies in between CLM4 and CLM4.5 and appears to result in the best match with observational estimates, for the historical period as well as the global carbon project era (1950-2012; Le Quéré et al., 2014). These results are also reflected by the comparatively high scores for the Global Net Ecosystem Carbon metric in ILAMB for CLM5 (Figure 3).

Although it is tempting to infer that the more realistic responses of CO_2 and N additions in CLM5 (Wieder et al., 2019) are responsible for the improved emergent behavior of the model with respect to the historical land C accumulation, historical C accumulation is a function of several sometimes-counteracting processes that control C fluxes and stocks and thus these changes should be interpreted cautiously. These processes include deforestation and wood harvest fluxes and the dependency of these fluxes on initial forest vegetation C stocks, C-uptake responses to increasing CO_2 and N-deposition trends, and vegetation and soil C responses to climate trends and variability. Furthermore and importantly, as noted above, Fisher et al. (2019) demonstrate that CLM5 responses to CO_2 and N fertilization exhibit strong sensitivity to several uncertain parameters. Nonetheless, the improvement in this important emergent behavior of the model is intriguing and is investigated in more depth in Bonan et al. (2019).

Also apparent in Figure 14 is a strong sensitivity to atmospheric forcing with accumulated land C for the period 1850 to 2014 differing between runs forced with GSWP3v1 and CRUNCEPv7 by 50, 20, and 10 Pg C for CLM4, CLM4.5, and CLM5 respectively. The divergence in C accumulation between runs with different forcing datasets arises early in the period, mainly prior to 1950, when CO_2 fertilization would have been relatively small and LULCC fluxes dominate. This implies, then, that LULCC C fluxes can differ substantially even within a model version forced with exactly the same LULCC time series but under different estimates of historical climate forcing. We hypothesize that the simulated pre-industrial (year 1850) vegetation C stocks and their regional distribution can impart a strong influence on historical LULCC C fluxes.

Finally, while the long-term land carbon accumulation agrees better with observed estimates, which are derived from atmospheric CO_2 and ocean C inventories, the interannual variation in land C accumulation appears to be degraded in CLM5 (larger low bias in variability), based on a comparison of interannual variability of atmospheric CO2 simulated from the CLM fluxes compared to that observed (see ILAMB CO_2 diagnostics). Throughout the Northern Hemisphere, interannual variability is at most one-third of that observed at NOAA marine boundary layer sites. The drivers and implications of this degradation from CLM4 to CLM4.5 to CLM5 require further investigation, since climate-driven variations at interannual timescales may provide useful information about future climate-driven changes in terrestrial carbon stocks (Cox et al., 2013; Keppel-Aleks et al., 2018). Preliminary investigation suggests that although the plant hydraulics scheme does tend to reduce variability in GPP and transpiration (see Section 4.4), it does not appear to be primarily responsible for the reduced C flux variability in CLM5, with the reduced variability potentially resulting from increased interannual synchronicity between NPP and ecosystem respiration.

4.9 Water use efficiency

Quantification of changes in water use efficiency (WUE, carbon uptake per unit of water loss) due to climate change and rising atmospheric CO_2 levels is challenging (Cheng et al., 2017). Changes in WUE will have strong implications for water availability, food and fiber production as well as the C sink capacity of terrestrial ecosystems. Though this topic has received considerable recent attention in the

literature (e.g., Cheng et al., 2017; Frank et al., 2015; Huang et al., 2015; Keenan et al., 2013), there is still no consensus on how the coupled terrestrial carbon and water cycles have changed or will change in the future.

A key feature of CLM5 is a more realistic coupling of N limitation and stomatal conductance, with stomatal conductance in CLM5 based on the N-limited photosynthesis (Ghimire et al., 2016) rather than on N-unlimited potential photosynthesis as it was in CLM4 and CLM4.5. This more realistic coupling has consequences for WUE and WUE trends since changes in N-limitation will propagate directly into simulated transpiration. The increase in global WUE (defined here as GPP / transpiration) over the historical period is considerably stronger in CLM5 compared to CLM4 and CLM4.5 (Figure 15). Global GPP trends are comparable across models, though CLM5 marginally exhibits the strongest increase, while CLM4 shows the weakest increase, at least partially due to the high N-limitation in that version (see Section 4.7). Global transpiration trends, on the other hand, diverge considerably across versions with CLM4 and CLM5 showing a declining trend in transpiration during 1980 to 2014 and CLM4.5 showing an increasing trend over the same period. Spatially, the increase in WUE is larger almost everywhere in CLM5 than in the other model versions, but the driver of the WUE change differs considerably by region. In the tropics, the CLM5 increase in WUE is driven by both increased GPP and somewhat reduced transpiration (Figure S6). In the boreal forest and across the mid- to high-northern latitudes, the historical increases in GPP are high, but transpiration is largely unchanged or is weakly increased. Deeper analysis of the WUE trends and its interaction with CO_2 fertilization and LAI, N-limitation, and soil moisture limitation trends across model versions and compared against available estimates of historical WUE trends is worthy of additional study, but is beyond the scope of this paper.

4.10 Crops

Agricultural management practices can have a considerable impact on climate (Bagley et al., 2015; Davin et al., 2014; Lombardozzi et al., 2018; Mueller et al., 2017; Thiery et al., 2017), highlighting the importance of representing agriculture in ESMs. CLM5 is the first version of CLM that includes transient representation of crop distribution and management, and the inclusion of managed agriculture in CLM5 does affect carbon, water, and energy fluxes from the land surface. The representation of crops in CLM5

also allows the model to track crop yields through time. The crop yields simulated by CLM5 increase from 1.1 tonnes ha^{-1} in 1850 to ~3 tonnes ha^{-1} in 2010 (Figure 16c). For the crop types represented in CLM5, the simulated yields match observations for the same crop types from the United Nations Food and Agriculture Organization (UN-FAO) from the start of available observations in 1961 through approximately 1980. Yields in CLM5 level off after that time, whereas the UN-FAO yields steadily increase, with the discrepancy likely due to the fact that crop representation in CLM5 does not include processes associated with intensification, such as increasing planting density. The spatial distribution of crop yields illustrates that CLM5 underestimates crop yields throughout the northern hemisphere compared to UN-FAO, particularly in Central US, Europe, and Southwestern Asia, but overestimates crop yields throughout much of the tropics (Figure 16a, b). Yields of individual crops are generally similar to UN-FAO estimates, though CLM5 underestimates corn yields throughout most temperate regions. The management techniques represented in CLM5 also impact the magnitude of crop yields. Globally, agricultural expansion and fertilization have large impacts on increasing crop yields, and irrigation has a smaller impact due to the fact that less than ~25% of cropland area is irrigated. Irrigation is quite important for crop yields within irrigated areas, however. Note that due to the inflexibility of the planting windows in CLM5, planting dates in some regions, such as India (too early), are unrealistic. A more flexible climate-driven planting date scheme is planned for future model versions.

4.11 Urban

To evaluate behavior of the updated urban model and building properties data, observations from five urban flux tower sites and a global anthropogenic heat flux (AHF) dataset were used. In simulations described in Oleson and Feddema (2019), radiative and turbulent fluxes, surface temperatures, and AHF were found to be generally improved compared to the previous version. The simulation of global and regional AHF is also significantly improved, mainly due to the new building energy model. For example, large positive biases in AHF over the U.S. and Europe, evident in the previous model version, are reduced such that simulated values are now within 1% and 11% of observations, respectively. The increased simulation fidelity and new capabilities of the model should enhance its utility for research into the combined effects of urbanization and global climate change.

5. Summary and Discussion

As with prior CLM versions, the development of CLM5 was an extensive community effort involving researchers from many different institutions and culminating with the integration of numerous disparate development efforts. The resulting updated model represents a significant advancement, relative to prior model versions. CLM5 includes new default and optional functionality, improved flexibility in model configurations and land cover transitions (natural vegetation ↔ glacier, natural vegetation ↔ crop), as well as more mechanistic and ecologically-relevant representations of the physics, biology, and human land management processes that govern terrestrial states and fluxes.

Benchmarking packages such as ILAMB mark a significant enhancement in our ability to evaluate land model representations of water, energy and carbon cycles. Broadly, ILAMB and other metrics presented here indicate that the simulation quality is improved in CLM5 over CLM4 and CLM4.5; although differences between CLM4.5 and CLM5 are less distinct, and particular variables or metrics show degraded performance. However, even with the deployment of advanced model assessment tools and metrics, in many cases a clear and unambiguous demonstration of improvement or degradation for a complex model such as CLM remains challenging. We find for example, perhaps unsurprisingly, that climate and weather forcing uncertainty confound the interpretation of impacts of model structural advances. The impact of parameter uncertainty, which is not assessed here (see Fisher et al. (2019) for partial parameter sensitivity assessment of CLM5). Nonetheless, we interpret the broad indications of improvement across multiple variables and metrics (>30) suggests genuine progress, which (hopefully) is grounded in the upgraded model parameterizations and more comprehensive process representation.

We stress, however, that model users should consider improvements or degradation identified in ILAMB or other metrics presented here with caution due to observed data limitations related to data scale applicability, measurement uncertainties, inconsistencies across multiple observational datasets for one or more variables (e.g., water and energy budgets derived from the available observationally-based ILAMB datasets do not close), as well as limitations in the metrics included in ILAMB. Improved methods within ILAMB to account for observational data uncertainty is critical and is a priority for the ILAMB project.

An improvement or a degradation for a particular variable or metric does not on its own imply that the model is suited or not suited for research related to that particular variable. For example, ILAMB indicates that snow water equivalent is degraded in CLM5 relative to CLM4. This apparent degradation occurs despite several mechanistic improvements to snow physics that have been introduced between CLM4 and CLM5. The lower ILAMB score for snow water equivalent for CLM5 could indicate a real model snow simulation performance degradation (due to structural or parametric problems introduced during development from CLM4 to CLM5), but it could also potentially be attributed due to inaccuracies in the forcing data or biases in the observed dataset used in ILAMB or limitations in the ILAMB metrics themselves. Consequently, CLM5 users interested in applying the model for research into snow processes will need to balance knowledge of the snow physics and snow physics structural advances against the ILAMB score decrease and against their own assessment of snow simulations to decide whether or not the model is 'fit-for-purpose.'

More explicit process representation enables new types of observations to be applied for evaluation of CLM. For example, since CLM5 implements prognostic, rather than prescribed, leaf photosynthetic traits, observations of V_{cmax25} and J_{max25} can be used as a means for assessing the model. Similarly, the introduction of plant hydraulics opens up the potential to employ several observational quantities that were previously not applicable to CLM including mid-day stomatal conductance, leaf water potential, and sap flow. This list could continue, but in general, the expansion of CLM science to more realistically represent physical and ecological processes opens also new opportunities to evaluate the model with diverse observational datasets. Identifying, developing, and applying these and other new data products to constrain the more realistic representations of physical and ecological processes is likely to be a fruitful avenue for research and model development going forward.

Open-source development of CLM is ongoing at https://github.com/escomp/ctsm. Model users and developers are encouraged to provide feedback, report bugs, and contribute model developments. New model features and parameterizations are in development for future versions of CLM including multiple lines of FATES development, explicit treatment of biomass heat storage (Swenson et al., 2019), a representative hillslope formulation that permits water to flow laterally within a grid cell according to topographic or water table gradients, and a multi-layer canopy parameterization (Bonan et al., 2018) as well as ongoing projects on agriculture (e.g., more realistic crop phenology and allocation (Peng et al.,

2018), tillage (Levis et al., 2014), biofuel crops), water management (e.g., multiple sources of irrigation water, reservoirs), and forestry. As these development projects come to fruition, they will be made available to the CLM research community for use.

Model availability

CLM5.0 is publicly available through the Community Terrestrial System Model (CTSM) git repository (https://github.com/ESCOMP/ctsm).

Author contributions

R.F., S.C.S., C.D.K., K.W.O., and D.M.L. were the lead model developers. D.M.L. designed the model experiment ensemble. K.W.O. completed the model experiments. D.M.L., R.A.F., W.W, D.K., S.C.S., P.J.L., H.L., D.L., C.D.K., D.K. and A.F. drafted the manuscript. R.A.F, C.D.K, K.W.O, S.S., B.G., L.K., D.K., E.K., F.L., H.L., D.L., W.J.R., M.S., M.V., C.X., A.A., A.M.B., G.B., M.A.B., J.B., M.C., A.C., K.D., B.D., J.B.F., M.F., P.G., R.K., J.L., L.R.L., W.H.L., Y.L., J.D.P., J.P., D.M.R., A.S. Z.M.S., J.T., S.B., M.V.M., B.M.S., and X.Z. contributed model developments or datasets for CLM4.5 and/or CLM5. N.C., F.H., J.T.R, G.K., D.M.L., G.B., S.K., A.P., and R.Q.T. developed ILAMB or contributed analyses for the manuscript. All authors participated in revision of the manuscript.

Acknowledgements

We would like to thank the reviewers for their insightful comments and helpful suggestions that improved the clarity and presentation of the manuscript. The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. DML was supported in part by the RUBISCO Scientific Focus Area (SFA), which is sponsored by the Regional and Global Climate Modeling (RGCM) Program in the Climate and Environmental Sciences Division (CESD) of the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science. DK and PG were supported by Columbia University

Presidential Fellowship. GB, DLL, WRW, and RQT were supported by the US Department of Agriculture NIFA Award number 2015-67003-23485. WRW and GKA were supported by the NASA Interdisciplinary Science Program award number NNX17AK19G. JBF and MS carried out the research in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. California Institute of Technology. Government sponsorship acknowledged. All rights reserved. JBF and MS were supported in part by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Terrestrial Ecosystem Science program under Award Numbers DE-SC0008317 and DE-SC0016188; the National Science Foundation Ecosystem Science program (DEB-1153401); and, NASA's CARBON and TE programs. All model data is archived and publicly available at the UCAR/NCAR Climate Data Gateway, https://doi.org/10.5065/d6154fwh.

References

Adole, T., Dash, J., & Atkinson, P. M. (2018). Large-scale prerain vegetation green-up across Africa. *Global Change Biology, 24*(9), 4054-4068.

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: Tansley review. *New Phytologist, 165*(2), 351-372.

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., et al. (2011). Emission factors for open and domestic biomass burning for use in atmospheric models. *Atmospheric Chemistry and Physics, 11*(9), 4039-4072.

Alemohammad, S. H., Fang, B., Konings, A. G., Aires, F., Green, J. K., Kolassa, J., et al. (2017). Water, Energy, and Carbon with Artificial Neural Networks (WECANN): a statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence. *Biogeosciences, 14*(18), 4101-4124. https://doi.org/10.5194/bg-14-4101-2017

Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschleger, S. D., Massoud, E. C., et al. (2016). A global scale mechanistic model of photosynthetic capacity (LUNA V1.0). *Geoscientific Model Development, 9*(2), 587-606. https://doi.org/10.5194/gmd-9-587-2016

Anderegg, W. R. L., Konings, A. G., Trugman, A. T., Yu, K., Bowling, D. R., Gabbitas, R., et al. (2018). Hydraulic diversity of forests regulates ecosystem resilience during drought. *Nature, 561*(7724), 538-541. https://doi.org/10.1038/s41586-018-0539-7

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). THE DATA ASSIMILATION RESEARCH TESTBED A Community Facility. *Bulletin of the American Meteorological Society, 90*(9), 1283-1296. https://doi.org/10.1175/2009bams2618.1

Andreae, M. O., & Merlet, P. (2001). Emission of trace gases and aerosols from biomass burning. *Global Biogeochemical Cycles, 15*(4), 955-966. https://doi.org/10.1029/2000gb001382

Atkin, O. K., Bloomfield, K. J., Reich, P. B., Tjoelker, M. G., Asner, G. P., Bonal, D., et al. (2015). Global variability in leaf respiration in relation to climate, plant functional types and leaf traits. *New Phytol, 206*(2), 614-636. https://doi.org/10.1111/nph.13253

Avitabile, V., Herold, M., Heuvelink, G. B., Lewis, S. L., Phillips, O. L., Asner, G. P., et al. (2016). An integrated pan-tropical biomass map using multiple reference datasets. *Global Change Biology, 22*(4), 1406-1420.

Badger, A. M., & Dirmeyer, P. A. (2015). Climate response to Amazon forest replacement by heterogeneous crop cover. *Hydrology and Earth System Sciences, 19*(11), 4547-4557. https://doi.org/10.5194/hess-19-4547-2015

Bagley, J. E., Miller, J., & Bernacchi, C. J. (2015). Biophysical impacts of climate-smart agriculture in the Midwest United States. *Plant Cell Environ, 38*(9), 1913-1930. https://doi.org/10.1111/pce.12485

Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. *Science, 359*(6375). https://doi.org/10.1126/science.aam8328

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-Biogeosciences, 116*(G2). https://doi.org/10.1029/2010jg001593

Bonan, G. B., & Levis, S. (2010). Quantifying carbon-nitrogen feedbacks in the Community Land Model (CLM4). *Geophysical Research Letters, 37*(7), n/a-n/a. https://doi.org/10.1029/2010gl042430

Bonan, G. B., Oleson, K. W., Fisher, R. A., Lasslop, G., & Reichstein, M. (2012). Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4. *Journal of Geophysical Research-Biogeosciences, 117*(G2). https://doi.org/10.1029/2011jg001913

Bonan, G. B., Patton, E. G., Harman, I. N., Oleson, K. W., Finnigan, J. J., Lu, Y. Q., & Burakowski, E. A. (2018). Modeling canopy-induced turbulence in the Earth system: a unified parameterization of turbulent exchange within plant canopies and the roughness sublayer (CLM-ml v0). *Geoscientific Model Development, 11*(4), 1467-1496. https://doi.org/10.5194/gmd-11-1467-2018

Bonan, G. B., Wieder, W., Lombardozzi, D. L., Lawrence, D. M., Fisher, R. A., & Oleson, K. W. (2019). Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). *Global Biogeochemical Cycles*. https://doi.org/doi.org/10.1029/2019GL083978

Brunke, M. A., Broxton, P., Pelletier, J., Gochis, D., Hazenberg, P., Lawrence, D. M., et al. (2016). Implementing and Evaluating Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5). *Journal of Climate, 29*(9), 3441-3461. https://doi.org/10.1175/Jcli-D-15-0307.1

Brzostek, E. R., Fisher, J. B., & Phillips, R. P. (2014). Modeling the carbon cost of plant nitrogen acquisition: Mycorrhizal trade-offs and multipath resistance uptake improve predictions of retranslocation. *Journal of Geophysical Research: Biogeosciences, 119*(8), 1684-1697.

Burns, S. P., Swenson, S. C., Wieder, W. R., Lawrence, D. M., Bonan, G. B., Knowles, J. F., & Blanken, P. D. (2018). A Comparison of the Diel Cycle of Modeled and Measured Latent Heat Flux During the Warm Season in a Colorado Subalpine Forest. *Journal of Advances in Modeling Earth Systems, 10*(3), 617-651. https://doi.org/10.1002/2017ms001248

Buzan, J. R., Oleson, K., & Huber, M. (2015). Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5. *Geoscientific Model Development, 8*(2), 151-170. https://doi.org/10.5194/gmd-8-151-2015

Campbell, J. E., Berry, J. A., Seibt, U., Smith, S. J., Montzka, S. A., Launois, T., et al. (2017). Large historical growth in global terrestrial gross primary production. *Nature, 544*(7648), 84-87. https://doi.org/10.1038/nature22030

Campioli, M., Vicca, S., Luyssaert, S., Bilcke, J., Ceschia, E., Chapin Iii, F., et al. (2015). Biomass production efficiency controlled by management in temperate and boreal ecosystems. *Nature Geoscience, 8*(11), 843.

Chen, L., Dirmeyer, P. A., Guo, Z. C., & Schultz, N. M. (2018). Pairing FLUXNET sites to validate model representations of land-use/land-cover change. *Hydrology and Earth System Sciences, 22*(1), 111-125. https://doi.org/10.5194/hess-22-111-2018

Cheng, L., Zhang, L., Wang, Y. P., Canadell, J. G., Chiew, F. H. S., Beringer, J., et al. (2017). Recent increases in terrestrial carbon uptake at little cost to the water cycle. *Nature Communications, 8*(1), 110. https://doi.org/10.1038/s41467-017-00114-5

Choat, B., Drayton, W. M., Brodersen, C., Matthews, M. A., Shackel, K. A., Wada, H., & McElrone, A. J. (2010). Measurement of vulnerability to water stress-induced cavitation in grapevine: a comparison of four techniques applied to a long-vesseled species. *Plant Cell Environ, 33*(9), 1502-1512. https://doi.org/10.1111/j.1365-3040.2010.02160.x

Christoffersen, B. O., Gloor, M., Fauset, S., Fyllas, N. M., Galbraith, D. R., Baker, T. R., et al. (2016). Linking hydraulic traits to tropical forest function in a size-structured and trait-driven model (TFS v.1-Hydro). *Geoscientific Model Development, 9*(11), 4227-4255. https://doi.org/10.5194/gmd-9-4227-2016

Clark, M. P., & Kavetski, D. (2010). Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes. *Water Resources Research, 46*(10).

Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., et al. (1999). Global patterns of terrestrial biological nitrogen (N2) fixation in natural ecosystems. *Global Biogeochemical Cycles, 13*(2), 623-645.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., et al. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *Journal of Advances in Modeling Earth Systems, 10*(11), 2731-2754. https://doi.org/10.1029/2018ms001354

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society, 137*(654), 1-28. https://doi.org/10.1002/qj.776

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., & Luke, C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature, 494*(7437), 341-344. https://doi.org/10.1038/nature11882

Dahlin, K., Fisher, R., & Lawrence, P. (2015). Environmental drivers of drought deciduous phenology in the Community Land Model. *Biogeosciences, 12*(16), 5061-5074.

Dahlin, K. M., Del Ponte, D., Setlock, E., & Nagelkirk, R. (2017). Global patterns of drought deciduous phenology in semi-arid and savanna-type ecosystems. *Ecography, 40*(2), 314-323. https://doi.org/10.1111/ecog.02443

Davin, E. L., Seneviratne, S. I., Ciais, P., Olioso, A., & Wang, T. (2014). Preferential cooling of hot extremes from cropland albedo management. *Proc Natl Acad Sci U S A, 111*(27), 9757-9761. https://doi.org/10.1073/pnas.1317323111

De Kauwe, M. G., Kala, J., Lin, Y.-S., Pitman, A. J., Medlyn, B. E., Duursma, R. A., et al. (2015). A test of an optimal stomatal conductance scheme within the CABLE land surface model. *Geoscientific Model Development, 8*(2), 431-452.

De Rooij, G. H. (2010). Comments on "Improving the numerical simulation of soil moisture–based Richards equation for land models with a deep or shallow water table". *Journal of Hydrometeorology, 11*(4), 1044-1050.

Defries, R. S., Hansen, M. C., Townshend, J. R., Janetos, A., & Loveland, T. R. (2000). A new global 1-km dataset of percentage tree cover derived from remote sensing. *Global Change Biology, 6*(2), 247-254.

DeLucia, E. H., Drake, J. E., Thomas, R. B., & GONZALEZ-MELER, M. (2007). Forest carbon use efficiency: is respiration a constant fraction of gross primary production? *Global Change Biology, 13*(6), 1157-1167.

Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., et al. (2016). Confronting weather and climate models with observational data from soil moisture networks over the United States. *J Hydrometeorol, 17*(4), 1049-1067. https://doi.org/10.1175/JHM-D-15-0196.1

Doughty, C. E., Goldsmith, G. R., Raab, N., Girardin, C. A. J., Farfan-Amezquita, F., Huaraca-Huasco, W., et al. (2018). What controls variation in carbon use efficiency among Amazonian tropical forests? *Biotropica, 50*(1), 16-25.

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

Fao/Iiasa/Isric/Isscas/Jrc. (2012). Harmonized world soil database (version 1.2). *FAO, Rome, Italy and IIASA, Laxenburg, Austria*.

Fisher, J. B., Sitch, S., Malhi, Y., Fisher, R. A., Huntingford, C., & Tan, S. Y. (2010). Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation. *Global Biogeochemical Cycles, 24*(1). https://doi.org/10.1029/2009gb003621

Fisher, R., Wieder, W., Sanderson, B., Koven, C., Oleson, K., Xu, C., et al. (2019). Parameteric controls on responses to biogeochemical forcing in the CLM5. *Journal of Advances in Modeling Earth Systems*. https://doi.org/doi.org/10.1029/2019MS001609

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., et al. (2018). Vegetation demographics in Earth System Models: A review of progress and priorities. *Glob Chang Biol, 24*(1), 35-54. https://doi.org/10.1111/gcb.13910

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., et al. (2015). Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geoscientific Model Development, 8*(11), 3593-3619. https://doi.org/10.5194/gmd-8-3593-2015

Flanner, M. G., Zender, C. S., Randerson, J. T., & Rasch, P. J. (2007). Present-day climate forcing and response from black carbon in snow. *Journal of Geophysical Research, 112*(112). https://doi.org/D11202, doi:10.1029/2006JD008003

Foken, T. (2006). 50 years of the Monin-Obukhov similarity theory. *Boundary-Layer Meteorology, 119*(3), 431-447. https://doi.org/10.1007/s10546-006-9048-6

Ford, B., Val Martin, M., Zelasky, S., Fischer, E., Anenberg, S., Heald, C., & Pierce, J. (2018). Future Fire Impacts on Smoke Concentrations, Visibility, and Health in the Contiguous United States. *GeoHealth, 2*(8), 229-247. https://doi.org/10.1029/2018GH000144

Fowler, D., Coyle, M., Skiba, U., Sutton, M. A., Cape, J. N., Reis, S., et al. (2013). The global nitrogen cycle in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1621), 20130164-20130164.

Fox, A. M., Hoar, T. J., Anderson, J. L., Arellano, A. F., Smith, W. K., Litvak, M. E., et al. (2018). Evaluation of a Data Assimilation System for Land Surface Models Using CLM4.5. *Journal of Advances in Modeling Earth Systems*.

Frank, D. C., Poulter, B., Saurer, M., Esper, J., Huntingford, C., Helle, G., et al. (2015). Water-use efficiency and transpiration across European forests during the Anthropocene. *Nature Climate Change, 5*(6), 579-582.

Franks, P. J., Berry, J. A., Lombardozzi, D. L., & Bonan, G. B. (2017). Stomatal Function across Temporal and Spatial Scales: Deep-Time Trends, Land-Atmosphere Coupling and Global Models. *Plant Physiol, 174*(2), 583-602. https://doi.org/10.1104/pp.17.00287

Franks, P. J., Bonan, G. B., Berry, J. A., Lombardozzi, D. L., Holbrook, N. M., Herold, N., & Oleson, K. W. (2018). Comparing optimal and empirical stomatal conductance models for application in Earth system models. *Glob Chang Biol*. https://doi.org/10.1111/gcb.14445

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., et al. (2011). The Community Climate System Model Version 4. *J. Climate*. https://doi.org/doi: 10.1175/2011JCLI4083.1

Ghimire, B., Riley, W. J., Koven, C. D., Mu, M. Q., & Randerson, J. T. (2016). Representing leaf and root physiological traits in CLM improves global carbon and nitrogen cycling predictions. *Journal of Advances in Modeling Earth Systems, 8*(2), 598-613. https://doi.org/10.1002/2015ms000538

Giardina, C. P., Coleman, M. D., Hancock, J. E., King, J. S., Lilleskov, E. A., Loya, W. M., et al. (2005). The response of belowground carbon allocation in forests to global change. In *Tree species effects on soils: implications for global change* (pp. 119-154): Springer.

Giardina, F., Konings, A. G., Kennedy, D., Alemohammad, S. H., Oliveira, R. S., Uriarte, M., & Gentine, P. (2018). Tall Amazonian forests are less sensitive to precipitation variability. *Nature Geoscience, 11*(6), 405-+. https://doi.org/10.1038/s41561-018-0133-5

Goldewijk, K. K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene - HYDE 3.2. *Earth System Science Data, 9*(2), 927-953. https://doi.org/10.5194/essd-9-927-2017

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development, 5*(6), 1471-1492. https://doi.org/10.5194/gmd-5-1471-2012

Handorf, D., Foken, T., & Kottmeier, C. (1999). The stable atmospheric boundary layer over an Antarctic ice sheet. *Boundary-Layer Meteorology, 91*(2), 165-189. https://doi.org/10.1023/A:1001889423449

Herridge, D. F., Peoples, M. B., & Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil, 311*(1-2), 1-18. https://doi.org/10.1007/s11104-008-9668-3

Hoffman, F. M., Randerson, J. T., Arora, V. K., Bao, Q., Cadule, P., Ji, D., et al. (2014). Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models. *Journal of Geophysical Research-Biogeosciences, 119*(2), 141-162. https://doi.org/10.1002/2013jg002381

Houlton, B. Z., Marklein, A. R., & Bai, E. (2015). Representation of nitrogen in climate change forecasts. *Nature Climate Change, 5*(5), 398-401. https://doi.org/10.1038/nclimate2538

Houlton, B. Z., Wang, Y. P., Vitousek, P. M., & Field, C. B. (2008). A unifying framework for dinitrogen fixation in the terrestrial biosphere. *Nature, 454*(7202), 327-330. https://doi.org/10.1038/nature07028

Huang, M., Piao, S., Sun, Y., Ciais, P., Cheng, L., Mao, J., et al. (2015). Change in terrestrial ecosystem water-use efficiency over the last three decades. *Glob Chang Biol, 21*(6), 2366-2378. https://doi.org/10.1111/gcb.12873

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., Kuhry, P., & Swanson, D. K. (2013). The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions. *Earth System Science Data, 5*(1), 3-13. https://doi.org/10.5194/essd-5-3-2013

Hungate, B. A., Stiling, P. D., Dijkstra, P., Johnson, D. W., Ketterer, M. E., Hymus, G. J., et al. (2004). CO2 elicits long-term decline in nitrogen fixation. *Science, 304*(5675), 1291. https://doi.org/10.1126/science.1095549

Huntzinger, D. N., Schwalm, C., Michalak, A. M., Schaefer, K., King, A. W., Wei, Y., et al. (2013). The North American Carbon Program Multi-Scale Synthesis and Terrestrial Model Intercomparison Project - Part 1: Overview and experimental design. *Geoscientific Model Development, 6*(6), 2121-2133. https://doi.org/10.5194/gmd-6-2121-2013

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The Community Earth System Model A Framework for Collaborative Research. *Bulletin of the American Meteorological Society, 94*(9), 1339-1360. https://doi.org/10.1175/Bams-D-12-00121.1

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., & Schulze, E. D. (1996). A global analysis of root distributions for terrestrial biomes. *Oecologia, 108*(3), 389-411. https://doi.org/10.1007/BF00333714

Janssens, I. A., Dieleman, W., Luyssaert, S., Subke, J. A., Reichstein, M., Ceulemans, R., et al. (2010). Reduction of forest soil respiration in response to nitrogen deposition. *Nature Geoscience, 3*(5), 315-322. https://doi.org/10.1038/ngeo844

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature, 496*(7445), 347-350. https://doi.org/10.1038/nature11983

Jones, P., & Harris, I. (2013). CRU TS3. 21: Climatic Research Unit (CRU) Time-Series (TS) version 3.21 of high resolution gridded data of month-by-month variation in climate (Jan. 1901–Dec. 2012). *NCAS British Atmospheric Data Centre*.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research-Biogeosciences, 116*. https://doi.org/10.1029/2010jg001566

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP-NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society, 77*(3), 437-471.

Kattge, J., Diaz, S., Lavorel, S., Prentice, C., Leadley, P., Bonisch, G., et al. (2011). TRY - a global database of plant traits. *Global Change Biology, 17*(9), 2905-2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Kattge, J., Knorr, W., Raddatz, T., & Wirth, C. (2009). Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology, 15*(4), 976-991.

Kavetski, D., Binning, P., & Sloan, S. W. (2001). Adaptive time stepping and error control in a mass conservative numerical solution of the mixed form of Richards equation. *Advances in Water Resources, 24*(6), 595-605. https://doi.org/Doi 10.1016/S0309-1708(00)00076-2

Keenan, T., Garcia, R., Friend, A. D., Zaehle, S., Gracia, C., & Sabate, S. (2009). Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO2 and water fluxes through combined in situ measurements and ecosystem modelling. *Biogeosciences, 6*(8), 1423-1444. https://doi.org/10.5194/bg-6-1423-2009

Keenan, T. F., Hollinger, D. Y., Bohrer, G., Dragoni, D., Munger, J. W., Schmid, H. P., & Richardson, A. D. (2013). Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature, 499*(7458). https://doi.org/10.1038/nature12291

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., da Costa, A. C. L., & Gentine, P. (2019). Implementing Plant Hydraulics in the Community Land Model, Version 5. *Journal of Advances in Modeling Earth Systems, 11*(2), 485-513. https://doi.org/10.1029/2018ms001500

Keppel-Aleks, G., Basile, S. J., & Hoffman, F. M. (2018). A Functional Response Metric for the Temperature Sensitivity of Tropical Ecosystems. *Earth Interactions, 22*(7), 1-20. https://doi.org/10.1175/Ei-D-17-0017.1

Konings, A. G., & Gentine, P. (2017). Global variations in ecosystem-scale isohydricity. *Glob Chang Biol, 23*(2), 891-905. https://doi.org/10.1111/gcb.13389

Konings, A. G., Yu, Y. F., Xu, L., Yang, Y., Schimel, D. S., & Saatchi, S. S. (2017). Active microwave observations of diurnal and seasonal variations of canopy water content across the humid African tropical forests. *Geophysical Research Letters, 44*(5), 2290-2299. https://doi.org/10.1002/2016gl072388

Koven, C. D., Hugelius, G., Lawrence, D. M., & Wieder, W. R. (2017). Higher climatological temperature sensitivity of soil carbon in cold than warm climates. *Nature Climate Change, 7*(11), 817-+. https://doi.org/10.1038/Nclimate3421

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., et al. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences, 10*(11), 7109-7131. https://doi.org/10.5194/bg-10-7109-2013

Kwon, Y., Yang, Z. L., Zhao, L., Hoar, T. J., Toure, A. M., & Rodell, M. (2016). Estimating Snow Water Storage in North America Using CLM4, DART, and Snow Radiance Data Assimilation. *Journal of Hydrometeorology, 17*(11), 2853-2874. https://doi.org/10.1175/Jhm-D-16-0028.1

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., et al. (2010). Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry & Physics Discussions, 10*, 4963-5019. https://doi.org/doi:10.5194/acpd-10-4963-2010

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., et al. (2016). The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale and experimental design. *Geoscientific Model Development, 9*(9), 2973-2998. https://doi.org/10.5194/gmd-9-2973-2016

Lawrence, D. M., Koven, C. D., Swenson, S. C., Riley, W. J., & Slater, A. G. (2015). Permafrost thaw and resulting soil moisture changes regulate projected high-latitude CO2 and CH4 emissions. *Environmental Research Letters, 10*(9), 094011. https://doi.org/Artn 094011
10.1088/1748-9326/10/9/094011

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., et al. (2011). Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst., 3*. https://doi.org/DOI: 10.1029/2011MS000045

Lawrence, D. M., & Slater, A. G. (2008). Incorporating organic soil into a global climate model. *Climate Dynamics, 30*. https://doi.org/doi:10.1007/s00382-007-0278-1

Lawrence, D. M., Slater, A. G., Romanovsky, V. E., & Nicolsky, D. J. (2008). The sensitivity of a model projection of near-surface permafrost degradation to soil column depth and inclusion of soil organic matter. *Journal of Geophysical Research, 113*(113). https://doi.org/F02011, doi:10.1029/2007JF000883

Lawrence, D. M., Thornton, P. E., Oleson, K. W., & Bonan, G. B. (2007). Partitioning of evaporation into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction. *Journal of Hydrometeorology, 8*, 862-880.

Lawrence, P. J., & Chase, T. N. (2007). Representing a new MODIS consistent land surface in the Community Land Model (CLM3.0). *Journal of Geophysical Research, 112*(112). https://doi.org/doi:10.1029/2006JG000168

Lawrence, P. J., Feddema, J. J., Bonan, G. B., Meehl, G. A., O'Neill, B. C., Oleson, K. W., et al. (2012). Simulating the Biogeochemical and Biogeophysical Impacts of Transient Land Cover Change and Wood Harvest in the Community Climate System Model (CCSM4) from 1850 to 2100. *Journal of Climate, 25*(9), 3071-3095. https://doi.org/10.1175/Jcli-D-11-00256.1

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., et al. (2018). Global Carbon Budget 2017. *Earth System Science Data, 10*(1), 405-448.

Le Quéré, C., Peters, G. P., Andres, R. J., Andrew, R. M., Boden, T. A., Ciais, P., et al. (2014). Global carbon budget 2013. *Earth System Science Data, 6*(1), 235-263.

LeBauer, D. S., & Treseder, K. K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology, 89*(2), 371-379. https://doi.org/Doi 10.1890/06-2057.1

Levis, S., Badger, A., Drewniak, B., Nevison, C., & Ren, X. L. (2018). CLMcrop yields and water requirements: avoided impacts by choosing RCP 4.5 over 8.5. *Climatic Change, 146*(3-4), 501-515. https://doi.org/10.1007/s10584-016-1654-9

Levis, S., Hartman, M., & Bonan, G. (2014). The Community Land Model underestimates land-use CO 2 emissions by neglecting soil disturbance from cultivation. *Geoscientific Model Development, 7*(2), 613-620.

Li, F., & Lawrence, D. M. (2017). Role of Fire in the Global Land Water Budget during the Twentieth Century due to Changing Ecosystems. *Journal of Climate, 30*(6), 1893-1908. https://doi.org/10.1175/Jcli-D-16-0460.1

Li, F., Levis, S., & Ward, D. S. (2013a). Quantifying the role of fire in the Earth system - Part 1: Improved global fire modeling in the Community Earth System Model (CESM1). *Biogeosciences, 10*(4), 2293-2314. https://doi.org/10.5194/bg-10-2293-2013

Li, F., Zeng, X., & Levis, S. (2012). A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model. *Biogeosciences, 9*(7), 2761-2780.

Li, H. Y., Wigmosta, M. S., Wu, H., Huang, M. Y., Ke, Y. H., Coleman, A. M., & Leung, L. R. (2013b). A Physically Based Runoff Routing Model for Land Surface and Earth System Models. *Journal of Hydrometeorology, 14*(3), 808-828. https://doi.org/10.1175/Jhm-D-12-015.1

Li, Q., Gentine, P., Mellado, J. P., & McColl, K. A. (2018). Implications of Nonlocal Transport and Conditionally Averaged Statistics on Monin-Obukhov Similarity Theory and Townsend's Attached Eddy Hypothesis. *Journal of the Atmospheric Sciences, 75*(10), 3403-3431. https://doi.org/10.1175/Jas-D-17-0301.1

Li, Y., Guan, K. Y., Gentine, P., Konings, A. G., Meinzer, F. C., Kimball, J. S., et al. (2017). Estimating Global Ecosystem Isohydry/Anisohydry Using Active and Passive Microwave Satellite Data. *Journal of Geophysical Research-Biogeosciences, 122*(12), 3306-3321. https://doi.org/10.1002/2017jg003958

Lin, Y. S., Medlyn, B. E., Duursma, R. A., Prentice, I. C., Wang, H., Baig, S., et al. (2015). Optimal stomatal behaviour around the world. *Nature Climate Change, 5*(5), 459-464. https://doi.org/10.1038/Nclimate2550

Lipscomb, W. H., Fyke, J. G., Vizcaino, M., Sacks, W. J., Wolfe, J., Vertenstein, M., et al. (2013). Implementation and Initial Evaluation of the Glimmer Community Ice Sheet Model in the Community Earth System Model. *Journal of Climate, 26*(19), 7352-7371. https://doi.org/10.1175/Jcli-D-12-00557.1

Lipscomb, W. H., Price, S. F., Hoffman, M. J., Leguy, G. R., Bennett, A. R., Bradley, S. L., et al. (2019). Description and evaluation of the Community Ice Sheet Model (CISM) v2. 1. *Geoscientific Model Development, 12*(1), 387-424. https://doi.org/doi.org/10.5194/gmd-12-387-2019

Liu, L., & Greaver, T. L. (2010). A global perspective on belowground carbon dynamics under nitrogen enrichment. *Ecol Lett, 13*(7), 819-828. https://doi.org/10.1111/j.1461-0248.2010.01482.x

Lombardozzi, D., Levis, S., Bonan, G., Hess, P. G., & Sparks, J. P. (2015a). The Influence of Chronic Ozone Exposure on Global Carbon and Water Cycles. *Journal of Climate, 28*(1), 292-305. https://doi.org/10.1175/Jcli-D-14-00223.1

Lombardozzi, D., Sparks, J. P., & Bonan, G. (2013). Integrating O-3 influences on terrestrial processes: photosynthetic and stomatal response data available for regional and global modeling. *Biogeosciences, 10*(11), 6815-6831. https://doi.org/10.5194/bg-10-6815-2013

Lombardozzi, D. L., Bonan, G. B., Smith, N. G., Dukes, J. S., & Fisher, R. A. (2015b). Temperature acclimation of photosynthesis and respiration: A key uncertainty in the carbon cycle-climate feedback. *Geophysical Research Letters, 42*(20), 8624-8631.

Lombardozzi, D. L., Bonan, G. B., Wieder, W., Grandy, A. S., Morris, C., & Lawrence, D. L. (2018). Cover Crops May Cause Winter Warming in Snow-Covered Regions. *Geophysical Research Letters, 45*(18), 9889-9897.

Luo, Y., Hui, D., & Zhang, D. (2006). ELEVATED CO ₂ STIMULATES NET ACCUMULATIONS OF CARBON AND NITROGEN IN LAND ECOSYSTEMS: A META-ANALYSIS. *Ecology, 87*(1), 53-63.

Malhi, Y., Doughty, C., & Galbraith, D. (2011). The allocation of ecosystem net primary productivity in tropical forests. *Philos Trans R Soc Lond B Biol Sci, 366*(1582), 3225-3245. https://doi.org/10.1098/rstb.2011.0062

Marschner, H. (1995). Mineral Nutrition of Higher Plants Academic Press London Google Scholar.
McGuire, A. D., Lawrence, D. M., Koven, C., Clein, J. S., Burke, E., Chen, G., et al. (2018). Dependence of the evolution of carbon dynamics in the northern permafrost region on the trajectory of climate change. *Proc Natl Acad Sci U S A, 115*(15), 3882-3887. https://doi.org/10.1073/pnas.1719903115
Medlyn, B. E., De Kauwe, M. G., Zaehle, S., Walker, A. P., Duursma, R. A., Luus, K., et al. (2016). Using models to guide field experiments: a priori predictions for the CO 2 response of a nutrient-and water-limited native Eucalypt woodland. *Global Change Biology, 22*(8), 2834-2851.
Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., et al. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology, 17*(6), 2134-2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x
Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., et al. (2015). Using ecosystem experiments to improve vegetation models. *Nature Climate Change, 5*(6), 528-534. https://doi.org/10.1038/nclimate2621
Meier, R., Davin, E. L., Lejeune, Q., Hauser, M., Li, Y., Martens, B., et al. (2018). Evaluating and improving the Community Land Model's sensitivity to land cover. *Biogeosciences, 15*(15), 4731-4757. https://doi.org/10.5194/bg-15-4731-2018
Mitchell, T. D., & Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology, 25*(6), 693-712.
Mueller, N. D., Rhines, A., Butler, E. E., Ray, D. K., Siebert, S., Holbrook, N. M., & Huybers, P. (2017). Global Relationships between Cropland Intensification and Summer Temperature Extremes over the Last 50 Years. *Journal of Climate, 30*(18), 7505-7528. https://doi.org/10.1175/Jcli-D-17-0096.1
Negrón-Juárez, R. I., Koven, C. D., Riley, W. J., Knox, R. G., & Chambers, J. Q. (2015). Observed allocations of productivity and biomass, and turnover times in tropical forests are not accurately represented in CMIP5 Earth system models. *Environmental Research Letters, 10*(6), 064017.
Niinemets, U., & Keenan, T. (2014). Photosynthetic responses to stress in Mediterranean evergreens: Mechanisms and models. *Environmental and Experimental Botany, 103*, 24-41. https://doi.org/10.1016/j.envexpbot.2013.11.008
Niu, G. Y., & Yang, Z. L. (2007). An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. *Journal of Geophysical Research, 112*(112). https://doi.org/D21101, doi:10.1029/2007JD008674
Niu, G. Y., Yang, Z. L., Dickinson, R. E., Gulden, L. E., & Su, H. (2007). Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. *Journal of Geophysical Research,* (112). Retrieved from doi:D07103, doi:10.1029/2006JD007522
Oleson, K., & Feddema, J. (2019). Parameterization and surface data improvements and new capabilities for the Community Land Model Urban (CLMU). *Submitted to JAMES*.
Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M., & Grimmond, C. S. B. (2008a). An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. *Journal of Applied Meteorology and Climatology, 47*(4), 1038-1060. https://doi.org/DOI: 10.1175/2007JAMC1597.1
Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., et al. (2013). *Technical Description of version 4.5 of the Community Land Model (CLM)* (NCAR Tech. Note NCAR/TN-503+STR). Retrieved from Boulder, CO:
Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., et al. (2010). *Technical description of version 4.0 of the Community Land Model* (NCAR Tech. Note NCAR/TN-478+STR). Retrieved from Boulder, CO:

Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., et al. (2008b). Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research-Biogeosciences, 113*(G1). https://doi.org/doi:10.1029/2007JG000563

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X. B., Troch, P. A., Niu, G. Y., et al. (2016). A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *Journal of Advances in Modeling Earth Systems, 8*(1), 41-65. https://doi.org/10.1002/2015ms000526

Peng, B., Guan, K. Y., Chen, M., Lawrence, D. M., Pokhrel, Y., Suyker, A., et al. (2018). Improving maize growth processes in the community land model: Implementation and evaluation. *Agricultural and Forest Meteorology, 250*, 64-89. https://doi.org/10.1016/j.agrformet.2017.11.012

Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles, 24*(1).

Powell, T. L., Galbraith, D. R., Christoffersen, B. O., Harper, A., Imbuzeiro, H. M., Rowland, L., et al. (2013). Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought. *New Phytol, 200*(2), 350-365. https://doi.org/10.1111/nph.12390

Quiring, S. M., Ford, T. W., Wang, J. K., Khong, A., Harris, E., Lindgren, T., et al. (2016). THE NORTH AMERICAN SOIL MOISTURE DATABASE Development and Applications. *Bulletin of the American Meteorological Society, 97*(8), 1441-+. https://doi.org/10.1175/Bams-D-13-00263.1

Reich, P. B., Hungate, B. A., & Luo, Y. Q. (2006). Carbon-nitrogen interactions in terrestrial ecosystems in response to rising atmospheric carbon dioxide. *Annual Review of Ecology Evolution and Systematics, 37*(1), 611-636. https://doi.org/10.1146/annurev.ecolsys.37.091305.110039

Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., et al. (2011). Barriers to Predicting Global Terrestrial Methane Fluxes: Analyses Using a Methane Biogeochemistry Model Integrated in CESM. *Accepted to JGR-Biogeosciences*.

Roesch, A., Wild, M., Gilgen, H., & Ohmura, A. (2001). A new snow cover fraction parametrization for the ECHAM4 GCM. *Climate Dynamics, 17*(12), 933-946. https://doi.org/10.1007/s003820100153

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., von Caemmerer, S., Dietze, M. C., et al. (2017). A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytol, 213*(1), 22-42. https://doi.org/10.1111/nph.14283

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S., & Helkowski, J. H. (2009). Effects of global irrigation on the near-surface climate. *Climate Dynamics, 33*(2-3), 159-175. https://doi.org/DOI 10.1007/s00382-008-0445-z

Sakaguchi, K., & Zeng, X. (2009). Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3. 5). *Journal of Geophysical Research: Atmospheres, 114*(D1).

Santoro, M., Beaudoin, A., Beer, C., Cartus, O., Fransson, J. B. S., Hall, R. J., et al. (2015). Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR. *Remote Sensing of Environment, 168*, 316-334. https://doi.org/10.1016/j.rse.2015.07.005

Schlesinger, W. H., & Jasechko, S. (2014). Transpiration in the global water cycle. *Agricultural and Forest Meteorology, 189*, 115-117. https://doi.org/10.1016/j.agrformet.2014.01.011

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology, 115*(1-2), 15-40. https://doi.org/10.1007/s00704-013-0860-x

Schultz, N. M., Lee, X., Lawrence, P. J., Lawrence, D. M., & Zhao, L. (2016). Assessing the use of subgrid land model output to study impacts of land cover change. *Journal of Geophysical Research-Atmospheres, 121*(11), 6133-6147. https://doi.org/10.1002/2016jd025094

Schuur, E. A., McGuire, A. D., Schadel, C., Grosse, G., Harden, J. W., Hayes, D. J., et al. (2015). Climate change and the permafrost carbon feedback. *Nature, 520*(7546), 171-179. https://doi.org/10.1038/nature14338

Shi, M., Fisher, J. B., Brzostek, E. R., & Phillips, R. P. (2016). Carbon cost of plant nitrogen acquisition: global carbon cycle impact from an improved plant nitrogen cycle in the Community Land Model. *Glob Chang Biol, 22*(3), 1299-1314. https://doi.org/10.1111/gcb.13131

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research-Biogeosciences, 116*(G4). https://doi.org/10.1029/2011jg001708

Slater, A. G., Lawrence, D. M., & Koven, C. D. (2017). Process-level model evaluation: a snow and heat transfer metric. *Cryosphere, 11*(2), 989-996. https://doi.org/10.5194/tc-11-989-2017

Storck, P., Lettenmaier, D. P., & Bolton, S. M. (2002). Measurement of snow interception and canopy effects on snow accumulation and melt in a mountainous maritime climate, Oregon, United States: FOREST CANOPY EFFECTS ON GROUND SNOW PACK DYNAMICS. *Water Resources Research, 38*(11), 5-1-5-16.

Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., & Lea, J. (2010). Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes. *Journal of Hydrometeorology, 11*(6), 1380-1394. https://doi.org/10.1175/2010jhm1202.1

Subin, Z. M., Riley, W. J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *Journal of Advances in Modeling Earth Systems, 4*(1). https://doi.org/10.1029/2011ms000072

Swenson, S., & Lawrence, D. (2014). Assessing a dry surface layer-based soil resistance parameterization for the Community Land Model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research: Atmospheres, 119*(17).

Swenson, S., & Lawrence, D. (2015). A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model. *Water Resources Research, 51*(11), 8817-8833.

Swenson, S. C., Burns, S. P., & Lawrence, D. M. (2019). The Impact of Biomass Heat Storage on the Canopy Energy Balance and Atmospheric Stability in the Community Land Model. *Journal of Advances in Modeling Earth Systems, 11*(1), 83-98. https://doi.org/10.1029/2018ms001476

Swenson, S. C., & Lawrence, D. (2012). A new fractional snow-covered area parameterization for the Community Land Model and its effect on the surface energy balance. *Journal of Geophysical Research: Atmospheres, 117*(D21).

Swenson, S. C., Lawrence, D. M., & Lee, H. (2012). Improved simulation of the terrestrial hydrological cycle in permafrost regions by the Community Land Model. *Journal of Advances in Modeling Earth Systems, 4*(3). https://doi.org/10.1029/2012ms000165

Terrer, C., Vicca, S., Stocker, B. D., Hungate, B. A., Phillips, R. P., Reich, P. B., et al. (2018). Ecosystem responses to elevated CO ₂ governed by plant-soil interactions and the cost of nitrogen acquisition. *New Phytologist, 217*(2), 507-522.

Teuling, A. J., Seneviratne, S. I., Williams, C., & Troch, P. A. (2006). Observed timescales of evapotranspiration response to soil moisture. *Geophysical Research Letters, 33*(23). https://doi.org/10.1029/2006gl028178

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

Thomas, R. Q., Bonan, G. B., & Goodale, C. L. (2013a). Insights into mechanisms governing forest carbon response to nitrogen deposition: a model-data comparison using observed responses to nitrogen addition. *Biogeosciences, 10*(6), 3869-3887. https://doi.org/10.5194/bg-10-3869-2013

Thomas, R. Q., Zaehle, S., Templer, P. H., & Goodale, C. L. (2013b). Global patterns of nitrogen limitation: confronting two global biogeochemical models with observations. *Glob Chang Biol, 19*(10), 2986-2998. https://doi.org/10.1111/gcb.12281

Thonicke, K., Spessa, A., Prentice, I., Harrison, S. P., Dong, L., & Carmona-Moreno, C. (2010). The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., & Mahowald, N. M. (2007). Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variability. *Global Biogeochemical Cycles, 21*(21). https://doi.org/GB4018, doi:10.1029/2006GB002868

Thornton, P. E., & Rosenbloom, N. A. (2005). Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecological Modelling, 189*(1-2), 25-48.

Tricht, K. V., Lhermitte, S., Gorodetskaya, I. V., & van Lipzig, N. P. (2016). Improving satellite-retrieved surface radiative fluxes in polar regions using a smart sampling approach. *The Cryosphere, 10*(5), 2379-2397.

Trugman, A. T., Medvigy, D., Mankin, J. S., & Anderegg, W. R. L. (2018). Soil Moisture Stress as a Major Driver of Carbon Cycle Uncertainty. *Geophysical Research Letters, 45*(13), 6495-6503. https://doi.org/10.1029/2018gl078131

Val Martin, M., Kahn, R., & Tosca, M. (2018). A Global Analysis of Wildfire Smoke Injection Heights Derived from Space-Based Multi-Angle Imaging. *Remote Sensing, 10*(10), 1609.

Val Martin, M., Logan, J. A., Kahn, R. A., Leung, F.-Y., Nelson, D. L., & Diner, D. J. (2010). Smoke injection heights from fires in North America: analysis of 5 years of satellite observations. *Atmospheric Chemistry and Physics, 10*(4), 1491-1510.

Van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., et al. (2016). LS3MIP (v1. 0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project-aims, setup and expected outcome. *Geoscientific Model Development, 9*, 2809-2832.

van Kampenhout, L., Lenaerts, J., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., & van den Broeke, M. R. (2017). Improving the representation of polar snow and firn in the Community Earth System Model. *Journal of Advances in Modeling Earth Systems, 9*(7), 2583-2600.

Vicca, S., Luyssaert, S., Peñuelas, J., Campioli, M., Chapin, F. S., Ciais, P., et al. (2012). Fertile forests produce biomass more efficiently: Forests' biomass production efficiency. *Ecology Letters, 15*(6), 520-526.

Vitousek, P. M., Menge, D. N., Reed, S. C., & Cleveland, C. C. (2013). Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philos Trans R Soc Lond B Biol Sci, 368*(1621), 20130119. https://doi.org/10.1098/rstb.2013.0119

Vogel, R. M., & Kroll, C. N. (1996). Estimation of Baseflow Recession Constants. *Water Resources Management, 10*(4), 303-320. https://doi.org/Doi 10.1007/Bf00508898

Walker, A. P., Hanson, P. J., De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Asao, S., et al. (2014). Comprehensive ecosystem model-data synthesis using multiple data sets at two temperate forest free-air CO ₂ enrichment experiments: Model performance at ambient CO ₂ concentration: FACE MODEL-DATA SYNTHESIS. *Journal of Geophysical Research: Biogeosciences, 119*(5), 937-964.

Walker, A. P., Ye, M., Lu, D., De Kauwe, M. G., Gu, L. H., Medlyn, B. E., et al. (2018). The multi-assumption architecture and testbed (MAAT v1.0): R code for generating ensembles with dynamic model structure and analysis of epistemic uncertainty from multiple sources. *Geoscientific Model Development, 11*(8), 3159-3185. https://doi.org/10.5194/gmd-11-3159-2018

Wang, A. H., & Zeng, X. (2009). Improving the Treatment of the Vertical Snow Burial Fraction over Short Vegetation in the NCAR CLM3. *Advances in Atmospheric Sciences, 26*(5), 877-886.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research, 50*(9), 7505-7514.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Osterle, H., et al. (2011). Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *Journal of Hydrometeorology, 12*(5), 823-848. https://doi.org/10.1175/2011jhm1369.1

Wieder, W. R., Cleveland, C. C., Lawrence, D. M., & Bonan, G. B. (2015). Effects of model structural uncertainty on carbon cycle projections: biological nitrogen fixation as a case study. *Environmental Research Letters, 10*(4), 044016. https://doi.org/10.1088/1748-9326/10/4/044016

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., et al. (2019). Beyond static benchmarking: Using experimental manipulations to evaluate land model assumptions. *Global Biogeochemical Cycles*. https://doi.org/doi.org/10.1029/2018GB006141

Xu, C., Fisher, R., Wullschleger, S. D., Wilson, C. J., Cai, M., & McDowell, N. G. (2012). Toward a mechanistic modeling of nitrogen limitation on vegetation dynamics. *PLoS One, 7*(5), e37914. https://doi.org/10.1371/journal.pone.0037914

Yoshimura, K., & Kanamitsu, M. (2008). Dynamical global downscaling of global reanalysis. *Monthly Weather Review, 136*(8), 2983-2998. https://doi.org/10.1175/2008mwr2281.1

Zaehle, S., & Friend, A. D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates: SITE-SCALE EVALUATION OF A C-N MODEL. *Global Biogeochemical Cycles, 24*(1), n/a-n/a.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., et al. (2014). Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO ₂ Enrichment studies. *New Phytologist, 202*(3), 803-822.

Zeng, X. B. (2001). Global vegetation root distribution for land modeling. *Journal of Hydrometeorology, 2*(5), 525-530. https://doi.org/10.1175/1525-7541(2001)002<0525:Gvrdfl>2.0.Co;2

Zeng, X. B., & Decker, M. (2009). Improving the Numerical Solution of Soil Moisture-Based Richards Equation for Land Models with a Deep or Shallow Water Table. *Journal of Hydrometeorology, 10*(1), 308-319.

Zhang, Y. F., Hoar, T. J., Yang, Z. L., Anderson, J. L., Toure, A. M., & Rodell, M. (2014). Assimilation of MODIS snow cover through the Data Assimilation Research Testbed and the Community Land Model version 4. *Journal of Geophysical Research-Atmospheres, 119*(12), 7091-7103. https://doi.org/10.1002/2013jd021329

Zhao, L., & Yang, Z. L. (2018). Multi-sensor land data assimilation: Toward a robust global soil moisture and snow estimation. *Remote Sensing of Environment, 216*, 13-27. https://doi.org/10.1016/j.rse.2018.06.033

Zhao, L., Yang, Z. L., & Hoar, T. J. (2016). Global Soil Moisture Estimation by Assimilating AMSR-E Brightness Temperatures in a Coupled CLM4-RTM-DART System. *Journal of Hydrometeorology, 17*(9), 2431-2454. https://doi.org/10.1175/Jhm-D-15-0218.1

Zhao, M. S., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment, 95*(2), 164-176. https://doi.org/10.1016/j.rse.2004.12.011

Zhu, Q., Riley, W. J., Tang, J., & Koven, C. D. (2016). Multiple soil nutrient competition between plants, microbes, and mineral surfaces: model development, parameterization, and example applications in several tropical forests. *Biogeosciences, 13*(1), 341-363. https://doi.org/10.5194/bg-13-341-2016

Lawrence David M. (Orcid ID: 0000-0002-2968-3023)
Fisher Rosie A. (Orcid ID: 0000-0003-3260-9227)
Koven Charles D. (Orcid ID: 0000-0002-3367-0065)
Oleson Keith W. (Orcid ID: 0000-0002-0057-9900)
Swenson Sean C. (Orcid ID: 0000-0002-2923-1203)
Bonan Gordon (Orcid ID: 0000-0003-0830-6437)
Collier Nathan (Orcid ID: 0000-0002-7367-3981)
van Kampenhout Leo (Orcid ID: 0000-0001-9047-5031)
Kennedy Daniel (Orcid ID: 0000-0001-9494-3509)
Lawrence Peter J. (Orcid ID: 0000-0002-4843-4903)
Li Fang (Orcid ID: 0000-0002-3686-2257)
Lombardozzi Danica (Orcid ID: 0000-0003-3557-7929)
Riley William J. (Orcid ID: 0000-0002-4615-2304)
Shi Mingjie (Orcid ID: 0000-0002-2469-4831)
Vertenstein Mariana (Orcid ID: 0000-0001-9486-9994)
Wieder William R. (Orcid ID: 0000-0001-7116-1985)
Ali Ashehad A. (Orcid ID: 0000-0002-7328-306X)
Badger Andrew, M. (Orcid ID: 0000-0003-4537-9993)
Bisht Gautam (Orcid ID: 0000-0001-6641-7595)
van den Broeke Michiel (Orcid ID: 0000-0003-4662-7565)
Brunke Michael A. (Orcid ID: 0000-0002-2677-7446)
Buzan Jonathan (Orcid ID: 0000-0002-8199-3264)
Clark Martyn (Orcid ID: 0000-0002-2186-2625)
Drewniak Beth (Orcid ID: 0000-0002-8193-6091)
Fisher Joshua B. (Orcid ID: 0000-0003-4734-9085)
Flanner Mark (Orcid ID: 0000-0003-4012-174X)
Fox Andrew M. (Orcid ID: 0000-0002-4027-4538)
Gentine Pierre (Orcid ID: 0000-0002-0845-8345)
Hoffman Forrest (Orcid ID: 0000-0001-5802-4134)
Keppel-Aleks Gretchen (Orcid ID: 0000-0003-2213-0044)
Kumar Sanjiv (Orcid ID: 0000-0002-0472-6074)
Lenaerts Jan (Orcid ID: 0000-0003-4309-4011)
Leung L. Ruby (Orcid ID: 0000-0002-3221-9467)

Table 1. List of changes in CLM4.5 and CLM5 (relative to previous versions of the model).

	CLM4.5	**CLM5**
Hydrology	● Ice impedance function; hydraulic properties for frozen soil determined by liquid water only; perched water table above icy permafrost soil ● Surface water store replaces wetland landunit	● Dry surface layer for ground evap ● Spatially variable soil depth (0.4 to 8.5m) ● Increased soil vertical resolution (20 soil layers + 5 bedrock layers) ● Remove unconfined aquifer, no flux lower boundary condition ● adaptive time-stepping solution of Richard's equation
Snow	● SCF parameterization accounting for hysteresis in accumulation and ablation ● Surface energy fluxes calculated separately for snow-covered and snow-free portions of each landunit	● Separate liquid and ice canopy water stores and radiative treatment, snow unloading due to T or wind ● Wind and T effects on fresh snow density ● Increase maximum snow layers from 5 to 12 and max SWE to 10m to allow for firn development
Glaciers	● Updated bare ice albedos	● Multiple elevation classes (10) with

		downscaling for glacier landunit
Rivers	• Variable flow velocity based on mean grid cell slope in RTM	• MOSART – Manning's equation-based model representing hillslope to tributary to main channel flow
Vegetation	• Revised canopy radiation scheme and canopy scaling of leaf processes • Co-limitation on photosynthesis among Rubisco-, light-, and export-limited rates; revised photosynthetic parameters for Rubisco kinetics and temperature response • BVOC emissions from MEGAN2.1	• Plant hydraulic stress model of water transport through vegetation, replaces empirical soil moisture stress formulation, hydraulic redistribution • Medlyn stomatal conductance replacing Ball-Berry • Consistent Jackson et al. (1996) rooting profile for water and carbon • Deepened rooting profile for Broadleaf Evergreen and Broadleaf Deciduous Tropical trees • Add antecedent rain requirement to stress deciduous phenology trigger
Carbon	• Vertically-resolved soil C and N • Vertical mixing of carbon due to bioturbation, cryoturbation, and diffusion • Litter and soil carbon and nitrogen pool structure based on Century model • 13C and 14C carbon isotopes • CH_4 production and emissions from natural wetlands	• Fixed allocation scheme replaces dynamic NPP-based allocation scheme • Weaker intrinsic decrease in decomposition rate with depth • Stronger soil moisture controls on decomposition • 13C and 14C carbon isotopes for crops
Nitrogen	• Nitrification and denitrification based on Century model • Biological fixation distributed more realistically over year	• Flexible plant C:N ratios, eliminate instantaneous downregulation of photosynthesis based on mineral N availability • Carbon cost for N-uptake and fixation (FUN) • Leaf N optimization for photosynthetic capacity • Free living fixation function of ET
Agriculture	• Temperate crop model (optional, not active by default) • Introduction of organ pools, crop yield	• Global crop model on in BGC default configuration with 8 temperate and tropical crop types • Updated irrigation trigger based on soil water deficit • Industrial fertilization from land use file; background manure • 1-yr grain product pool
LULCC		• Wood harvest by mass rather than area • Land unit weights can change during simulation

		• Updated 1850 and CMIP6 LUMIP transient PFT-distribution
Fire	• New fire model with anthropogenic triggers and suppression; agricultural, deforestation, and peat fires	• Modified dependence of fire occurrence and spread on fuel wetness
Lakes	• Updated lake model including deep and shallow lakes	
Urban	• Multiple urban density classes	• Building space heating and air conditioning • Heat stress indices

Table 2. Simulations performed for CLM analysis.

	CLM4				CLM4.5				CLM5			
Forcing	SP	BGC	+CO_2, +N	no LUC	SP	BGC	+CO_2, +N	no LUC	SP	BGC	+CO_2, +N	no LUC
GSWP3v1	✓o	✓o★	✓	✓	✓o★	✓	✓	✓	✓o	✓o★	✓	✓
CRUNCEPv7		✓				✓			✓	✓★		✓
WATCH/ WFDEI									✓ WF	✓ W		

Note: ✓ indicates historical simulation (1850-2014, ^{W}1850-2001, ^{WF}1979-2014). ★ indicates projection period simulation (RCP8.5, 2015-2300). o indicates daily and hourly output for selected years. Abbreviations: SP – Satellite Phenology, BGC – biogeochemistry, no LUC – land-use and land-cover change turned off, +N – nitrogen addition, +CO_2 – CO_2 addition. Except where explicitly noted, all simulations include all historical or future forcings as applicable to the version of the model (CO_2, N-deposition, aerosol deposition, and LUC time series). Note that CLM5 is the only CLM version with capability to run with the dynamic crop model. See Section 3.1 for further details on the experimental design. Except for CLM4SP, CLM4.5SP, and CLM5SP, all simulations are with active BGC.

Table 3: PFT-level quantities.

	Area	LAI RMSE[a] (m^2 m^{-2})			Survival (%)			V_{cmax25} (µmol CO_2 m^{-2} s^{-1})			
	(10^6 km^2)	CLM4	CLM4.5	CLM5	CLM4	CLM4.5	CLM5	Obs[b]	CLM4	CLM4.5	CLM5
NET temperate	4.9	1.9	2.3	2.0	93	91	87	63	61	63	42
NET boreal	6.0	1.7	3.6	2.7	91	91	92	63	54	63	55
NDT boreal	2.1	1.0	1.3	1.1	67	10	67	39	57	39	52
BET tropical	14.3	1.9	1.5	1.1	92	76	92	29	72	55	38
BET temperate	1.3	1.4	1.4	3.8	88	74	99	62	72	62	39
BDT tropical	2.0	2.4	2.2	2.5	63	24	78	29	52	41	35
BDT temperate	3.0	1.9	2.1	1.1	85	82	81	58	52	58	50
BDT boreal	1.0	1.5	1.6	1.0	92	88	78	58	52	58	50
BES temperate	0.03	3.2	2.3	1.7	83	68a	67	62	72	62	43
BDS temperate	4.6	0.9	0.6	0.6	26	18	30	54	52	54	44
BDS boreal	4.9	1.1	1.3	1.8	35	56	88	54	52	54	58
C_3 Arctic grass	8.7	1.4	2.7	2.5	61	78	83		52	78	52
C_3 grass	16.0	1.6	1.6	1.8	60	78	85	78	52	78	36
C_4 grass	19.9	2.1	2.4	1.2	99	98	90		52	52	50

Note: PFT area, LAI RMSE and V_{cmax25} are calculated for year 2010. Survival percentage is for year 1850. For V_{cmax25}, data is only included for plants that survive (i.e., annual maximum LAI > 0.01 m^2 m^{-2}). For V_{cmax25}, CLM4 and CLM4.5 parameter values are prescribed. V_{cmax25} is a prognostic quantity in CLM5. All data shown is from GSWP3v1-forced BGC simulations. Note that survival rates for CRUNCEPv7-forced simulations are generally within 5% of those with GSWP3v1. PFT abbreviations: N - needleleaf; B - broadleaf; E- evergreen; D - deciduous; T - tree; S - shrub.

[a] LAI observations are from MODIS.

[b] V_{cmax25} observational estimates are from Kattge et al. (2009). Note that TRY database reports two values for Tropical trees, 29 for trees on oxisol soil, 41 for trees on nonoxisol soils.

Table 4: Summary of water cycle variables.

	P ($mm\ d^{-1}$)	ET ($mm\ d^{-1}$)	Total Runoff ($mm\ d^{-1}$)	Transpiration (%)	Soil Evaporation (%)	Canopy Evaporation (%)
Obs	2.22	1.17	0.83	60		
CLM4SP	2.22	1.27	0.86	53	32	14
CLM4.5SP	2.22	1.24	0.84	53	34	13
CLM5SP	2.22	1.20	0.89	60	26	14
CLM4BGC	2.22(2.08)	1.34(1.42)	0.77(0.52)	60	25	15
CLM4.5BGC	2.22(2.08)	1.25(1.33)	0.81(0.58)	56 (53)	29 (26)	15 (20)
CLM5BGC	2.22(2.08)	1.17(1.30)	0.91(0.58)	54	31	15

Note: Sum of evapotranspiration components may not add up to 100% due to rounding. Values in parentheses are from CRUNCEPv7 forced simulations. All other values are from GSWP3v1 forced simulations. Note that observed values are for the area of intersection where both observations and model are land as reported in ILAMB (Collier et al., 2018). Observations are GPCC for precipitation, GLEAM for ET, Dai and Trenberth (2007) for runoff, and WECANN.

Table 5: Summary of global C fluxes and stocks.

	GPP	NPP	CUE	AR	HR	ER	Ecosys C	Veg C	Soil C 1m
Obs	(118)[a]					(94)[a]		(449)[b]	(1320)[c]
CLM4BGC	130 (122)	45	0.40	85	41	126 (119)	1043	469 (453)	500 (493)
CLM4.5BGC	113 (137)	47	0.48	66	43	107 (100)	2379	450 (432)	879 (862)
CLM5BGC	125 (119)	49	0.43	70	42	112 (103)	2574	492 (460)	1057 (1040)

Note: Fluxes include gross and net primary productivity (GPP and NPP), carbon use efficiency (CUE), autotrophic, heterotrophic, and ecosystem respiration (AR, HR and ER). Units are Pg C yr^{-1} for C fluxes, Pg C for C stocks, and a unitless ratio for CUE. Data are averages for the period 1995-2014 from the GSWP3v1-forced simulations. Values in parentheses indicate ILAMB values for the area of intersection between observations and model.

[a] GPP and ER observed estimates from FLUXNET-MTE upscaling (Jung et al., 2011).

[b] Forest vegetation C stock observed estimate from GEOCARBON project (Avitabile et al., 2016; Santoro et al., 2015).

[c] Soil C to 1 m depth observed estimate from Harmonized World Soil Database (Fao/Iiasa/Isric/Isscas/Jrc, 2012).

Table 6: Summary of nitrogen fluxes.

N Flux	**Obs estimate**	**CLM4BGC**	**CLM4.5BGC**	**CLM5BGC**
N Inputs				
N deposition	70[a]	63.1	63.1	63.1
Symbiotic N fix	58[b]*			57.9
Free living N fix		107.7	96.6	38.5
Soy N fix	60[c]			6.0[e]
Crop N fertilization	120[a]			106.7
N Losses				
Denitrification	100[a]	117.8	51.2	100.1
Hydrologic N losses	80[a]	0.01	10.2	33.2
LULCC N losses		5.6	2.6	19.2
Wood harvest		1.4	1.2	2.6
Crop harvest				32.7
Fire losses	14.1[d]	21.8	43.3	32.0
N_2O	13[a]		2.7	6.3

Note: Units are TgN yr^{-1}. Data are averages for the period 1995-2014 from the GSWP3v1-forced simulations.

[a] data from synthesis by Fowler et al. 2013.

[b] Preindustrial estimate (Vitousek et al., 2013), range 40-100 TgN yr^{-1}. *Note, this includes both free-living and symbiotic N fixation.

[c] Globally upscaled estimate for agricultural N fixation, range 50-70 TgN yr^{-1} (Herridge et al., 2008).

[d] Global biomass burning estimates of NO_x and NH_3, converted to Tg N yr^{-1} (Lamarque et al., 2010).

[e] Soy N fix is also included in global estimate of symbiotic N fixation listed above in the table. Data here are from CFT output for nitrogen fixation.

Figure 1: Schematic representation of primary processes and functionality in CLM5. Abbreviations: SCF – snow cover fraction; BVOC – biogenic volatile organic compounds; C/N – carbon and nitrogen. For Biogeochemical Cycles, black arrow denotes carbon flux, purple arrow denotes nitrogen flux. Note that not all soil levels are shown. Not all processes are depicted. Optional features that are not active in default configurations are italicized. Updated from Figure 1 in Lawrence et al. (2011).

Figure 2. Standard configuration of the CLM5 subgrid hierarchy. Box in upper right shows hypothetical subgrid distribution for a single grid cell. Note that the Crop land unit is only used when the model is run with the crop model active. Abbreviations: TBD – Tall Building District; HD – High Density; MD – Medium Density, G – Glacier, L – Lake, U – Urban, C – Crop, V – Vegetated, PFT – Plant Functional Type, Irr – Irrigated, Rnfd – Rainfed. Red arrows indicate allowed land unit transitions. Purple arrows indicate allowed patch-level transitions.

Figure 3: ILAMB summary diagram for CLM4BGC, CLM4.5BGC, and CLM5BGC for GSWP3v1 and CRUNCEPv7 forcing. A version of this figure designed for colorblind readers is available within the ILAMB results.

Figure 4: ILAMB variable-to-variable comparison summary diagram for CLM4BGC, CLM4.5BGC, CLM5BGC for GSWP3v1 forcing. See Collier et al. (2018) for details on this metric. Right panels show example ILAMB relationship plot for a particular variable-to-variable comparison between climatological annual precipitation and LAI. Black line, repeated in each plot, is the observationally-derived relationship. Error bars indicate the +- 1 standard deviation of LAI for all grid cells that lie within that precipitation bin. Values in parentheses indicate ILAMB score for that comparison.

Figure 5: ILAMB summary diagram for CLM5SP (prescribed vegetation, left) and CLM5BGC (prognostic vegetation and carbon cycle, right) forced with three alternative forcing datasets (GSWP3v1, CRUNCEPv7, WFDEI/WATCH). Note that the CLM5BGC WATCH-forced runs only run through year 2001 which means that CLM5BGC-WATCH runs are evaluated over different set of observational years. Grey color for CLM5-WATCH for Terrestrial Water Storage is because there are not enough years of overlap between observations and model. Note that a different set of forcing datasets is used for SP versus BGC simulations (WFDEI for SP and WATCH for BGC) which affects the relative scores even for forcing variables such as precipitation which is the same for CRUNCEPv7 and GSWP3v1 in SP versus BGC.

Figure 6: Maps of biases in annual monthly maximum LAI ($m^3\ m^{-3}$) for CLM4BGC, CLM4.5BGC, and CLM5BGC with GSWP3 forcing for two PFTs, temperate needleleaf evergreen trees and C4 grasses. Weighted area-average RMSE is shown in upper left of each plot. Data is shown only where the individual PFT fraction for a particular grid cell is > 0.

Figure 7: Plots of CLM4SP, CLM4.5SP, and CLM5SP (GSWP3) evapotranspiration (upper panels, mm $month^{-1}$) and soil moisture (lower panels, fraction of saturation) with depth for two years for two example grid cells in (a) southwest US (248°E, 40°N; upper plots) and (b) southeast US (278.5°E, 36°N; lower plots). Observations are from GBAF latent heat flux converted to evapotranspiration (mm $month^{-1}$).

RMSD between model and observations shown for ET comparison. Black lines in soil moisture plots indicate water table position. Grey shading in CLM5 soil moisture plots indicates hydrologically inactive bedrock; white area in CLM4 and CLM4.5 indicates the unconfined aquifer.

Figure 8: (a) Soil moisture residence time (SMRT, May to October) as simulated by CLM5SP (contours) and as derived from soil moisture timeseries at 928 North American Soil Moisture Database (NASMD) stations (filled circles, note that ARM and OK-Mesonet sites are not included because they exhibit large high soil moisture residence time biases (Dirmeyer et al., 2016). (b) Percentage change in SMRT from CLM4.5SP to CLM5SP: ($CLM5SP_{SMRT}$ - $CLM4.5SP_{SMRT}$)*100/$CLM5_{SMRT}$. (c) Change in SMRT from CLM4SP to CLM5SP.

Figure 9: Impact of PHS on variability of GPP and ET. Coefficient of variation (CV, unitless) of GPP and ET with two CLM5SP model configurations: PHS and SMS. CV data are derived from monthly model output over fifty years (1964-2013), after removing the trend and seasonal cycle. GPP analyses exclude grid cells with average GPP less than 0.5 umol m^{-2} s^{-1}. ET analyses exclude grid cells with average ET less than 5 W m^{-2}. Bin widths for all bar plots are 0.05 (CV, unitless), with each bar representing the percent of global average GPP or ET where CV falls within the corresponding range.

Figure 10: Annual transpiration fraction (transpiration / total ET) as estimated with WECANN dataset (upper left) and the difference between model and WECANN for CLM4, CLM4.5, and CLM5 (GSWP3) for SP (left) and BGC (right) configurations.

Figure 11: Permafrost maximum active layer thickness (top panels) and February snow density (lower panels), averaged over years 1990-2010 for CLM4.5BGC and CLM5BGC with GSWP3v1 and CRUNCEPv7 forcing. Total permafrost area in millions of km^2 shown under title bar. CLM4 results (not shown) are similar to CLM4.5.

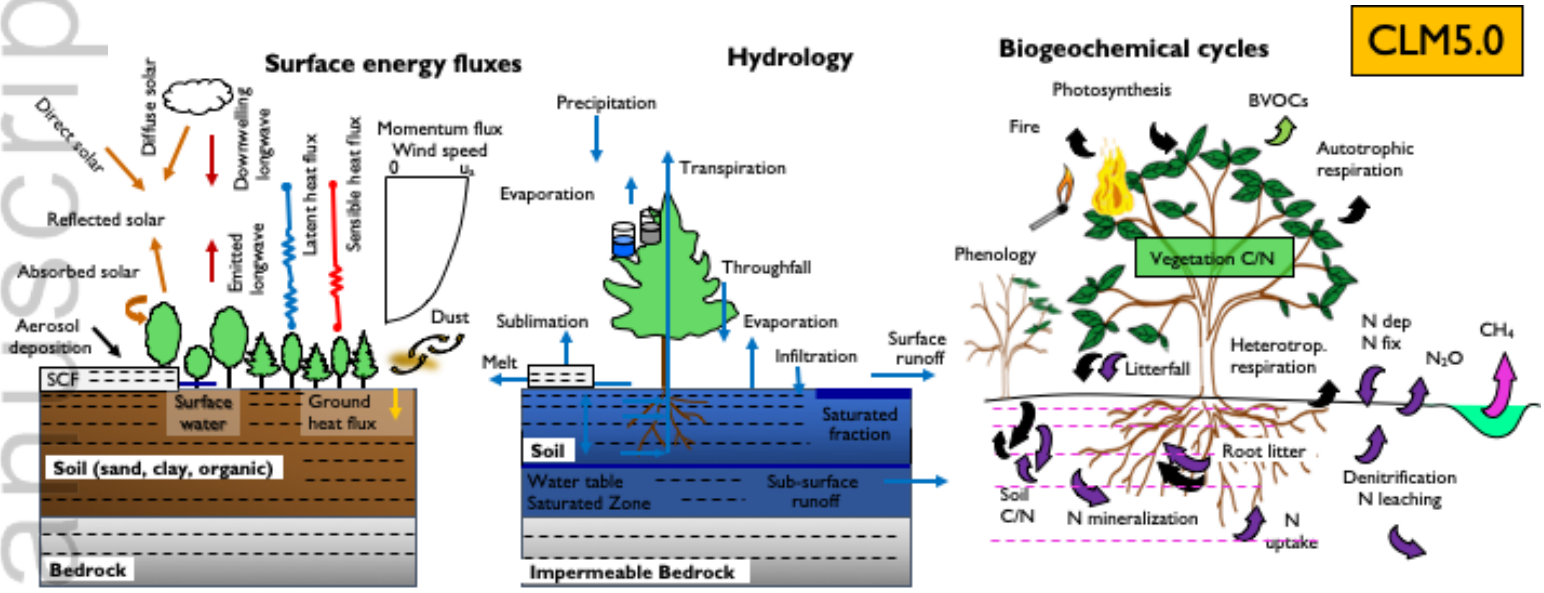
Figure 12: Metric for apparent soil carbon turnover time versus mean air temperature, as in Koven et al. (2017) for observations and CLMBGC model versions. Turnover time is calculated in observations and models as ratio of mean carbon stocks (SOM) over climatological annual mean carbon inputs (NPP). Each dot represents one grid cell, color coded by mean annual precipitation. The best-fit regression curve for the observational data with 50% prediction intervals is shown as black lines for the models. RMSE represents the agreement with the best-fit curves. See Koven et al. (2017) for full description of this metric. Observations for soil organic matter (SOM) are merged from Harmonized World Soil Database (Fao/Iiasa/Isric/Isscas/Jrc, 2012) and Northern Circumpolar Soil Carbon Database (Hugelius et al., 2013). Observed NPP estimate is from MODIS (Zhao et al., 2005).

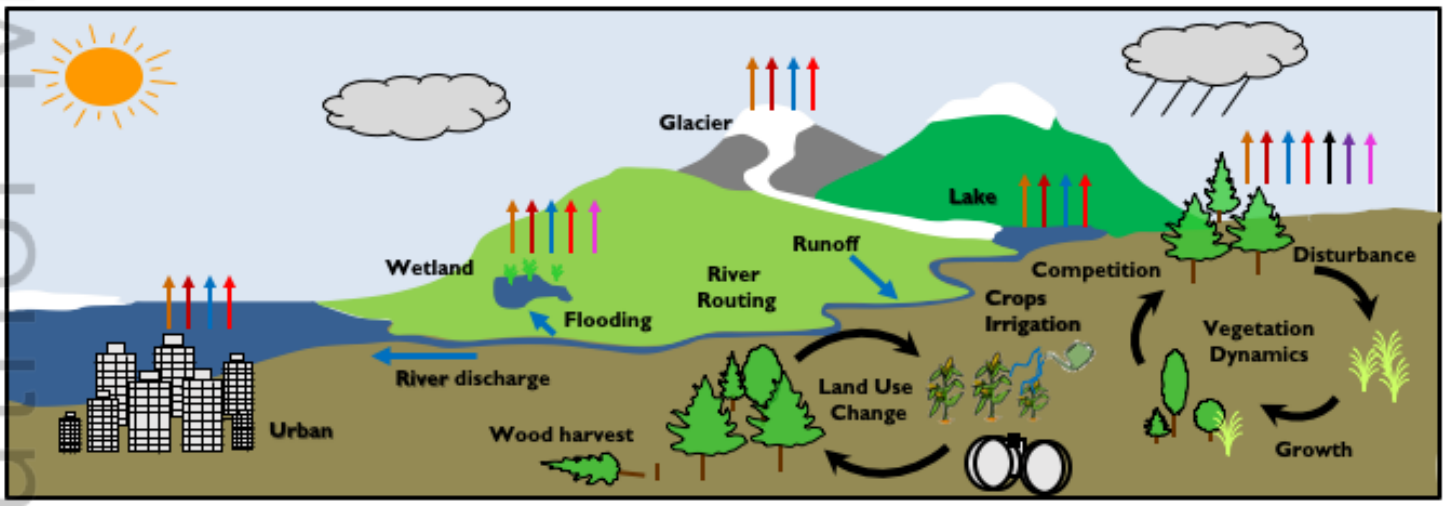
Figure 13: Simulated effect sizes of nitrogen versus CO_2 enrichment on global rates of net primary productivity (NPP) that was calculated for CLM4BGC, CLM4.5BGC, and CLM5BGC (brown, turquoise, and purple symbols, respectively; GSWP3 simulations). Observational constraints for the nitrogen response (above-ground NPP from LeBauer and Treseder (2008)) and CO_2 response (dry matter production from Ainsworth and Long (2004)) are shown with the vertical and horizontal lines, respectively (mean +/- 95% confidence interval). Figure reproduced from Figure 7a in Wieder et al. (2019).

Figure 14: Land carbon accumulation for the period (a)1850-2014 and (b)1960-2014 for CLM4BGC, CLM4.5BGC, and CLM5BGC simulations forced with GSWP3v1 and CRUNCEPv7. Observationally-constrained model reconstructions (black lines, uncertainty estimates in grey) are from (a) Hoffman et al. (2014) and (b) Global Carbon Project (Le Quéré et al. 2014).

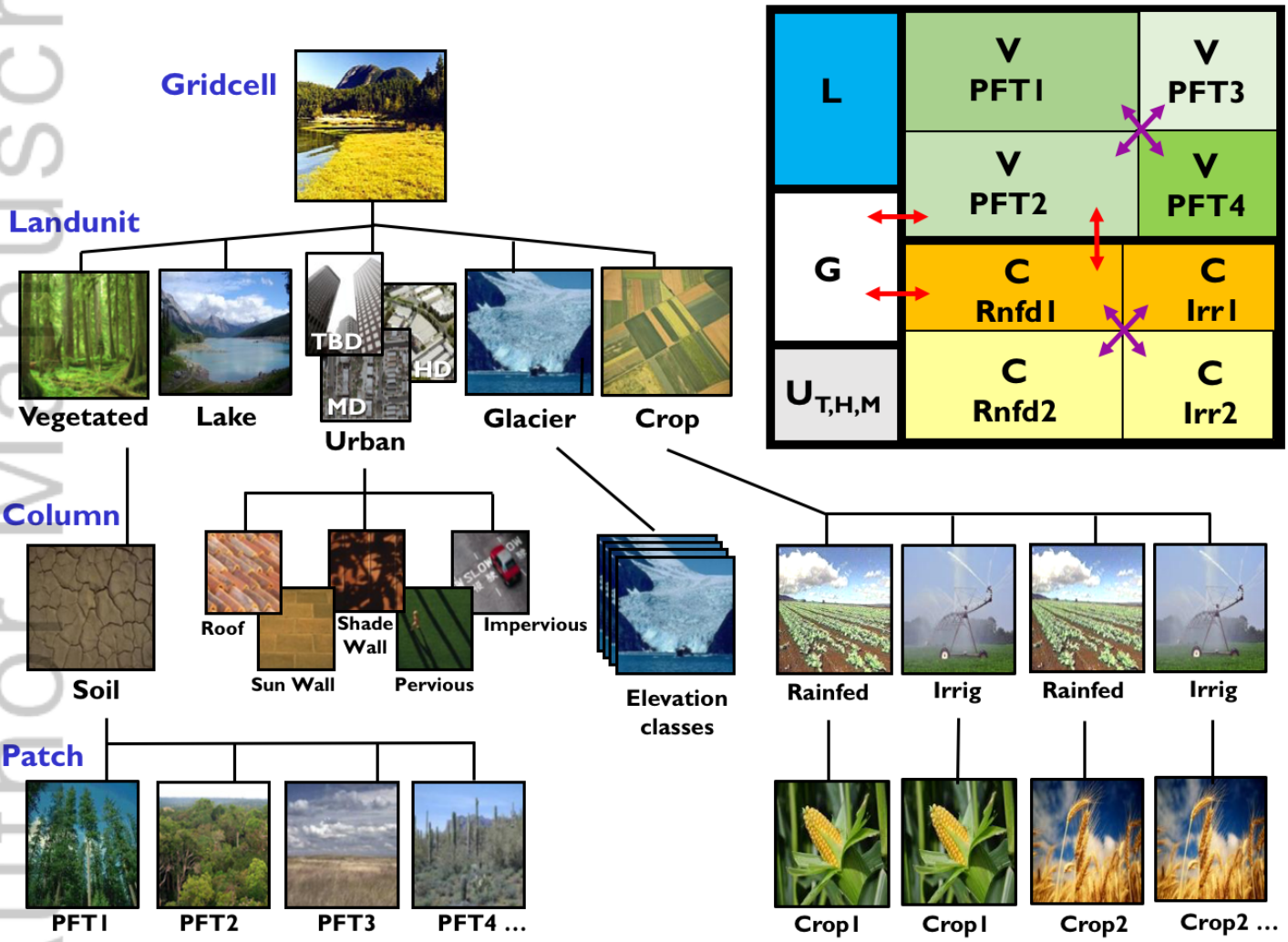
Figure 15: Global time series of the change in (a) mean water use efficiency (WUE = GPP/Transpiration), (b) GPP (weighted sum) and (c) transpiration. Note that only grid cells with annual maximum LAI > 0.1 $mm^2\ mm^{-2}$ are included in the averaging. (d) Global maps of change in WUE over the historical period (2005-2014 minus 1850-1859) for GSWP3v1-forced CLM4BGC, CLM4.5BGC, and CLM5BGC. Grey color denotes regions where annual maximum LAI < 0.1 $mm^2\ mm^{-2}$.

Figure 16: CLM5BGC crop yields were evaluated against data from the United Nations Food and Agriculture Organization (UNFAO), which were downscaled using EarthStat data. The spatial distribution of crop yields averaged from 1990 - 2010 are plotted in (a) for UNFAO and (b) for CLM5 and illustrate the summed yields for all explicitly represented crop types (temperate and tropical soybean, temperate and tropical corn, rice, sugarcane, cotton, and wheat). The time series of globally-summed crop yields for all explicitly represented crop types is plotted in (c) as simulated by CLM5 (black line) over the full time period and evaluated against data from UNFAO (red line) starting in 1961, when data are first available.

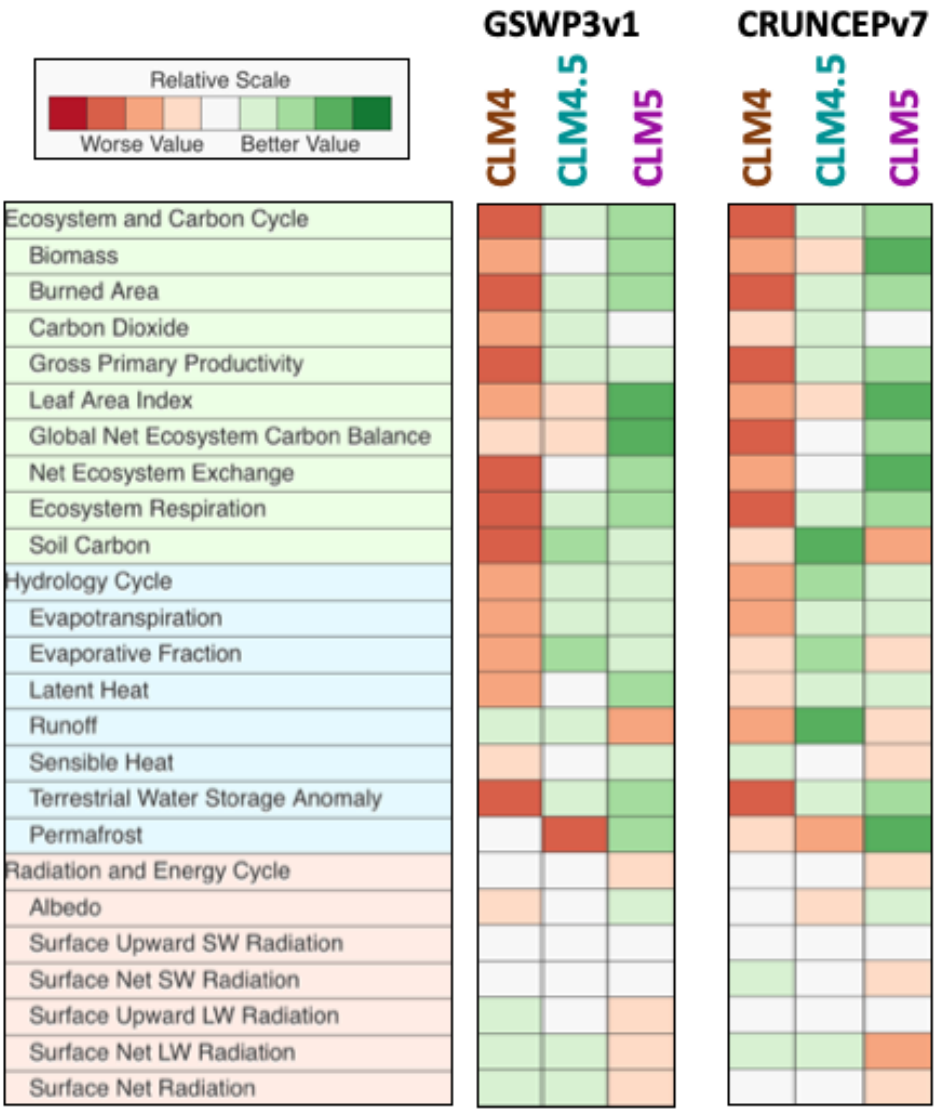




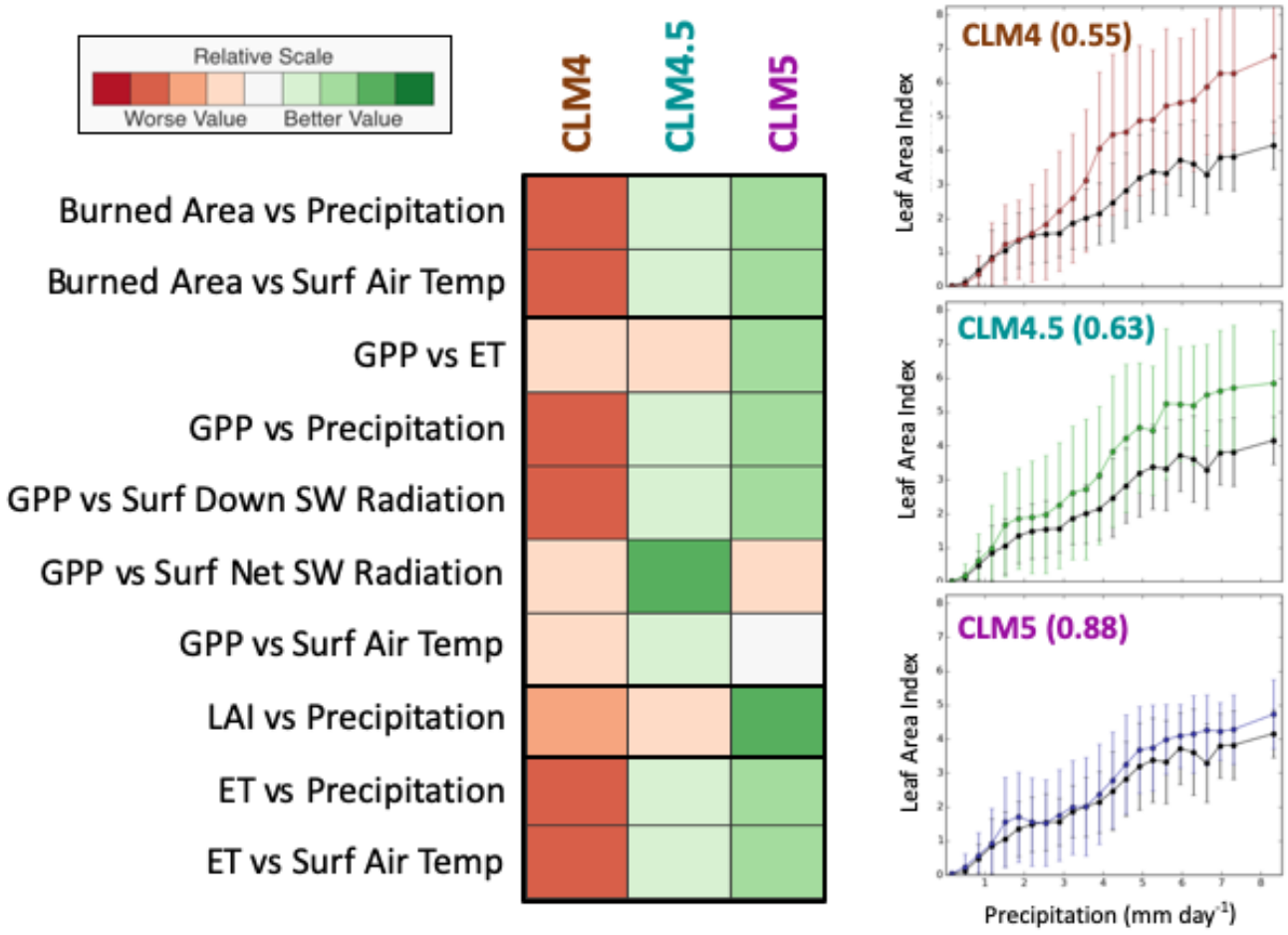
2018MS001583-f01-z-.tif



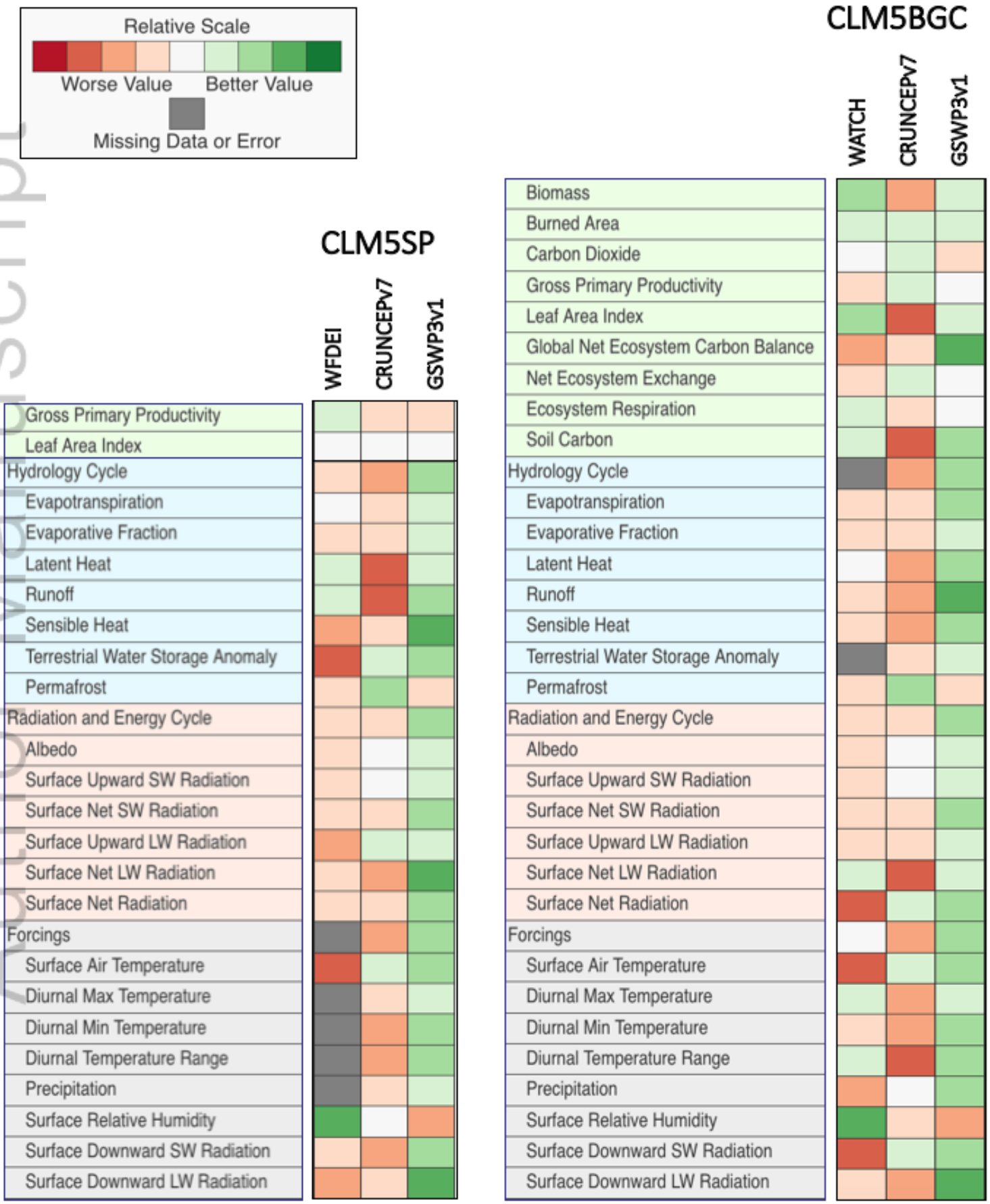
2018MS001583-f02-z-.png



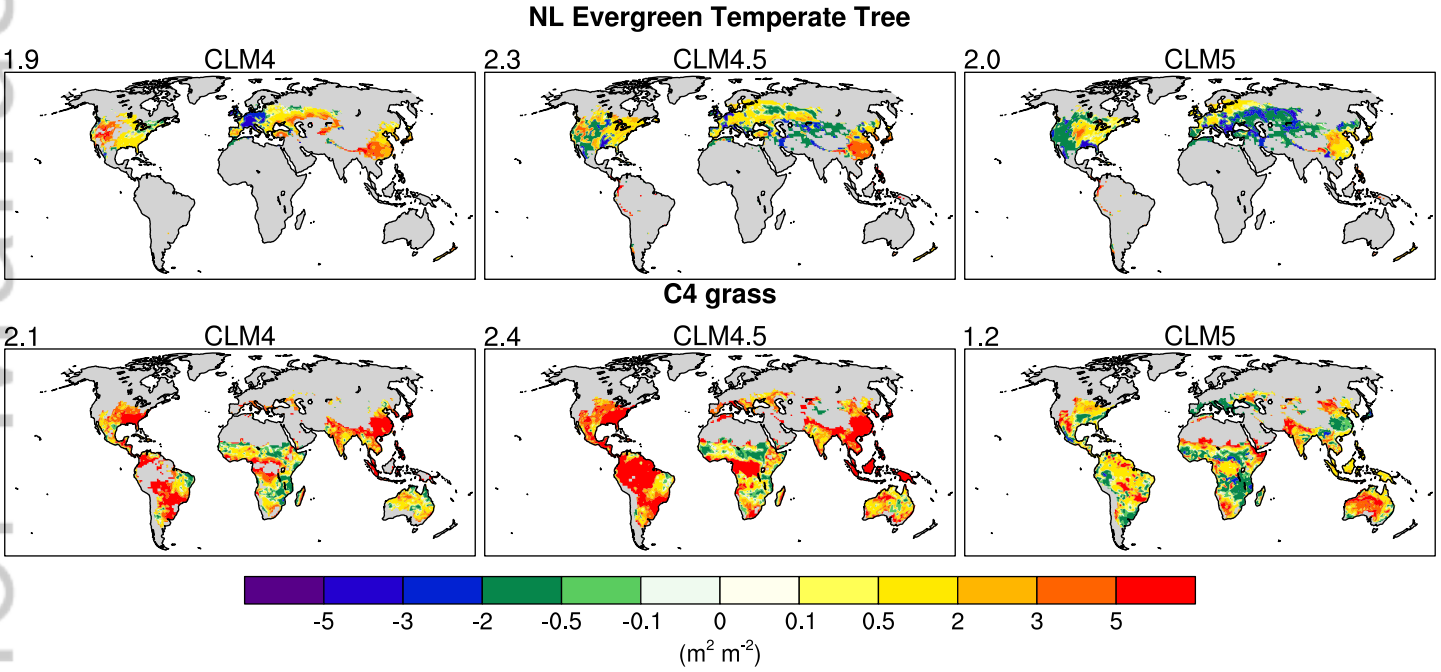
2018MS001583-f03-z-.png

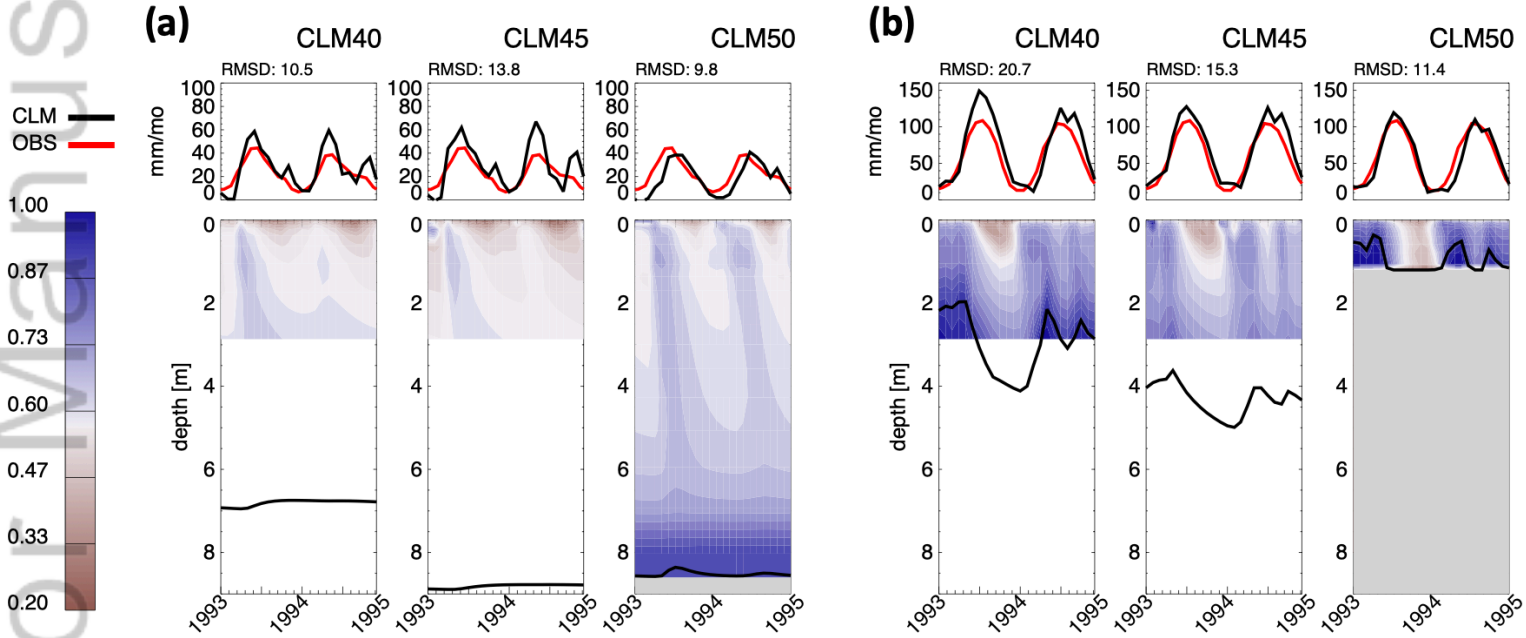


2018MS001583-f04-z-.png

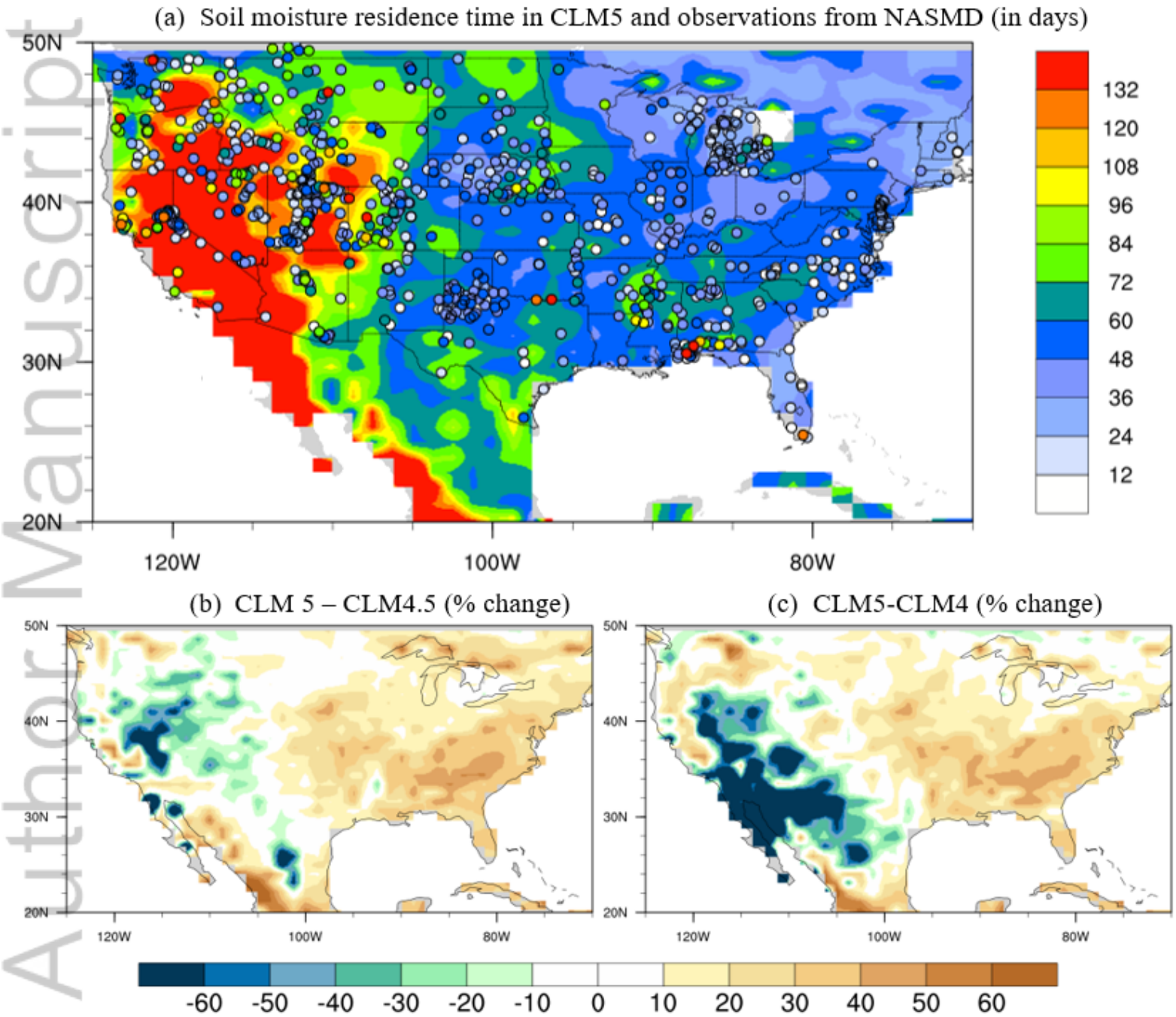


2018MS001583-f05-2.png

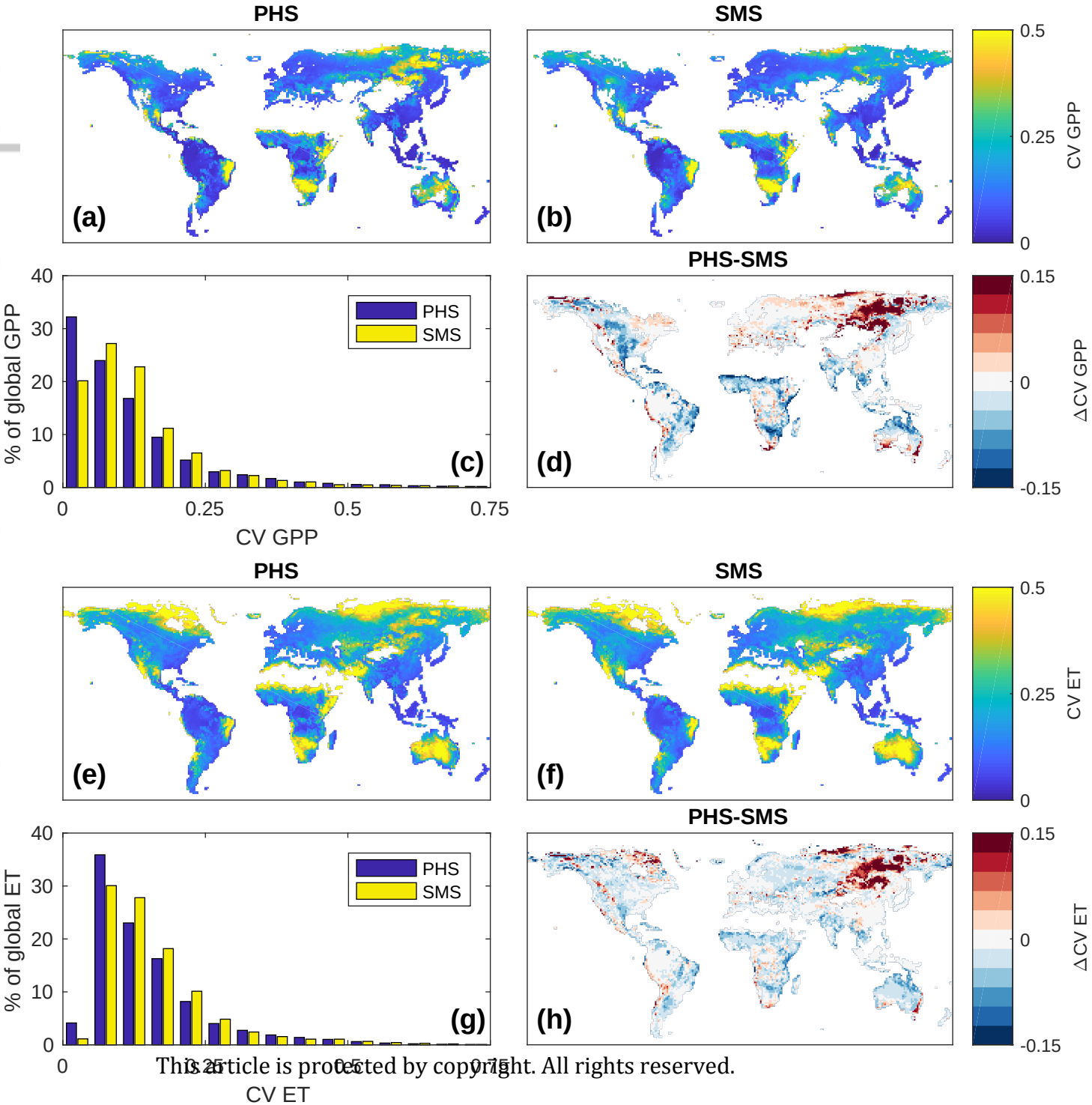


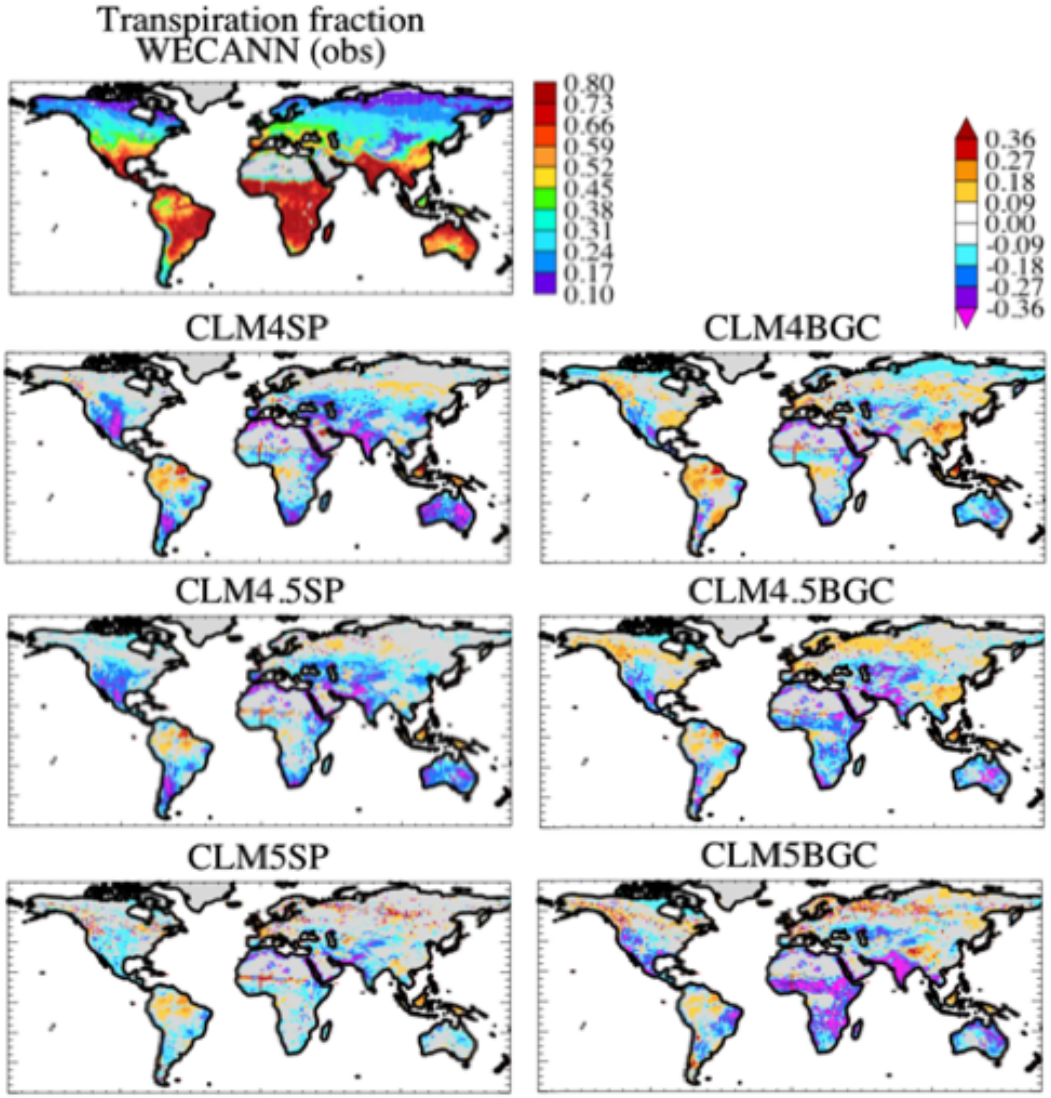


2018MS001583-f07-z-.png

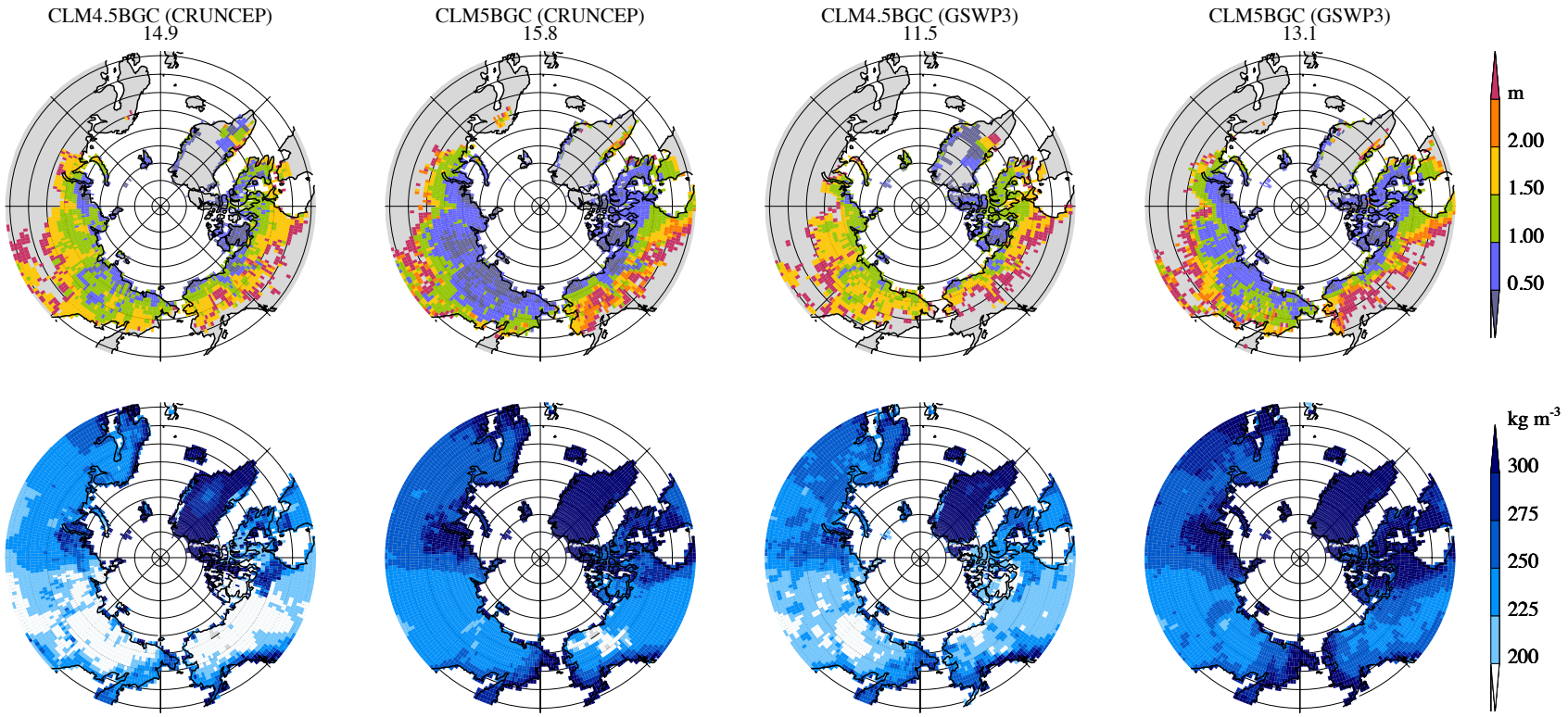


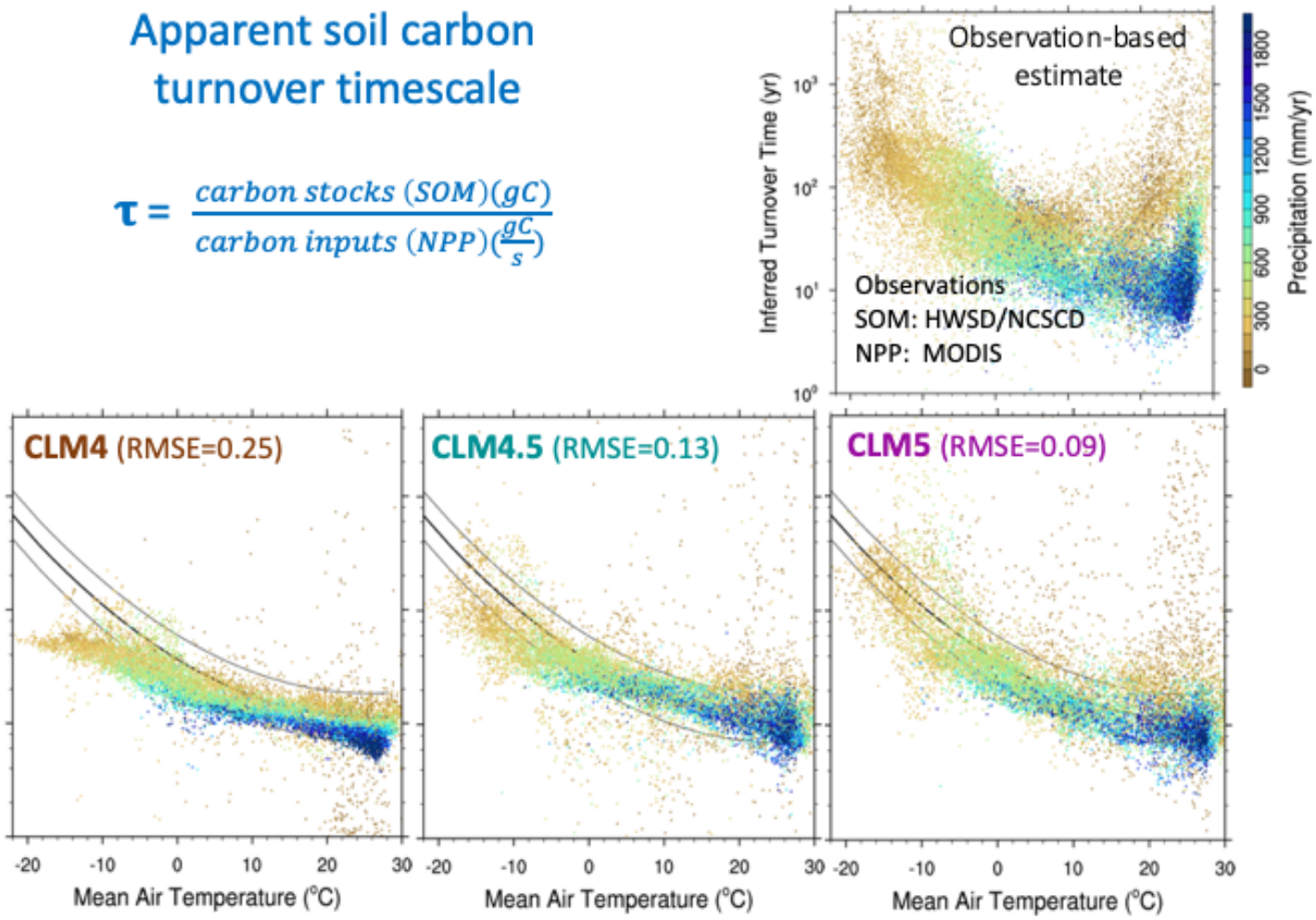
2018MS001583-f08-z-.png





2018MS001583-f10-z-.png





2018MS001583-f12-z-.png