Validation of Terrestrial Biogeochemistry in CMIP6 Earth System Models: A Review

Lynsay Spafford[1,2], Andrew H. MacDougall[1]

[1] Climate and Environment, Saint Francis Xavier University, Antigonish, Canada

[2] Environmental Sciences, Memorial University, St. John's, Canada

Correspondence to: Lynsay Spafford (lspaffor@stfx.ca)

Abstract. The vital role of terrestrial biogeochemical cycles in influencing global climate change is explored by modelling groups internationally through Land Surface Models (LSMs) coupled to atmospheric and oceanic components within Earth System Models (ESMs). The sixth phase of the Coupled Model Intercomparison Project (CMIP6) provided an opportunity to compare ESM output by providing common forcings and experimental protocols. Despite these common experimental protocols, a variety of terrestrial biogeochemical cycle validation approaches were adopted by CMIP6 participants, leading to ambiguous model performance assessment and uncertainty attribution across ESMs. In this review we summarize current methods of terrestrial biogeochemical cycle validation utilized by CMIP6 participants and concurrent community model comparison studies. We focus on variables including: the dimensions of evaluations, observation-based reference datasets, and metrics of model performance. To ensure objective and thorough validations for the seventh phase of CMIP (CMIP7) we recommend the use of a standard validation protocol employing a broad suite of certainty-weighted observation-based reference datasets, targeted model performance metrics, and comparisons across a range of spatiotemporal scales.

1 Introduction

The terrestrial biosphere is presently responsible for sequestering about one quarter of anthropogenic carbon emissions, substantially reducing the severity of ongoing climate change (Friedlingstein et al., 2020). The future capacity of the terrestrial biosphere to sequester CO_2 emissions is uncertain due to non-linear feedbacks such as CO_2 fertilization, growing season extension in cold-limited regions, enhanced heterotrophic respiration, and potentially other feedbacks, as well as environmental and physiological constraints such as moisture availability, nutrient limitations and stomatal closure (Fleischer et al., 2019; Green et al., 2019; Xu et al., 2016; Wieder et al., 2015). Earth system models (ESMs) are a means to simulate past, present, and future terrestrial biogeochemical cycles, examine the influence of changes in climate and atmospheric CO_2 concentration on CO_2 uptake, explore feedbacks and limitations, and estimate anthropogenic carbon emissions compatible with avoiding a given threshold in global temperature change. ESMs simulate global exchanges of matter and energy through the coupling of land, atmospheric, and oceanic components. Through concerted efforts, successive generations of ESMs have improved in

30   terms of spatiotemporal resolution, complexity, and process representation (Anderson et al., 2016). Despite this progress, terrestrial biogeochemical cycles remain a major source of uncertainty in future climate projections (Arora et al., 2020; Lovenduski and Bonan, 2017). This uncertainty stems from limited process understanding, lacking observational constraints, inherent cycle variability, temporal discrepancy between forcings and responses (Sellar et al., 2019; Ciais et al., 2013), and uncertain stock quantifications (Ito et al., 2020; Wieder et al., 2015) which together compound uncertainty within models.

35   Among models, this uncertainty is amplified by artefacts in the form of inconsistent model structure, boundary conditions, forcing datasets, experimental protocols, and benchmarking observational datasets, which is magnified by the increasing number, diversity, and complexity of ESMs (Eyring et al., 2020). Subsequently, a study on uncertainty in projected terrestrial carbon uptake based upon 12 Coupled Model Intercomparison Project phase 5 (CMIP5) ESMs indicated that uncertainty stemming from model structure may be four times greater than uncertainty from different emission scenarios and internal

40   variability (Lovenduski and Bonan, 2017). Some progress has been made in addressing the large uncertainty associated with the terrestrial biogeochemistry in ESMs, as comparison of the carbon-climate and carbon-concentration feedback among ESMs participating in the sixth phase of CMIP (CMIP6) by Arora et al. (2020) shows a reduced model spread amongst models which included a nitrogen cycle, which provided a realistic constraint on photosynthesis in the context of elevated atmospheric CO_2 concentration. However, the spread in estimated feedback parameters across ESMs overall has not been significantly reduced

45   from CMIP6 relative to CMIP5 (Arora et al., 2020; 2013).

        To answer scientific questions regarding climate change, the CMIP was initiated in 1995 by the World Climate Research Programme's (WCRP) Working Group of Coupled Modelling (WCRP, 2020). The CMIP designates standard experimental protocols, model output formats, and model forcings to diagnose climate change variability, predictability, and uncertainty following various scenarios within a multi-model framework. CMIP6 began in 2013 with three years of planning

50   and community consultation to address knowledge gaps, prior to the conduction of simulations and analyses in 2016 and onwards. Model validation in the context of CMIP consists of demonstrating sufficient agreement between model output data and historical observation-based reference data following model development and is a crucial process in model advancement. Such comparison facilitates model improvement by identifying model limitations in performance or sources of model-data uncertainty (Lovenduski and Bonan, 2017), and informs the weighting of different ESMs in influencing climate projections

55   and policy (Eyring et al., 2019). CMIP6 specified detailed experimental protocols for modelling group participants to facilitate objective comparisons of the output of different models with common forcings (Eyring et al., 2016a).

        Here we focus on validations of the stocks and biological fluxes of fully coupled ESMs and associated LSM releases from 2017 onwards with explicit terrestrial biogeochemical cycle representation contributed by CMIP6 participating modelling groups (hereafter participants; Table 1; Arora et al., 2020). Validations are analyzed in terms of variables included,

60   spatiotemporal scales, reference datasets, and metrics of performance. Section 2 compares the methods of historical terrestrial biogeochemical cycle validation used by participants, Section 3 summarizes the methods used in community analyses of CMIP5 era models, and a critique of these methods. A future outlook is presented in Section 4.

Table 1: Modelling group contributions to C4MIP of CMIP6 from Arora et al. (2020).

Modelling Group	ESM	Land Surface Model Biogeochemistry Component	Explicit N Cycle	Dynamic Vegetation	Prognostic LAI	Prognostic Leaf Phenology	Reference(s)
CSIRO	ACCESS-ESM1.5	CABLE2.4	Yes	No	Yes	No	Ziehn et al., 2020
BCC	BCC-CSM2-MR	BCC-AVIM2	No	No	Yes	Yes (for deciduous)	Wu et al., 2019; Li et al., 2019
CCCma	CanESM5	CLASS-CTEM	No	No	Yes	Yes	Swart et al., 2019
CESM	CESM2	CLM5	Yes	No	Yes	Yes	Danabasoglu et al., 2020; Lawrence et al., 2019
CNRM	CNRM-ESM2-1	ISBA-CTRIP	No	No	Yes	Yes (from leaf carbon balance)	Séférian et al., 2019; Delire et al., 2020
GFDL	GFDL-ESM4	LM4.1	No	Yes	-	-	Dunne et al., 2020
IPSL	IPSL-CM6A-LR	ORCHIDEE, version 2.0	No	No	Yes	Yes	Boucher et al., 2020; Vuichard et al., 2019
JAMSTEC	MIROC-ES2L	VISIT-e	Yes	No	Yes	Yes	Hajima et al., 2020
MPI	MPI-ESM1.2-LR	JSBACH3.2	Yes	Yes	Yes	Yes	Mauritsen et al., 2019; Goll et al., 2017
NCC	NorESM2-LM	CLM5	Yes	No	Yes	Yes	Seland et al., 2020
UK	UKESM1-0-LL	JULES-ES-1.0	Yes	Yes	Yes	Yes	Sellar et al., 2019

65

2. Participant Methods of Validating Terrestrial Biogeochemical Cycles

To participate in CMIP6, participants had to submit four Diagnosis, Validation, and Characterization of Klima (DECK) experimental simulations which included a control simulation with prescribed idealized pre-industrial (1850) forcing for at least 500 years to demonstrate stability in global climate and biogeochemical exchanges. Additionally, participants had to conduct historical simulations from 1850-2014 using designated CMIP6 forcings (available at https://esgf-node.llnl.gov/search/input4MIPs/, last access: February 8th, 2021) as well as initialization from the pre-industrial forcing control run (Eyring et al., 2016a). Each modelling group demonstrated stability in the global carbon cycle, with global net

70

carbon exchange below the suggested limit of 0.1 PgC year^{-1} by Jones et al. (2016), while no suitable pre-industrial simulation global nitrogen or phosphorus flux was specified for CMIP6 though these were generally below 2.0 Pg year^{-1} (Ziehn et al., 2020). Each modelling group validated terrestrial biogeochemical cycle components for the historical simulation in a unique fashion, which is summarized below and detailed in Appendix A.

2.1 Variables Included in Validations

The number of terrestrial biogeochemical cycle variables evaluated against observation-based estimates by participants varied considerably from 0 to 21, with a total of 38 unique variables evaluated by all participants combined. The variable validated most often was Gross Primary Production (GPP), which was validated by all but one participant. The next nine most validated variables in descending order were soil carbon, the global land carbon sink, leaf area index (LAI), vegetation carbon, ecosystem respiration, global land-atmosphere CO_2 flux, surface CO_2 concentrations, total biomass, and burned area (Fig. 1). For a list of variable definitions, see Table 2.

Table 2: Terms associated with terrestrial biogeochemical cycles and their definitions as used by participants.

Term	CMIP6 Definition
Gross Primary Production (GPP)	The quantity of CO_2 removed from the atmosphere by vegetation.
Net Primary Productivity (NPP)	The quantity of CO_2 removed from the atmosphere by vegetation minus the quantity of CO_2 from autotrophic respiration.
Autotrophic Respiration (AR)	The quantity of CO_2 from cellular respiration in plants.
Ecosystem Respiration (ER)	The quantity of CO_2 from autotrophic respiration and heterotrophic respiration.
Heterotrophic Respiration (HR)	The quantity of CO_2 from cellular respiration by heterotrophs.
Net Ecosystem Production (NEP)	The quantity of CO_2 removed from the atmosphere by vegetation minus the quantity of CO_2 from autotrophic and heterotrophic respiration.
Net Biome Production (NBP)	The net rate of organic carbon accumulation minus autotrophic and heterotrophic respiration as well as non-respiratory losses from disturbance.
Net Ecosystem Carbon Balance (NECB)	The net rate of organic carbon accumulation in an ecosystem, independent of scale.
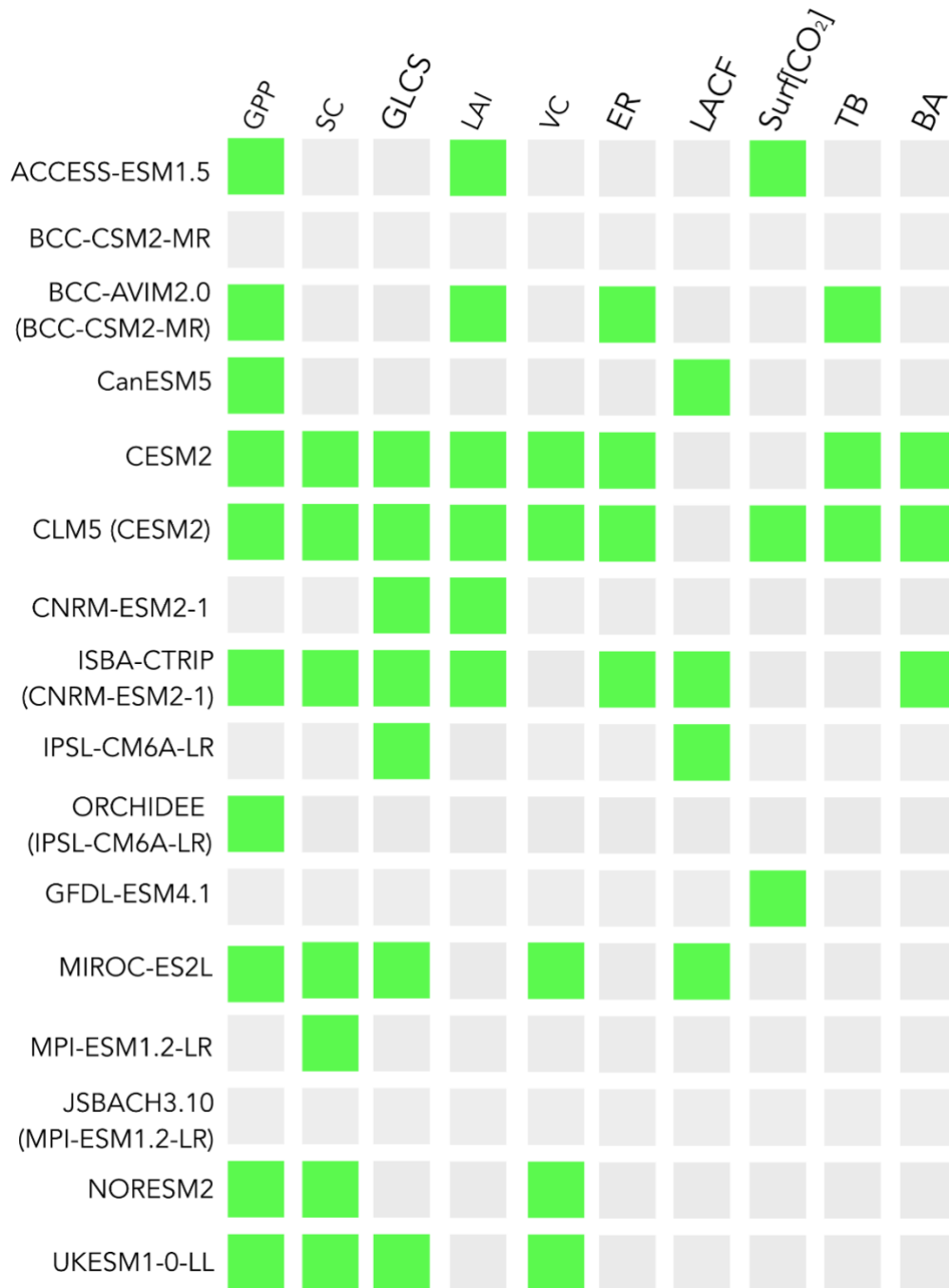
Figure 1: Validation (green) or omission (grey) of the ten most frequently validated variables by participants (treating ESMs and LSMs separately), including Gross Primary Productivity (GPP), Soil Carbon (SC), Global Land Carbon Sink (GLCS), Leaf Area Index (LAI), Vegetation Carbon (VC), Ecosystem Respiration (ER), Land-Atmosphere CO_2 Flux (LACF), Surface CO_2 concentrations (Surf[CO_2]), Total Biomass (TB), and Burned Area (BA).

90

The majority of variables were validated by just one or two participants (Fig. 2). Danabasoglu et al. (2020) and Lawrence et al. (2019) validated a relatively extensive suite of variables with the International Land Model Benchmarking (ILAMB) package version 2.1 (ILAMBv2.1; Collier et al., 2018, Fig. 3), including an explicit uncertainty analysis of the influences of interannual variability, forcing datasets, and model structure in the form of prescribed versus prognostic vegetation phenology. While no nitrogen cycle variable was validated by more than one group, soil N_2O flux and total N_2O emissions were evaluated by Hajima et al. (2020) and Lawrence et al. (2019), respectively.
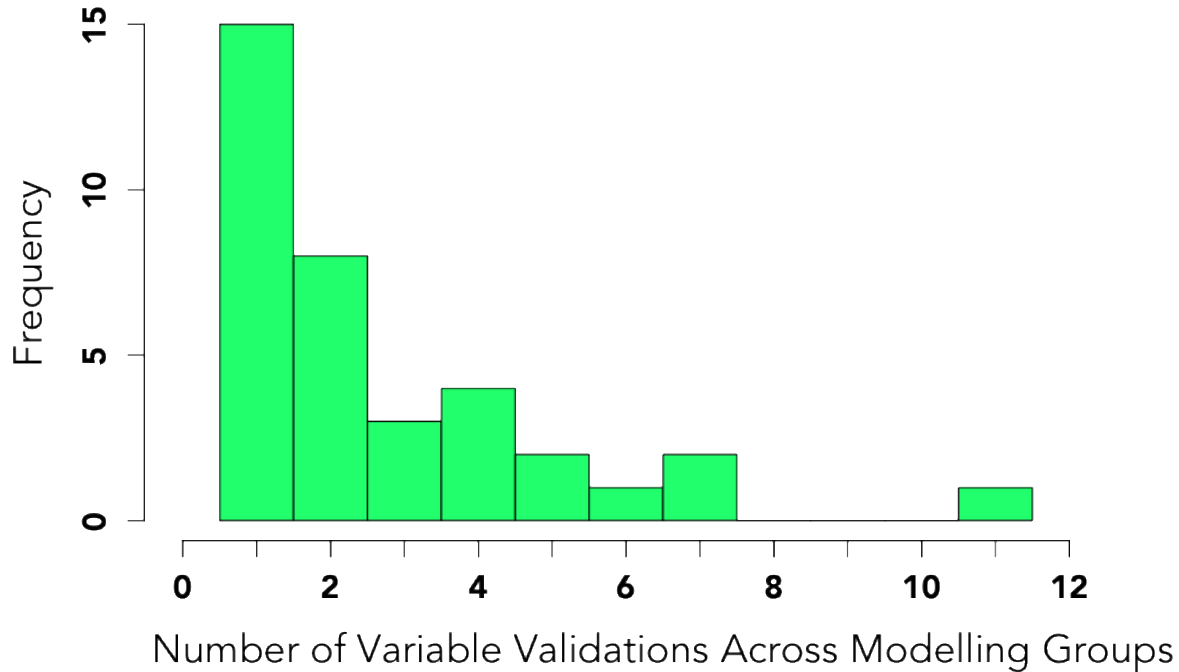
Figure 2: Frequency of a given variable being validated across participants (treating ESMs and LSMs separately). Most variables were validated only once across participants (leftmost x-axis), while GPP was validated by 11 participants (rightmost bar).

Relative Scale

Worse Value — Better Value

Missing Data or Error

CLM5SP

Columns: WFDEI, CRUNCEPv7, GSWP3v1

	WFDEI	CRUNCEPv7	GSWP3v1
Gross Primary Productivity			
Leaf Area Index			
Hydrology Cycle			
Evapotranspiration			
Evaporative Fraction			
Latent Heat			
Runoff			
Sensible Heat			
Terrestrial Water Storage Anomaly			
Permafrost			
Radiation and Energy Cycle			
Albedo			
Surface Upward SW Radiation			
Surface Net SW Radiation			
Surface Upward LW Radiation			
Surface Net LW Radiation			
Surface Net Radiation			
Forcings			
Surface Air Temperature			
Diurnal Max Temperature			
Diurnal Min Temperature			
Diurnal Temperature Range			
Precipitation			
Surface Relative Humidity			
Surface Downward SW Radiation			
Surface Downward LW Radiation			

CLM5BGC

Columns: WATCH, CRUNCEPv7, GSWP3v1

	WATCH	CRUNCEPv7	GSWP3v1
Biomass			
Burned Area			
Carbon Dioxide			
Gross Primary Productivity			
Leaf Area Index			
Global Net Ecosystem Carbon Balance			
Net Ecosystem Exchange			
Ecosystem Respiration			
Soil Carbon			
Hydrology Cycle			
Evapotranspiration			
Evaporative Fraction			
Latent Heat			
Runoff			
Sensible Heat			
Terrestrial Water Storage Anomaly			
Permafrost			
Radiation and Energy Cycle			
Albedo			
Surface Upward SW Radiation			
Surface Net SW Radiation			
Surface Upward LW Radiation			
Surface Net LW Radiation			
Surface Net Radiation			
Forcings			
Surface Air Temperature			
Diurnal Max Temperature			
Diurnal Min Temperature			
Diurnal Temperature Range			
Precipitation			
Surface Relative Humidity			
Surface Downward SW Radiation			
Surface Downward LW Radiation			

A variety of spatiotemporal scales of these variables were considered in validations both within and among participants. Spatial scales consisted of site-level, model grid cell, degree of latitude, region, and global, with the latter being the most common across participants. Temporal scales included daily, seasonal, annual, decadal, select periods, and long-term trends, accumulations, or averages over the whole historical simulation period from 1850-2014. For more detail on the spatiotemporal scales of validation used by each participant, readers should refer to Appendix A. Dynamic variables such as LAI were subject to a detailed assessment, including annual maximum and minimum magnitude (Séférian et al., 2019) and month (Li et al., 2019), seasonality (Ziehn et al., 2020), and seasonal average, as well as global averages. GPP was also evaluated across a variety of scales, including in terms of the daily, seasonal, and annual magnitude on a plant functional type (PFT), spatial, and global basis against site-level observations (Vuichard et al., 2019), as well as globally in terms of functional relationships with temperature and precipitation (Swart et al., 2019) and the relative contribution of drivers of variation (Vuichard et al., 2019). Biomass and carbon stock variables were evaluated in terms of spatial distributions or global averages over chosen time periods, often on a decadal scale (Li et al., 2019). Global vegetation and soil carbon turnover times were also evaluated for selected time periods (Delire et al., 2020; Lawrence et al., 2019).

2.2 Reference Datasets

For variables which were validated by more than one modelling group, such as GPP, a variety of observation-based reference datasets were utilized. For example, across participants, several different GPP reference datasets were used (Table 3), though most participants utilized model tree ensemble (MTE) machine-learning upscaled ground eddy-covariance, meteorological, and satellite observation-based estimates of GPP from Jung et al. (2011). Interestingly one group, Centre National de Recherches Météorologiques (CNRM; Delire et al., 2020) used a more recent Fluxnet-based GPP dataset (FluxComv1; Jung et al., 2016; Tramontana et al., 2016), and further used the mean of 12 products therein. CNRM along with the Institut Pierre Simon Laplace (IPSL, Vuichard et al., 2019) were the only groups to include a comparison to site-level GPP observations. A variety of reference datasets were also utilized for the second most frequently validated variable, soil carbon (Table 4), spanning a 12-year publication range (Batjes, 2016; Global Soil Data Task Group, 2002). Several participants used more than one reference dataset for evaluation of soil carbon depending upon regional or global focus, such as the Northern Circumpolar Soil Carbon Database provided by Hugelius et al. (2013) for mid-high latitudes, while global soil carbon estimates were obtained from Batjes (2016), Carvalhais et al. (2014), Todd-Brown et al. (2013), and FAO (2012). While biomass and carbon stocks were predominantly compared to present day observations, Delire et al. (2020) used records from the Global Database of Litterfall Mass and Litter Pool Carbon and Nutrients database which extends from 1827-1997 (Holland et al., 2015).

Table 3: The source for Gross Primary Production (GPP) data referenced by each modelling group for ESM or LSM simulations. Adjacent contributions from the same modelling group are banded in a common fashion for readability. LSM-focused validations by each modelling group are presented with the associated ESM in brackets.

Model Validation	GPP Reference Data
ACCESS-ESM1.5	Jung et al., 2011; Ziehn et al., 2011; Beer et al., 2010
BCC-CSM2-MR	-
BCC-AVIM2.0 (BCC-CSM2-MR)	Jung et al., 2011
CanESM5	Jung et al., 2009
CESM2	Jung et al., 2011
CLM5 (CESM2)	Jung et al., 2011
CNRM-ESM2-1	-
ISBA-CTRIP (CNRM-ESM2-1)	Jung et al., 2016; Tramontana et al., 2016; Joetzjer et al., 2015
IPSL-CM6A-LR	-
ORCHIDEE (IPSL-CM6A-LR)	Jung et al., 2011
GFDL-ESM4.1	-
MIROC-ES2L	Jung et al., 2011
MPI-ESM1.2-LR	-
JSBACH3.10 (MPI-ESM1.2-LR)	-
NORESM2	Jung et al., 2011
UKESM1-0-LL	Jung et al., 2011

Table 4: The source for soil carbon data referenced by each modelling group for ESM or LSM simulations. Adjacent contributions from the same modelling group are banded in a common fashion for readability. LSM-focused validations by each modelling group are presented with the associated ESM in brackets.

Model Validation	Soil Carbon Reference Data
ACCESS-ESM1.5	-
BCC-CSM2-MR	-
BCC-AVIM2.0 (BCC-CSM2-MR)	-
CanESM5	-
CESM2	Hugelius et al., 2013; Todd-Brown et al., 2013
CLM5 (CESM2)	FAO, 2012
CNRM-ESM2-1	-
ISBA-CTRIP (CNRM-ESM2-1)	FAO, 2012
IPSL-CM6A-LR	-
ORCHIDEE (IPSL-CM6A-LR)	-

GFDL-ESM4.1	-
MIROC-ES2L	Batjes, 2016; Hugelius et al., 2013; Todd-Brown et al., 2013
MPI-ESM1.2-LR	Goll et al., 2015
JSBACH3.10 (MPI-ESM1.2-LR)	-
NORESM2	FAO, 2012
UKESM1-0-LL	Batjes, 2016; Carvalhais et al., 2014; Global Soil Data Task Group, 2002

2.3 Statistical Metrics of Model Performance

A variety of statistical metrics were used to quantify model performance in simulating historical variables in comparison to
150   observations, though chosen metrics were more consistent than selected variables. The comparison of simulated and observation-based averages calculated over space and time was the most common metric used by all but two participants (Table 5). The next most commonly used metric was root mean squared error (RMSE), followed by bias (simulated – observed) on a spatial or global basis. Evaluations of global accumulations, seasonal phase, seasonal maximum and or minimum, as well as global totals were also used. The Taylor diagram, which geometrically combines spatiotemporal correlation, standard
155   deviations, and root mean square (RMS) difference (Taylor, 2001) was used to summarize model performance by three participants (Li et al., 2019; Collier et al., 2018; Goll et al., 2017). The correlation coefficient (r) was also used by three participants (Swart et al., 2019; Mauritsen et al., 2019; Goll et al., 2017). RMSE normalized by the standard deviation of observations (NRMSE) was only used by Swart et al. (2019), while the coefficient of determination (r^2) was only used by Mauritsen et al. (2019). A targeted metric in the form of dissected mean squared deviation (Kobayashi and Salam, 2000), the
160   sum of squared bias, squared difference between standard deviations, and lack of correlation weighted by standard deviation, was used to distinguish model sources of error by Vuichard et al. (2019). In addition to quantitative metrics, the qualitative aspects of simulations were compared to observational reference data, such as in demonstrating source or sink behaviour over time (Danabasoglu et al., 2020), or in visual comparison of spatial distribution maps.

Table 5: Model performance metrics used by each modelling group for ESM or LSM simulations. Adjacent contributions from the
165   **same modelling group are banded in a common fashion for readability. LSM-focused validations by each modelling group are presented with the associated ESM in brackets.**

Model Validation	Presented Model Performance Assessment Metrics
ACCESS-ESM1.5	Space-time averages, seasonal amplitude, timing and magnitude of annual maximums and minimums

BCC-CSM2-MR	-
BCC-AVIM2.0 (BCC-CSM2-MR)	Average annual cycle phase, global mean bias, RMSE, Taylor score
CanESM5	Space-time averages, geographic distribution of time averages and bias, latitudinal averages, correlation coefficient (r), RMSE, NRMSE (RMSE ÷ standard deviation of observations), change in NRMSE
CESM2	Space-time averages, seasonal cycles, spatial distributions, time series, interannual variability, global accumulations, functional relationships, relative bias, RMSE, ILAMB relative scale
CLM5 (CESM2)	Space-time averages, seasonal cycles, annual monthly maximum, spatial distributions, global totals, turnover time, time series, interannual variability, functional relationships, bias, relative bias, RMSE, ILAMB relative scale
CNRM-ESM2-1	Average annual maximums and minimums, spatial distribution, bias, RMSE, model correlation between spatial pattern of error
ISBA-CTRIP (CNRM-ESM2-1)	Geographic distribution of time averages and bias, latitudinal averages, global accumulations, bias, spatial correlation, turnover time, average annual maximums, seasonal cycle amplitude and phase
IPSL-CM6A-LR	Global annual averages and accumulations over time
ORCHIDEE (IPSL-CM6A-LR)	Daily, seasonal, annual averages, spatial distribution, regional averages, global averages, RMSE, NRMSE, dissected mean squared deviation (squared bias, squared difference between standard deviations, lack of correlation weighted by standard deviations from Kobayashi and Salam (2000)), relative drivers of variation
GFDL-ESM4.1	Spatial distribution of seasonal amplitude, interannual variability, RMSE, correlation coefficient (r), coefficient of determination (r^2)
MIROC-ES2L	Space-time averages, latitudinal averages, spatial distribution, gradient, seasonality, density, global accumulations
MPI-ESM1.2-LR	Spatial variability, latitudinal average density, global accumulations
JSBACH3.10 (MPI-ESM1.2-LR)	Space-time averages, spatial variability, frequency distribution, response ratio, correlation coefficient (r), RMSE, Taylor score
NORESM2	Global averages and totals
UKESM1-0-LL	Space-time averages, spatial distribution, latitudinal averages, global accumulations and totals

3. Community Methods of Validating Terrestrial Biogeochemical Cycles

A variety of software and projects have been dedicated to the communal evaluation of ESM (Gleckler et al., 2016) and LSM

170 performance (Kumar et al., 2012; Gulden et al., 2008), with CMIP6-era collaborative efforts including the Earth System Model Evaluation Tool version 2 (ESMValToolv2.0; Eyring et al., 2016b) and ILAMBv2.1 (Danabasoglu et al., 2020; Lawrence et al., 2019; Collier et al., 2018). Both ESMValToolv2.0 and ILAMBv2.1 are openly available tools for the evaluation of a variety of model output against re-processed observations (cmip-esmvaltool.dkrz.de; https://pypi.org/project/ILAMB/; Eyring et al., 2020; 2016b; Collier et al., 2018). The observation-based reference datasets for each are displayed in Table 6. For

175 ESMValToolv2.0 dataset re-processing for compatible comparison in space as well as masking of missing observations is detailed in Righi et al. (2020). The analysis of the land carbon cycle in ESMValToolv2.0 (Eyring et al., 2020) is based upon the approach of Anav et al. (2013) in considering long-term trends, interannual variability, and seasonal cycles. A variety of model performance tailored metrics are available with ESMValToolv2.0 (Eyring et al., 2020). The relative space-time root-mean square deviation (RMSD) indicates model success relative to the multi-model median in simulating the seasonal cycle

180 of key variables, originally from Flato et al. (2013), and allows simultaneous comparison to more than one observational reference for each simulated variable, where available. ESMValTool2.0's AutoAssess function provides a highly resolved model performance evaluation for 300 individual variables, originally developed by the UK Met Office. Further, land cover can be comprehensively evaluated with ESMValToolv2.0 in terms of areas, mean fractions, and biases on a regional and global basis, accommodating different model representations of land cover. ILAMBv2.1 was used to validate terrestrial

185 biogeochemical cycle components in CESM2 (Danabasoglu et al., 2020) and CLM5 (Fig. 3; Lawrence et al., 2019). ILAMBv2.1 was also used to demonstrate the absolute and relative performance of DGVMs within several iterations of the Global Carbon Project (Friedlingstein et al., 2020; 2019; Le Quéré et al., 2018). In addition to variables presented in Table 6, functional relationships between these variables and temperature and precipitation are provided for validation purposes in ILAMBv2.1. ILAMBv2.1 employs a weighting system to assign scores to observation-based datasets, which encompasses

190 certainty measures, spatiotemporal scale appropriateness, and process implications. In computing statistical model performance scores, ILAMBv2.1 acknowledges how reference observations represent discontinuous constants in time and space. For example, if a reference dataset contains average information across a span of years, the annual cycle of such a dataset is assumed to be undefined and is therefore not used as a reference. The calculation of averages over time in ILAMBv2.1 addresses spatiotemporally discontinuous data by performing calculations over specific intervals for which data

195 are considered valid. For each variable evaluation, ILAMBv2.1 generates a series of graphical diagnostics, including spatial contour maps, time series plots, and Taylor diagrams (Taylor, 2001), as well as statistical model performance scores including period mean, bias, RMSE, spatial distribution, interannual coefficient of variation, seasonal cycle, and long-term trend. These scores are then scaled based upon the weighting of reference observation-based datasets, and for multi-model comparisons are presented across metrics and datasets to provide a single score.

200

Table 6: Select observation-based reference dataset sources for ESMValToolv2.0 (Eyring et al., 2020) and ILAMBv2.1 (Collier et al., 2018), including Net Biome Production (NBP), Leaf Area Index (LAI), Land Cover (LC), Gross Primary Production (GPP), Net Ecosystem Exchange (NEE), Soil Carbon (SC), Vegetation Carbon (VC), Ecosystem Carbon Turnover (ECT), Vegetation Biomass (VB), and Burned Area (BA). Note that vegetation carbon is dependent upon vegetation biomass.

Variables	ESMValToolv2.0	ILAMBv2.1
NBP	Le Quéré et al., 2018; Maki et al., 2010	Le Quéré et al., 2016; Hoffman et al., 2014
LAI	Zhu et al., 2013; Baret et al., 2007	De Kauwe et al., 2011; Myneni et al., 1997
LC	Defourny et al., 2016	
GPP, NEE	Jung et al., 2019; 2011	Lasslop et al., 2010; Jung et al., 2010

SC	Wieder, 2014	Todd-Brown et al., 2013; Hugelius et al., 2013
VC	Gibbs, 2006	-
ECT	Carvalhais et al., 2014	-
VB	-	Saatchi et al., 2011; Kellndorfer et al., 2013; Blackard et al., 2008
BA	-	Giglio et al., 2010

205

4. Critique of Validation Approaches

While standard protocols were used by participants for historical simulations in CMIP6, no standard protocol in terms of variables evaluated, reference data, performance metrics, or acceptable performance threshold was adopted for terrestrial biogeochemical cycle validation. The validation of particular variables by different participants occasionally employed the
210 same datasets, though in many cases inconsistent reference datasets were used for the same variable, and the spatial and temporal dimension of validations was often distinct. This contrasts with other works employing multiple models such as the Global Carbon Project (Friedlingstein et al., 2020; 2019; Le Quéré et al., 2018) which provide explicit validation criteria, such as simulating recent historical net land-atmosphere carbon flux within a particular range and being within the 90% confidence interval of specified observations. The stringency of such criteria must be carefully chosen to acknowledge the role of
215 observational uncertainty as well as uncertainty stemming from potential model tuning to forcing datasets. The use of different validation approaches impedes the comparison of performance across models, however it also provides a diverse collection of example methods.

4.1 Variable Choice

A comprehensive validation of a process-based model should include all simulated interacting variables for which a reliable empirical reference is available. Improvement in the simulation of one variable through altered parameters, structure, or
220 algorithms may translate into degradation for other variables, which would be otherwise obscured in a restricted variable analysis (Deser et al., 2020; Ziehn et al., 2020; Lawrence et al., 2019). Given the scope of CMIP6 publications in demonstrating model improvements relative to previous versions as well as the results of CMIP6 experiments, it is understandable that most participants validated a few select variables, and more extensive validations may be in preparation. Essential Climate Variables
225 (ECVs) prioritized for land evaluation in the ESMValToolv2.0 included GPP, LAI, and NBP (Eyring et al., 2020; 2016b), as these variables intersect with other ESM components in matter and energy exchanges (Reichler and Kim, 2008). Contrarily, LAI and NBP were not as frequently validated as GPP by CMIP6 participants (Fig. 1), though the third most validated variable, the global land carbon sink, is equivalent to NBP minus land use emissions. The most common variable chosen for validation by participants was GPP, which is advantageous as it represents a crucial carbon cycle flux. GPP designates the quantity of
230 CO_2 removed from the atmosphere and assimilated into structural and non-structural carbohydrates during photosynthesis by

vegetation, part of which is later respired back to the atmosphere. This quantity is limited by nutrient availability, light, soil moisture, stomatal response to atmospheric CO_2 concentration, and other environmental factors (Davies-Barnard et al., 2020), and is the largest carbon flux between the land biosphere and atmosphere (Xiao et al., 2019). Over- or under-estimations of GPP can lead to biases in carbon stocks, which are exacerbated through time (Carvalhais et al., 2014).

235 An emergent ecosystem property which integrates a variety of influential model processes is carbon turnover time calculated as the ratio of a long-term average total carbon stock compared to GPP or NPP (Eyring et al., 2020; Yan et al., 2017; Carvalhais et al., 2014). Carbon turnover times can be the source of pervasive uncertainty within ESMs, and their misrepresentation can lead to long-term drifts in carbon stocks, fluxes, and feedbacks (Koven et al., 2017). The evaluation capacity of turnover times was seldom utilized by CMIP6 participants, despite soil carbon being a relatively commonly validated variable. Many CMIP5
240 models were found to underestimate turnover times both globally and on a latitudinal basis (Eyring et al., 2020; Fan et al., 2020), while two participants here reported overestimated carbon turnover times, Delire et al. (2020) and Lawrence et al. (2019), though demonstrate improvement from previous models.

Another approach to validation which combines high-level variables and re-parameterization efforts is the assessment of functional relationships or emergent constraints, such as the relationship between GPP or turnover times and temperature,
245 moisture, growing season length, and nutrient stoichiometry (Danabasoglu et al., 2020; Swart et al., 2019; Anav et al., 2015; McGroddy et al., 2004). Physically interpretable emergent constraints can aid in identifying model components which are particularly influential to climate projections (Eyring et al., 2019), such as the temperature control on carbon turnover in the top metre of soil in cold climates (Koven et al., 2017), GPP responses to soil moisture availability (Green et al., 2019), or regional carbon-climate feedbacks (Yoshikawa et al., 2008). With the goal of realistically simulating Earth system processes
250 to develop informed predictions of future climate, large scope variables which inherit uncertainty from an amalgamation of processes are often prioritized for validation. Several participants focused on comparing simulated long-term trends or accumulations in global land carbon fluxes to observation-based estimates from the Global Carbon Project (Friedlingstein et al., 2019; Le Quéré et al., 2018; 2016). While this summation approach can signal a large bias (Eyring et al., 2020; 2016b; Reichler and Kim, 2008) and reduce the effect of sub-scale noise, it does not identify sources of model error or may even
255 obscure model error. For example, if simulated land-atmosphere carbon flux from the pre-industrial era to the 2010s is found to concur with observation-based estimates, this could be due in part to compounding underlying biases which neutralize one another over time (Fisher et al., 2019; Yoshikawa et al., 2008), or alternatively suitable global averages may be susceptible to antagonistic regional biases, such as between the tropics and northern high latitudes. Plant functional type-level evaluations, such as that of the maximum rate of rubisco carboxylation and canopy height by Lawrence et al. (2019) demonstrate the
260 performance of underlying variables in influencing large-scale carbon fluxes and stocks. Several participants included latitudinal-scale evaluations (Delire et al., 2020; Hajima et al., 2020; Mauritsen et al., 2019), which are both informative and readily comparable to observations. A comprehensive validation should therefore encompass a range of scales and a variety of variables to demonstrate model performance in not only producing suitable averages or accumulations, though also in representing processes.

4.2 Reference Datasets

Satellite-based remote sensing of terrestrial biogeochemical components has been conducted for almost 50 years, with the launch of the Landsat satellite in 1972 (Xiao et al., 2019; Mack, 1990), while field-based experimental and observational data has been available since at least the early 19th century (Holland et al., 2015). In terms of just satellite-based observational data products there are currently thousands available (Waliser et al., 2020). Despite this seeming wealth in observational data and observation-based data products, the implementation of a variety of observation-based references for validation of terrestrial biogeochemical cycles within ESMs and LSMs is challenging for several reasons. These include the specifications required for direct model output comparison, inconsistent spatial and temporal domains, missing observations, logistical biases, and large uncertainty in global scale data products (Delire et al., 2020; Collier et al., 2018; Lovenduski and Bonan, 2017). The incomplete coverage of observational datasets in space-time dimensions has led to significant bias in comparisons of model data and observation data previously (de Mora et al., 2013), though was not generally discussed in validation exercises by CMIP6 participants. Observational discontinuity has been addressed previously in a LSM validation by Orth et al. (2017) which excluded daily observation reference averages when more than one hour of data from a 24-hr period was missing, and through exclusion criteria in Collier et al. (2018). For example, the compilation of satellite observations to develop a LAI data product with one observation-based estimate every 15 days by Zhu et al. (2013) for monthly average or seasonal extrema comparison would require careful consideration for comparison to model averages computed from more resolved output. In an analysis of how sparse historical measurements compare to continuous model output, de Mora et al. (2013) demonstrate that where data are lacking in time or space, the discrete comparison of model output to records from site-level measurements may provide a strategic assessment of model performance over time, especially in producing interannual variability. Site-level comparisons of GPP and or CO_2 concentrations were performed by Delire et al. (2020), Dunne et al. (2020), and Vuichard et al. (2019), while Collier et al. (2018) caution against the use of spatially sparse data, though indicate that inclusion of site level evaluations is a key future focus for the ILAMB project.

Another approach to overcome spatial discontinuity may be to compare broad gradients or trends in a given variable with reference datasets, such as regional and functional type trends in forest carbon stocks rather than a global summation or average (Thurner et al., 2014), to investigate whether or not the model captures enduring spatial patterns. In addition, some observational methods may invoke inherent bias, such as satellite-based observation estimates of LAI in mid to high latitudes seasonally underestimating LAI due to snow cover, leading to ambiguous model performance assessment (Ziehn et al., 2020; Liu et al., 2018). Observational uncertainty can be addressed by applying a weighting to reference datasets as in ILAMBv2.1, as well as by using more than one observational reference when available (Eyring et al., 2020; Sellar et al., 2019; Collier et al., 2018). Careful consideration of spatiotemporal discontinuity in observations and inherent bias is warranted in future validations, which can be achieved through filtered exclusions, site-level comparisons, pattern comparison, certainty weighting of datasets, and the use of more than one reference dataset.

The globally gridded 1982-2008 GPP data product frequently used for GPP validation by CMIP6 participants was developed from machine learning upscaling of site-level eddy-covariance Fluxnet observations with model tree ensembles based on remote sensing vegetation indices, meteorological data, and land use (Jung et al., 2011). Observation-based estimates of GPP can be obtained through satellite-derived vegetation indices such as the normalized difference vegetation index (NDVI; Phillips et al., 2008) and solar induced chlorophyll fluorescence (Zhang et al., 2020), in addition to ground-based monitoring of turbulent CO_2 fluxes with the eddy covariance technique (Jung et al., 2009). Logistical challenges with eddy covariance-based techniques of estimating GPP can result in potentially extensive data gaps and systematic omission of diel cycle observations (Rodda et al., 2021; Erkkilä et al., 2017; Jung et al., 2011; Lasslop et al., 2010; 2008; Desai et al., 2008). For example, in a study of eddy-covariance monitoring of CO_2 flux, Jonsson et al. (2008) report only 34% data coverage of a growing season period, of which 54% was discarded as it did not demonstrate energy balance closure. To address these challenges Jung et al. (2011) employ Bowen ratio corrections of energy imbalance (Twine et al., 2000), quality control criteria to exclude sites with more than 20% missing observations, and monthly averages to alleviate noise. Where NEE observations are missing in space over time driver relationships can be utilized for multi-decadal extrapolation, though only 38% and 60% of Fluxnet sites with less than 15 years of observations capture mean conditions and interannual variability of drivers sufficiently well for this extrapolation as of 2015, and most have been operating for less than five years (Chu et al., 2017). While the site-level observations from Jung et al. (2011) originate from 212 sites, presenting a globally extensive network, regions with an important contribution to overall carbon stocks and fluxes are underrepresented (Jung et al., 2020), and even the recent global Fluxnet GPP data product by Jung and Tramontana et al. (2016) has just 14 tropical and 5 Arctic sites. GPP observations from Fluxnet products currently do not account for fire and waterbody emissions, which prompts regional and interannual bias (Jung et al., 2020). Despite these caveats, such global-scale data products provide a critical resource to the CMIP community in conducting model validation (Collier et al., 2018), and the relatively common use of Jung et al. (2011) for validations by CMIP6 participants coincidentally reduces the influence of observational contradiction (Xie et al., 2020; Anav et al., 2015). Site-level GPP evaluation with observations from the tropics by Delire et al. (2020) and Vuichard et al. (2019) demonstrates a strategic approach to addressing the representation bias in GPP validations. Site-level evaluations often benefit from a wealth of available information including spatially consistent meteorological forcing, and avoid the influence of spatial extrapolation error. While Jung et al. (2011) do not provide uncertainty measures, several forms of uncertainty are explicitly presented for the Fluxnet2015 dataset by Pastorello et al. (2020). Therefore the utility of Fluxnet GPP data products could be improved with standardized use by participants in junction with other independent data products, select site-level evaluations, explicit uncertainty quantifications, and improved ecological representation in underlying site-level data.

4.3 Statistical Metrics and Validation Approaches

Several participants relied primarily on residual-based metrics such as bias (simulated-observed) for validation of terrestrial biogeochemical cycle model components. On a spatial basis bias can identify significant regional over- or under-estimations of a given variable. However, the attribution of model error from global maps of bias can be ambiguous, as the displayed bias

is the combined result of different forms of uncertainty, including model structural representations, unforced variability, and spatial disagreement (Deser et al., 2020; Lovenduski and Bonan, 2017; Koch et al., 2016). Such residual-based metrics may not indicate how well the model would perform in simulating future conditions beyond the current contextual envelope of observations (Gulden et al., 2008), and neglect the contribution of uncertainty from observations. These limitations are considerable in the context of ESMs and LSMs as tools for predicting terrestrial biogeochemical function. A more contextualized bias assessment is the Wilcoxon test as applied by Swart et al. (2019) to filter insignificant bias. In a LSM evaluation, Orth et al. (2017) provides an observationally robust bias assessment by subtracting mean seasonal cycles from each grid cell and correlating the resulting anomalies between observation-based datasets and model output. In addition, RMSE normalized by the mean or standard deviation of the observed quantity, NRMSE, contextualizes the difference between simulated and observed variable quantities in terms of the magnitude or inherent variability of the variable of interest (Swart et al., 2019; Fan et al., 2018), which is advantageous for variables such as GPP with large interannual variability.

Beyond these, a variety of targeted model skill metrics have been published for process-based modelling which provide detailed assessments of different forms of model uncertainty (Collier et al., 2018; Orth et al., 2017; Eyring et al., 2016b; Koch et al., 2016; Law et al., 2015; Kumar et al., 2012; Taylor, 2001; Kobayashi and Salam, 2000). Mean squared deviation, the sum of squared bias, squared difference between standard deviations, and lack of correlation weighted by standard deviations, presented by Kobayashi and Salam, (2000), was used by Vuichard et al. (2019). This metric is readily applicable to the objective validation and improvement of mechanistic models, as its dissection allows for the accurate attribution of different sources of model errors. Additionally, a Taylor diagram (Fig. 4, Taylor, 2001) conveys several dimensions of model error and allows for the concise simultaneous display of variables and models and was utilized in the evaluation of BCC-AVIM2 (Li et al., 2019), and NORESM2 (Seland et al., 2020), as well as several LSMs and ESMs by Anav et al. (2015) and is incorporated into ILAMBv2.1 (Collier et al., 2018). The Taylor diagram was designed for simultaneous performance comparison of several simulated variables and serves as a concise and informative validation tool. Caution is warranted however in the evaluation of fully coupled model output due to the inability of fully coupled models to reproduce the timing of internal climate variability phenomena such as El Niño-Southern Oscillation (ENSO; Flato et al., 2013). While the magnitude of observed and simulated internal climate variability may be statistically consistent, bias, RMSE, and NRMSE assessments of fully coupled model output should encompass decadal or longer periods to address the influence of temporal mismatches in simulated internal climate variability relative to observational records. Alternatively, as offline simulations can be directly forced with historical observation data, the output of offline simulations can be validated on a finer temporal scale.
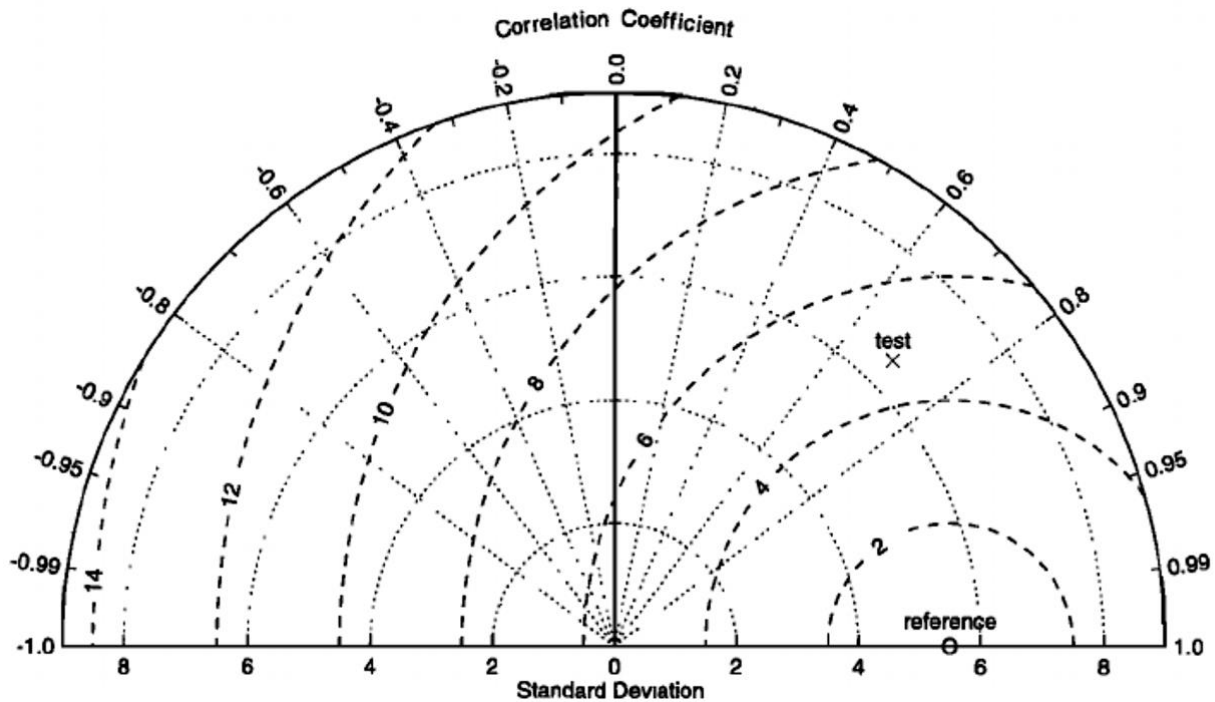
Figure 4: Taylor diagram from Taylor, (2001). The standard deviation of model fields is displayed as the radial distance from the origin and can be visually compared to the observed (reference) point, which is indicated by a circle on the abscissa. The correlation between the model and observed fields decreases with azimuthal angle (dotted lines), and the root-mean-square difference between the model and observed fields is proportional to the distance from the reference point (quantified by dashed contours).

For example, Taylor diagrams of global and regional NPP by Anav et al. (2015) demonstrated consistent low correlation and high standard deviation for model estimates in the tropics which is substantially reduced in the extratopics and globally, warranting focus on tropical NPP. The validation process of terrestrial biogeochemical cycles and dissection of model uncertainty may also be enhanced through offline simulations or models with intermediate complexity as these allow for a greater replication of simulations with different initializations, forcing datasets, and model configurations, due to their computational affordability (Bonan et al., 2019; Umair et al., 2018; Orth et al., 2017). Offline simulations also reduce the potential for incidental compounding error from coupling components, though this leads to an underestimation in uncertainty for equivalent fully coupled simulations. Replicate simulations with different initial conditions allow for the attribution of uncertainty from unforced variability, such as performed by Danabasoglu et al. (2020), which accounted for half of the inter-model spread in key variables previously (Deser et al., 2020; Eyring et al., 2019). In addition, replicate simulations with different forcing datasets can indicate the role of forcing uncertainty (Wei et al., 2018), which Lawrence et al. (2019) found to be significant. Further, sensitivity analyses or perturbed parameter analyses involving replicated simulations with one or more variables fixed as performed by Hajima et al. (2020) and Lawrence et al. (2019) illuminate structural uncertainty. The use of well-established statistical and model performance metrics in addition to strategic simulations facilitates a detailed analysis of model uncertainty.

4.4 Moving Forward

A model can only be expected to perform well in simulating past, present, and future conditions if provided with high quality observational constraints. Lovenduski and Bonan (2017) suggest that obtaining accurate observations and improving process understanding should take precedence over reducing model spread, as constraining models to uncertain observations does not improve their predictive capacity, and even models which agree well with observations can prompt divergent projections. Several of the challenges inherent in implementing observations in model validation and development are now a key focus of the Observations for Model Intercomparison Project (obs4MIPs; Waliser et al., 2020) which strives to deliver long-term, high-quality observations from international efforts. An obs4MIPs meeting held in preparation for CMIP6 with more than 50 satellite data and global climate modelling experts identified underutilized observation products and recommended new efforts to address knowledge gaps, including an expanded inventory of datasets, higher-frequency datasets and model output, more reliable uncertainty measures, more datasets tailored to offline simulations, and more explicit metadata for modellers (Waliser et al., 2020). Further, recent satellite missions such as the Sentinel2A twin satellite launched in 2015 have unprecedented spectral, spatial, and temporal resolution combinations, which can be used alone or in combination with other satellite-based observations to provide higher fidelity references for validation (Vafaei et al., 2018). Field experimental data provide unique insight as to the functional responses of vegetation to elevated CO_2 concentration (Goll et al., 2017), temperature change (Richardson et al., 2018), moisture availability (Williams et al., 2019; Hovenden and Newton, 2018), and nutrient limitations (Fleischer et al., 2019), outside the current context of observations. The integration of experimental findings in evaluations is challenging given the environmentally rapid application of treatments and limited ecological representation (Nowak et al., 2004), though sophisticated relationship-based techniques such as used by Goll et al. (2017) alleviate some of these issues. Increased collaboration between field and model researchers in designing experiments could improve the applicability of future experiments. In addition, enhanced field and remote sensing collaboration would allow for higher fidelity calibrated global data products (Orth et al., 2017; Verger et al., 2016). Thus future CMIPs will benefit from forthcoming collaborations and reference data products tailored for validation.

A standard protocol for the validation of terrestrial biogeochemical variables would facilitate a thorough and objective assessment of model performance within and among participants. Further, the collective merits and limitations of the current variety of approaches utilized by participants could be consolidated and addressed in a comprehensive protocol. In the interest of model improvement and weighting for predictions, validation with an exhaustive assessment of variables across a range of spatiotemporal scales against all available peer-recommended observation-based references is optimal. Dataset-specific expertise is also warranted to correctly implement reference datasets in these evaluations (Waliser et al., 2020; Liu et al., 2018). The procurement and application of reference datasets within validations is demanding for participants, considering their presiding obligation to continuously refine model components and participate in CMIP with computationally expensive ESM simulations. Additionally, the universal inclusion of often overlooked processes such as moisture limitation, nitrogen and phosphorus cycles, dynamic vegetation, prognostic leaf phenology, and natural disturbance regimes should be a priority focus

for participants in developing diagnostic models as these processes are highly influential on terrestrial biogeochemistry and physics (Eyring et al., 2020; Fleisher et al., 2019; Piao et al., 2019; Wieder et al., 2015; Achard et al., 2014; Richardson et al., 2013; Heimann and Reichstein, 2008; Tucker et al., 1986), and their omission contributes to widespread bias (Green et al., 2019; Anav et al., 2015). While outside the focus of this review, equal attention should be applied to the physical components of terrestrial biogeochemical cycles, including explicit representation of permafrost and riverine carbon transport dynamics. In fact, a study including four CMIP5 ESMs found that soil moisture variability prompted variability in terrestrial NBP on the order of gigatonnes, with non-linear responses to both moisture scarcity and excess (Green et al., 2019). Further, many of the merits and limitations of the validation approaches discussed herein apply to the validation of these physical components as well.

The communal use of software packages such as ESMValToolv2.0 and ILAMBv2.1 (Eyring et al., 2020; Collier et al., 2018) could liberate time and computational resources for modellers. In addition, this would standardize validation protocols, address long-overlooked model uncertainty distinctions (Deser et al., 2020), and avoid terminology confusion (Lovett et al., 2006). While these packages include extensive suites of peer-verified observational reference datasets and performance metrics, these packages do not yet include evaluation of nitrogen and phosphorus cycles, which may be due to the combined scarcity of observations, upscaling approaches, and model representations (Lawrence et al., 2019; Zhu et al., 2018; Wieder et al., 2015; Zaehle and Dalmonech, 2011). The strategic situation of nitrogen, phosphorus, and soil moisture monitoring which coincides with current Fluxnet sites (Jung et al., 2020) could provide high fidelity insight as to nutrient and environmental limitations on GPP, coherent turnover time assessments, and broadly applicable functional relationships to facilitate upscaling. The co-situation of multiple observational monitoring objectives at Fluxnet sites would enhance the utility of each site-level dataset and alleviate errors due to spatiotemporal inconsistencies between datasets in both performing evaluations and developing large scale data products. Following increased collaboration between empirical and modelling communities to strategically expand observations, and their inclusion in a comprehensive evaluation software, the CMIP-designated use of such software would standardize, conserve, and augment validation efforts.

5. Conclusion

The current generation of ESMs which participated in the sixth phase of the Coupled Model Intercomparison Project adopted a broad assortment of approaches to validate historically simulated terrestrial biogeochemical cycles. Validations which encompassed a large suite of variables over a range of spatiotemporal scales in conjunction with informative model performance metrics demonstrated relatively comprehensive assessments of model performance. Across CMIP6 participants, the variety of variables, reference datasets, evaluation dimensions, and statistical metrics utilized make general assessments of model performance in simulating terrestrial biogeochemistry challenging. To address this inconsistency and alleviate the immense responsibilities of participants, we recommend the designation of a standard validation protocol for CMIP participants, which is consolidated in an open-source software (such as the Earth System Model Evaluation Tool version 2

(ESMValToolv2.0) or the International Land Model Benchmarking version 2.1 (ILAMBv2.1)). This protocol should utilize a comprehensive suite of certainty-weighted observational reference datasets, targeted model performance metrics, and comparisons across a range of spatiotemporal dimensions. The insights from a universally adopted validation protocol would precisely attribute model uncertainty and aid in directing future observational efforts to improve crucial process understanding

450 within terrestrial biogeochemical cycles.

Data availability

The data used to generate figures 1 and 2 are openly available from the literature.

Author contributions

455 LS and AHMD both initiated the research and significantly contributed to the writing of the paper. LS conducted the analysis and wrote the original draft. AHMD provided supervisory support.

Competing interests

All authors declare no competing interests.

Acknowledgements

460 We wish to thank Drs. Susan Ziegler, Hugo Beltrami, Entcho Demirov, Joel Finnis, and Evan Edinger for their insightful comments on an early draft of this manuscript. This research was supported by the National Science and Engineering Research Council of Canada (NSERC) Discovery Grant Program. LS is grateful for support from a NSERC Canada Graduate Scholarships - Doctoral Program Scholarship.

References

Achard, F., Beuchle, R., Mayaux, P., Stibig, H. J., Bodart, C., Brink, A., and Simonetti, D.: Determination of tropical deforestation rates and related carbon losses from 1990 to 2010, Glob Change Biol, 20, 2540-2554, https://doi.org/10.1111/gcb.12605, 2014.

470   Amthor, J. S.: The McCree–de Wit–Penning de Vries–Thornley respiration paradigms: 30 years later, Ann Bot-London, 86, 1– 20, https://doi.org/10.1006/anbo.2000.1175, 2000.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., and Zhu, Z.: Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models, J Climate, 26, 6801-6843,
475   https://doi.org/10.1175/JCLI-D-12-00417.1, 2013.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Rev Geophys, 53, 785-818, https://doi.org/10.1002/2015RG000483, 2015.

480   Anderson, T. R., Hawkins, E., and Jones, P. D.: CO_2, the greenhouse effect and global warming: from the pioneering work of Arrhenius and Callendar to today's Earth System Models, Endeavour, 40, 178-187, https://doi.org/10.1016/j.endeavour.2016.07.002, 2016.

Arora, V. K., and Boer, G. J.: A parameterization of leaf phenology for the terrestrial ecosystem component of climate
485   models, Glob Change Biol, 11, 39-59, https://doi.org/10.1111/j.1365-2486.2004.00890.x, 2005.

Arora, V. K., and Boer, G. J.: Uncertainties in the 20th century carbon budget associated with land use change, Glob Change Biol, 16, 3327-3348, https://doi.org/10.1111/j.1365-2486.2010.02202.x, 2010.

490   Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon–concentration and carbon–climate feedbacks in CMIP5 Earth system models, J Climate, 26, 5289-5314, https://doi.org/10.1175/JCLI-D-12-00494.1, 2013.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher,
495   O., Cadule, P., Chamberlain, M., Christian, J., Delire, C., Fisher, R., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C., Krasting, J., Law, R., Lawrence, D., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models, Biogeosciences, 17, 4173-4222, https://doi.org/10.5194/bg-17-4173-2020, 2020.

500   Avitabile, V., Herold, M., Heuvelink, G. B. M., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., Bayol, N., Berry, N. J., Boeckx, P., de Jong, B. H. J., DeVries, B., Girardin, C. A. J., Kearsley, E., Lindsell, J. A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E. T. A., Nagy, L., Qie, L., Quinones, M. J., Ryan, C. M., Ferry, S. J. F., Sunderland, T., Laurin, G. V., Gatti, R. C., Valentini, R., Verbeeck, H., Wijaya, A., and Willcock, S.: An integrated pan-

tropical biomass map using multiple reference datasets, Glob Change Biol, 22, 1406-1420, https://doi.org/10.1111/gcb.13139, 2016.

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., and Houghton, R. A.: Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps, Nat Clim Change, 2, 182-185, https://doi.org/10.1038/nclimate1354, 2012.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., and Leroy, M.: LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. Remote Sens Environ, 110, 275-286, https://doi.org/10.1016/j.rse.2007.02.018, 2007.

Batjes, N. H.: Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks, Geoderma, 269, 61-68, https://doi.org/10.1016/j.geoderma.2016.01.034, 2016.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science, 329, 834-838, https://doi.org/10.1126/science.1184984, 2010.

Blackard, J. A., Finco, M. V., Helmer, E. H., Holden, G. R., Hoppus, M. L., Jacobs, D. M., Lister, A. J., Moisen, G. G., Nelson, M. D., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D. L., Winterberger, K. C., Brandeis, T. J., Czaplewski, R. L., McRoberts, R. E., Patterson, P. L., and Tymcio, R. P.: Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information, Remote Sens Environ, 112, 1658–1677, https://doi.org/10.1016/j.rse.2007.08.021, 2008.

Bonan, G. B., Lombardozzi, D. L., Wieder, W. R., Oleson, K. W., Lawrence, D. M., Hoffman, F. M., and Collier, N.: Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014), Global Biogeochem Cy, 33, 1310-1326, https://doi.org/10.1029/2019GB006175, 2019.

Bond-Lamberty, B., Bailey, V. L., Chen, M., Gough, C. M., and Vargas, R.: Globally rising soil heterotrophic respiration over recent decades, Nature, 560, 80-83, https://doi.org/10.1038/s41586-018-0358-x, 2018.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, L., Guilyardi, É., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Khadre Traore, A., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and evaluation of the IPSL-CM6A-LR climate model, J Adv Model Earth Sy, 12, e2019MS002010, https://doi.org/10.1029/2019MS002010, 2020.

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Mu, M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J. T., and Reichstein, M.: Global covariation of carbon turnover times with climate in terrestrial ecosystems, Nature, 514, 213-217, https://doi.org/10.1038/nature13731, 2014.

Chu, H., Baldocchi, D. D., John, R., Wolf, S., and Reichstein, M.: Fluxes all of the time? A primer on the temporal representativeness of Fluxnet, J Geophys Res-Biogeo, 122, 289-307, https://doi.org/10.1002/2016JG003576, 2017.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, available at: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter06_FINAL.pdf 2013.

Ciais, P., Tan, J., Wang, X., Roedenbeck, C., Chevallier, F., Piao, S.-L., Moriarty, R., Broquet, G., Le Quéré, C., Canadell, J. G., Peng, S., Poulter, B., Liu, Z., and Tans, P.: Five decades of northern land carbon uptake revealed by the interhemispheric CO_2 gradient, Nature, 568, 221– 225, https://doi.org/10.1038/s41586-019-1078-6, 2019.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description–Part 2: carbon fluxes and vegetation dynamics, Geosci Model Dev, 4, 701-722, https://doi.org/10.5194/gmd-4-701-2011, 2011.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., and Randerson, J. T.: The International Land Model Benchmarking (ILAMB) system: design, theory, and implementation, J Adv Model Earth Sy, 10, 2731-2754, https://doi.org/10.1029/2018MS001354, 2018.

575 Dai, M., Yin, Z., Meng, F., Liu, Q., and Cai, W. J.: Spatial distribution of riverine DOC inputs to the ocean: an updated global synthesis. Curr Opin Env Sust, 4, 170-178, https://doi.org/10.1016/j.cosust.2012.03.003, 2012.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L.K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M., Large, W., Lauritzen, P., Lawrence, D., Lenaerts, J., Lindsay, K.,
580 Lipscomb, W., Mills M. J., Neale, R., Oleson, K., Otto-Bliesner, B., Phillips, A., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J., Kinnison, D., Kushner, P., Larson, V., Long, M., Mickelson, S., Moore, J., Nienhouse, E., Polvani, L., Rasch, P., and Strand, W.: The community earth system model version 2 (CESM2). J Adv Model Earth Sy, 12, https://doi.org/10.1029/2019MS001916, 2020.

585 Davies-Barnard, T., Meyerholt, J., Zaehle, S., Friedlingstein, P., Brovkin, V., Fan, Y., Wiltshire, A. J.: Nitrogen cycling in CMIP6 land surface models: progress and limitations. Biogeosciences, 17, 5129-5148, https://doi.org/10.5194/bg-17-5129-2020, 2020.

Defourny, P., Boettcher, M., Bontemps, S., Kirches, G., Lamarche, C., Peters, M., Santoro, M., and Schlerf, M.: Land cover
590 CCI Product user guide version 2, Technical report, European Space Agency, 2016.

de Kauwe, M. G., Disney, M. I., Quaife, T., Lewis, P., and Williams, M.: An assessment of the MODIS Collection 5 leaf area index product for a region of mixed coniferous forest, Remote Sens Environ, 115, 767– 780, https://doi.org/10.1016/j.rse.2010.11.004, 2011.
595
Delire, C., Séférian, R., Decharme, B., Alkama, R., Calvet, J. C., Carrer, D., Gibelin, A., Joetzjer, E., Morel, X., Rochner, M., and Tzanos, D.: The global land carbon cycle simulated with ISBA-CTRIP: Improvements over the last decade, J Adv Model Earth Sy, 12, e2019MS001886, https://doi.org/10.1029/2019MS001886, 2020.

600 de Mora, L. D., Butenschön, M., and Allen, J. I.: How should sparse marine in situ measurements be compared to a continuous model: An example, Geosci Model Dev, 6, 533-548, https://doi.org/10.5194/gmd-6-533-2013, 2013.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., and Stauch, V. J.: Cross-site evaluation of eddy covariance GPP and RE decomposition techniques, Agr Forest Meteorol, 148, 821-838, https://doi.org/10.1016/j.agrformet.2007.11.012, 2008.

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., Ting, M.: Insights from Earth system model initial-condition large ensembles and future prospects, Nat Clim Change, 10, 277-286, https://doi.org/10.1038/s41558-020-0731-2, 2020.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik1, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics, J Adv Model Earth Sy, 12, e2019MS002015, https://doi.org/10.1029/2019MS002015, 2020.

Ehlers, I., Augusti, A., Betson, T. R., Nilsson, M. B., Marshall, J. D., and Schleucher, J.: Detecting long-term metabolic shifts using isotopomers: CO_2-driven suppression of photorespiration in C3 plants over the 20th century, P Natl Acad Sci USA, 112, 15585-15590, https://doi.org/10.1073/pnas.1504493112, 2015.

Erkkilä, K. M., Ojala, A., Bastviken, D., Biermann, T., Heiskanen, J. J., Lindroth, A., and Mammarella, I.: Methane and carbon dioxide fluxes over a lake: comparison between eddy covariance, floating chambers and boundary layer method, Biogeosciences, 15, 429-445, https://doi.org/10.5194/bg-15-429-2018, 2018.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci Model Dev, 9, 1937-1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016a.

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K.-D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., Ulft, L. H. van, Walton, J., Wang, S. and Williams, K. D.: ESMValTool (v1. 0)–a community diagnostic and performance metrics tool for routine validation of Earth system models in CMIP, Geosci Model Dev, 9, 1747-1802, https://doi.org/10.5194/gmd-9-1747-2016, 2016b.

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E.,

640   Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C., Simpson, I. R., Stouffer, R. J., and Williamson, M. S.: Taking climate model validation to the next level. Nat Clim Change, 9, 102-110, https://doi.org/10.1038/s41558-018-0355-y, 2019.

Eyring, V., Bock, L., Lauer, A., Righi, M., Schlund, M., Andela, B., Arnone, E., Bellprat, O., Brötz, B., Caron, L.-P.,

645   Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., Davin, E., Davini, P., Debeire, K., de Mora, L., Deser, C., Docquier, D., Earnshaw, P., Ehbrecht, C., Gier, B. K., Gonzalez-Reviriego, N., Goodman, P., Hagemann, S., Hardiman, S., Hassler, B., Hunter, A., Kadow, C., Kindermann, S., Koirala, S., Koldunov, N., Lejeune, Q., Lembo, V., Lovato, T., Lucarini, V., Massonnet, F., Müller, B., Pandde, A., Pérez-Zanón, N., Phillips, A., Predoi, V., Russell, J., Sellar, A., Serva, F., Stacke, T., Swaminathan, R., Torralba, V., Vegas-Regidor, J., von Hardenberg, J., Weigel, K., and Zimmermann, K.: Earth System Model

650   Evaluation Tool (ESMValTool) v2.0 - an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP, Geosci Model Dev, 13, 3383-3438, https://doi.org/10.5194/gmd-13-3383-2020, 2020.

Fan, J., Chen, B., Wu, L., Zhang, F., Lu, X., and Xiang, Y.: Evaluation and development of temperature-based empirical

655   models for estimating daily global solar radiation in humid regions, Energy, 144, 903-914, https://doi.org/10.1016/j.energy.2017.12.091, 2018.

Fan, N., Koirala, S., Reichstein, M., Thurner, M., Avitabile, V., Santoro, M., and Carvalhais, N.: Apparent ecosystem carbon turnover time: uncertainties and robust features, Earth Syst Sci Data, 12, 2517-2536, https://doi.org/10.5194/essd-12-2517-

660   2020, 2020.

FAO.: Harmonized World Soil Database v 1.2, Retrieved January 29, 2021, from http://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/, 2012.

665   Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., and Lawrence, D. M.: Parametric controls on vegetation responses to biogeochemical forcing in the CLM5, J Adv Model Earth Sy, 11, 2879-2895, https://doi.org/10.1029/2019MS001609, 2019.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V.,

670   Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of climate models,

In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

675

Fleischer, K., Rammig, A., De Kauwe, M. G., Walker, A. P., Domingues, T. F., Fuchslueger, L., Lapola, D. M.: Amazon forest response to CO_2 fertilization dependent on plant phosphorus acquisition, Nat Geosci, 12, 736-741, https://doi.org/10.1038/s41561-019-0404-9, 2019.

680 Fowler, D., Coyle, M., Skiba, U., Sutton, M. A., Cape, J. N., Reis, S., Sheppard, L. J., Jenkins, A., Grizzetti, B., Galloway, J. N., Vitousek, P., Leach, A., Bouwman, A. F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M., and Voss, M.: The global nitrogen cycle in the twenty-first century, Philos T R Soc B, 368, 20130164–20130164, https://doi.org/10.1098/rstb.2013.0164, 2013.

685 Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, 690 P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global carbon budget 2019, Earth Syst Sci Data, 11, 1783-1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

695

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, 700 K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan,

705   W., Yue, X., and Zaehle, S.: Global carbon budget 2020, Earth Syst Sci Data, 12, 3269-3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., and Vo, C. J.: Nitrogen cycles:
710   past, present, and future, Biogeochemistry, 70, 153-226, https://doi.org/10.1007/s10533-004-0370-0, 2004.

Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., and Sutton, M. A.: Transformation of the nitrogen cycle: recent trends, questions, and potential solutions, Science, 320, 889-892, https://doi.org/10.1126/science.1136674, 2008.
715

Galloway, J. N., Leach, A. M., Bleeker, A., and Erisman, J. W.: A chronology of human understanding of the nitrogen cycle, Philos T R Soc B, 368, 20130120, https://doi.org/10.1098/rstb.2013.0120, 2013.

Gibbs, H. K.: Olson's Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation: An Updated Database Using
720   the GLC2000 Land Cover Product NDP-017b, Oak Ridge National Laboratory, Oak Ridge, TN, https://doi.org/10.3334/CDIAC/lue.ndp017.2006, 2006.

Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, Biogeosciences, 7, 1171– 1186,
725   https://doi.org/10.5194/bg-7-1171-2010, 2010.

Gleckler, P. J., Doutriaux, C., Durack, P. J., Taylor, K. E., Zhang, Y., Williams, D. N., and Servonnat, J.: A more powerful reality test for climate models, EOS, 97, 20-24, available at: https://eos.org/science-updates/a-more-powerful-reality-test-for-climate-models, 2016.
730

Global Monitoring Laboratory.: Global monitoring Laboratory - carbon cycle greenhouse gases, available at: https://www.esrl.noaa.gov/gmd/ccgg/trends/, 2005.

Global Soil Data Task Group.: Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS), Tech. rep., available
735   at: https://doi.org/10.3334/ORNLDAAC/569, 2002.

GLOBAL VIEW-CO$_2$.: Cooperative Global Atmospheric Data Integration Project, updated annually, Multi-laboratory compilation of synchronized and gap-filled atmospheric carbon dioxide records for the period 1979-2012, NOAA, Boulder, CO, https://doi.org/10.3334/OBSPACK/1002, 2013.

740

Goll, D. S., Brovkin, V., Liski, J., Raddatz, T., Thum, T., and Todd-Brown, K. E.: Strong dependence of CO$_2$ emissions from anthropogenic land cover change on initial land cover and soil carbon parametrization, Global Biogeochem Cy, 29, 1511-1523, https://doi.org/10.1002/2014GB004988, 2015.

745   Goll, D. S., Winkler, A. J., Raddatz, T., Dong, N., Prentice, I. C., Ciais, P., and Brovkin, V.: Carbon–nitrogen interactions in idealized simulations with JSBACH (version 3.10), Geosci Model Dev, 10, 2009-2030, https://doi.org/10.5194/gmd-10-2009-2017, 2017.

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., and Gentine, P.: Large influence
750   of soil moisture on long-term terrestrial carbon uptake, Nature, 565, 476-479, https://doi.org/10.1038/s41586-018-0848-x, 2019.

Gruber, N., and Galloway, J. N.: An Earth-system perspective of the global nitrogen cycle, Nature, 451, 293-296, https://doi.org/10.1038/nature06592, 2008.

755

Gulden, L. E., Rosero, E., Yang, Z. L., Wagener, T., and Niu, G. Y.: Model performance, model robustness, and model fitness scores: A new method for identifying good land-surface models, Geophys Res Lett, 35, https://doi.org/10.1029/2008GL033721, 2008.

760   Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, Geosci Model Dev, 13, 2197-2244, https://doi.org/10.5194/gmd-13-2197-2020, 2020.

765   Harper, A. B., Wiltshire, A. J., Cox, P. M., Friedlingstein, P., Jones, C. D., Mercado, L. M., Sitch, S., Williams, K., and Duran-Rojas, C.: Vegetation distribution and terrestrial carbon cycle in a carbon cycle configuration of JULES4. 6 with new plant functional types, Geosci Model Dev, 11, 2857-2873, https://doi.org/10.5194/gmd-11-2857-2018, 2018.

Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K., and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12, 4121-4132, https://doi.org/10.5194/bg-12-4121-2015, 2015.

He, Y., Piao, S. L., Li, X. Y., Chen, A. P., and Qin, D. H.: Global patterns of vegetation carbon use efficiency and their climate drivers deduced from MODIS satellite data and process-based models, Agr Forest Meteorol, 256–257, 150– 158, https://doi.org/10.1016/j.agrformet.2018.03.009, 2018.

Heimann, M., and Reichstein, M.: Terrestrial ecosystem carbon dynamics and climate feedbacks, Nature, 451, 289-292, https://doi.org/10.1038/nature06591, 2008.

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

Hoffman, F. M., Randerson, J. T., Arora, V. K., Bao, Q., Cadule, P., Ji, D., Jones, C.D., Kawamiya, M., Khatiwala, S., Lindsay, K., and Wu, T.: Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models, J Geophys Res-Biogeo, 119, 141-162, https://doi.org/10.1002/2013JG002381, 2014.

Holland, E. A., Post, W. M., Matthews, E., Sulzman, J. M., Staufer, R., and Krankina, O. N.: A global database of litterfall mass and litter pool carbon and nutrients, ORNL DAAC, available at: https://daac.ornl.gov/VEGETATION/guides/Global_Litter_Carbon_Nutrients.html, 2015.

Houlton, B. Z., Marklein, A. R., and Bai, E.: Representation of nitrogen in climate change forecasts, Nat Clim Change, 5, 398-401, https://doi.org/10.1038/nclimate2538, 2015.

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., Jam, A., Musat, I., Idelkadi, A., Fairhead, L., Foujols, M.-A., Mellul, L., Traore, A.-K., Ghattas, J., Gastineau, G., Dufresne, J.-L., Boucher, O., Lefebvre, M.-P., Millour, E., Vignon, E., Jouaud, J., Bint Diallo, F., Bonazzola, M. and Lott, F.: LMDZ6: Improved atmospheric component of the IPSL coupled model, J. Adv. Model. Earth Sy, 12, e2019MS001892, https://doi.org/10.1029/2019MS001892, 2020.

Hovenden, M., and Newton, P.: Plant responses to CO2 are a question of time, Science, 360, 263-264, https://doi.org/10.1126/science.aat2481, 2018.

Hugelius, G., Bockheim, J. G., Camill, P., Elberling, B., Grosse, G., Harden, J. W., Johnson, K., Jorgenson, T., Koven, C. D., Kuhry, P., Michaelson, G., Mishra, U., Palmtag, J., Ping, C.-L., O'Donnell, J., Schirrmeister, L., Schuur, E. A. G., Sheng, Y., Smith, L. C., Strauss, J., and Yu, Z.: A new data set for estimating organic carbon storage to 3 m depth in soils of the northern circumpolar permafrost region, Earth Syst Sci Data, 5, 393-402, https://doi.org/10.5194/essd-5-393-2013, 2013.

Huntzinger, D. N., Schwalm, C., Michalak, A. M., Schaefer, K., King, A. W., Wei, Y., Jacobson, A., Liu, S., Cook, R. B., Post, W. M., Berthier, G., Hayes, D., Huang, M., Ito, A., Lei, H., Lu, C., Mao, J., Peng, C. H., Peng, S., Poulter, B., Riccuito, D., Shi, X., Tian, H., Wang, W., Zeng, N., Zhao, F., and Zhu, Q.: The north american carbon program multi-scale synthesis and terrestrial model intercomparison project–part 1: Overview and experimental design, Geosci Model Dev, 6, 2121-2133, https://doi.org/10.5194/gmd-6-2121-2013, 2013.

Ito, A.: A historical meta-analysis of global terrestrial net primary productivity: are estimates converging?, Glob Change Biol, 17, 3161-3175, https://doi.org/10.1111/j.1365-2486.2011.02450.x, 2011.

Ito, A., and Inatomi, M.: Use of a process-based model for assessing the methane budgets of global terrestrial ecosystems and validation of uncertainty, Biogeosciences, 9, 759-773, https://doi.org/10.5194/bg-9-759-2012, 2012.

Ito, A., Hajima, T., Lawrence, D. M., Brovkin, V., Delire, C., Guenet, B., Jones, C., Malyshev, S., Materia, S., McDermid, S., Peano, D., Pongratz, J., Robertson, E., Shevliakova, E., Vuichard, N., Warlind, D., Wiltshire, A., and Ziehn, T.: Soil carbon sequestration simulated in CMIP6-LUMIP models: implications for climatic mitigation, Environ Res Lett, 15, 124061, https://doi.org/10.1088/1748-9326/abc912, 2020.

Joetzjer, E., Delire, C., Douville, H., Ciais, P., Decharme, B., Carrer, D., Verbeeck, H., de Weirdt, M., and Bonal, D.: Improving the ISBA CC land surface model simulation of water and carbon fluxes and stocks over the Amazon forest, Geosci Model Dev, 8, 1709-1727, https://doi.org/10.5194/gmd-8-1709-2015, 2015.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J.G., and Zaehle, S.: C4MIP–The coupled climate–carbon cycle model intercomparison project: Experimental protocol for CMIP6, Geosci Model Dev, 9, 2853-2880, https://doi.org/10.5194/gmd-9-2853-2016, 2016.

Jonsson, A., Åberg, J., Lindroth, A., and Jansson, M.: Gas transfer rate and CO_2 flux between an unproductive lake and the atmosphere in northern Sweden, J Geophys Res-Biogeo, 113, https://doi.org/10.1029/2008JG000688, 2008.

835   Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of Fluxnet eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6, 2001-2013, https://doi.org/10.5194/bg-6-2001-2009, 2009.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., De Jeu,
840   R., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951-954, https://doi.org/10.1038/nature09396, 2010.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L.,
845   Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J Geophys Res-Biogeo, 116, https://doi.org/10.1029/2010JG001566, 2011.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P.,
850   Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rödenbeck, C., Tramontana, G., Viovy, N., Wang, Y.-P., Weber, U., Zaehle, S., and Zeng, N.: FLUXCOM (RS+METEO) Global Land Carbon Fluxes using CRUNCEP climate data. FLUXCOM Data Portal, https://doi.org/10.17871/FLUXCOM_RS_METEO_CRUNCEPv6_1980_2013_v1, 2016.

855   Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rödenbeck, C., Tramontana, G., Viovy, N., Wang, Y., Weber, U., Zaehle, S., and Zeng, N.: Compensatory water effects link yearly global land CO_2 sink changes to temperature, Nature, 541, 516– 520, https://doi.org/10.1038/nature20780, 2017.

860   Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A. K., Liu, J., Lombardozzi, D., Nabel, J. E. M. S., Nelson, J. A., O'Sullivan, M., Pallandt, M., Papale, D., Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U., and Reichstein, M.: Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach, Biogeosciences, 17, 1343-1365, https://doi.org/10.5194/bg-17-1343-
865   2020, 2020.

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

870

Kellndorfer, J., Walker, W., Kirsch, K., Fiske, G., Bishop, J., Lapoint, L., Hoppus, M., Westfall, J.: NACP aboveground biomass and carbon baseline data, V.2 (NBCD 2000), U.S.A., 2000, http://dx.doi.org/10.3334/ORNLDAAC/1161, 2013.

Kindermann, G., McCallum, I., Fritz, S., and Obersteiner, M.: A global forest growing stock, biomass and carbon map based
875 on FAO statistics, Silva Fenn, 42, 387-396, https://doi.org/10.14214/sf.244, 2008.

Kobayashi, K., and Salam, M. U.: Comparing simulated and measured values using mean squared deviation and its components, Agron J, 92, 345-352, https://doi.org/10.2134/agronj2000.922345x, 2000.

880 Koch, J., Siemann, A., Stisen, S., and Sheffield, J.: Spatial validation of large-scale land surface models against monthly land surface temperature patterns using innovative performance metrics, J Geophys Res-Atmos, 121, 5430-5452, https://doi.org/10.1002/2015JD024482, 2016.

Koven, C. D., Hugelius, G., Lawrence, D. M., and Wieder, W. R.: Higher climatological temperature sensitivity of soil carbon
885 in cold than warm climates, Nat Clim Change, 7, 817-822, https://doi.org/10.1038/nclimate3421, 2017.

Kowalczyk, E. A., Stevens, L., Law, R. M., Dix, M., Wang, Y. P., Harman, I. N., Haynes, K., Srbinovsky, J., Pak, B., and Ziehn, T.: The land surface model component of ACCESS: description and impact on the simulated surface climatology, Aust. Meteorol. Oceanogr. J, 63, 65-82, http://www.bom.gov.au/jshess/docs/2013/kowalczyk_hres.pdf, 2013.

890

Kowalczyk, E. A., Wang, Y. P., Law, R. M., Davies, H. L., McGregor, J. L., and Abramowitz, G.: The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model, CSIRO Marine and Atmospheric Research Paper, 13, 42, http://www.cmar.csiro.au/e-print/open/kowalczykea_2006a.pdf, 2006.

895 Kumar, S. V., Peters-Lidard, C. D., Santanello, J., Harrison, K., Liu, Y., and Shaw, M.: Land surface Verification Toolkit (LVT)–a generalized framework for land surface model evaluation, Geosci Model Dev, 5, 869-886, https://doi.org/10.5194/gmd-5-869-2012, 2012.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen,
900 B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald,

N., McConnell, J. R., Naik, V., Riahi, K., and Van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, Atmos Chem Phys, 10, 4963– 5019, https://doi.org/10.5194/acp-10-7017-2010, 2010.

905 Lasslop, G., Reichstein, M., Kattge, J., and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, Biogeosciences, 5, 1311-1324, https://doi.org/10.5194/bg-5-1311-2008, 2008.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., Stoy, P., Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: Critical issues and global

910 evaluation, Glob Change Biol, 16, 187– 208, https://doi.org/10.1111/j.1365-2486.2009.02041.x, 2010.

Law, K., Stuart, A., and Zygalakis, K.: Data assimilation: A Mathematical Introduction, Texts in Applied Mathematics, Cham, Switzerland: Springer, 141, 242, https://doi.org/10.1007/978-3-319-20325-6, 2015.

915 Lawrence, P. J., Feddema, J. J., Bonan, G. B., Meehl, G. A., O'Neill, B. C., Oleson, K. W., Levis, S., Lawrence, D.M., Kluzek, E., Lindsay, K., and Thornton, P. E.: 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, J Climate, 25, 3071-3095, https://doi.org/10.1175/JCLI-D-11-00256.1, 2012.

920 Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., and Zeng, X.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty, J Adv Model Earth Sy, 11, 4245-4287, https://doi.org/10.1029/2018MS001583, 2019.

925 Le Quéré, C., Peters, GP., Andres, RJ., Andrew, RM., Boden, TA., Ciais, P., Friedlingstein, P., Houghton, RA., Marland, G., Moriarty, R., Sitch, S., Tans, P., Arneth, A., Arvanitis, A., Bakker, DCE., Bopp, L., Canadell, JG., Chini, LP., Doney, SC., Harper, A., Harris, I., House, JI., Jain, AK., Jones, SD., Kato, E., Keeling, RF., Klein, Goldewijk, K., Körtzinger, A., Koven, C., Lefèvre, N., Maignan, F., Omar, A., Ono, T., Park, G-H., Pfeil, B., Poulter, B., Raupach, MR., Regnier, P., Rödenbeck, C., Saito, S., Schwinger, J., Segschneider, J., Stocker, BD., Takahashi, T., Tilbrook, B., van, Heuven, S., Viovy, N., Wanninkhof,

930 R., Wiltshire, A., and Zaehle, S.: Global carbon budget 2013, Earth Syst Sci Data, 6, 235–263, https://doi.org/10.5194/essd-6-235-2014, 2014.

Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P., Manning, A. C., Boden, T. A., Tans, P. P., Houghton, R. A., Keeling, R. F., Alin, S., Andrews, O. D., Anthoni, P., Barbero, L., Bopp, L., Chevallier, F., Chini, L.

935 P., Ciais, P., Currie, K., Delire, C., Doney, S. C., Friedlingstein, P., Gkritzalis, T., Harris, I., Hauck, J., Haverd, V., Hoppema, M., Klein Goldewijk, K., Jain, A. K., Kato, E., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Melton, J. R., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., O'Brien, K., Olsen, A., Omar, A. M., Ono, T., Pierrot, D., Poulter, B., Rödenbeck, C., Salisbury, J., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Sutton, A. J., Takahashi, T., Tian, H., Tilbrook, B., van der Laan-Luijkx, I. T., van

940 der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., and Zaehle, S.: Global carbon budget 2016, Earth Syst Sci Data, 8, 605-649, https://doi.org/10.5194/essd-8-605-2016, 2016.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney,

945 S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook,

950 B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global carbon budget 2018, Earth Syst Sci Data, 10, 2141-2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

Li, W., Zhang, Y., Shi, X., Zhou, W., Huang, A., Mu, M., Qiu, B., and Ji, J.: Development of land surface model

955 BCC_AVIM2.0 and its preliminary performance in LS3MIP/CMIP6, J Meteorol Res-Prc, 33, 851-869, https://doi.org/10.1007/s13351-019-9016-y, 2019.

Liang, J., Qi, X., Souza, L., and Luo, Y.: Processes regulating progressive nitrogen limitation under elevated carbon dioxide: a meta-analysis, Biogeosciences, 13, 2689-2699, https://doi.org/10.5194/bg-13-2689-2016, 2016.

960

Liu, Y. Y., Van Dijk, A. I., De Jeu, R. A., Canadell, J. G., McCabe, M. F., Evans, J. P., and Wang, G.: Recent reversal in loss of global terrestrial biomass, Nat Clim Change, 5, 470-474, https://doi.org/10.1038/nclimate2581, 2015.

Liu, Y., Xiao, J., Ju, W., Zhu, G., Wu, X., Fan, W., and Zhou, Y.: Satellite-derived LAI products exhibit large discrepancies

965 and can lead to substantial uncertainty in simulated carbon and water fluxes, Remote Sens Environ, 206, 174-188, https://doi.org/10.1016/j.rse.2017.12.024, 2018.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. W. M. J., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, Int J Remote Sens, 21, 1303-1330, https://doi.org/10.1080/014311600210191, 2000.

970

Lovenduski, N. S., and Bonan, G. B.: Reducing uncertainty in projections of terrestrial carbon uptake, Environ Res Lett, 12, 044020, https://doi.org/10.1088/1748-9326/aa66b8, 2017.

975 Lovett, G. M., Cole, J. J., and Pace, M. L.: Is net ecosystem production equal to ecosystem carbon accumulation?, Ecosystems, 9, 152-155, https://doi.org/10.1007/s10021-005-0036-3, 2006.

Mack, P. E.: Viewing the Earth: The social construction of the Landsat satellite system, MIT Press, https://books.google.ca/books?id=Pk7WtI2MJPgC, 1990.

980

Maki, T., Ikegami, M., Fujita, T., Hirahara, T., Yamada, K., Mori, K., Takeuchi, A., Tsutsumi, Y., Suda, K., and Conway, T. J.: New technique to analyse global distributions of CO_2 concentrations and fluxes from non-processed observational data, Tellus B, 62, 797–809, https://doi.org/10.1111/j.1600-0889.2010.00488.x, 2010.

985 Malhi, Y., Aragao, L. E. O., Metcalfe, D. B., Paiva, R., Quesada, C. A., Almeida, S., Anderson, L., Brando, P., Chamber, J. Q., da Costa, A. C. L., Hutyra, L. R., Oliveira, P., Patino, S., Pyle, E., Robertson, A., and Teixeira, L.: Comprehensive assessment of carbon productivity, allocation and storage in three Amazonian forests, Glob Change Biol, 15, 1255-1274, https://doi.org/10.1111/j.1365-2486.2008.01780.x, 2009.

990 Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., 995 Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.- S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing CO2, J Adv Model Earth Sy, 11, 998-1038, https://doi.org/10.1029/2018MS001400, 2019.

1000

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H. W., Bouwman, A. F., Fekete, B. M., Kroeze, C., and Van Drecht, G.: Global nutrient export from WaterSheds 2 (NEWS 2): model development and implementation, Environ Modell Softw, 25, 837-853, https://doi.org/10.1016/j.envsoft.2010.01.007, 2010.

1005 McGroddy, M. E., Daufresne, T., and Hedin, L. O.: Scaling of C: N: P stoichiometry in forests worldwide: Implications of terrestrial redfield-type ratios, Ecology, 85, 2390-2401, https://doi.org/10.1890/03-0351, 2004.

Mitchell, T. D., and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, International Journal of Climatology: A Journal of the Royal Meteorological Society, 25,

1010 693-712, https://doi.org/10.1002/joc.1181, 2005.

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

1015

Mouillot, F., and Field, C. B.: Fire history and the global carbon budget: A $1° \times 1°$ fire history reconstruction for the 20th century, Glob Change Biol, 11, 398–420, https://doi.org/10.1111/j.1365-2486.2005.00920.x, 2005.

Myneni, R. B., Ramakrishna, R., Nemani, R., and Running, S. W.: Estimation of global leaf area index and absorbed PAR

1020 using radiative transfer models, IEEE Transactions on Geoscience and remote sensing, 35, 1380-1393, https://doi.org/10.1109/36.649788, 1997.

NASA LP DAAC.: MOD17A3 Terra/MODIS net primary production yearly L4 global 1 km, NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota

1025 (https://lpdaac.usgs.gov), https://doi.org/10.5067/ASTER/AST_L1T.003, 2017.

Nowak, R. S., Ellsworth, D. S., and Smith, S. D.: Functional responses of plants to elevated atmospheric CO_2–do photosynthetic and productivity data from FACE experiments support early predictions?, New Phytol, 162, 253-280, https://doi.org/10.1111/j.1469-8137.2004.01033.x, 2004.

1030

Norby, R. J., DeLucia, E. H., Gielen, B., Calfapietra, C., Giardina, C. P., King, J. S., and Oren, R.: Forest response to elevated CO2 is conserved across a broad range of productivity, P Natl Acad Sci USA, 102, 18052-18056, https://doi.org/10.1073/pnas.0509478102, 2005.

1035   Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R.: Terrestrial Ecoregions of the World: A New Map of Life on Earth, BioScience, 51, 933-938, https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2, 2006.

1040   Orth, R., Dutra, E., Trigo, I. F., and Balsamo, G.: Advancing land surface model development with satellite-based Earth observations, Hydrol Earth Syst Sc, 21, 2483-2495, https://doi.org/10.5194/hess-21-2483-2017, 2017.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., Li, Y.: The Fluxnet2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Sci Data, 7, 1-27, https://doi.org/10.1038/s41597-020-0534-3, 2020.

1045

Phillips, L. B., Hansen, A. J., and Flather, C. H.: Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production, Remote Sens Environ, 112, 4381-4392, https://doi.org/10.1016/j.rse.2008.04.012, 2008.

1050   Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Zhu, X.: Plant phenology and global climate change: Current progresses and challenges, Glob Change Biol, 25, 1922-1940, https://doi.org/10.1111/gcb.14619, 2019.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant 1055   functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative, Geosci Model Dev, 8, 2315-2328, https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Randerson, J.T., G.R. van der Werf, L. Giglio, G.J. Collatz, and P.S. Kasibhatla.: Global fire emissions database, version 4.1 (GFEDv4), ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1293, 2017.

1060

Reichler, T., and Kim, J.: How well do coupled models simulate today's climate?, B Am Meteorol Soc, 89, 303-312, https://doi.org/10.1175/BAMS-89-3-303, 2008.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M.: Climate change, phenology, and 1065   phenological control of vegetation feedbacks to the climate system, Agr Forest Meteorol, 169, 156-173, https://doi.org/10.1016/j.agrformet.2012.09.012, 2013.

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Furze, M. E., Seyednasrollah, B., Hanson, P. J.: Ecosystem warming extends vegetation activity but heightens vulnerability to cold temperatures, Nature, 560, 368-371,
1070 https://doi.org/10.1038/s41586-018-0399-1, 2018.

Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., Vegas-Regidor, J., Bock, L., Brötz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Loosveldt Tomas, S., and Zimmermann, K.: Earth System Model Validation Tool (ESMValTool) v2. 0–technical overview, Geosci Model Dev, 13, 1179-1199,
1075 https://doi.org/10.5194/gmd-13-1179-2020, 2020.

Rodda, S. R., Thumaty, K. C., Praveen, M. S. S., Jha, C. S., and Dadhwal, V. K.: Multi-year eddy covariance measurements of net ecosystem exchange in tropical dry deciduous forest of India, Agr Forest Meteorol, 301, 108351, https://doi.org/10.1016/j.agrformet.2021.108351, 2021.
1080
Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., and Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, P Natl Acad Sci USA, 108, 9899-9904, https://doi.org/10.1073/pnas.1019576108, 2011.

1085 Santoro, M., Beaudoin, A., Beer, C., Cartus, O., Fransson, J. B. S., Hall, R. J., Pathe, C., Schmullius, C., Schepaschenko, D., Shvidenko, A., Thurner, M., and Wegmüller, U.: Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR, Remote Sens Environ, 168, 316-334, https://doi.org/10.1073/pnas.1019576108, 2015.

1090 Saugier, B., and Roy, J.: Estimations of global terrestrial productivity: Converging towards a single number?, In Physiological Ecology, Global Terrestrial Productivity, Academic Press, 543-557, https://doi.org/10.1016/B978-012505290-0/50024-7, 2001.

Schlesinger, W. H.: Biogeochemistry: An analysis of global change (2nd ed.), Oxford: Academic,
1095 https://doi.org/10.1017/S0016756898231505, 1997.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Madec, G.: Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, J Adv Model Earth Sy, 11, 4182-4227, https://doi.org/10.1029/2019MS001791, 2019.
1100

Seland, Ø., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schancke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Hafsahl Karset, I. H., Landgren, O., Liakka, J., Onsum Moseid, K., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iverson, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geosci Model Dev, 13, 6165-6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., Zerroukat, M.: UKESM1: Description and evaluation of the UK Earth System Model, J Adv Model Earth Sy, 11, 4513-4558, https://doi.org/10.1029/2019MS001739, 2019.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, J Climate, 19, 3088-3111, https://doi.org/10.1175/JCLI3790.1, 2006.

Shevliakova, E., Malyshev, S., Martinez-Cano, I., Milly, P. C. D., Pacala, S. W., Ginoux, P., Dunne, K. A., Dunne, J. P., Dupius, C., Findell, K., Ghannam, K., Horowitz, L. W., John, J. G., Knutson, T. R., Krasting, J. P., Naik, V., Zadeh, N., Zeng, F., and Zeng, Y.: The land component LM4. 1 of the GFDL Earth System Model ESM4. 1: biophysical and biogeochemical processes and interactions with climate, J Adv Model Earth Syst, 2019MS002040, in review, 2020.

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A.: Mapping forest canopy height globally with spaceborne lidar, J Geophys Res-Biogeo, 116, https://doi.org/10.1029/2011JG001708, 2011.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D. and Winter, B.: The Canadian Earth System Model version 5 (CanESM5. 0.3), Geosci Model Dev, 12, 4823-4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J Geophys Res-Atmos, 106, 7183-7192, https://doi.org/10.1029/2000JD900719, 2001.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R., and Schmullius, C.: Carbon stock and density of northern boreal and temperate forests, Global Ecol Biogeogr, 23, 297-310, https://doi.org/10.1111/geb.12125, 2014.

1135 Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., Arneth, A., Chang, J., Chen, G., Ciais, P., Gerber, S., Ito, A., Huang, Y., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., Saikawa, E., Thompson, R. L., Vuichard, N., Winiwarter, W., Zaehle, S., Zhang, B., Zhang, K., and Zhu, Q.: The global N_2O model intercomparison project, B Am Meteorol Soc, 99, 1231-1251, https://doi.org/10.1175/BAMS-D-17-0212.1, 2018.

1140 Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717-1736, https://doi.org/10.5194/bg-10-1717-2013, 2013.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, 1145 A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global Fluxnet sites with regression algorithms, Biogeosciences, 13, 4291– 4313, https://doi.org/10.5194/bg-13-4291-2016, 2016.

Tucker, C. J., Fung, I. Y., Keeling, C. D., Gammon, R. H.: Relationship between atmospheric CO_2 variations and a satellite-1150 derived vegetation index, Nature, 319, 195-199, https://doi.org/10.1038/319195a0, 1986.

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P., Meyers, T. P., and Wesely, M. L.: Correcting eddy-covariance flux underestimates over a grassland, Agr Forest Meteorol, 103, 279-300, https://doi.org/10.1016/S0168-1923(00)00123-4, 2000.

1155

Umair, M., Kim, D., Ray, R. L., and Choi, M.: Estimating land surface variables and sensitivity analysis for CLM and VIC simulations using remote sensing products, Sci Total Environ, 633, 470-483, https://doi.org/10.1016/j.scitotenv.2018.03.138, 2018.

1160 Vafaei, S., Soosani, J., Adeli, K., Fadaei, H., Naghavi, H., Pham, T. D., and Tien Bui, D.: Improving accuracy estimation of Forest Aboveground Biomass based on incorporation of ALOS-2 PALSAR-2 and Sentinel-2A imagery and machine learning: A case study of the Hyrcanian forest area (Iran), Remote Sensing, 10, 172, https://doi.org/10.3390/rs10020172, 2018.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., and Sheffield, J.: LS3MIP (v1. 0) contribution 1165 to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project–aims, setup and expected outcome, Geosci Model Dev, 9, 2809-2832, https://doi.org/10.5194/gmd-9-2809-2016, 2016.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth Syst Sci Data, 9, 697– 720, https://doi.org/10.5194/essd-9-697-2017, 2017.

Verger, A., Filella, I., Baret, F., and Peñuelas, J.: Vegetation baseline phenology from kilometric global LAI satellite products, Remote Sens Environ, 178, 1-14, https://doi.org/10.1016/j.rse.2016.02.057, 2016.

Vitousek, P. M., Menge, D. N., Reed, S. C., and Cleveland, C. C.: Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems, Philos T R Soc B, 368, 20130119, https://doi.org/10.1098/rstb.2013.0119, 2013.

Vuichard, N., and Papale, D.: Filling the gaps in meteorological continuous data measured at Fluxnet sites with ERA-Interim reanalysis, Earth Syst Sci Data, 7, 157-171, https://doi.org/10.5194/essd-7-157-2015, 2015.

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Peylin, P.: Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): multi-scale validation of gross primary production, Geosci Model Dev, 12, 4751-4779, https://doi.org/10.5194/gmd-12-4751-2019, 2019.

Waliser, D., Gleckler, P. J., Ferraro, R., Taylor, K. E., Ames, S., Biard, J., Bosilovich, M. G., Brown, O., Chepfer, H., Cinquini, L., Durack, P., Eyring, V., Mathieu, P.-P., Lee, T., Pinnock, S., Potter, G. L., Rixen, M., Saunders, R., Shulz, J., Thepaut, J.-N., and Tuma, M.: Observations for model Intercomparison project (Obs4MIPs): Status for CMIP6, Geosci Model Dev, 13, 2945-2958, https://doi.org/10.5194/gmd-13-2945-2020, 2020.

WCRP.: CMIP Phase 6 (CMIP6), Retrieved January 23, 2021, from https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6, 2020.

Wei, J., Dirmeyer, P. A., Yang, Z. L., and Chen, H.: Effect of land model ensemble versus coupled model ensemble on the simulation of precipitation climatology and variability, Theor Appl Climatol, 134, 793-800, https://doi.org/10.1007/s00704-017-2310-7, 2018.

Wieder, W.: Regridded Harmonized World Soil Database v1.2, ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1247, 2014.

Wieder, W. R., Cleveland, C. C., Smith, W. K., and Todd-Brown, K.: Future productivity and carbon storage limited by terrestrial nutrient availability, Nat Geosci, 8, 441-444, https://doi.org/10.1038/ngeo2413, 2015.

Williams, K. E., Harper, A. B., Huntingford, C., Mercado, L. M., Mathison, C. T., Falloon, P. D., Cox, P. M., and Kim, J.: How can the First ISLSCP Field Experiment contribute to present-day efforts to evaluate water stress in JULESv5.0?, Geosci. Model Dev., 12, 3207–3240, https://doi.org/10.5194/gmd-12-3207-2019, 2019.

1205

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., Zhang, F., Zhang, Y., Wu, F., Li, J., Chu, M., Wang, Z., Shi, X., Liu, X., Wei, M., Huang, A., Zhang, Y., and Liu, X.: The Beijing Climate Center climate system model (BCC-CSM): The main progress from CMIP5 to CMIP6, Geosci Model Dev, 12, 1573-1600, https://doi.org/10.5194/gmd-12-1573-2019, 2019.

1210

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., and Zhang, X.: Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years, Remote Sens Environ, 233, 111383, https://doi.org/10.1016/j.rse.2019.111383, 2019.

1215 Xie, X., Li, A., Tan, J., Lei, G., Jin, H., and Zhang, Z.: Uncertainty analysis of multiple global GPP datasets in characterizing the lagged effect of drought on photosynthesis, Ecological Indicators, 113, 106224, https://doi.org/10.1016/j.ecolind.2020.106224, 2020.

Xu, Z., Jiang, Y., Jia, B., and Zhou, G.: Elevated-CO_2 response of stomata and its dependence on environmental factors, Front
1220 Plant Sci, 7, 657, https://doi.org/10.3389/fpls.2016.00657, 2016.

Yan, Y., Zhou, X., Jiang, L., and Luo, Y.: Effects of carbon turnover time on terrestrial ecosystem carbon storage, Biogeosciences, 14, 5441-5454, https://doi.org/10.5194/bg-14-5441-2017, 2017.

1225 Yoshikawa, C., Kawamiya, M., Kato, T., Yamanaka, Y., and Matsuno, T.: Geographical distribution of the feedback between future climate change and the carbon cycle, J Geophys Res-Biogeo, 113, https://doi.org/10.1029/2007JG000570, 2008.

Zaehle, S., and Dalmonech, D.: Carbon–nitrogen interactions on land at global scales: current understanding in modelling climate biosphere feedbacks, Curr Opin Env Sust, 3, 311-320, https://doi.org/10.1016/j.cosust.2011.08.008, 2011.

1230

Zhang, Y. J., Yu, G. R., Yang, J., Wimberly, M. C., Zhang, X. Z., Tao, J., Jiang, Y. B., and Zhu, J. T.: Climate-driven global changes in carbon use efficiency, Global Ecol Biogeogr, 23, 144– 155, https://doi.org/10.1111/geb.12086, 2014.

Zhang, Z., Zhang, Y., Zhang, Y., Gobron, N., Frankenberg, C., Wang, S., and Li, Z.: The potential of satellite FPAR product for GPP estimation: An indirect evaluation using solar-induced chlorophyll fluorescence, Remote Sens Environ, 240, 111686, https://doi.org/10.1016/j.rse.2020.111686, 2020.

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

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R.R., and Myneni, R. B.: Global data sets of vegetation leaf area index (LAI) 3g and fraction of photosynthetically active radiation (FPAR) 3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3g) for the period 1981 to 2011, Remote sens, 5, 927-948, https://doi.org/10.3390/rs5020927, 2013.

Zhu, Q., Castellano, M. J., and Yang, G.: Coupling soil water processes and the nitrogen cycle across spatial scales: Potentials, bottlenecks and solutions, Earth-Sci Rev, 187, 248-258, https://doi.org/10.1016/j.earscirev.2018.10.005, 2018.

Ziehn, T., Kattge, J., Knorr, W., and Scholze, M.: Improving the predictability of global CO2 assimilation rates under climate change, Geophys Res Lett, 38, https://doi.org/10.1029/2011GL047182, 2011.

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.P., Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, Journal of Southern Hemisphere Earth Systems Science, 70, 193-214, https://doi.org/10.1071/ES19035, 2020.

Appendices

A. Technical Summary of Validation Activities by Participants

A.1 CSIRO

1260   The Australian Community Climate and Earth System Simulator (ACCESS-ESM1.5) was developed by the Australian modelling group Commonwealth Scientific and Industrial Research Organization (CSIRO) for participation in CMIP6 (Ziehn et al., 2020). The land surface model used in ACCESS-ESM1.5 is the Community Atmosphere Biosphere Land Exchange (CABLE) model (Kowalczyk et al., 2013; 2006) version 2.4. Ziehn et al. (2020) compared ACCESS-ESM1.5 simulated land carbon cycle variables against observation-based estimates for the 1986-2005 period. The spatial distribution

1265 of simulated average annual GPP was compared to upscaled Fluxnet observations from Jung et al. (2011), while average annual global GPP was compared to observation-based estimates from Beer et al. (2010) and Ziehn et al. (2011). Simulated LAI magnitude and seasonality was compared to global and regional estimates based on Moderate Resolution Imaging Spectroradiometer (MODIS) and Advanced Very High-Resolution Radiometer (AVHRR) data from Zhu et al. (2013). Simulated surface CO_2 concentrations in terms of mean seasonal cycle amplitude and timing were compared to four

1270 NOAA/Earth System Research Laboratory station flask samples provided in the GLOBAL VIEW data product (GLOBAL VIEW-CO_2 2013).

A.2 BCC

        The Beijing Climate Centre (BCC) participated in CMIP6 with the BCC Climate System Model version 2 with medium resolution (BCC-CSM2-MR; Wu et al., 2019). Land biogeochemistry in BCC-CSM2-MR was simulated through the

1275 BCC Atmosphere and Vegetation Interactive Model version 2.0 (BCC-AVIM2; Li et al., 2019). While Wu et al. (2019) did not provide validation results for terrestrial biogeochemistry from BCC-CSM2-MR, a detailed validation with offline simulations of BCC-AVIM2 was provided by Li et al. (2019) using the Princeton global forcing dataset (Sheffield et al., 2006). Li et al. (2019) compared the annual peak month, seasonal average, and global average of LAI to satellite observations from 1982-2010 by the AVHRR (Myneni et al., 1997). Surface carbon fluxes including GPP and ER were compared to upscaled

1280 Fluxnet observations from Jung et al. (2011). Above ground biomass was compared to Avitabile et al. (2016), while global total biomass carbon from 1990-2010 was compared to Saatchi et al. (2011). The performance of BCC-AVIM2 in estimating each of these variables was assessed through bias, RMSE, and Taylor diagram metrics (Taylor, 2001).

A.3 CCCma

        The Canadian Centre for Climate Modelling and Analysis (CCCma) participated in CMIP6 with the CCCma fifth

1285 generation Earth System model (CanESM5; Swart et al., 2019). The land biogeochemistry component of CanESM5 is the

Canadian Terrestrial Ecosystem Model (CTEM; Arora and Boer, 2010; 2005). Swart et al. (2019) compared CanESM5 simulated GPP from 1982-2009 with observation-based estimates from Jung et al. (2009) in terms of geographical distribution, zonal averages, as well as functional relationships with air temperature and precipitation. Several metrics were used to illustrate CanESM5's performance in simulating GPP, including the correlation coefficient (r) between simulated and observed spatial patterns in GPP, bias (simulated – observed), as well as root mean squared error (RMSE) normalized (NRMSE) by observed spatial standard deviation. Global average decadal land-atmosphere CO_2 flux as well as net cumulative atmosphere-land CO_2 flux from 1850-2014 were compared to observation-based estimates from the Global Carbon Project (GCP; Le Quéré et al., 2018), the latter by subtracting cumulative land use emissions from cumulative land carbon uptake.

A.4 Climate and Global Dynamics Laboratory NCAR

The Community Earth System Model version 2 (CESM2) was developed by the Climate and Global Dynamics Laboratory at the American National Centre for Atmospheric Research (NCAR) for participation in CMIP6 (Danabasoglu et al., 2020). The land component of CESM2 is the Community Land Model Version 5 (CLM5; Lawrence et al., 2019). Danabasoglu et al. (2019) and Lawrence et al. (2019) comprehensively assessed terrestrial biogeochemical cycle variable outputs from simulations of CESM2 and CLM5, respectively, with the International Land Model Benchmarking package (ILAMBv2.1; Collier et al., 2018), including an explicit analysis of interannual variability with a three member ensemble from different pre-industrial control initialization years (CESM2), the influence of forcing through the use of three forcing datasets (CLM5), and the influence of prescribed versus prognostic vegetation phenology (CLM5). ILAMBv2.1 utilizes a suite of data products weighted by certainty. These included vegetation biomass (tropical: Saatchi et al., 2011; global: Kellndorfer et al., 2013; Blackard et al., 2008), burned area (Giglio et al., 2010), CO_2 concentrations, GPP (Fluxnet: Lasslop et al., 2010; Global biosphere-atmosphere flux: Jung et al., 2010), LAI (AVHRR: Myneni et al., 1997; MODIS: de Kauwe et al., 2011), global net ecosystem carbon balance (GCP: Le Quéré et al., 2014; Hoffman et al., 2014), net ecosystem exchange (Fluxnet: Lasslop et al., 2010; GBAF: Jung et al., 2010), NBP, ER, NEP (equivalent to GPP-ER), soil carbon (Harmonized World Soil Database (HWSD): Todd-Brown et al., 2013; Northern Circumpolar Soil Carbon Database (NCSCDV22): Hugelius et al., 2013), as well as 10 functional relationships. Lawrence et al. (2019) also compared the relationship between apparent soil carbon turnover times versus air temperature to observation-based estimates developed from HWSD, NCSDV22, and MODIS. Lawrence et al. (2019) additionally compared maximum monthly LAI and average V_{cmax25} (maximum rubisco carboxylation rate at 25°C and high irradiance per unit leaf area in $\mu mol \cdot m^{-2} \cdot s^{-1}$) at the PFT-level for the year 2010 to Zhao et al. (2005) and Kattge et al. (2009), respectively, as well as canopy height for the year 2005 for tree PFTs to Simard et al. (2011). Nitrogen cycle variables evaluated by Lawrence et al. (2019) with observational references included nitrogen deposition (Fowler et al., 2013), symbiotic fixed nitrogen (Vitousek et al., 2013), soy fixed nitrogen (Herridge et al., 2008), crop nitrogen fertilization (Fowler et al., 2013), denitrification (Fowler et al., 2013), hydrologic nitrogen losses (Fowler et al., 2013), fire losses (Lamarque et al., 2010), and N_2O flux (Fowler et al., 2013). Different climate forcing datasets and anthropogenic forcings were utilized to examine the effect of climate, CO_2 emissions, land use change, and nitrogen additions on carbon cycle variables as well as three CLM

model versions to partition total uncertainty into forcing and model contributions using fixed-effect analysis of variance, with additional PFT-level analysis and prognostic versus prescribed vegetation and carbon cycling for CLM5. In addition to the ILAMB validation, Danabasoglu et al. (2019) and Lawrence et al. (2019) compared simulated global net biome production (NBP) and cumulative land carbon sink to observation-based estimates from 1850-2014 from the GCP for 1959-2014 (Le Quéré et al., 2016), and from Hoffman et al. (2014) for 1850-2010. Observation-based GPP, ER, and NEP (equivalent to GPP-ER), comparison data were obtained from Jung et al. (2011; 2010). Vegetation carbon was evaluated relative to observations for the tropics from Saatchi et al. (2011), as well as GEOCARBON and GlobalCarbon datasets (Collier et al., 2018; Avitabile et al., 2016; Santoro et al., 2015). ILAMBv2.1 results from these investigations comprised a collection of statistical metrics for annual mean, bias, relative bias, RMSE, seasonal cycle phase, spatial distribution, and interannual variability, in addition to functional relationships. Bonan et al. (2019) provides a detailed analysis on the role of climate forcing uncertainty in influencing CLM5 output.

A.5 CNRM and CERFACS

The Centre National de Recherches Météorologiques (CNRM) and Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) contributed the CNRM-ESM2-1 to CMIP6 (Séférian et al., 2019). The land component in CNRM-ESM2-1 is the Ineraction Soil-Biosphere-Atmosphere with Total Runoff Integrating Pathways with carbon cycling (ISBA-CTRIP; Delire et al., 2020). Séférian et al. (2019) compared CNRM-ESM2-1 simulated annual minimum and maximum LAI to AVHRR observations from 1998-2011 (Zhu et al., 2013). The simulated land carbon sink from 1982-2010 was compared to a multi-model estimate by Huntzinger et al. (2013). These validations included spatial bias, global mean bias, RMSE, as well as spatial error correlation between CNRM ESM versions to distinguish model sources of error. Delire et al. (2020) validated offline ISBA-CTRIP simulated GPP, NPP, autotrophic respiration, and ER from 1980-2010 with estimates with the mean of 12 products from the FluxComv1 dataset (Jung et al., 2017; 2016; Tramontana et al., 2016), and a satellite product from the Numerical Terradynamic Simulation Group: MODIS17A3 (NASA LP DAAC, 2017; Zhao et al., 2005), with reference autotrophic respiration calculated as the mean of FLUXCOM GPP products minus MODIS17A3 NPP. Simulated crop NPP for the 2000s was compared to the Harvested Area and Yield dataset (Monfreda et al., 2008). Carbon use efficiency (CUE), calculated as the ratio of NPP to GPP, was evaluated with observation and model-based estimates for tropical evergreen forest from Malhi et al. (2009), and tropical deciduous, temperate, and boreal forests from He et al. (2018), Zhang et al. (2014), and theoretical derivations by Amthor, (2000). Simulated heterotrophic respiration was evaluated with a data product from Hashimoto et al. (2015) which combines global and Amazonian in situ observations from the Soil Respiration database (Bond-Lamberty et al., 2018) and Malhi et al. (2009), respectively, and global gridded climate data. The simulated burned area and fire CO_2 emissions were compared to Mouillot and Field (2005) and the Global Fire Emissions Database version 4.1 (Randerson et al., 2017; van der Werf et al., 2017). Simulated dissolved organic carbon yield leached from soil was compared to model results of Mayorga et al. (2010), and observations by Dai et al. (2012). Simulated global aboveground biomass carbon was validated with observation-based estimates from 1993-2012 from Liu et

al. (2015), regional datasets for mid-high northern latitudes from Thurner et al. (2014), and tropical datasets from Saatchi et al. (2011) and Baccini et al. (2012). Simulated above ground litter carbon was compared to site measurements from 1827-1997 from the Global Database of Litterfall Mass and Litter Pool Carbon and Nutrients (Holland et al., 2015). Simulated belowground organic carbon was validated with the HWSDv1.2 (FAO, 2012). Vegetation turnover time calculated as biomass divided by NPP and soil turnover time calculated as the combination of litter and soil carbon divided by NPP for 1984-2014 were also computed for validation. Delire et al. (2020) also used local scale Fluxnet data from Joetzjer et al. (2015) to assess ISBA-CTRIP performance. Each variable was validated through comparison of the distribution of simulated and observation-based estimates of annual averages, as well as zonal averages, and the spatial distribution of the bias (simulated minus observed). Average simulated carbon fluxes from 2006-2015 and the trend from 1960-2015 were also compared to observation-based estimates from the GCP (Le Quéré et al., 2018) and Ciais et al. (2019).

A.6 IPSL

The Institut Pierre Simon Laplace (IPSL) participated in CMIP6 with IPSL-CM6A-LR, the land component of which was the ORCHIDEE land surface model version 2.0 (Boucher et al., 2020; Hourdin et al., 2020). Boucher et al. (2020) evaluated IPSL-CM6A-LR simulated average annual carbon fluxes from 1990-1999 and 2009-2018 resulting from land cover change, fossil fuel emissions, the terrestrial sink, and total net land fluxes (the terrestrial sink minus land cover change) with observation-based estimates from the 2019 GCP (Friedlingstein et al., 2019). Vuichard et al. (2019) validated ORCHIDEE simulated GPP in terms of the mean annual, seasonal, and daily simulated GPP on a PFT, spatial, and global basis against observations from 78 Fluxnet sites (Vuichard and Paple, 2015) and the global-scale MTE-GPP product based upon upscaled Fluxnet observations for 1982-2008 (Jung et al., 2011). RMSE as well as dissected mean squared deviation (MSE; which is the sum of squared bias, squared difference between standard deviations, and lack of correlation weighted by standard deviations; based on Kobayashi and Salam, (2000)), metrics were used to attribute different sources of uncertainty. The relative contribution of drivers of variation in present-day GPP were also assessed, including seasonal variability in NO_x and NH_x deposition as well as leaf carbon: nitrogen ratio. The sensitivity of ORCHIDEE output to model structure in terms of MSE was also analyzed on a global and PFT-level basis, including fixed and dynamic fully coupled carbon-nitrogen cycles.

A.7 GFDL

The American National Oceanic and Atmospheric Administration Geophysical Fluid Dynamics Laboratory (GFDL) participated in CMIP6 with GFDL-ESM4.1 (Dunne et al., 2020), in which land biogeochemistry is simulated with the GFDL Land Model version 4.1 (LM4.1; Shevliakova et al., 2020). Dunne et al. (2020) validated GFDL-ESM4.1's simulated spatial distribution of seasonal amplitude in CO_2 concentrations and interannual variability of CO_2 concentrations compared to NOAA Global Monitoring Division sites with at least 15-year long records (Global Monitoring Laboratory, 2005) using RMSE and the coefficient of determination (r^2), as well as the correlation coefficient (r) for individual sites.

A.8 JAMSTEC, University of Tokyo, and National Institute for Environmental Studies

The Japanese Agency for Marine-Earth Science and Technology (JAMSTEC), University of Tokyo, and National Institute for Environmental Studies participated in CMIP6 with the Model for Interdisciplinary Research on Climate Earth System version 2 for Long-term simulations (MIROC-ES2L; Hajima et al., 2020). The land biogeochemical component in MIROC-ES2L is Vegetation Integrative Simulator for Trace gases model (VISIT-e; Ito and Inatomi, 2012). Hajima et al. (2020) evaluated MIROC-ES2L simulated terrestrial carbon gain with and without land use, as well as land use emissions from 1850-2014 in comparison to multi-model estimates from the GCP (Le Quéré et al., 2018). Observational-based data products used for other comparisons included 1) the spatial pattern, gradient across biomes, magnitude, seasonality, and length of growing season of global gridded GPP from 1986-2005 from Fluxnet (Jung et al., 2011), 2) the magnitude and density of forest carbon stock (Kindermann et al., 2008) and 3) global and regional soil organic carbon from the harmonized soil property values for broad-scale modelling (WISE30Sec; Batjes, 2016), the northern high latitudes from the Northern Circumpolar Soil Carbon Database version 2 (NCSCDv2; Hugelius et al., 2013), and an estimate from Todd-Brown et al. (2013) developed from the HWSD version 1.3 (FAO, 2012). Hajima et al. (2020) also compared simulated and observation-based estimates of annual biological nitrogen fixation (BNF) from 1850-2014 (Gruber and Galloway, 2008), present-day BNF (Galloway et al., 2008; Herridge et al., 2008), annual unperturbed state terrestrial N_2 flux (Gruber and Galloway, 2008), and change in annual soil nitrous oxide emissions from 1850-2014 relative to a model comparison study by Tian et al. (2018).

A.9 MPI

The Max Planck Institute for Meteorology (MPI) Earth System Model version 1.2 Low Resolution (MPI-ESM1.2-LR) was developed for participation in CMIP6 (Mauritsen et al., 2019) by the MPI, the land component of which is JSBACH3.2 (Goll et al., 2017). Mauritsen et al. (2019) compared the spatial variability and zonally averaged density of MPI-ESM1.2-LR simulated soil and litter carbon stocks to estimates by Goll et al. (2015) developed from the Harmonized World Soil Database. The simulated evolution in global total land carbon from 1850-2013 was compared to estimates provided by Ciais et al. (2013). Additionally, simulated land use change carbon emissions from 1860-2013 were compared to estimates provided by Ciais et al. (2013). In a model description paper of JSBACH version 3.10, which was set to be used in CMIP6, Goll et al. (2017) compare JSBACH3.1 simulated present-day NPP to Ito (2011), while simulated present-day biomass carbon was compared to Saugier and Roy, (2001) and Ciais et al. (2013). The simulated response of NPP and GPP to increases in atmospheric CO_2 were compared to experimentally observed estimates from four free-air CO_2 enrichment (FACE) experiments (Norby et al., 2005) and an intramolecular isotope distribution examination of plant metabolic shifts (Ehlers et al., 2015). Simulated present-day biomass nitrogen was compared to Schlesinger (1997) while simulated present-day total nitrogen was compared to Galloway et al. (2013). Simulated values of pre-industrial (1850) and present-day leaching and BNF were compared to Galloway et al. (2013; 2004), Vitousek et al. (2013), and short-term experimental results from a meta-analysis by Liang et al. (2016), while simulated present-day denitrification was compared to Galloway et al. (2013). Goll et al. (2017) also

1415 verified the simulated spatial variability in reactive nitrogen-loss pathways using a compilation of nitrogen-15 isotopic data (Houlton et al., 2015) with the statistical metrics r, RMSE, and Taylor score (Taylor 2001).

A.10 NCC

The Norwegian Earth System Model (NORESM2) was developed for participation in CMIP6 (Seland et al., 2020) by the Norwegian Climate Consortium (NCC), and is based on CESM2. As in CESM2, the land model in NORESM2 is CLM5
1420 (Lawrence et al., 2019). The performance of NORESM2 was validated through a three-member ensemble of historical simulations from 1850-2014 with slightly varying initial conditions. Simulated carbon cycle variables which were compared to observation variables included GPP, soil carbon, and vegetation carbon, from Jung et al. (2011), FAO, (2012), and Avitabile et al. (2016) and Santoro et al. (2015), respectively. Seland et al. (2020) NORESM2 results in terms of carbon stocks and fluxes broadly agree with those of Lawrence et al. (2019) while conducting land-only simulations of CLM5.

1425 ### A.11 NERC and Met Office

The United Kingdom Community Earth System Model (UKESM1-0-LL) was developed for participation in CMIP6 by the United Kingdom Natural Environmental Research Council (NERC) and National Meteorological Service (Met Office; Sellar et al., 2019). The land component in UKESM1-0-LL is an updated version of the Joint UK Land Environment Simulator (JULES; Clark et al., 2011) with an additional PFT updated competition scheme (Harper et al., 2018). Sellar et al. (2019)
1430 evaluated UKESM1-0-LL simulated global GPP magnitude and evolution in time through comparisons to recent decadal GPP from the Fluxnet model tree ensemble data product (Jung et al., 2011). The areal land cover of aggregated plant functional types (PFTs) was validated with satellite observation-based datasets from the European Space Agency Climate Change Initiative Land Cover data (Poulter et al., 2015) as well as the International Geosphere-Biosphere Programme (IGBP) Land Use and Cover Change project (Loveland et al., 2000) using the model year 2005. The coverage of PFTs were validated using
1435 these observation-based datasets as references both spatially and as a fraction of biomes based upon regions defined by Olson et al. (2006). The simulated vegetation carbon distribution was validated on a latitudinal basis with observation-based estimates from GEOCAROBON (Avitabile et al., 2016) and Saatchi et al. (2011), while the spatial distribution of soil carbon was validated with observation-based estimates WISE30sec (Batjes, 2016), IGBP-DIS (Global Soil Data Task Group, 2002), and Carvalhais et al. (2014). The magnitude of simulated global total soil carbon was compared to whole soil profile observation-
1440 based estimates from Carvalhais et al. (2014) and upper 2 m observation-based estimates from Batjes, (2016). Cumulative carbon uptake and land use emissions from 1850-2014 was compared to observation-based estimates from the GCP (Le Quéré et al., 2018).