Nitrogen Cycling in CMIP6 Land Surface Models: Progress and Limitations

Taraka Davies-Barnard[1,2], Johannes Meyerholt[2], Sönke Zaehle[2], Pierre Friedlingstein[1,3], Victor Brovkin[4], Yuanchao Fan[5,6], Rosie A. Fisher[7,8], Chris D. Jones[9], Hanna Lee[5], Daniele Peano[10], Benjamin Smith[11,12], David Wårlind[11,12], and Andy Wiltshire[9]

[1]University of Exeter, Exeter, UK
[2]Max Planck Institute for Biogeochemistry, Jena, Germany
[3]Laboratoire de Meteorologie Dynamique, Institut Pierre-Simon Laplace, CNRS-ENS-UPMC-X, Departement de Geosciences, Ecole Normale Superieure, 24 rue Lhomond, 75005 Paris, France
[4]Max Planck Institute for Meteorology, Hamburg, Germany
[5]NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen, Norway
[6]Harvard University, Cambridge, USA
[7]National Center for Atmospheric Research, Boulder, Colorado, USA
[8]Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France
[9]Met Office Hadley Centre, Exeter, UK
[10]Fondazione Centro euro-Mediterraneo sui Cambiamenti Climatici, Bologna, Italy
[11]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
[12]Hawkesbury Institute for the Environment, Western Sydney University, Richmond, Australia

Correspondence to: T. Davies-Barnard (t.davies-barnard@exeter.ac.uk)

Abstract. The nitrogen cycle and its effect on carbon uptake in the terrestrial biosphere is a recent progression in earth system models. As with any new component of a model, it is important to understand the behaviour, strengths, and limitations of the various process representations. Here we assess and compare five models with nitrogen cycles that ~~will be~~are used as the terrestrial components of some of the earth system models in CMIP6. We use a historical control simulation and two perturbations to assess the models' nitrogen-related performance: a simulation with atmospheric carbon dioxide 200 ppm higher, and one with nitrogen deposition increased by 50 kg N ha^{-1} yr^{-1}. ~~We find that, despite differing nitrogen cycle representations, all models simulate recent global trends~~There is generally greater variability in ~~terrestrial productivity and net carbon uptake commensurate with observations. The between-model variation is likely more influenced by other, non-nitrogen parts of the models. Globally, the~~ productivity response across models to increased nitrogen than to carbon dioxide ~~is commensurate with~~. Compared to observations ~~for four~~, two models of the ~~five models, but highly spatially variable within and between~~ models. ~~The~~ considered here have low productivity response to nitrogen, and another one a low response to elevated atmospheric carbon dioxide. In all five models individual grid cells tend toward bimodality, with either a strong response to increased nitrogen ~~is significantly lower than observed in two of the five models.~~or atmospheric carbon dioxide, but rarely to both to an equal extent. However, this local effect does not scale to either the regional or global level. The global and tropical ~~values~~responses are generally better represented than boreal, tundra, or other high latitude areas. These results are due to divergent though valid choices in the representation of key nitrogen cycle

Style Definition: Bibliography: Space After: 12 pt, Line spacing: single

Formatted: Superscript

Formatted: Superscript

processes. They show the need for better understanding and more provision of observational constraints of nitrogen processes, especially nitrogen-use efficiency and biological nitrogen fixation.

1 Introduction

The terrestrial carbon (C) cycle currently removes around a third of anthropogenic carbon emissions from the atmosphere (Friedlingstein et al., 2019; Le Quéré et al., 2018). Changes in this uptake will affect the allowable emissions for targets such as limiting warming to 1.5°C (Millar et al., 2017; Müller et al., 2016). Nitrogen (N) is required to synthesise new plant tissue (biomass) out of plant-assimilated ~~carbon~~C, in differing ratios across biomes and tissue types (McGroddy et al., 2004)~~.~~ ~~Therefore, future projections of terrestrial carbon uptake and allowable emissions are dependent on N availability, particularly under high atmospheric carbon dioxide (CO_2) conditions (Arora et al., 2019; Wieder et al., 2015b; Zaehle et al., 2014b). Earth System Models (ESMs) are tools to project the responses of the coupled earth system and its components such as the atmosphere, biosphere, and oceans to anthropogenic and natural forcings~~ . Therefore, future projections of terrestrial C uptake and allowable emissions are dependent on N availability, particularly under high atmospheric carbon dioxide (CO_2) conditions (Arora et al., 2019; Meyerholt et al., 2020; Wieder et al., 2015b; Zaehle et al., 2014b). A key tool for projections of allowable emissions are Earth System Models (ESMs), which project the responses of the coupled earth system (Anav et al., 2013; Arora et al., 2013; Friedlingstein et al., 2006; Jones et al., 2013). The Fifth Phase of the Coupled Model Intercomparison Project (CMIP5, Taylor et al., 2012) had numerous ~~models~~ESMs with a global C cycle but only two, based on the same land component, with terrestrial N cycling (Thornton et al., 2009)~~. However, predictions of terrestrial C storage would decrease by 37 – 58% if ESMs accounted for the constraints on terrestrial C sequestration imposed by the tight coupling of the C cycle with the N cycle (Wieder et al., 2015b; Zaehle et al., 2014b)~~. A number of studies with stand-alone terrestrial biosphere models (Sokolov et al., 2008; Wårlind et al., 2014; Zaehle et al., 2010; Zhang et al., 2013) as well as post-hoc assessments of CMIP5 projections suggest that predictions of terrestrial C storage would decrease by 37 – 58% if ESMs accounted for N constraints (Wieder et al., 2015b; Zaehle et al., 2014b).

The latest generation of models in CMIP6 (Eyring et al., 2016) have at least six ~~models~~ESMs that incorporate the N cycle (Arora et al., 2019). These models employ a range of assumptions and process formulations, reflecting divergent theory and significant knowledge gaps (Zaehle and Dalmonech, 2011). Since N availability is an important source of uncertainty for the C cycle (Meyerholt et al., 2020) assessment of the sensitivity of the N cycle to changes in atmospheric CO_2 and N inputs is required. Because of the tight coupling of C and N dynamics, a direct evaluation of the N effects on simulated C cycle dynamics using conventional model benchmarking approaches (Collier et al., 2018; Luo et al., 2012) is challenging. More insights into the magnitude of a N effect can be gained by comparing model simulations against perturbation experiments

that provide evidence for the responses of terrestrial ecosystems to changes in the C and N availability (Thomas et al., 2013; Wieder et al., 2019; Zaehle et al., 2010).

In this study, we test five land surface models (LSMs) with N cycles employed in the latest generation of ESMs used in European Earth System modelling centres that contribute to CMIP6, we use a set of standardised model forcing and protocol to simulate historical changes in the C and N balance, as well as the response to N and C perturbations. The perturbation experiments (described in the methods) are designed to approximate field experiments undertaken to understand the effects of elevated CO_2 or N. These simulations reveal the overall pattern of response of the model to these forcings. ~~Since variation in N is a main source of uncertainty for the carbon cycle (Song et al., 2019), assessment of the models' N cycle processes is required to understand the headline performance and sensitivity of the N cycle to changes in atmospheric CO_2 and N inputs. In this study, we evaluate land surface models (LSMs) with N cycles employed in the latest generation of European ESMs that contribute to CMIP6. We use a range of upscaled field-based~~We use a range of upscaled meta-analyses of observations, satellite observations, and model-to-model comparisons to assess the behaviour and performance of the models. Comparisons between models alone can also provide useful insight into the models' behaviour. The approach of assessing ESM N cycles via their corresponding offline LSMs, driven by a standardised set of model forcing, has the advantage of making model projections directly comparable while giving a representative view of the latest N cycle developments.

2 Methods

2.1 Models

We ~~user~~an simulations with five LSMs that are the land components of five different European ESMs taking part in CMIP6. ~~The models employ a diverse set of process representations of the N cycle and its coupling to the C cycle.~~ The key N process formulations are summarized in Table 1. A brief description of each model follows.

The Community Land Model version 4.5 (CLM4.5; Koven et al., 2013; Oleson et al., 2010) is used in the Euro-Mediterranean Centre on Climate Change coupled climate model (CMCC-CM2; Cherchi et al., 2019). The N component is described in Koven et al., (2013)~~. CLM4 was the first N model for ESMs, used in CMIP5. The N cycling component of CLM4.5 is similar to CLM4, but other features of the model~~ and was the first N model for ESMs, used in CMIP5 (Thornton et al., 2007, 2009). While the N cycling component of CLM4.5 is similar to CLM4 features of CLM4.5, such as leaf physiological traits (Bonan et al., 2012), were modified~~.~~ and there is a vertically resolved soil biogeochemistry scheme (Koven et al., 2013) as opposed to the single-layer box modelling scheme for CN in CLM4.

~~The Community Land Model version 5 (CLM5; Lawrence et al., 2020) is used in the Norwegian Earth System Model version 2 (NorESM2). CLM5 is the latest version of CLM and represents a suite of developments on top of CLM4.5. CLM5 prescribes crop management including nitrogen fertilizer use and irrigation fraction at annual time step.~~The Community Land Model version 5 (CLM5; Lawrence et al., 2019) is used in the Norwegian Earth System Model version 2 (NorESM2; Seland et al., 2020). CLM5 is the latest version of CLM and represents a suite of developments on top of CLM4.5. The N

component is described in Fisher et al., (2010); and Shi et al., (2016). The key difference for the N cycle compared to CLM4 is the implementation of a C cost basis for acquiring N, derived from the Fixation and Uptake of Nitrogen (FUN) approach (Fisher et al., 2010).

~~The~~ JSBACH version 3.20 model (Goll et al., 2017) is used in the Max Planck Earth System Model version 1.2 (MPI-ESM; ~~Mauritsen et al., 2019)~~Mauritsen et al., 2019). The N component is described in Goll et al., (2017).

The Joint UK Land Environment Simulator version 5.4 (JULES-ES; Best et al., 2011; Clark et al., 2011) is used in the UK Earth System Model (UKESM1; ~~Sellar et al., 2020.)~~Sellar et al., 2020.). The N component is described in Wiltshire et al. (forthcoming) and ~~Sellar et al., (2020)~~Sellar et al., (2020).

The Lund-Potsdam-Jena General Ecosystem Simulator version 4.0 (LPJ-GUESS; Olin et al., 2015; Smith et al., 2014) is used in the European community Earth-System Model (EC-Earth; Hazeleger et al., 2012). The N component is described in Smith et al., (2014).

2.2 Forcing Data and Model Initialisation

All ~~models ran a global spin-up for all ecosystem~~models' pools were spun-up to ~~the year 1860,~~equilibrium forced by pre-industrial conditions. This comprised of a constant atmospheric CO_2 concentration of 287.14 ppm, cycling global climate data at 0.5° x 0.5° resolution for the years 1901-1930 from the CRU-NCEP dataset version 7.0 (New et al., 2000), ~~and assuming constant 1860 land-use from the Hurtt et al., (2011) database. Next, transient~~assuming constant 1860 land cover from the Hurtt et al., (2020) database, and 1860s nitrogen deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (Lamarque et al., 2013). Next, transient historical runs were performed for the 1861-1900 period with the same climate forcing as the spin-up, but now including varying atmospheric CO_2 concentrations from synthesized ice core and National Oceanic and Atmospheric Administration (NOAA) measurements, as well as annually varying land-use from ~~Hurtt et al., (2011). The simulations were then continued for 1901-2015 under fully dynamic forcing including climate.~~Hurtt et al., (2020). The N deposition is taken from the Atmospheric Chemistry and Climate Model Intercomparison Project (Lamarque et al., 2013). The simulations were then continued for 1901-2015 under fully dynamic forcing including climate.

The models applied their individual soil and vegetation spin-ups according to their respective conventions. The goal of the spin-up procedure is to obtain quasi-steady states of the ecosystem pools in relation to climate, avoiding drifting pool sizes due to lack of equilibrium, especially ~~by~~for slow-turnover soil organic matter pools. Because of differences among the models, pool sizes after spin-up are not identical. ~~However, we presume that model responses to perturbations do not depend strongly on the initial model states if those remain within a reasonable range.~~

Formatted: Font: Bold

2.3 Model Experiments

In addition to the ~~Control~~historical run described above~~,~~ (referred to hereafter as the Control), two experiments were run for the period 1996-2015: increased CO_2 (+CO2) and increased N (+N). These two experimental runs are compared to the corresponding 1996-2015 simulations from the unperturbed Control runs. SI Table 1 provides a summary of the experiments.

For the increased CO_2 experiment (+CO2) the atmospheric CO_2 concentration was abruptly increased to constant 550 ppm. This is almost twice the pre-industrial atmospheric CO_2 of 280ppm or a 200ppm increase compared to the 1996 atmospheric CO_2 of ~350 ppm, similar to free-air CO_2 enrichment experiments performed in the 1990s (Norby et al., 2005).

For the increased N experiment (+N) N deposition was abruptly increased by 50 kg N ha^{-1} yr^{-1}, which is roughly equivalent to a number of forest N fertilisation trials (Thomas et al., 2013) and around 5 – 10 times higher than typical background N deposition (Zak et al., 2017).

2.4 Analytical Framework

The response of the terrestrial productivity (and with it terrestrial C storage) to changes in the N cycle is in principle controlled by two components: (i) the net ecosystem balance of N, which determines the change in the N capital available for plant growth and soil organic matter decay, and (ii) the ratio of carbon production per unit N availability, which can be most effectively be described as the N-use efficiency of growth.

Because the individual processes and pools considered varies between the five models (Table 1), we use a simplified N budget to assess the annual change in the terrestrial N store (ΔN, including soil and plants):

$$\Delta N = N_{dep} + BNF - N_{loss} \quad (1)$$

where N_{dep} is the N deposition, BNF is the biological N fixation, and N_{loss} is the N lost from gaseous, leaching, and other pathways, as declared by the models. This paradigm assumes that increased ecosystem N input from deposition or fixation enters the soil and then becomes available for plant uptake. In a similar way, plant N uptake (N_{up}) could led to reduced N losses, which would (assuming constant N inputs) result in an apparent increase in the ecosystem N capital. Note that crop fertilisation is not included here, as it is assumed to remain constant between the 3 simulations.

Whether and how this change in N capital affects plant growth is dependent on the magnitude of the change in plant N uptake, as well as relationship between N_{up} and NPP (whole-plant nitrogen-use efficiency; NUE; (Zaehle et al., 2014a))

~~For the increased CO_2 experiment (+CO2) the atmospheric CO_2 concentration was abruptly increased to constant 550 ppm. This is almost twice the pre-industrial atmospheric CO_2 of 280 ppm or a 200 ppm increase compared to the 1996 atmospheric CO_2 of ~350 ppm.~~

For the increased N experiment (+N) N deposition was abruptly increased by 50 kg N ha^{-1} yr^{-1}., which is roughly equivalent to typical N fertilisation of crops and around 5 – 10 times higher than typical background N deposition (Zak et al., 2017).

2.4 Analytical Framework

To help generalize the key C-N processes for each model under the +CO2 and +N scenarios, we use a simple framework that relates N-limited plant growth (NPP) to whole-plant N use efficiency (NUE; newly constructed biomass C per unit N taken up) and plant N uptake (Zaehle et al., 2014a):

$$NPP = NUE * N_{up} \quad , \qquad (1= \frac{NPP}{N_{up}} \quad . \qquad (2)$$

where N_{up} includes plant uptake of soil inorganic N of any origin, i.e. atmospheric deposition, fertilization, decomposition of plant litter, or biological nitrogen fixation (BNF). NUE is the outcome of the product of tissue stoichiometry and fractional allocation of NPP to different tissue types, and therefore varies with changes in the allocation fractions and tissue C:N.

This paradigm assumes that ecosystem N input from deposition or fixation enters the soil and then becomes available for plant uptake.

As a consequence of Eq. 1, we consider changes in plant growth induced from perturbations as resulting from changes in NUE or N_{up}:

$$\Delta NPP = \Delta NUE * \Delta N_{up} \quad . \qquad (2)$$

Our focus is on how models differ in inorganic N availability to plants and plant NUE. If these are constant, changes would be limited to shifts in C allocation between plant organs with different C:N. Flexible C:N within plant organs allows for more complex interactions through changes in internal stoichiometry under perturbation (Meyerholt and Zaehle, 2015).

We also use a simplified N budget to assess the change in N:

2.5 Observations for Comparison

We utilise a range of observation-based metrics for comparison to the models at global and regional scales, detailed in Table 2. Most of these are based on small numbers of field studies upscaled or averaged to give an approximate global value with confidence intervals. While these upscaled values need to be interpreted with proportional caution, in the absence of more robust comparators they are useful benchmarks that can provide real-world context in addition to field scale comparisons and inter-model comparisons.

$\Delta N = N_{dep} + BNF - N_{loss}$. (3)

Where N_{dep} is the N deposition, BNF is the biological nitrogen fixation, and N_{loss} is the N lost from gaseous and leaching pathways as declared by the models. Note that crop fertilisation is not included here, as it is assumed to remain constant between the 3 simulations.

3 Results

3.1 Control Run Results

TheGlobal C and N cycle acts primarily as a restriction on the productivity of the C cycle, so theoretically the addition of an N cycle should increase the accuracy of modelled GPP. This generation of N models are generally consistent within observational constraints, showing an improvement compared to CMIP5 N models (Arora et al., 2013; Shao et al., 2013). All the models have global total GPP within the error margin of the estimate based on upscaled flux tower data (MTE) by Jung et al., (2011) and are close to the central estimate (see Fig. 1f). Large areas within all the models are well represented. For key regions, such as the Amazon basin, all models are within ~30% of the observed GPP.budgets

The global scale agreement with total observed GPP obscures considerable regional variability and errors in latitudinal patterns of GPP (SI Fig. 1). Consistent with having the highest global GPP of the models, LPJ-GUESS has widespread higher than observed GPP, particularly in the northern hemisphere outside Western Europe (SI Fig. 2). Combined with lower than observed GPP in parts of the tropics, LPJ-GUESS exhibits a too little latitudinal gradient of GPP. CLM5 also has a reduced gradient trend, though to a lesser extent (SI Fig. 2). JSBACH has the opposite issue, with higher than observed GPP in the parts of the tropics, particularly the Amazonia area of northern South America (Fig. 1 and SI Fig. 2). This gives an overall steeper than observed latitudinal gradient of GPP in JSBACH.

The spatial variability in N limitation (LeBauer and Treseder, 2008; Schulte-Uebbing and Vries, 2018) might suggest that models with an N cycle could represent more N limited areas better. The model errors compared with MTE in GPP is largest in the tropics, which have low N limitation (ibid). However, in the far northern high latitudes, which are highly N limited (ibid), the model-MTE agreement is also weak (Fig. 1a-e). As a proportion of observed GPP, all the models show either over or underestimations of GPP in tundra and boreal areas of Eurasia and America. All the models except JSBACH have extensive areas of higher than observed GPP in these high-latitude regions. Where the models are overestimating Control run GPP in N limited areas (LeBauer and Treseder, 2008; Schulte-Uebbing and Vries, 2018) suggests the N cycle model may not be correctly simulating key processes that limit growth in these cold and N limited areas.

Despite having some shared components, CLM5 and CLM4.5 show different GPP patterns, particularly in the high latitudes. CLM5 models photosynthesis rates and GPP differently from CLM4.5 due to the implementation of a flexible C:N scheme (versus fixed C:N in CLM4.5) and the Leaf Use of Nitrogen for Assimilation model (LUNA, Xu et al., 2012 and Ali et al., 2016). Other key differences include stomatal conductance in CLM5 (Medlyn et al., 2011) is based on the N-limited net

Formatted: Heading 2

photosynthesis rather than on potential photosynthesis as in CLM4.5. Additionally, GPP can be more affected by non-N cycle aspects of the models, such as the new plant hydraulic stress function (Kennedy et al., 2019) added in CLM5.

The global net ecosystem productivity (NEP), i.e. the net balance of C assimilation through photosynthesis and C losses to heterotrophic and autotrophic respiration as well as land use changes is calculated via annual bookkeeping (Global Carbon Budget, Friedlingstein et al., 2019). The long-term record (1960-2018) provides a constraint on the global simulated C cycle. All the models are on the correct order of magnitude for NEP and capture the upward trend post 1980 (Fig. 2). JULES is the most accurate model by the GCB measure of NEP, having more years within the bookkeeping constraints than the other models.

The non-N model structure strongly influences NEP; CLM5 and CLM4.5 are the two lowest representations of NEP and follow a similar trajectory (Fig. 2). The CLM models share many model components and this common functioning appears to dominate the signal for NEP. This contrast with GPP, where despite the similarities between CLM4.5 and CLM5 they are not as close in global total as in NEP and have significant differences in spatial pattern (Figs. 1 and 3). This is because while the GPP is heavily influenced by the N cycles in these models, the respiration terms are similarly affected by N and thus result in similar patterns of NEP.

Across the ensemble there is a slight correlation between the global GPP total and NEP. LPJ-GUESS is the highest GPP and highest NEP in the majority of years in the period 1970–2010. Similarly, although CLM5 has the lowest NEP in the same period, CLM4.5 has only slightly more NEP. JSBACH is the second highest global total amongst the models for NEP and GPP. Compared to GPP, including the respiration in this global metric increases the spread of models, reflecting increased uncertainties with increased processes considered.

3.1.1 Global C and N budget

Looking at a range of pools and fluxes emphasises the similarities between the models (see Fig. 3). We use the closest comparable observation-based data to look at the flow of model C and N. The C input of GPP is generally better represented by the models than this flow chart would indicate. In the interests of consistency, we've used 1996–2005 as the baseline for GPP, but when the directly comparable period is used (as in Fig. 1) all the models are well within the uncertainty boundaries. [Fig 3. Flux/Pool flow chart]

The N inputs of deposition and BNF show that most of the uncertainty for N input comes from BNF (Fig. 3). Deposition is a prescribed input with very little variation. BNF on the other hand has a wider observed range, and one model is still outside of it. The three models with the highest BNF (JSBACH, CLM5, and JULES) use NPP based function to calculate BNF. JULES and JSBACH are based on Cleveland (1999)'s empirical large-scale correlation with net primary productivity (NPP). LPJ-GUESS, the lowest BNF model, also uses an empirical correlation from Cleveland (1999), based on evapotranspiration rather than NPP. Thus, even BNF functions from the same source can have very different results (Wieder et al., 2015a), due to the large range of BNF functions and differences in how they are implemented. BNF dominates N input variability both

because of lack of process understanding to constrain model structures and the continued large uncertainty in available observations.

The respiration terms are consistent across all the models, with two errors cancelling each other out. Autotrophic respiration is overestimated in all the models, by up to ~50% of the observed value (Luyssaert et al., 2014; Piao et al., 2010). Heterotrophic respiration is underestimated in all the models, by as much as ~20% compared to the observed values from Bond-Lamberty and Thomson, (2010). The observed value (Bond-Lamberty and Thomson, 2010) is, however, reduced by 33% to account for root respiration, in line with Bowden et al., (1993). Overall, the underestimation of heterotrophic respiration fully compensates for the overestimate of autotrophic respiration. Looking at total observed respiration (102 – 128 PgC yr^{-1}) all the models are within that range.

N losses via gaseous loss to the atmosphere and leaching to groundwater / rivers are not well constrained globally (Galloway et al., 2004). However, the models have a limited range of N loss values. Looking at inputs and losses excluding anthropogenic N addition (BNF + N Deposition – N Loss), all the models have a surplus of N and could be said to be 'open' systems with regard to N balance.

The stocks of C and N are less well constrained and have a larger inter-model range than the fluxes. The Vegetation C is generally high and the Soil+Litter C is generally low, compared to observational estimates from Carvalhais et al., (2014). Despite the large range, three of five models are under observation range (white arrows, Fig. 3). The C and N in the vegetation are not strongly correlated, with the model with the highest VegC having a middle of the model range VegN (JULES). Comparing the C:N of Soil+Litter global total weight the ratios are similar across models, around 11-13:1 C:N and JSBACH 18:1. The higher ratio for JSBACH is due to the 10:1 ratio for slowly decomposing soil carbon (humus) and larger ratio for litter. Overall, the flux values are both better known and represented, while the pools are more variable and have fewer global comparators.

3.1 Modelled Responses to +N and +CO2 Experiments

The +N and +CO2 perturbation experiments (described in the methods) are designed to mimic field experiments undertaken to understand the effects of elevated CO_2 or N. A range of pools and fluxes from the models compared to the closest comparable observation-based data show a good performance overall and emphasises similarities between the models (Fig. 1). For GPP, all the models compare well to the MTE data (Jung et al., 2011) and when the directly comparable time period is used (see SI Fig. 2) the models are all within the MTE range. The global GPP value is underlain by some regional variations between models (SI Fig. 2 and 3).

The total respiration term is similar across all the models and within a range of estimates based of the statistical upscaling of field measurements (102 – 128 Pg C yr^{-1}) (Bond-Lamberty and Thomson, 2010; Bowden et al., 1993; Luyssaert et al., 2007; Piao et al., 2010) but the partitioning between the autotrophic and heterotrophic respiration differs. Autotrophic respiration is overestimated by up to ~50% in all the models (Luyssaert et al., 2007; Piao et al., 2010), while heterotrophic respiration is underestimated by as much as ~20% (Bond-Lamberty and Thomson, 2010). The value from Bond-Lamberty and Thomson,

(2010) was reduced by 33% to account for root respiration in line with Bowden et al., (1993), and without this adjustment the discrepancy would be larger.

Despite similarities in GPP, N inputs differ strongly between the models because of widely varying biological nitrogen fixation (BNF, Fig. 1). N deposition is a prescribed input with small variations resulting from differences in the land-sea mask of the individual models and does not reflect uncertainties in the simulated efficiency of ecosystems to capture nitrogen deposition. BNF on the other hand has a wide range among models. An upscaled meta-analysis of BNF covering the period of approximately 1990 – 2019 (Davies-Barnard and Friedlingstein, 2020) has a range of 52 – 130 Tg N yr^{-1} and only one model is outside of that range. The three models with the highest BNF (JSBACH, CLM5, and JULES-ES) use an NPP based function. While CLM5's process based function includes NPP, JULES-ES and JSBACH use an empirical large-scale correlation with NPP (Cleveland et al., (1999). LPJ-GUESS, the lowest BNF model, also uses an empirical correlation from Cleveland et al., (1999), based on evapotranspiration rather than NPP. Thus, even BNF functions from the same source (Cleveland et al., 1999) can have very different results (Wieder et al., 2015a), due to the large range of BNF functions within the source and differences in how they are implemented (Meyerholt et al., 2016). BNF dominates N input variability both because of lack of process understanding to constrain model structures and the continued uncertainty in available observations.

3.2 Modelled NPP Responses to +CO2 Experiment

A meta-analysis of NPP responses to +200 ppm CO_2 suggests a positive response of 15.6 ± 12.8% (Song et al., 2019) and all the models are within this range (Table 3.). Other meta-analyses of productivity changes with elevated CO_2 give higher ranges of response (Table 3.) and suggest a lower limit of around 12%. Therefore, the fact that the models fall within the uncertainty bounds of the observations is equally indicative of the biases and lack of precision in the observational estimates and their upscaling as the fidelity with which the models can predict local and global response to elevated CO_2.

CLM4.5 has a notably lower NPP response to +CO2 than the other models, despite areas where the absolute values of NPP are low and therefore the proportional changes are large (Fig. 2). This lower response can also be seen in the absolute changes (SI Fig. 4), where the changes are consistently less than the other four models. The low response in CLM4.5 is due to a lack of mechanisms to ameliorate N limitation when C supply increases, for instance via variable C:N ratios or increased BNF (as is the case for CLM5) (Fisher et al., 2018; Wieder et al., 2019).

Despite the seeming agreement of the NPP response to +CO2 at the global scale, the regional patterns in response vary considerably for key biomes (Fig. 2). In high latitude tundra areas, the +CO2 response ranges between near zero (JULES-ES), very low in CLM4.5, JSBACH and LPJ-Guess to high (CLM5). In most models, this region shows sparse vegetation cover and nitrogen availability, allowing for only little increase in response to elevated CO_2, whereas the increased BNF in CLM5 facilitates a response to increasing CO_2 levels. With the exception of JULES-ES, most models predict a large +CO2 response in very dry ecosystems with marginal productivity.

The NPP response of the equatorial region overall (SI Table 2 and SI Fig. 2) to +CO2 ranges from 4% for CLM4.5 to 18% for CLM5 and JSBACH. Looking at latitudinal averages (SI Fig. 4) we can see the overall trends are consistent across most models, and while the percent change varies a lot, the absolute change in NPP shows considerable agreement between models, with the exception of CLM4.5. Model responses of NPP to +CO2 in greater Amazonia however, do not~~These simulations reveal the overall pattern of response of the model to these forcings. Ideally the models would capture the global and regional signals, as indicated by the observed data. However, comparisons between models alone can also provide useful insight into the models' behaviour.~~

~~All the models except CLM4.5 (which has a low response) capture the global mean response of NPP to +CO2 within the range of observations (Fig. 4). CLM4.5 has proportionally high response in arid areas of North Africa and central Asia, but the absolute values of NPP in these regions are low.~~

~~Regional patterns in response vary considerably for key biomes. +CO2 response ranges between none (JULES) and high (CLM5) for high latitude tundra areas which have little vegetation carbon in the present day but are critical in future climate change scenarios. Model responses of NPP to +CO2 in greater Amazonia also don't~~ reach a consensus. Comparing the response in the Amazonia region with that of coastal regions of northern South America, the JSBACH response is lower, CLM5 and LPJ-GUESS higher, and JULES-ES and CLM4.5 are approximately the same. ~~The~~ JSBACH's dip in +CO2 NPP response ~~of Amazonia~~at the equator (compared to ~~+CO2 ranges from ~10%~~surrounding areas) can also be seen in the absolute values averaged by latitude (SI Fig. 4). The process responsible for ~~CLM4.5 to ~30%~~this spatial pattern is currently unclear, but may be associated with the strongly enhanced GPP simulated by the model for ~~CLM5 and LPJ-GUESS. Therefore, although the models reach a majority consensus on +CO2 NPP effects overall, the important regional details are still contradictory.~~this region compared to observation-derived estimates (SI Fig. 2).

~~The response to +N in the models shows a wider spread than for +CO2 or the control scenario, with 3/5 models outside of the observed limits for NPP response (Fig. 5). JSBACH and JULES have global NPP responses to +N of 2.3% and 2.6% respectively. For JSBACH, this is explained by the concept of CO_2-induced nitrogen limitation (Goll et al., 2017) which assumes an absence of N limitation in pre-industrial state due to long-term plant adaptation to environment during the Holocene. CLM4.5 has the highest response, on account of its very high initial N limitation, at 25%. However, even within models with similar levels of +N response, there are regional differences in distribution.~~

~~The most prevalent spatial trend is regions either having sensitivity to +N or +CO2, but not both (see Fig. 4 and 5). This is most pronounced in JULES, where high latitude regions have a response to +N but no response to +CO2. However, to a lesser extent this is also seen in CLM4.5, CLM5, and LPJ-GUESS. Wieder et al., (2019) also found that there was a trade-off between +N impact and +CO2 impact in CLM4, CLM4.5 and CLM5. There is spatial pattern and global total similarity between CLM4.5 and CLM5 in their NPP response to +N, in contrast to their differing responses to +CO2. Only JSBACH doesn't appear to have a dichotomy between +N and +CO2 response.~~

~~Although the models generally have lower than observed global NPP response to +N, regionally there are areas of over and underestimation. Tundra biome response is high in CLM5 and JULES, and present in LPJ-GUESS and CLM4.5 (Fig. 5).~~

~~This is an overestimation of the observed 35% N fertilisation response (LeBauer and Treseder, 2008). Conversely, the +N response of temperate grassland (53%) (LeBauer and Treseder, 2008) is underestimated by all the models, with JULES and JSBACH showing less than 20% response in most temperate areas.~~

3.3 Modelled NPP Responses to +N Experiment

The response to +N in the models shows a binary distribution, with models either having a high (>17%) or low (<3%) level of response (Fig. 3).

A meta-analysis of NPP responses to +50 kg N ha^{-1} yr^{-1} suggests a positive response of 3 – 10.5% (Song et al., 2019) but none of the models are within this range (Table 3.). Other meta-analyses of productivity changes with increased N give higher ranges of response (11 – 39.8%), encompassing three of the five models (Table 3). As both a percent change and absolute change (see SI Fig. 5) JULES and JSBACH show much lower +N NPP response than the other models considered here. CLM4.5 has the highest response (23%), on account of its high initial N limitation (Koven et al., 2013).

The tundra biome response is high in CLM5 and JULES-ES, and lower but present in LPJ-GUESS and CLM4.5 (Fig. 3 and SI Fig. 5). If low NPP is excluded then the tundra mean response across models is 2 – 9% (SI Table 2), much lower than the average of observations compiled by LeBauer and Treseder, (2008) of 35% (95% confidence interval 12 – 64%). There is a high response to +N in Africa & Australia in CLM4.5, CLM5, and LPJ-GUESS, despite aridity likely limiting increase in NPP in absolute, if not relative, terms~~., but insufficient observations to make meaningful comparisons.~~ One area of agreement between the models is the lack of +N response ~~in the tropics, especially over~~of the Amazonian region~~. This~~ (Fig. 3) which is ~~also~~ consistent with observations which show just a 5% non-significant +N response in ~~the tropics~~tropical forests (Schulte-Uebbing and Vries, 2018). ~~With~~However, when other tropical regions are included the ~~exception of the tropics, the models are underestimating the global~~ +N NPP response rise to 18 – 27% in LPJ-GUESS, CLM4.5 and ~~showing the response in biomes inconsistent~~CLM5, with ~~the observations.~~ JULES-ES and JSBACH remaining low (SI Table 2).

3.4 Comparison of NPP +N and +CO2 Responses

It might be anticipated that there would be a relationship between the +N and +CO2 responses, as an ecosystem (model) that is less N limited could respond more strongly to increased atmospheric CO_2 (Meyerholt et al., 2020). Since a lack of response could indicate sufficient supply or saturation of either N or CO_2, this could enable increased NPP if the area were limited by the other (C or N) nutrient. This is the case in the models at small model scales, but does not scale to either the regional or global level. The prevalent grid cell level spatial trend is bimodal, with grid cells either having a strong sensitivity to +N or +CO2, but not both (see Fig. 4). Comparing percent change emphasises the dichotomy of +N and +CO2 effects, with most values clustered near either zero for +N or zero for +CO2, but SI Fig. 6 shows that there is no positive relationship or heterogeneous distribution in the absolute values either. The bias toward +CO2 is clear for JSBACH and JULES-ES, with most values varying in +CO2 sensitivity but not +N (this can also be seen in the absolute anomalies, see SI

Fig. 6). A slight tendency towards the reverse is true for CLM4.5, CLM5, and LPJ-GUESS, with more points having a strong +N response and a weaker +CO2 response (Fig. 4). Altogether, LPJ-GUESS and CLM5 show the most areas with both +N and +CO2 sensitivity. Wieder et al., (2019) found that there was a trade-off between +N impact and +CO2 impact in CLM4, CLM4.5 and CLM5, and this seems to be true for our ensemble of models too.

The latitudinal distribution of response shows similarities across models, with high latitudes (shown in purple in Fig. 4) generally more +N sensitive, and the mid latitudes (red to orange on Fig. 4) more +CO2 sensitive. While negative NPP values are present in both +N and +CO2 simulations they occur in different places, with negative NPP occurring in hot arid areas for +N and cold arid areas for +CO2 (Fig. 2, 3, and 4). In hot arid areas +N increases simulates GPP and plant growth but also plant respiration, which then exceed the additional productivity, giving a decrease in NPP. Such model behaviour has been noted before (Meyerholt et al., 2020), however, it is not evident that such a process would occur in nature. The negative values in all models except CLM4.5 also appear to have a regional bias, with a small number of grid cells responding negatively to both +CO2 and +N in CLM5, JSBACH, and JULES-ES in the subtropics and a larger number of negative values in the subtropics in LPJ-GUESS (Fig. 4). These arid areas appear to be sensitive to neither +N nor +CO2, due to water limitation.

We can gain further insights by considering the relationship between responses to +CO2 and +N (Fig. 5) by forest biome. The ideal for the models is to be in the area where the observations for +N and +CO2 intersect. Two of the models achieve this partially, JSBACH and CLM5, by having tightly clustered forest vegetation C (VegC) response to +N and forest NPP response to +CO2. The dichotomy between +N and +CO2 NPP response is averaged out at this scale and the models show little of the relationship between the +N response and +CO2 response seen at the grid cell level (Fig. 4 and 5).

According to observations from N addition experiments collated in we would expect models to have biome level variation in +N response (LeBauer and Treseder, 2008; Schulte-Uebbing and Vries, 2018). Schulte-Uebbing and Vries, (2018) show that tropical forest +N VegC response is lowest and boreal and temperate forest response higher (Fig. 5). While LPJ-GUESS and CLM4.5 capture some variation between averaged biomes, none of the models have the biome responses in the correct order (Fig. 5). However, all the models except LPJ-GUESS and CLM4.5 have tropical +N response in the correct range. LPJ-GUESS is the only model to have the boreal +N response in the correct range. It is the boreal response that seems to be the main issue, as most of the models show increased +N response compared to the tropics for temperate regions, but dampened response for boreal regions. Therefore, although the global values of response are acceptable, the relative spatial patterns show limitations in the reliability of all the models.

3.5 N Budget Responses to +N and +CO2

The models' responses in different ~~parts~~components of the N budget reflect and affect their overall N sensitivity (Fig. 6). N inputs of BNF and N deposition and loss (we only consider the sum of leaching and gaseous loss so as to be consistent between models) are ~~very~~ similar between all the models in the Control simulation~~. This leads to small differences in the N~~

balance between the models. The N balance appears to be weakly anti-correlated with total GPP (see (Fig. 1f and Fig. 6a). However, theThe uptake of N by vegetation varies a lot more strongly between models, reflecting differing levels of N mineralisation, model structure, and assumed N requirements for growth, as also reflected by the different amounts of C and N pools depicted in Fig. 1.

The largest responses to +N and +CO2 of input and loss do not necessarily correlate with either N uptake or changes to productivity. The weak response of NPP to +CO2 in CLM4.5 would suggest only small changes in uptake compared to the other models (Fig. 4). However, the +CO2 induced changes in BNF, loss, N balance and uptake CLM4.5 are approximately middle of the model range (Fig. 6). CLM5 has a large increase in N balance from BNF and decreased loss (positive change in gaseous loss and leaching in Fig. 6b) but decreased uptake. This is due to a model feature that CLM considers BNF as N delivery to the plant by nodules, and thus doesn't count it as uptake from the soil. CO_2 fertilisation increases BNF in CLM5 due to the cost based portioning of C expenditure on N uptake that reduces the active uptake of mineralised N and increases BNF when soil N stocks are low (on account of increased growth from +CO2) (Fisher et al., 2010, 2018).

TheChanges in the N budget components to +CO2 and +N (Fig. 6b and 6c) are not straightforwardly related to changes to productivity (Fig. 2 and 3). For instance, the weak response of NPP to +CO2 in CLM4.5 would suggest only small changes in uptake compared to the other models (Fig. 2 and 6), however, the +CO2 induced changes in uptake CLM4.5 are higher than that of LPJ-GUESS (Fig. 6b). Similarly, CLM5 has the largest increase in N balance for +CO2 (Fig. 6b) amongst the models, but this does not correspond to a larger response of NPP or uptake response to elevated CO_2. Nevertheless, Fig. 6b reveals a number of important characteristics of the N cycle response to +CO2 underlying the NPP response presented in Section 3.2. For all models except CLM5, which shows a strong response of BNF to elevated CO_2, reduced N losses are an important reason for the increased N balance of the ecosystem, which facilitates an increase in NPP in the absence of changes in ecosystem stoichiometry. For all models except CLM5, plant N uptake under elevated CO_2 is more enhanced than this change in the N balance of the ecosystem, implying a net transfer of nitrogen from the soil to vegetation.

Conversely, the N uptake changes in JULES-ES and JSBACH reflect their sensitivity of productivity to +N and +CO2. These models have the highest uptake changes for +CO2 (Figs. 2,3, and the smallest uptake changes for +N (Fig. 6). For JULES-ES we can see that this is driven by changes in loss, particularly for +N, which leads to a much smaller increase in N balance in JULES-ES than the other models. In common with all the models, in JULES-ES the N loss term is a fixed fraction of the mineralisation flux and the soil N pool size. In contrastHowever, JSBACH has less than half the increase in N loss of JULES-ES in the +N simulation (Fig. 6c) and almost no change in NUE (Fig. 7d). This suggests that in both JULES-ES and JSBACH there is effectively very little unmet N demand in the Control scenario andbut whereas JULES-ES loses the extra N, JSBACH retains it in the soil.

Two of the most important factors for plants' use of N are the availability and demand for N use. The variability of these processes is determined primarily by the BNF and NUE respectively, which are both known to be affected by increased CO_2 and N. Neither BNF nor NUE are well constrained by observations and thus show heterogeneity of response between

models. The models come to quite similar global results (e.g. for GPP (Fig. 1) or NPP +CO2 response (Fig. 4)) through very different NUE and BNF behaviours (Fig. 7).

The BNF responses to +CO2 ofN in the models differ from the average response recorded in a globalin magnitude (Fig. 7b) and mostly are smaller than a meta-analysis of CO_2 manipulation suggests (Liang et al., 2016) (Fig. 7a). The observation here is an upscaling of studies rather than a representative global observation and therefore the comparison needs to be interpreted accordingly.. Only JULES'JULES-ES' responses, in all regions except for boreal forests, and CLM5's boreal response are within the range of the meta-analysis. of observations. CLM5 is a clear outlier, with a large increase in BNF as it. CLM5 takes a C cost approach to estimated BNF, which is different to the other models (Table 1).), and BNF can be acquired for a relatively fixed amount of C (Houlton et al., 2008), when C availability increases under +CO2 the BNF, in this case more than field data would suggest. Fisher et al., (2018) conduct a parameter sensitivity analysis of both +CO2 and +N fertilization and illustrate that both responses are sensitive to the maximum fraction of carbon which is available for fixation (a proxy for the fraction of N fixing plants) and thus this response might be due to this.

The models' BNF response to +N shows one of two responses: a small increase in JULES, CLM4.5, and JSBACH; or a large decrease in CLM5 and LPJ-GUESS (Fig. 7). The latter models capture the correct BNF sign of response to +N of a decrease, though the amplitude is too high. The former models estimate BNF as a function of NPP which results in increased BNF whatever the source of the additional NPP is. This gives a counter-intuitive BNF response of more BNF being produced, even though there is already sufficient N. Overall, there is little evidence for any of the BNF algorithms performing well.

The NUE responses allow comparison between models, though comparisons with observations are limited by lack of field studies. The NUE responses of the models are similar to that of BNF, with CLM5 having the largest response to +CO2 and CLM4.5 the smallest (Fig. 7). All models have an increase in NUE with +CO2, which is in line with current theory (Walker et al., 2015). CLM4.5 has low NUE response to +CO2 due to fixed C:N ratios, which allow little change in NUE. The other models allow either more allocation to wood or flexible C:N that results in the larger increases of NUE. CLM5 has large changes in NUE because BNF is not considered as uptake.

There is regional variation in NUE response to +N between biomes in some models. CLM5 and LPJ-GUESS are distinct in their response to +N compared to the other models, but do not share the same geographical spread of response. JSBACH, LPJ-GUESS, and CLM4.5 all have globally consistent responses. JULES and CLM5 have more spatial variation, with different responses between boreal and tropical regions. All the models except JULES reduce NUE in response to +N. Though no empirical measurements are currently available for NUE response to +N, it seems more likely that NUE would increase when N was scarce and vice versa. The large variations in signal and sign of BNF and NUE response between models suggests there is still progress to be made.

We can gain further insights by considering the relationship between the response to +CO2 and +N (Fig. 8). and thus when C availability increases under +CO2 the BNF in CLM5 increases. Fisher et al., (2018) conducted a parameter sensitivity analysis of both +CO2 and +N fertilization which illustrates that both responses are sensitive to the maximum fraction of C which is available for fixation (a proxy for the fraction of N fixing plants and their efficiency). However, the correct

parametrisation of this fraction of C available for fixation is not well known and further field studies are required. The BNF +CO2 response in the other four models is determined by their simple empirical BNF equations (see Table 1) based on NPP or evapotranspiration. However, new analysis suggests that simple empirical relationships cannot well represent BNF (Davies-Barnard and Friedlingstein, 2020).

The models' BNF response to +N shows one of two responses: a small increase in JULES-ES, CLM4.5, and JSBACH; or a large decrease in CLM5 and LPJ-GUESS (Fig. 7b). The latter models capture the correct BNF sign of response to +N of a decrease according to the meta-analysis of Zheng et al., (2019), though the amplitude is too large. The former models estimate BNF as a function of NPP resulting in increased BNF whatever the source of the additional NPP is. Thus even when there is sufficient N more BNF is produced in these models when NPP increases, despite observational evidence (Zheng et al., 2019) showing this is wrong response and that facultative BNF reduces and obligate BNF is out-competed in these circumstances (Menge et al., 2009). Overall, there is little evidence for any of the BNF functions performing well, primarily due to a deficit of robust model parameterisations and parameter values.

The NUE responses allow comparison between models, though comparisons with observations are limited by a lack of field studies. With the exception of JULES-ES in the boreal region (Fig. 7c) all models have an increase in NUE with +CO2 in line with the current theory of Walker et al., (2015). Since the BNF in JULES is directly related to NPP, so the reduction in NUE indicates excess N in the system from mineralisation, possibly related to soil warming, in boreal regions with +CO2, leading to decreased N uptake. CLM4.5 has low NUE response to +CO2 due to fixed C:N ratios, which allow little change in NUE. The other models allow either more allocation to wood or flexible C:N that results in the larger increases of NUE. CLM5 has large changes in NUE, and like JULES' boreal region this indicates a source of N other than BNF.

There is regional variation in models' NUE response to +N between biomes but all the models in our ensemble reduce NUE in response to +N. CLM5 and LPJ-GUESS are distinct in their larger NUE response to +N compared to the other models, but do not share the same geographical spread of response. There is little consistency between models as to which regions have the largest change in NUE. CLM5 has the largest NUE change in the temperate region, whereas in JULES it occurs in the boreal region. No empirical measurements are currently available for NUE response to +N. On the basis that scarcity encourages more frugal use of scarce resource a hypothesis could be that NUE could decrease with increased N availability, as the models show. However, water-use efficiency suggests an alternative hypothesis, as it tends to reduce during drought (Yu et al., 2017). The large variations in signal and sign of BNF and NUE responses to +N treatment between models suggest there is considerable uncertainty in our understanding.

~~The ideal for the models is to be in the area where the observations for +N and +CO2 intersect. Two of the models achieve this partially, JSBACH and CLM5, by having tightly clustered vegetation C (VegC) response to +N and NPP response to +CO2. One model, CLM4.5 has all three biomes outside of both the +N and +CO2 observational ranges.~~

~~[Figure 8 +N vs +CO2]~~

~~It might be anticipated there would be a relationship between the two responses, as an ecosystem (model) that is less N limited would be expected to respond more strongly to increased atmospheric CO_2. However, the models show little~~

~~relationship between the +N response and +CO2 response (Fig. 8). At a regional level the models either vary in +N response but not +CO2 (LPJ-GUESS and CLM4.5) or vary little in either (JULES, JSBACH and CLM5).~~

~~According to observations we would expect models to have biome differences in +N response, but none of the models correctly capture this (Fig. 8). All the models except LPJ-GUESS and CLM4.5 have tropical +N response in the correct range (Fig. 8). LPJ-GUESS is the only model to have the boreal +N response in the correct range. The observations from N addition experiments (Schulte-Uebbing and Vries, 2018) show lower tropical +N response than boreal or temperate, but none of the models have the biome responses in the correct order. It is the boreal response that seems to be the main issue, as most of the models show increased +N response compared to the tropics for temperate regions, but dampened response for boreal regions. Therefore, although the global values of response are acceptable, the relative spatial patterns show limitations in the reliability of all the models.~~

4 Discussion

~~This second generation of N models in CMIP do well at reproducing key global metrics, such as GPP and NEP and generally show magnitudes of responses to increased atmospheric CO_2 commensurate with those recorded in field studies. The response to N addition is more varied. Though there is considerable variation in the values for many fluxes and pools, the models tend to represent net fluxes well where they are constrained by observations. In the case of respiration, overestimates of autotropic and underestimates of heterotrophic respiration combine to give good overall agreement. The fluxes are generally better represented by the models than the pools, possibly due to the tighter observational constraints on fluxes. However, the N processes have considerable heterogeneity and there are regional differences in the robustness of response.~~

~~Two models of the five (JULES and JSBACH) have virtually no productivity response to increased N availability, suggesting they don't have any significant N limitation (Fig. 5). There are four substantial similarities between these two models (Table 1): the use of NPP to determine BNF; GPP is not directly affected by N; the use of dynamic (as oppose to prescribed) vegetation; and the assumption of no pre-industrial N limitation. Since we lack observations of N limitation for the pre-industrial period, the assumption that increases in atmospheric CO_2 are the cause of N limitation in the present day is difficult to disprove. However, this assumption drives other model decisions (such as N limitation not being incorporated into the GPP equation, see Table 1), which appear to lead to models being under-sensitive to N compared to observations.~~

In this paper, we investigated the performance of five nitrogen-enabled land surface models that are part of current generation Earth System Models used in the framework of CMIP6 (Eyring et al., 2016). These new N-enabled land surface models in CMIP6 reproduce key global carbon cycle metrics. Nevertheless, despite the importance of N availability for regional productivity, there is large and unconstrained uncertainty in the magnitude of the global and regional N fluxes (Fig. 1).

We have focused on three general ~~parts~~components of N-enabled models that affect the plant N uptake and eventual productivity: N inputs via BNF; NUE; and the N losses. We find that all three show considerable heterogeneity of response

between models. Previous studies suggest that stoichiometric controls and the processing of soil organic matter are important for a realistic +CO2 response (Zaehle et al., 2014a). These are essentially contributory factors to NUE, where we find large variation between models. (Fig. 7). The lack of well constrained observations for global and biome-level NUE and N loss responses make these areas that need more work. N loss is particularly challenging, as there are multiple paths (leaching, flooding, gaseous loss, fire, land use change, etc.) and forms (N_2O, N_2, etc.) of loss and each model represents these in different ways. More observational studies and syntheses of existing observations are needed to quantify the nitrogen cycle in different biomes. In particular, better constraints are needed for the N cycle response to perturbations.

The models mostly represent high latitude northern hemisphere regions less well than other parts of the world, in part because of the unique challenges these areas set for models. We see in Fig. 8 that the N response of the boreal biome is underestimated, and in Fig. 5 the conflicting predictions of N response in the very high latitudes.All the models show a global average productivity response to increased atmospheric CO_2 commensurate with those recorded in field studies. However, the regional responses and mechanisms behind this response vary widely, resulting from the interaction of the instantaneous physiological response to elevated CO_2 (e.g. Ainsworth and Long, (2005), which is embedded in all five models (but see Rogers et al., (2017)), with limitations imposed by temperature, water, light, and nitrogen, as well as the response-time of vegetation dynamics. This large regional variance highlights the need for a more comprehensive observational data-base to constrain responses to elevated CO_2, particularly in under-sampled regions such as the high arctic and tropical semi-arid regions (Song et al., 2019). Tundra and arctic responses vary widely and are associated with the representation of BNF. In LPJ-Guess and CLM5 the responses in semi-arid tropical ecosystems is smaller than that of temperate ecosystems and the other models, suggesting a combined effect of water- and nitrogen-limitation of soil organic matter decomposition and thus nitrogen availability that is not compensated for by changes in BNF.

The growth response to N addition across models is more varied. Two of the five models (JULES-ES and JSBACH) have little productivity response to increased N availability, indicating that they do not have any significant limitation of the carbon cycle by N availability (Fig. 3). There are four substantial similarities between these two models (Table 1): (i) the use of NPP to determine BNF; (i) a direct control of NPP by N availability, whereas photosynthetic C uptake (GPP) is not directly affected by N; (iii) the use of dynamic (as oppose to prescribed) vegetation; and (iv) the assumption that N availability in pre-industrial times was sufficient to sustain the carbon cycle everywhere on land, and that observed present-day N limitation was a result of anthropogenic changes, most notably increased CO_2 (Goll et al., 2017).

The hypothesis behind the assumption of no pre-industrial N limitation is that prior to industrial times, the conditions of natural terrestrial ecosystems were stable for sufficient time to permit any lack of N availability to be filled by biological nitrogen fixation (Thomas et al., 2015). Consequently, JULES-ES and JSBACH first simulate a reference run without N interactions, mimicking the dynamics of the C-cycle only version of these models. After a C equilibrium has been reached these models add a second spin-up simulation, in which C and N dynamics are coupled. While there is evidence for wide-spread (co-) limitation of NPP in recent decades (LeBauer and Treseder, 2008; Song et al., 2019; Vitousek and Howarth, 1991), there is insufficient data to test the hypothesis of no pre-industrial N limitation. A summary by Thomas et al., (2015)

suggests reasons that pre-industrial productivity of terrestrial ecosystems was affected by ecosystem N availability, e.g. the presence of unavoidable losses to denitrification, or the competitive exclusion of nitrogen fixing species as ecosystems mature. The inability of JULES-ES and JSBACH to simulate observed N addition responses comparable to models without this assumption suggests that models assuming that pre-industrial N availability does not limit vegetation growth may be missing an important component of the nitrogen cycle constraint on the global carbon cycle. No pre-industrial N limitation also drives other model decisions (such as N limitation not being incorporated into the GPP equation, see Table 1), which may further contribute to the models being under-sensitive to N compared to observations.

The models mostly represent changes in productivity from +N in high latitude northern hemisphere regions less well than other parts of the world as a percentage, as covered in the results section 3.3, Fig. 5, and SI Table 2. While the low NPP of these regions make them more likely to have high percentage increases, when these are excluded the mean +N response across the models is 2 – 9%, well below the range of a meta-analysis of observations. Thus the model response is either too low or too high. High latitude tundra is a critical but difficult to model biome because of the potential for release of methane (Nauta et al., 2015), permafrost C and N release (Anisimov, 2007; Burke et al., 2012; O'Connor et al., 2010), and albedo changes with vegetation expansion (Myers-Smith et al., 2011) and the difficulty in representing large amounts of C stored in soil. ~~This partly accounts for why tundra NPP response to +N is overestimated in all the models, because the processes in this biome are complex, but differing vegetation representation may also contribute. If +CO2 does increase NPP in tundra areas, the substantial biogeophysical albedo changes may be larger than the carbon changes (Betts, 2000; Cohen et al., 2013). Tundra is a low productivity area and though the absolute change from +N or +CO2 is small, as a proportion it is large and could have wide ranging effects. A fully integrated model that accounts correctly for all of these is not yet possible.~~This complexity in C and N cycle is not always well understood or represented in models and therefore could limit the ability of models to provide accurate responses to perturbation. A fully integrated model that accounts correctly for all of these is not yet possible but is necessary to reduce uncertainties.

~~The northern part of South America of the~~The greater Amazon basin is a critical area of interest for ~~LSMs~~the future of the terrestrial carbon balance under climate change. ~~The~~Our simulations show that for most models, NPP in this area increases ~~in some of the models for~~with +CO2, but all the models find a small or no change in NPP with +N. The small +N response is ~~small because~~consistent with the idea that large rates of BNF in tropical rain forests make these ecosystems not strongly N limited. This result supports the idea that favourable climate conditions ~~already give~~cause a high leaf area index (LAI) in this part of the tropics, ~~and~~such that there is little margin for increased NPP ~~(Fisher et al., 2018).~~from +N (Fisher et al., 2018). For +CO2 there is the potential for increased NPP because ~~the~~of either increase in NUE ~~increases~~or decreases in N losses, giving productivity increase without an increase in LAI. ~~However, reducing~~Reducing the uncertainty in NPP response to +CO2 is important, as the moist tropics represent a significant proportion of the world's aboveground biomass and therefore the size of the overall terrestrial sink will be influenced by the CO_2 uptake in this biome.

Part of the uncertainty in the models comes from the reanalysis climate dataset used to drive the models. CRU-NCEP was chosen for the good spatial and temporal coverage, but some biases exist in the data compared to climatologies such as

WATCH (Weedon et al., 2011). Offline simulations driven by low forcing frequency (six-hourly) CRU-NCEP data significantly overestimate evapotranspiration in regions with convective rainfall types and thereby could affect stomatal conductance and photosynthesis (Fan et al., 2019). This does not affect all the models equally, as some are ~~more sensitive to the driving climatology. JSBACH's performance is particularly affected by CRU-NCEP, and JULES and LPJ-GUESS may also be strongly affected due to its dynamic vegetation. Lawrence et al. (2019)~~known to be sensitive to the driving climatology. JSBACH, JULES-ES and LPJ-GUESS may be particularly strongly affected due to their dynamic vegetation. Lawrence et al., (2019) show that CLM5 corresponds best to benchmarks with GSWP3 forcing dataset (Hurk et al., 2016)~~.~~ and work with JULES shows that climate forcing is the biggest cause of variance of those considered (Ménard et al., 2015). Responses may partially be shaped by other limiting factors such as water availability, which will be handled differently between models limiting the insight on the exact processes that control model responses to change. ~~The climatology should have a small effect on the model response but may skew the global values the models with dynamic vegetation since the biomes may be incorrectly represented globally.~~

As well as uncertainty in the models, the observational data also has uncertainties and limitations. Global benchmarks are approximate measures, as multi-faceted process mechanics are integrated over large domains and generalized, e.g., over climate zones that are inherently variable. Of the limited global or regional observations available, many use interpolation or proxies such as satellite data to upscale relatively small amounts of direct observational data. In particular, the perturbed responses have uncertainties beyond the spread of the observed responses because of the small observation basis and potential biases in the geographical sampling. ~~The~~One of the +N global ~~response~~responses cited is based on 126 values from LeBauer and Treseder, (2008) but may over-estimate the global response by including high responses from young tropical soils. The NPP response to +CO2 response for woody plants total above ground biomass (Fig. 5) is based on just ~~23~~16 experiments~~, of which only few have been conducted at field scale~~ (Baig et al., 2015)~~. Similarly, the accuracy of global GPP data from satellite data is based on modelling assumptions that convert observations to GPP values (Koirala et al. forthcoming).~~, making the upscaling to biome scale less reliable than if more data were available. These meta-analyses combine measurements from a range of time periods and places, and different conditions (e.g. graduated or instantaneous perturbations) and thus global model experiments cannot be expected to be identical. Hence statements about the marginal issues of model accuracy are unlikely to be robust as further observational constraints may alter the perspective.

5 Conclusions

~~This is the first systematic comparison of the responses to N and CO_2 in LSMs with terrestrial N cycles used within the CMIP6 endeavour. CMIP6 models with interactive N cycling show a reduction in carbon feedback uncertainty comparing to a carbon-only models (Arora et al., 2019). Our analysis elucidates some of the strengths and weaknesses of LSMs that include an N cycle. The five models considered here have good overall agreement with global and tropical observations but are less robust in high latitude regions.~~

This is the first systematic comparison of the responses to increased N (+N) and CO_2 (+CO2) in LSMs with terrestrial N cycles used within the CMIP6 endeavour. The five models considered here have fair overall agreement with global and tropical observations but are less robust in high latitude regions.

The models are not equally sensitive to either +CO2 or +N, with ~~responses varying by up to an order of magnitude.~~individual grid cells tending to either be +N or +CO2 sensitive. However, at the regional and global scale this pattern is averaged away and there is little correlation. Within this ensemble there is clear distinction between models that show strong N limitation, e.g. CLM4.5, which has ~~very~~a low NPP response to +CO2, and models that show very weak N limitation, e.g. JULES-ES and JSBACH ~~has~~which have a low NPP response to +N. ~~Though~~The two models with intermediate N limitation (CLM5 and LPJ-GUESS) capture the global scale response to +CO2 and +N reasonably well. However, although CLM5 performs well by many metrics, it ~~has large~~is an outlier compared to other models or observations in its BNF and the NUE ~~responses to CO_2.~~response to CO_2 appears to be larger than supported by observations. Similarly, LPJ-GUESS captures NPP responses to +CO2 and +N well at the global level but overestimates the ~~VegC~~vegetation C response to +N in ~~non-~~forested tropical and temperate biomes.

The wide range of empirical or semi-mechanistic representations for key processes such as BNF, NUE, and N loss, show how important further process understanding is for many parts of the N cycle. These parts of the models are influential, but because N cycle components are a recent addition to LSMs, fewer data are available than for carbon cycle components. ~~Despite a larger range of observational datasets, the carbon cycle pools are not well constrained and have large differences between models, especially for the soil carbon pools.~~ Consequently, better observational constraints are required to understand whether models are working appropriately, even ~~where~~when the process understanding is improved.

Acknowledgements

Data used to generate the figures, plots, and tables in this paper will be archived on the University of Exeter repository upon publication, in line with the appropriate guidelines.

~~TDB acknowledges~~Authors acknowledge the work ~~and analysis~~ by Dr~~.~~ Johannes Meyerholt~~, PhD~~ (deceased 2020) that formed the basis of this paper.

Authors acknowledge funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 641816 Coordinated Research in Earth Systems and Climate: Experiments kNowledge, Dissemination and Outreach (CRESCENDO).

SZ acknowledges support by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 647204 (QUINCY).

PF ~~acknowledges~~and SZ acknowledge funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 821003 (4C project).

RF was supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement 1852977.

BS acknowledges this study is a contribution to the Strategic Research Area MERGE.

References

Ainsworth, E. A. and Long, S. P.: What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2, New Phytologist, 165(2), 351–372, doi:10.1111/j.1469-8137.2004.01224.x, 2005.

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(18), 6801–6843, doi:10.1175/JCLI-D-12-00417.1, 2013.

Anisimov, O. A.: Potential feedback of thawing permafrost to the global climate system through methane emission, Environ. Res. Lett., 2(4), 045016, doi:10.1088/1748-9326/2/4/045016, 2007.

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(15), 5289–5314, doi: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, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C., Krasting, J., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T. and Ziehn, T.: Carbon-concentration and carbon-climate feedbacks in CMIP6 models, and their comparison to CMIP5 models, Biogeosciences Discussions, 1–124, doi:https://doi.org/10.5194/bg-2019-473, 2019.

Baig, S., Medlyn, B. E., Mercado, L. M. and Zaehle, S.: Does the growth response of woody plants to elevated CO2 increase with temperature? A model-oriented meta-analysis, Global Change Biology, 21(12), 4303–4319, doi:10.1111/gcb.12962, 2015.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A. and Gedney, N.: The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes, Geoscientific Model Development, 4(3), 677–699, 2011.

~~Betts, R. A.: Offset of the potential carbon sink from boreal forestation by decreases in surface albedo, Nature, 408(6809), 187–190, doi:10.1038/35041545, 2000.~~

Bonan, G. B., Oleson, K. W., Fisher, R. A., Lasslop, G. and Reichstein, M.: Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4, Journal of Geophysical Research: Biogeosciences, 117(G2), doi:10.1029/2011JG001913, 2012.

Bond-Lamberty, B. and Thomson, A.: A global database of soil respiration data, Biogeosciences, 7(6), 1915–1926, doi:https://doi.org/10.5194/bg-7-1915-2010, 2010.

Bowden, R. D., Nadelhoffer, K. J., Boone, R. D., Melillo, J. M. and Garrison, J. B.: Contributions of aboveground litter, belowground litter, and root respiration to total soil respiration in a temperate mixed hardwood forest, Can. J. For. Res., 23(7), 1402–1407, doi:10.1139/x93-177, 1993.

Burke, E. J., Hartley, I. P. and Jones, C. D.: Uncertainties in the global temperature change caused by carbon release from permafrost thawing, The Cryosphere, 6(5), 1063–1076, doi:10.5194/tc-6-1063-2012, 2012.

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(7521), 213–217, doi:10.1038/nature13731, 2014.

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A. and Navarra, A.: Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model, Journal of Advances in Modeling Earth Systems, 11(1), 185–209, doi:10.1029/2018MS001369, 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, Geoscientific Model Development, 4(3), 701–722, doi:10.5194/gmd-4-701-2011, 2011.

~~Cohen, J., Pulliainen, J., Ménard, C. B., Johansen, B., Oksanen, L., Luojus, K. and Ikonen, J.: Effect of reindeer grazing on snowmelt, albedo and energy balance based on satellite data analyses, Remote Sensing of Environment, 135, 107–117, doi:10.1016/j.rse.2013.03.029, 2013.~~

Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., Perakis, S. S., Latty, E. F., Fischer, J. C. V., Elseroad, A. and Wasson, M. F.: Global patterns of terrestrial biological nitrogen (N2) fixation in natural ecosystems, Global Biogeochemical Cycles, 13(2), 623–645, doi:10.1029/1999GB900014, 1999.

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, Journal of Advances in Modeling Earth Systems, 10(11), 2731–2754, doi:10.1029/2018MS001354, 2018.

Davies-Barnard, T. and Friedlingstein, P.: The Global Distribution of Biological Nitrogen Fixation in Terrestrial Natural Ecosystems, Global Biogeochemical Cycles, 34(3), e2019GB006387, doi:10.1029/2019GB006387, 2020.

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, Geoscientific Model Development, 9(5), 1937–1958, doi:https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fan, Y., Meijide, A., Lawrence, D. M., Roupsard, O., Carlson, K. M., Chen, H.-Y., Röll, A., Niu, F. and Knohl, A.: Reconciling Canopy Interception Parameterization and Rainfall Forcing Frequency in the Community Land Model for Simulating Evapotranspiration of Rainforests and Oil Palm Plantations in Indonesia, Journal of Advances in Modeling Earth Systems, 11(3), 732–751, doi:10.1029/2018MS001490, 2019.

Fisher, J. B., Sitch, S., Malhi, Y., Fisher, R. A., Huntingford, C. and Tan, S.-Y.: Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation, Global Biogeochem. Cycles, 24(1), GB1014, doi:10.1029/2009GB003621, 2010.

Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., Fisher, J. B., Shi, M., Walker, A. P. and Lawrence, D. M.: Parametric Controls on Vegetation Responses to Biogeochemical Forcing in the CLM5, Journal of Advances in Modeling Earth Systems, 2879–2895, doi:10.1029/2019MS001609@10.1002/(ISSN)1942-2466.CESM2, 2018.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M. and Fung, I.: Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison, Journal of Climate, 19(14), 3337–3353, 2006.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Quéré, C. L., 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, 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., Werf, G. R. van der, Wiltshire, A. J. and Zaehle, S.: Global Carbon Budget 2019, Earth System Science Data, 11(4), 1783–1838, doi:https://doi.org/10.5194/essd-11-1783-2019, 2019.

~~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 Vöosmarty, C. J.: Nitrogen Cycles: Past, Present, and Future, Biogeochemistry, 70(2), 153–226, doi:10.1007/s10533-004-0370-0, 2004.~~

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), Geoscientific Model Development, 10(5), 2009–2030, doi:https://doi.org/10.5194/gmd-10-2009-2017, 2017.

Hazeleger, W., Wang, X., Severijns, C., Ştefănescu, S., Bintanja, R., Sterl, A., Wyser, K., Semmler, T., Yang, S., van den Hurk, B., van Noije, T., van der Linden, E. and van der Wiel, K.: EC-Earth V2.2: description and validation of a new seamless earth system prediction model, Clim Dyn, 39(11), 2611–2629, doi:10.1007/s00382-011-1228-5, 2012.

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

Hurk, B. van den, Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M. and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome, Geoscientific Model Development, 9(8), 2809–2832, doi:https://doi.org/10.5194/gmd-9-2809-2016, 2016.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., ~~Betts, R., Feddema~~Bodirsky, B. L., Calvin, K., Doelman, J., ~~Fischer, G.~~ C., Fisk, J., ~~Hibbard, K., Houghton, R., Janetos, A., Jones, C., Kindermann, G., Kinoshita, T., Klein~~ Fujimori, S., Goldewijk, K~~.~~. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J., Kennedy, J., Kristzin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., ~~Smith, S.,~~ Stehfest, E., ~~Thomson, A.,~~ Thornton, P., ~~van~~ Tubiello, F. N., Vuuren, D. P. van and ~~Wang, Y~~Zhang, X.: Harmonization of ~~land-use scenarios~~Global Land-Use Change and Management for the ~~period 1500~~ Period 850–2100~~: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Climatic Change, 109(~~ (LUH2) for CMIP6, Geoscientific Model Development Discussions, 1~~), 117–161~~ 65, doi:https://doi.org/10.~~1007/s10584-011-0153-2, 2011~~5194/gmd-2019-360, 2020.

Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C. H., Roelandt, C., Segschneider, J. and Tjiputra, J.: Twenty-First-Century Compatible CO2 Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways, J. Climate, 26(13), 4398–4413, doi:10.1175/JCLI-D-12-00554.1, 2013.

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., 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, Journal of Geophysical Research: Biogeosciences, 116(G3), doi:10.1029/2010JG001566, 2011.

~~Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Costa, A. C. L. da and Gentine, P.: Implementing Plant Hydraulics in the Community Land Model, Version 5, Journal of Advances in Modeling Earth Systems, 11(2), 485–513, doi:10.1029/2018MS001500, 2019.~~

Kottek, M., Grieser, J., Beck, C., Rudolf, B. and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, [online] Available from: https://www.ingentaconnect.com/content/schweiz/mz/2006/00000015/00000003/art00001 (Accessed 16 April 2020), 2006.

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

Lamarque, J.-F., Shindell, D. T., Josse, B., Young, P. J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A. and Zeng, G.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, Geoscientific Model Development, 6(1), 179–206, doi:https://doi.org/10.5194/gmd-6-179-2013, 2013.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L. van, Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke, M. van den, Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Martin, M. V. and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, Journal of Advances in Modeling Earth Systems, ~~n/a(n/a),~~11(12), 4245–4287, doi:10.1029/2018MS001583, ~~20202~~2019.

Le Quéré, C. L., 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, 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, B., Tubiello, F. N., Laan-Luijkx, I. T. van der, Werf, G. R. van der, Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S. and Zheng, B.: Global Carbon Budget 2018, Earth System Science Data, 10(4), 2141–2194, doi:https://doi.org/10.5194/essd-10-2141-2018, 2018.

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

Liang, J., Qi, X., Souza, L. and Luo, Y.: Processes regulating progressive nitrogen limitation under elevated carbon dioxide: a meta-analysis, Biogeosciences, 13(9), 2689–2699, doi:https://doi.org/10.5194/bg-13-2689-2016, 2016.

~~Luyssaert, S., Jammet, M., Stoy, P. C., Estel, S., Pongratz, J., Ceschia, E., Churkina, G., Don, A., Erb, K., Ferlicoq, M., Gielen, B., Grünwald, T., Houghton, R. A., Klumpp, K., Knohl, A., Kolb, T., Kuemmerle, T., Laurila, T., Lohila, A., Loustau, D., McGrath, M. J., Meyfroidt, P., Moors, E. J., Naudts, K., Novick, K., Otto, J., Pilegaard, K., Pio, C. A., Rambal, S., Rebmann, C., Ryder, J., Suyker, A. E., Varlagin, A., Wattenbach, M. and Dolman, A. J.: Land management and land-cover change have impacts of similar magnitude on surface temperature, Nature Clim. Change, 4(5), 389–393, doi:10.1038/nclimate2196, 2014.~~

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., Niu, S. L., Norby, R., Piao, S. L., Qi, X., Peylin, P., Prentice, I. C., Riley, W., Reichstein, M., Schwalm, C., Wang, Y. P., Xia, J. Y., Zaehle, S. and Zhou, X. H.: A framework for benchmarking land models, Biogeosciences, 9(10), 3857–3874, doi:https://doi.org/10.5194/bg-9-3857-2012, 2012.

Luyssaert, S., Inglima, I., Jung, M., Richardson, A. D., Reichstein, M., Papale, D., Piao, S. L., Schulze, E.-D., Wingate, L., Matteucci, G., Aragao, L., Aubinet, M., Beer, C., Bernhofer, C., Black, K. G., Bonal, D., Bonnefond, J.-M., Chambers, J., Ciais, P., Cook, B., Davis, K. J., Dolman, A. J., Gielen, B., Goulden, M., Grace, J., Granier, A., Grelle, A., Griffis, T., Grünwald, T., Guidolotti, G., Hanson, P. J., Harding, R., Hollinger, D. Y., Hutyra, L. R., Kolari, P., Kruijt, B., Kutsch, W., Lagergren, F., Laurila, T., Law, B. E., Maire, G. L., Lindroth, A., Loustau, D., Malhi, Y., Mateus, J., Migliavacca, M., Misson, L., Montagnani, L., Moncrieff, J., Moors, E., Munger, J. W., Nikinmaa, E., Ollinger, S. V., Pita, G., Rebmann, C., Roupsard, O., Saigusa, N., Sanz, M. J., Seufert, G., Sierra, C., Smith, M.-L., Tang, J., Valentini, R., Vesala, T. and Janssens, I. A.: CO2 balance of boreal, temperate, and tropical forests derived from a global database, Global Change Biology, 13(12), 2509–2537, doi:10.1111/j.1365-2486.2007.01439.x, 2007.

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., 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., Storch, J.-S. von, 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, Journal of Advances in Modeling Earth Systems, 11(4), 998–1038, doi:10.1029/2018MS001400, 2019.

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(9), 2390–2401, doi:10.1890/03-0351, 2004.

~~Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., Crous, K. Y., Angelis, P. D., Freeman, M. and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance, Global Change Biology, 17(6), 2134–2144, doi:10.1111/j.1365-2486.2010.02375.x, 2011.~~

Ménard, C. B., Ikonen, J., Rautiainen, K., Aurela, M., Arslan, A. N. and Pulliainen, J.: Effects of Meteorological and Ancillary Data, Temporal Averaging, and Evaluation Methods on Model Performance and Uncertainty in a Land Surface Model, J. Hydrometeor, 16(6), 2559–2576, doi:10.1175/JHM-D-15-0013.1, 2015.

Menge, D. N. L., Levin, S. A. and Hedin, L. O.: Facultative versus Obligate Nitrogen Fixation Strategies and Their Ecosystem Consequences., The American Naturalist, 174(4), 465–477, doi:10.1086/605377, 2009.

Meyerholt, J. and, Zaehle, S.: The role of stoichiometric flexibility in modelling forest ecosystem, and Smith, M. J.: Variability of projected terrestrial biosphere responses to elevated levels of atmospheric CO_2 due to uncertainty in biological nitrogen fertilization, New Phytologist, 208(4), 1042–1055, doi:10.1111/nph.13547, 2015fixation, Biogeosciences, 13(5), 1491–1518, doi:https://doi.org/10.5194/bg-13-1491-2016, 2016.

Meyerholt, J., Sickel, K. and Zaehle, S.: Ensemble projections elucidate effects of uncertainty in terrestrial nitrogen limitation on future carbon uptake, Global Change Biology, n/a(n/a), doi:10.1111/gcb.15114, 2020.

Millar, R. J., Fuglestvedt, J. S., Friedlingstein, P., Rogelj, J., Grubb, M. J., Matthews, H. D., Skeie, R. B., Forster, P. M., Frame, D. J. and Allen, M. R.: Emission budgets and pathways consistent with limiting warming to 1.5 °C, Nature Geoscience, 10(10), 741–747, doi:10.1038/ngeo3031, 2017.

Müller, C., Stehfest, E., Minnen, J. G. van, Strengers, B., Bloh, W. von, Beusen, A. H. W., Schaphoff, S., Kram, T. and Lucht, W.: Drivers and patterns of land biosphere carbon balance reversal, Environ. Res. Lett., 11(4), 044002, doi:10.1088/1748-9326/11/4/044002, 2016.

Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., Macias-Fauria, M., Sass-Klaassen, U., Lévesque, E., Boudreau, S., Ropars, P., Hermanutz, L., Trant, A., Collier, L. S., Weijers, S., Rozema, J., Rayback, S. A., Schmidt, N. M., Schaepman-Strub, G., Wipf, S., Rixen, C., Ménard, C. B., Venn, S., Goetz, S., Andreu-Hayles, L., Elmendorf, S., Ravolainen, V., Welker, J., Grogan, P., Epstein, H. E. and Hik, D. S.: Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities, Environ. Res. Lett., 6(4), 045509, doi:10.1088/1748-9326/6/4/045509, 2011.

Nauta, A. L., Heijmans, M. M. P. D., Blok, D., Limpens, J., Elberling, B., Gallagher, A., Li, B., Petrov, R. E., Maximov, T. C., van Huissteden, J. and Berendse, F.: Permafrost collapse after shrub removal shifts tundra ecosystem to a methane source, Nature Climate Change, 5(1), 67–70, doi:10.1038/nclimate2446, 2015.

New, M., Hulme, M. and Jones, P.: Representing Twentieth-Century Space–Time Climate Variability. Part II: Development of 1901–96 Monthly Grids of Terrestrial Surface Climate, J. Climate, 13(13), 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.

Norby, R. J., DeLucia, E. H., Gielen, B., Calfapietra, C., Giardina, C. P., King, J. S., Ledford, J., McCarthy, H. R., Moore, D. J. P., Ceulemans, R., Angelis, P. D., Finzi, A. C., Karnosky, D. F., Kubiske, M. E., Lukac, M., Pregitzer, K. S., Scarascia-Mugnozza, G. E., Schlesinger, W. H. and Oren, R.: Forest response to elevated CO2 is conserved across a broad range of productivity, PNAS, 102(50), 18052–18056, doi:10.1073/pnas.0509478102, 2005.

O'Connor, F. M., Boucher, O., Gedney, N., Jones, C. D., Folberth, G. A., Coppell, R., Friedlingstein, P., Collins, W. J., Chappellaz, J., Ridley, J. and Johnson, C. E.: Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review, Reviews of Geophysics, 48(4), doi:10.1029/2010RG000326, 2010.

Oleson, K. W., Lawrence, D. M., B, G., Flanner, M. G., Kluzek, E., J, P., Levis, S., Swenson, S. C., Thornton, E., Feddema, J., Heald, C. L., Lamarque, J., Niu, G., Qian, T., Running, S., Sakaguchi, K., Yang, L., Zeng, X. and Zeng, X.: Technical Description of version 4.0 of the Community Land Model (CLM)., 2010.

Olin, S., Lindeskog, M., Pugh, T. a. M., Schurgers, G., Wårlind, D., Mishurov, M., Zaehle, S., Stocker, B. D., Smith, B. and Arneth, A.: Soil carbon management in large-scale Earth system modelling: implications for crop yields and nitrogen leaching, Earth System Dynamics, 6(2), 745–768, doi:https://doi.org/10.5194/esd-6-745-2015, 2015.

Piao, S., Luyssaert, S., Ciais, P., Janssens, I. A., Chen, A., Cao, C., Fang, J., Friedlingstein, P., Luo, Y. and Wang, S.: Forest annual carbon cost: a global-scale analysis of autotrophic respiration, Ecology, 91(3), 652–661, doi:10.1890/08-2176.1, 2010.

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., Caemmerer, S. von, Dietze, M. C., Kattge, J., Leakey, A. D. B., Mercado, L. M., Niinemets, Ü., Prentice, I. C., Serbin, S. P., Sitch, S., Way, D. A. and Zaehle, S.: A roadmap for improving the representation of photosynthesis in Earth system models, New Phytologist, 213(1), 22–42, doi:10.1111/nph.14283, 2017.

Schulte-Uebbing, L. and Vries, W. de: Global-scale impacts of nitrogen deposition on tree carbon sequestration in tropical, temperate, and boreal forests: A meta-analysis, Global Change Biology, 24(2), e416–e431, doi:10.1111/gcb.13862, 2018.

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.: The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations, Geoscientific Model Development Discussions, 1–68, doi:https://doi.org/10.5194/gmd-2019-378, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., Mora, L. de, Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G. and Zerroukat, M.: UKESM1: Description and ~~evaluation~~Evaluation of the ~~UK~~U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, ~~n/a(n/a),~~11(12), 4513–4558, doi:10.1029/2019MS001739, ~~2020~~2019.

~~Shao, P., Zeng, X., Sakaguchi, K., Monson, R. K. and Zeng, X.: Terrestrial Carbon Cycle: Climate Relations in Eight CMIP5 Earth System Models, J. Climate, 26(22), 8744–8764, doi:10.1175/JCLI-D-12-00831.1, 2013.~~

Shi, M., Fisher, J. B., Brzostek, E. R. and Phillips, R. P.: Carbon cost of plant nitrogen acquisition: global carbon cycle impact from an improved plant nitrogen cycle in the Community Land Model, Global Change Biology, 22(3), 1299–1314, doi:10.1111/gcb.13131, 2016.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J. and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11(7), 2027–2054, doi:https://doi.org/10.5194/bg-11-2027-2014, 2014.

Sokolov, A. P., Kicklighter, D. W., Melillo, J. M., Felzer, B. S., Schlosser, C. A. and Cronin, T. W.: Consequences of Considering Carbon–Nitrogen Interactions on the Feedbacks between Climate and the Terrestrial Carbon Cycle, J. Climate, 21(15), 3776–3796, doi:10.1175/2008JCLI2038.1, 2008.

Song, J., Wan, S., Piao, S., Knapp, A. K., Classen, A. T., Vicca, S., Ciais, P., Hovenden, M. J., Leuzinger, S., Beier, C., Kardol, P., Xia, J., Liu, Q., Ru, J., Zhou, Z., Luo, Y., Guo, D., Adam Langley, J., Zscheischler, J., Dukes, J. S., Tang, J., Chen, J., Hofmockel, K. S., Kueppers, L. M., Rustad, L., Liu, L., Smith, M. D., Templer, P. H., Quinn Thomas, R., Norby, R. J., Phillips, R. P., Niu, S., Fatichi, S., Wang, Y., Shao, P., Han, H., Wang, D., Lei, L., Wang, J., Li, X., Zhang, Q., Li, X., Su, F., Liu, B., Yang, F., Ma, G., Li, G., Liu, Y., Liu, Y., Yang, Z., Zhang, K., Miao, Y., Hu, M., Yan, C., Zhang, A., Zhong, M., Hui, Y., Li, Y. and Zheng, M.: A meta-analysis of 1,119 manipulative experiments on terrestrial carbon-cycling responses to global change, Nature Ecology & Evolution, 3(9), 1309–1320, doi:10.1038/s41559-019-0958-3, 2019.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Thomas, R. Q., Zaehle, S., Templer, P. H. and Goodale, C. L.: Global patterns of nitrogen limitation: confronting two global biogeochemical models with observations, Global Change Biology, 19(10), 2986–2998, doi:10.1111/gcb.12281, 2013.

Thomas, R. Q., Brookshire, E. N. J. and Gerber, S.: Nitrogen limitation on land: how can it occur in Earth system models?, Glob Change Biol, n/a-n/a, doi:10.1111/gcb.12813, 2015.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A. and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variability, Global Biogeochemical Cycles, 21(4), doi:10.1029/2006GB002868, 2007.

Thornton, P. E., Doney, S. C., Lindsay, K., Moore, J. K., Mahowald, N., Randerson, J. T., Fung, I., Lamarque, J.-F., Feddema, J. J. and Lee, Y.-H.: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences, 6(10), 2099–2120, doi:https://doi.org/10.5194/bg-6-2099-2009, 2009.

Vitousek, P. M. and Howarth, R. W.: Nitrogen limitation on land and in the sea: How can it occur?, Biogeochemistry, 13(2), 87–115, doi:10.1007/BF00002772, 1991.

Walker, A. P., Zaehle, S., Medlyn, B. E., Kauwe, M. G. D., Asao, S., Hickler, T., Parton, W., Ricciuto, D. M., Wang, Y.-P., Wårlind, D. and Norby, R. J.: Predicting long-term carbon sequestration in response to CO2 enrichment: How and why do current ecosystem models differ?, Global Biogeochemical Cycles, 29(4), 476–495, doi:10.1002/2014GB004995, 2015.

Wang, Y., Ciais, P., Goll, D., Huang, Y., Luo, Y., Wang, Y.-P., Bloom, A. A., Broquet, G., Hartmann, J., Peng, S., Penuelas, J., Piao, S., Sardans, J., Stocker, B. D., Wang, R., Zaehle, S. and Zechmeister-Boltenstern, S.: GOLUM-CNP v1.0: a data-driven modeling of carbon, nitrogen and phosphorus cycles in major terrestrial biomes, Geoscientific Model Development, 11(9), 3903–3928, doi:https://doi.org/10.5194/gmd-11-3903-2018, 2018.

Wårlind, D., Smith, B., Hickler, T. and Arneth, A.: Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model, Biogeosciences, 11(21), 6131–6146, doi:https://doi.org/10.5194/bg-11-6131-2014, 2014.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O. and Best, M.: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, J. Hydrometeor., 12(5), 823–848, doi:10.1175/2011JHM1369.1, 2011.

Wieder, W. R., Cleveland, C. C., Lawrence, D. M. and Bonan, G. B.: Effects of model structural uncertainty on carbon cycle projections: biological nitrogen fixation as a case study, Environ. Res. Lett., 10(4), 044016, doi:10.1088/1748-9326/10/4/044016, 2015a.

Wieder, W. R., Cleveland, C. C., Smith, W. K. and Todd-Brown, K.: Future productivity and carbon storage limited by terrestrial nutrient availability, Nature Geoscience, 8(6), 441–444, doi:10.1038/ngeo2413, 2015b.

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., Grandy, A. S., Koven, C. D., Lombardozzi, D. L., Oleson, K. W. and Thomas, R. Q.: Beyond Static Benchmarking: Using Experimental Manipulations to Evaluate Land Model Assumptions, Global Biogeochemical Cycles, 33(10), 1289–1309, doi:10.1029/2018GB006141, 2019.

Yu, Z., Wang, J., Liu, S., Rentch, J. S., Sun, P. and Lu, C.: Global gross primary productivity and water use efficiency changes under drought stress, Environ. Res. Lett., 12(1), 014016, doi:10.1088/1748-9326/aa5258, 2017.

Zaehle, S. and Dalmonech, D.: Carbon–nitrogen interactions on land at global scales: current understanding in modelling climate biosphere feedbacks, Current Opinion in Environmental Sustainability, 3(5), 311–320, doi:10.1016/j.cosust.2011.08.008, 2011.

Zaehle, S., Friend, A. D., Friedlingstein, P., Dentener, F., Peylin, P. and Schulz, M.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 2. Role of the nitrogen cycle in the historical terrestrial carbon balance, Global Biogeochem. Cycles, 24(1), GB1006, doi:10.1029/2009GB003522, 2010.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., Luo, Y., Wang, Y.-P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warlind, D., Weng, E., Parton, W., Iversen, C. M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P. J., Prentice, I. C., Oren, R. and Norby, R. J.: Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate Free-Air CO2 Enrichment studies, New Phytol, 202(3), 803–822, doi:10.1111/nph.12697, 2014a.

Zaehle, S., Jones, C. D., Houlton, B., Lamarque, J.-F. and Robertson, E.: Nitrogen Availability Reduces CMIP5 Projections of Twenty-First-Century Land Carbon Uptake, J. Climate, 28(6), 2494–2511, doi:10.1175/JCLI-D-13-00776.1, 2014b.

Zak, D. R., Freedman, Z. B., Upchurch, R. A., Steffens, M. and Kögel-Knabner, I.: Anthropogenic N deposition increases soil organic matter accumulation without altering its biochemical composition, Global Change Biology, 23(2), 933–944, doi:10.1111/gcb.13480, 2017.

Zhang, Q., Pitman, A. J., Wang, Y. P., Dai, Y. J. and Lawrence, P. J.: The impact of nitrogen and phosphorous limitation on the estimated terrestrial carbon balance and warming of land use change over the last 156 yr, Earth System Dynamics, 4(2), 333–345, doi:https://doi.org/10.5194/esd-4-333-2013, 2013.

Zheng, M., Zhou, Z., Luo, Y., Zhao, P. and Mo, J.: Global pattern and controls of biological nitrogen fixation under nutrient enrichment: A meta-analysis, Global Change Biology, 25(9), 3018–3030, doi:10.1111/gcb.14705, 2019.

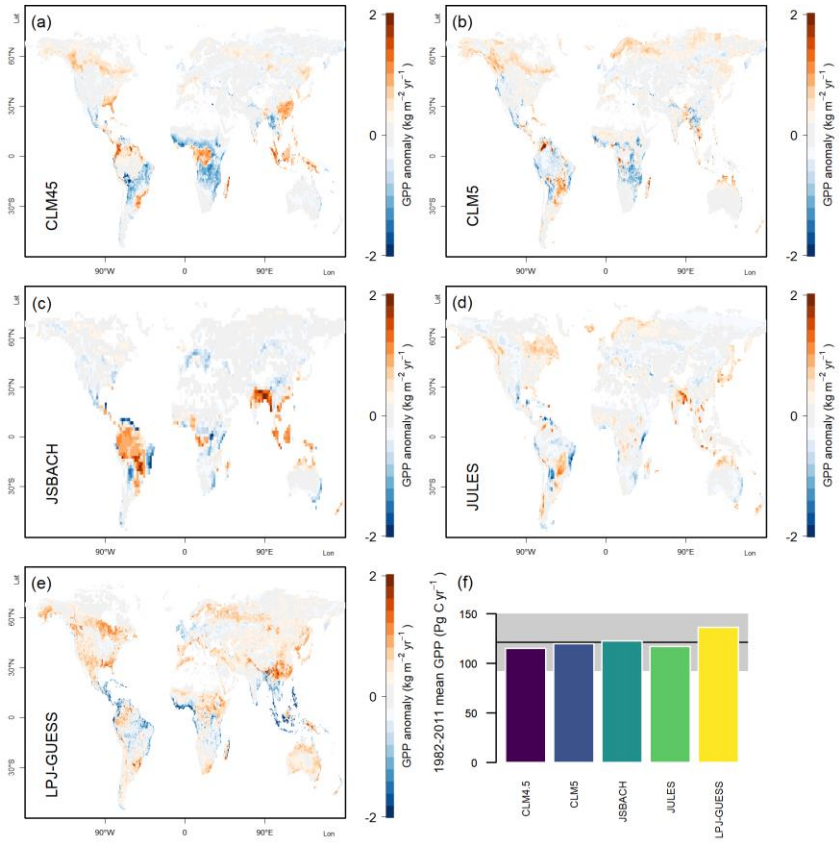
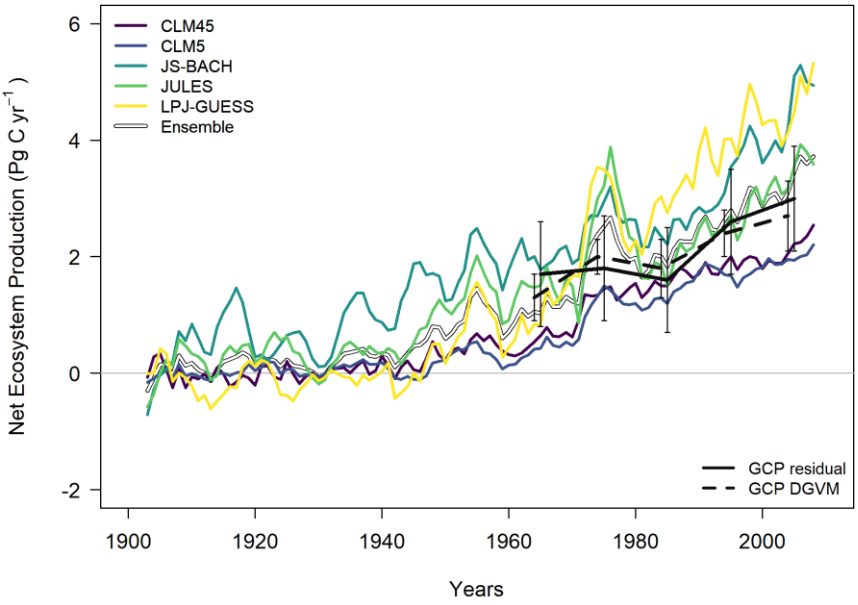


Figure 1. Model output of 1982-2011 mean gross primary productivity (GPP). (a) – (e) Model estimates, shown as anomaly of the corresponding observation-based estimate (MTE) published by Jung et al. (2011). (f) Globally integrated estimates. Black line indicates the global average from the observation-based source; grey area indicates the globally integrated standard deviation from the global average in the model tree ensemble applied to obtain the global average.


CLM45
CLM5
JS-BACH
JULES
LPJ-GUESS
Ensemble
GCP residual
GCP DGVM
Net Ecosystem Production (Pg C yr^{-1})
Years
6
4
2
0
-2
1900
1920
1940
1960
1980
2000

~~Figure 2.~~

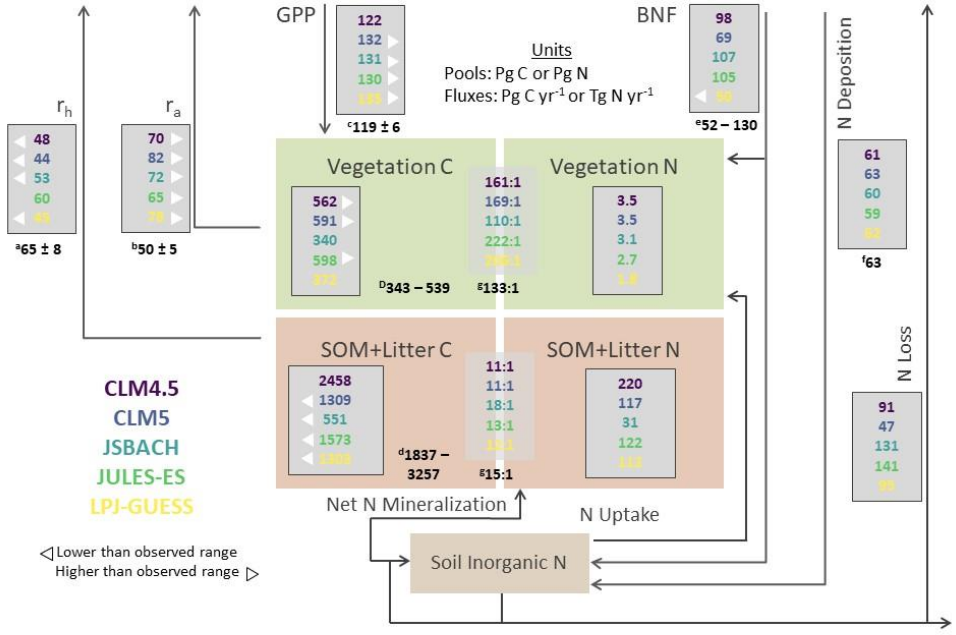


Figure 1. 1996-2005 mean model estimates of the major ecosystem C and N component pools and fluxes in comparison with observation-based estimates from the literature. C = Carbon; N = Nitrogen; rh = Heterotrophic respiration; ra = Autotrophic respiration; GPP = Gross primary productivity; SOM = Soil organic matter; BNF = Biological nitrogen fixation; The N uptake flux refers to root uptake of inorganic N. Ranges shown represent the 95% confidence intervals, standard deviation, or similar uncertainty metrics, where available. N loss is the loss via gaseous loss and leaching. The black numbers indicate observation-based estimates from the literature: a) Heterotrophic respiration: Bond-Lamberty and Thomson, (2010), soil respiration estimate for 2008. To account for the included root respiration, we reduced the literature estimate by 33% according to (Bowden et al., 1993); b) Autotrophic respiration: Piao et al., (2010), Luyssaert et al., (2007), present day estimate for forests from 2007; c) GPP: Jung et al., (2011), averaged estimate for 1982-2011; d) SOM+Litter, and Vegetation C: Carvalhais et al., (2014), present day estimate from 2014; e) BNF: (Davies-Barnard and Friedlingstein, 2020) upscaled averages for 1980-2019; f) N deposition: (Lamarque et al., 2013), estimate for 2000; g) C:N ratios: Wang et al., (2018).

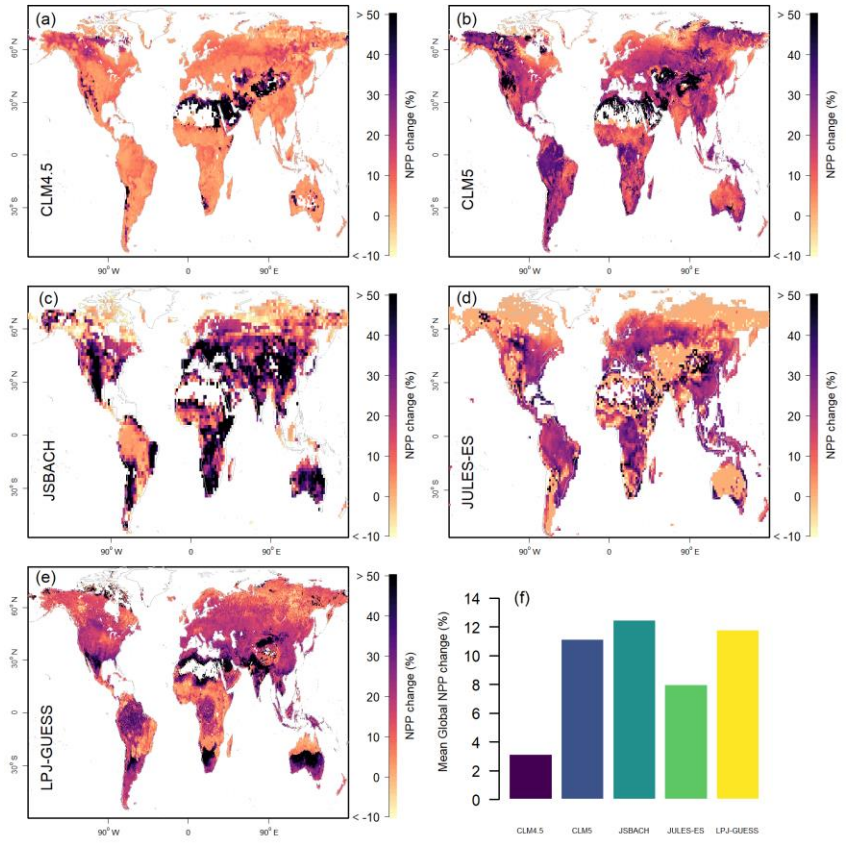
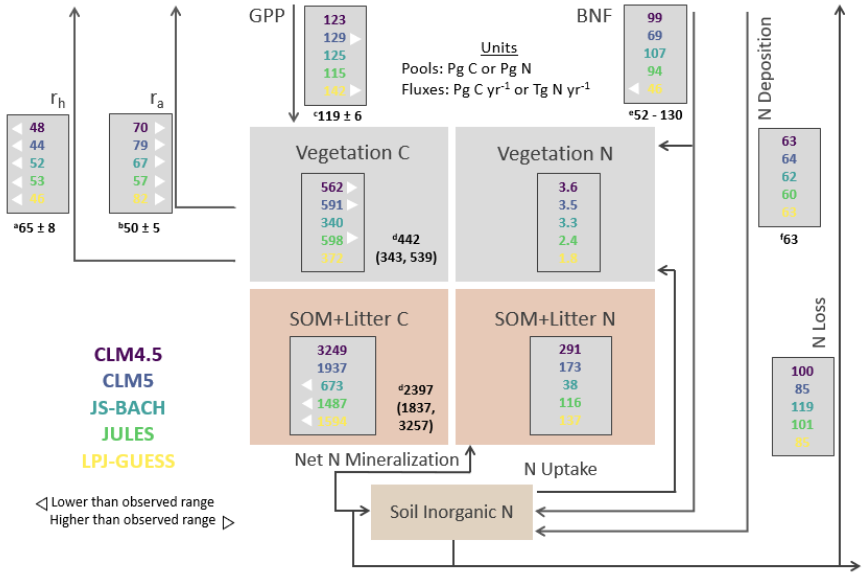
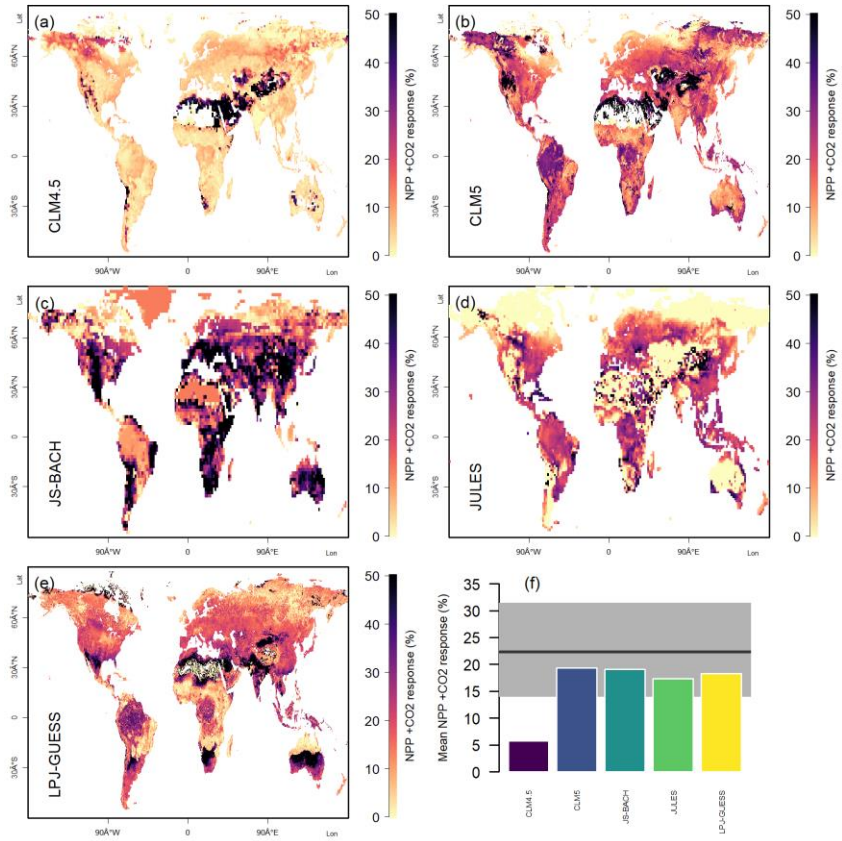


Figure 2. ~~Model predictions (five year moving averages) of 1901-2010 global net ecosystem production (NEP). Coloured curves show model predictions as anomalies to the 1901-1910 mean, "Ensemble" shows the model average. Black curves and error bars indicate assessments published by the Global Carbon Project (Friedlingstein et al., 2019). "GCP residual" shows the estimates and uncertainty resulting from the consideration of all constituting global fluxes as a "residual sink". "GCP-DGVM" shows the combined estimate ± 1 standard deviation from an ensemble of global vegetation models.~~

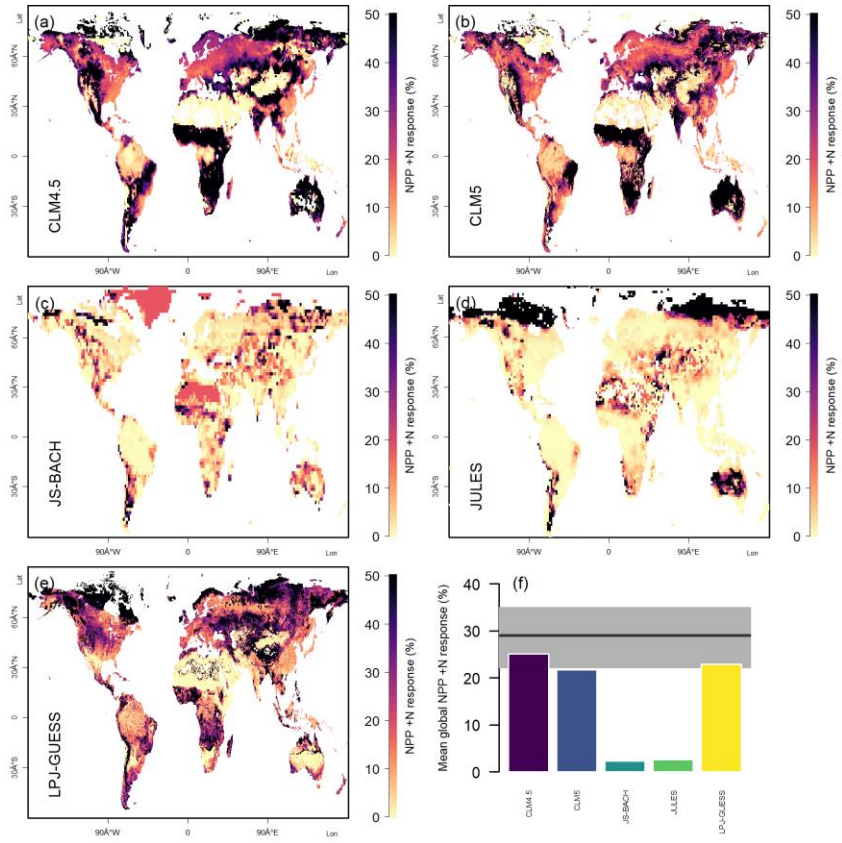


~~Figure 3. 1996-2005 mean model estimates of the major ecosystem C and N component pools and fluxes in comparison with observation based estimates from the literature. C = Carbon; N = Nitrogen; rh = Heterotrophic respiration; ra = Autotrophic respiration; GPP = Gross primary productivity; SOM = Soil organic matter; BNF = Biological nitrogen fixation; The N uptake flux refers to root uptake of inorganic N. N loss is the loss via gaseous loss and leaching. The black numbers indicate observation-based estimates from the literature: a) Heterotrophic respiration: Bond-Lamberty & Thomson (2010), soil respiration estimate for 2008. To account for the included root respiration, we reduced the literature estimate by 33% according to Bowden et al. (1993); b) Autotrophic respiration: Piao et al. (2010), Luyssaert et al. (2007), present day estimate for forests from 2007; c) GPP: Jung et al. (2011), averaged estimate for 1982-2011; d) SOM+Litter: Carvalhais et al. (2014), present day estimate from 2014; e) BNF: Davies-Barnard and Friedlingstein, (forthcoming) global total BNF; f) N deposition: Lamarque et al. (2010), estimate for 2000.~~



~~Figure 4.~~ **Model estimates of 1996-2005 mean net primary productivity (NPP) response to +CO2. (a) – (e) Model estimates, shown as the anomaly compared to the model control scenario. Values above 50% are given the 50% colour. (f) Globally integrated** ~~estimates. Black line indicates the global average observed NPP responses to +CO2 (Baig et al., 2015); grey area indicates the uncertainty range from the observations.~~ values.

Formatted: Font: 10 pt


(a)
CLM4.5
(b)
CLM5
(c)
JS-BACH
(d)
JULES
(e)
LPJ-GUESS
(f)
NPP +N response (%)
Mean global NPP +N response (%)
Lat
Lon
90Â°W
0
90Â°E
60Â°N
30Â°N
30Â°S
CLM 5
CLM5
JS-BACH
JULES
LPJ-GLESS

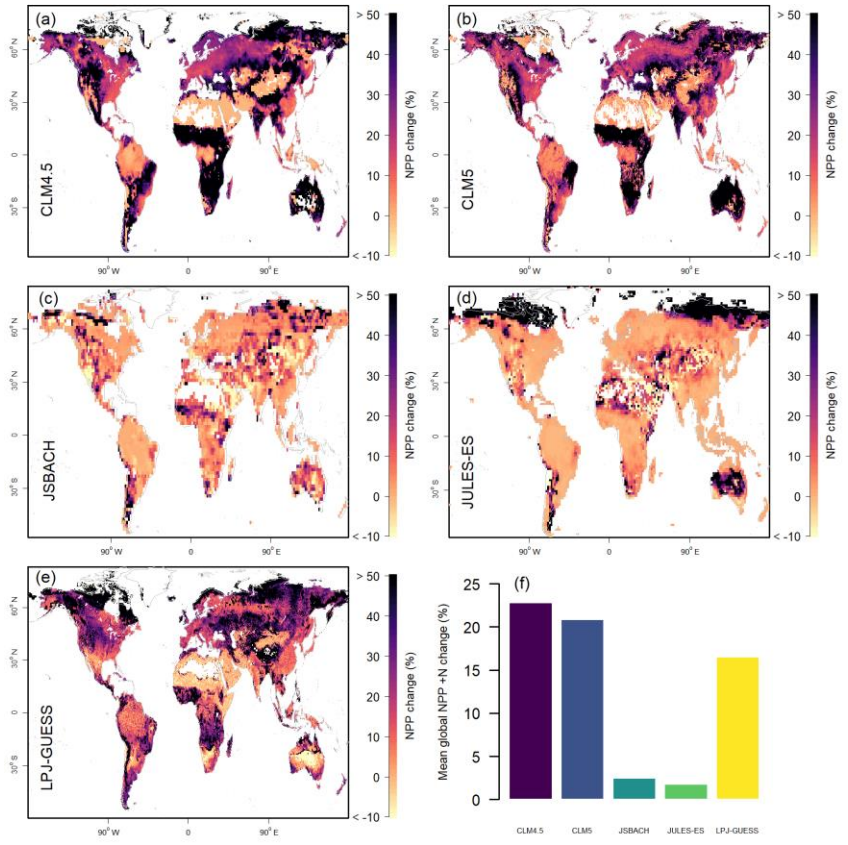
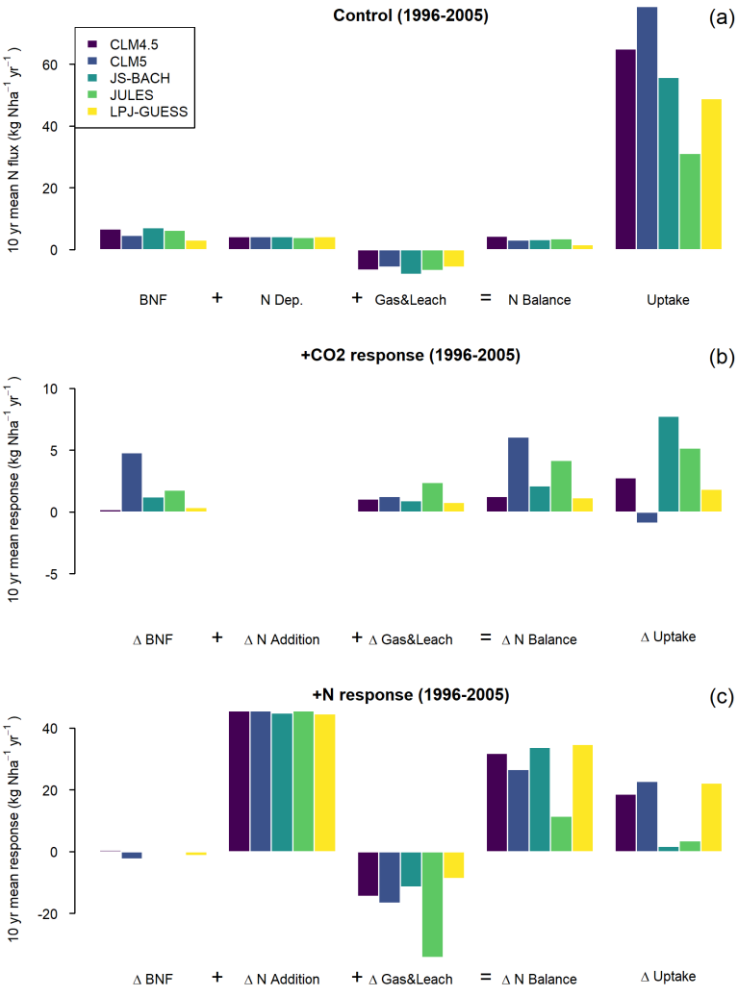


Figure ~~53~~. Model estimates of 1996-2005 mean net primary productivity (NPP) response to +N. (a) – (e) Model estimates, shown as the anomaly compared to the model control scenario. Values above 50% are given the 50% colour. (f) Globally integrated ~~estimates. Black line indicates the global average observed NPP responses to +N (LeBauer and Treseder, 2008); grey area indicates the uncertainty range from the observations.~~ values.


Control (1996-2005)
(a)
CLM4.5
CLM5
JS-BACH
JULES
LPJ-GUESS
10 yr mean N flux (kg Nha^{-1} yr^{-1})
60
40
20
0
BNF
+
N Dep.
+
Gas&Leach
=
N Balance
Uptake
+CO2 response (1996-2005)
(b)
10 yr mean response (kg Nha^{-1} yr^{-1})
10
5
0
-5
Δ BNF
+
Δ N Addition
+
Δ Gas&Leach
=
Δ N Balance
Δ Uptake
+N response (1996-2005)
(c)
10 yr mean response (kg Nha^{-1} yr^{-1})
40
20
0
-20
Δ BNF
+
Δ N Addition
+
Δ Gas&Leach
=
Δ N Balance
Δ Uptake

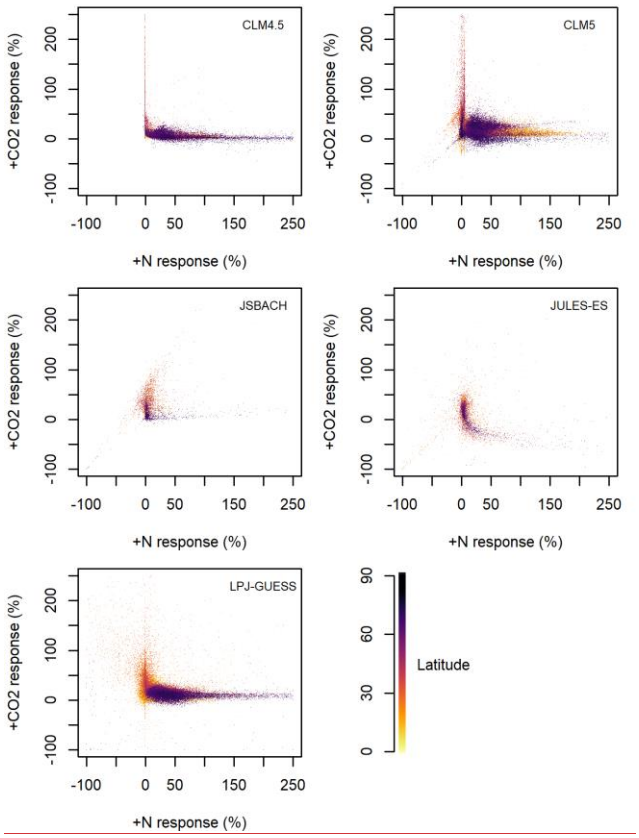


Figure 4. Model estimates of 1996-2005 mean net primary productivity (NPP) response to +N vs +CO2, as a percent anomaly of the control scenario. Each grid box is plotted against the corresponding grid box for the other variable. The percent change is capped at 250% and values above are not plotted. The colour of the points indicates the latitude.

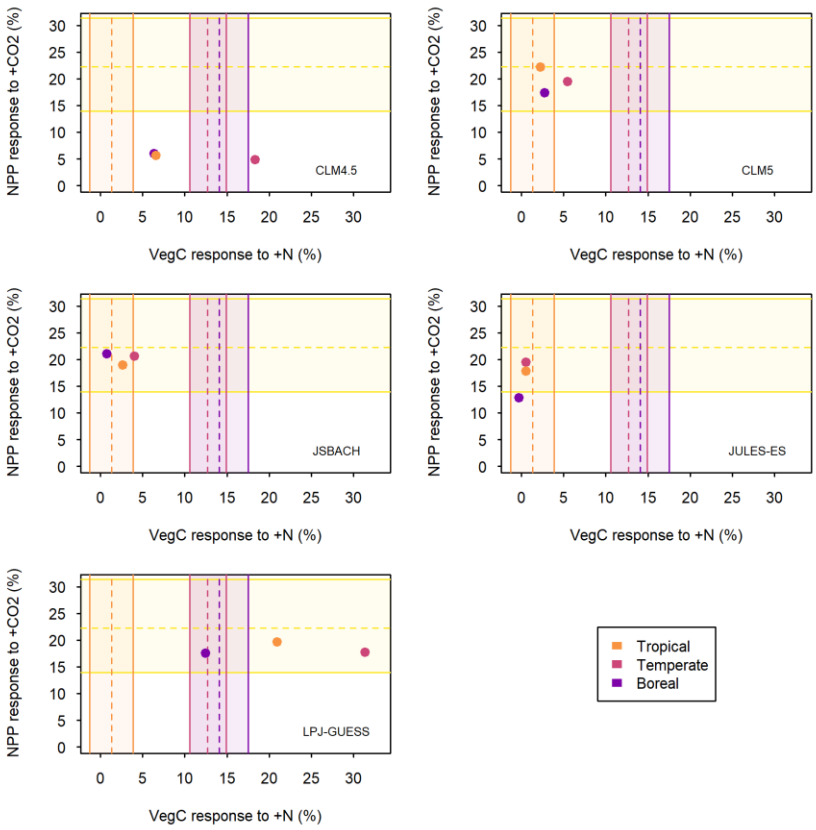


Figure 5. Average 1996-2005 model predictions of forest NPP responses to +CO2 (y-axis) and aboveground vegetation C pool size responses to nitrogen (N) addition (x-axis) for each of the models (as labelled). Area outlined in yellow indicates synthesis of observed forest NPP responses to +CO2 (Baig et al., 2015). Other coloured areas indicate biome-wise estimates of aboveground forest C change per added N (Schulte-Uebbing and Vries, 2018). For +CO2, NPP is restricted to simulated vegetation with NPP > 0.2 kg C m^{-2} yr^{-1} to exclude non-forest areas. For +N, forest VegC in CLM5, CLM4.5, and LPJ-GUESS is taken from wood C and N, whereas all C and N is included for JULES-ES and JSBACH due to model output limitations. The biomes are allocated according to Köppen-Geiger climate classification (Kottek et al., 2006).

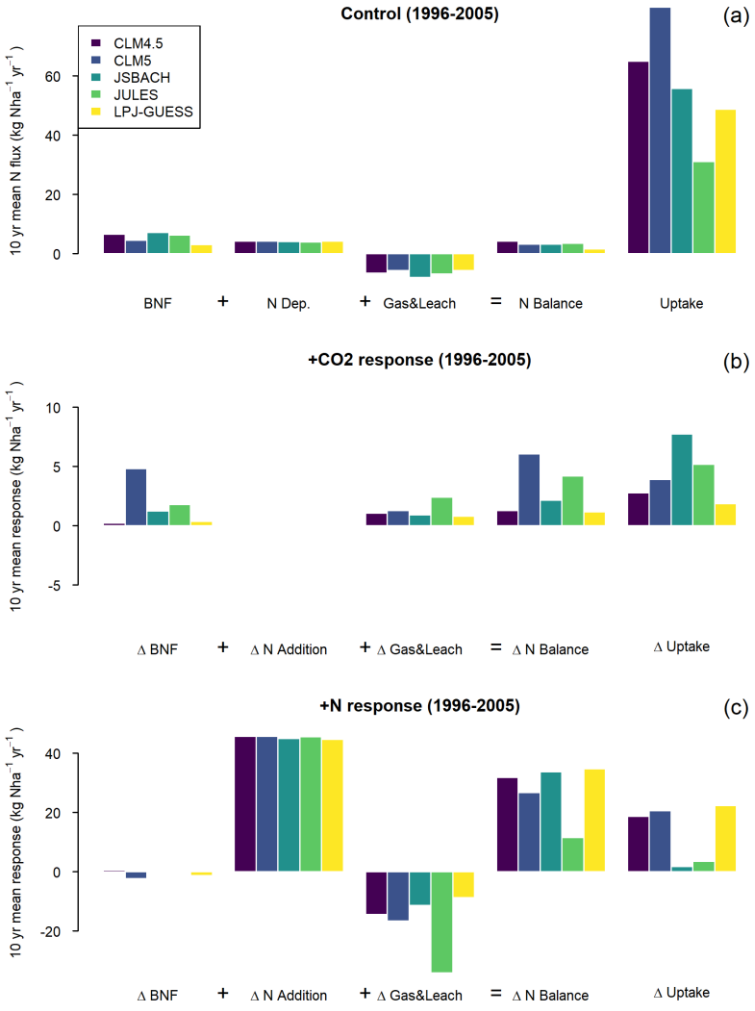


Figure 6. Global averaged 1996-2005 biological nitrogen fixation (BNF), N deposition, N loss via gaseous and leaching, the balance of those three inputs/losses, and the N uptake of the models. The top panel represents the Control scenario, and the second and third panels the response to +CO2 and +N perturbations (~~Methods~~see methods). Note that the y-axis scale is 4x smaller for +CO2 response than the Control or +N response.

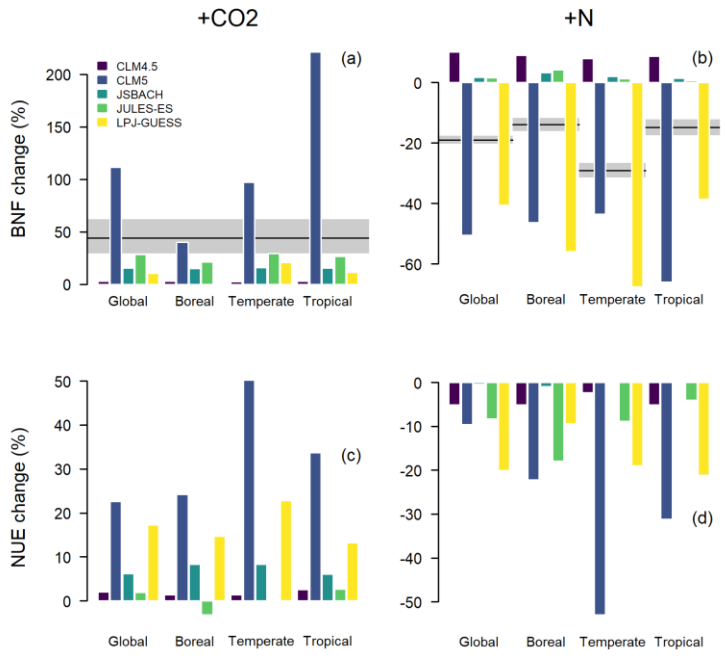
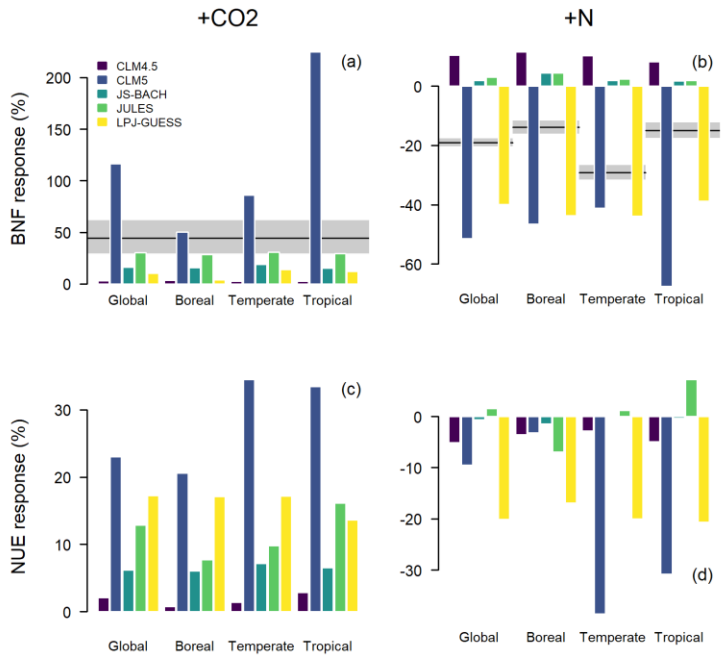
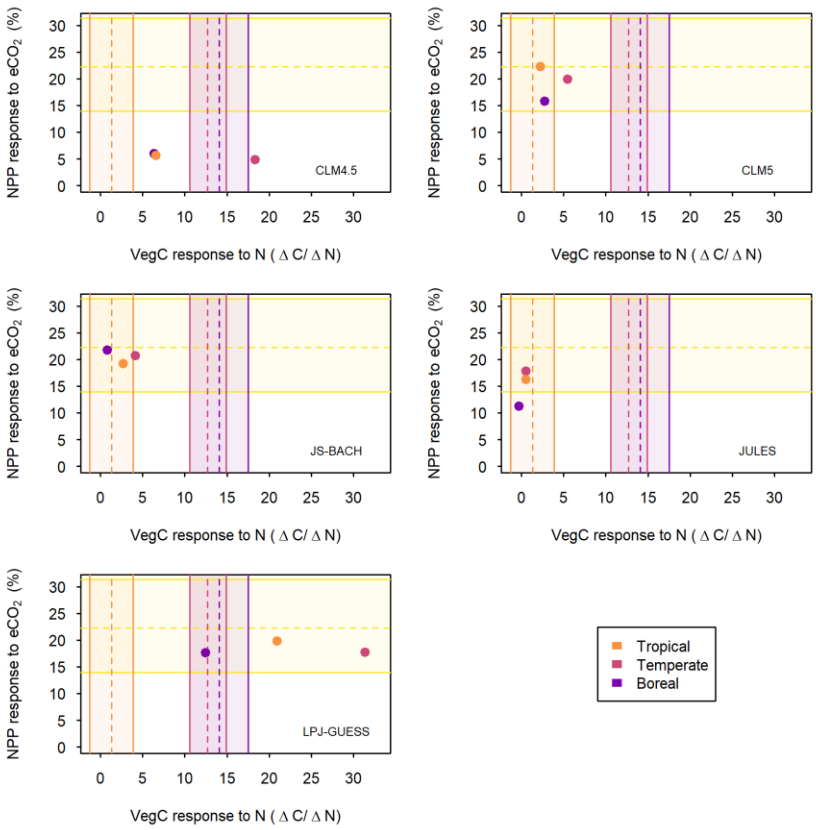


Figure 7. Averaged 1996-2005 responses in biological nitrogen fixation (BNF) and nitrogen-use efficiency (NUE; see Eq. 1) to +CO2 and +N perturbations for the global (all vegetation types) or forest region averages. (a) Model BNF responses to +CO2. Black line and grey area indicate mean and 95% CI of the global estimate published by Liang et al., (2016). (b) Model BNF responses to +N. Black lines and grey areas indicate means and 95% confidence intervals of the forest estimates published by Zheng et al., (2019). (c) Model NUE responses to +CO2. (d) Model NUE responses to +N. Forest biomes are according to Köppen-Geiger climate classification (Kottek et al., 2006), see SI Fig. 1.

Table 1. Key nitrogen cycle algorithms applied by the models. C = Carbon; N = Nitrogen; GPP = gross primary productivity; NPP = net primary productivity; PFT = plant functional type



~~Figure 7. Averaged 1996-2005 responses in biological nitrogen fixation (BNF) and nitrogen use efficiency (NUE; see Eq. 1) to +CO2 and +N perturbations (Methods). (a) Model BNF responses to +CO2. Black line and grey area indicate mean and 95% CI of the global estimate published by Liang et al. (2016). (b) Model BNF responses to +N. Black lines and grey areas indicate means and 95% confidence intervals of the estimates published by Zheng et al. (2019). (c) Model NUE responses to +CO2. (d) Model NUE responses to +N. Biomes according to Köppen-Geiger climate classification (Kottek et al., 2006).~~



~~Figure 8. Average 1996-2005 model predictions of net primary productivity (NPP) responses to +CO2 (y-axis) and aboveground vegetation carbon (C) pool size responses to nitrogen (N) addition (x-axis) for each of the models (as labelled). Area outlined in yellow indicates synthesis of observed NPP responses to +CO2 (Baig et al., 2015). Other coloured areas indicate biome-wise estimates of aboveground forest C change per added N (Schulte-Uebbing et al., 2018). Model results restricted to simulated vegetation with NPP > 0.2 kg C m^{-2} yr^{-1}. Biomes according to Köppen-Geiger climate classification (Kottek et al., 2006).~~

	CLM4.5	CLM5	JSBACH	JULES-ES	LPJ-GUESS
Key references	Oleson et al. (2013)	Lawrence et al. (2020)	Goll, et al. (2017), Mauritsen et al. (2019)	Wiltshire et al. (forthcoming)	Smith et al. (2014)
N effect on GPP	Downregulation of GPP to match stoichiometric constraint from allocable N	Leaf N compartmentalized into different pools to co-regulate photosynthesis according to the LUNA model	No direct effect	No direct effect	Reduction of rubisco capacity in case of N stress
N effect on autotrophic respiration	N content-dependent tissue-level maintenance respiration	Updated PFT-specific N-dependent leaf respiration scheme	No direct effect	N content-dependent maintenance respiration for roots and stems	N content-dependent maintenance respiration for roots and stems; leaf respiration reduced under N stress
Vegetation pool C:N stoichiometry	Fixed for all pools	Flexible for all pools	Fixed for all pools except labile	Flexible leaf stoichiometry from which root and stem C:N are scaled with fixed fractions	Flexible for leaves and fine roots, fixed otherwise
Retranslocation of N from shed leaves	Fraction of leaf N moved to mobile plant N pool prior to shedding. Fraction depends on PFT-specific fixed live leaf and leaf litter C:N ratios.	Fraction of leaf N moved to mobile plant N prior to shedding via two pathways: a free retranslocation, or a paid-for retranslocation dependent on PFT-specific dynamic leaf C:N range and minimum leaf litter C:N and available carbon to spend for	Fraction of leaf N moved to mobile plant N pool prior to shedding	Fraction of leaf N moved to labile store with PFT specific retranslocation coefficient	Fraction of leaf N moved to mobile plant N pool prior to shedding. Fraction depends on N stress.

Formatted Table

		extraction in FUN model			
Biological N fixation	Monotonically increasing function of NPP	Symbiotic N fixation according to the FUN model, asymbiotic N fixation linearly dependent on evapotranspiration	Non-linear function of NPP	Linear function of NPP, 0.0016 kg N per kg C NPP	Linear function of ecosystem evapotranspiration, 0.102 cm yr^{-1} ET +0.524 per kg N ha^{-1}
Ecosystem N loss	Denitrification loss as fraction of gross N mineralization + fraction of soil inorganic N pool in case of N saturation (CLM-CN) / Denitrification as fraction of nitrification (CENTURY) Leaching as function of soil inorganic N pool size Fractional fire loss as fraction of vegetation and litter pools	Denitrification as fraction of nitrification (CENTURY) Leaching as function of soil inorganic N pool size Fractional fire loss as fraction of vegetation and litter pools	Denitrification proportional to soil inorganic N pool and soil moisture Leaching proportional to soil inorganic N pool and drainage	Denitrification is a fixed fraction (1%) of mineralization flux Leaching of nitrogen is a function of soil inorganic N pool, drainage, and a parameter representing the effective solubility of nitrogen	Denitrification as fixed fraction of mineralization flux Leaching as function of soil inorganic N pool and drainage N loss from fire events
Plant N uptake	Function of plant N demand, soil inorganic N availability, and competition with heterotrophs	Soil uptake of inorganic N according to the FUN model	Plant N demand-based, limited by soil inorganic N availability	Demand based on GPP and limited by soil inorganic N availability	Determined to maintain optimal leaf N for photosynthesis, limited by soil inorganic N availability, fine root mass, soil temperature

					and plant N status

Table 2. Observational datasets used for comparison with model results

Variable/effect	Dataset	Reference	Number of measurements
+CO2 effect on NPP	meta-analysis of total above ground biomass of woody plants	Baig et al., (2015)	16
	meta-analysis for whole plant NPP and aboveground NPP (ANPP)	Song et al., (2019)	unspecified, maximum of 103
+N effect on NPP	meta-analysis on NPP changes	LeBauer and Treseder, (2008)	126, incl. tundra (10), tropics (8), arid land (3)
	meta-analysis for whole plant NPP and aboveground NPP (ANPP)	Song et al., (2019)	unspecified, maximum of 429
BNF responses to +CO2	global meta-analysis estimate	Liang et al. (2016).	89
BNF responses to +N	meta-analysis	Zheng et al., (2019),	tropical forest (92), temperate forest (52), boreal forest (37)
Biomass response to +N	aboveground forest biomass C change per added N from meta-analysis	Schulte-Uebbing and Vries, (2018)	tropical (17), temperate (41), boreal (12)
GPP (SI Fig. 2)	Flux tower data model tree ensemble	Jung et al., (2011)	unknown
Biome allocation (SI Fig. 1)	Köppen-Geiger climate classification	Kottek et al., 2006)	n/a

Table 3. Percent change in mean global NPP from perturbations. The observations come from meta-analyses which may not be directly comparable, but which provide a useful context.

	+CO2	+N

CLM4.5	3.2%	22.8%
CLM5	11.2%	20.9%
JSBACH	12.5%	2.5%
JULES-ES	8.0%	1.8%
LPJ-GUESS	11.8%	16.6%
Mean whole plant NPP values based on meta-analyses of field scale measurements	15.6% (2.8 – 28.4%) (Song et al., 2019)	6.5% (3 – 10.5%) (Song et al., 2019)
Mean productivity values based on meta-analyses of field scale measurements	26% (12.2 – 39.8%) (Song et al., 2019) (ANPP) 22.3% (13.9 – 31.4%) (Baig et al., 2015) (total woody plant biomass) 21.4% (11 – 32.8%) (Baig et al., 2015) (above-ground woody plant biomass)	20% (7.5 – 32.5%) (Song et al., 2019) (ANPP) 29% (22 -35%) (LeBauer and Treseder, 2008) (ANPP)

~~Table 1. Key nitrogen cycle algorithms applied by the models. C = Carbon; N = Nitrogen; GPP = gross primary productivity; NPP = net primary productivity; PFT = plant functional type~~