Simulation of the biogeochemical cycle of phosphorus in the ORCHIDEE land surface model : evaluation against local and global observational data

Yan Sun

Thèse de doctorat

NNT : 2021UPASJ001

université PARIS-SACLAY

Simulation du cycle biogéochimique du phosphore dans le modèle de surface terrestre ORCHIDEE-CNP: évaluation par rapport à des données d'observation locales et globales

Simulation of the biogeochemical cycle of phosphorus in the ORCHIDEE land surface model: evaluation against local and global observational data

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 129: Sciences de l'environnement d'Ile-de-France (SEIF)
Spécialité de doctorat: Météorologie, océanographie, physique de l'environnement
Unité de recherche : Paris-Saclay University, CNRS, CEA, UVSQ, Laboratory of Climate and Environmental Sciences, 91191, Gif-sur-Yvette, France
Référent : Université de Versailles -Saint-Quentin-en-Yvelines

Thèse présentée et soutenue à Paris-Saclay, le 23 mars 2021, par

Yan SUN

Composition du Jury

M. Philippe BOUSQUET Professeur des universités, Université de Versailles Saint-Quentin-en-Yvelines	Président
M. Christian BEER Professeur, Universität Hamburg	Rapporteur & Examinateur
Mme. Anja RAMING Professeur, Technical University of Munich	Rapporteur & Examinatrice
Mme. Ana BASTOS Chef de groupe, Max Planck Institute for Biogeochemistry	Examinatrice
Mme. Josette A. GARNIER Directeur de recherche, CNRS, Université de Paris-Saclay	Examinatrice
M. Soenke ZAEHLE Professeur, Max Planck Institute for Biogeochemistry	Examinateur

Direction de la thèse

M. Philippe CIAIS Directeur de recherche, LSCE	Directeur de thèse
M. Daniel S. GOLL Ingénieur de recherche, LSCE	Co-Directeur de thèse
M. Nicolas DELPIERRE Assistant Professeur, Laboratoire Écologie, Systématique, Évolution	Invité

ÉCOLE DOCTORALE
Sciences de l'Environnement d'Île-de-France (SEIF)

Titre : Simulation du cycle biogéochimique du phosphore dans le modèle de surface terrestre ORCHIDEE-CNP: évaluation par rapport à des données d'observation locales et globales

Mots clés : la séquestration du carbone, écosystèmes terrestres, limitations par la disponibilité en phosphore, apprentissage automatique, modèle global d'écosystèmes

Résumé : Le phosphore (P) joue un rôle essentiel dans le contrôle des processus métaboliques, de la dynamique de la matière organique du sol et de la productivité des écosystèmes, affectant ainsi le bilan des gaz à effet de serre (GES) des écosystèmes terrestres. Un nombre croissant de modèles numériques d'écosystèmes terrestres (LSMs) ont incorporé le cycle du phosphore mais leurs prévisions des bilans de GES restent incertaines. Les raisons sont: (1) le manque de données de référence pour les processus clés liés au P, (2) le manque d'approche intégrée globale d'évaluation adaptée aux processus spécifiques à P et aux interactions entre le cycle de P et celui du carbone (C) et de l'azote (N), et (3) le calibrage insuffisant des modèles, limité par le coût de calcul élevé pour simuler des cycles CNP couplés sur des échelles de temps allant de quelques minutes à plusieurs millénaires. Pour lever ces verrous de recherche, j'applique une combinaison de méthodes statistiques (apprentissage automatique), de LSMs et de données d'observation à différentes échelles.

Premièrement (chapitre 2), pour compléter les données de référence de l'évaluation des modèles. J'ai appliqué deux méthodes d'apprentissage automatique afin de produire des cartes spatiales de l'activité de la phosphatase acide (AP) à l'échelle continentale en extrapolant les observations sur sites de l'activité potentielle de AP. Le AP sécrété par les mycorhizes, les bactéries et les racines des plantes jouent un rôle important dans le recyclage du P du sol en transformant le P organique non disponible en phosphate assimilable. La méthode du réseau artificiel de neurones (BPN) a expliqué 58% de la variabilité spatiale de AP et peut reproduire les gradients en AP le long de trois transects représentatifs en Europe. Les éléments nutritifs du sol et les variables climatiques ont été détectés comme étant les principaux facteurs influençant les variations de la AP dans l'espace.

Deuxièmement (chapitre 3), j'ai évalué les performances de la version globale du LSM ORCHIDEE-CNP(v1.2) en utilisant les données du chapitre 2 ainsi que des données issues de la télédétection, des réseaux de mesure au sol et différentes bases de données. Les composantes simulées du cycle N et P à différents niveaux d'agrégation sont en bon accord avec les estimations empiriques. Nous avons toutefois identifié des biais du modèle, sur la stoechiométrie des feuilles et du sol et de l'efficacité d'utilisation des plantes P, qui suggèrent une sous-estimation de la disponibilité de P aux hautes latitudes, et donc des limitations peut être surestimées sur la productivité primaire. Sur la base de cette analyse, nous proposons des moyens de corriger les biais du modèle en donnant la priorité à une meilleure représentation des processus de minéralisation du P organique du sol et de la transformation du P inorganique du sol.

Enfin (chapitre 4) j'ai conçu et testé une procédure basée sur l'apprentissage automatique (ML) pour l'accélération de l'équilibration des cycles biogéochimiques dans le modèle ORCHIDEE-CNP (spin-up) en réponse à des conditions aux limites stationnaires, un problème qui est la source d'une faible efficacité de calcul des LSMs représentants les couplages entre P et autres éléments, dont le spin-up n'a pas de solution analytique simple. Cette approche d'accélération basée sur le ML requiert de ne faire tourner qu'un petit sous-ensemble de pixels du globe (14,1%) à partir desquels l'état d'équilibre des pixels restants est estimé par ML. La méthode prédit suffisamment bien l'état d'équilibre des stocks de C, N et P du sol, de la biomasse et de la litière C, N et P, comme l'indique l'erreur mineure introduite dans la simulation du bilan actuel du C terrestre. La consommation de temps de calcul du MLA est un ordre de grandeur inférieure à l'approche 'brute' actuellement utilisée, ce qui rend possible l'assimilation de données et l'optimisation de paramètres à l'aide de données d'observation en constante augmentation.

Dans les perspectives, je discute des applications spécifiques de l'approche MLA et des priorités de recherche futures pour améliorer encore la fiabilité et la robustesse des LSMs incluant le cycle de P et d'autres nutriments.

Maison du doctorat de l'Université Paris-Saclay
2ème étage aile ouest, Ecole normale supérieure Paris-Saclay
4 avenue des Sciences,
91190 Gif sur Yvette, France

Title : Simulation of the biogeochemical cycle of phosphorus in the ORCHIDEE land surface model: evaluation against local and global observational data

Keywords : Phosphorus limitation, terrestrial ecosystem, carbon storage, machine learning, land surface model

Abstract : Phosphorus (P) plays a critical role in controlling metabolic processes, soil organic matter dynamics, plant growth and ecosystem productivity, thereby affecting greenhouse gas balance (GHG) of land ecosystems. A small number of land surface models (LSMs) have incorporated P cycles but their predictions of GHG balances remain highly uncertain. The reasons are: (1) scarce benchmarking data for key P-related processes (e.g. continental to global scale gridded datasets), (2) lack of comprehensive global evaluation strategy tailored for d P processes and interlinkages with carbon (C) and nitrogen (N) cycles, and (3) insufficient model calibration limited by the high computation cost to simulate coupled CNP cycles which operate on timescales of minutes to millennia. Addressing those research gaps, I apply a combination of statistical methods (machine learning), LSMs and observational data among various scales.

Firstly (Chapter 2), to address the lack of benchmarking data, I applied two machine-learning methods with the aim to produce spatial gridded maps of acid phosphatase (AP) activity on continental scale by scaling up scattered site observations of potential AP activity. AP secreted by fungi, bacteria and plant roots play an important role in recycling of soil P via transforming unavailable organic P into assimilable phosphate. The back-propagation artificial network (BPN) method that was chosen explained 58% of AP variability and was able to identify the gradients in AP along three transects in Europe. Soil nutrients (total N, total P and labile organic P) and climatic controls (annual precipitation, mean annual temperature and temperature amplitude) were detected to be the dominant factors influencing AP variations in space.

Secondly (Chapter 3), I evaluated the performance of the global version of the LSM ORCHIDEE-CNP (v1.2) using the data from chapter 2 as well as additional data from remote-sensing, ground-based measurement networks and ecological databases. Simulated components of the N and P cycle at different levels of aggregation (from local to global) are in good agreement with data-driven estimates. We identified model biases, in the simulated large-scale patterns of leaf and soil stoichiometry and plant P use efficiency, which point towards an underestimation of P availability towards the poles. Based on our analysis, we propose ways to address the model biases by giving priority to better representing processes of soil organic P mineralization and soil inorganic P transformation.

Lastly (Chapter 4), I designed and tested a Machine Learning (ML)-based procedure for acceleration of the equilibration of biogeochemical cycles to boundary conditions which is causing the low computational efficiency of current P-enabled LSMs. This ML-based acceleration approach (MLA) requires to spin-up only a small subset of model pixels (14.1%) from which the equilibrium state of the remaining pixels is estimated by ML. MLA predicts the equilibrium state of soil, biomass and litter C, N and P on both PFT and global scale sufficiently well as indicated by the minor error introduced in simulating current land C balance. The computational consumption of MLA is about one order of magnitude less than the currently used approach, which opens the opportunity of data assimilation using the ever-growing observation datasets.

In the outlook, specific applications of the MLA approach and future research priorities are discussed to further improve the reliability and robustness of phosphorus-enabled LSMs.

Maison du doctorat de l'Université Paris-Saclay
2ème étage aile ouest, Ecole normale supérieure Paris-Saclay
4 avenue des Sciences,
91190 Gif sur Yvette, France

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Philippe Ciais, for the opportunities he provided me to do this study, and for the continuous support for my Ph.D study and research. During the past three years, he guided my research and provided many chances for me to in-depth study and communicate with other researchers. I benefited a lot from his immense knowledge and creative mind. His guidance helped me in the research and writing of this thesis.

I would specially thank Daniel S. Goll, my Ph.D. co-supervisors, for his direct guidance on me in scientific questions, technique issues, and English writing, and for organizing weekly meetings to discuss the science problems I encountered. Every time talking with him I learn something new. He is very encouraging and supportive. I've always got quick responses when I turned to him for advice and help. I'm also inspired by his passion and rigorous attitude towards science and his sharp scientific mind.

I would be very grateful to Yuanyuan Huang, Jinfeng Chang and Yilong Wang, who has given me many insightful ideas and comments on this work. I would also like to express my appreciation to my Ph.D. monitoring committee members, Ying-ping Wang, Cory Cleveland and Martin De Kauwe, who ensured the smooth running of my Ph.D. research by giving very helpful advice. I would thank all the co-authors and reviewers of the chapters for their help in improving the manuscript. I would also thank all the jury members for their agreement to join in and their time to review the work presented in this thesis.

I would also thank my Chinese friends, Hui Yang, Chunjing Qiu, Jinghui Lian, Yuan Zhang, Yitong Yao, Haicheng Zhang, Yang Xu, Yilin Cai, Bo Zheng, Kailiang Yu, Yuanhong Zhao, Yuanyuan Huang, Jinfeng Chang and Yilong Wang for all of the help in my life and study. The enjoyable time with them made my life in a foreign country easier and more colorful.

I am thankful to my master supervisors Prof. Shilong Piao and Prof. Shushi Peng for their recommendations, and for their care about my life in France.

Last but not the least, I want to thank my parents and grandparents. They always encourage me and give me unconditional love.

Table of Contents

Chapter 1 General Introduction

The atmospheric CO_2 concentration has been steadily increasing since the industrial revolution, largely contributed by the increasing anthropogenic CO_2 emissions. Doubling of the CO_2 concentration (from 270ppm to 355ppm) have exerted an effect on the under-going climate changes, including global warming, melting glaciers and ice sheets, increasing sea level, and more frequent extreme events (IPCC, 2019), which have significantly threatened natural and human systems. Terrestrial ecosystem serves as a CO_2 sink that offset about one third of the total anthropogenic CO_2 emissions (Friedlingstein et al., 2019). Although this land carbon (C) sink was projected to be sustained in the future (i.e. 2100) by most of the earth system models (i.e. 2100), large uncertainties still remain (Friedlingstein et al., 2014).

One of the uncertainty sources is our limited understanding of the interactions between nutrients and the land C cycle. Nitrogen (N) and phosphorus (P), two key macro elements required by plants, regulate the photosynthesis capacity and the mineralization of organic matter, and should down-regulate the ecosystem C fluxes when their availability is insufficient. During past decades, great efforts have been put on investigating the role of N in affecting C cycles, while few focused on the effect of phosphorus.

P is an essential nutrient for all living organisms (White & Hammond, 2008). It serves as a structural component in DNA and RNA, as a metabolic energy unit in ATP for energy transfer (Hawkesford et al., 2012), and enters in other structural components such as phospholipids in cell membranes. Thus, P plays a critical role in controlling metabolic processes, soil organic matter dynamics, and ecosystem productivity (Elser et al., 2007; Bradford et al., 2008; Cleveland et al., 2013; Hou et al., 2019). P deficiency as indicated by reduced photosynthesis (Kattge et al., 2009) and low productivity (Elser et al., 2007; Vitousek et al., 2010; Hou et al., 2019) was inferred to be occurring in natural tropical and subtropical ecosystems (e.g. Gleason et al., 2009; Ellsworth et al., 2017; Wright et al., 2011), natural temperate ecosystems (Hou et al., 2019) and agricultural ecosystems for low fertilization levels (e.g. MacDonald et al., 2011). On the other hand, oversupply of P from fertilizers is common in many agricultural systems, leading to environmental problems (Sharpley et al., 1994). The consequences of P limitation or excess on greenhouse gas balance on ecosystem and global level remain unclear, in particular for future conditions (Sun et al., 2017).

The response of photosynthesis and biomass production on ecosystem scale to elevated atmospheric CO_2 concentration is also likely constrained in many ecosystems by soil P availability (e.g. Terrer et al., 2019). It is thus critical to account for P limitation, and its interactions with global change drivers (e.g. nutrient deposition, climate change, elevated CO_2 concentration) in estimating future land carbon (C) fluxes and stocks (Wieder et al., 2015; Sun et al., 2017). A small number of land surface C models (LSMs) with P cycles and P interactions with C and N elements have been developed and applied (e.g. Wang et al., 2010; Yang et al., 2014; Goll et al., 2017). They found that the inclusion of the phosphorus cycle affects simulated C fluxes. Results from those LSMs describing C, N and P interactions showed that P availability limits primary productivity and carbon stocks, in particular on highly weathered soils of the tropics (Wang et al., 2010; Yang et al., 2014), and one study also suggested that P limitations could occur in the northern hemisphere (Goll et al., 2012). However, there exist large uncertainties for those LSMs in assessing P limitations due to uncertain model representations of P interactions (Fleischer et al., 2019), lack of P specific biogeochemistry data on relevant scales for model evaluation, and incomplete integration of available data due to the low computational efficiency and the complexity of global models.

1.1 Phosphorus constraints land C cycling

1.1.1 Phosphorus transfers between soils and plants

The abundance of P in the earth's crust is 0.1% (Rumble, 2020), of which 95% exist in apatite (Jahnke, 1992). Weathering of rocks and minerals slowly releases P as soluble form into the soil solution, and it is the primary P source for terrestrial natural ecosystems. Atmospheric P deposition from dust minerals containing P and from fossil fuel combustion anthropogenic aerosols over the last century is another source, which is more important for marine ecosystems or for land regions with very low weathering inputs, i.e. highly weathered soils (Okin et al., 2011; Yu et al., 2015). Soluble P in the soil is either taken up and utilized by plants and microbes to support their growth and activity, or is absorbed on mineral surface and can become occluded in secondary minerals (Walker & Syers, 1976). Soil P is continuously transported to water (rivers, lakes and ocean) via leaching, erosion and river run-off and is ultimately buried in deep ocean sediments. On very long geological time scales, marine sediments can be re-exposed to the surface by tectonic uplift and undergo weathering again (Buendia et al., 2010), closing the P-cycle (Filippelli, 2008) (Figure 1.1a). The cycling of P between the lithosphere, hydrosphere and biosphere occurs on a time scale going from few hours for plant uptake of P in the soil solution, up to millions of years for ocean transfers and rock weathering (Walker & Syers 1976; Vitousek, 2010).

In the terrestrial ecosystems, plant roots take up soil inorganic P (PO_4^{3+}) and mainly transport it to active tissues (e.g. fine root and leaf) to support the P demand of metabolic processes (Figure 1.1b). When plant tissue senesce, a small fraction of P in plants goes into litter, while most of it is re-translocated and stored to be invested again into alive tissues (i.e. resorption). Organic P in litters and soils is transformed back into assimilable phosphate, catalyzed by enzymes excreted by plant roots such as phosphatases, fungi and bacteria (i.e. mineralization). Thus, P is recycled within terrestrial ecosystems and resorption and accelerated mineralization are two key P availability enhancement mechanisms that plants use to reduce their dependence on very limited external new P supply i.e. P deposition from atmospheric aerosols and bedrock weathering release (Cleveland et al., 2013; Figure 1.1b). Most of the phosphorus demand of biota is thus met by recycling P rather than by using new inputs from weathering and deposition (Wang et al., 2010; Goll et al., 2012; Cleveland et al., 2013). Thus, the land P cycle is considered as a tight cycle, in contrast to the carbon and nitrogen cycles which have large fluxes connecting land reservoirs with the atmosphere.

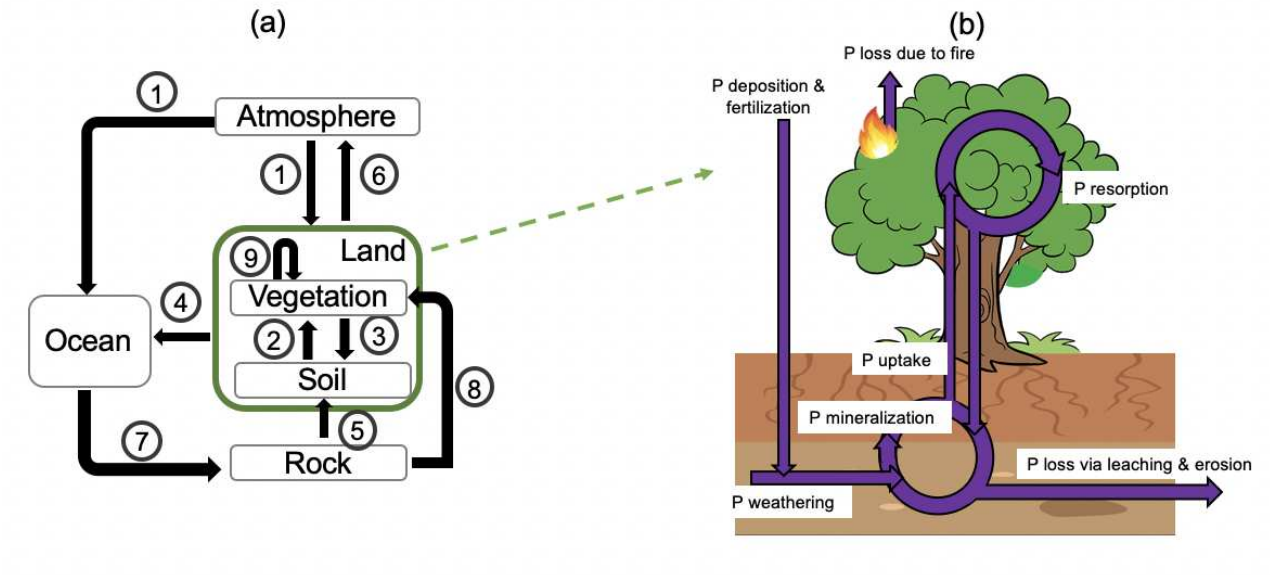


Figure 1.1 (a) P cycles among the lithosphere, hydrosphere and biosphere. Atmospheric P deposition (①) and rock P weathering (⑤) provide available P for ocean and land biomes. For land ecosystems, plants take up available P from soil (②) and return parts of it to soil organic matter pools

with dead tissue via litter fall and decomposition (③). The rest is recycled internally by plants (⑨). A small fraction of P in biomass goes into the atmosphere via burning (⑥) and gets redeposited as fly-ash (①). Soil dissolved P goes into river and ocean through run-off, erosion of particulate organic matter and leaching (④). Over long geological scales, marine sediments get re-exposed to the surface by tectonic uplift (⑦). Phosphorus mined from rock goes into the land ecosystem via the inputs of chemical fertilizer (⑧). (b) illustration of P turnover in a terrestrial ecosystem with the main loops of uptake, allocation and resorption in plant tissues and mineralization of soil organic matter.

1.1.2 Phosphorus availability and land ecosystem productivity

The mobility of P in soil is low compared to other major nutrients, like nitrogen, due to the high reactivity of phosphate ions with numerous soil constituents (e.g. $Ca^{2+,}$ Al^{3+} and Fe^{3+}), thus soil P is considered to be a major or even the prime limiting factor for plant growth (Hinsinger, 2001). Land ecosystems growing on old highly weathered soils tend to have lower P availability than ones on younger soils due to more P-depleted parental material and more strong P occlusion effect for ageing soils (Walker & Syer, 1976; Vitousek et al., 2010). According to theory based on limited data from soil development chronosequences, rock-derived nutrients (P) progressively get locked into long-lived plant biomass, soil organic matter and soil secondary minerals as ecosystems develop. As a consequence, the availability of rock-derived nutrients (P) is high on young soils and declines as soils age (Walker & Syer, 1976). Thus, tropical forests and savannas growing on old soils are regarded to be mainly restricted by P, while temperate and boreal ecosystem with youngers soils are considered to be mainly limited by N which originates primarily from the fixation of atmospheric N_2 (Walker & Syer, 1976; Vitousek et al., 2004; Vitousek et al., 2010; Du et al., 2020). However, this prevalent view has been challenged by recent meta-analysis study which reported a significant P limitation also in temperate regions (Hou et al., 2019) and modelling results which indicated a more attenuated role of soil age (Buendía et al., 2010).

Low P availability reduces plant photosynthesis and biomass productivity (Figure 2). Leaf-level observations show a significant positive correlation between maximum electron transport rate (J_{max}) and leaf P concentration (Walker et al., 2014; Ellsworth et al., 2015; Norby et al., 2017). Moreover, increasing leaf P concentration substantially increased the sensitivity of maximal Rubisco carboxylation rate (V_{cmax}) to leaf N concentration (Walker et al., 2014). Net primary productivity (NPP) in the Amazon region is positively correlated with leaf P concentration (Šímová et al., 2019). Although a pervasive phosphorus limitation for some individual tree species was measured at one intensive tropical forest measurement site, it did not translate to an emerging limitation on ecosystem level (Terrer et al., 2019). A large number of studies based on field experiments show that P fertilization significantly enhance aboveground plant production across tropical, subtropical, temperate (Elser et al., 2007; Li et al, 2016; Hou et al., 2019) and arctic regions (Hou et al., 2019).

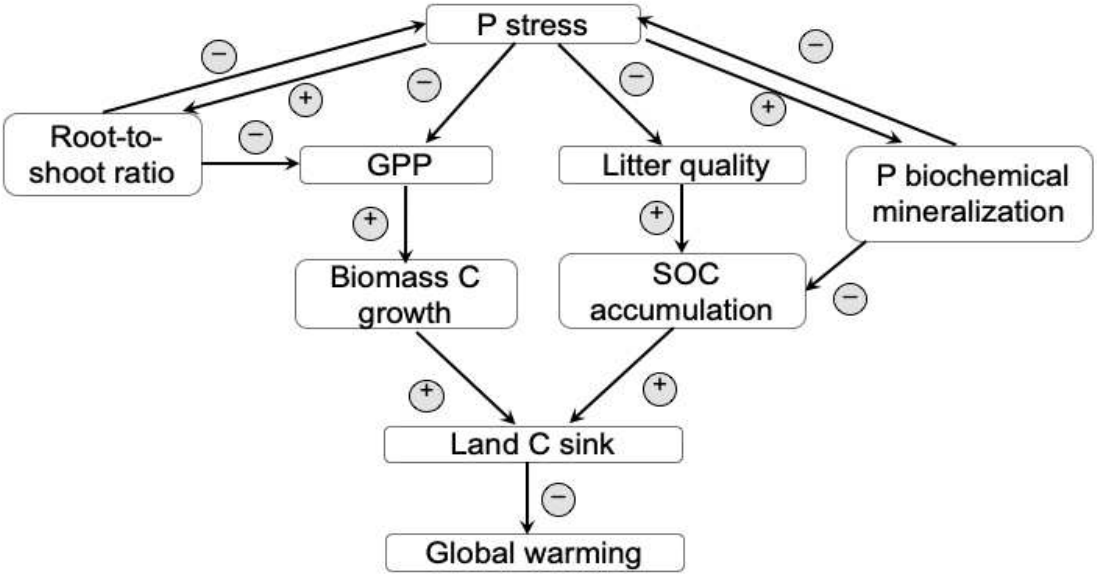


Figure 1.2 Feedbacks between C and P cycling affecting climate change. The positive and negative feedbacks are represented by '+' and '-' respectively. Low P availability (indicated as P-stress) reduces plant photosynthesis (GPP) and biomass C growth. Low P availability also reduces the litter quality (higher C:P ratios) and the amount of P released via litter decomposition (Knops, et al., 2010). Relative higher soil respiration (i.e. reducing SOC accumulation) was found in case of lower soil P availability in subtropical forests due to the lower microbial biomass (Fang et al., 2019). Besides, a series of strategies are taken by plants to adapt to P stress and increase P availability. Plants under P limitation have a reduced GPP. Still they can increase their C investment into roots at the expense of investments into other tissues (i.e. increasing root-to-shoot ratio; Hermans et al., 2006), which dampens that negative feedback on GPP by alleviating P limitation. Biochemical mineralization of P is one of the key pathways of P acquisition from soil organic P via phosphatase enzyme production (Wang et al., 2010). This process, driven indirectly by ecosystem P demand, is enhanced under P stress (Vance et al., 2003). As a consequence of higher P release from SOM, the accumulation of SOC is suppressed. Overall, P limitation is supposed to greatly affect the land C cycle and ultimately exert a non-ignorable feedback effect on global warming.

1.1.3 Phosphorus cycling and carbon interactions during the human-induced perturbation of the carbon cycle

Significant global human-made changes have been occurring over the last centuries: increasing atmospheric CO_2 concentration, increasing surface air temperature, shifts in precipitation patterns (Stocker et al., 2013), increasing chemical fertilizer inputs for agriculture and altered atmospheric N and P deposition (Peñuelas et al., 2013). Global changes have dramatically affected the biogeochemical cycles of C, N and P elements of Earth's ecosystems (Vitousek et al., 1997; Gruber & Galloway, 2008; Yuan & Chen, 2015; Jonard et al., 2015), and therefore their functioning, structure and ecosystem services (Peñuelas et al., 2013; Sardans et al., 2016; Schmitz et al., 2019; Peñuelas et al., 2020).

Global changes influence P availability by affecting weathering-P release rates, P mineralization and plant P recycling. The chemical weathering of P increases with global warming due to higher soil temperature (Goll et al., 2014; Hartmann et al., 2014). The biogeochemical mineralization of P increases with warming and increasing water availability (Sardans et al., 2006) due to the enhanced microbial activity under approaching optimal hydro-thermal conditions. The change of soil P availability under global changes is still uncertain. On the one hand, more P is locked in accumulating biomass, which reduces soil P availability. One the other hand, the increasing plant belowground investment in labile C due to the 'CO_2 fertilization', which improves soil P availability (Hasegawa et

al., 2016). P availability within plants, as indicated by foliar P resorption is observed to be reduced by warming and/or rainfall reduction (Prieto et al., 2020), but increased with increasing soil N content (See et al., 2015) and increasing nutrient deposition (Jonard et al., 2015).

Global changes influence the P demand of ecosystems (Sun et al., 2017), which causes response shifts in plant investment into different tissues, and their respective N and P concentration / stoichiometry. Meta-analysis shows stoichiometric changes in different directions and to varying extents due to exposure to elevated atmospheric CO_2 concentration, warming, changes in precipitation and increasing N and P fertilization (Yuan & Chen, 2015; Xu et al., 2020). For example, increasing atmospheric CO_2 concentration tends to decrease leaf N and P concentrations through a 'dilution effect' (Luo et al., 2004; Yuan & Chen, 2015; Peñuelas et al., 2020), whereas the response of leaf N:P ratio to elevated CO_2 is still uncertain (Deng et al., 2015; Yuan & Chen, 2015; Peñuelas et al., 2020; Prieto et al., 2020).

Changes in P cycling influences the land carbon balance. Elevated atmospheric CO_2 concentration is shown to stimulate plant growth and biomass productivity, but not always. The CO_2 'fertilization effect' has been observed to be attenuated ~15% in one low-P ecosystem compared to high-P ecosystems (Jiang et al., 2020a). The only Free-Air Carbon Enrichment on P poor soil (eucalypt) found no increase in aboveground biomass under elevated CO_2 levels (Ellsworth et al., 2017; Jiang et al., 2020b). A fertilization experiment nearby the EucFACE site showed that additional P input can increase ~50% of the stem basal area under ambient CO_2 (Crous et al., 2015). The extent to which this finding is representative for other regions, as well as the extent to which plants can upregulate P recycling to benefit from the elevated CO_2 remains largely unknown (Fleischer et al., 2019).

1.2 P limitation of future land C uptake

The terrestrial biosphere served as a net carbon (C) sink over the past few decades, which removed about one third of the atmosphere CO_2 emitted by anthropogenic activities (Friedlingstein et al., 2020). Estimation of the future land C uptake capacity on global scale under different GHG emission scenarios is of great significance for guiding human social and economic development. Most of the Earth System Models (ESMs) project increases in terrestrial carbon storage during the 21st century (Jones et al., 2013). However, part of the predicted future land C sink may not be realized due to the constraints by N and P, which are associated with large uncertainties or largely omitted in case of P in ESMs.

There are two types of approaches to predict the future land C uptake constrained by P limitation: (a) diagnosing it from the C-only models (i.e. with book-keeping methods; see below and Sun et al., 2017) and (b) simulating the CNP-interaction process-based models.

The unrealized increase in land C storage due to N and P shortage can be estimated from 'unlimited' C fluxes (i.e. net primary productivity; Wieder et al., 2015) or C storages (C storage-based; Hungate et al., 2003; Peñuelas et al., 2013) simulated by C-only ESMs using observations on the stoichiometry of plant and soil C pools and several assumptions (in the following 'book-keeping method'). Two previous studies used 'stoichiometric' book-keeping approaches to diagnose the P constraints on land C sink and came to different conclusions: C storage-based approach estimated the amount of P supporting increases of C storage lies within the estimates of available soil P (Peñuelas et al., 2013), while a C flux-based approach concluded that P limitation could reduce land C storage and turn land into a C source by end of the century (Wieder et al., 2015). The contradiction between results by the two different book-keeping methods has been reconciled by my work that was started before this Ph.D. and finished during its first year (Sun et al., (2017). I demonstrated that the estimation of land C storage increase using book-keeping methods is highly dependent on the assumed (and poorly constrained by observations) soil P availability on decadal to centennial timescale (Brovkin and Goll, 2015; Sun et al., 2017), rather than on the choice of a book-keeping method (e.g. flux vs. storage-based). Book-keeping approaches are primarily subject to three sources

of uncertainty related to their assumptions on ecosystem P recycling, soil P availability and ecosystem stoichiometry. Those uncertainties greatly influence the estimates of future land P deficit and realizable increase in C storage (Sun et al., 2017). The main uncertainties for book-keeping approaches are listed below.

(1) Non-readily available P in litter and SOM is recycled via decomposition and made available for plant uptake (Schachtman et al., 1998). The rate at which ecosystems can recycle P from litter and soil organic matter is poorly quantified by observations (Gill and Finzi, 2016). It depends on various biotic and abiotic factors (e.g. climate, litter quality, physicochemical soil properties; Davidson and Janssens, 2006; Doetterl et al., 2015) and varies in time and space. Under the assumption of medium soil labile P amount (resin and bicarbonate inorganic P) to define P availability, the land C sink that can be realized by 2100 considering P limitation is 2-12 Pg C lower for a low P turnover scenario than for high P turnover scenario (Sun et al., 2017).

(2) Soil "labile P", defined as the sum of resin-extractable P, bicarbonate-extractable inorganic P and organic P (Hedley et al., 1982; Cross and Schlesinger, 1995), was considered to be the primary P source for plant uptake on daily to decadal timescale. However, this labile P was recently found to have a high turnover rate (Helfenstein et al., 2018) which varies with environmental conditions or biological conditions (see 1.1.3; Buendía et al., 2014; Hasegawa et al., 2016). Thus, the use of those measurable quantities as a proxy for plant P availability is debatable even on an annual time scale (Johnson et al., 2003).

(3) Changes in the stoichiometry of vegetation biomass and SOM are observed and attributed to global changes drivers (see 1.1.3). Adjustments of C:P ratios for different plant tissue, or growth allocation among tissue of contrasting stoichiometries (e.g. wood vs leaves) can theoretically alleviate or circumvent P limitation by adjusting the P demand. How flexible stoichiometry on ecosystem level actually is remains a key uncertainty for the role of P in controlling the land C cycle.

The limitations of this diagnostic approach of estimating future C limitations led me to turn to process modeling for the rest of my PhD dissertation.

1.3 Current state-of-the-art of land models involving the phosphorus cycle

Considering the large uncertainty and shortcomings of book-keeping approaches, the inclusion of P cycling into land surface models (LSMs) is needed (Sun et al., 2017), which would significantly improve our capacity to test hypotheses and forecast interactions between biogeochemical cycles and a changing climate (Reed et al., 2015).

1.3.1 Simulated P constraints on land carbon cycling in global land models

Several LSMs incorporated different parameterizations of C-N interactions (e.g. Thornton et al., 2007; Zaehle et al., 2014) but very few global models have included C-N-P interactions. The few models which do include C-N-P interactions predict that P availability limits primary productivity and carbon stocks on highly weathered soils of the tropics (Wang et al., 2010; Goll et al., 2012; Yang et al., 2014) and one study also suggested that P limitations could also occur in the northern hemisphere in the near future (Goll et al., 2012).

The inclusion of the phosphorus cycle in LSMs may improve model performances with regard to reproducing observed C fluxes and storages (Wang et al., 2010; Goll et al., 2012; Yang et al., 2014; Thum et al., 2019), especially for the C cycles of current state in tropical regions (Goll et al., 2018; Yang et al., 2019). For example, the inclusion of P and water interactions in ORCHIDEE simulate a more realistic temporal variation in GPP for Amazon forests (Goll et al., 2018). And the

implementation of phosphorus (P) cycle and P limitation in the version 1 of the E3SM land model (ELM v1) improves simulated spatial pattern of wood productivity in Amazon forests (Yang et al., 2019). However adding a new cycle with many poorly known parameters is also expected to degrade the fit to conventional C data, compared to a C-only model that has been extensively calibrated. In other words, a new cycle ensures a more realistic representation of key mechanisms in the long run, but in the short run, it may also degrade the performance of a LSM compared to a previously tuned less complex version.

Moreover, P limitation could become increasingly important for tropical regions as CO_2 increases in the future, since the uptake of extra C into plant tissues inevitably requires either more P, or a more efficient use of P per mole of C fixed. Fleischer et al. (2019) conducted a model comparison in Amazon forest, and found that terrestrial biosphere models that explicitly consider C–N-P interactions show a reduced CO_2 fertilization effect on plant biomass C compared to CN-only models (~46%) and C-only models (~51 %) (Fleischer et al., 2019). However, the estimates of the size of this P effect on elevated CO_2 for Amazon forests vary among different models as a result of variability in the representation of the processes that determine P controls over C storage.

1.3.2 Uncertainties and limitations

Adding new but uncertain P-related processes in LSMs does not grant an automatic improvement of model predictions. First, more (nutrient-related) equations with more uncertain parameters should result in less robust predictions. Second, models ignoring nutrients were often calibrated on available carbon data, so that a new model with nutrients inevitably needs a parameter recalibration to reach the similar performances than the same model without nutrients. Third, for evaluating a large-scale model resolving both nutrient and carbon biogeochemistry, one should look for specific nutrient related datasets which are more scarce than classical biomass, productivity, soil carbon data used for benchmarking carbon only models.

Currently, model representations of P interactions are highly uncertain since the critical processes are poorly constrained by current observational data. In particular, the desorption of P from soil minerals surface and the enhancement of P availability for plants by phosphatase enzymes secreted by plant roots and microbes were identified to be critical but poorly constrained (Fleischer et al., 2019). Those large uncertainties in CNP models call for 1) a comprehensive model evaluation for the N- and P-related processes and 2) data-assimilation using the growing data observations.

The evaluation for N and P together with C cycling in global LSM models remains very limited and incomplete (Wang et al., 2010; Goll et al., 2012) but recent ground-based measurements and ecological datasets offer the opportunity to make progress. With recent meta-analysis of site-level nutrient fertilization experiments (e.g. Yuan and Chen, 2015; Wright, 2019), data-driven assimilation schemes to constrain nutrient budgets (Wang et al., 2018), new knowledge about the critical P-processes of sorption (Helfenstein et al., 2018; 2020), global datasets of leaf nutrient content (Butler et al., 2017), and empirical constraints on the CO_2 fertilization effect on land carbon storage (Terrer et al., 2019; Liu et al., 2019), a better evaluation of C, N, P models is feasible. In addition to direct comparison with nutrient datasets, it is also possible to diagnose emerging model responses in terms of ecosystem resource use efficiencies (RUE) and confront them to observations for identifying how ecosystems adjust and optimize nutrient, water, light, and carbon resource availabilities (Fernández-Martínez et al., 2014; Hodapp et al., 2019). In particular, modeled N and P use efficiencies can be compared to observation-based estimates at ecosystem scale (Gill and Finzi, 2016) and at biome scale (Wang et al., 2018).

Soil organic matter stores a large amount of phosphorus, and the degree to which its recycling can be maintained or enhanced under future conditions is critical for predicting future P availability and C sequestration (Yang et al., 2014). Soil phosphatases secreted by fungi, bacteria and plant roots play an important role in recycling of P in organic matter via transforming complex and unavailable forms

of organic P into assimilable phosphate (Caldwell, 2005). Previous studies measured the magnitude of potential phosphatase activity in the laboratory and investigated its responses to environment changes (e.g. Sardans et al., 2006; Marklein and Houlton, 2012). However, these local studies do not give information on the phosphatase activity's distribution on large spatial scales and do not provide hints of understanding the drivers of this distribution. This limited understanding for phosphatase activity on large spatial scales further hampers the global efforts of including P cycles in the land surface models (Reed et al., 2015; Jiang et al., 2019), and thus induce large uncertainty in investigating the future changes in land P recycling and the corresponding C storage under future global changes.

The application of model-data integration systems and uncertainty assessments relies on the realization of large numbers of simulations. However, this process is hampered by the high computation cost of spin-up simulations for reaching the initial equilibrium states (i.e. spin-up). Previous acceleration approaches for spin-up based on linear differential equations and fixed C-to-N ratio of pools or fluxes (e.g. Thornton and Rosenbloom, 2005; Xia et al., 2012; Shi et al., 2013; Koven et al., 2013; Fang et al., 2014) are not applicable to LSMs which include flexible stoichiometry and non-linear soil nutrient processes (Wang et al., 2012, Goll et al., 2017) such as inorganic soil P transformation. New acceleration approaches of spin-up for LSMs with C-N-P interactions are necessary to facilitate data assimilation, parameter calibration, and uncertainties reduction.

1.4 The aim and research questions of this thesis

A small number of LSMs have incorporated P cycles and C-N-P interactions to investigate the role of P in influencing land C cycle. However, modelling the P cycle in current LSMs has limitations due to (1) scarce benchmarking data for key P-related processes (e.g. continental to global scale gridded datasets (2) scarce long term site datasets and process studies with all important variables measured, (3) lack of comprehensive global evaluation strategies including both N and P processes, and (3) insufficient model calibration limited by the high computation cost for the equilibration of the coupled C, N and P cycles (i.e. spin-up). The aim of this thesis is to provide measures to improve the simulation of coupled cycles of C, N and P by addressing those bottlenecks using a combination of statistical methods (machine learning), observations and LSMs.

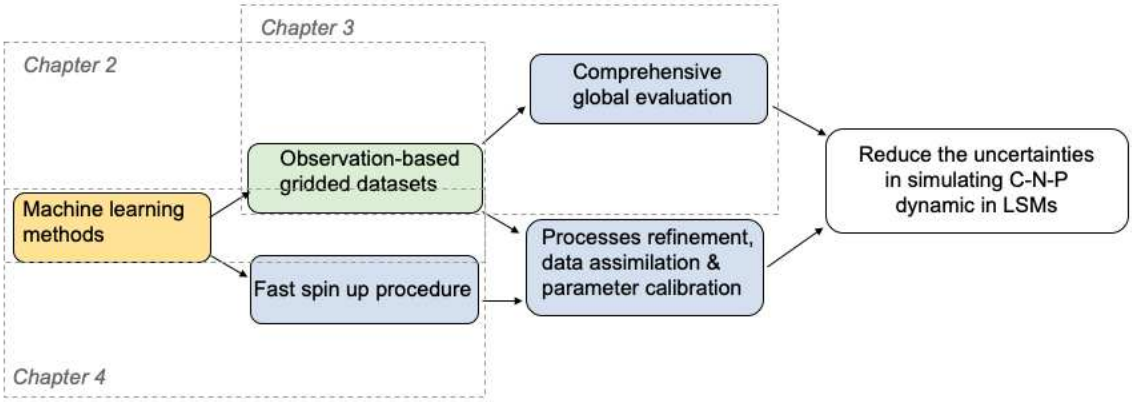


Figure 1.3 The work in thesis on reducing the uncertainties in simulated C, N and P dynamics.

The main research questions and objectives of my thesis are:

(1) Phosphatase is a key 'accelerator' of ecosystem P turnover. However, its spatial variations are not yet explored. The questions addressed in Chapter 2 are: Is it possible to produce a spatially explicit dataset of soil phosphatase activity from current scarce field data ? What are the major drivers of

dataset of soil phosphatase activity from current scarce field data ? What are the major drivers of phosphatase spatial variations?

(2) Although suitable observation-based datasets have been accumulating, current LSM predictions of N and P fluxes and stocks are not systematically evaluated against them - leaving uncertainties in simulating the effect of P (and N) on C cycles. The questions addressed in Chapter 3 are: How well does the LSM ORCHIDEE-CNP reproduce observation-based nutrient-related properties of the land biosphere? What are the processes which should be preferentially improved in future model development?

(3) The calibration of parameters in complex CNP models is hindered by the high computation cost for the equilibration of the coupled C, N and P cycles to boundary conditions (i.e. the model spin-up). The question addressed in Chapter 4 is: Is it possible to develop an acceleration approach for the spin-up?

Each of these objectives are addressed by work described specific chapters:

In Chapter 2, I apply machine learning methods to provide a quantification of the spatial variation in potential acid phosphatase activity on large spatial scales, across the European continent, and its drivers, which can help to reduce the uncertainty in our understanding of bio-availability of soil P and serve as a benchmarking data for current LSMs involving P cycle.

In Chapter 3, I evaluate the performance of the global nutrient enabled version land surface model ORCHIDEE-CNP v1.2 (r5986) using data from remote-sensing, ground-based measurement networks and ecological databases as well as the maps of acid phosphatase activity from 'Chapter 2'. This study will help propose ways to address the model biases.

In Chapter 4, I present the design and results from a Machine Learning (ML)-based procedure for the acceleration of spin-up of a complex LSM that has an imbrication of short and very long time scales. The approach is applied to the derivation of the pre-industrial spin-up state of the ORCHIDEE-CNP v1.2 model. The increase in computational efficiency for spin-up opens the opportunity of data assimilation and sensitivity assessments of this model, and is applicable to other similar models.

Finally, at the end of the document, Chapter 5 is a section summarizing the main finding of this Ph.D., and I propose some future research pathways.

References

Bradford, M. A., Fierer, N., & Reynolds, J. F. (2008). Soil carbon stocks in experimental mesocosms are dependent on the rate of labile carbon, nitrogen and phosphorus inputs to soils. Functional Ecology, 22(6), 964-974.

Brovkin, V., & Goll, D. (2015). Land unlikely to become large carbon source. Nature Geoscience, 8(12), 893-893.

Buendía, C., Arens, S., Hickler, T., Higgins, S. I., Porada, P., & Kleidon, A. (2014). On the potential vegetation feedbacks that enhance phosphorus availability–insights from a process-based model linking geological and ecological timescales. Biogeosciences, 11, 3661-3683.

Buendia, C., Kleidon, A., & Porporato, A. (2010). The role of tectonic uplift, climate and vegetation in the long-term terrestrial phosphorous cycle. *Biogeosciences*, *7*(6), 2025-2038.

Buendía, C., Kleidon, A., & Porporato, A. (2010). The role of tectonic uplift, climate and vegetation in the long-term terrestrial phosphorous cycle. Biogeosciences, 7(6), 2025-2038.

Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., ... & Blonder, B. (2017). Mapping local and global variability in plant trait distributions. Proceedings of the National Academy of Sciences, 114(51), E10937-E10946.

Caldwell, B. A. (2005). Enzyme activities as a component of soil biodiversity: a review. Pedobiologia, 49(6), 637-644.

Cleveland, C. C., Houlton, B. Z., Smith, W. K., Marklein, A. R., Reed, S. C., Parton, W., ... & Running, S. W. (2013). Patterns of new versus recycled primary production in the terrestrial biosphere. Proceedings of the National Academy of Sciences, 110(31), 12733-12737.

Crous, K. Y., Wujeska-Klause, A., Jiang, M., Medlyn, B. E., & Ellsworth, D. S. (2019). Nitrogen and phosphorus retranslocation of leaves and stemwood in a mature Eucalyptus forest exposed to 5 years of elevated CO_2. Frontiers in Plant Science, 10, 664.

Crowley, K. F., McNeil, B. E., Lovett, G. M., Canham, C. D., Driscoll, C. T., Rustad, L. E., ... & Goodale, C. L. (2012). Do nutrient limitation patterns shift from nitrogen toward phosphorus with increasing nitrogen deposition across the northeastern United States?. Ecosystems, 15(6), 940-957.

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. Nature, 440(7081), 165-173.

Deng, Q., Hui, D., Luo, Y., Elser, J., Wang, Y. P., Loladze, I., ... & Dennis, S. (2015). Down-regulation of tissue N: P ratios in terrestrial plants by elevated CO_2. Ecology, 96(12), 3354-3362.

Doetterl, S., Kearsley, E., Bauters, M., Hufkens, K., Lisingo, J., Baert, G., ... & Boeckx, P. (2015). Aboveground vs. belowground carbon stocks in African tropical lowland rainforest: drivers and implications. PloS one, 10(11), e0143209.

Du, E., Terrer, C., Pellegrini, A. F., Ahlström, A., van Lissa, C. J., Zhao, X., ... & Jackson, R. B. (2020). Global patterns of terrestrial nitrogen and phosphorus limitation. Nature Geoscience, 13(3), 221-226.

Elser, J. J., Bracken, M. E., Cleland, E. E., Gruner, D. S., Harpole, W. S., Hillebrand, H., ... & Smith, J. E. (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecology letters, 10(12), 1135-1142.

Ellsworth, D. S., Anderson, I. C., Crous, K. Y., Cooke, J., Drake, J. E., Gherlenda, A. N., ... & Tjoelker, M. G. (2017). Elevated CO 2 does not increase eucalypt forest productivity on a low-phosphorus soil. Nature Climate Change, 7(4), 279-282.

Ellsworth, D. S., Crous, K. Y., Lambers, H., & Cooke, J. (2015). Phosphorus recycling in photorespiration maintains high photosynthetic capacity in woody species. Plant, Cell &

Environment, 38(6), 1142-1156.

Fang, Y., Liu, C., Huang, M., Li, H., & Leung, L. R. (2014). Steady state estimation of soil organic carbon using satellite-derived canopy leaf area index. Journal of Advances in Modeling Earth Systems, 6(4), 1049-1064.

Fang, X. M., Zhang, X. L., Chen, F. S., Zong, Y. Y., Bu, W. S., Wan, S. Z., ... & Wang, H. (2019). Phosphorus addition alters the response of soil organic carbon decomposition to nitrogen deposition in a subtropical forest. Soil Biology and Biochemistry, 133, 119-128.

Fernández-Martínez, M., Vicca, S., Janssens, I. A., Luyssaert, S., Campioli, M., Sardans, J., ... & Peñuelas, J. (2014). Spatial variability and controls over biomass stocks, carbon fluxes, and resource-use efficiencies across forest ecosystems. Trees, 28(2), 597-611.

Filippelli, G. M. (2008). The global phosphorus cycle: past, present, and future. *Elements*, 4(2), 89-95.

Fleischer, K., Rammig, A., De Kauwe, M. G., Walker, A. P., Domingues, T. F., Fuchslueger, L., ... & Haverd, V. (2019). Amazon forest response to CO_2 fertilization dependent on plant phosphorus acquisition. *Nature Geoscience*, 12(9), 736-741.

Friedlingstein, P., Jones, M., O'sullivan, M., Andrew, R., Hauck, J., Peters, G., ... & DBakker, O. (2019). Global carbon budget 2019. *Earth System Science Data*, 11(4), 1783-1838.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate, 27(2)*, 511-526.

Gleason, S. M., Read, J., Ares, A., & Metcalfe, D. J. (2009). Phosphorus economics of tropical rainforest species and stands across soil contrasts in Queensland, Australia: understanding the effects of soil specialization and trait plasticity. Functional Ecology, 23(6), 1157-1166.

Gill, Allison L., and Adrien C. Finzi. "Belowground carbon flux links biogeochemical cycles and resource-use efficiency at the global scale." Ecology Letters 19.12 (2016): 1419-1428.

Goll, D. S., Brovkin, V., Parida, B. R., Reick, C. H., Kattge, J., Reich, P. B., ... & Niinemets, Ü. (2012). Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling. Biogeosciences, 9, 3547-3569.

Goll, D. S., Joetzjer, E., Huang, M., & Ciais, P. (2018). Low phosphorus availability decreases susceptibility of tropical primary productivity to droughts. Geophysical Research Letters, 45(16), 8231-8240.

Goll, D. S., Moosdorf, N., Hartmann, J., & Brovkin, V. (2014). Climate-driven changes in chemical weathering and associated phosphorus release since 1850: Implications for the land carbon balance. Geophysical Research Letters, 41(10), 3553-3558.

Gruber, N., & Galloway, J. N. (2008). An Earth-system perspective of the global nitrogen cycle. Nature, 451(7176), 293-296.

Hartmann, J., Moosdorf, N., Lauerwald, R., Hinderer, M., & West, A. J. (2014). Global chemical weathering and associated P-release—The role of lithology, temperature and soil properties. Chemical Geology, 363, 145-163.

Hasegawa, S., Macdonald, C. A., & Power, S. A. (2016). Elevated carbon dioxide increases soil nitrogen and phosphorus availability in a phosphorus-limited Eucalyptus woodland. Global Change Biology, 22(4), 1628-1643.

Hawkesford, M., Horst, W., Kichey, T., Lambers, H., Schjoerring, J., Møller, I. S., & White, P. (2012).

IPCC, 2019: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial

ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.

Functions of macronutrients. In Marschner's mineral nutrition of higher plants (pp. 135-189). Academic Press.

Helfenstein, J., Tamburini, F., von Sperber, C., Massey, M. S., Pistocchi, C., Chadwick, O. A., ... & Frossard, E. (2018). Combining spectroscopic and isotopic techniques gives a dynamic view of phosphorus cycling in soil. Nature communications, 9(1), 1-9.

Helfenstein, J., Pistocchi, C., Oberson, A., Tamburini, F., Goll, D. S., & Frossard, E. (2020). Estimates of mean residence times of phosphorus in commonly considered inorganic soil phosphorus pools. Biogeosciences, 17(2), 441-454.

Hermans, C., Hammond, J. P., White, P. J., & Verbruggen, N. (2006). How do plants respond to nutrient shortage by biomass allocation?. Trends in plant science, 11(12), 610-617.

Hinsinger, P. (2001). Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: a review. Plant and soil, 237(2), 173-195.

Hodapp, D., Hillebrand, H., & Striebel, M. (2019). “Unifying” the concept of resource use efficiency in ecology. Frontiers in Ecology and Evolution, 6, 233.

Hou, E., Lu, X., Jiang, L., Wen, D., & Luo, Y. (2019). Quantifying soil phosphorus dynamics: a data assimilation approach. Journal of Geophysical Research: Biogeosciences, 124(7), 2159-2173.

Hungate, B. A., Dukes, J. S., Shaw, M. R., Luo, Y., & Field, C. B. (2003). Nitrogen and climate change. Science, 302(5650), 1512-1513.

Izquierdo, J. E., Houlton, B. Z., & van Huysen, T. L. (2013). Evidence for progressive phosphorus limitation over long-term ecosystem development: examination of a biogeochemical paradigm. Plant and soil, 367(1-2), 135-147.

Jahnke, R. A. (1992). 14 The Phosphorus Cycle. In International Geophysics (Vol. 50, pp. 301-315). Academic Press.

Jiang, M., Caldararu, S., Zhang, H., Fleischer, K., Crous, K. Y., Yang, J., ... & Zaehle, S. (2020a). Low phosphorus supply constrains plant responses to elevated CO_2: A meta-analysis. Global change biology, 26(10), 5856-5873.

Jiang, M., Caldararu, S., Zaehle, S., Ellsworth, D. S., & Medlyn, B. E. (2019). Towards a more physiological representation of vegetation phosphorus processes in land surface models. New Phytologist, 222(3), 1223-1229.

Jiang, M., Medlyn, B. E., Drake, J. E., Duursma, R. A., Anderson, I. C., Barton, C. V., ... & Crous, K. Y. (2020b). The fate of carbon in a mature forest under carbon dioxide enrichment. Nature, 580(7802), 227-231.

Johnson, A. H., Frizano, J., & Vann, D. R. (2003). Biogeochemical implications of labile phosphorus in forest soils determined by the Hedley fractionation procedure. Oecologia, 135(4), 487-499.

Jonard, M., Fürst, A., Verstraeten, A., Thimonier, A., Timmermann, V., Potočić, N., ... & Ponette, Q. (2015). Tree mineral nutrition is deteriorating in Europe. Global change biology, 21(1), 418-430.

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

Knops, J. M., Wedin, D. A., & Naeem, S. (2010). The role of litter quality feedbacks in terrestrial nitrogen and phosphorus cycling. Open Ecol J, 3, 14–25.

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., ... & Swenson, S. C. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. Biogeosciences, 10(11), 7109.

Li, J., Guo, Q., Zhang, J., Korpelainen, H., & Li, C. (2016). Effects of nitrogen and phosphorus supply on growth and physiological traits of two Larix species. Environmental and Experimental Botany, 130, 206-215.

Liu, Y., Piao, S., Gasser, T., Ciais, P., Yang, H., Wang, H., ... & Wang, K. (2019). Field-experiment constraints on the enhancement of the terrestrial carbon sink by CO 2 fertilization. Nature Geoscience, 12(10), 809-814.

Luo, Y., Su, B. O., Currie, W. S., Dukes, J. S., Finzi, A., Hartwig, U., ... & Pataki, D. E. (2004). Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. Bioscience, 54(8), 731-739.

MacDonald, G. K., Bennett, E. M., Potter, P. A., & Ramankutty, N. (2011). Agronomic phosphorus imbalances across the world's croplands. Proceedings of the National Academy of Sciences, 108(7), 3086-3091.

Marklein, A. R., & Houlton, B. Z. (2012). Nitrogen inputs accelerate phosphorus cycling rates across a wide variety of terrestrial ecosystems. New Phytologist, 193(3), 696-704.

Norby, R. J., Gu, L., Haworth, I. C., Jensen, A. M., Turner, B. L., Walker, A. P., ... & Winter, K. (2017). Informing models through empirical relationships between foliar phosphorus, nitrogen and photosynthesis across diverse woody species in tropical forests of Panama. New Phytologist, 215(4), 1425-1437.

Okin, G. S., Baker, A. R., Tegen, I., Mahowald, N. M., Dentener, F. J., Duce, R. A., ... & Prospero, J. M. (2011). Impacts of atmospheric nutrient deposition on marine productivity: Roles of nitrogen, phosphorus, and iron. Global Biogeochemical Cycles, 25(2).

Penuelas, J., Poulter, B., Sardans, J., Ciais, P., Van Der Velde, M., Bopp, L., ... & Nardin, E. (2013). Human-induced nitrogen–phosphorus imbalances alter natural and managed ecosystems across the globe. Nature communications, 4(1), 1-10.

Penuelas, J., Fernández-Martínez, M., Vallicrosa, H., Maspons, J., Zuccarini, P., Carnicer, J., ... & Ciais, P. (2020). Increasing atmospheric CO 2 concentrations correlate with declining nutritional status of European forests. Communications biology, 3(1), 1-11.

Perring, M. P., Hedin, L. O., Levin, S. A., McGroddy, M., & De Mazancourt, C. (2008). Increased plant growth from nitrogen addition should conserve phosphorus in terrestrial ecosystems. Proceedings of the National Academy of Sciences, 105(6), 1971-1976.

Prieto, I., & Querejeta, J. I. (2020). Simulated climate change decreases nutrient resorption from senescing leaves. Global Change Biology, 26(3), 1795-1807.

Reed, S. C., Yang, X., & Thornton, P. E. (2015). Incorporating phosphorus cycling into global modeling efforts: a worthwhile, tractable endeavor. New Phytologist, 208(2), 324-329.

Rumble J (2020). ABUNDANCE OF ELEMENTS IN THE EARTH'S CRUST AND IN THE SEA, CRC Handbook of Chemistry and Physics, 100th edition (2019–2020).

Sardans, J., Alonso, R., Carnicer, J., Fernández-Martínez, M., Vivanco, M. G., & Peñuelas, J. (2016). Factors influencing the foliar elemental composition and stoichiometry in forest trees in Spain. Perspectives in Plant Ecology, Evolution and Systematics, 18, 52-69.

Sardans, J., Peñuelas, J., & Estiarte, M. (2006). Warming and drought alter soil phosphatase activity and soil P availability in a Mediterranean shrubland. Plant and Soil, 289(1-2), 227-238.

Schachtman, D. P., Reid, R. J., & Ayling, S. M. (1998). Phosphorus uptake by plants: from soil to cell. Plant physiology, 116(2), 447-453.

Schmitz, A., Sanders, T. G., Bolte, A., Bussotti, F., Dirnböck, T., Johnson, J., ... & Verstraeten, A. (2019). Responses of forest ecosystems in Europe to decreasing nitrogen deposition. Environmental pollution, 244, 980-994.

See, C. R., Yanai, R. D., Fisk, M. C., Vadeboncoeur, M. A., Quintero, B. A., & Fahey, T. J. (2015). Soil nitrogen affects phosphorus recycling: foliar resorption and plant–soil feedbacks in a northern hardwood forest. Ecology, 96(9), 2488-2498.

Sharpley, A. N., Chapra, S. C., Wedepohl, R., Sims, J. T., Daniel, T. C., & Reddy, K. R. (1994). Managing agricultural phosphorus for protection of surface waters: Issues and options. Journal of environmental quality, 23(3), 437-451.

Shi, M., Yang, Z. L., Lawrence, D. M., Dickinson, R. E., & Subin, Z. M. (2013). Spin-up processes in the Community Land Model version 4 with explicit carbon and nitrogen components. Ecological Modelling, 263, 308-325.

Šímová, I., Sandel, B., Enquist, B. J., Michaletz, S. T., Kattge, J., Violle, C., ... & Wiser, S. K. (2019). The relationship of woody plant size and leaf nutrient content to large-scale productivity for forests across the Americas. Journal of Ecology, 107(5), 2278-2290.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., et al. (2013), Climate change 2013: The physical science basis, Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, Cambridge: Cambridge University Press.

Sun, Y., Goll, D. S., Ciais, P., Peng, S., Margalef, O., Asensio, D., ... & Peñuelas, J. (2020). Spatial pattern and environmental drivers of acid phosphatase activity in Europe.

Sun, Y., Peng, S., Goll, D. S., Ciais, P., Guenet, B., Guimberteau, M., ... & Poulter, B. (2017). Diagnosing phosphorus limitations in natural terrestrial ecosystems in carbon cycle models. Earth's future, 5(7), 730-749.

Terrer, C., Jackson, R. B., Prentice, I. C., Keenan, T. F., Kaiser, C., Vicca, S., ... & Peñuelas, J. (2019). Nitrogen and phosphorus constrain the CO 2 fertilization of global plant biomass. Nature Climate Change, 9(9), 684-689.

Thornton, P. E., Lamarque, J. F., Rosenbloom, N. A., & Mahowald, N. M. (2007). Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. Global biogeochemical cycles, 21(4).

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

Thum, T., Caldararu, S., Engel, J., Kern, M., Pallandt, M., Schnur, R., ... & Zaehle, S. (2019). A new terrestrial biosphere model with coupled carbon, nitrogen, and phosphorus cycles (QUINCY v1. 0; revision 1772). Geoscientific Model Development, 12, 4781-4802.

Vance, C. P., Uhde-Stone, C., & Allan, D. L. (2003). Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New phytologist, 157(3), 423-447.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., ... & Tilman, D. G. (1997). Human alteration of the global nitrogen cycle: sources and consequences. Ecological applications, 7(3), 737-750.

Vitousek, P. M., Ladefoged, T. N., Kirch, P. V., Hartshorn, A. S., Graves, M. W., Hotchkiss, S. C., ... & Chadwick, O. A. (2004). Soils, agriculture, and society in precontact Hawaii. Science, 304(5677), 1665-1669.

Vitousek, P. M., Porder, S., Houlton, B. Z., & Chadwick, O. A. (2010). Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen–phosphorus interactions. Ecological applications, 20(1), 5-15.

Walker, A. P., Beckerman, A. P., Gu, L., Kattge, J., Cernusak, L. A., Domingues, T. F., ... & Woodward, F. I. (2014). The relationship of leaf photosynthetic traits–Vcmax and Jmax–to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. Ecology and evolution, 4(16), 3218-3235.

Walker, T. W., & Syers, J. K. (1976). The fate of phosphorus during pedogenesis. Geoderma, 15(1), 1-19.

Wang, Y. P., Law, R. M., & Pak, B. (2010). A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences, 7(7).

Wang, Y., Ciais, P., Goll, D. S., Huang, Y., Luo, Y., Wang, Y. P., ... & Penuelas, J. (2018). GOLUM-CNP v1. 0: a data-driven modeling of carbon, nitrogen and phosphorus cycles in major terrestrial biomes.

White, P. J., & Hammond, J. P. (2008). Phosphorus nutrition of terrestrial plants. In The ecophysiology of plant-phosphorus interactions (pp. 51-81). Springer, Dordrecht.

Wieder, W. R., Cleveland, C. C., Smith, W. K., & Todd-Brown, K. (2015). Future productivity and carbon storage limited by terrestrial nutrient availability. Nature Geoscience, 8(6), 441-444.

Wright, S. J., Yavitt, J. B., Wurzburger, N., Turner, B. L., Tanner, E. V., Sayer, E. J., ... & Garcia, M. N. (2011). Potassium, phosphorus, or nitrogen limit root allocation, tree growth, or litter production in a lowland tropical forest. Ecology, 92(8), 1616-1625.

Wright, S. J. (2019). Plant responses to nutrient addition experiments conducted in tropical forests. Ecological Monographs, 89(4), e01382.

Xia, J. Y., Luo, Y. Q., Wang, Y. P., Weng, E. S., & Hararuk, O. (2012). A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. Geoscientific Model Development, 5(5), 1259-1271. https://doi.org/10.5194/gmd-5-1259-2012.

Xu, S., Sardans, J., Zhang, J., & Peñuelas, J. (2020). Variations in foliar carbon: nitrogen and nitrogen: phosphorus ratios under global change: a meta-analysis of experimental field studies. Scientific reports, 10(1), 1-11.

Yang, X., Thornton, P. E., Ricciuto, D. M., & Post, W. M. (2014). The role of phosphorus dynamics in tropical forests--a modeling study using CLM-CNP. Biogeosciences, 11(6).

Yu, H., Chin, M., Yuan, T., Bian, H., Remer, L. A., Prospero, J. M., ... & Zhang, Z. (2015). The fertilizing role of African dust in the Amazon rainforest: A first multiyear assessment based on data from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations. Geophysical Research Letters, 42(6), 1984-1991.

Yuan, Z. Y., & Chen, H. Y. (2015). Decoupling of nitrogen and phosphorus in terrestrial plants associated with global changes. Nature Climate Change, 5(5), 465-469.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., ... & Jain, A. (2014). Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate Free-Air CO_2 Enrichment studies. New Phytologist, 202(3), 803-822.

Zhang, Q., Wang, Y. P., Matear, R. J., Pitman, A. J., & Dai, Y. J. (2014). Nitrogen and phosphorous limitations significantly reduce future allowable CO_2 emissions. Geophysical Research Letters, 41(2), 632-637.

Chapter 2 Spatial patterns of acid phosphatase activity and their environmental drivers: a new method for upscaling field data applied to the European continent

Summary

The lack of benchmarking data for key processes for phosphorus (P) cycling to evaluate land surface models (LSMs) leads to large uncertainties in modeled P cycling and effects of P availability changes on the land carbon (C) cycle. The recycling of P from soil organic matter is enhanced by phosphatase enzymes produced by plants and microbes, as phosphatase helps to cleave organic phosphorus and liberate phosphate ions used by microbes and plants. Phosphatase production is thus a key process in models, and depending on its representation, it can lead to strong or no productivity limitation in simulations. A quantification of the spatial variation in phosphatase activity in soils does not exist. Having such a dataset of soil phosphatase activity in a gridded form will not solve all the problems of models, but it is a first step to evaluate them. The goal of this chapter is to create such a new dataset based on a large compilation of point field samples of acid phosphatase activity (AP). Initially, I wanted to create a dataset for the globe, but the extrapolation of scarce measurements to under-sampled regions has proven to be too uncertain. Therefore, most of the results are presented for the European continent, where the sampling is denser.

I applied two machine-learning methods (Random forests and back-propagation artificial networks) to simulate the spatial patterns of potential acid phosphatase (AP) activity across Europe by scaling up 126 site observations of potential AP activity from field samples measured in the laboratory, using 12 environmental drivers as predictors. The back-propagation artificial network (BPN) method explained 58% of AP variability, more than the regression tree model (49%). In addition, BPN was able to identify the gradients in AP along three transects in Europe. Partial correlation analysis revealed that soil nutrients (total nitrogen, total P and labile organic P) and climatic controls (annual precipitation, mean annual temperature and temperature amplitude) were the dominant factors influencing AP variations in space. Higher AP occurred in regions with higher mean annual temperature, precipitation and higher soil total nitrogen. Soil TP and Po were non-monotonically correlated with modelled AP for Europe, indicating different strategies of P utilization in arid and humid ecosystems.

This study provided spatial gridded maps of AP in Europe as well as relationships with environmental drivers which can be used to evaluate models. A shortcoming of the study is related to scarce data from tropical regions, thus a robust map could only be achieved for Europe. As more measurements of AP for tropical regions become available, our approach could be updated to understand the physiological basis of P-use strategies by AP production in natural soils worldwide.

This chapter has been published as Sun, Y., Goll, D. S., Ciais, P., Peng, S., Margalef, O., Asensio, D., Sardans, J., and Peñuelas, J.: Spatial 1560 pattern and environmental drivers of acid phosphatase activity in Europe, Front. Big Data, http://doi.org/10.3389/fdata.2019.00051, 2020.

frontiers in Big Data

ORIGINAL RESEARCH
published: 23 January 2020
doi: 10.3389/fdata.2019.00051

Spatial Pattern and Environmental Drivers of Acid Phosphatase Activity in Europe

Yan Sun[1], Daniel S. Goll[1,2], Philippe Ciais[1], Shushi Peng[3], Olga Margalef[4,5], Dolores Asensio[4,5], Jordi Sardans[4,5] and Josep Peñuelas[4,5]*

[1] Laboratoire des Sciences du Climat et de 1'Environnement, CEA-CNRS-UVSQ, Gif sur Yvette, France, [2] Institute of Geography, University of Augsburg, Augsburg, Germany, [3] Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing, China, [4] CSIC, Global Ecology Unit, Centre de Recerca Ecològica i Aplicacions Forestals, Consejo Superior de Investigaciones Cientificas, UAB, Bellaterra, Spain, [5] Centre de Recerca Ecològica i Aplicacions Forestals, Cerdanyola del Vallès, Spain

Acid phosphatase produced by plants and microbes plays a fundamental role in the recycling of soil phosphorus (P). A quantification of the spatial variation in potential acid phosphatase activity (AP) on large spatial scales and its drivers can help to reduce the uncertainty in our understanding of bio-availability of soil P. We applied two machine-learning methods (Random forests and back-propagation artificial networks) to simulate the spatial patterns of AP across Europe by scaling up 126 site observations of potential AP activity from field samples measured in the laboratory, using 12 environmental drivers as predictors. The back-propagation artificial network (BPN) method explained 58% of AP variability, more than the regression tree model (49%). In addition, BPN was able to identify the gradients in AP along three transects in Europe. Partial correlation analysis revealed that soil nutrients (total nitrogen, total P, and labile organic P) and climatic controls (annual precipitation, mean annual temperature, and temperature amplitude) were the dominant factors influencing AP variations in space. Higher AP occurred in regions with higher mean annual temperature, precipitation and higher soil total nitrogen. Soil TP and Po were non-monotonically correlated with modeled AP for Europe, indicating diffident strategies of P utilization by biomes in arid and humid area. This study helps to separate the influences of each factor on AP production and to reduce the uncertainty in estimating soil P availability. The BPN model trained with European data, however, could not produce a robust global map of AP due to the lack of representative measurements of AP for tropical regions. Filling this data gap will help us to understand the physiological basis of P-use strategies in natural soils.

Keywords: back-propagation artificial network, Europe, phosphorus cycling, partial correlation analysis, regression tree, soil acid phosphatase

OPEN ACCESS

Edited by:
Forrest M. Hoffman, Oak Ridge National Laboratory (DOE), United States

Reviewed by:
Jitendra Kumar, Oak Ridge National Laboratory (DOE), United States
Qingyuan Zhang, Universities Space Research Association (USRA), United States

****Correspondence:***
Yan Sun
ysun@lsce.ipsl.fr

Specialty section:
This article was submitted to Data-driven Climate Sciences, a section of the journal Frontiers in Big Data

Received: *31 May 2019*
Accepted: *26 December 2019*
Published: *23 January 2020*

Citation:
Sun Y, Goll DS, Ciais P, Peng S, Margalef O, Asensio D, Sardans J and Peñuelas J (2020) Spatial Pattern and Environmental Drivers of Acid Phosphatase Activity in Europe. Front. Big Data 2:51. doi: 10.3389/fdata.2019.00051

INTRODUCTION

Phosphorus (P) is an essential nutrient for all living organisms (White and Hammond, 2008). As weathering of minerals and the deposition of atmospheric dust are minor sources of P (Walker and Syers, 1976; Vitousek et al., 2010; Wang et al., 2015), the recycling of organic P from litter and soil organic matter is of utter importance for plant growth and microbial activity in terrestrial ecosystems. In P-poor ecosystems, limited P recycling may dampen the response of plant growth to

elevated CO_2 concentration (Ellsworth et al., 2017). Yang et al. (2014) reported in a modeling study that the effect of elevated CO_2 on plant productivity in the Amazon Basin critically dependent on assumptions regarding the P-recycling efficiency within soils, which was strongly related to the parameterization of phosphatase production in their model.

The rate at which ecosystems can recycle P from litter and soil organic matter is poorly quantified by observation (Gill and Finzi, 2016). Soil phosphatases secreted by fungi, bacteria, and plant roots play an important but poorly quantified role in transforming complex and unavailable forms of organic P into assimilable phosphate (Caldwell, 2005). Potential phosphatase activity in soils, which can be measured in the lab from soil samples, is an indicator of the capacity of enzyme communities to cleave organic molecules containing P (Krämer and Green, 2000), and serves as a surrogate for the lacking measurements of P mineralization in the soil.

Potential phosphatase activity measured in the laboratory under optimal conditions (optimal temperature, well-mixed soil, no water limitation; Eivazi and Tabatabai, 1977) provides an upper limit of their actual activity in a soil (Nannipieri et al., 2011; Margalef et al., 2017), which cannot be directly measured. Acid phosphatases (AP) are more widespread than alkaline phosphatases at soil pH values representative of most natural soils (Margalef et al., 2017), which justifies our focus on AP in this study.

Many experiments have investigated the responses of potential AP activity to fertilization (Marklein and Houlton, 2012; Maistry et al., 2015), temperature changes (Sardans et al., 2006), and water availability (Sardans and Peñuelas, 2005; Sardans et al., 2006; Zhou et al., 2013) or to other disturbances under controlled conditions (Sinsabaugh et al., 2008). Gradients of AP along transects have been measured in few regions (Brockett et al., 2012; Huang et al., 2013; Kitayama, 2013). They found that warming, increasing soil water and nitrogen availability can enhance AP activity. However, these studies were limited to site or small region scale and only considering a subset of potential environmental factors. Recently, Margalef et al. (2017) addressed this gap, by compiling a global data set of phosphatase activity and using correlation analysis, regression analysis and structural equation models (SEMs) to provide insights of the drivers of phosphatase activity distribution on global scale. However, the approach used by Margalef et al. (2017) cannot account for non-linear responses to different variables (Ma et al., 2017) and omitted important variable like soil labile P.

The benchmarking of the growing number of land models which include phosphorus cycling is currently hampered by the lack of spatial explicit information on AP on regional to global scale. The limited understanding of the drivers responsible for differences in AP across different ecosystems and climatic and soils conditions further hampers the global efforts of including P cycles in the land surface models (Reed et al., 2015).

Machine learning (ML) is a family of approaches which has been increasingly used to identify patterns in complex ecological data sets and scale up site measurements (e.g., Papale and Valentini, 2003; Golubiewski, 2006; Wiesmeier et al., 2011; Keenan et al., 2012; Were et al., 2015), but have not been used for upscaling the spatial patterns of AP.

We used two ML methods in combination with gridded fields of environmental factors to upscale site data of potential AP (Margalef et al., 2017) to gridded AP fields for continental Europe. Then we identify the main drivers behind the spatial variation of upscaled AP in Europe. Finally, we used the best ML model trained by European data in a first attempt to produce a global map of AP, and this map is cross-validated using non-European data.

METHODS

AP measurements compiled by Margalef et al. (2017) are combined with gridded fields of 12 environmental factors to model spatial pattern of annual AP across Europe and the globe on a spatial resolution of 10 km. This was done by using back-propagation artificial network (BPNs) and Regression Trees (RT). **Figure 1** shows the main procedure of model calibration and evaluation.

Datasets

We used 296 measurements of AP from 139 published studies (Margalef et al., 2017). This dataset contained 54 tropical sites and 242 ex-tropical sites (155 in Europe). Location, soil pH and soil nutrient contents (total soil C, soil organic C, total nitrogen, and total P) for each site were obtained from the original publications.

We separated the data set into four subsets as shown in **Figure 1**: (1) data used to train the ML models (*Data set A*), (2) data used for cross-validating the spatial gradients of AP modeled for Europe (*Data set B*), (3) data used to evaluate AP patterns in temperate regions outside Europe (*Data set C*), and (4) data used to evaluate AP patterns in tropical regions (*Data set D*). *Dataset A* contains 126 European sites with complete information for the all the predictors listed below. *Data set B* contains all 155 European sites. *Data set C* and *Data set D* contained 87 temperate sites outside Europe and only 54 tropical sites, respectively.

We selected 12 variables as predictors for the upscaling based on findings from manipulation experiments (Hogan et al., 2010; Huang et al., 2011; Marklein and Houlton, 2012), observation time series (Sardans et al., 2006) and previous regression analyses (Margalef et al., 2017): soil pH (SoilpH), soil clay content (Clay), soil organic carbon (OC), soil total nitrogen (TN), soil total phosphorus (TP) and soil labile organic phosphorus (Po), net primary productivity (NPP), mean annual temperature (MAT), amplitude of yearly temperature (AMP), mean annual precipitation (MAP), soil type (SoilType), and vegetation type (VegType).

We extracted predictors from each original publication when available. In case information on predictors were not reported, we extracted the missing data from gridded datasets (**Table 1**) based on the geographical coordinates of each measurement site (Shangguan et al., 2014; Ballabio et al., 2016, 2019; Fick and Hijmans, 2017; Hengl et al., 2017; MOD12Q2 and MOD17). The databases for each predictor variables are listed in **Table 1**. Detailed information about the databases

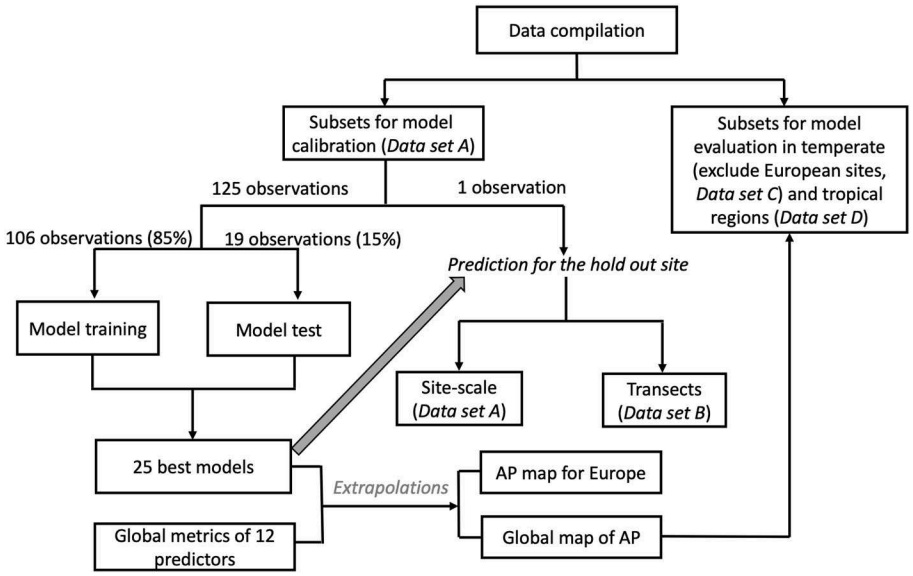


FIGURE 1 | Flow chart for the calibration and evaluation of the prediction models.

TABLE 1 | Predictors used in prediction models and its source.

Predictor name	Abbreviation	Source	Spatial resolution	References
Soil organic carbon	OC	Soilgrids	250 m	Hengl et al., 2017
Soil pH	SoilpH	LUCAS	500 m	Ballabio et al., 2019
Soil total nitrogen	TN			
Soil total phosphorus	TP			
Soil clay content	Clay			Ballabio et al., 2016
Soil labile organic phosphorus	Po	Global maps of the soil P contents for different P forms TP from LUCAS USDA soil types and ratio of labile inorganic P and labile organic P	1 km	Yang et al., 2013; Hengl et al., 2017; Sun et al., 2017; Ballabio et al., 2019
Net primary productivity	NPP	MODIS-NPP (MOD17A3; mean value during 2000–2014)	1 km	Running et al., 2004; Zhao et al., 2005; Turner et al., 2006
Mean annual temperature	MAT	WorldClim	1 km	Fick and Hijmans, 2017
Yearly temperature amplitude	AMP			
Mean annual precipitation	MAP			
Soil type (categorical variable)	SoilType	Soilgrids and USDA class	250 m	Hengl et al., 2017
Vegetation type (categorical variable)	VegType	MODIS	1 km	Friedl et al., 2010

are provided in **Supporting Information Appendix S1**. Details on the gap filling of predictors can be found in **Supporting Information Appendix S2**. AP, NPP, AMP, and MAP, which followed a lognormal distribution, were log-transformed before build the prediction model (**Supporting Information Appendix S2**, **Figure S2**).

Back-Propagation Artificial Networks

Back-propagation training algorithm (Kelley, 1960) is the most frequently used neural-network method (Were et al., 2015). These algorithms train networks until some targeted minimal error is achieved between the predicted and observed outputs (Kelley, 1960). We applied a BPN constituting a four-layer neural network: one input layer, two hidden layers and one output layer. Two categorical variables (vegetation type and soil type) were converted into vector form as BPN are not designed to handle categorical values following Mason et al. (2017). For example, the soil type has eight possible categories: (1) Entisol, (2) Inceptisol, (3) Aridsol, (4) Mollisol, (5) Alfisol, (6) Spodosol, (7) Ultisol, (8) Oxisol. An Entisol would have a vector value of [1 0 0 0 0 0 0 0], whereas a Alfisol would have a vector value of [0 0 0 0 1 0 0 0]. The input layer contained 34 nodes (10 numeral predictors

and two categorical predictors; **Table 1**). The numbers of neurons for the two hidden layers were 10 and 5 for building the BPN model (**Supporting Information Appendix S3**, **Figure S3**). Each layer of the BPN was linked to prior and forward layers by weights that were determined using a gradient-descent learning algorithm, such as Widrow-Hoff learning rule (Kelley, 1960). We used the resilient back-propagation algorithm to update weights and biases along the negative of the gradient of the performance function (Riedmiller and Braun, 1993).

Each BPN runs multiple epochs until either of three convergence criteria is satisfied: (1) the number of training epochs reach up to 500, (2) the predicting bias is lower than 0.2, or (3) the number of validation checks failure reached 100 times (i.e., the predicting bias for validation sets doesn't decrease for 100 times). Criterion (1) and (2) helps to avoid the over-fitting issue. Criterion (3) helps to avoid additional simulations without any improvement of BPN model.

Regression Trees

BPN models cannot rate the importance of predictors for accounting for differences in observed AP, so we also used RTs to build regression models. The input predictors, training and testing subset selection, leave-one-out validation and best model selection for the RT were the same as those used in the BPN method. The response data at each binary split were grouped into two descendant nodes to maximize homogeneity, and the best binary split was selected. The trees were grown to their maximum sizes following the tree template. We used an RT tree template where the minimal number of parents was 6, and the minimal number of observations at the terminal nodes of the trees was 3.

Leave-One-Out Validation Method and BPN and RT Model Selection

The leave-one-out (LOO) cross validation method is best suited for model calibration and evaluation when observational data are scarce (Allen, 1974). LOO can test the reproducibility of prediction models using independent data (Cawley, 2006). Out of the 126 European sites (*Data set A*), 125 observations were used to train and select models and the remaining observation to verify the model. This procedure was repeated 126 times by excluding each observation once.

Within the BPN and RT models training and selection, the 126 observations were randomly separated into training (106 sites, 85%) and testing (19 sites, 15%) data sets (**Figure 1**). The BPN models were trained 5,000 times using different training subsets, and the RT models were trained 500 times. We used R^2 and RMSE to assess the performance of the BPN and RT models on the training and testing data sets for each train-test possibility. We removed BPN and RT models with training R^2 <0.6, testing R^2 <0.5, training RMSE >15 μmol g^{-1} h^{-1} and testing RMSE >10 μmol g^{-1} h^{-1}. Then, we selected the 25 models with the highest test-R^2 for observed AP < 25 μmol g^{-1} h^{-1} and lowest RMSE for observed AP >25 μmol g^{-1} h^{-1}. This selection algorithm ensured that the BNP and RT can accurately reproduce AP in both the training and testing data sets.

Extrapolation of AP for Europe

We resampled all of those gridded datasets of predictors (**Table 1**, **Table S1**) on a 10 km resolution using area-weighted mean methods for numerical predictors, and by using mode resampling for categorical predictors (i.e., vegetation types and soil types). Then we used the 25 selected model ensembles (25 best models × 126 LOO rounds) that produced AP with the highest correlation with observed AP in combination with gridded maps of the predictors at a resolution of 10 km (**Figure 2**) to scale up point observations to continental Europe (36–75°N and 20°W to 32°E). We first used the median of the 25 estimates for each LOO round and then derived the median estimates of AP across 126 LOO rounds to define the AP maps and their uncertainties across the European continent.

Attribution of AP Pattern to Drivers Across Europe

ML models like BPN are able to extract hidden, complex, and generally subtle, non-linear relationships from scattered data and multiple predictors and use these relationships for up-scaling. The up-scaled AP maps are necessary but difficult to be interpreted by ML models due to the numbers of predictors and their interactions. We thus applied partial correlation analysis, a simple diagnostic analysis, on modeled maps of AP to gain a first order understanding of the most important predictors emerging from the model results in different regions.

The spatial partial correlation analysis for Europe was conducted at spatial resolution of half degree (∼50 km). To do so, predictors were firstly resampled into half degree by using the same methods as described above. Then we performed the partial correlation analysis between AP and predictors (only numerical predictors) using a spatial moving window of 4.5 × 4.5°. This analysis partly avoided the collinearity among the predictors but still failed in separating effects of strongly collinear factors (e.g., NPP and climatic variables). Therefore, we performed (1) a partial correlation analysis between AP and all abiotic factors, and (2) a partial correlation analysis between AP, soil conditions and NPP without climatic variables. The median estimate of AP and the predictor metrics for Europe were first resampled to 0.5°. We thus selected 81 pixels surrounding each 0.5 grid cell for calculating the coefficients of partial correlation for AP and the predictors.

RESULTS

Model Performance Across European Sites and Along Spatial Gradients

We assessed the performance of the BPN and RT models for each of the 126 European sites (*Data set A*) by LOO cross-validation. Excluding nine outliers with absolute bias >20 μmol g^{-1} h^{-1} and relative bias >50% (**Supporting Information Appendix S4**, **Figures S4, S5**), the BPN models explained 58% of AP variability across all sites, with an RMSE of 6.83 μmol g^{-1} h^{-1} (**Figure 3**, **Table S2**). The BPN models reproduced 93% of the variance of AP for the remaining 117 sites. The observations of AP for 68 sites were within the uncertainty quantile range of 10–90%

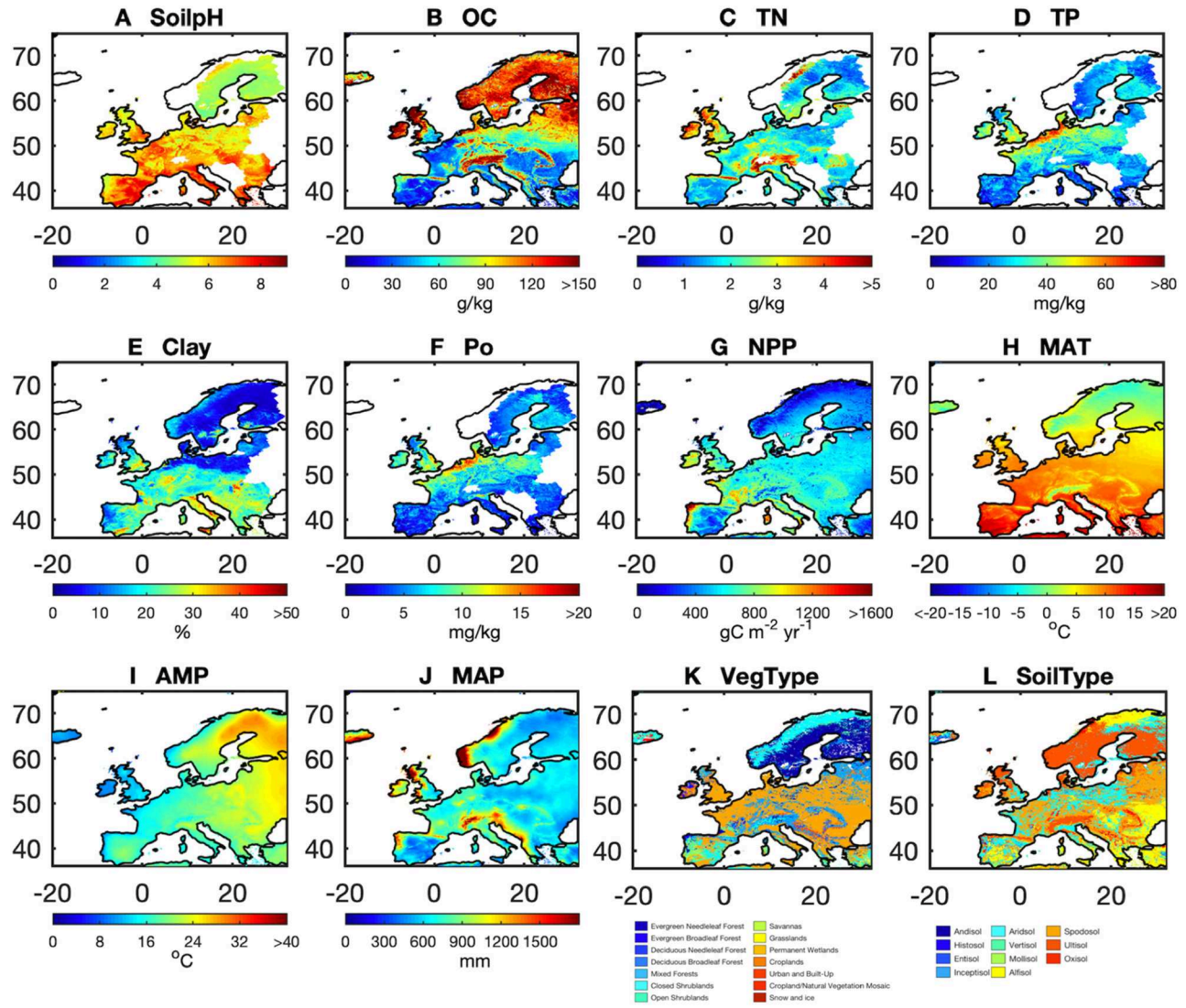


FIGURE 2 | Spatial pattern of the 12 predictors: **(A)** soil pH (soilpH), **(B)** soil organic carbon content (OC), **(C)** soil total nitrogen content (TN), **(D)** soil total phosphorus content (TP), **(E)** clay content (Clay), **(F)** soil labile organic phosphorus content (Po), **(G)** net primary productivity (NPP), **(H)** mean annual temperature (MAT), **(I)** amplitude of yearly temperature (AMP), **(J)** mean annual precipitation (MAP), **(K)** vegetation cover (VegType), and **(L)** soil type (SoilType).

of AP simulated with the ensemble of 25 BPN models, and 32 site observations were within the 25–75% quantile range (**Figure 3**). The BPN model generally underestimated AP, but overestimated AP at sites with low activity (<20 μmol g^{-1} h^{-1}), while AP at sites with intermediate and high observed activities (>20 μmol g^{-1} h^{-1}) tended to be systematically underestimated. The BPN models outperformed the RT models for LOO validation (**Table S2**). With excluding 13 outliers, RT models explained 48.7% of the spatial variance of AP over all measurements, accompanied by a higher RMSE of 7.85 μmol g^{-1} h^{-1} (**Figure S6**, **Table S2**).

Three transects (*Data set B*) were selected to determine if the BPN model could reproduce large scale observed spatial gradients (**Figures 4C–E**). *Transect 1* spanned across Spain, France, Germany, Poland and Sweden. Median estimates of modeled AP along this transect gradually decrease from 8.5 μmol g^{-1} h^{-1} in Spain to 5.6 μmol g^{-1} h^{-1} for Germany, which is roughly consistent with the observations (decrease from 11.8 to 2.3 μmol g^{-1} h^{-1}). Modeled AP for sites in Poland and Sweden were all higher than observed values, though. *Transect 2* (across England and France to Italy) shows observed values going from 7.8 μmol g^{-1} h^{-1} (6.6–9.9 μmol g^{-1} h^{-1}) in England to 4.2 μmol g^{-1} h^{-1} (0.4–9.8 μmol g^{-1} h^{-1}) in Italy. Predicted AP along *Transect 2* give a smaller decrease than observed, going from 7.4 μmol g^{-1} h^{-1} (6.1–9.3 μmol g^{-1} h^{-1}) to 7.2 μmol g^{-1} h^{-1} (5.5–8.8 μmol g^{-1} h^{-1}). Modeled AP along *Transect 3* (from Western and Eastern Spain, Italy and Turkey) decreased from 8.5 μmol g^{-1} h^{-1} (5.1–10.8 μmol g^{-1} h^{-1}) for Western Spain to 5.4 μmol

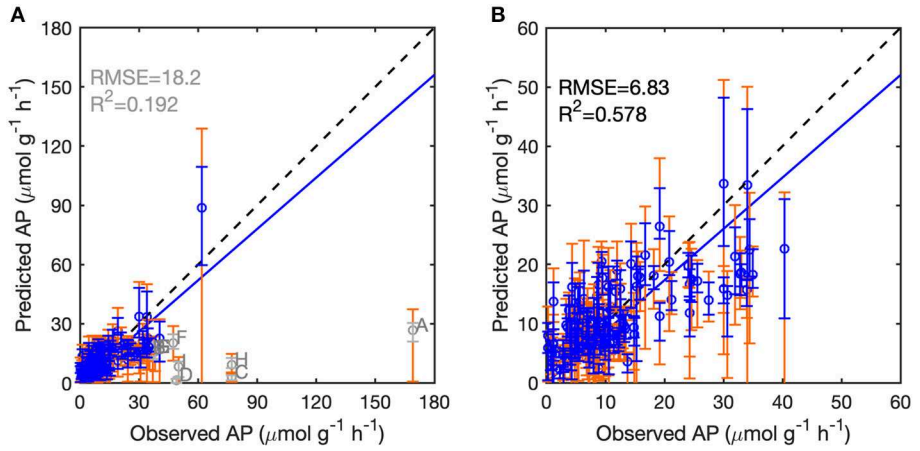


FIGURE 3 | Performance of the BPN models on 126 European sites following the leave-one-out framework. **(A)** Predicted vs. observed acid phosphatase activity (AP) for all sites. **(B)** Predicted vs. observed AP with excluding seven outliers (gray circles). The dashed lines indicate the 1:1 lines. The blue lines indicate the regression lines between predicted and observed AP. The blue and gray error bars indicate the 25 and 75% quantiles of predicted AP, and the orange error bars indicate the 10 and 90% quantiles of predicted AP. The nine outliers are marked in gray: *Site A* and *D* are located in central England, *Site B* is in Northern Italy, *Site C* in Eastern Spain, *Site E* in Poland, *Site* F in Norway, and *Site G* in Western Spain.

g^{-1} h^{-1} for Turkey, a smaller decrease than the observations. This bias could be mainly attributed to the underestimation of modeled AP for sites with high values in Spain.

In summary, The BPN models are found in reproduce the spatial variance of AP for Europe and the direction of the AP gradients across ecological zones and climate, despite negative or positive biases compared to individual site-scale data points. Median BPN modeled estimate of AP across Europe increases from north to south, and also from east to west (**Figure 4A**). AP was highest for Northern England and Western Spain (10–40 μmol g^{-1} h^{-1}), and lowest for Northern and Eastern Europe (<2 μmol g^{-1} h^{-1}) (**Figure 4**). The uncertainty in the AP estimates (the difference between 75 and 25% quantiles; **Figure S7**) was low for most of Europe (<3 μmol g^{-1} h^{-1}) except for England and Northern Europe (>3 μmol g^{-1} h^{-1}).

Relationship Between Modeled AP and Climate and Soil Conditions

The partial correlation analysis indicated that correlations between gridded TP and Po and gridded climatic factors (MAT, AMP, and MAP) with gridded BPN estimates of AP for Europe were stronger than with the other factors (**Figure 5**). The proportions of land areas with significant ($p < 0.1$) correlations between AP and MAT, AMP and MAP were 76, 79, and 71%, respectively. AP was generally correlated with both MAT and MAP (**Figures 6G–I**), being positively associated with spatially increasing MAT, and negatively correlated with AMP.

When excluding climatic variables from partial correlation analysis, NPP has a significant correlation ($p < 0.1$) with AP for 58% of land areas. Significant positive correlations between NPP and AP are mainly found in Southern and Western Europe, while significant negative correlations are in Sweden and Finland (**Figure 6K**). The emerging relationship between modeled AP and NPP is consistent with those found between AP and climatic variables (**Figures 6G–I,K**) which reflects strong collinearity between NPP and climatic variables (**Figure S9**).

Modeled AP is found to be negatively correlated with soil pH at high soil pH values (6–9), which is common for Europe, especially in the Southern and the Western of the domain (**Figure 6A**). Soil nutrient contents also partly accounted for the spatial patterns of modeled AP. AP is found to be positively correlated with soil organic C content (OC) (**Figures 6B–D,F**). More specifically, AP and OC are significantly positively correlated ($p < 0.1$) in Southern and Western Europe but not significantly negatively correlated for Northern and central Europe. AP and TN are strongly positively correlated with AP in Southern and the Western Europe. AP is significantly ($p < 0.1$) positively correlated with TP but negatively correlated with Po in aridity area (P/PET < 0.7) (**Figure 6**; **Figure S11**), while AP is significantly ($p < 0.1$) negatively correlated with TP but positively correlated with Po in wetter region (P/PET > 1; **Figure 6**; **Figure S11**).

Modeled AP is found to be higher in highly weathered soils (Ultisols and Oxisols, 7.8–8.3 μmol g^{-1} h^{-1}) compared to slightly and moderately weathered soils (1.2–7.2 μmol g^{-1} h^{-1}) (**Figure S8A**). AP was the highest (8.9 μmol g^{-1} h^{-1}) for Evergreen Broadleaf Forests and the lowest for Evergreen Needleleaf Forests among all of the biomes (**Figure S8B**).

Upscaling of Site Data to a Global Map of AP

We used the BPN models trained on European data in a first attempt to scale up AP to the global scale. The modeled global spatial pattern of AP show a clear latitudinal gradient of increasing AP from high latitudes to the equator (**Figure 7**). Modeled AP is also much higher for coastal than inland areas. The global up-scaled distribution of AP is however more uncertain than the European distribution.

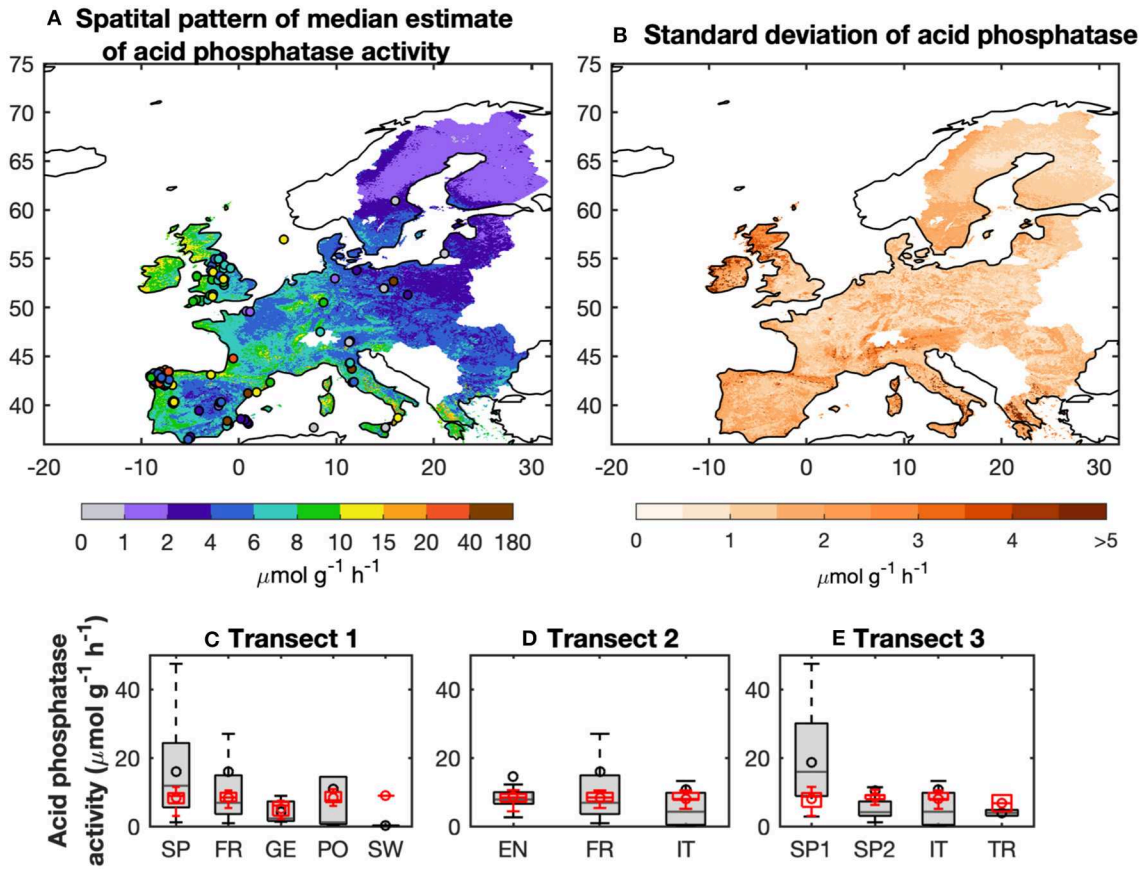


FIGURE 4 | Spatial pattern of median estimates of acid phosphatase activity (AP) from the BPN models for Europe **(A)**. The observed AP are also shown. **(B)** Uncertainties of AP defined as the difference between the 75 and 25% quantiles of predicted AP from the BPN models. **(C–E)** show the changes of observed and predicted AP along three transects: **(C)** Spain (SP), France (FR), Germany (GE), Poland (PO), and Sweden (SW); **(D)** England (EN), France (FR), and Italy (IT); and **(E)** Western Spain (west of 5°W; SP1), Eastern Spain (east of 5°W; SP2), Italy (IT), and Turkey (TR).

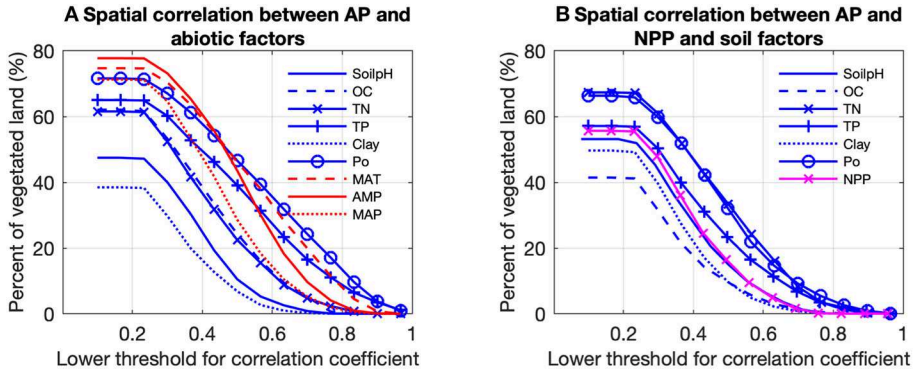


FIGURE 5 | Proportions of vegetated land in Europe controlled by **(A)** climatic (red) and soil correlation analysis between predicted acid phosphatase activity (AP) and all numerical abiotic factors. **(B)** Was based on the partial correlation analysis between AP and NPP and soil factors. Pixels with $p > 0.1$ are excluded. The soil factors are: soil pH (SoilpH), soil organic carbon content (OC), soil total nitrogen content (TN), soil total phosphorus content (TP), soil clay content (Clay), soil labile organic phosphorus content (Po). The climatic factors are: mean annual temperature (MAT), amplitude of yearly temperature (AMP), and mean annual precipitation (MAP). For example, AP for about 48% of the vegetated land is controlled by soil pH, with partial correlation coefficients between AP and soil pH > 0.2.

The BPN models explained only a small fraction (<15%) of AP (**Supporting Information Appendix S5**) variance over all measurements for both temperate sites outsides Europe (*Data set C*) and tropical regions (*Data set D*) (**Figures S12A–D**). The low prediction capacity for these sites may be partly attributed to the low accuracy of extracted gridded TN from ISRIC-WISE

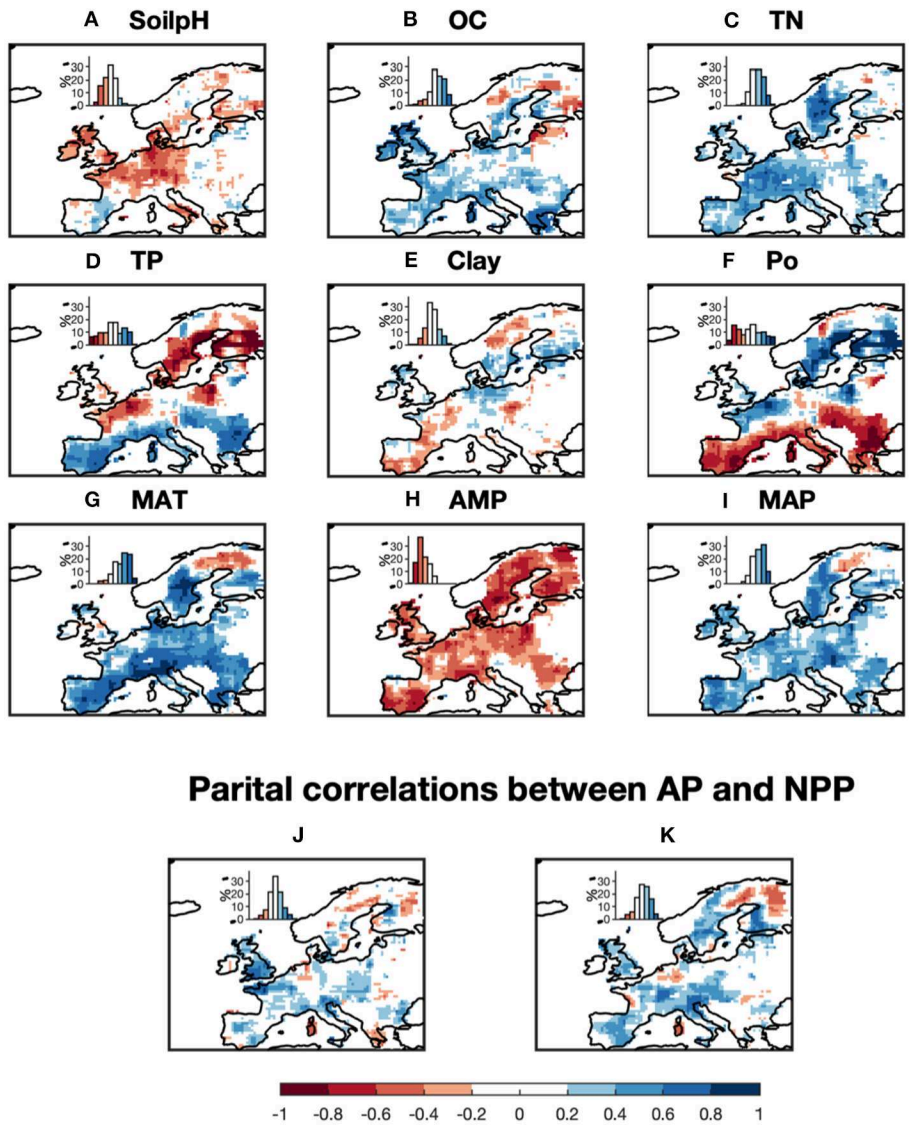


FIGURE 6 | Partial correlations between the median estimates of acid phosphatase activity (AP) and 9 abiotic factors: **(A)** soil pH, **(B)** soil organic carbon, **(C)** soil total nitrogen content, **(D)** soil total phosphorus content, **(E)** soil clay content, **(F)** soil labile organic phosphorus content, **(G)** mean annual temperature, **(H)** amplitude of yearly temperature, and **(I)** mean annual precipitation. **(J,K)** Are the partial correlations between AP and NPP when including climatic variables and excluding climatic variables. Pixels with $p > 0.1$ were excluded. Frequency histograms showing the areal coverages correspond to positive and negative correlations, estimated as proportions of the total study area, are provided for each map (insets).

(**Supporting Information Appendix S1**). Excluding the sites with gap-filled TN, the BPN models explained 34% of AP variability for the rest of temperate sites outside Europe (**Figures S12B,C**). Besides, MAT of many sites fall out of the range of the Europe training sites, which also lead to the poor performance of our BNP models on sites outside Europe. Excluding sites with gap-filled TN and sites with MAT falling out the ranges between 5 and 20°C, the BPN models explained 42% of AP variability for temperate sites outside Europe within the MAT range of the training sites (5°C < MAT < 20°C) (**Figure 3**). This indicates that the BPN can predict spatial gradients of AP with a reasonable accuracy ($R^2 = 0.58$) over the European continent where the training data originate from, and still has some predictive ability at cross validation sites that originate from other temperate regions, mainly USA and China (**Figure 7**).

DISCUSSION

Extrapolation of AP From Site to Regional and Global Scales

This study is a first attempt to extrapolate AP site measurements from Europe to regional and global scales using ML methods. The BPN method successfully reproduced the spatial variance

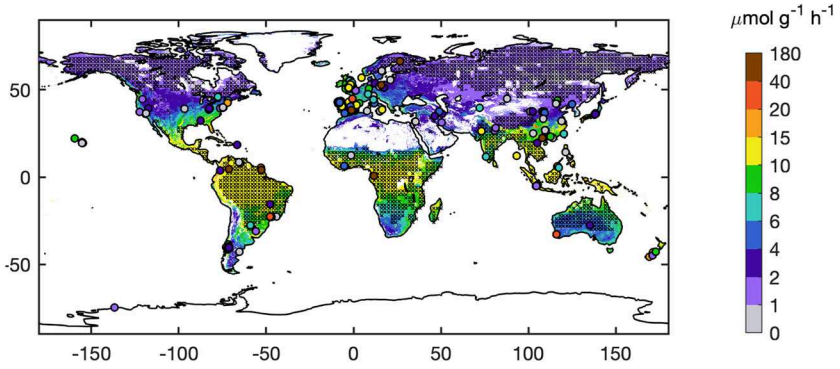


FIGURE 7 | Global pattern of AP from the BPN models (median). "x" indicated area with mean annual temperature (MAT) < 5°C or MAT > 20°C which falling out of the ranges of BPN built based on Europe sites. The observed AP are shown as dots.

of AP across Europe. The performance and reproducibility of BPN models strongly depend on the quantity, quality, and representativeness of the training data set. We attribute the failure in extrapolation to global scale (**Figure 8**) to an insufficient number of representative training AP measurements, especially for tropical soils and climate.

AP activity in tropical regions varies widely from 0 to 80 μmol g^{-1} h^{-1}. The BPN model trained across Europe was not able to capture this large range. A low representativeness of observations in Europe for vegetation types and climatic conditions outside Europe can partly account for this low prediction accuracy. The more homogenous climatic conditions in the tropics than the extra-tropics may increase the importance of soil conditions in accounting for AP variation (**Supporting Information Appendix S6**). Soil properties in the tropics are very heterogeneous and were shown to drive large scale patterns of ecosystem properties (Quesada et al., 2012). More AP measurements in tropical regions would be needed to help us understand the differences between soil types and P-use strategies of ecosystems in this region.

Drivers of the Spatial Patterns of AP in Europe

The assessment of the major factors influencing potential AP help us to understand (1) the controls on AP across Europe and (2) the failure of our model to extrapolate to a global scale. Simple regression analyses were applied in previous studies to investigate the factors influencing AP (Sinsabaugh et al., 2008; Huang et al., 2013; Margalef et al., 2017). Such approaches, however, cannot manage collinearity among predictors nor account for non-linear responses to different variables (Ma et al., 2017). We applied partial correlation diagnostic to the output of the BPN model, which partly overcomes these shortcomings in separating the effect of each predictor on AP for different climatic zones and states of soil nutrients.

Relationships Between Modeled AP and Climatic Variables

Modeled AP is found to be positively correlated with MAT and MAP for most of Europe (**Figures 6G,I**) in line with results from manipulation experiments (soil warming and irrigation). Warming had a positive effect on AP in experimental studies for both temperate grassland soil (0–10 cm depth; Zhou et al., 2013) and Mediterranean shrubland (Sardans et al., 2006). A possible mechanism is that increasing temperature can accelerate plant growth and P acquisition by plants, thus reducing soil P content and stimulating AP production (Sardans et al., 2006). The depletion of soil P can stimulate the production of AP by both plants and soil microbes (Clarholm, 1993; Olander and Vitousek, 2000). Drought experiments in the Mediterranean reported no response (Sardans and Peñuelas, 2005) or reduction in AP activity (Sardans et al., 2006) under conditions of low soil-water content, associated with low water availability and plant growth and ultimately with a reduction in the P demand of plants and microbes. Zhou et al. (2013) reported that increased precipitation significantly increased soil AP in a temperate grassland at depths of 10–20 cm. Measurements based on spatial sampling have also indicated that a higher availability of soil water leads to a larger microbial biomass (Brockett et al., 2012), and hence increased activities of soil extracellular enzymes. Increasing precipitation should also promote substrate diffusion, which brings more soil NO_3-N and NH_4-N to the subsurface and increases AP (Zhou et al., 2013).

Relationships Between Modeled AP and Soil Variables

Nutrient availability, pH, organic C, and clay content are also found to play important roles in the modeled spatial variance of AP from west to east Europe (**Figures 6A–F**). Modeled AP was positively correlated with SOC. The substrate for AP, soil organic P, is tightly linked to SOC (Tipping et al., 2016), but not included in our analysis due to lack of data. Therefore, the correlation between AP and SOC can be interpreted as a control of substrate on AP. This finding was consistent with results from a meta-analysis of soil phosphatase activity and transect measurements (Sinsabaugh et al., 2008; Huang et al., 2013).

TN was also positively correlated with modeled AP for most of Europe (**Figure 6C**). This finding is consistent with results by Margalef et al. (2017) and a meta-analysis of N-fertilization experiments by Marklein and Houlton (2012). Higher soil N

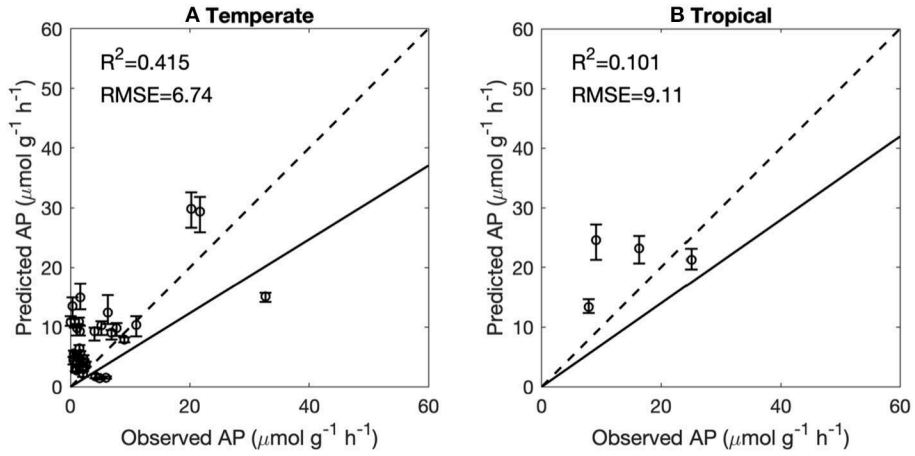


FIGURE 8 | Performance of the back-propagation networks on temperate sites (excluding Europe) **(A)** and tropical sites **(B)**. Temperate and tropical sites with original TN and 5°C < MAT < 20°C were reproduced using BPNs and were then compared to observations. The solid lines indicate the regression lines between predicted vs. observed AP. The dashed lines indicate the 1:1 lines. The error bars indicate the 10 and 90% quantiles of predicted AP.

availability is supposed to strongly increase AP activity in both plant roots and bulk soils across a wide range of ecosystems (Marklein and Houlton, 2012).

In contrast to the monotonic behavior of AP with TN and climatic variables, soil TP and Po were non-monotonically correlated with modeled AP for Europe (**Figures 6D,F**). AP in wet regions (P/PET > 0.7) was negatively correlated with TP but positively correlated with Po. This could indicate that biological mineralization of Po via AP is an essential strategy to maintain the bioavailable P pool under humid conditions (Izquierdo et al., 2013). AP in aridity area (P/PET < 0.7) is positively correlated with TP but negatively correlated with Po (**Figure 6**). Compared with humid soils, arid soils typically contain lower SOM, thus the importance of minerals as source of P for biota is suggested to become more important (Lajtha and Schlesinger, 1988; Cross and Schlesinger, 1995; Delgado-Baquerizo et al., 2013). Our results show a lower investment into AP under semi-dry to dry conditions providing support for such shifts in nutritional strategies on large spatial scales (e.g., Feng et al., 2016).

Soil pH was negatively correlated with modeled AP across Europe (**Figure 6A**), in agreement with the AP measurements based on acidic soils (Dick et al., 2000). Soils with high pHs are generally associated with fewer amino acid functional groups that are essential for catalysis (Dick et al., 2000). Clay content had a non-linear but weak relationship with modeled AP across Europe: increasing clay content increased AP in soils with low clay content (<10%) but decreased AP in soils with high clay content (>10%) (**Figure 6E**). Clay particles are characterized by many reactive binding sites for P and AP leading to the deactivation of the latter (Rao et al., 2000). Thus, other strategies than AP, like extraction of chelating agents and acids, might be advantageous in clay rich soils.

Relationships Between Modeled AP and Net Primary Production

The relationships between abiotic (climate, soils) factors and AP generally are in qualitative agreement with the results from most experimental studies, which provided evidence of a positive association between AP and environmental resources and conditions (water, temperature, and nutrients). Our analysis showed that MODIS-NPP, which covaries with climate, is also statistically associated with modeled AP (**Figure 6K**). Generally, our results are consistent with the mechanism that adequate resources (temperature, water and nutrients) can sustain a higher production and P demand for plants and microbes, thus larger AP activity to increase mineralization rates (Margalef et al., 2017). However, the co-relationship between AP and NPP is weaker than that with climatic variables and TN (**Figures 6G–I,K**), which indicating that the AP pattern is not only determined by P demand by biota but also by climate and soil nutrients.

Relationships Between Modeled AP and Vegetation and Soil Types

Modeled AP differed among the soil types, with higher activities in Oxisols and Ultisols (**Figure S8**) in line with earlier findings (Acosta-Martinez et al., 2007). These soils are commonly poor in P and thus a high recycling efficiency of organic P is an adaption to these conditions (Vance et al., 2003). We found that AP was higher in evergreen broadleaf forest than deciduous forests (**Figure S8**), consistent with the finding by Redel et al. (2008) of higher litter accumulation and lower P cycling in deciduous forests. Margalef et al. (2017) found that soil phosphatase was significantly higher in angiosperm than gymnosperm forests, likely because gymnosperm tend to grow on soils which have a lower biological activity as an adaptation to harsh conditions and thus have the associated disadvantage of low N-use efficiency.

Implication and Uncertainties

The BPN models successfully extrapolated the site AP measurements on the Europe region. However, there are some inevitable uncertainties due to the gap-filling of input predictors which were not reported in the original literature.

The extracted values from gridded datasets are biased to some extent, e.g., due to the coarse spatial resolution of predictor maps as some variables like TN, TP, and Clay vary strongly within very short (cm-m) distance (Lechowicz and Bell, 1991; Iqbal et al., 2005). As a consequence, the covariation between observed and extracted values of TN, TP, and Clay are not very strong (**Figure S1**).

However, non-linear BPN models are considered to be more sensitive to the orders of predictor values rather than linear variation among predictor values. This advantage of BPN compared to linear models reduces the influence of biases in gap-filling on AP predictions. The spearman rank correlations between observed and extracted values for Europe sites are found to be significant for all variables except for Clay (**Figure S1**). Thus, the data quality seems in general sufficient for the BPN model. Still, the BPN model might underestimate the driving effect of Clay on AP pattern due to the large biases. High-quality gridded datasets especially for soil clay content are required, to reduce the uncertainty in the extrapolations based on machine learning methods.

CONCLUSIONS

This study is the first quantification of AP on continental scale by extrapolating the site observations of AP to a spatial explicit map of AP which can be used to benchmark land surface models for temperate regions. This study also provided the comprehensive overview of the factors driving the spatial variations of phosphatase activity on continental scale. The extent to which plants can upregulate P recycling under future climate changes could be limited by climatic and soil factors which are currently not accounted for in global models (e.g., Goll et al., 2012). Further, we identified the need for more measurements in tropical, arid, boreal, or alpine ecosystems, so as to be able to extend our approach to global scale.

DATA AVAILABILITY STATEMENT

The datasets of predictors used in this study are publicly available and can be accessed by the links we provided in **Supporting Information**. The code for BPN and RT can be accessed on Github: https://github.com/yansun8/AP_extropolation. Other data are available upon request from the authors.

AUTHOR CONTRIBUTIONS

YS, DG, and SP designed the research. YS and DG wrote the paper. YS did the extrapolation of AP and analyzed the data. OM and DA collected the dataset of site measurements. PC, JS, and JP helped revise this paper.

FUNDING

This work was supported by the European Research Council Synergy project SyG-2013-610028 IMBALANCE-P.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdata.2019.00051/full#supplementary-material

REFERENCES

Acosta-Martinez, V., Cruz, L., Sotomayor-Ramirez, D., and Pérez-Alegría, L. (2007). Enzyme activities as affected by soil properties and land use in a tropical watershed. *Appl. Soil Ecol.* 35, 35–45. doi: 10.1016/j.apsoil.2006.05.012

Allen, L. H. (1974). Model of light penetration into a wide-row crop 1. *Agron. J.* 66, 41–47. doi: 10.2134/agronj1974.00021962006600010011x

Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., et al. (2019). Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. *Geoderma* 355:113912. doi: 10.1016/j.geoderma.2019.113912

Ballabio, C., Panagos, P., and Monatanarella, L. (2016). Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma* 261, 110–123. doi: 10.1016/j.geoderma.2015.07.006

Brockett, B. F. T., Prescott, C. E., and Grayston, S. J. (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biol. Biochem.* 44, 9–20. doi: 10.1016/j.soilbio.2011.09.003

Caldwell, B. A. (2005). Enzyme activities as a component of soil biodiversity: a review. *Pedobiologia* 49, 637–644. doi: 10.1016/j.pedobi.2005.06.003

Cawley, G. C. (2006). "Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings* (Vancouver, BC), 1661–1668. doi: 10.1109/IJCNN.2006.246634

Clarholm, M. (1993). Microbial biomass P, labile P, and acid phosphatase activity in the humus layer of a spruce forest, after repeated additions of fertilizers. *Biol. Fertil. Soils* 16, 287–292. doi: 10.1007/BF00369306

Cross, A. F., and Schlesinger, W. H. (1995). A literature review and evaluation of the Hedley fractionation: applications to the biogeochemical cycle of soil phosphorus in natural ecosystems. *Geoderma* 64, 197–214. doi: 10.1016/0016-7061(94)00023-4

Delgado-Baquerizo, M., Maestre, F. T., Gallardo, A., Bowker, M. A., Wallenstein, M. D., Quero, J. L., et al. (2013). Decoupling of soil nutrient cycles as a function of aridity in global drylands. *Nature* 502, 672–676. doi: 10.1038/nature12670

Dick, W. A., Cheng, L., and Wang, P. (2000). Soil acid and alkaline phosphatase activity as pH adjustment indicators. *Soil Biol. Biochem.* 32, 1915–1919. doi: 10.1016/S0038-0717(00)00166-8

Eivazi, F., and Tabatabai, M. A. (1977). Phosphatases in soils. *Soil Biol. Biochem.* 9, 167–172. doi: 10.1016/0038-0717(77)90070-0

Ellsworth, D. S., Anderson, I. C., Crous, K. Y., Cooke, J., Drake, J. E., Gherlenda, A. N., et al. (2017). Elevated CO_2 does not increase eucalypt forest productivity on a low-phosphorus soil. *Nat. Clim. Chang.* 7:279. doi: 10.1038/nclimate3235

Feng, J., Turner, B. L., Lü, X., Chen, Z., Wei, K., Tian, J., et al. (2016). Phosphorus transformations along a large-scale climosequence in arid and semiarid grasslands of northern China. *Glob. Biogeochem. Cycles* 30, 1264–1275. doi: 10.1002/2015GB005331

Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. doi: 10.1002/joc.5086

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., et al. (2010). MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Rem. Sens. Environ.* 114, 168–182. doi: 10.1016/j.rse.2009.08.016

Gill, A. L., and Finzi, A. C. (2016). Belowground carbon flux links biogeochemical cycles and resource-use efficiency at the global scale. *Ecol. Lett.* 19, 1419–1428. doi: 10.1111/ele.12690

Goll, D. S., Brovkin, V., Parida, B. R., Reick, C. H., Kattge, J., Reich, P. B., et al. (2012). Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling. *Biogeosciences* 9, 3547–3569. doi: 10.5194/bg-9-3547-2012

Golubiewski, N. E. (2006). Urbanization increases grassland carbon pools: effects of landscaping in Colorado's front range. *Ecol. Appl.* 16, 555–571. doi: 10.1890/1051-0761(2006)016[0555:uigcpe]2.0.co;2

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* 12:e0169748. doi: 10.1371/journal.pone.0169748

Hogan, E. J., Minnullina, G., Smith, R. I., and Crittenden, P. D. (2010). Effects of nitrogen enrichment on phosphatase activity and nitrogen: phosphorus relationships in Cladonia portentosa. *New Phytol.* 186, 911–925. doi: 10.1111/j.1469-8137.2010.03222.x

Huang, W., Liu, J., Wang, Y. P., Zhou, G., Han, T., and Li, Y. (2013). Increasing phosphorus limitation along three successional forests in southern China. *Plant Soil* 364, 181–191. doi: 10.1007/s11104-012-1355-8

Huang, W., Liu, J., Zhou, G., Zhang, D., and Deng, Q. (2011). Effects of precipitation on soil acid phosphatase activity in three successional forests in southern China. *Biogeosciences* 8, 1901–1910. doi: 10.5194/bg-8-1901-2011

Iqbal, J., Thomasson, J. A., Jenkins, J. N., Owens, P. R., and Whisler, F. D. (2005). Spatial variability analysis of soil physical properties of alluvial soils. *Soil Sci. Soc. Am. J.* 69, 1338–1350. doi: 10.2136/sssaj2004.0154

Izquierdo, J. E., Houlton, B. Z., and Huysen, T. L. V. (2013). Evidence for progressive phosphorus limitation over long-term ecosystem development: examination of a biogeochemical paradigm. *Plant Soil* 367, 135–147. doi: 10.1007/s11104-013-1683-3

Keenan, T. F., Davidson, E., Moffat, A. M., Munger, W., and Richardson, A. D. (2012). Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Glob. Change Biol.* 18, 2555–2569. doi: 10.1111/j.1365-2486.2012.02684.x

Kelley, H. J. (1960). Gradient theory of optimal flight paths. *ARS J.* 30, 947–954. doi: 10.2514/8.5282

Kitayama, K. (2013). The activities of soil and root acid phosphatase in the nine tropical rain forests that differ in phosphorus availability on Mount Kinabalu, Borneo. *Plant Soil* 367, 215–224. doi: 10.1007/s11104-013-1624-1

Krämer, S., and Green, D. M. (2000). Acid and alkaline phosphatase dynamics and their relationship to soil microclimate in a semiarid woodland. *Soil Biol. Biochem.* 32, 179–188. doi: 10.1016/S0038-0717(99)00140-6

Lajtha, K., and Schlesinger, W. H. (1988). The biogeochemistry of phosphorus cycling and phosphorus availability along a desert soil chronosequence. *Ecology* 69, 24–39. doi: 10.2307/1943157

Lechowicz, M. J., and Bell, G. (1991). The ecology and genetics of fitness in forest plants. II. Microspatial heterogeneity of the edaphic environment. *J. Ecol.* 79, 687–696. doi: 10.2307/2260661

Ma, Z., Liu, H., Mi, Z., Zhang, Z., Wang, Y., Xu, W., et al. (2017). Climate warming reduces the temporal stability of plant community biomass production. *Nat. Commun.* 8, 1–7. doi: 10.1038/ncomms15378

Maistry, P. M., Muasya, A. M., Valentine, A. J., and Chimphango, S. B. M. (2015). Increasing nitrogen supply stimulates phosphorus acquisition mechanisms in the fynbos species *Aspalathus linearis*. *Funct. Plant Biol.* 42, 52–62. doi: 10.1071/FP14100

Margalef, O., Sardans, J., Fernández-Martínez, M., Molowny-Horas, R., Janssens, I. A., Ciais, P., et al. (2017). Global patterns of phosphatase activity in natural soils. *Sci. Rep.* 7, 1–13. doi: 10.1038/s41598-017-01418-8

Marklein, A. R., and Houlton, B. Z. (2012). Nitrogen inputs accelerate phosphorus cycling rates across a wide variety of terrestrial ecosystems. *New Phytol.* 193, 696–704. doi: 10.1111/j.1469-8137.2011.03967.x

Mason, C., Twomey, J., Wright, D., and Whitman, L. (2017). Predicting engineering student attrition risk using a probabilistic neural network and comparing results with a backpropagation neural network and logistic regression. *Res. High Educ.* 59, 382–400. doi: 10.1007/s11162-017-9473-z

Nannipieri, P., Giagnoni, L., Landi, L., and Renella, G. (2011). "Role of phosphatase enzymes in soil," in *Phosphorus in Action*, eds E. Bünemann, A. Oberson, and E. Frossard (Berlin; Heidelberg: Springer), 215–243. doi: 10.1007/978-3-642-15271-9_9

Olander, L. P., and Vitousek, P. M. (2000). Regulation of soil phosphatase and chitinase activity by N and P availability. *Biogeochemistry* 49, 175–191. doi: 10.1023/A:1006316117817

Papale, D., and Valentini, R. (2003). A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. *Glob. Change Biol.* 9, 525–535. doi: 10.1046/j.1365-2486.2003.00609.x

Quesada, C. A., Phillips, O. L., Schwarz, M., Czimczik, C. I., Baker, T. R., Patiño, S., et al. (2012). Basin-wide variations in Amazon forest structure and function are mediated by both soils and climate. *Biogeosciences* 9, 2203–2246. doi: 10.5194/bg-9-2203-2012

Rao, M. A., Violante, A., and Gianfreda, L. (2000). Interaction of acid phosphatase with clays, organic molecules and organo-mineral complexes: Kinetics and stability. *Soil Biol. Biochem.* 32, 1007–1014. doi: 10.1016/S0038-0717(00)00010-9

Redel, Y., Rubio, R., Godoy, R., and Borie, F. (2008). Phosphorus fractions and phosphatase activity in an Andisol under different forest ecosystems. *Geoderma* 145, 216–221. doi: 10.1016/j.geoderma.2008.03.007

Reed, S. C., Yang, X., and Thornton, P. E. (2015). Incorporating phosphorus cycling into global modeling efforts: a worthwhile, tractable endeavor. *New Phytol.* 208, 324–329. doi: 10.1111/nph.13521

Riedmiller, M., and Braun, H. (1993). "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *Proceedings of the IJCNN* (San Francisco, CA: IEEE Press), 586–591.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* 54, 547–560. doi: 10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2

Sardans, J., and Peñuelas, J. (2005). Drought decreases soil enzyme activity in a Mediterranean. *Soil Biol. Biochem.* 37, 455–461. doi: 10.1016/j.soilbio.2004.08.004

Sardans, J., Peñuelas, J., and Estiarte, M. (2006). Warming and drought alter soil phosphatase activity and soil P availability in a Mediterranean shrubland. *Plant Soil* 289, 227–238. doi: 10.1007/s11104-006-9131-2

Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H. (2014). A global soil data set for earth system modeling. *J. Adv. Model. Earth. Syst.* 6, 249–263. doi: 10.1002/2013MS000293

Sinsabaugh, R. L., Lauber, C. L., Weintraub, M. N., Ahmed, B., Allison, S. D., Crenshaw, C., et al. (2008). Stoichiometry of soil enzyme activity at global scale. *Ecol. Lett.* 11, 1252–1264. doi: 10.1111/j.1461-0248.2008.01245.x

Sun, Y., Peng, S., Goll, D. S., Ciais, P., Guenet, B., Guimberteau, M., et al. (2017). Diagnosing phosphorus limitations in natural terrestrial ecosystems in carbon cycle models. *Earth's Fut.* 5, 730–749. doi: 10.1002/2016EF000472

Tipping, E., Somerville, C. J., and Luster, J. (2016). The C: N: P: S stoichiometry of soil organic matter. *Biogeochemistry* 130, 117–131. doi: 10.1007/s10533-016-0247-z

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., et al. (2006). Evaluation of MODIS NPP and GPP products across multiple biomes. *Rem. Sens. Environ.* 102, 282–292. doi: 10.1016/j.rse.2006.02.017

Vance, C. P., Uhde-Stone, C., and Allan, D. L. (2003). Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. *New Phytol.* 157, 423–447. doi: 10.1046/j.1469-8137.2003.00695.x

Vitousek, P. M., Porder, S., Houlton, B. Z., and Chadwick, O. (2010). Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen–phosphorus interactions. *Ecol. Appl.* 20, 5–15. doi: 10.1890/08-0127.1

Walker, T., and Syers, J. (1976). The fate of phosphorus during pedogenesis. *Geoderma* 15, 1–19. doi: 10.1016/0016-7061(76)90066-5

Wang, R., Balkanski, Y., Boucher, O., Ciais, P., Peñuelas, J., and Tao, S. (2015). Significant contribution of combustion-related emissions to the atmospheric phosphorus budget. *Nat. Geosci*. 8:48. doi: 10.1038/ngeo2324

Were, K., Bui, D. T., Dick, Ø. B., and Singh, B. R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecol. Indic.* 52, 394–403. doi: 10.1016/j.ecolind.2014.12.028

White, P. J., and Hammond, J. P. (2008). "Phosphorus nutrition of terrestrial plants," in *The Ecophysiology of Plant-Phosphorus Interactions*, eds P. J. White and J. P. Hammond (Dordrecht: Springer), 51–81. doi: 10.1007/978-1-4020-8435-5_4

Wiesmeier, M., Barthold, F., Blank, B., and Kögel-Knabner, I. (2011). Digital mapping of soil organic matter stocks using random forest modeling in a semi-arid steppe ecosystem. *Plant Soil* 340, 7–24, 2011. doi: 10.1007/s11104-010-0425-z

Yang, X., Post, W. M., Thornton, P. E., and Jain, A. (2013). The distribution of soil phosphorus for global biogeochemical modeling. *Biogeosciences* 10, 2525–2537. doi: 10.5194/bg-10-2525-2013

Yang, X., Thornton, P. E., Ricciuto, D. M., and Post, W. M. (2014). The role of phosphorus dynamics in tropical forests–a modeling study using CLM-CNP. *Biogeosciences* 11, 1667–1681. doi: 10.5194/bg-11-1667-2014

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

Zhou, X., Chen, C., Wang, Y., Xu, Z., Han, H., Li, L., et al. (2013). Warming and increased precipitation have differential effects on soil extracellular enzyme activities in a temperate grassland. *Sci. Total. Environ*. 444, 552–558. doi: 10.1016/j.scitotenv.2012.12.023

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers

Supplementary Material

S1. Datasets sources for 12 predictors used in this study

We used the best fitted data-driven metrics for data gap-filling for input predictors and AP extrapolation for whole Europe region based on the comparison of rank correlation (ρ) between observed and extracted values (Table 1; Figrue S1). SoilpH, TN, TP and Clay are gathered by using Land Use and Coverage Area Frame Survey (LUCAS) topsoil (0~20cm) database with spatial resolution of 500m, which is one of the world's largest and most comprehensive, harmonized continental-scale soil database generated by using hybrid approaches (Ballabio et al., 2016, 2019; https://esdac.jrc.ec.europa.eu/resource-type/datasets). OC is gathered by using Soilgrids data with spatial resolution of 250m which contains predictions of OC at seven standard depths: 0 cm, 5 cm, 15 cm, 30 cm, 60 cm, 100 cm and 200 cm (Hengl et al., 2017; https://files.isric.org/soilgrids/data/recent/). We first calculated the weighted average of OC (wt %) within 0~15 cm soil depth as below, according to Hengl et al. (2017):

$$OC = \frac{1}{15} \cdot \frac{1}{2} \sum_{k=1}^{N-1} (D_{k+1} - D_k)(OC_k + OC_{k+1})$$

where N is the number of depths, D_k is the k-th depth and OC_k is the OC at depth D_k.

Climatic information (MAT, AMP and MAP) was obtained from WorldClim datasets (version 2.0; http://worldclim.org/version2), which was produced by interpolation from weather stations combined with satellite datasets (Fick & Hijmans, 2017). We used average NPP across 2000-2013 from MODIS (MOD17A3) (Running et al., 2004; Zhao et al., 2005; Turner et al., 2006; http://files.ntsg.umt.edu/data/NTSG_Products/MOD17/MOD17A3). SoilType was extracted from global map of soil suborders derived by World Soil Information (ISRIC; http://www.isric.org/content/faq-soilgrids). VegType was obtained from the MODIS Land Cover Type product (MOD12Q2; Friedl & Sulla-Menashe, 2015; https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q2). Since there is no available Po dataset, we estimated the Po from global map of soil labile inorganic phosphorus (labile Pi; half degree; Yang et al., 2013), the only available global datasets for the content of soil P forms,

multiplied by ratio of the labile Pi and labile Po. The ratios for each soil suborder are derived from collection of soil P measurements at depth of 50 cm by Yang et al. (2013) (Sun et al., 2017).

For the extrapolation of AP map for Europe on 10km, we used gridded datasets above which was resampled into 10km. For the extrapolation of AP map for global region exclude Europe region on 10km, we used global gridded datasets from ISRIC-WISE Global Dataset of Derived Soil Properties for TN and Clay (Batjes et al., 2015; https://data.isric.org/geonetwork/srv/chi/catalog.search#/metadata/dc7b283a-8f19-45e1-aaed-e9bd515119bc), and used Global Soil Dataset for Earth System Modelling (GSDE) for TP (Shangguan et al., 2014; http://globalchange.bnu.edu.cn/research/soilw). We used the same gridded datasets with that for Europe for other predictors (USDA-soiltype, MODIS-NPP and WorldClim data, see above) (Table S1). Both of those datasets have a spatial resolution of 1km and was resample into 10km before the extrapolation.

S2. Gap filling and pre-processes of observation data sets and predictors

We used total C content instead of sites with soil pH <6 for sites with no OC data, because inorganic C is largely in carbonate forms that do not occur in acid soils (Nelson & Sommers, 1996). The predictors SoilpH, OC, TN, TP and Clay were reported together with AP in the original publications takes proportion of 95%, 28%, 92%, 25% and 23% of 126 European sites *(Data set A*). In other cases, these variables were obtained from gridded databases with different spatial resolutions from 250m~1 km based on geographical coordinates of each measurement site (LUCAS; Ballabio et al., 2016, 2019). Climatic variables (MAT, MAP and AMP) (WorldClim; Fick & Hijmans, 2017), NPP (MOD17), SoilType (ISRIC; Hengl et al., 2017) and VegType (MOD12Q2) were extracted from global databases by using the coordinates of the measurement site.

Environmental data often have skewed distributions, such as soil nutrient concentrations (Blackwood, 1992). Fitting lognormal distributions is commonly used to represent this kind of data in statistical analysis (Blackwood, 1992). We tested the distribution of 296 measurements of AP and of the 10 numerical predictors for all global pixels (Figure S2). The acid phosphatase activity was found to follow a lognormal distribution, so that AP was log-transformed. The content of soil nutrients (OC, TN, TP and Po) are also log-normally distributed. Additionally, NPP, AMP and MAP also follow lognormal distribution. Therefore, all these predictors were also log-transformed.

S3. Optimal number of neurons for BPN

The performances of a BPN is partly determined by the number of neurons in each hidden layer. To decide the optimal number of neurons, we used exhaustive combinations of 1-20 neurons for two hidden layers to evaluate the performance of BPN models with respect to number of neurons. The BPN models were trained 1000 times using different training-testing subsets (85% for training datasets, 15% for test datasets). We found that the lowest root mean square error (RMSE) for test datasets occurs in BPN framework with about 10 neurons for hidden layer 1 and 5 neurons for hidden layer 2 (Figure S3)

S4. Detection of outliers

There is a very low accuracy of 19% of explained variance of AP over all measurements, accompanied by a very high root mean square error (RMSE) of 18.2 μmol g^{-1} h^{-1} (Table 2). 9 outliers (*Site A-I*) with extreme bias (absolute bias >20 μmol g^{-1} h^{-1} and the relative bias >50%, Figure 3) were detected. Those seven outliers are located in central England (*Site A*), Eastern Spain (*Site C, H, I*), Poland (*Site D*), and Western Spain (*Site B, E, F, G*). Those extreme biases are attributed to that BPN is insufficient to reproduce the AP observations for particular sites of which predictors falling in the edge of their distributions. For example, *Site A* is out of the margin of Clay. *Site C* is out of the the margin of Clay, OC and TN. The information of the marginal predictors for each outlier are listed in Table S3.

S5. Compare the performance of the BPN on tropical sites and temperate sites outside Europe

To access if the AP extrapolation for global region is acceptable, we compared the performance of the BPN on tropical sites and temperate sites outside Europe. The predicted AP were generated (1) by reproducing the AP for temperate and tropical sites with complete information using and (2) by extracting from the global pattern of predicted AP according to the coordinates of the measurement sites (Figure S12). Note that TN was detected to be largely affect AP pattern (Figure 8) but have a low accuracy for extracted TN from global gridded dataset ISRIC-WISE compared with the values reported by literature (Figure S1), we only reproduced the sites with original TN for (1).

S6. Understanding the failure to extrapolate AP to tropical region

The relationships between AP and environmental factors reported for tropical regions differed greatly from our model for Europe, which could account for the failure to extrapolate the BPN model trained using European sites to tropical regions. Climates (precipitation and temperature) in tropical rainforest are generally not factors limiting growth or more homogenous than in extra-tropics. Soil

conditions instead of climate are thus considered to be the most important factors in driving biological activity and thus AP.

Highly weathered tropical soil in particular always has large amounts of highly reactive minerals (Vitousek et al., 2010) that can interact with and deactivate AP. Because these enzymes interact with these minerals (Dick & Tabatabai, 1987) and can potentially lead to differences between AP from microbes and plants (Kitayama, 2013) that are absent in Europe. Gridded datasets of soil variables, such as Po, are associated with large uncertainties, which strongly limits BPN extrapolation for tropical regions. Besides, soil environment (pH, nutrient contents, texture) can also vary within short distances. The representativeness of samples will thus strongly influence model calibration and may lead to large uncertainties in AP estimates, which cannot be solved by our BPN model. More measurements of AP in tropical regions are needed to help us understand the physiological basis of P-use strategies in tropical forest ecosystems and reduce the uncertainty in estimates of tropical AP estimates.

References

Blackwood, L. G. (1992). The lognormal distribution, environmental data, and radiological monitoring. *Environmental monitoring and assessment*, *21*(3), 193-210.

Nelson, D. W., & Sommers, L. E. (1996). Total carbon, organic carbon, and organic matter. *Methods of soil analysis part 3—chemical methods, (methodsofsoilan3)*, 961-1010.

Table S1 Predictors used in extrapolation of AP for regions outside of Europe.

Predictor Name	Abbreviation	Source	References
Soil organic carbon	OC	Soilgrids	Hengl et al., 2017
Soil pH	SoilpH	ISRIC-WISE	Batjes et al., 2015
Soil total nitrogen	TN		
Soil clay content	Clay		
Soil total phosphorus	TP	GSDE	Shangguan et al., 2014
Soil labile organic phosphorus	Po	Global maps of the soil P contents for different P forms and USDA soil types and ratio of labile inorganic P and labile organic P	Yang et al., 2013; Hengl et al., 2017; Sun et al., 2017
Net primary productivity	NPP	MODIS-NPP (MOD17A3; mean value during 2000-2014)	Running et al., 2004; Zhao et al., 2005; Turner et al., 2006
Mean annual temperature	MAT	WorldClim	Fick and Hijmans, 2017
Yearly temperature amplitude	AMP		
Mean annual precipitation	MAP		
Soil type (categorical variable)	SoilType	Soilgrids & USDA class	Hengl et al., 2017
Vegetation type (categorical variable)	VegType	MODIS	Friedl et al, 2015

Table S2 Regression statistics for Observed vs Predicted acid phosphatase activity derived from (BPN) and (RT) models. Statistics that excluding outliers were also provided. Numbers in parenthesis indicate the number of outliers.

	R^2		RMSE ($\mu mol\ g^{-1}\ h^{-1}$)	
	BPN	RT	BPN	RT
With outliers	0.192	0.096	18.2	19.7
Without outliers	0.578 (9)	0.487 (13)	6.83	7.85

Table S2 Information of the marginal predictors for 9 outliers detected by BPN.

Sites	Predictors falling in the edge of their distributions
A	Clay
B	NPP, Soiltype
C	Clay, pH, TP, OC, TN
D	Labile Po, NPP, OC, TN, Soiltype
E	NPP
F	TP, labile Po
G	NPP
H	TN
I	TP

Figure S1 The comparison of Pearson correlation and rank correlation (ρ) between observed predictors and extracted values from 4 different gridded metrics.

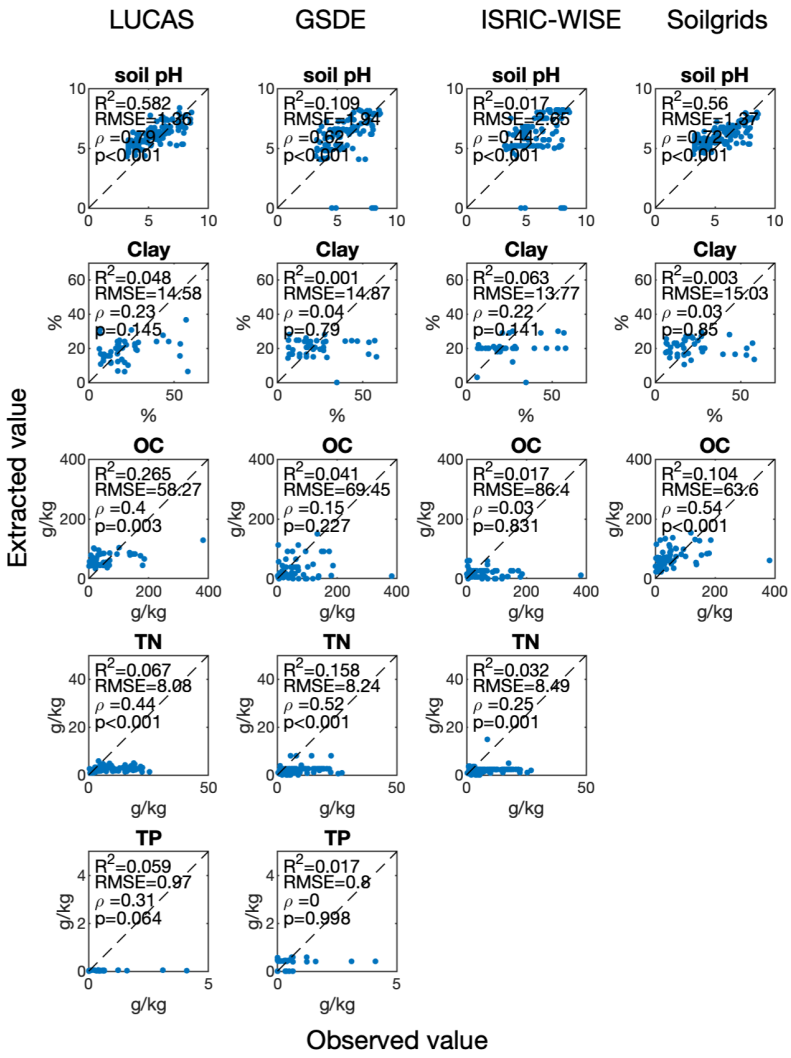


Figure S2 The distribution of 10 numerical predictors used in this analysis for global pixels.

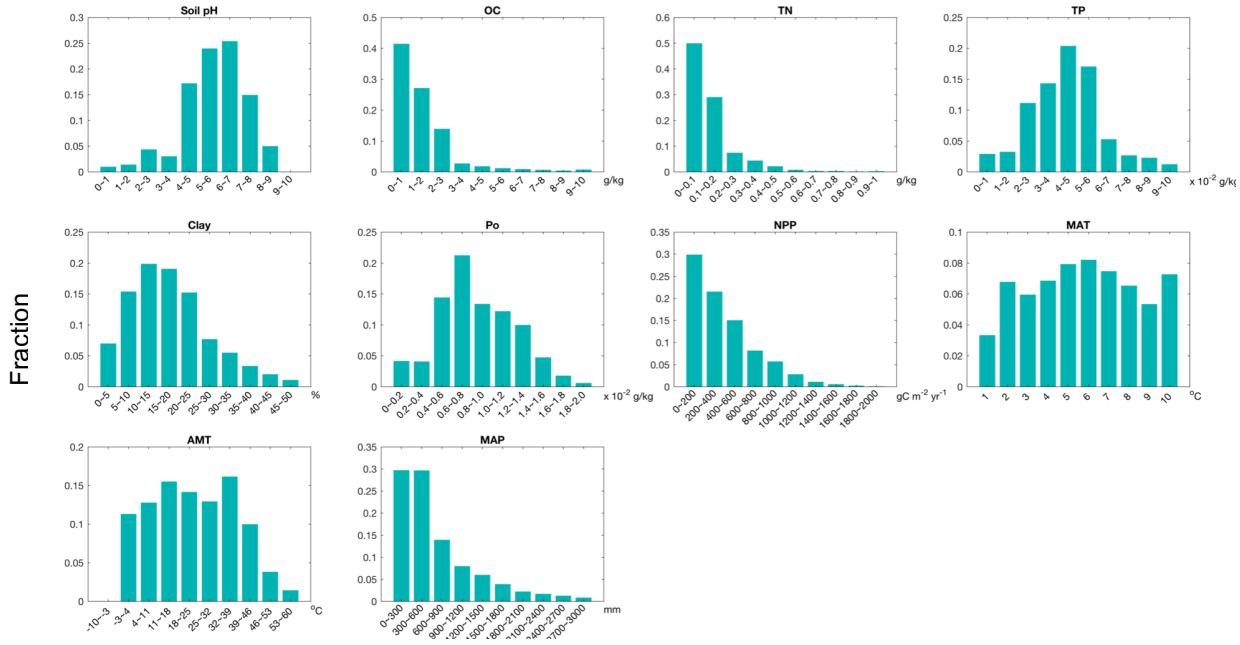


Figure S3 The root mean square error (RMSE; μmol g^{-1} h^{-1}) on train and test datasets by using different number of neurons in two hidden layers for back-propagation models.

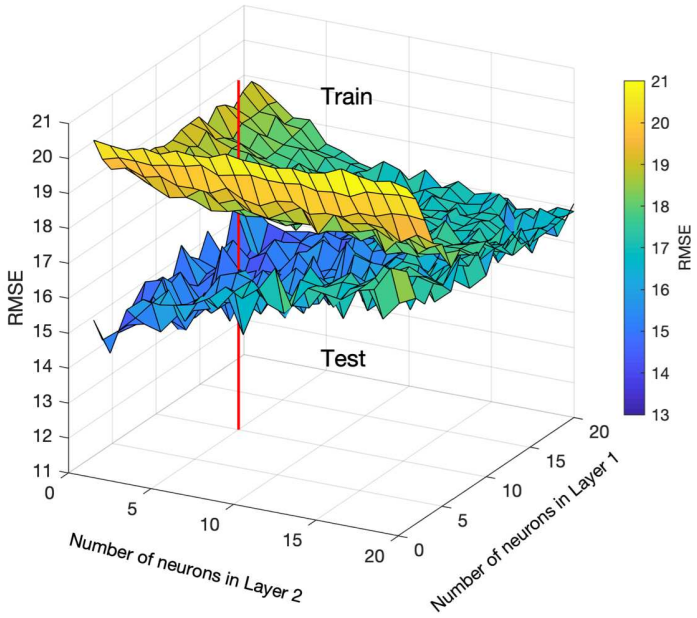


Figure S4 The distribution of 9 sites with very biased estimates of AP by BPN.

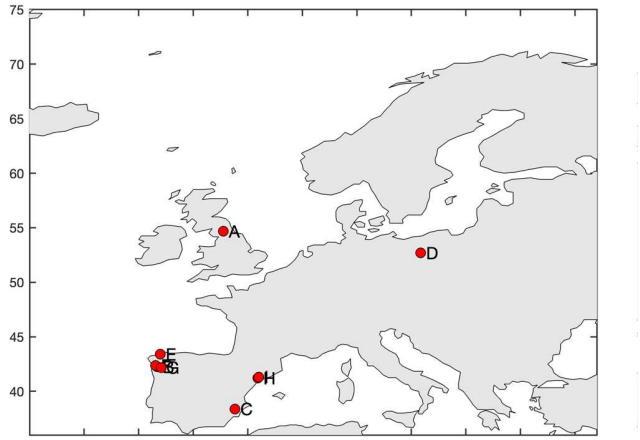


Figure S5 The distribution of 12 predictors for all of 126 European sites. 8 soil types and 11 biome types are considered: Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Broadleaf Forest (DBF), Mixed forest (MF), Open Shrublands (OS), Woody Savannas (WS), Savannas (S), Grasslands (GR). 9 sites with very biased estimates of AP are highlighted by grey dots and alphabet corresponding to Figure S4.

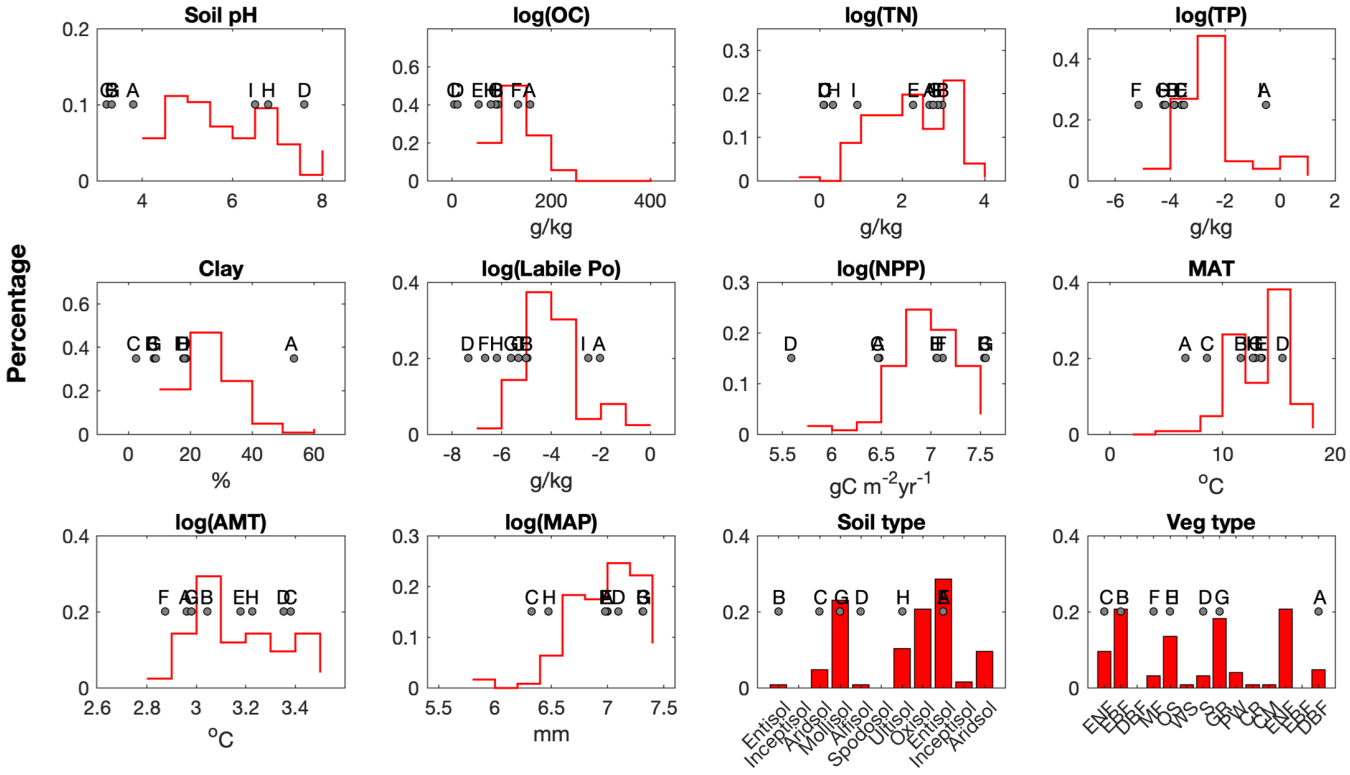


Figure S6 Performance of Regression Tree models (RT) on 126 European sites following hold-one-out framework. (A) shows the Predicted vs Observed acid phosphatase activity (AP) for all sites, while (B) shows the Predicted vs Observed AP with excluding 7 outliers (grey circles). Dashed lines indicate the 1:1 line. Blue lines refer to the regression line between Predicted vs Observed AP. Blue and grey error bars show the 25%~75% quantiles of predicted AP, while orange error bars show the 10%~90% quantiles of predicted AP. (C) shows the distribution of 13 outliers with very biased estimates of AP by RT.

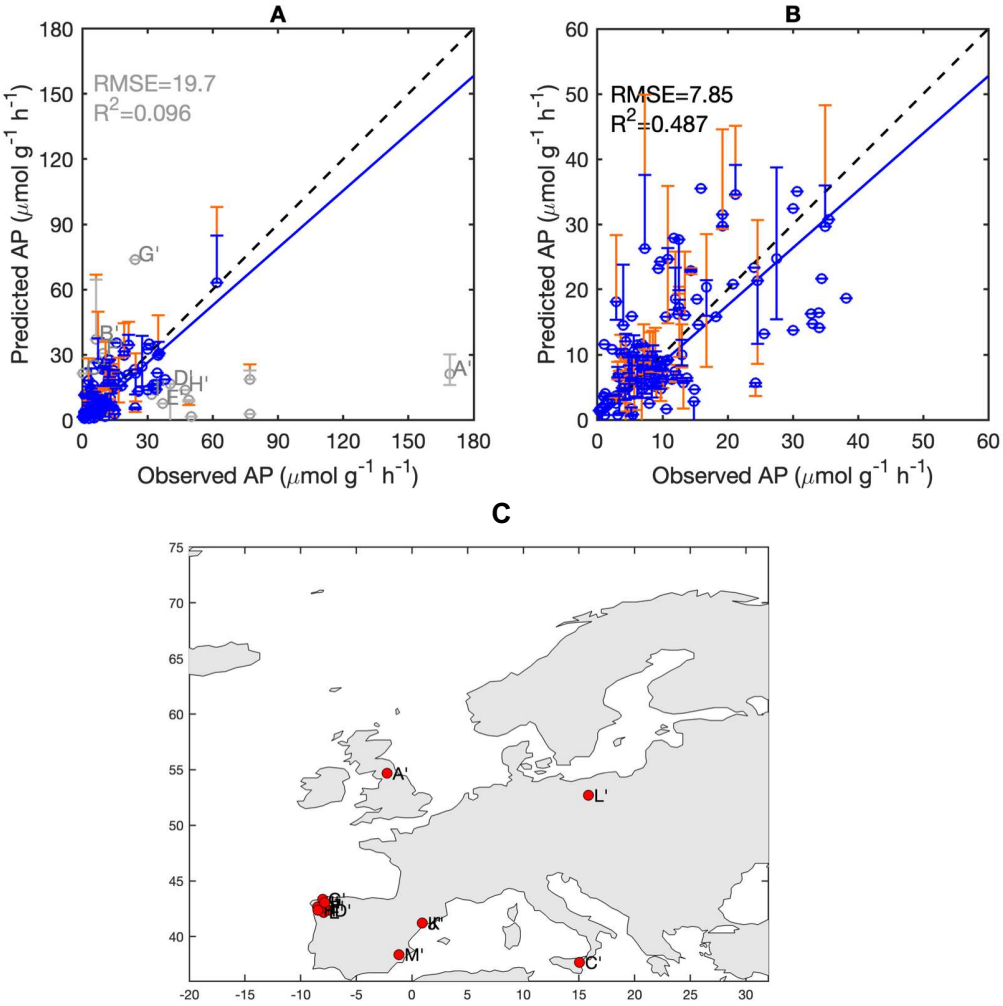


Figure S7 The extrapolated spatial pattern of 25% (A) and 75% (B) quantile estimates of acid phosphatase activity by BPN on Europe.

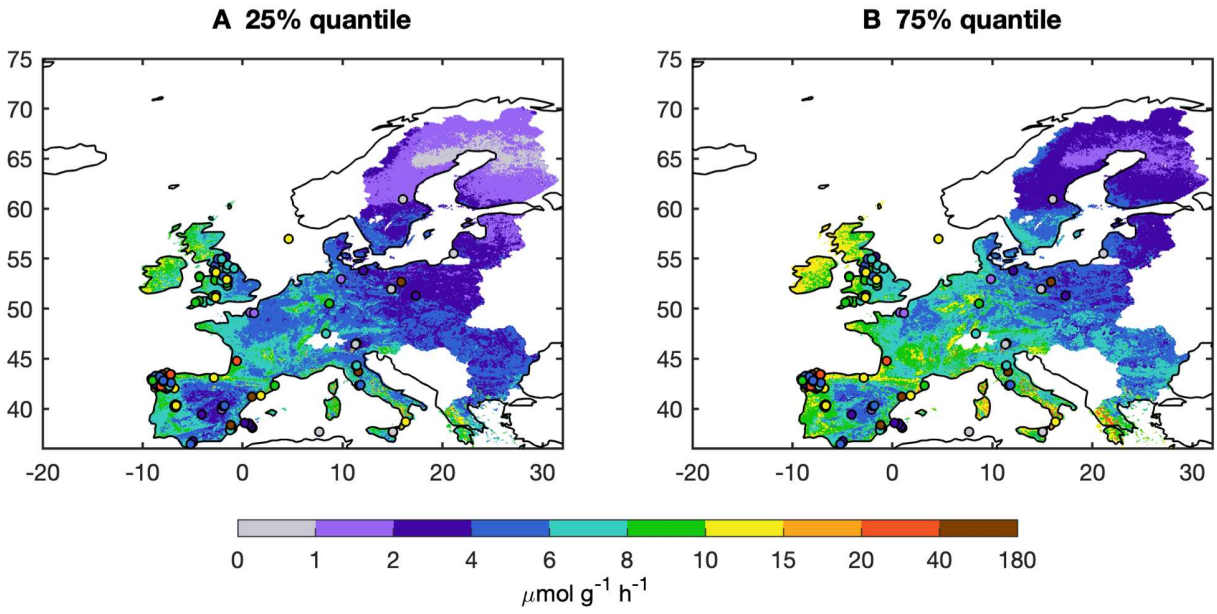


Figure S8 Median estimates of AP for different 8 soil types and 11 biome types: Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Broadleaf Forest (DBF), Mixed forest (MF), Open Shrublands (OS), Woody Savannas (WS), Savannas (S), Grasslands (GR). Error bars show the 25%~75% quantiles of AP for same soil (or biome) type.

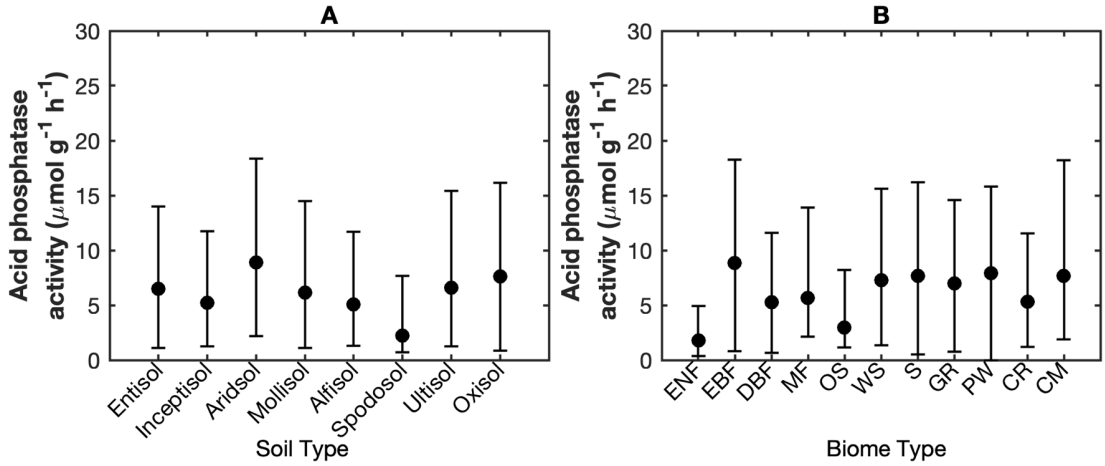


Figure S9 Explanation of climates (MAT, AMT, MAP) to NPP for Europe region. The regression analysis was conducted by using a spatial moving window of 4.5° × 4.5°, of which climates and NPP were firstly resample into half degree by using area-weighted mean methods.

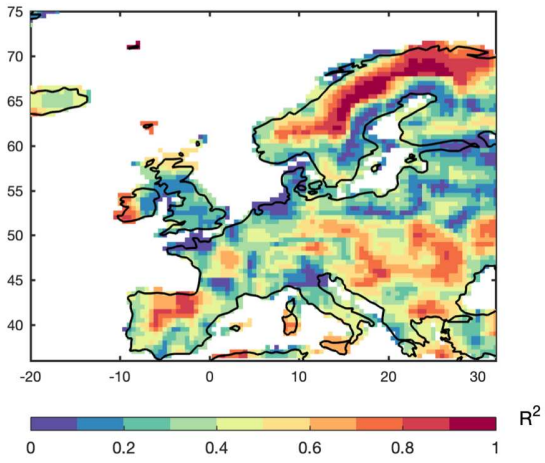


Figure S10 The distribution of 12 predictors for all of 126 European sites (*Data set A*; black) and 87 temperate sites outsides Europe (*Data set C*; blue) and 54 tropical sites (*Data set D*; red). 8 soil types and 11 biome types are considered: Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Broadleaf Forest (DBF), Mixed forest (MF), Open Shrublands (OS), Woody Savannas (WS), Savannas (S), Grasslands (GR).

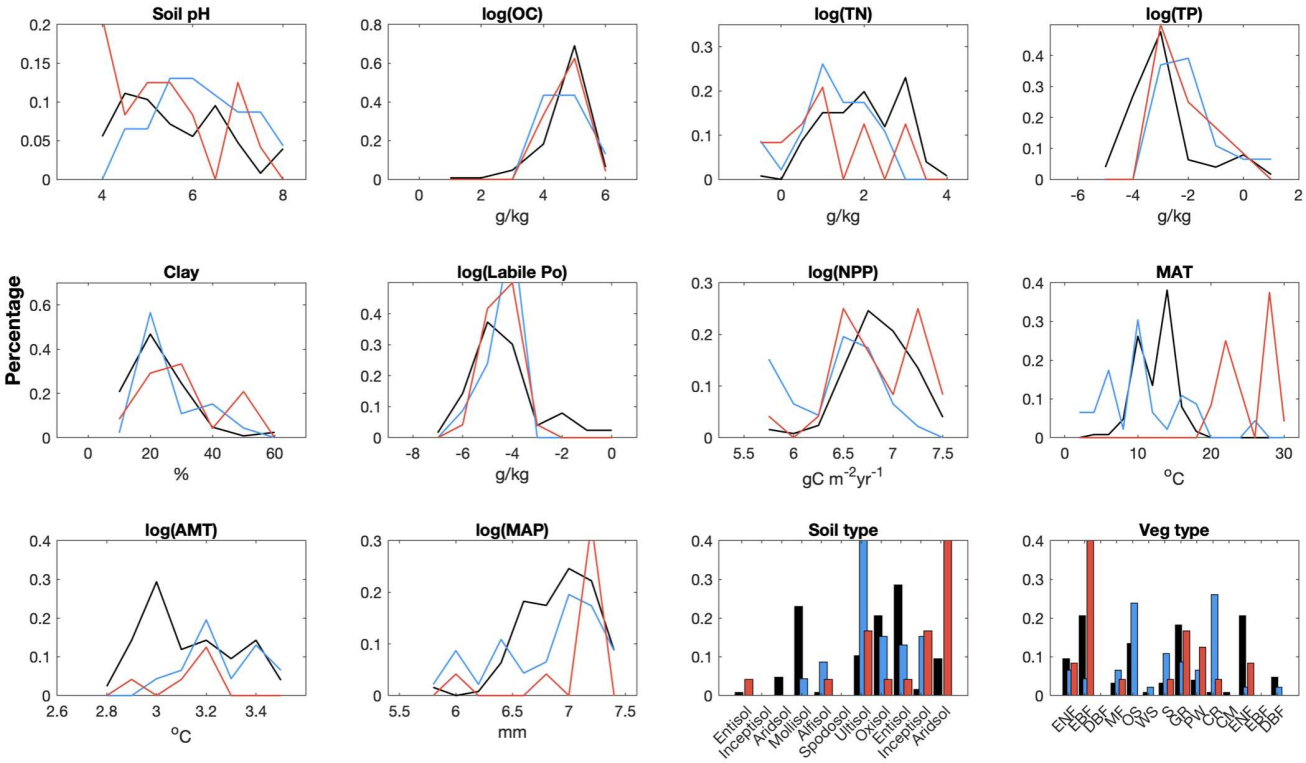


Figure S11 The relationship between partial correlation coefficients of AP to TP (blue) and labile Po (red) and aridity index (P/PET). The aridity index is calculated as the ratio of mean annual precipitation and mean potential evapotranspiration across 1981-2018 that derived from Climate Research Unit (CRU) datasets (CRU TSv4.03). Blue and red lines indicate the regression between partial correlation coefficients and aridity index. Grey shading indicates where the partial correlation relationship is not significant (p>0.1).

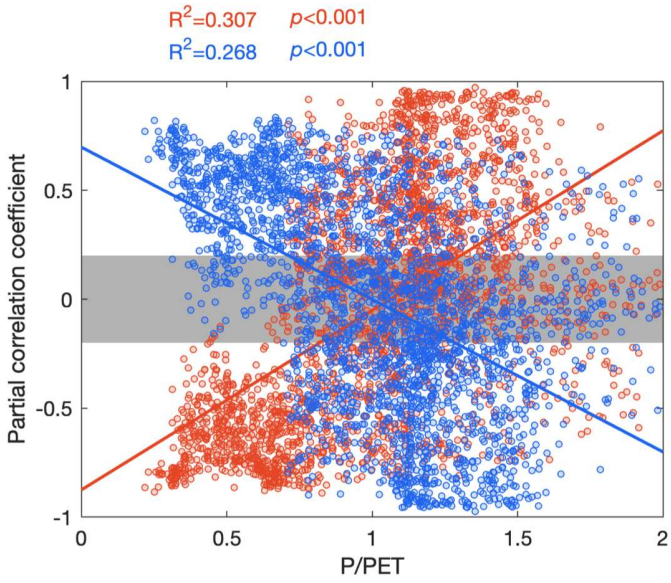
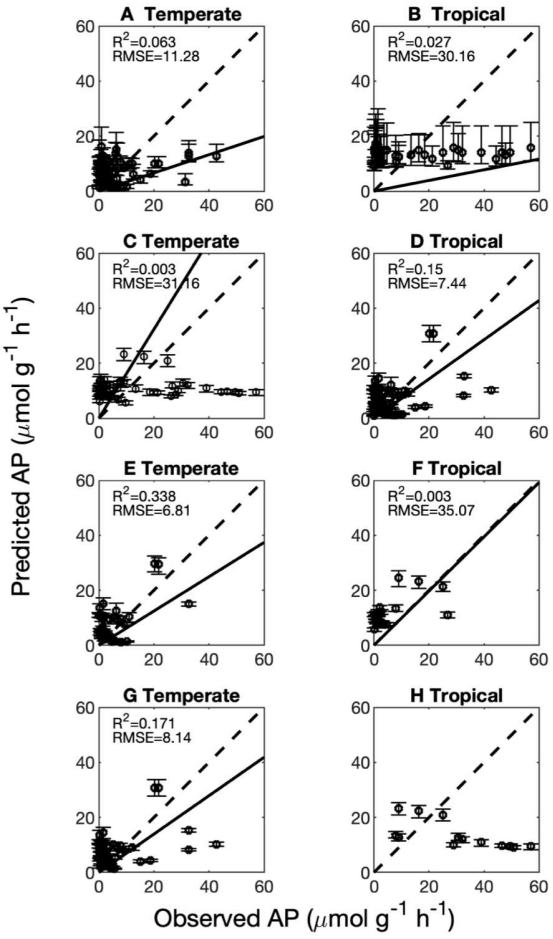


Figure S12 Performance of the back-propagation networks (BPN) on temperate sites (excluding Europe) (A, C, E, G) and tropical sites (B, D, F, H). A and B show the comparisons between observed acid phosphatase activity (AP) and AP extracted from the extrapolated global pattern of AP (Figure 5). C-H show the comparisons between observed AP and predicted AP by using BPN models and predictor information on site scale. C and D show the performance of BPN models on all sites (*Data set C* and *D*). E and F exclude the sites with gap-filled soil total nitrogen (TN), while G and H exclude the sites with mean annual temperature (MAT) <5^{o}C and MAT>20^{o}C. The solid lines



indicate the regression lines between predicted vs observed AP. The dashed lines indicate the 1:1 lines. The error bars indicate the 10% and 90% quantiles of predicted AP.

Chapter 3 Tailored evaluation for a phosphorus enabled land surface model

Summary

A small but increasing number of land surface models (LSMs) have incorporated phosphorus (P) cycles and carbon (C) – nitrogen (N) - P interactions but their evaluation remains very limited. Recent increase in ground-based measurements and ecological datasets, e.g. gridded AP dataset (chapter 2), offer the opportunity to design a tailored model evaluation strategy to assess the quality of predictions of CNP models. This tailored evaluation aiming for model emerging ecosystem properties rather than just fluxes and pools provide more comprehensive understanding for the impact of interactions between C, N, P and water.

In this chapter, I evaluated the performance of the global version of the land surface model ORCHIDEE-CNP (v1.2) which explicitly simulates N and P biogeochemistry in terrestrial ecosystems coupled with carbon, water and energy transfers. I developed a strategy based on remote-sensing, ground-based measurement networks and ecological databases.

Components of the N and P cycle at different levels of aggregation (from local to global) are in good agreement with data-driven estimates. When integrated for the period 1850 to 2017 forced with variable climate, rising CO_2 and land use change, we show that ORCHIDEE-CNP underestimates the land carbon sink in the North Hemisphere (NH) during the recent decades, despite an a priori realistic gross primary productivity (GPP) response to rising CO_2. This result suggests either that other processes than CO_2 fertilization which are omitted in ORCHIDEE-CNP, such as changes in biomass turnover, are predominant drivers of the northern land sink, and/or that the model parameterizations produce too strict emerging nutrient limitations on biomass growth in northern areas. In line with the latter, we identified biases in the simulated large-scale patterns of leaf and soil stoichiometry and plant P use efficiency pointing towards a too severe P limitations towards the poles.

Based on our analysis of ecosystem resource use efficiencies and nutrient cycling, I propose ways to address the model biases by giving priority to better representing processes of soil organic P mineralization and soil inorganic P transformation, followed by refining the biomass production efficiency under increasing atmospheric CO_2, phenology dynamics and canopy light absorption. This chapter was accepted and will be published in the journal *Geoscientific Model Development (GMD): https://doi.org/10.5194/gmd-2020-93*.

Geosci. Model Dev., 14, 1–24, 2021
https://doi.org/10.5194/gmd-14-1-2021

Geoscientific Model Development
Open Access
EGU

Global evaluation of the nutrient-enabled version of the land surface model ORCHIDEE-CNP v1.2 (r5986)

Yan Sun[1], Daniel S. Goll[1,2], Jinfeng Chang[3], Philippe Ciais[1], Betrand Guenet[1,4], Julian Helfenstein[5], Yuanyuan Huang[1,6], Ronny Lauerwald[1,7], Fabienne Maignan[1], Victoria Naipal[1,8], Yilong Wang[1,9], Hui Yang[1], and Haicheng Zhang[1,7]

[1]Laboratoire des Sciences du Climat et de l'Environnement/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif sur Yvette, 91191, France
[2]Department of Geography, University of Augsburg, Augsburg, Germany
[3]Ecosystems Services and Management Program, International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, 2361 Laxenburg, Austria
[4]Laboratoire de Géologie, UMR 8538, Ecole Normale Supérieure, PSL Research University, CNRS, Paris, France
[5]Agroecology and Environment, Agroscope, Reckenholzstrasse 191, 8046 Zurich, Switzerland
[6]CSIRO Oceans and Atmosphere, Aspendale 3195, Australia
[7]Department Geoscience, Environment & Society, Universite libre de Bruxelles, 1050 Brussels, Belgium
[8]Department of Geography, Ludwig-Maximilian University, Munich, Germany
[9]Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

Correspondence: Yan Sun (ysun@lsce.ipsl.fr) and Daniel S. Goll (dsgoll123@gmail.com)

Received: 6 April 2020 – Discussion started: 28 July 2020
Revised: 14 February 2021 – Accepted: 23 February 2021 – Published:

Abstract. The availability of phosphorus (P) and nitrogen (N) constrains the ability of ecosystems to use resources such as light, water and carbon. In turn, nutrients impact the distribution of productivity, ecosystem carbon turnovers and their net exchange of CO_2 with the atmosphere in response to variation of environmental conditions in both space and time. In this study, we evaluated the performance of the global version of the land surface model ORCHIDEE-CNP (v1.2), which explicitly simulates N and P biogeochemistry in terrestrial ecosystems coupled with carbon, water and energy transfers. We used data from remote sensing, ground-based measurement networks and ecological databases. Components of the N and P cycle at different levels of aggregation (from local to global) are in good agreement with data-driven estimates. When integrated for the period 1850 to 2017 forced with variable climate, rising CO_2 and land use change, we show that ORCHIDEE-CNP underestimates the land carbon sink in the Northern Hemisphere (NH) during recent decades despite an a priori realistic gross primary productivity (GPP) response to rising CO_2. This result suggests either that processes other than CO_2 fertilization, which are omitted in ORCHIDEE-CNP such as changes in biomass turnover, are predominant drivers of the northern land sink and/or that the model parameterizations produce emerging nutrient limitations on biomass growth that are too strict in northern areas. In line with the latter, we identified biases in the simulated large-scale patterns of leaf and soil stoichiometry as well as plant P use efficiency, pointing towards P limitations that are too severe towards the poles. Based on our analysis of ecosystem resource use efficiencies and nutrient cycling, we propose ways to address the model biases by giving priority to better representing processes of soil organic P mineralization and soil inorganic P transformation, followed by refining the biomass production efficiency under increasing atmospheric CO_2, phenology dynamics and canopy light absorption.

Published by Copernicus Publications on behalf of the European Geosciences Union.

1 Introduction

Nitrogen (N) and phosphorus (P) are key macronutrients that control metabolic processes and plant growth and constrain ecosystem-level productivity (Elser et al., 2007; Norby et al., 2010; Cleveland et al., 2013). The amount and stability of soil carbon (C) stock are also affected by N and P through their regulating role in the mineralization of litter and soil organic matter (Gärdenäs et al., 2011; Melillo et al., 2011). The availability of N and P is likely to limit future carbon storage under climate change and rising atmospheric CO_2. Empirical stoichiometry observations were applied in the posteriori estimates of future carbon storage from land surface models (LSMs) lacking an explicit simulation of N and P biogeochemistry, which consistently led to an overestimation of future carbon storage in LSMs (Hungate et al., 2003; Wang and Houlton, 2009; Zaehle et al., 2015; Wieder et al., 2015). Nevertheless, this approach has large uncertainties (Penuelas et al., 2013; Sun et al., 2017) and relies on unproven assumptions (Brovkin and Goll, 2015).

An alternative is to directly represent the complex interactions between N, P and carbon in an LSM. Several LSMs incorporated different parameterizations of N interactions (e.g., Thornton et al., 2007; Zaehle et al., 2014), but very few global models have included P interactions. The few models accounting for P limitation in plant growth showed that P availability limits primary productivity and carbon stocks on highly weathered soils in the tropics (Wang et al., 2010; Yang et al., 2014), and one study also suggested that P limitations could also occur in the Northern Hemisphere in the near future (Goll et al., 2012). Model representations of P interactions are highly uncertain since the critical processes are poorly constrained by current observational data. In particular, the desorption of P from soil mineral surfaces and the enhancement of P availability for plants by phosphatase enzymes secreted by plant roots and microbes were identified to be critical but poorly constrained (Fleischer et al., 2019).

Previous studies (Wang et al., 2010; Goll et al., 2012; Yang et al., 2014; Thum et al., 2019) have suggested that the inclusion of the phosphorus cycle improves model performances with regard to reproducing observed C fluxes. But adding new and uncertain P-related processes does not grant an automatic improvement in an LSM in general. First, more (nutrient-related) equations with more uncertain parameters can result in less robust predictions. Second, models ignoring nutrients were often calibrated on available carbon data so that a new model with nutrients inevitably needs a parameter recalibration to reach similar performances as the same model without nutrients. Third, for evaluating a large-scale model resolving both nutrient and carbon biogeochemistry, one needs specific nutrient-related datasets, which are more scarce than classical biomass, productivity and soil carbon data used for benchmarking carbon-only models.

The evaluation for N and P, together with carbon cycling in global LSMs, remains very limited (Wang et al., 2010; Goll et al., 2012), but recent advances in ground-based measurements, ecological datasets and process understanding have made a better evaluation of C, N and P models feasible. The available nutrient datasets have allowed for meta-analyses of site-level nutrient fertilization experiments (e.g., Yuan and Chen, 2015; Wright, 2019), data-driven assimilation schemes to constrain nutrient budgets (Wang et al., 2018), new knowledge about the critical P processes of sorption (Helfenstein et al., 2018, 2020) and phosphatase-mediated mineralization (Sun et al., 2020), global datasets of leaf nutrient content (Butler et al., 2017), and empirical constraints on the CO_2 fertilization effect on land carbon storage (Terrer et al., 2019; Liu et al., 2019). In addition to direct comparison with nutrient datasets, it is also possible to diagnose emerging model responses in terms of ecosystem resource use efficiencies (RUE) and confront them with observations for identifying how ecosystems adjust and optimize nutrient, water, light and carbon resource availabilities (Fernández-Martínez et al., 2014; Hodapp et al., 2019). In particular, modeled N and P use efficiencies can be compared to observation-based estimates at ecosystem scale (Gill and Finzi, 2016) and at biome scale (Wang et al., 2018).

Here we evaluate the global cycles of C, N and P in the nutrient-enabled version of the LSM ORCHIDEE, ORCHIDEE-CNP (v1.2). The model has been previously evaluated for tropical sites (Goll et al., 2017a, 2018) and for coarse-scale global carbon fluxes and stocks using the International Land Model Benchmarking system iLAMB by, e.g., Friedlingstein et al. (2019). The results from this evaluation showed a slightly worse performance for ORCHIDEE-CNP (v.1.2) than the carbon-only version of ORCHIDEE, which has been extensively calibrated (Friedlingstein et al., 2019). In this study, we perform a detailed evaluation of ORCHIDEE-CNP focusing on four ecosystem characteristics that were found to be critical for the response of land C cycling to increasing CO_2 and climate change: (1) vegetation resource use efficiencies, (2) the response of plant productivity to increasing CO_2, (3) ecosystem N and P turnover and openness, and (4) large-scale patterns of ecosystem stoichiometries. Points (1) and (2) control the response of vegetation carbon storage operating on timescales of years to decades, while points (3) and (4) control the carbon storage potential on an ecosystem level, which determines the response on much longer timescales. Further, the implications of including nutrient cycles on the simulated land C cycling are discussed.

2 Modeling

2.1 Model description

ORCHIDEE-CNP simulates the exchange of greenhouse gases (i.e., carbon dioxide, nitrous oxide), water and energy at the land surface and features a detailed representation of

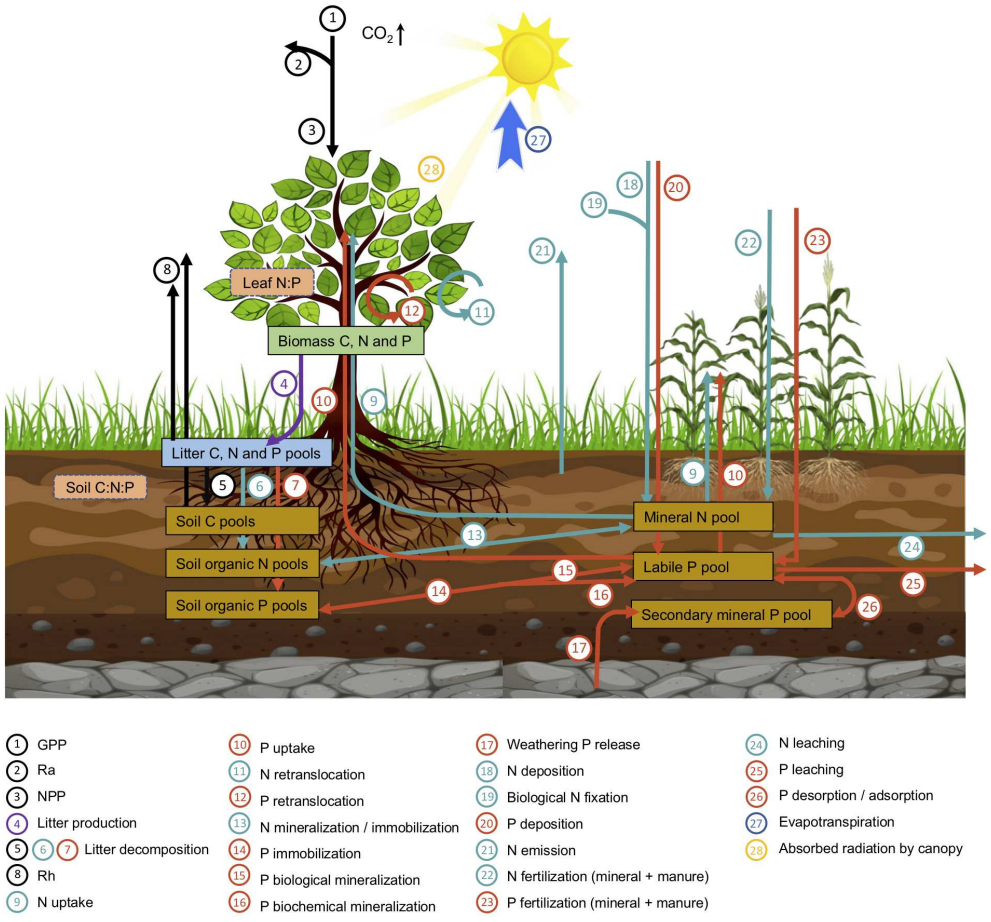


Figure 1. Schematic of C, N and P cycles considered in ORCHIDEE-CNP.

the root uptake of dissolved N and P, the allocation of N and P among tissues, and the N and P turnover in litter and soil organic matter (Goll et al., 2017a, 2018) (Fig. 1). In this study, we present the first global application of the model and an evaluation against global carbon and nutrient datasets. ORCHIDEE-CNP simulates the cycles of C, N and P, which are described in detail elsewhere (Krinner et al., 2005; Zaehle and Friend, 2010; Goll et al., 2014, 2017a, 2018). We give a brief overview here. P enters the ecosystem through release from minerals into the soil solution, whereas N is biologically fixed from an ample reservoir of dinitrogen. Dissolved nutrients are either taken up by vegetation, converted into soil organic matter or absorbed onto soil particles. Losses occur as leaching of dissolved nutrients, gaseous soil N emissions or occlusion of P in secondary minerals. When nutrients are taken up by vegetation they are either stored internally or used to build new plant tissue driven by the availability of C, N and P in vegetation. The nutrient concentration of plant tissue varies within a prescribed range depending on the relative availability of C, N and P. Before plant tissue is shed, depending on the tissue a fixed fraction of the nutrients is recycled. The nutrients contained in dead plant tissue and organic matter are mineralized and released back into the soil solution. The model version applied in this study is based on Goll et al. (2017a, 2018) and referred to as ORCHIDEE-CNP v1.2. Major modifications compared to v1.1 are described as follows (details can be found in Sect. S1 in the Supplement).

The original formulation of photosynthetic capacity in ORCHIDEE-CNP v1.1 assumed leaf N to be the sole regulator of leaf photosynthetic characteristics (Kattge et al., 2009). Here, we applied a new empirical function that relates photosynthetic capacity to both the leaf N and P concentration based on data from 451 species from 83 different plant families (D. S. Ellsworth et al., unpublished data). Leaf C : N : P ratios that were prescribed in ORCHIDEE-CNP v1.1 a priori in a narrow range specific to the plant functional type (PFT) are now given a larger range common to all PFTs (Table S1), allowing for the prediction of variation in leaf stoichiometry across climate and soil gradients independently of the prescribed vegetation (PFT) map.

In ORCHIDEE-CNP v1.1, an empirical function, $f(T_{\mathrm{soil}})$, was used to reduce biochemical mineralization and plant nutrient uptake at low soil temperature (Eq. 5 in Goll et al., 2017a), which was adopted from the N-enabled version of

ORCHIDEE (Zaehle and Friend, 2010) to avoid an unrealistic accumulation of N within plants when temperatures are low. We found that this function was not needed when P uptake is accounted for and was thus removed. It should be noted that this temperature dependence is different from the one that describes the temperature dependence of soil organic matter (SOM) and litter decomposition. For grasslands and croplands, we implemented root dormancy, which is triggered by drought or low temperatures. During dormancy, root maintenance respiration is reduced by 90 % following Shane et al. (2009), but root acquisition of soil nutrients continues as long as root biomass exists (Malyshev and Henry, 2012). It should be noted that total root loss can occur for extremely long droughts or cold periods when maintenance respiration depletes root carbon.

Several parameters were recalibrated, i.e., the coefficient relating maintenance respiration to biomass and the leaf to sapwood ratio, or corrected in the case of the turnover of sapwood for tropical evergreen broadleaf forest (TREBF) and tropical rain-green broadleaf forest (TRDBF) to achieve more realistic wood growth rates for those forests (not shown). We also adjusted the recycling efficiency of nutrients from the root ($f^{\mathrm{N}}_{\mathrm{trans,root}}$, $f^{\mathrm{P}}_{\mathrm{trans,root}}$) and leaf ($f^{\mathrm{N}}_{\mathrm{trans,leaf}}$, $f^{\mathrm{P}}_{\mathrm{trans,leaf}}$) according to data compilations from Freschet et al. (2010) and Vergutz et al. (2012). The new values of these parameters and their sources are given in Supplement (Sect. S1).

2.2 Simulation setup

We performed a global simulation at $2^\circ \times 2^\circ$ spatial resolution for the historical period (1700–2017) by adapting the TRENDY version 6 protocol (Sitch et al., 2015; Le Quéré et al., 2018). The simulation was performed using historical climate forcing, land cover changes and management (i.e., mineral fertilizer application, crop harvest; see Sect. 3.1.6), and atmospheric CO_2 concentrations (S3-type simulation). Prior to the historical simulation, we performed a model spin-up to equilibrate the C, N and P pools and fluxes (Sect. S1A in the Supplement) by forcing the model with cycled climate forcing of 1901–1920 and the land cover map and land management corresponding to the year 1700. To disentangle the effect of introducing nutrient cycles into ORCHIDEE, we performed the same simulation with ORCHIDEE (revision 5375), which has no nutrient cycles and a comparable parameterization for other processes. ORCHIDEE was run at a higher spatial resolution ($0.5^\circ \times 0.5^\circ$) than ORCHIDEE-CNP. Prior to the analysis, the data from ORCHIDEE were remapped to the resolution of ORCHIDEE-CNP.

2.2.1 Meteorological data

The model was forced by CRU-JRA-55 meteorological data provided at a spatial resolution of $0.5^\circ \times 0.5^\circ$ and upscaled to a resolution of $2^\circ \times 2^\circ$. These data comprise global 6-hourly climate forcing data providing observation-based temperature, precipitation and incoming surface radiation. They are derived from Climatic Research Unit (CRU) TS3.1 monthly data (Harris et al., 2014) and the Japanese 55-year Reanalysis (JRA-55) data (Kobayashi et al., 2015), covering the period 1901 to 2017. This climate dataset was provided by the TRENDY v6 model intercomparison project (Le Quéré et al., 2018).

2.2.2 Land cover

The historic land cover change maps were based on the European Space Agency Climate Change Initiative (ESA-CCI) land cover data (Bontemps et al., 2013). To be used by global vegetation models like ORCHIDEE-CNP, ESA-CCI land cover data were aggregated to $2^\circ \times 2^\circ$ and grouped into PFTs using the reclassification method from Poulter et al. (2011, 2015). The fraction of cropland and pasture in the PFT map was further constrained by the cropland area and the sum of pasture and rangeland area for the year 2010, respectively, in the History Database of the Global Environment land use dataset (HYDE 3.2; Klein Goldewijk et al., 2017a, b), which were also aggregated to $2^\circ \times 2^\circ$. The above processes produced a reference ESA-CCI-based PFT map for the year 2010. The land use changes derived from Land-Use Harmonization (LUH) v2 (http://luh.umd.edu/data.shtml, last access: 3 September 2018; an update release of Hurtt et al., 2011) were aggregated to $2^\circ \times 2^\circ$ and then applied to this reference PFT map to constrain the land cover changes of forest, grassland, pasture and rangeland, and cropland during the period 1700–2017 using the backward natural land cover reconstruction method of Peng et al. (2017). As a result, a set of historic PFT maps suitable for global vegetation models was established, distinguishing global land cover changes for the period of 1700–2017 at $2^\circ \times 2^\circ$ resolution.

2.2.3 Soil and lithology datasets

ORCHIDEE-CNP v1.2 is forced by information on soil texture, pH, bulk density and soil types (Goll et al., 2017a). We used a global gridded map of three soil texture classes from Zobler (1986) to derive soil-texture-specific parameters for soil water capacity, hydraulic conductivity and thermal conductivity. We used global gridded data on bulk density from the Harmonized World Soil Database (HWSD; FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) and soil pH from the International Geosphere–Biosphere Programme Data Information System Soil Data (Global Soil Data Task Group, 2000). Soil pH forcing maps are needed to simulate the dynamics of NH_3 and NH_4^+ in soil in ORCHIDEE (Zaehle and Friend, 2010). We used a global gridded map with the dominant soil orders (following the USDA Soil Taxonomy) at $1^\circ \times 1^\circ$ resolution to derive soil-order-specific soil phosphorus sorption parameters (Goll et al., 2017a).

The release of P from chemical weathering of rocks is computed dynamically following Goll et al. (2017a) and depends on the lithology types and soil shielding (discontinuation of the active soil zone from the bedrock) (Hartmann et al., 2014). We used the global lithological map (GLiM) of Hartmann and Moosdorf (2012) upscaled to $1^\circ \times 1^\circ$ resolution, which accounts for the lithology fractional coverage of 16 classes on a sub-grid scale. We also used a spatially explicit map of soil shielding at $1^\circ \times 1^\circ$ resolution (Hartmann et al., 2014).

2.2.4 Atmospheric nitrogen and phosphorus deposition

Global gridded monthly atmospheric N and P deposition during 1860–2017 was derived from a reconstruction based on the global aerosol chemistry–climate model LMDZ-INCA (Wang et al., 2017). LMDZ-INCA was driven by emission data, which included sea salt and dust for P, primary biogenic aerosol particles for P, oceanic emissions for N (NH_3), vegetation emissions for N (NO), agricultural activities (including fertilizer use and livestock) for N, and fuel combustion for both N (NO_y and NH_x) and P. Reconstructions for the years 1850, 1960, 1970, 1980 and 1990, as well as each year from 1997 to 2013, were linearly interpolated to derive a time series for 1850–2013. For the period before 1850, we assumed N and P deposition rates of the year 1850. For the period after 2013, we assumed rates of the year 2013. In ORCHIDEE-CNP, atmospheric N and P deposition is added to the respective soil mineral N and P pools without considering interception by the canopy.

2.2.5 Nutrient management

For croplands, we used yearly gridded mineral N and P fertilizer application data from Lu and Tian (2017) available for the period 1960 to 2017. This dataset is based on national-level data on crop-specific fertilizer application amounts from the International Fertilizer Industry Association (IFA) and the FAO. N and P mineral fertilization between 1900 and 1959 was linearly extrapolated assuming that fertilizer applications for 1900 are zero and that there were no N and P fertilizers applied before 1900. For pasture, we used global gridded datasets of mineral N fertilizer application rates from Lu and Tian (2017), developed by combining country-level statistics (FAO) and land use datasets (HYDE 3.2) (Xu et al., 2019). For both cropland and pasture, N and P in mineral fertilizer were assumed to go directly into soil mineral pools, where all mineral N fertilizer was assumed to be in the form of ammonium nitrate, with half of N as ammonium (NH_4^+) and half as nitrate (NO_3^-).

Manure applications are also included as a model forcing, given their significant input contribution to agricultural soils. For cropland, we used gridded annual manure N application data for the period 1860–2014 from Zhang et al. (2017) compiled and downscaled based on country-specific annual livestock population data from FAOSTAT. For the period before 1860, we assumed N and P deposition rates of the year 1860. For pasture, we used global gridded datasets of N manure application rates from Lu and Tian (2017). The application of manure P in cropland and pasture was derived from manure N assuming a manure P : N ratio of 0.2. This ratio is a weighted value by the amount of manure N applied to soil and derived from ruminants ($14.4\,\mathrm{Tg\,N\,yr^{-1}}$) and monogastric animals ($10.1\,\mathrm{Tg\,N\,yr^{-1}}$) from FAOSTAT for the year 2000 with P : N ratios of 0.165 in ruminant manure (mean of 0.15–0.18 from Lun et al., 2018) and 0.26 in monogastric manure (mean of 0.24–0.28 from Lun et al., 2018). For manure applied to cropland and pasture, we assumed a typical slurry application with 90 % of the N in the liquid part of the slurry (like urine) going into the soil $NH4^+$ pool. For the solid part of the slurry, we assumed it goes into a litter pool with a C : N ratio of 10 : 1 following Soussana and Lemaire (2014).

Mineral and manure N and P fertilizers in cropland were applied at day of year (DOY) 120 for the Northern Hemisphere (30–90° N), DOY 180 for tropical regions (30° N–30° S) and DOY 240 for the Southern Hemisphere (30–90° S).

3 Evaluation

We evaluated the performance of ORCHIDEE-CNP v1.2 based on four major aspects (Fig. 1). Firstly, we evaluated the global C, N and P flows and storages. In the absence of robust spatially resolved estimates of N and P fluxes, we used the data-driven reconstruction of steady-state C, N and P fluxes on the biome level from the Global Observation-based Land-ecosystems Utilization Model of Carbon, Nitrogen and Phosphorus (GOLUM-CNP) v1.0 (Wang et al., 2018) (Table 1). Secondly, we evaluated plant resource use efficiencies (RUEs) of light, water, C, N and P on global and biome scales. RUEs reflect how ecosystems adjust and adapt to the availability of nutrient, water, light, and carbon resources (Fernández-Martínez et al., 2014; Hodapp et al., 2019). For this, we used estimates from site measurements and observation-based gridded datasets. Thirdly, we evaluated the response of gross primary productivity (GPP) to elevated CO_2 to assess the response of plant productivity to changing resource availability (i.e., CO_2) and historical perturbation C fluxes. For this, we used observation-based estimates (Ehlers et al., 2015; Campbell et al., 2017). Fourthly, we evaluated large-scale patterns of vegetation and soil N : P ratios as well as the N and P openness and turnover rates on the ecosystem level to assess spatial variation in nutrient limitation and the underlying drivers. For this, we used estimates from GOLUM-CNP, site measurements and observation-based gridded datasets (Kerkhoff et al., 2005; McGroddy et al., 2004; Reich and Oleksyn, 2004; Tipping et al., 2016; Butler et al., 2017; Wang et al., 2018). Finally,

Table 1. Main information on datasets used for global evaluation of ORCHIDEE-CNP.

Dataset	Variable	Resolution	Period	Uncertainties	References
MODIS	GPP, NPP, CUE	1 km	2000–2015	Bias against local measurements for GPP and NPP	Running et al. (2004); Zhao et al. (2005); Turner et al. (2006)
MTE	GPP, WUE	0.5°	1982–2011	25 ensemble trees for GPP and ET	Jung et al. (2009, 2011)
BESS	GPP	0.5°	2001–2015	Bias against local measurements	Ryu et al. (2012); Jiang and Ryu (2016)
BETHY	NPP	0.008°	2000–2009	–	Tum et al. (2016); Wißkirchen et al. (2013)
GIMMS	NPP	0.5°	1982–2015	Using different climate inputs	Smith et al. (2016)
TRENDY v6	NBP	0.5°	1959–2016	1σ standard deviation	Sitch et al. (2015)
JENA_inversion	NBP	1°	1985–2016	–	Rödenbeck et al. (2003)
CAMS inversion	NBP	1.875° × 3.75°	1979–2016	–	Chevallier et al. (2005)
CTracker inversion	NBP	1°	2001–2016	–	van der Laan-Luijkx et al. (2017)
Peng BNF	BNF	biome	2001–2009	–	Peng et al. (2019)
Sullivan BNF	BNF	biome	1999, 2009	–	Sullivan et al. (2014)
Mayorga	N and P leaching	polygon	2000	–	Mayorga et al. (2010)
Helfenstein	K_m	soil order	–	–	Helfenstein et al. (2018)
Sun	Pasae activity	10 km	–	–	Sun et al. (2020)
GOLUM-CNP	C, N and P fluxes, N and P openness and turnover rate, PUE, NUE	0.25°	2001–2010	–	Wang et al. (2018)
Global SeaWiFS Level-3 data and MTE GPP	LUE	0.01°	1997–2006	–	Gobron et al. (2006a, b)
Butler	Leaf N : P ratio	1 km		100 estimates by Bayes' method	Butler et al. (2017)
Site leaf measurements	Leaf N : P ratio	site	–	–	Kerkhoff et al. (2005); McGroddy et al. (2004); Reich and Oleksyn (2004)
Tipping	SOM C, N and P	site	–	–	Tipping et al. (2016)
Site measurements of NUE and PUE	NUE and PUE	site	–	–	Gill and Finzi (2016)

we showed the implications of ORCHIDEE-CNP for C cycling by evaluating the spatiotemporal patterns of terrestrial C fluxes and pools of the two versions of ORCHIDEE. For this, we used observation-based products of GPP and atmospheric inversions of the net land–atmosphere CO_2 flux excluding fossil fuel emissions (Table 1). Each dataset is summarized in Table 1 and described in detail in the Supplement. All the gridded datasets with high spatial resolutions (Table 1) were resampled to the 2° × 2° resolution of the model output using area-weighted mean methods.

3.1 Ecosystem productivity

Different data-driven maps of NPP and GPP based on remote sensing and climate datasets were used (Table 1), thereby accounting for the uncertainty of each product and for the uncertainty from the spread between different products. Uncertainties of each NPP and GPP product were derived according to the original publications. We used a 20 % uncertainty of gridded GPP from the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Breathing Earth System Simulator (BESS) (Sect. S1C in the Supplement; Turner et al., 2006; Jiang and Ryu, 2016) at a 2° scale. This is a coarse extrapolation of uncertainty reported at the grid cell scale, since none of these products reported spatial error covariance information, allowing us to upscale this uncertainty at 2° resolution. Further, for some products, uncertainty was defined as the bias against local measurements (Turner et al., 2006) and for others by using different climate input fields (Table 1). For multi-tree ensemble (MTE) GPP (Table 1), we used the spread (1σ standard deviation) from an ensemble of 25 members produced by different machine-learning methods (Jung et al., 2009). For MODIS NPP (Table 1), we used a 19 % uncertainty as assessed by Turner et al. (2006). For BETHY NPP we do not have an uncertainty (Tum et al., 2016). For Global Inventory Modeling and Mapping Studies (GIMMS) NPP (Table 1), we used the variance of three sets of products (Table 1) based on different climate datasets (Smith et al., 2016).

Two statistical indices were used to summarize the performance of ORCHIDEE and ORCHIDEE-CNP with respect to the interannual and seasonal variability of GPP and the interannual variability of net biome productivity (NBP) (Sect. 4.6): the coefficient of determination (R^2) and relative mean square deviation (rMSE). The rMSE is defined as

$$\mathrm{rMSE} = \frac{\sum_{j=1}^{n} \left(X_{\mathrm{model},j} - X_{\mathrm{ref},j}\right)^2}{\sum_{j=1}^{n} \left(X_{\mathrm{ref},j} - \hat{X}_{\mathrm{ref},j}\right)^2}. \quad (1)$$

X_{model} and X_{ref} are values from models (i.e., ORCHIDEE and ORCHIDEE-CNP) and referenced datasets (i.e., MTE and BESS; Sect. S1C), respectively, and $\hat{X}_{\mathrm{ref},j}$ is the mean value across all years (for interannual variability evaluation) or all months (for seasonality evaluation).

3.2 Resource use efficiencies

The definition of resource use efficiencies is explained in Sect. 4.2. Observation-based light use efficiency (LUE) was calculated using MTE GPP, downward shortwave radiation from CRUJRA, and the fraction of absorbed photosynthetically active radiation (fAPAR) from the Global SeaWiFS Level-3 data (Gobron et al., 2006a, b). Uncertainty was derived from 25 ensemble members of MTE GPP. Observation-based water use efficiency (WUE) was calculated as the ratio between MTE GPP and MTE ET (Table 1); its uncertainties were calculated using a Monte Carlo resampling procedure in which 25 different members of GPP and ET were randomly selected. Observation-based carbon use efficiency (CUE) was calculated from the ratio of MODIS NPP to MODIS GPP. It should be noted that MODIS NPP is based on a calibrated version of the BIOME-BGC model (Turner et al., 2006) so that CUE is not strictly an observation-based quantity. CUE uncertainties were calculated using a Monte Carlo method given 20 % and 19 % uncertainty for MODIS GPP and NPP products at 2° resolution, respectively.

4 Results

4.1 Carbon, nitrogen, and phosphorus flows and storages

We compared the simulated fluxes of C, N and P within natural ecosystems for the period 2001–2010 to the data-driven estimates from GOLUM-CNP (Table 1; Sect. S1B in Supplement) on the global scale and for natural ecosystems at biome scale. Modeled global C, N and P fluxes in ORCHIDEE-CNP are comparable with the estimates by GOLUM-CNP (Fig. 2). One exception is that ORCHIDEE-CNP simulates a fourfold lower P leaching from soils ($3.7 \pm 9.7\,\mathrm{mg\,P\,m^{-2}\,yr^{-1}}$) than GOLUM-CNP ($23\,\mathrm{mg\,P\,m^{-2}\,yr^{-1}}$) (Fig. 2), which mainly occurs in forest ecosystems (Fig. S1). Note that GOLUM-CNP presents the steady-state C, N and P cycles in natural biomes, omitting human perturbations that have strongly altered the flows of C, N and P during the recent past. The impact of such perturbations on the nutrient flows are analyzed in detail in Sects. S2 and S3 in the Supplement.

In terms of C and nutrient storages, ORCHIDEE-CNP simulated comparable soil C, N and P storage (soil organic matter and litter) but higher vegetation C, N and P than GOLUM-CNP. Detailed comparisons for the spatial pattern of soil organic carbon (SOC) and forest aboveground C against observation-based datasets can be found in Figs. S2 and S3.

4.2 Resource use efficiencies

Here we evaluate the resource use efficiencies of GPP for light (L), water (W), C, N and P defined by

$$\mathrm{LUE} = \frac{\mathrm{GPP}}{\mathrm{f\,APAR} \times \mathrm{PAR}}, \quad (2)$$

$$\mathrm{WUE} = \frac{\mathrm{GPP}}{\mathrm{ET}}, \quad (3)$$

$$\mathrm{CUE} = \frac{\mathrm{GPP}}{\mathrm{GPP}}, \quad (4)$$

$$\mathrm{NUE} = \frac{\mathrm{GPP}}{F_{\mathrm{N}}}, \quad (5)$$

$$\mathrm{PUE} = \frac{\mathrm{GPP}}{F_{\mathrm{P}}}, \quad (6)$$

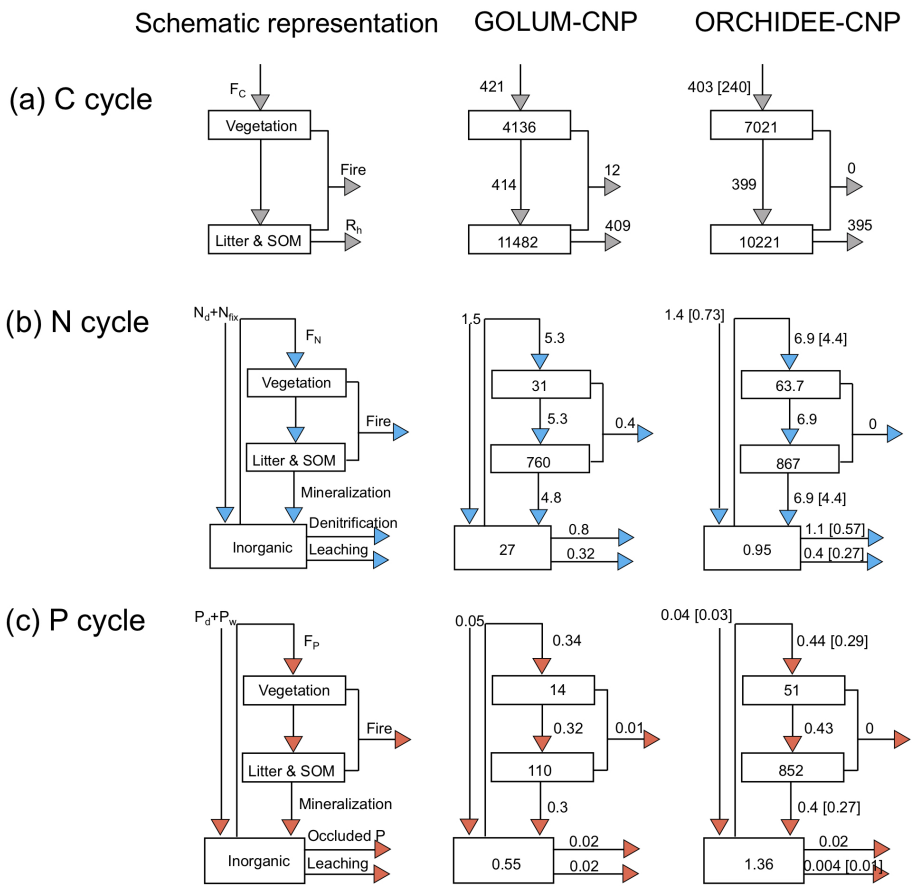


Figure 2. Flowchart of mean flows and storages per area of C, N and P ($g\,C\,g^{-1}$, $N\,g^{-1}$, $P\,m^{-2}\,yr^{-1}$) in natural biomes for GOLUM-CNP and ORCHIDEE-CNP. GOLUM-CNP stands for Global Observation-based Land-ecosystems Utilization Model of Carbon, Nitrogen and Phosphorus (GOLUM-CNP) v1.0, which is a data-driven model of steady-state C, N and P cycles for present-day (2001–2016) conditions. C, N and P losses via fire in ORCHIDEE-CNP are ignored. Numbers in square brackets indicate the standard deviations to account for the spatial spread of C, N and P fluxes.

where GPP is the annual gross primary productivity ($g\,C\,m^{-2}\,yr^{-1}$), fAPAR the fraction of absorbed photosynthetically active radiation (%), PAR the annual photosynthetically active radiation ($W\,m^{-2}\,yr^{-1}$), ET the annual evapotranspiration ($mm\,m^{-2}\,yr^{-1}$), and F_N and F_P the total N uptake ($g\,N\,m^{-2}\,yr^{-1}$) and P uptake by plants ($g\,P\,m^{-2}\,yr^{-1}$), respectively. We calculated fAPAR in ORCHIDEE-CNP and ORCHIDEE as a function of leaf area index (LAI): $fAPAR = 1 - \exp(-0.5 \cdot LAI)$ (Ito et al., 2004).

Compared to observed LUE (Sect. S1E in the Supplement), ORCHIDEE-CNP modeled median values at the biome level are generally lower but still within the ranges of uncertainties of observation-based datasets (Sect. 3.2) except for tropical (TRF) and temperate deciduous forest (TEDF). In comparison to ORCHIDEE, ORCHIDEE-CNP simulated LUEs closer to observations for four out of six biomes (TECF, BOCF, TEG, TRG) (Fig. 3a).

Compared to observed WUE, the ORCHIDEE and ORCHIDEE-CNP simulated values fall within the uncertainty range of observations (Fig. 3b). However, the WUE values from ORCHIDEE-CNP are on the high end of the range for temperate conifers (TECF) and BOCF and on the low end for temperate and tropical grasslands (TEG and TRG). The highest median WUEs were correctly simulated in temperate forests by ORCHIDEE-CNP (Fig. 3b), but the lowest WUE values were simulated in temperate instead of tropical forests.

Compared with observed CUE, ORCHIDEE-CNP simulated comparable values for TEDF and TECF but lower values for TRF, BOCF and grasslands. Both ORCHIDEE-CNP and ORCHIDEE cannot capture the increase in CUE from tropical to boreal forests apparent in the observation-based products (Fig. 3c) and in measurements from forest sites (Piao et al., 2010). In comparison to ORCHIDEE, ORCHIDEE-CNP simulated CUEs closer to observations for four out of six biomes (TEDF, TECF, BOCF, TEG) with respect to the median and spread.

Consistent with site observations of NUE from Gill and Finzi (2016) and GOLUM-CNP outputs, ORCHIDEE-CNP correctly simulated the high values of TECF and the low val-

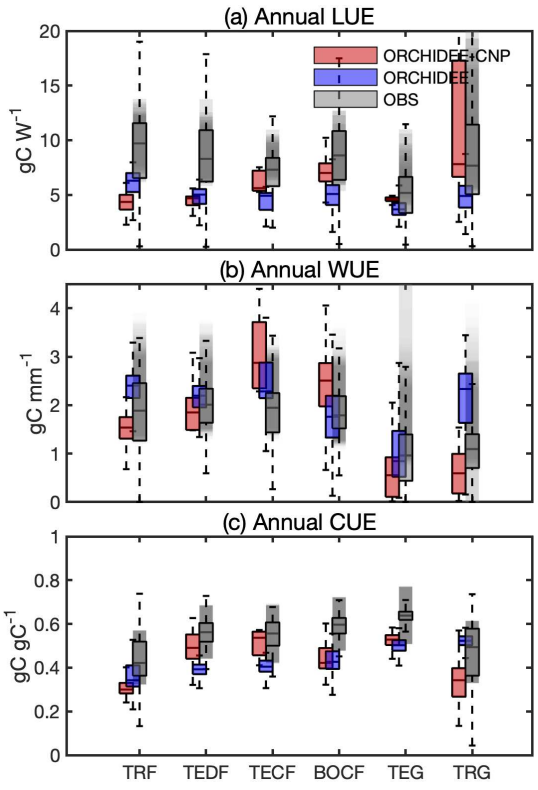


Figure 3. Comparison of annual use efficiencies of light (LUE), water (WUE) and carbon (CUE) between ORCHIDEE-CNP, ORCHIDEE and satellite-based estimations for six biomes: tropical rainforest (TRF), temperate deciduous forest (TEDF), temperate conifer forest (TECF), boreal conifer forest (BOCF), temperate grass (TEG) and tropical grass (TRG). The whiskers indicate the interquartile (box) and 95 % confidence intervals (dashed lines). Grey boxes indicate the satellite-based estimations (referenced). The grey shaded areas indicate the uncertainties of resource use efficiencies given by referenced estimations, which involve uncertainties for multi-estimations and spatial variability for each estimation.

ues of tropical forests (Fig. 4a). However, compared with site observations of PUE from Gill and Finzi (2016), showing a PUE decrease from tropical to boreal regions, ORCHIDEE-CNP simulated a rather flat value (Fig. 4b). This suggests a P limitation that is too strong in high-latitude ecosystems, consistent with the fact that the model underestimates peak northern GPP and the northern land sink (Sect. 4.6). Nevertheless, the model-simulated PUE values fall in the range of GOLUM-CNP estimates. Tropical C_4 grasslands have higher simulated NUE and PUE than temperate C_3 grasslands, consistent with GOLUM-CNP (Fig. 4).

4.3 CO_2 fertilization effect

We compare the simulated response of plant productivity to increasing CO_2 during the historical period (i.e., CO_2 fertilization effect Eco_2) with observation-based estimates for C_3 plants from the historical change in deuterium isotopomers in leaf herbarium samples (Ehlers et al., 2015). For global (C_3 and C_4) vegetation we compare to indirect evidence from carbonyl sulfide (COS) atmospheric ice-core observations (Campbell et al., 2017). The CO_2 fertilization effect is defined here by the GPP ratio (E_{CO_2}):

$$E_{CO_2} = \frac{GPP_{396}}{GPP_{296}}, \tag{7}$$

where GPP_{296} indicates pre-industrial GPP ($g\,C\,m^{-2}\,yr^{-1}$) under a CO_2 concentration of 296 ppm and GPP_{396} under a current CO_2 concentration of 396 ppm. Those CO_2 concentrations of 296 and 396 ppm correspond to tropospheric mixing ratios of CO_2 in the years $\sim$ 1900 and 2013, respectively, similar to values used for estimating the response of GPP to a $\sim$ 100 ppm CO_2 increase in Ehlers et al. (2015) and Campbell et al. (2017).

Modeled E_{CO_2} by ORCHIDEE-CNP of natural biomes ranges between 1.0 and 1.3 for most regions (Fig. 5a), slightly lower than global E_{CO_2} derived from COS of 1.26–1.36 (Campbell et al., 2017). Modeled E_{CO_2} values for C_3 plants (Figs. 5c, S4) are also consistent with E_{CO_2} from herbarium samples (Ehlers et al., 2015), equal to 1.23. When compared to ORCHIDEE without nutrient cycles, we found that ORCHIDEE-CNP simulates smaller and more realistic values of E_{CO_2} (Fig. 5c, d) but with lower values in boreal regions that could not be checked against observations (Fig. S5).

4.4 Ecosystem nutrient openness and nutrient turnover

Nutrients taken up by plants are either recycled within the ecosystem or acquired from external sources (P weathering of primary and secondary minerals, atmospheric N and P deposition, biological nitrogen fixation – BNF, and N and P fertilizer addition to cultivated lands). Wang et al. (2018) calculated an indicator of the openness of N and P cycling in natural ecosystems as the ratio of external inputs of N and P into the ecosystem to the total amount of N and P that plants use for GPP. Similarly, we diagnosed the openness for N and P (O_N and O_P) from the ORCHIDEE-CNP output by

$$O_x = \frac{I_x}{F_x + RSB_x}, \tag{8}$$

where I_x is the annual external nutrient input ($g\,X\,m^{-2}\,yr^{-1}$), F_x the annual plant uptake of soil nutrients ($g\,X\,m^{-2}\,yr^{-1}$) and RSB_x the flux of nutrients recycled within plants ($g\,X\,m^{-2}\,yr^{-1}$) by foliar nutrient resorption prior to leaf shedding. External nutrient inputs include atmospheric N deposition and BNF, as well as P deposition and P release from rock weathering.

Modeled O_N in natural biomes by ORCHIDEE-CNP showed only a small variance across the globe, whereas GOLUM-CNP predicts a higher O_N in tropical and temperate regions than in boreal regions (Fig. 6a, b). O_P values

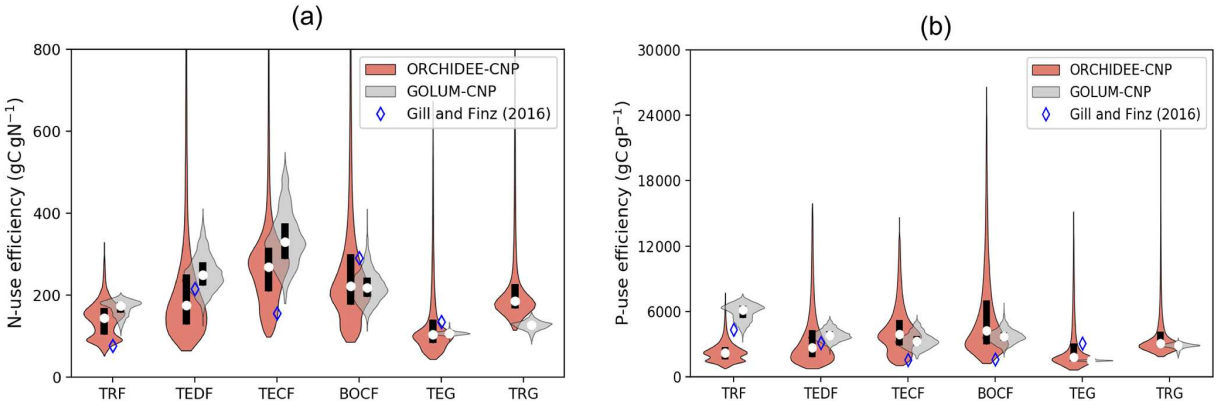


Figure 4. Violin plots of nitrogen use efficiency (NUE; **a**) and phosphorus use efficiency (PUE; **b**) by ORCHIDEE-CNP, GOLUM-CNP and observations (Gill and Finzi, 2018) for six biomes: tropical rainforest (TRF), temperate deciduous forest (TEDF), temperate conifer forest (TECF), boreal conifer forest (BOCF), temperate grass (TEG) and tropical grass (TRG). Open circles are the medians of all grid cells within each biome, with balloons representing the probability density distribution of each value. Black whiskers indicate the interquartile.

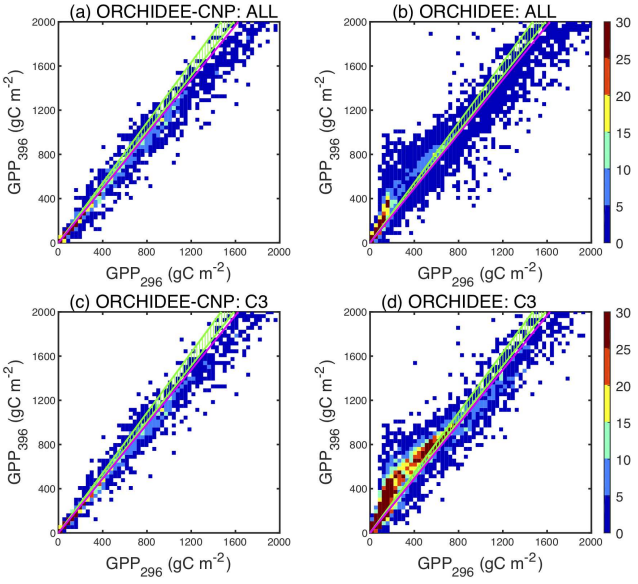


Figure 5. Comparisons between pre-industrial GPP with an atmospheric CO_2 concentration of 296 ppm (GPP_{296}) and current GPP with an atmospheric CO_2 concentration of 396 ppm (GPP_{396}) for all natural plants **(a, b)** and natural C_3 plants **(c, d)** by ORCHIDEE-CNP **(a, c)** and ORCHIDEE **(b, d)**. The color scale shows the point density. Different point densities and patch sizes for ORCHIDEE and ORCHIDEE-CNP are due to the different spatial resolution ($2^\circ \times 2^\circ$ for ORCHIDEE-CNP and $0.5^\circ \times 0.5^\circ$ for ORCHIDEE). The ratio between GPP_{396} and GPP_{296} indicates the CO_2 fertilization effects (E_{CO_2}). Green dashed areas indicate the observed E_{CO_2} from the Campbell et al. (2017) COS records. Pink lines indicate the observed E_{CO_2} from Ehlers et al. (2015).

are below 15 % in ORCHIDEE-CNP for most biomes, of a similar order of magnitude as in GOLUM-CNP (Fig. 6c, d). ORCHIDEE-CNP simulates a lower O_N in tropical natural biomes than GOLUM-CNP, which is mainly due to lower but more realistic tropical BNF in ORCHIDEE-CNP compared to GOLUM-CNP (Sect. S4 in the Supplement). ORCHIDEE-CNP simulates a higher O_N in high-latitude grassland (Fig. 6a, b) than GOLUM-CNP, which is due to overestimation of BNF in NH in ORCHIDEE-CNP (Sect. S4 in the Supplement). Modeled O_P in natural biomes by ORCHIDEE-CNP compares well with GOLUM-CNP except for central Africa (Fig. 6c, d). This is primarily because ORCHIDEE-CNP used a lower P deposition forcing than GOLUM-CNP.

Residence time quantifies the average time it takes for an N (or P) molecule to enter and leave the ecosystem (τ_N and τ_P). In this study, we adopted the approach of Carvalhais et al. (2014) for the carbon residence time. We define the residence time of N and P as the ratio of total respective nutrient stock in the ecosystem to their respective total input flux:

$$\tau_N = \frac{\sum_{i=1}^{5} N_i + N_{inorg}}{N_d + BNF}, \quad (9)$$

$$\tau_P = \frac{\sum_{i=1}^{5} P_i + P_{inorg}}{P_d + P_w}, \quad (10)$$

where N_i indicates the N mass (g N m^{-2}) in organic matter pools i (with i = plant, litter, SOM pools), N_{inorg} is the sum of all inorganic N pools, and N_d and BNF are N deposition and biological N fixation rates, respectively (g N m^{-2} yr^{-1}). Similarly, P_i is the P mass (g P m^{-2}) in organic matter pools, P_{inorg} the sum of inorganic P pools, and P_d and P_w are P deposition and P weathering release rates, respectively (g P m^{-2} yr^{-1}).

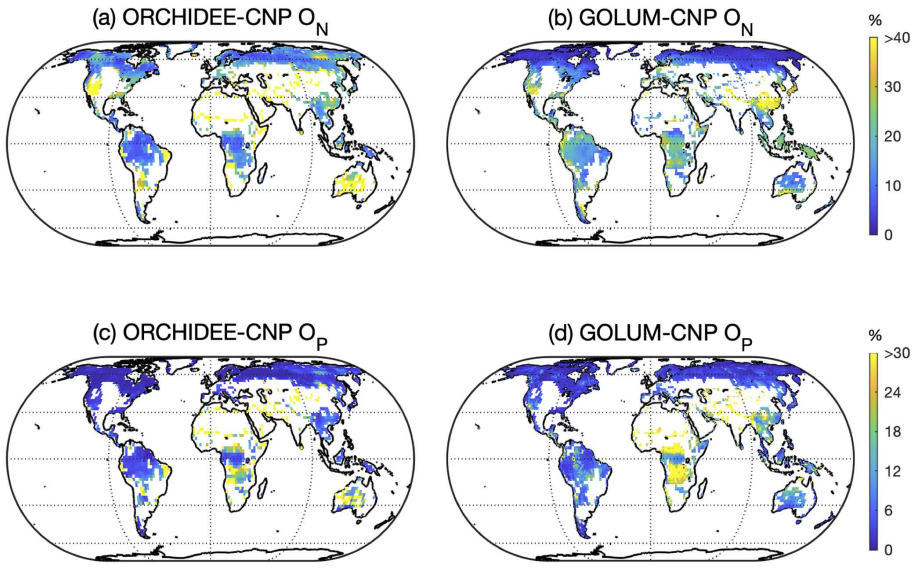


Figure 6. Global pattern of N (O_N, **a, b**) and P openness and that (O_P, **c, d**) simulated by ORCHIDEE-CNP **(a, c)** and GOLUM-CNP **(b, d)**. Pixels with managed lands $> 50\,\%$ in ORCHIDEE-CNP were masked. The same area was masked from the pattern of O_N and O_P for GOLUM-CNP.

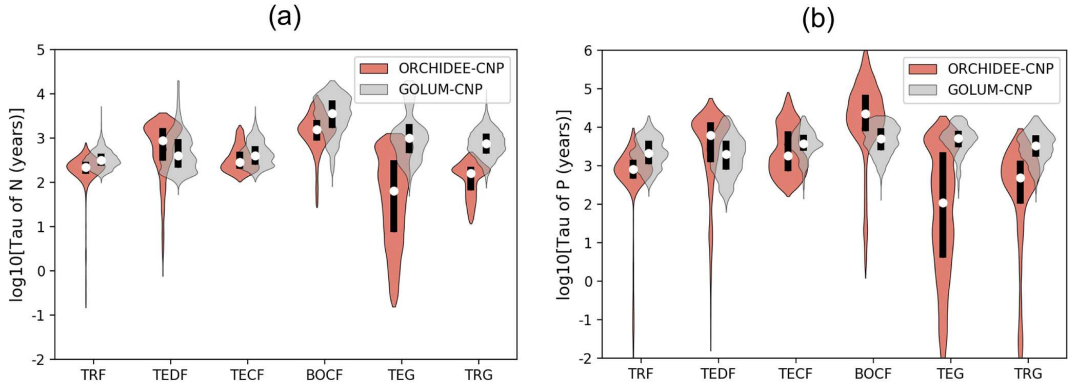


Figure 7. Violin plots of the residence time of **(a)** N and **(b)** P cycles for six biomes: tropical rainforest (TRF), temperate deciduous forest (TEDF), temperate coniferous forest (TECF), boreal coniferous forest (BOCF), temperate grass (TEG) and tropical grass (TRG). Open circles are the medians of all grid cells within each biome, with balloons representing the probability density distribution of each value. Black whiskers indicate the interquartile.

Modeled median τ_N of natural biomes in ORCHIDEE-CNP varies between 56 and 1585 years, while τ_P varies within a large range of 101 to 223 870 years (Fig. 7). ORCHIDEE-CNP captured the order of magnitude of τ_N and τ_P for forests found in GOLUM-CNP. Longer median τ_N (1585 years) and τ_P (1 223 870 years) are simulated for boreal forest compared to temperate and tropical forests (251–794 years for τ_N and 891–7080 years for τ_P) and grassland (56–158 years for τ_N and 101–468 years for τ_P) by ORCHIDEE-CNP, consistent with results from GOLUM-CNP. However, for grasslands, simulated τ_N (56-158 years) and τ_P (101–468 years) are 5–11-fold shorter than in GOLUM-CNP (Fig. 7).

4.5 Stoichiometry

4.5.1 Foliar stoichiometry

Leaf N : P ratios for natural biomes predicted by ORCHIDEE-CNP vary between 15 and 25 (Fig. 8a). The observed decline in median leaf N : P ratios with increasing latitude was not reproduced by the model (Sect. S1E1 in the Supplement; Fig. 8e), although the modeled latitudinal distribution of leaf N : P ratios remained within the 10–90th quantiles of the site-level data (Kerkhoff et al., 2005; McGroddy et al., 2004; Reich and Oleksyn, 2004). Further, the simulated leaf N : P ratios fall within the interquartile of upscaled site measurements by Butler

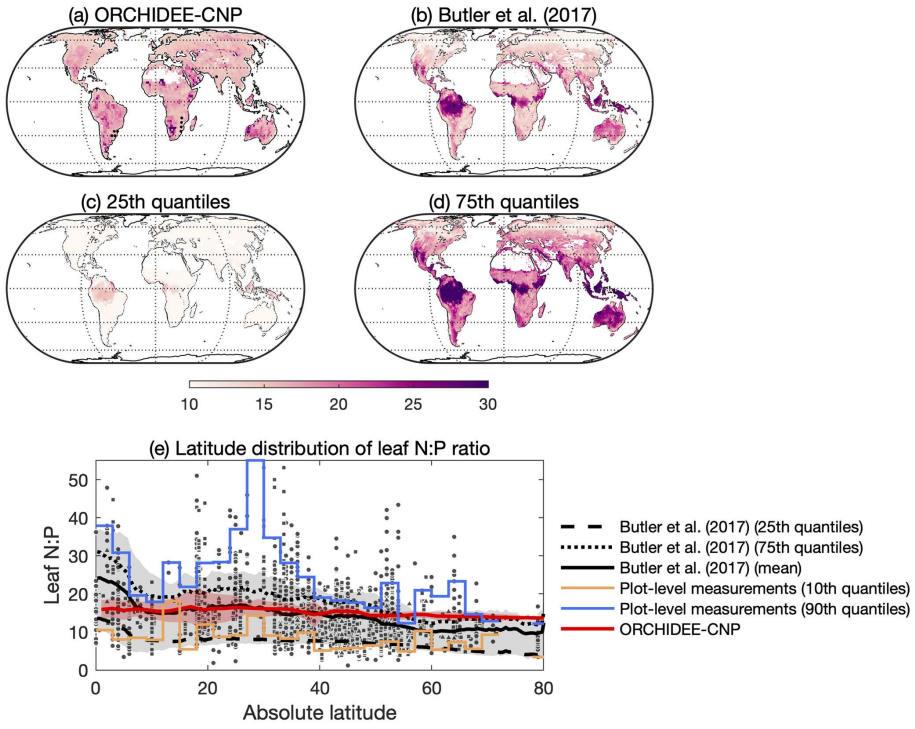


Figure 8. Comparisons of the leaf N : P ratio between ORCHIDEE-CNP, data-driven estimates and observations. **(a)** The global pattern of the mean leaf N : P ratio over 2001–2016 for ORCHIDEE and **(b)** for mean leaf N : P in Butler et al. (2017). Panels **(c)** and **(d)** are the 25th and 75th percentile, respectively, of the leaf N : P ratio by Butler et al. (2017). Dots in **(a)** indicate the area with a leaf N : P ratio from ORCHIDEE-CNP falling into the 25th–75th percentiles of the Butler et al. (2017) estimation. **(d)** The latitude distributions of the leaf N : P ratio for ORCHIDEE-CNP, the Butler et al. (2017) estimation and site measurements. Red shared area indicates the uncertainty from latitudinal spreads of the leaf N : P ratio for ORCHIDEE-CNP. The grey shaded area indicates the uncertainty from both the estimations and latitudinal spreads for Butler et al. (2017). Blue and yellow lines indicate the 10th and 90th percentiles, respectively, of measured leaf N : P ratios in each bin of 3° latitude.

et al. (2017) for most of the globe, with the exception of regions north of 55° N where leaf N : P ratios are outside the observation-based range, suggesting a P constraint relative to N that is too strong (Fig. 8).

4.5.2 Soil stoichiometry

Here we evaluate the modeled C : N, C : P and N : P ratios of soil organic matter for different biomes against data from the large compilation of measurements for soils (0–60 cm depth) by Tipping et al. (2016). Modeled C : N ratios fall into much more narrow ranges (7.8–11.8 for the widest interquartile range) compared to the observations (11.1–20.5; Fig. 9a) as a result of prescribing constant C : N ratios in ORCHIDEE-CNP (Goll et al., 2017a). SOM P content varies in ORCHIDEE-CNP as a consequence of varying biochemical phosphorus mineralization rates (Sect. S7), and thus C : P and N : P ratios of SOM show pronounced variation in space. ORCHIDEE-CNP simulates comparable N : P ratios as measurements in terms of both the median value and distributions for tropical forests, but it overestimates the observed N : P ratios by 108 %–327 % in temperate forest, tropical and temperate grassland soils (Fig. 9b, c). The higher observed C : P and N : P in forest compared to grassland soils are not captured by ORCHIDEE-CNP (Fig. 9b, c). We also compared ORCHIDEE-CNP N : P ratios to the results of GOLUM-CNP, which were based on data from Zechmeister-Boltenstern et al. (2015) that are more limited than Tipping et al. (2016), and found an overestimation for temperate forests, tropical forests and temperate grasslands.

4.6 Nutrient effects on carbon cycling

We analyze the performance of ORCHIDEE-CNP v1.2 and ORCHIDEE without nutrient cycles with respect to the spatiotemporal patterns of GPP, NPP and net biome productivity.

Global GPP and NPP simulated by ORCHIDEE-CNP averaged over the period 2001–2010 are 119 and 48 Pg C yr^{-1}, respectively, which are both within ranges of the data-driven products listed in Table 1 (Sect. S1C in the Supplement; Table S2). GPP and NPP simulated by ORCHIDEE-CNP are lower than those simulated by ORCHIDEE (140 Pg C yr^{-1} for GPP and 60 Pg C yr^{-1} for NPP). The values from ORCHIDEE are on the high end of the range of estimates from

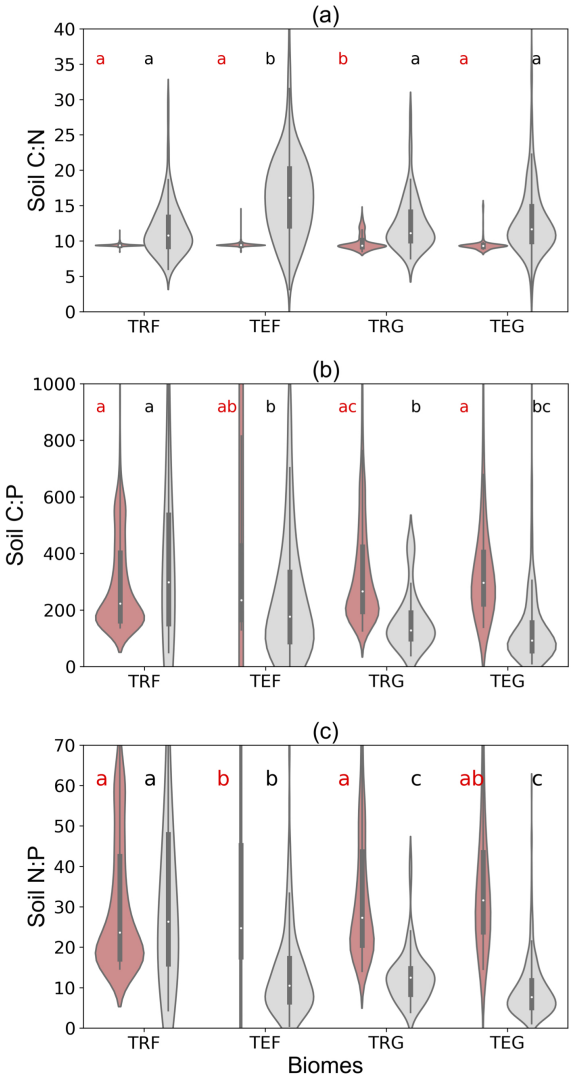


Figure 9. C : N, C : P and N : P ratios of soil organic matter by ORCHIDEE-CNP and plot-level measurements by Tipping et al. (2016) for four biomes: tropical forest (TRF), temperate forest (TEF), tropical grass (TRG) and temperate grass (TEG). Soil C : N : P ratios for ORCHIDEE-CNP are calculated for the total soil pool and include soil passive, slow and active pools, while measurements by Tipping et al. (2016) are for soils of 0–60 cm depth. The letters "a", "b" and "c" indicate the significance of differences among biomes from the analysis of variance (ANOVA).

the data-driven products in Table 1. ORCHIDEE-CNP simulated comparable GPP values for most parts of the globe (Fig. S6a) and comparable NPP values for most of the northern high latitudes (Fig. S6b), which lie within the range given by the data-driven products.

Interannual and seasonal variations of GPP reflect the response of ecosystems to interannual or seasonal climatic variability, as well as the effects of natural (e.g., fires, wind throw, insect outbreaks and storms) and anthropogenic disturbances (e.g., land management and land cover change) (Anav et al., 2015). Regarding the interannual anomalies of de-trended GPP (GPP_{int}) for the period 2001–2011, estimations on a global scale from ORCHIDEE-CNP show rather good correlation with the observation-driven model BESS GPP ($R^2 = 0.71$) but not with MTE GPP ($R^2 = 0.11$) (Fig. 10a). ORCHIDEE performs somewhat worse on a global scale than ORCHIDEE-CNP, primarily due to its low performance in the NH. We find that inclusion of nutrients in ORCHIDEE leads to a lower model prediction error on a global scale and for all latitudinal bands irrespective of the observation-based product (Fig. 10a).

Regarding the seasonal variation of GPP over the period 2001–2011, the predictions of ORCHIDEE-CNP are in good agreement with observation-based estimates and show no significant differences when compared to ORCHIDEE, except for tropical regions (Fig. 10b). Here, the model errors in seasonal variations of GPP are substantially larger for ORCHIDEE-CNP than for ORCHIDEE (Fig. 10b).

Net biome productivity (NBP) is defined as the net C exchange between the atmosphere and the terrestrial biosphere, which is the sum of net primary productivity, heterotrophic respiration and emissions due to disturbances; positive values denote a land carbon sink. Compared to the three sets of atmospheric inversions (CAMS, JENA and CTracker), ORCHIDEE(-CNP) performs slightly worse than the mean of predictions from 16 land surface models from TRENDY ensembles (v6) (Fig. 10c). ORCHIDEE-CNP shows a worse performance in interannual variability of NBP than ORCHIDEE when compared against inversion datasets at global scale and for the Northern Hemisphere. However, ORCHIDEE-CNP improved the performance of the interannual variability of NBP against inversion datasets relative to ORCHIDEE for tropical regions (higher R^2 and lower rMSE), with closer or even better fitness against inversion datasets than the mean value of TRENDY ensemble models (Fig. 10c).

5 Discussion

We performed a detailed evaluation of ORCHIDEE-CNP in terms of four nutrient-related ecosystem properties that control ecosystem gas exchanges and carbon storage: vegetation resource use efficiencies, CO_2 fertilization effect, ecosystem N and P turnover and openness, and large-scale pattern of ecosystem stoichiometries.

We find that the inclusion of nutrients tends to lead to improvements in simulated resource use efficiency of plant resources (light, carbon, water) on a biome scale (Sect. 5.1). In line with changes in resource use efficiency, the sensitivity of GPP to variations in climate is improved, leading to improved interannual variation in GPP, in particular for the Northern Hemisphere (Sect. 5.5). In addition, the response of GPP to an increasing atmospheric CO_2 concentration is improved (Sect. 5.2). However, model biases in C fluxes re-

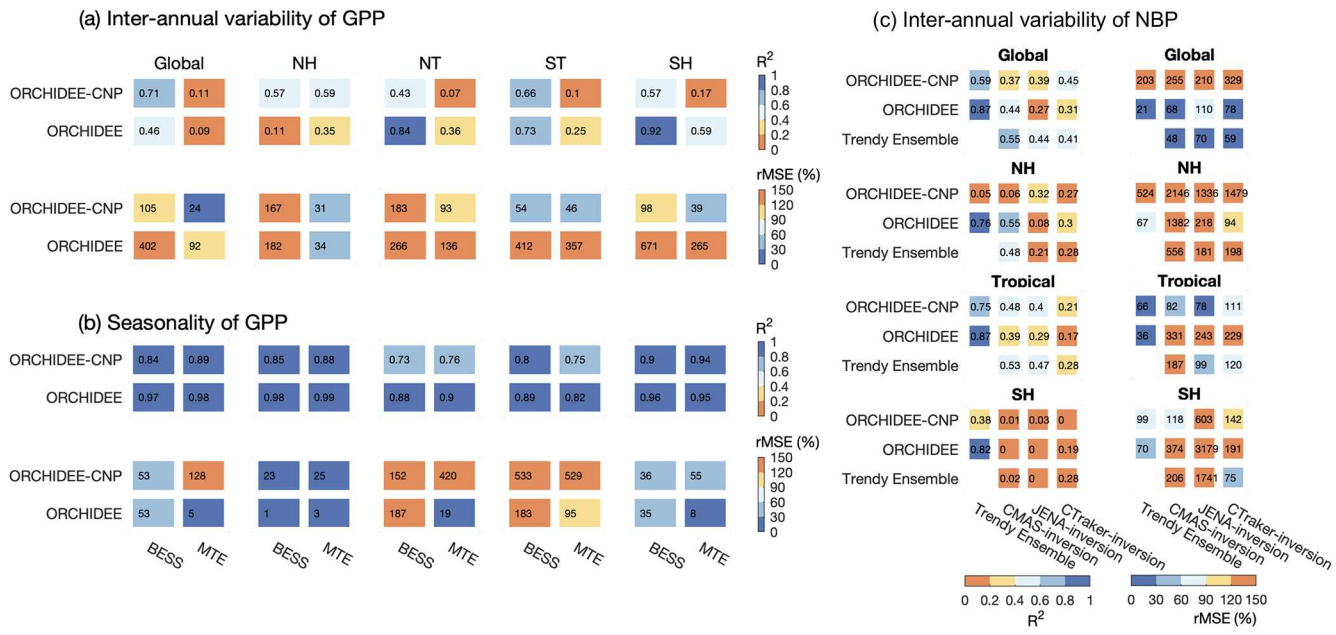


Figure 10. The performances of ORCHIDEE and ORCHIDEE-CNP on the interannual variability of de-trended anomalies of GPP during 2001–2010 **(a)**, the seasonal variability of mean GPP across 2001–2010 **(b)**, and the interannual variability of net biome productivity (NBP) **(c)**. Two statistics were used to represent the model performance: coefficient of determination (R^2) and relative mean square deviation (rMSE). For **(a)** and **(b)**, the evaluations are for the globe, the Northern Hemisphere (30–90° N; NH), the northern tropics (0–30° N; NT), the southern tropics (0–30° S; ST) and the Southern Hemisphere (30–90° S; SH). Two sets of observation-based GPP products (BESS GPP and MTE GPP) were used for the comparison. For **(c)**, the evaluations are for the globe, the Northern Hemisphere (30–90° N, NH), the tropics (30° S–30° N) and the Southern Hemisphere (30–90° S; SH). The mean value across TRENDY ensemble models (v6) and three sets of NBP from inversion datasets were used as the reference database for the comparison with different available periods (TRENDY Ensemble: 1959–2016; CAMS: 1979–2016; JENA: 1985–2016; CTracker: 2001–2016).

mained or increased, for example, in the NBP of the Northern Hemisphere. The analysis of nutrient use efficiencies (Sect. 5.1), stoichiometry (Sect. 5.4), and ecosystem openness and turnover of nutrients (Sect. 5.3) reveals biases in boreal regions that might be related to issues with soil organic matter accumulation that is too strong and the dependency of photosynthesis on leaf nutrients in needleleaf PFTs. On a seasonal scale, we found a general deterioration of the simulated seasonal cycle of GPP due to the inclusion of nutrient cycles (Sect. 5.5).

In the following, we discuss in more detail the model performance with respect to nutrient cycles and their effects on simulated C fluxes, and we propose ways to address model biases.

5.1 Inclusion of nutrient cycling improves use efficiencies of other plant resources

Resource use efficiency (RUE) is an ecological concept that measures the proportion of supplied resources that support plant productivity; i.e., it relates realized to potential productivity (Hadapp et al., 2019). It is therefore a critical ecosystem property that relates resource availability to ecosystem productivity, as well as being affected by resource availability.

With the inclusion of the additional plant resources nitrogen and phosphorus, changes in the simulated vegetation use efficiencies of resources like water (WUE), light (LUE) and carbon (CUE) are expected. Indeed, the annual use efficiencies on the biome scale differ between ORCHIDEE-CNP and ORCHIDEE. In comparison to observation-based estimates, the inclusion of nutrient cycles tends to improve simulated LUE and CUE and WUE (Fig. 3).

Both ORCHIDEE-CNP and ORCHIDEE generally underestimate annual LUE for forest biomes (Fig. 3a), which is due to a high bias in fAPAR in both models (28 %–380 % for ORCHIDEE and 80 %–173 % for ORCHIDEE-CNP) (Fig. S4a, b). Although the bias in LUE for TRF is higher, the bias in GPP is largely reduced, whereas the bias in fAPAR is similar in ORCHIDEE-CNP compared to ORCHIDEE (Fig. S4a, b), indicating general issues in ORCHIDEE with respect to how light is transferred within the canopy in tropical forest. Both versions assume a constant canopy light extinction coefficient of 0.5, omitting variations among biomes due their distinctive canopy architectures (Ito et al., 2004). Improving this part of the model requires a canopy light transfer scheme that better accounts for canopy structure (Naudts et al., 2015)

and the inclusion of different light components including diffuse incoming, scattered and direct light (Zhang et al., 2020).

ORCHIDEE-CNP simulated a lower WUE than ORCHIDEE with the exception of coniferous biomes (Fig. 3b). The improvement of WUE in TRF is related to improvements in GPP and ET, while the overestimation of WUE in coniferous-dominated biomes by ORCHIDEE-CNP is related to an overestimation of GPP (Fig. S4c). The latter is likely a result of the application of a relationship between photosynthetic capacity and the leaf nutrient concentration, which is based on measurements from broadleaf species for all PFTs. Kattge et al. (2009) showed that coniferous PFTs have a $\sim 40\,\%$ lower carboxylation capacity for a given leaf nitrogen concentration than other PFTs. The omission of this could explain the bias in coniferous GPP in ORCHIDEE-CNP. Uncertainties in evaluation datasets hamper a more detailed evaluation of the variations of WUE among biome types.

We found that the inclusion of nutrient cycles improved the spatial variability in simulated CUE, but general biases remain (Fig. 3c), and uncertainties in observation-based estimates are large. Improvements are mainly found in temperate biomes (TEDF, TECF and TEG), indicating that the allocation of GPP to respiration and biomass growth, which is controlled by nutrient availability, works reasonably well. ORCHIDEE-CNP underestimates CUE for tropical biomes (TRF and TRG) more strongly than ORCHIDEE, despite substantially reduced biases in NPP and GPP (Fig. S4d). However, we should be cautious in drawing conclusions considering the large uncertainty in MODIS CUE (He et al., 2018).

NUE and PUE on the biome scale compare well to estimates (Fig. 4), indicating that ORCHIDEE-CNP is able to simulate the coupling strength between C, N and P cycles. However, ORCHIDEE-CNP underestimates PUE in tropical forests. A sensitivity analysis by GOLUM-CNP indicated that NUE and PUE were most sensitive to the NPP allocation fractions (especially to woody biomass) and foliar stoichiometry (Wang et al., 2018). Therefore, we attribute the biases in PUE to the biases in foliar stoichiometry (Fig. 8) and to issues in plant internal P allocation in ORCHIDEE-CNP (Fig. S1).

5.2 Inclusion of nutrient cycling improves CO_2 fertilization effect

The effect of CO_2 fertilization on terrestrial ecosystem productivity is thought to be the dominant driver behind the current land carbon sink. The strength of the fertilization effect on GPP differs strongly between LSMs (Friedlingstein et al., 2014). We used proxies of the historical increase in GPP for an indirect model evaluation of the CO_2 fertilization effect from COS and deuterium measurements of herbarium samples (Ehlers et al., 2015; Campbell et al., 2017), and we found that ORCHIDEE-CNP has smaller and more realistic E_{CO_2} than the same model without nutrients (Fig. 5), in particular for C_3 plants and in boreal regions (Fig. S5). Both ORCHIDEE-CNP and ORCHIDEE simulated an E_{CO_2} for C_4 grass of ~ 1, as the carboxylation of C_4 plants is weakly influenced by elevated CO_2 (Osmond et al., 1982; Pearcy and Ehleringer, 1984; Bowes, 1993). This indicates that the inclusion of N and P constraints on GPP leads to a more realistic CO_2 fertilization effect in ORCHIDEE-CNP.

5.3 Ecosystem nutrient turnover and openness indicates model biases in boreal phosphorus availability

The capacity of ecosystems to sequester and store additional carbon depends on their ability to supply nutrients for the buildup of organic matter. Enhanced internal nutrient recycling and the accumulation of nutrients over time in ecosystems are theoretically possible mechanisms through which nutrients can be supplied. Therefore, it is important for simulating changes in land carbon storage on decadal timescales and longer that models capture the dependency of ecosystem production to external nutrient sources (i.e., openness of N and P cycles) (Cleveland et al., 2013) and the residence time of nutrients within ecosystems. Besides being related to each other, openness and residence times are also related to the inflows and outflows of nutrients (Eqs. 9 and 10) as well as the turnover time of nutrients in specific ecosystem compartments.

We find that ORCHIDEE-CNP simulates the openness of nutrient cycles, including differences among biomes that are close to estimates from the model–data fusion framework GOLUM-CNP (Fig. 6; Sect. 4.4). There are differences in the openness of N (O_N) in tropical natural biomes and the openness of P (O_P) in central Africa, which are related to lower but more realistic tropical BNF in ORCHIDEE-CNP (Sect. S4) and a difference in the prescribed P deposition compared to GOLUM-CNP. Simulated nutrient losses due to aquatic transport are generally in good agreement with independent estimates (Sect. S5).

Residence times of N and P (τ_N and τ_P) in ORCHIDEE-CNP compare generally well to estimates from GOLUM-CNP: ORCHIDEE-CNP simulates shorter τ_N and τ_P in tropical and temperate biomes compared to boreal ones, in line with GOLUM-CNP (Fig. 7). This indicates that ORCHIDEE-CNP is able to reproduce large-scale patterns in the nutrient residence time of biomes, with one exception. In boreal regions, we find that ORCHIDEE-CNP simulates higher τ_P for BOCF due to the higher standing P stocks of biomass and soil organic matter than in GOLUM-CNP (Fig. S1). This indicates that ORCHIDEE-CNP is likely underestimating P availability in boreal regions. The underlying processes of biochemical P mineralization (Sect. S7) and sorption of P to soil particles (Sect. S6) are reasonably well captured in ORCHIDEE-CNP.

5.4 Model biases in stoichiometry indicate need for refinement of process representation

Leaf and soil stoichiometry are key indexes to characterize the relative ecosystem N and P limitation (e.g., Güsewell, 2004). Measurements show a decrease in foliar N : P ratios from low to high latitudes in natural ecosystems (McGroddy et al., 2004; Reich and Oleksyn, 2004; Kerkhoff et al., 2005). This is seen as evidence for tropical vegetation being generally more P- than N-limited, in contrast to extratropical vegetation (Reich and Oleksyn, 2004). The observed trend of foliar N : P ratios was not reproduced by ORCHIDEE-CNP (Fig. 8), which simulated a flat foliar N : P latitudinal profile. In contrast to the majority of global models, wherein leaf N : P ratios are either prescribed (Goll et al., 2012) or vary within a PFT-specific range (Wang et al., 2010), we conservatively assumed a globally uniform range to let the model freely calculate leaf N : P stoichiometry. It is not trivial to pin down the failure of the model to capture the latitudinal trend in leaf N : P ratios, which could be due to (1) omitted variability in leaf P resorption efficiencies, which varies among biomes between 46 % and 66.6 % (Reed et al., 2012) but was set to 65 % in ORCHIDEE-CNP, (2) the simplistic parameterization of nutrient investment into different plant tissues, (3) and the omission of the diversity of nutrient acquisition pathways (e.g., mycorrhizal association) and rooting strategies (Warren et al., 2015). Testing new formulations for plant growth based on optimality principles (Kvakić et al., 2020) and the refinement of nutrient acquisition pathways (Sulman et al., 2017) are ways forward to improve the model.

Regarding soil stoichiometry, measurements show that tropical biomes have lower soil C : N and higher soil C : P and soil N : P than temperate biomes (Tipping et al., 2016), echoing the pattern of leaf stoichiometry. ORCHIDEE-CNP fails to capture these patterns (Fig. 9). Modeled soil N : P and C : P for tropical forests are comparable to measurements but are too low in temperate forest, tropical forest and temperate grass, which is most likely related to a nutrient immobilization that is too strong in accumulating soil organic matter (Figs. S1); this tends to push systems into P limitation rather than N limitation as O_N is larger than O_P (Fig. 6). In general, the spread in soil P concentration is well represented by ORCHIDEE-CNP. The rudimentary representation of organic matter decomposition and the lack of nutrient effects on decomposer carbon use efficiency (see Zhang et al., 2018, for possible improvements; Sect. 5.5) are likely contributing to the biases. New developments, including explicit representation of decomposer communities and soil organic matter stabilization (Zhang et al., 2020), will be included in the next model version.

5.5 Nutrient effects on carbon cycling

In the following we discuss the implications for the simulated carbon fluxes of changes in plant resource use efficiencies and the sensitivity of plant productivity to increasing CO_2 due to the inclusion of nutrient cycles. We link biases in the simulated carbon fluxes to biases in nutrient cycling, which allows us to prioritize follow-up model development.

5.5.1 Inclusion of nutrient cycling improves the interannual variability of GPP

To what extent nutrient effects on vegetation affect the sensitivity of ecosystem CO_2 fluxes to climatic variation is unclear (Goll et al., 2018). For instance, drought can reduce nutrient use by decreasing GPP, but it also slows down decomposition, which supplies nutrients for plant uptake. Further, N : P stoichiometry is also strongly modified by drought and warming towards increased N : P in whole plant biomass (Yuan and Chen, 2015). Here we found that the inclusion of N and P cycles in ORCHIDEE affects the interannual variability of GPP for all vegetation types. In ORCHIDEE-CNP, the interannual variation (IAV) of GPP is better correlated with that of observation-based datasets than in ORCHIDEE globally and for the NH, but less correlated for other regions (Fig. 10a). Observation-based GPP estimates are uncertain, as some of them ignore soil-moisture-induced reductions of GPP during drought (Stocker et al., 2019), as well as soil thaw and snow-related effects (Jiang and Ryu, 2016). Thus, at the moment, it is difficult to falsify one model version over another and to constrain nutrient effects on the variation of GPP based on current observation-based GPP.

In order to further explore the underlying reasons for the general improvement in the IAV of GPP due to the inclusion of nutrient cycles, we analyzed the sensitivity of GPP anomalies to anomalies of temperature (S_T), precipitation (S_P) and shortwave radiation (S_R), all with mean annual values (Sect. S1I). We found that S_P by ORCHIDEE-CNP compares well with BESS GPP and MTE GPP, while it is overestimated in ORCHIDEE (Figs. S7 and S8). Thus, the difference in S_P is likely the major reason for the differences in IAV in NH between model versions, as S_T and S_R show only minor differences there. This provides confidence that the improvement in IAV of GPP in the NH is due to an improved sensitivity towards a climatic driver (i.e., S_P). For tropical regions, ORCHIDEE-CNP simulates more realistic S_P but higher biases in S_R than in ORCHIDEE, while observation-based estimates of S_T disagree on the sign and model versions show only minor differences (Fig. S7). Therefore, the deterioration of the IAV of tropical GPP by the inclusion of nutrient cycles is likely caused by enhanced biases in S_R due to a lowering of LUE of GPP (Sects. 4.2 and 5.1).

5.5.2 Inclusion of nutrient cycling deteriorates phenology and seasonality of GPP

The performance in reproducing seasonal variations of GPP was deteriorated by the inclusion of N and P nutrient cycles in ORCHIDEE (Fig. 10b). We found that biases in GPP are

related to biases in the seasonality of the LAI introduced in ORCHIDEE-CNP (Figs. S9a and S10a). For the NH, the delayed increase in LAI in ORCHIDEE-CNP could be partly caused by nutrient shortage during the first half of the growing season, as indicated by the increasing leaf nutrient concentration throughout the growing season (Fig. S11). Several factors could lead to a supply of nutrients that is too low at the beginning of the growing season: an insufficient internal plant nutrient reserve due to resorption of nutrients that is too low prior to leaf shedding or an underestimation of nutrient uptake during the dormant season, an insufficient investment into root growth to acquire nutrients, and an overestimation of soil nutrient losses during the dormant season leaving the soil nutrient depleted at the beginning of the growing season. Many of the related processes (e.g., root phenology, mineralization, nutrient resorption, growth allocation) are only rudimentarily represented. For tropical regions, ORCHIDEE-CNP simulates a quasi-flat seasonal cycle of GPP, in contrast to a peak of GPP during the wet season in MTE GPP and BESS GPP, which is correctly captured by ORCHIDEE (Fig. S9b, c). The reduction of seasonal GPP in ORCHIDEE-CNP compared to ORCHIDEE is more pronounced in the dry season ($\sim 100\,\mathrm{g\,C\,m^{-2}}$) than in the wet season (Fig. S9b, c), concurrent with a larger reduction of LAI in the dry season (Fig. S10b, c). Tropical phenology is currently only rudimentarily represented in ORCHIDEE(-CNP) (Chen et al., 2020), causing a suboptimal allocation of nutrients to leaves that could cause the biases in the seasonal cycle of GPP and LAI. Model–data assimilation of phenology (Williams et al., 2009; MacBean et al., 2018; Bacour et al., 2019) and efforts to better characterize processes related to plant resource investment into different tissues and symbioses (Prentice et al., 2015; Warren et al., 2015; Jiang et al., 2019) as well as leaf age effects during the year for evergreen forests (Chen et al., 2020) should help to reduce tropical phenology biases in future versions of ORCHIDEE-CNP.

5.5.3 Inclusion of nutrient cycling leads to an underestimation of the land carbon sink

Current LSMs unanimously conclude that CO_2 fertilization is the main driver of the land carbon sink and its trend (Friedlingstein et al., 2014), but it remains unclear to what extent other drivers (i.e., climate change, land management, nutrient deposition) contribute to the sink as well. Also, it remains unclear how commonly omitted dynamics (climate and management induced effects on tree mortality, nutrients) lead to overestimation of the contribution of CO_2 fertilization in models (Ellsworth et al., 2017; Fleischer et al., 2019). ORCHIDEE-CNP simulates a land carbon sink over the past decades that is lower than other dynamic global vegetation models (DGVMs) and atmospheric inversions (Fig. S12), despite the fact that the response of GPP to CO_2 in ORCHIDEE-CNP is in line with proxy data (Fig. 5; Sect. 5.2). In particular, the NH carbon sink, which has persistently increased over the last 50 years (Ciais et al., 2019), is strongly underestimated. The few Free Air Carbon Enrichment (FACE) studies that have experimentally applied elevated CO_2 levels in mature stands found no increase in biomass production (Bader et al., 2013; Klein et al., 2016; Körner et al., 2005; Sigurdsson et al., 2013; Ellsworth et al., 2017); thus, an increase in GPP does not necessarily translate into an increase in biomass production, whereas in most DGVMs wherein mortality is constant and growth follows GPP, biomass production is inevitably coupled to GPP. Based on upscaling of data from FACE experiments, Terrer et al. (2019) suggested that the effect of elevated CO_2 on biomass may be severely overestimated (on average by a factor of 3.6) in LSMs that ignore nutrients. It would be tempting to conclude from this study that ORCHIDEE-CNP is "right" in its underestimation of the carbon sink, whereas other models are "wrong" because they miss processes such as forest regrowth (Pugh et al., 2019) from, e.g., decreased harvesting pressure (Ciais et al., 2008) and thus have a realistic NH land sink for the wrong reasons. We also showed that ORCHIDEE-CNP underestimates peak GPP (Fig. S12b) and overestimates P limitations in the NH (Sect. 5.1, 5.3 and 5.4); thus, another explanation is that the NH sink in this study is too low because of P limitations that are too strong in this region. These two hypotheses explaining why we underestimate the NH sink (missing forest regrowth vs. overly strong nutrient limitations in the NH) are examined below.

The overly small NH carbon sink in ORCHIDEE-CNP may be explained by an immobilization of nutrients that is too strong in accumulating nutrient-rich organic matter, which leads to a reduction of plant-available nutrients, the so-called "progressive nutrient limitation" proposed by Luo et al. (2004), and subsequently to reduced biomass production. The amount of accumulated N and P immobilized into SOM in the NH during 1850–2016 reaches up to $75.3\,\mathrm{g\,N\,m^{-2}}$ and $2.4\,\mathrm{g\,P\,m^{-2}}$, respectively, which is twice as much as the accumulated respective nutrient inputs to ecosystems in this region during the same period ($37.8\,\mathrm{g\,N\,m^{-2}}$ and $1.6\,\mathrm{g\,P\,m^{-2}}$; Figs. S13 and S14). This suggests a strong progressive nutrient limitation in the model. The omission of nutrient controls on litter and SOM decomposition in the soil module of ORCHIDEE-CNP could have favored the immobilization of nutrients in accumulating SOM (Zhang et al., 2018). Microbe incubation and N fertilization experiments showed that a low availability of nutrients can hamper the buildup of SOM as more carbon gets respired by decomposers due to elevated energetic requirements of processing low-quality substrate (Recous et al., 1995; Janssens et al., 2010; Allison et al., 2009) and overall lower microbial activity (Wang et al., 2011; Knorr et al., 2005). Uncertainties with respect to the capability of ecosystems to up-regulate P mineralization when P becomes scarce could have contributed to the decline in plant-available nutrients with increasing SOM stocks. The inclusion of nutrient effects on decomposition and microbial dynamics in ORCHIDEE-CNP is ongoing (Zhang et al., 2018,

2020), but the lack of quantification of the ability of ecosystems to enhance P recycling hampers model developments.

The overly small NH carbon sink in ORCHIDEE-CNP may also be explained by the lack of representation of effects of forest age and management on biomass turnover and biomass production efficiency (i.e., CUE). Pugh et al. (2019) found that old-growth forests in the NH have a much smaller C sink than regrowing forests ($< 0.1\,\mathrm{Pg\,C\,yr^{-1}}$ compared to $0.86\,\mathrm{Pg\,C\,yr^{-1}}$) for the period 2001–2010. Forest management effects on biomass production efficiency and biomass turnover are only rudimentarily represented in ORCHIDEE(-CNP). ORCHIDEE-CNP prescribes constant tree mortality rates (i.e., the fraction of total carbon in wood lost to litter), whereas in reality tree mortality rates change with management and climate conditions (Peng et al., 2011). Moreover, ORCHIDEE(-CNP) omits the effect of forest age on C uptake. Compared to data-driven estimates for C storage (Sect. S1G and S1H), ORCHIDEE-CNP simulates a higher global aboveground forest biomass (387 Pg C; 283 Pg C for GlobBiomass, Santoro, 2018; 221 Pg C for GEOCARBON, Operational Global Carbon Observing System, Avitabile et al., 2016; Fig. S2) but lower global soil organic carbon (801 Pg C; 4387 Pg C for Soilgrids, Hengl et al., 2017; 1680 Pg C for GSDE – Global Soil Dataset for use in Earth system models, Shangguan et al., 2014; Fig. S3).

6 Concluding remarks

In this study, we evaluated the performance of ORCHIDEE-CNP and found that the model has sufficient skills in capturing observed patterns in (1) vegetation resource use efficiencies, (2) CO_2 vegetation fertilization, (3) ecosystem N and P openness and turnover, and (4) leaf and soil stoichiometry. The inclusion of nutrients improves the simulation of the sensitivity of plant productivity to increasing CO_2 and to interannual variation in precipitation. However, the nutrient-enabled version cannot capture the current land carbon sink in the NH. This suggests that either the land carbon sink might be less a consequence of the CO_2 fertilization effect than of other processes that are currently not well resolved in global models (e.g., biomass turnover, land management) or that ORCHIDEE-CNP underestimates the ability of ecosystems to circumvent nutrient constraints on biomass built up under elevated CO_2. We propose the following focus to improve ORCHIDEE in the next model versions: (1) refine the canopy light absorption processes; (2) use model–data assimilation frameworks (like ORCHIDAS) to better calibrate root phenology, mineralization, nutrient resorption and growth allocation; (3) better represent soil processes related to decomposition, stabilization of soil organic matter (e.g., Zhang et al., 2018, 2020) and inorganic P transformation (e.g., Helfenstein et al., 2020); and (4) refine the dynamics of biomass turnover and biomass production efficiency, including effects of forest management and climate. Continued improvements of nutrient cycle representations will further reduce uncertainties in predicting the land carbon sink under climate change and rising atmospheric CO_2.

Code and data availability. The source code is freely available online via the following address: http://forge.ipsl.jussieu.fr/orchidee/wiki/GroupActivities/CodeAvalaibilityPublication/ORCHIDEE-CN-P_v1.2_r5986 (last access: 8 May 2020, Goll, 2020). Please contact the corresponding author if you plan an application of the model and envisage longer-term scientific collaboration.

Primary data and scripts used in the analysis and other supplementary information that may be useful in reproducing the authors' work can be obtained at https://doi.org/10.17632/f54v9zcgbf.1(Sun et al., 2020).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/gmd-14-1-2021-supplement.

Author contributions. YS and DSG carried out the simulation of ORCHIDEE-CNP. YS, JC, JH, VN, YW and HY analyzed the model outputs. YS, DSG, and PC prepared the paper with contributions from all coauthors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Yan Sun, Daniel S. Goll, Philippe Ciais, Haicheng Zhang and Yuanyuan Huang are funded by the IMBALANCE-P project of the European Research Council (ERC-2013-SyG610028). Julian Helfenstein is supported by the Swiss National Science Foundation (project number 200021_162422).

Financial support. This research has been supported by the IMBALANCE-P project of the European Research Council (grant no. ERC-2013-SyG610028).

Review statement. This paper was edited by Chiel van Heerwaarden and reviewed by two anonymous referees.

References

Allison, S. D., LeBauer, D. S., Ofrecio, M. R., Reyes, R., Ta, A.-M., and Tran, T. M.: Low levels of nitrogen addition stimulate decomposition by boreal forest fungi, Soil Biol. Biochem., 41, 293–302, https://doi.org/10.1016/j.soilbio.2008.10.032, 2009.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and

Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Rev. Geophys., 53, 785–818, https://doi.org/10.1002/2015RG000483, 2015.

Avitabile, V., Herold, M., Heuvelink, G. B. M., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., Bayol, N., Berry, N. J., Boeckx, P., de Jong, B. H. J., DeVries, B., Girardin, C. A. J., Kearsley, E., Lindsell, J. A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E. T. A., Nagy, L., Qie, L., Quinones, M. J., Ryan, C. M., Ferry, S. J. W., Sunderland, T., Laurin, G. V., Gatti, R. C., Valentini, R., Verbeeck, H., Wijaya, A., and Willcock, S.: An integrated pan-tropical biomass map using multiple reference datasets, Glob Change Biol, 22, 1406–1420, https://doi.org/10.1111/gcb.13139, 2016.

Bacour, C., Maignan, F., Peylin, P., MacBean, N., Bastrikov, V., Joiner, J., Köhler, P., Guanter, L., and Frankenberg, C.: Differences Between OCO-2 and GOME-2 SIF Products From a Model-Data Fusion Perspective, J. Geophys. Res.-Biogeosci., 124, 3143–3157, https://doi.org/10.1029/2018JG004938, 2019.

Bader, M. K. F., Leuzinger, S., Keel, S. G., Siegwolf, R. T., Hagedorn, F., Schleppi, P., and Körner, C.: Central European hardwood trees in a high-CO_2 future: synthesis of an 8-year forest canopy CO_2 enrichment project, J. Ecol., 101, 1509–1519, https://doi.org/10.1111/1365-2745.12149, 2013.

Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., and Kirches, G.: Consistent global land cover maps for climate modelling communities: current achievements of the ESA's land cover CCI, Proceedings of the ESA Living Planet Symposium, Edimburgh, 2013, 9–13, 2013.

Bowes, G.: Facing the inevitable: plants and increasing atmospheric CO_2, Annu. Rev. Plant. Biol., 44, 309–332, https://doi.org/10.1146/annurev.arplant.44.1.309, 1993.

Brovkin, V. and Goll, D.: Land unlikely to become large carbon source, Nat. Geosci., 8, 893–893, https://doi.org/10.1038/ngeo2598, 2015.

Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini, B. E. L., Cornelissen, J. H. C., Craine, J. M., Craven, D., de Vries, F. T., Díaz, S., Domingues, T. F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W. N., Hickler, T., Jansen, S., Kramer, K., Kraft, N. J. B., Kurokawa, H., Laughlin, D. C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N. A., Spasojevic, M. J., Sosinski, E., Thornton, P. E., Valladares, F., van Bodegom, P. M., Williams, M., Wirth, C., and Reich, P. B.: Mapping local and global variability in plant trait distributions, P. Natl. Acad. Sci. USA, 114, E10937–E10946, https://doi.org/10.1073/pnas.1708984114, 2017.

Campbell, J., Berry, J., Seibt, U., Smith, S. J., Montzka, S., Launois, T., Belviso, S., Bopp, L., and Laine, M.: Large historical growth in global terrestrial gross primary production, Nature, 544, 84–87, https://doi.org/10.1038/nature22030, 2017.

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

Chen, X., Maignan, F., Viovy, N., Bastos, A., Goll, D., Wu, J., Liu, L., Yue, C., Peng, S., and Yuan, W.: Novel representation of leaf phenology improves simulation of Amazonian evergreen forest photosynthesis in a land surface model, J. Adv. Model Earth Syst., 12, e2018MS001565, https://doi.org/10.1029/2018MS001565, 2020.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., Chédin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res.-Atmos., 110, https://doi.org/10.1029/2005JD006390, 2005.

Ciais, P., Schelhaas, M.-J., Zaehle, S., Piao, S., Cescatti, A., Liski, J., Luyssaert, S., Le-Maire, G., Schulze, E.-D., Bouriaud, O., Freibauer, A., Valentini, R., and Nabuurs, G. J.: Carbon accumulation in European forests, Nature, 1, 425–429, https://doi.org/10.1038/ngeo233, 2008.

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

Cleveland, C. C., Houlton, B. Z., Smith, W. K., Marklein, A. R., Reed, S. C., Parton, W., Del Grosso, S. J., and Running, S. W.: Patterns of new versus recycled primary production in the terrestrial biosphere, P. Natl. Acad. Sci. USA, 110, 12733–12737, https://doi.org/10.1073/pnas.1302768110, 2013.

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

Ellsworth, D. S., Anderson, I. C., Crous, K. Y., Cooke, J., Drake, J. E., Gherlenda, A. N., Gimeno, T. E., Macdonald, C. A., Medlyn, B. E., Powell, J. R., Tjoelker, M. G., and Reich, P. B.: Elevated CO_2 does not increase eucalypt forest productivity on a low-phosphorus soil, Nat. Clim. Chang., 7, 279–282, https://doi.org/10.1038/nclimate3235, 2017.

Elser, J. J., Bracken, M. E., Cleland, E. E., Gruner, D. S., Harpole, W. S., Hillebrand, H., Ngai, J. T., Seabloom, E. W., Shurin, J. B., and Smith, J. E.: Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems, Ecol. Lett., 10, 1135–1142, https://doi.org/10.1111/j.1461-0248.2007.01113.x, 2007.

FAO/IIASA/ISRIC/ISSCAS/JRC: Harmonized World Soil Database (version 1.2), available at: http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/ (last access: 13 May 2015), 2012.

Fernández-Martínez, M., Vicca, S., Janssens, I. A., Sardans, J., Luyssaert, S., Campioli, M., Chapin III, F. S., Ciais, P., Malhi, Y., Obersteiner, M., Papale, D., Piao, S. L., Reichstein, M., Rodà, F., and Peñuelas, J.: Nutrient availability as the key regulator of global forest carbon balance, Nat. Clim. Chang., 4, 471–476, https://doi.org/10.1038/nclimate2177, 2014.

Fleischer, K., Rammig, A., De Kauwe, M. G., Walker, A. P., Domingues, T. F., Fuchslueger, L., Garcia, S., Goll, D. S., Grandis, A., Jiang, M., Haverd, V., Hofhansl, F., Holm, J. A., Kruijt, B., Leung, F., Medlyn, B. E., Mercado, L. M., Norby, R. J., Pak, B., von Randow, C., Quesada, C. A., Schaap, K. J., Valverde-

Barrantes, O. J., Wang, Y., Yang, X., Zaehle, S., Zhu, Q., and Lapola, D. M.: Amazon forest response to CO_2 fertilization dependent on plant phosphorus acquisition, Nat. Geosci., 12, 736–741, https://doi.org/10.1038/s41561-019-0404-9, 2019.

Freschet, G. T., Cornelissen, J. H., Van Logtestijn, R. S., and Aerts, R.: Evidence of the "plant economics spectrum" in a subarctic flora, J. Ecol., 98, 362–373, https://doi.org/10.1111/j.1365-2745.2009.01615.x, 2010.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J. Clim., 27, 511–526, https://doi.org/10.1175/JCLI-D-12-00579.1, 2014.

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

Gärdenäs, A. I., Ågren, G. I., Bird, J. A., Clarholm, M., Hallin, S., Ineson, P., Kätterer, T., Knicker, H., Nilsson, S. I., Näsholm, T., Ogle, S., Paustian, K., Persson, T., and Stendahl, J.: Knowledge gaps in soil carbon and nitrogen interactions–from molecular to global scale, Soil Biol. Biochem., 43, 702–717, https://doi.org/10.1016/j.soilbio.2010.04.006, 2011.

Gill, A. L. and Finzi, A. C.: Belowground carbon flux links biogeochemical cycles and resource-use efficiency at the global scale, Ecol. Lett., 19, 1419–1428, https://doi.org/10.1111/ele.12690, 2016.

Global Soil Data Task Group: Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS), Data set, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, 2000.

Gobron, N., Pinty, B., Aussedat, O., Chen, J., Cohen, W. B., Fensholt, R., Gond, V., Huemmrich, K. F., Lavergne, T., Mélin, F., Privette, J. L., Sandholt, I., Taberner, M., Turner, D. P., Verstraete, M. M., and Widlowski, J.: Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations, J. Geophys. Res.-Atmos, 111, D13110, https://doi.org/10.1029/2005JD006511, 2006a.

Gobron, N., Pinty, B., Taberner, M., Mélin, F., Verstraete, M., and Widlowski, J.-L.: Monitoring the photosynthetic activity of vegetation from remote sensing data, Adv. Space Res., 38, 2196–2202, https://doi.org/10.1016/j.asr.2003.07.079, 2006b.

Goll, D. S., Brovkin, V., Parida, B. R., Reick, C. H., Kattge, J., Reich, P. B., van Bodegom, P. M., and Niinemets, Ü.: Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling, Biogeosciences, 9, 3547–3569, https://doi.org/10.5194/bg-9-3547-2012, 2012.

Goll, D. S., Moosdorf, N., Hartmann, J., and Brovkin, V.: Climate-driven changes in chemical weathering and associated phosphorus release since 1850: Implications for the land carbon balance, Geophys. Res. Lett., 41, 3553–3558, https://doi.org/10.1002/2014GL059471, 2014.

Goll, D. S., Vuichard, N., Maignan, F., Jornet-Puig, A., Sardans, J., Violette, A., Peng, S., Sun, Y., Kvakic, M., Guimberteau, M., Guenet, B., Zaehle, S., Penuelas, J., Janssens, I., and Ciais, P.: A representation of the phosphorus cycle for ORCHIDEE (revision 4520), Geosci. Model Dev., 10, 3745–3770, https://doi.org/10.5194/gmd-10-3745-2017, 2017a.

Goll, D. S., Joetzjer, E., Huang, M., and Ciais, P.: Low phosphorus availability decreases susceptibility of tropical primary productivity to droughts, Geophys. Res. Lett., 45, 8231–8240, https://doi.org/10.1029/2018GL077736, 2018.

Güsewell, S.: N : P ratios in terrestrial plants: variation and functional significance, New Phytol., 164, 243–266, https://doi.org/10.1111/j.1469-8137.2004.01192.x, 2004.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset, Int. J. Climatol., 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Hartmann, J. and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at the Earth surface, Geochem. Geophys. Geosyst., 13, Q12004, https://doi.org/10.1029/2012GC004370, 2012.

Hartmann, J., Moosdorf, N., Lauerwald, R., Hinderer, M., and West, A. J.: Global chemical weathering and associated P-release—The role of lithology, temperature and soil properties, Chem. Geol., 363, 145–163, https://doi.org/10.1016/j.chemgeo.2013.10.025, 2014.

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

Helfenstein, J., Jegminat, J., McLaren, T. I., and Frossard, E.: Soil solution phosphorus turnover: derivation, interpretation, and insights from a global compilation of isotope exchange kinetic studies, Biogeosciences, 15, 105–114, https://doi.org/10.5194/bg-15-105-2018, 2018.

Helfenstein, J., Pistocchi, C., Oberson, A., Tamburini, F., Goll, D. S., and Frossard, E.: Estimates of mean residence times of phosphorus in commonly considered inorganic soil phosphorus pools, Biogeosciences, 17, 441–454, https://doi.org/10.5194/bg-17-441-2020, 2020.

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G., Walsh, M. G. and Gonzalez, M. R.: SoilGrids1km – global soil information based on automated mapping, PLoS One, 9, e114788, https://doi.org/10.1371/journal.pone.0105992, 2014.

Hodapp, D., Hillebrand, H., and Striebel, M.: "Unifying" the concept of resource use efficiency in ecology, Front. Ecol. Evol., 6, 233, https://doi.org/10.3389/fevo.2018.00233, 2019.

Hungate, B. A., Dukes, J. S., Shaw, M. R., Luo, Y., and Field, C. B.: Nitrogen and climate change, Science, 302, 1512–1513, https://doi.org/10.1126/science.1091390, 2003.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Clim. Change, 109, 117–161, https://doi.org/10.1007/s10584-011-0153-2, 2011.

Ito, A. and Oikawa, T.: Global mapping of terrestrial primary productivity and light-use efficiency with a process-based model, Global Environmental Change in the Ocean and on Land, 343–358, 2004.

Janssens, I. A. and Vicca, S.: Soil carbon breakdown, Nat. Geosci., 3, 823–824, https://doi.org/10.1038/ngeo1024, 2010.

Jiang, C. and Ryu, Y.: Multi-scale evaluation of global gross primary productivity and evapotranspiration products derived from Breathing Earth System Simulator (BESS), Remote Sens. Environ., 186, 528–547, https://doi.org/10.1016/j.rse.2016.08.030, 2016.

Jiang, M., Caldararu, S., Zaehle, S., Ellsworth, D. S., and Medlyn, B. E.: Towards a more physiological representation of vegetation phosphorus processes in land surface models, New Phytol., 222, 1223–1229, https://doi.org/10.1111/nph.15688, 2019.

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

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Damiano, G., 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, J. Geophys. Res.-Biogeo., 116, G00J07, https://doi.org/10.1029/2010JG001566, 2011.

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

Kerkhoff, A. J., Enquist, B. J., Elser, J. J., and Fagan, W. F.: Plant allometry, stoichiometry and the temperature-dependence of primary productivity, Glob. Ecol. Biogeogr., 14, 585–598, https://doi.org/10.1111/j.1466-822x.2005.00187.x, 2005.

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene – HYDE 3.2, Earth Syst. Sci. Data, 9, 927–953, https://doi.org/10.5194/essd-9-927-2017, 2017a.

Klein Goldewijk, K., Dekker, S. C., and van Zanden, J. L.: Per-capita estimations of long-term historical land use and the consequences for global change research, J. Land Use Sci., 12, 313–337, https://doi.org/10.1080/1747423X.2017.1354938, 2017b.

Klein, T., Siegwolf, R. T., and Körner, C.: Belowground carbon trade among tall trees in a temperate forest, Science, 352, 342–344, https://doi.org/10.1126/science.aad6188, 2016.

Knorr, M., Frey, S., and Curtis, P.: Nitrogen additions and litter decomposition: A meta-analysis, Ecology, 86, 3252–3257, https://doi.org/10.1890/05-0150, 2005.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, J. Meteorol. Soc. Jpn, Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Körner, C., Asshoff, R., Bignucolo, O., Hättenschwiler, S., Keel, S. G., Peláez-Riedl, S., Pepin, S., Siegwolf, R. T., and Zotz, G.: Carbon flux and growth in mature deciduous forest trees exposed to elevated CO_2, Science, 309, 1360–1362, https://doi.org/10.1126/science.1113977, 2005.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Glob. Biogeochem. Cy., 19, GB1015, https://doi.org/10.1029/2003GB002199, 2005.

Kvakić, M., Tzagkarakis, G., Pellerin, S., Ciais, P., Goll, D., Mollier, A., and Ringeval, B.: Carbon and phosphorus allocation in annual plants: an optimal functioning approach, Front. Plant Sci., 11, 149 pp., https://doi.org/10.3389/fpls.2020.00149, 2020.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, 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., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth Syst. Sci. Data, 10, 2141–2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

Liu, Y., Piao, S., Gasser, T., Ciais, P., Yang, H., Wang, H., Keenan, T. F., Huang, M., Wan, S., and Song, J.: Field-experiment constraints on the enhancement of the terrestrial carbon sink by CO_2 fertilization, Nat. Geosci., 12, 809–814, https://doi.org/10.1038/s41561-019-0436-1, 2019.

Lu, C. and Tian, H.: Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hot spots and nutrient imbalance, Earth Syst. Sci. Data, 9, 181–192, https://doi.org/10.5194/essd-9-181-2017, 2017.

Lun, F., Liu, J., Ciais, P., Nesme, T., Chang, J., Wang, R., Goll, D., Sardans, J., Peñuelas, J., and Obersteiner, M.: Global and regional phosphorus budgets in agricultural systems and their implications for phosphorus-use efficiency, Earth Syst. Sci. Data, 10, 1–18, https://doi.org/10.5194/essd-10-1-2018, 2018.

Luo, Y. Q., Su, B., Currie, W. S., Dukes, J. E., Finzi, A., Hartwig, U., Hungate, B., Mc Murtrie, R. E., Oren, R., Parton, W. J., Pataki, D. E., Rebecca Shaw, M., Zak, D. R., and Field, C. B.: Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide, BioScience, 54, 731–739, https://doi.org/10.1641/0006-3568(2004)054[0731:PNLOER]2.0.CO;2, 2004.

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., Köhler, P., Gómez-Dans, J., and Disney, M.: Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data, Sci. Rep., 8, 1973, https://doi.org/10.1038/s41598-018-20024-w, 2018.

Malyshev, A. V. and Henry, H. A.: N uptake and growth responses to sub-lethal freezing in the grass Poa pratensis L, Plant Soil, 360, 175–185, https://doi.org/10.1007/s11104-012-1233-4, 2012.

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

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

Melillo, J. M., Butler, S., Johnson, J., Mohan, J., Steudler, P., Lux, H., Burrows, E., Bowles, F., Smith, R., Scott, L., Vario, C., Hill, T., Burton, A., Zhou, Y., and Tang, J.: Soil warming, carbon–nitrogen interactions, and forest carbon budgets, P. Natl. Acad. Sci. USA, 108, 9508–9512, https://doi.org/10.1073/pnas.1018189108, 2011.

Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., Bönisch, G., Campioli, M., Ghattas, J., De Groote, T., Haverd, V., Kattge, J., MacBean, N., Maignan, F., Merilä, P., Penuelas, J., Peylin, P., Pinty, B., Pretzsch, H., Schulze, E. D., Solyga, D., Vuichard, N., Yan, Y., and Luyssaert, S.: A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes, Geosci. Model Dev., 8, 2035–2065, https://doi.org/10.5194/gmd-8-2035-2015, 2015.

Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E., and McMurtrie, R. E.: CO_2 enhancement of forest productivity constrained by limited nitrogen availability, P. Natl. Acad. Sci. USA, 107, 19368–19373, https://doi.org/10.1073/pnas.1006463107, 2010.

Osmond, C., Winter, K., and Ziegler, H.: Functional significance of different pathways of CO_2 fixation in photosynthesis, in: Physiological plant ecology II, Springer, 479–547, 1982.

Pearcy, R. W. and Ehleringer, J.: Comparative ecophysiology of C3 and C4 plants, Plant Cell Environ., 7, 1–13, https://doi.org/10.1111/j.1365-3040.1984.tb01194.x, 1984.

Peng, C., Ma, Z., Lei, X., Zhu, Q., Chen, H., Wang, W., Liu, S., Li, W., Fang, X., and Zhou, X.: A drought-induced pervasive increase in tree mortality across Canada's boreal forests, Nat. Clim. Chang., 1, 467–471, https://doi.org/10.1038/NCLIMATE1293, 2011.

Peng, J., Wang, Y. P., Houlton, B. Z., Dan, L., Pak, B., and Tang, X.: Global carbon sequestration is highly sensitive to model-based formulations of nitrogen fixation, Glob. Biogeochem. Cy., 34, e2019GB006296, https://doi.org/10.1029/2019GB006296, 2019.

Peng, S., Ciais, P., Maignan, F., Li, W., Chang, J., Wang, T., and Yue, C.: Sensitivity of land use change emission estimates to historical land use and land cover mapping, Glob. Biogeochem. Cycle, 31, 626–643, https://doi.org/10.1002/2015GB005360, 2017.

Penuelas, J., Poulter, B., Sardans, J., Ciais, P., Van Der Velde, M., Bopp, L., Boucher, O., Godderis, Y., Hinsinger, P., Llusia, J., Nardin, E., Vicca, S., Obersteiner, M., and Janssens, I. A.: Human-induced nitrogen–phosphorus imbalances alter natural and managed ecosystems across the globe, Nat. Commun., 4, 1–10, https://doi.org/10.1038/ncomms3934, 2013.

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, 652–661, https://doi.org/10.1890/08-2176.1, 2010.

Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., and Zimmermann, N. E.: Plant functional type mapping for earth system models, Geosci. Model Dev., 4, 993–1010, https://doi.org/10.5194/gmd-4-993-2011, 2011.

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

Prentice, I. C., Liang, X., Medlyn, B. E., and Wang, Y.-P.: Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmos. Chem. Phys., 15, 5987–6005, https://doi.org/10.5194/acp-15-5987-2015, 2015.

Pugh, T. A., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., and Calle, L.: Role of forest regrowth in global carbon sink dynamics, P. Natl. Acad. Sci. USA, 116, 4382–4387, https://doi.org/10.1073/pnas.1810512116, 2019.

Recous, S., Robin, D., Darwis, D., and Mary, B.: Soil inorganic N availability: effect on maize residue decomposition, Soil Biol. Biochem., 27, 1529–1538, https://doi.org/10.1016/0038-0717(95)00096-W, 1995.

Reed, S. C., Townsend, A. R., Davidson, E. A., and Cleveland, C. C.: Stoichiometric patterns in foliar nutrient resorption across multiple scales, New Phytol., 196, 173–180, https://doi.org/10.1111/j.1469-8137.2012.04249.x, 2012.

Reich, P. B. and Oleksyn, J.: Global patterns of plant leaf N and P in relation to temperature and latitude, P. Natl. Acad. Sci. USA, 101, 11001–11006, https://doi.org/10.1073/pnas.0403588101, 2004.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, https://doi.org/10.5194/acp-3-1919-2003, 2003.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H.: A continuous satellite-derived measure of global terrestrial primary production, BioScience, 54, 547–560, https://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2, 2004.

Ryu, Y., Verfaillie, J., Macfarlane, C., Kobayashi, H., Sonnentag, O., Vargas, R., Ma, S., and Baldocchi, D. D.: Continuous observation of tree leaf area index at ecosystem scale using upward-

pointing digital cameras, Remote Sens. Environ., 126, 116–125, https://doi.org/10.1016/j.rse.2012.08.027, 2012.

Santoro, M.: GlobBiomass – global datasets of forest biomass, PANGAEA, https://doi.org/10.1594/PANGAEA.894711, 2018.

Shane, M. W., McCully, M. E., Canny, M. J., Pate, J. S., Ngo, H., Mathesius, U., Cawthray, G. R., and Lambers, H.: Summer dormancy and winter growth: root survival strategy in a perennial monocotyledon, New Phytol., 183, 1085–1096, https://doi.org/10.1111/j.1469-8137.2009.02875.x, 2009.

Sigurdsson, B. D., Medhurst, J. L., Wallin, G., Eggertsson, O., and Linder, S.: Growth of mature boreal Norway spruce was not affected by elevated [CO_2] and/or air temperature unless nutrient availability was improved, Tree Physiol., 33, 1192–1205, https://doi.org/10.1093/treephys/tpt043, 2013.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., Yuan, H.: A global soil data set for earth system modeling, J. Adv. Model Earth Syst., 6.1, 249–263, https://doi.org/10.1002/2013MS000293, 2014.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, 2015.

Smith, W. K., Reed, S. C., Cleveland, C. C., Ballantyne, A. P., Anderegg, W. R., Wieder, W. R., Liu, Y. Y., and Running, S. W.: Large divergence of satellite and Earth system model estimates of global terrestrial CO_2 fertilization, Nat. Clim. Chang., 6, 306–310, https://doi.org/10.1038/NCLIMATE2879, 2016.

Soussana, J.-F. and Lemaire, G.: Coupling carbon and nitrogen cycles for environmentally sustainable intensification of grasslands and crop-livestock systems, Agric. Ecosyst. Environ., 190, 9–17, https://doi.org/10.1016/j.agee.2013.10.012, 2014.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., and Peñuelas, J.: Drought impacts on terrestrial primary production underestimated by satellite monitoring, Nat. Geosci., 12, 264–270, https://doi.org/10.1038/s41561-019-0318-6, 2019.

Sullivan, B. W., Smith, W. K., Townsend, A. R., Nasto, M. K., Reed, S. C., Chazdon, R. L., and Cleveland, C. C.: Spatially robust estimates of biological nitrogen (N) fixation imply substantial human alteration of the tropical N cycle, P. Natl. Acad. Sci. USA, 111, 8101–8106, https://doi.org/10.1073/pnas.1320646111, 2014.

Sulman, B. N., Brzostek, E. R., Medici, C., Shevliakova, E., Menge, D. N., and Phillips, R. P.: Feedbacks between plant N demand and rhizosphere priming depend on type of mycorrhizal association, Ecol. Lett., 20, 1043–1053, https://doi.org/10.1111/ele.12802, 2017.

Sun, Y.: Global evaluation of the nutrient enabled version of land surface model ORCHIDEE-CNP (v1.2), Mendeley Data, https://doi.org/10.17632/f54v9zcgbf.1, 2020.

Sun, Y., Peng, S., Goll, D. S., Ciais, P., Guenet, B., Guimberteau, M., Hinsinger, P., Janssens, I. A., Peñuelas, J., Piao, S., Poulter, B., Violette, A., Yang, X., Yin, Y., and Zeng, H.: Diagnosing phosphorus limitations in natural terrestrial ecosystems in carbon cycle models, Earths Future, 5, 730–749, https://doi.org/10.1002/2016EF000472, 2017.

Sun, Y., Goll, D. S., Ciais, P., Peng, S., Margalef, O., Asensio, D., Sardans, J., and Peñuelas, J.: Spatial pattern and environmental drivers of acid phosphatase activity in Europe, Front. Big Data, 2, 51, https://doi.org/10.3389/fdata.2019.00051, 2020.

Terrer, C., Jackson, R. B., Prentice, I. C., Keenan, T. F., Kaiser, C., Vicca, S., Fisher, J. B., Reich, P. B., Stocker, B. D., Hungate, B. A., Peñuelas, J., McCallum, I., Soudzilovskaia, N. A., Cernusak, L. A., Talhelm, A. F., van Sundert, K., Piao, S., Newton, P. C. D., Hovenden, M. J., Blumenthal, D. M., Liu, Y. Y., Müller, C., Winter, K., Field, C. B., Viechtbauer, W., van Lissa, C. J., Hoosbeek, M. R., Watanabe, M., Koike, T., Leshyk, V. O., Polley, H. W., and Franklin, O.: Nitrogen and phosphorus constrain the CO_2 fertilization of global plant biomass, Nat. Clim. Chang., 9, 684–689, https://doi.org/10.1038/s41558-019-0545-2, 2019.

Thornton, P. E., Lamarque, J. F., Rosenbloom, N. A., and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability, Glob. Biogeochem. Cy., 21, GB4018, https://doi.org/10.1029/2006GB002868, 2007.

Thum, T., Caldararu, S., Engel, J., Kern, M., Pallandt, M., Schnur, R., Yu, L., and Zaehle, S.: A new model of the coupled carbon, nitrogen, and phosphorus cycles in the terrestrial biosphere (QUINCY v1.0; revision 1996), Geosci. Model Dev., 12, 4781–4802, https://doi.org/10.5194/gmd-12-4781-2019, 2019.

Tipping, E., Somerville, C. J., and Luster, J.: The C : N : P : S stoichiometry of soil organic matter, Biogeochemistry, 130, 117–131, https://doi.org/10.1007/s10533-016-0247-z, 2016.

Tum, M., Zeidler, J., Günther, K. P., and Esch, T.: Global NPP and straw bioenergy trends for 2000–2014, Biomass Bioenerg., 90, 230–236, https://doi.org/10.1016/j.biombioe.2016.03.040, 2016.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., and Ahl, D.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sens. Environ., 102, 282–292, https://doi.org/10.1016/j.rse.2006.02.017, 2006.

van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: implementation and global carbon balance 2001–2015, Geosci. Model Dev., 10, 2785–2800, https://doi.org/10.5194/gmd-10-2785-2017, 2017.

Vergutz, L., Manzoni, S., Porporato, A., Novais, R. F., and Jackson, R. B.: Global resorption efficiencies and concentrations of carbon and nutrients in leaves of terrestrial plants, Ecol. Monogr., 82, 205–220, https://doi.org/10.1890/11-0416.1, 2012.

Wang, C., Han, G., Jia, Y., Feng, X., Guo, P., and Tian, X.: Response of litter decomposition and related soil enzyme activities to different forms of nitrogen fertilization in a subtropical forest, Ecol. Res., 26, 505–513, https://doi.org/10.1007/s11284-011-0805-8, 2011.

Wang, R., Goll, D., Balkanski, Y., Hauglustaine, D., Boucher, O., Ciais, P., Janssens, I., Penuelas, J., Guenet, B., Sardans, J., Bopp, L., Vuichard, N., Zhou, F., Li, B., Piao, S., Peng, S., Huang, Y., and Tao, S.: Global forest carbon uptake due to nitrogen and phosphorus deposition from 1850 to 2100, Glob. Change Biol., 23, 4854–4872, https://doi.org/10.1111/gcb.13766, 2017.

Wang, Y. P., Law, R. M., and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, Biogeosciences, 7, 2261–2282, https://doi.org/10.5194/bg-7-2261-2010, 2010.

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, Geosci. Model Dev., 11, 3903–3928, https://doi.org/10.5194/gmd-11-3903-2018, 2018.

Wang, Y. P. and Houlton, B. Z.: Nitrogen constraints on terrestrial carbon uptake: Implications for the global carbon-climate feedback, Geophys. Res. Lett., 36, L24403, https://doi.org/10.1029/2009GL041009, 2009.

Warren, J. M., Hanson, P. J., Iversen, C. M., Kumar, J., Walker, A. P., and Wullschleger, S. D.: Root structural and functional dynamics in terrestrial biosphere models–evaluation and recommendations, New Phytol., 205, 59–78, https://doi.org/10.1111/nph.13034, 2015.

Wieder, W. R., Cleveland, C. C., Smith, W. K., and Todd-Brown, K.: Future productivity and carbon storage limited by terrestrial nutrient availability, Nat. Geosci., 8, 441–444, https://doi.org/10.1038/NGEO2413, 2015.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y.-P.: Improving land surface models with FLUXNET data, Biogeosciences, 6, 1341–1359, https://doi.org/10.5194/bg-6-1341-2009, 2009.

Wißkirchen, K., Tum, M., Günther, K. P., Niklaus, M., Eisfelder, C., and Knorr, W.: Quantifying the carbon uptake by vegetation for Europe on a 1 km^2 resolution using a remote sensing driven vegetation model, Geosci. Model Dev., 6, 1623–1640, https://doi.org/10.5194/gmd-6-1623-2013, 2013.

Wright, S. J.: Plant responses to nutrient addition experiments conducted in tropical forests, Ecol. Monogr., 89, e01382, https://doi.org/10.1002/ecm.1382, 2019.

Xu, X., Jain, A. K., and Calvin, K. V.: Quantifying the biophysical and socioeconomic drivers of changes in forest and agricultural land in South and Southeast Asia, Glob. Change Biol., 25, 2137–2151, https://doi.org/10.1111/gcb.14611, 2019.

Yang, X., Thornton, P. E., Ricciuto, D. M., and Post, W. M.: The role of phosphorus dynamics in tropical forests – a modeling study using CLM-CNP, Biogeosciences, 11, 1667–1681, https://doi.org/10.5194/bg-11-1667-2014, 2014.

Yuan, Z. and Chen, H. Y.: Decoupling of nitrogen and phosphorus in terrestrial plants associated with global changes, Nat. Clim. Chang., 5, 465–469, https://doi.org/10.1038/NCLIMATE2549, 2015.

Zaehle, S. and Friend, A.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Glob. Biogeochem. Cy., 24, GB1005, https://doi.org/10.1029/2009GB003521, 2010.

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. Clim., 28, 2494–2511, https://doi.org/10.1175/JCLI-D-13-00776.1, 2015.

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 terrestrial carbon–nitrogen cycle models against observations from two temperate Free-Air CO_2 Enrichment studies, New Phytol., 202, 803–822, https://doi.org/10.1111/nph.12697, 2014.

Zhang, B., Tian, H., Lu, C., Dangal, S. R. S., Yang, J., and Pan, S.: Global manure nitrogen production and application in cropland during 1860–2014: a 5 arcmin gridded global dataset for Earth system modeling, Earth Syst. Sci. Data, 9, 667–678, https://doi.org/10.5194/essd-9-667-2017, 2017.

Zhang, H., Goll, D. S., Manzoni, S., Ciais, P., Guenet, B., and Huang, Y.: Modeling the effects of litter stoichiometry and soil mineral N availability on soil organic matter formation using CENTURY-CUE (v1.0), Geosci. Model Dev., 11, 4779–4796, https://doi.org/10.5194/gmd-11-4779-2018, 2018.

Zhang, H., Goll, D. S., Wang, Y. P., Ciais, P., Wieder, W. R., Abramoff, R., Huang, Y., Guenet, B., Prescher, A. K., Viscarra Rossel, R. A., Barré, P., Chenu, C., Zhou, G., and Tang, X.: Microbial dynamics and soil physicochemical properties explain large-scale variations in soil organic carbon, Glob. Change Biol., 26, 2668–2685, https://doi.org/10.1111/gcb.14994, 2020.

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., and Shen, Y.: Managing nitrogen for sustainable development, Nature, 528, 51–59, https://doi.org/10.1038/nature15743, 2015.

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

Zobler, L.: A World Soil File for Global Climate Modelling. NASA Technical Memorandum 87802, NASA Goddard Institute for Space Studies, New York, New York, USA, 32 pp., 1986.

Supplementary

S1. Evaluation datasets

A: Spin-up and pre-industrial simulations

The historic simulation is preceded by three steps of spin-up to equilibrate plant and soil organic C, N and P to pre-industrial conditions. The spin-up steps are forced by looped climate fields over the period 1901-1920, fixed pre-industrial atmospheric CO_2 (286 ppm) and land cover maps. We run the *spin-up 1* over 200 years with forced C:N:P ratios of plant and soil, followed by two iterations of an analytic spin-up (40 years for each iteration) (Vuichard et al., 2019). This facilitates the accumulation of soil organic C, N and P and approaches the equilibrium for plant productivity. We then run the *spin-up 2* from the last year of the *spin-up 1* with dynamic C:N:P ratios of plant and soils for 10k years, to reach a quasi-equilibrium for plant and soil N and P pools. In the last step, crop management (harvest and fertilizer) was activated for 200 years.

B: Global Observation-based Land-ecosystems Utilization Model of Carbon, Nitrogen and Phosphorus (GOLUM-CNP)

Global Observation-based Land-ecosystems Utilization Model of Carbon, Nitrogen and Phosphorus (GOLUM-CNP) v1.0 is a data-driven modeling of steady-state carbon, nitrogen and phosphorus cycles for present day (2001-2016) conditions. It combines the CARbon DAta MOdel fraMework (CARDAMOM) data-constrained C-cycle analysis with spatially explicit data-driven estimates of N and P inputs and losses and with observed stoichiometric ratios. We extracted the following variables from GOLUM: global GPP and NPP, N and P in- and out-fluxes (e.g. BNF, N and P deposition, N and P leaching) on global and biome scale, biome-specific N and P resource use efficiencies, N and P openness, and N and P turnover rate. GOLUM-CNP only computed fluxes for non-cropland biomes. To compare the global total stocks and flux budgets of C, N and P from GOLUM-CNP with those from ORCHIDEE-CNP, we multiply the stock and flux in each grid cell of GOLUM-CNP by the non-cropland fraction from the MODIS land cover map (version MCD12C1v006, https://lpdaac.usgs.gov/products/mcd12c1v006/) and sum over all the grid cells.

C: Evaluation datasets for the C budget

C1 GPP and NPP from MODIS

MODIS GPP is derived from climate and satellite data. It was estimated at 1 km spatial resolution from a light-use efficiency (LUE) model, as part of the operational MODIS algorithms (Running et al., 2004) using meteorological data from NASA Data Assimilation Office and detailed vegetation information (land cover and FPAR) derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite from 2000 to present (Running et al., 2004; Zhao et al., 2005; Turner et al., 2006).

C2 MTE GPP

MTE GPP is a data-oriented product derived by extrapolating the GPP estimates from a network of flux-tower of 178 sites in space and time using climate data and remotely sensed fAPAR data from 1982 to 2008 (Jung et al., 2009). This latter monthly GPP product covers the period 1982-2011 at 0.5° spatial resolution, but many of the underlying flux tower observations cover much shorter periods.

C3 BESS-GPP

Breathing Earth System Simulator (BESS) is a simplified process-based model that couples atmosphere and canopy radiative transfers, canopy photosynthesis, transpiration, and energy balance. This process-based model uses the MODIS Atmosphere products to calculate atmospheric radiative transfer. This product provides gridded GPP on a monthly scale from 2000-2015 at 0.5° spatial resolution.

C4 BETHY-NPP

Biosphere Energy-Transfer Hydrology (BETHY)/DLR is a Soil-Vegetation-Atmosphere (SVAT) model to simulate photosynthesis of terrestrial ecosystems, which is operated at the German Aerospace Center (DLR). It was adapted from the BETHY scheme to be driven by remote sensing data (leaf area index (LAI) and land cover information) and meteorology (Wißkirchen et al., 2013). This product provides gridded global NPP on a monthly scale from 2000-2009 at 0.008° spatial resolution.

C5 GIMMS-NPP

GIMMS-NPP is a NPP product derived from the MODIS NPP algorithm driven by long-term Global Inventory Modeling and Mapping Studies (GIMMS) FPAR and LAI data (Smith et al., 2016). The uncertainty bounds were given by using a wide range of parameter combinations. This global satellite-derived NPP is available from 1982-2011 with a spatial resolution of 0.5° x 0.5°.

C6 Net biome productivity from GCP and inversion data

Estimates are based on 16 DGVMs involved in GCP Trendy v6 were forced by CRU-NCEP temperature, precipitation, and cloud cover fields (transformed into incoming surface radiation) based on observations and provided on a 0.5° × 0.5° grid and updated to 2016. We compared ORCHIDEE-CNP with Trendy v6 simulations instead of v7 because ORCHIDEE-CNP was part of the v7 ensemble.

We used three sets of CO_2 inversion data from CarbonTracker (van der Laan-Luijkx et al., 2017), Jena CarboScope (Rödenbeck, 2005), and CAMS (Chevallier et al., 2005). They all used atmospheric inversion methods (based on the same Bayesian inversion principles) and atmospheric CO_2 observations from various flask and in situ networks to constrain the location of the combined total surface CO_2 fluxes from all sources. The three systems used different transport models, which was demonstrated to be a driving factor behind differences in atmospheric-based flux estimates, and specifically for the global distribution (Stephens et al., 2007).

D: Evaluation datasets for the N and P budget

D1 Biological Nitrogen fixation (BNF) from Peng et al. (2019)

We used the global gridded BNF (both symbiotic and asymbiotic) from Peng et al., (2019) for 2001-2010. They calculated BNF as a function of soil N, LAI, labile inorganic soil P, temperature, etc (Wang et al., 2007; Houlton et al., 2008) which were constrained by the observation-based estimates of C:N ratios of the various plant pools, soil microbial biomass and organic matter under present conditions.

D2 N and P loads to rivers on basin scale by GlobalNEWS2 model

Global Nutrient Export from Water-Sheds (GLOBNEWS) model generates global spatially explicit of N and P exports by rivers based on a mass-balance approach for the land surface (watershed) and river system for year 2000 (Mayorga et al., 2010).

D3 Inorganic P turnover rate

Soil solution P turnover (Km) were derived from isotopic exchange kinetic (IEK) experiments and aggregated on soil order level (Helfenstein et al., 2018).

D4 Data-driven acid phosphatase activity

This global metric is an extrapolation of 126 European site observations of potential acid phosphatase activity from field samples by using back-propagation artificial network 12 environmental drivers as predictors (Sun et al., 2020).

E: Evaluation datasets for leaf stoichiometry and resource use efficiencies

E1 Gridded leaf N:P ratios

Global maps of leaf N and P concentrations datasets (Butler et al., 2017) were derived from global plant trait database Bayesian modeling. We used 100 sets of estimates of leaf N and P concentrations provided by Butler et al. (2017) to generate 10000 sets of global gridded leaf N:P ratio. Mean values and 25%~75% percentiles of leaf N:P ratios over those 10000 datasets were used to evaluate leaf N:P ratios by ORCHIDEE-CNP.

E2 Remote-sensing fAPAP and PAR datasets

Global SeaWiFS Level-3 generates estimate of the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) over the land surface through an optimized index which values are computed on the basis of top of atmosphere bidirectional reflectance factor values, as well as information on the geometry of illumination and observation (Gobron et al., 2006a, b). This dataset is available for period 1997-2006 with spatial resolution of 0.01^{o}.

E3 Evapotranspiration and GPP from MTE

MTE-ET is data-oriented product derived by extrapolating the flux-tower observations from 4678 sites in space and time using climate data and remotely sensed data from 1982 to 2008 (Jung et al., 2011). This latter monthly GPP product now covers the period 1982-2011 at 0.5° spatial resolution, but many of the flux tower observations only cover much shorter periods.

F: Evaluation datasets for LAI

The GIMMS-LAI3g data were derived from the Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Differential Vegetation Index (NDVI) using the neural network algorithm (Zhu et al., 2013). This monthly dataset is available for period 1982-2015 with spatial resolution of 1/12 degree.

G: Evaluation datasets for forest above-ground biomass

The GlobBiomass dataset of above-ground biomass (AGB) density (Fig. S2b) were generated by combining spaceborne Synthetic Aperture Radar (SAR), LiDAR (Light Detection And Ranging) and optical remote sensing observations. This data is available for the year of 2010

and with a spatial resolution of 100m. The global forest above-ground biomass density from GEOCARBON project (Fig. S2c) were generated from three existing datasets (Saatchi et al., 2011; Baccini et al., 2012; Santoro et al., 2015). The AGB estimate at 0.01^{o} spatial resolution in the tropics was the weighted averages of data used in Saatchi et al. (2011) and Baccini et al. (2012), while that in the temperate and boreal forest corresponds to the biomass estimated in Santoro et al. (2015). These three datasets (i.e. Saatchi, Baccini and Santoro) were based on EO images acquired in different epochs (i.e. year 2000, 2007-2008 and year 2010). Simulated above-ground forest biomass carbon for the year of 2010 is compared with observation-based datasets.

H: Evaluation datasets for SOC

H1 Global Soil Dataset for use in Earth System Models (GSDE)

GSDE provides global gridded dataset of soil organic C (SOC) with a spatial resolution of 30 seconds and soil depth of 2.3m. This dataset is constructed based on the Soil Map of the World and various regional and national soil databases, including soil attribute data and soil maps (Shangguan et al., 2014).

H2 SoilGrids

SoilGrids provides global gridded dataset of SOC with a spatial resolution of 250m and soil depth of 2m. This dataset is extrapolated from over 230 000 soil profile from the WoSIS database by using state-of-the-art machine learning methods (Hengl et al., 2017).

I: sensitivity of GPP anomalies to anomalies of mean annual temperature (S_T), annual precipitation (S_P) and shortwave radiation (S_R)

In this study, we estimated the response of GPP to climate variability by employing the multiple regression approach from Piao et al. (2013):

$$y = S_T x_T + S_P x_P + S_R x_R + \varepsilon \qquad \text{(S4)}$$

where y is the de-trended anomaly of GPP, x_T is the detrended mean annual temperature anomaly, x_P is the de-trended annual precipitation anomaly, and x_R is the detrended mean annual radiation anomaly. The fitted regression coefficients S_T, S_P, S_R define the apparent GPP sensitivity to inter-annual variations in temperature, precipitation and radiation, respectively, and ε represents the residual error term. ORCHIDEE-CNP and ORCHIDEE used the same forcing data of meteorology from CRU-JRA-55. But BESS and MTE used climate datasets from CRU-NCEP. For analyzing the sensitivity of GPP anomaly to climates, we used CRU-JRA-55 for ORCHIDEE-CNP and ORCHIDEE and CRU-NCEP for BESS and MTE.

S2. Influences of agricultural activities on N and P fluxes

The residence time of N and P in fertilized agricultural ecosystems is completely overridden by human input, so that model biases for those ecosystems should be related to how added nutrients are partitioned between their incorporation in biomass and losses, and soil P accumulation for P through sorption. Total N fertilizer input and harvest N for cropland by ORCHIDEE-CNP are 125 Tg N yr^{-1} and 75 Tg N yr^{-1} in 2010, which gives a high N use efficiency for crop biomass production equal to 0.6 (NUE; defined as the ratio of harvest N yield to total N inputs). Modelled N harvest by ORCHIDEE-CNP is very close with the FAOSTAT-based estimates of 74 Tg N yr^{-1} (Zhang et al., 2015), but our total N input prescribed from Zhang et al. (2015) is lower than FAOSTAT-based estimate of 174 Tg N yr^{-1}. This

difference is partly due to the different method to allocate fertilizer to crop and pasture areas by Zhang et al. (2015) (Conant et al., 2013; Lassaletta et al., 2014) and Lu and Tian (2017) (Monfreda et al., 2008). Further, biological nitrogen fixation (BNF) in croplands is about ~30 Tg N yr^{-1} according to Bodirsky et al., (2012) and 39 Tg N yr^{-1} for both cropland and pasture in Bouwman et al. (2013b), which was not accounted for in ORCHIDEE-CNP but was accounted in Zhang et al. (2015). A lower estimation in total N input by ORCHIDEE-CNP leads to a higher NUE compared to Zhang et al. (2015) (0.42).

Averaged total P input for cropland over 2002-2010 is estimated to be 20.4 Tg P yr^{-1}, which is lower than the estimate by Lun et al. (2018) (25.5 Tg P yr^{-1}). This is because we used a lower manure P input (4.2 Tg P yr^{-1}) than Lun et al. (2018) (7.1 Tg P yr^{-1}) and omitted sludge P input (1.4 Tg P yr^{-1} by Lun et al., 2018). Averaged modelled P harvest over 2002-2010 is 5.8 Tg P yr^{-1}, which is also lower than that of Lun et al. (2018) (11.7 Tg P yr^{-1}). The stoichiometry of crop harvest differs among crop types, while ORCHIDEE-CNP currently only distinguishes between three major crop types (maize, rice and wheat), which may explain the bias of P harvest whereas N harvest rates are in good agreement with independent estimates.

Modelled N and P leaching from croplands and pasture contributes 39% and 88% of global total N and P leaching in ORCHIDEE-CNP, respectively. The simulated global P leaching from managed land (3.3 Tg P yr^{-1}) is lower than the estimation by Lun et al. (2018) of 5.4 Tg P yr^{-1}) but higher than in Mekonnen and Hoekstra (2018) (0.6 Tg P yr^{-1}). This large range of published estimates is due to the different methodologies and the assumptions of the ratio of P loss to P input (Liu et al., 2018). For example, Lun et al. (2018) assumed that a constant 12.5% of total P input was leached, while Mekonnen and Hoekstra (2018) estimated gridded P load to freshwater based on a much smaller erosion-runoff-leaching fraction (2.8% of total P input) for diffuse sources (Table S3).

S3. Biome-scale nutrient budgets

In a model resolving nutrients like ORCHIDEE-CNP, as changes in climate and CO_2 drive an increased land carbon storage, N and P tend to accumulate over time in biomass and soil carbon pools. We analyzed the modelled N and P storage changes for 4 managed and 10 natural biomes for the period of the 2000s from the ORCHIDEE-CNP output. We found a net N accumulation in natural biomes mainly fueled by BNF (Fig. S15a), except for boreal forests and natural grasslands where the contribution of N deposition is the dominant source of new N inputs. A net N accumulation of 14.4 Tg N yr^{-1} is found in pastures, whereas a net N loss of 6.5 Tg N yr^{-1} is found in croplands (Fig. S15a). N fertilizers application accounts for 20~33% of the total pasture N input, and dominates N inputs in cropland (80~95%, 149 Tg N yr^{-1} in total). Only 55~60% of N input from fertilizers application is harvested in crop biomass. The rest of these inputs being lost in gaseous form (45.7 Tg N yr^{-1}) and from drainage and surface runoff leaching (20.7 Tg N yr^{-1}).

We found a net P accumulation of 0.8 Tg P yr^{-1} in tropical and temperate forests, but a net P loss of -0.16 Tg P yr^{-1} in natural grasslands due to increased soil P fixation exceeding P deposition and weathering inputs. Boreal forests show a nearly balanced P budget (Fig. S15b). For tropical broad-leaved evergreen forests, inputs from deposition and weathering of primary and secondary minerals represent 55% and 45% of total P inputs, respectively. For other natural biomes, deposition dominates (>75%) the P inputs. P losses to soil P fixation are the major P loss pathway for all biomes (Fig. S15b).

A net P accumulation of 1.2 Tg P yr^{-1} is simulated for fertilized pasture and croplands, respectively (Fig. S15b). Manure fertilization dominates the total pasture P input (69%~93%) but mineral fertilizers dominates the total cropland P inputs (>95%). In croplands, 31% of

cropland P input from fertilizers is harvested in crop biomass (5.9 Tg P yr^{-1}), 36% is lost to soil P fixation (4.3 Tg P yr^{-1}) and drainage and surface runoff (2.6 Tg P yr^{-1}). The net P accumulation in pasture is mainly attributed to the omission of P output by grazing in ORCHIDEE-CNP.

S4. Biological nitrogen fixation

Biological nitrogen fixation (BNF) is the major natural input of N to terrestrial ecosystems (Vitousek et al., 2013) and its temperature dependence is critical for large-scale patterns of N limitations (Du et al., 2020) while its response to changing conditions critically affects the occurrence of N limitation under increasing CO_2 (Goll et al., 2017b). Here we used BNF estimates from Sullivan et al. (2014) for tropical forests and a model-based global gridded BNP from Peng et al. (2019) to evaluate BNF in ORCHIDEE-CNP. A special attention is given to tropical forests where interactions between N and P are potentially controlling BNF (Houlton et al., 2008).

Sullivan et al. (2014) estimated symbiotic and free-living BNF of tropical forest sites based on acetylene reduction assays with nodules, soil and litter at 12 sites, using a spatial sampling method to generate unbiased estimates of mean symbiotic BNF independent of legume trees abundance, which more robustly captures the irregular distribution of nodules in the landscape. They found plot-level measurements of BNF rates which were five times lower than used in previous empirical upscaling to the tropical forest biomes (Cleveland et al. 1999; Wang and Houlton, 2009). Recalibrating the empirical relationship between BNF and either NPP or evapotranspiration from Cleveland et al. (1999) and Wang and Houlton et al. (2009) based on their plot level estimates of BNF, Sullivan et al. (2014) upscaled plot-level BNF to all tropical forests to 4.9~6.3 kg N ha^{-1} yr^{-1} and 3.7~7.8 kg N ha^{-1} yr^{-1} for first and third quantiles, respectively. Peng et al. (2019) developed a global BNF model constrained by observed C:N ratios of various plant pools (Sect. S1D1 in the supplement), and they simulated a higher BNF for tropical forests than Sullivan et al. (2014) with an interquartile range (IQR) of 14~43 kg N ha^{-1} yr^{-1}. BNF by ORCHIDEE-CNP in tropical forests is comparable in median value (5.6 kgN ha^{-1} yr^{-1}) but has a wider IQR of 2.5-7.8 kg N ha^{-1} yr^{-1} than the upscaled estimation according to Sullivan et al. (2014) (Fig. S16), but is lower than Peng et al. (2019)'s estimate, even lower than their first quartile. For all biomes, ORCHIDEE-CNP simulated a higher total BNF (153 Tg N yr^{-1}) than Peng et al. (2019) (116 Tg N yr^{-1}) and a roughly comparable spatial range (IQR = 3.1-18 kg N ha^{-1} yr^{-1} compared to 0.23-15 kg N ha^{-1} yr^{-1}) (Fig. S16). Specifically, ORCHIDEE-CNP simulated a tropical BNF which is close to estimates derived from measurements, but a higher global BNF of 153 Tg N yr-1 than previous estimates (40-128 Tg N yr-1) (Galloway et al., 2004; Vitousek et al., 2013), primarily due to an overestimation of extratropical BNF rates (Fig. S16). The latter is likely related to the omission of a direct temperature control on BNF in ORCHIDEE-CNP which was shown to be able to explain low BNF rates at higher latitudes (Houlton et al., 2008).

S5. N and P load rates to river catchment scale

We used the export of N and P from terrestrial systems from the Global Nutrient Export from Water-Sheds(GlobalNEWS2) model (Table 1). This model gives N and P loading to rivers by surface runoff and drainage based on a mass-balance approach (Mayorga et al., 2010). As ORCHIDEE-CNP lacks a representation of nutrient soil infiltration and groundwater networks, we had to approximate N and P load from ORCHIDEE output variables. Since most of dissolved N in groundwater enters the river network via drainage (Bouwman et al., 2013a), we extracted annual N exported via drainage and surface runoff from the ORCHIDEE-CNP output to be compared with the GlobalNEWS2 load rates (N_{load}) on catchments scale:

$$N_{load} = N_{runoff} + N_{drainage}\ , \quad \text{(S1)}$$

where N_{runoff} is the simulated annual N flux lost by surface runoff, and $N_{drainage}$ the annual N flux lost by drainage both in kgN km^{-2} yr^{-1}.

Dissolved P which drains in deeper soil layers (P_{dra}) is mostly adsorbed by soil minerals and only a minor fraction enters the groundwater. Thus, the annual soil P loss via surface runoff (kg P km^{-2} yr^{-1}) from the ORCHIDEE-CNP output were extracted, and were compared with the GlobalNEWS2 load rates (P_{load}):

$$P_{load} = P_{runoff}\ , \quad \text{(S2)}$$

ORCHIDEE-CNP simulated a global N_{load} calculated over the period 2002-2010 of 55 Tg N yr^{-1} with N_{load} from natural biomes being 35 Tg N yr^{-1}, close to the estimate by GOLUM-CNP (38 Tg N yr^{-1}) but higher than one by GlobalNEWS2 (28 Tg N yr^{-1} in year 2000). Simulated N_{load} from managed land, namely croplands and pastures, is 21 Tg N yr^{-1}, about half the estimate from GlobalNEWS2 (40 Tg N yr^{-1} in year 2000; Mayorga et al., 2010). This underestimation by ORCHIDEE-CNP is likely due to the lower manure N fertilizer scenario, as part of the model forcing, than used in GlobalNEWS2 (Fig. S21), due to the omission of point sources such as livestock farms and sewage systems. On catchment scale, N_{load} in the year 2000 simulated by ORCHIDEE-CNP have a comparable range (0~2500 kg N km^{-2} yr^{-1}) but generally lower values than GlobalNEWS2 (Fig. S17a to c), excepted for some basins dominated by Oxisols such as the Amazon (Fig. S18). The overestimation of N_{load} in some tropical basins dominated by Oxisols soils (e.g. Amazon) (Fig. S17) is mainly attributed to our assumptions regarding the sum of N_{runoff} and N_{dra} to approximate N_{load} (Eq. S1). Tropical basins have strong N denitrification losses from the groundwater as a result of large flux and high N concentration of groundwater (Bouwman et al., 2013a), thereby a part of the N is lost before the water enters the river (i.e. N_{dra}). Denitrification losses from groundwater are not accounted for in ORCHIDEE-CNP. In addition, ORCHIDEE-CNP applies a globally uniform ammonium sorption capacity of soils, calibrated to a limited number of soil measurements (Zaehle and Friend, 2010) which might not be applicable to highly weathered and low pH soils in the tropics.

Global total P^*_{load} ($P^*_{load} = P_{dra} + P_{runoff}$) is estimated at 3.8 Tg P yr^{-1} averaged over the period 2002-2010 in ORCHIDEE-CNP. Natural biomes account for only ~10% of global total P^*_{load} (0.39 Tg P yr^{-1}), about an order of magnitude less than in GOLUM-CNP (Table S4). This lower total P^*_{load} for natural biomes mainly occurs in forest ecosystems (Fig. S17) which can be mainly attributed to the low substrate availability (i.e. dissolved P concentration) due to strong P immobilization into biomass and SOM for forests (Fig. S1a, g) in ORCHIDEE-CNP. Croplands and pastures account for most of the global total P^*_{load} with a value of 3.3 Tg P yr^{-1} which is rather close to the estimate of 4 Tg P yr^{-1} by Bouwman et al. (2013b), but lower than the 5.4 Tg P yr^{-1} reported by Lun et al. (2018) and higher than in Mekonnen and Hoekstra (2018) (0.6 Tg P yr^{-1}) (Table S3). ORCHIDEE-CNP cannot capture the high P losses from extremely high livestock densities in small areas where soil P retention capacities have been reached. Besides, Lun et al. (2018) considered sludge inputs (1.4 Tg P yr^{-1} in year 2000), which are omitted in ORCHIDEE-CNP. The simulated values of P_{load} (Eq. S2) at catchment scale are in the same order of magnitude than estimates from GlobalNEWS2 (Mayorga et al., 2010). However, ORCHIDEE-CNP simulates generally lower P_{load} over the Amazon and Central African basins compared to GlobalNEWS2 (Mayorga et al., 2010) (Fig. S17d, e). This underestimation can be partly attributed to the strong P fixation capacity of soil (i.e. occlusion) in tropical regions in ORCHIDEE-CNP, which is related to a fast turnover of soil inorganic P

(Sect. S6 in the supplement). Note that P_{dra} is small but non-negligible in surface and shallow subsurface runoff (Mayorga et al., 2010), and ignoring this flux for P_{load} may lead to our underestimation of P_{load} (Eq. S2). Simulated P_{load} is also higher than GlobalNEWS2 in European catchments, where pasture is widespread (Fig. 17d, e), which can be attributed to either a less effective soil P sorption capacity or the omission of grazing which enhances the mineralization of organic P and as a consequence inorganic P losses.

S6. Soil solution inorganic phosphorus turnover rate

The simulated soil P fluxes cannot be evaluated against direct observations (Frossard et al., 2011). Here we used pool turnover times instead of fluxes to evaluate P dynamics, as the turnover of P in the soil solution is an important component of P bioavailability in soil (Helfenstein et al., 2018). We compared soil solution P turnover in ORCHIDEE-CNP with empirical data from isotope exchange kinetic (IEK) experiments (Fardeau et al. 1991; Frossard et al. 2011). IEK experiments measure the exchange rate of inorganic P between soil solution and the soil solid phase (K_m, unit: min^{-1}) in laboratory conditions, omitting biological processes. The inorganic exchange processes captured in K_m dominate short-term P fluxes.

For a straightforward comparison (excluding biological uptake process from model output to match the experimental method used for measuring), we diagnosed K_m (unit: min^{-1}) in the simulations from the net exchange between dissolved and sorbed labile P:

$$K_m = \frac{P_{dissolved}}{\Delta P_{sorbed} - F_{dss}}, \quad \text{(S3)}$$

Where $P_{dissolved}$ is the labile P in soil solution (g P m^{-2}) and ΔP_{sorbed} the change in labile P adsorbed (g P m^{-2} t^{-1}) with t=30 minutes (model time step) corrected for the loss term of P_{sorbed} related to strongly sorption. In ORCHIDEE-CNP, the exchange between $P_{dissolved}$ and P_{sorbed} (F_{dss}) is calculated using a Freundlich adsorption isotherm by assuming a chemical equilibrium between both pools is reached on the model time step (Goll et al., 2018).

We used data from a global compilation of empirical K_m data from 218 different soil samples, mainly from Europe and North America but also covering Asian and Africa (Helfenstein et al., 2018). We found that simulated $K_m$are of similar magnitude with median value of 24~468 min^{-1} across soil orders compared to those empirical data with simulated 5~990 min^{-1} (Fig. S19). However, the turnover rates for different soil types did not match empirical data. The model overestimated K_m for Alfisols, Aridisols, Entisols, Mollisols, Spodosols, Ultisols, and underestimated K_m for Oxisols and Vertisols.

ORCHIDEE-CNP cannot capture variations in K_m among different soil orders as derived from measurements (Helfenstein et al., 2018), that is because that ORCHIDEE-CNP distinguishes only between three classes of soils with respect to parameter controlling inorganic P sorption dynamics: Oxisols, Molisols and other soils (Table S5; Goll et al., 2017a). The difference in biases of K_m between Oxisols and other soil suborders suggests that our parameterization of sorption is the primary cause for the model bias, as all other processes affecting dissolved P are not a function of soil type. In the P-enabled LSMs, inorganic P processes operating on longer timescale (occlusion, strong sorption) are only simply represented (Wang et al., 2010; Yang et al., 2014; Goll et al., 2017b). This processes in LSMs is primarily based on calibration rather than data driven, which remain a large source of uncertainty regarding changes in P availability under elevated CO_2 (Goll et al., 2012). Recent advances in the quantification of inorganic soil P turnover (Helfenstein et al., 2020) should be used to improve this part of models.

S7. P biochemical mineralization from phosphatase enzymes

Biochemical phosphorus mineralization (BPM) mediated by extracellular phosphatase enzymes produced plants and microbes through cleaving P out of organic matter, is a potentially significant pathway for mineralizing P (Wang et al., 2007). ORCHIDEE-CNP includes an empirical parameterization of BPM which accounted for an enhanced mineralization when plants experience suboptimal P-to-N availabilities, as an approximation of the stoichiometric status of the whole ecosystem, including the effect of substrate availability on mineralization (McGill and Cole, 1981) (see Eq. 18 in Goll et al., 2017b). Potential phosphatase activity (P_{ases}) can be measured in the lab from soil samples, and serves as a surrogate for the lack of direct BPM measurements in real-world soil conditions (Sun et al., 2020). We compare the global pattern of simulated P mineralization fluxes in ORCHIDEE-CNP with the map of P_{ases} extrapolated from site measurements (Margalef et al., 2017) produced by Sun et al. (2020). The modelled BPM is higher in tropical regions than in desert and boreal regions, consistent with the P_{ases} pattern (Fig. S20), although the scarcity of tropical measurements makes the high values observed in these regions more uncertain (Sun et al., 2020).

Global organic P mineralization fluxes were evaluated against potential phosphatase activity (P_{ases}) data (Margalef et al., 2017; Sun et al., 2020), and actual bulk P mineralization measurements based on radioisotopic approaches (Bünemann, 2015; Wanek et al., 2019), to our knowledge the first attempt to evaluate a model for this process. ORCHIDEE-CNP can capture the general global pattern of P_{ases} (Fig. S20b), which gives us some confidence that large-scale differences with respect to the decoupling of P mineralization from the mineralization of C and N can be roughly captured by the model.

Direct measurements of bulk organic P mineralization using radio-isotopic approaches are only available for few soils. If up-scaled to yearly rates, measured organic P mineralization rates range between 10^1 to 10^3 g P m^{-2} yr^{-1} (Bünemann, 2015; Wanek et al., 2019) which are orders of magnitude higher than simulated by ORCHIDEE-CNP (0.53 ± 0.48 g P m^{-2} yr^{-1}; Fig. S20a). It has to be noted that lab measurements of P mineralization rates and potential phosphatase activity are conducted under optimal conditions, whereas in the field, temperature and water hamper enzyme activity for large parts of the year. This illustrates that the evaluation of model performance in representing soil organic P mineralization remains a challenge since quantification of apparent rates under field conditions are currently missing.

References

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

Bodirsky, B., Popp, A., Weindl, I., Dietrich, J., Rolinski, S., Scheiffele, L., Schmitz, C., and Lotze-Campen, H.: N2O emissions from the global agricultural nitrogen cycle - current state and future scenarios, Biogeosciences, 9, 4169-4197, http://doi.org/10.5194/bg-9-4169-2012, 2012.

Bouwman, A., Beusen, A., Griffioen, J., Van Groenigen, J., Hefting, M., Oenema, O., Van Puijenbroek, P., Seitzinger, S., Slomp, C., and Stehfest, E.: Global trends and uncertainties in terrestrial denitrification and N2O emissions, Philos. Trans. R. Soc. B-Biol. Sci., 368, http://doi.org/10.1098/rstb.2013.0112, 2013a.

Bouwman, L., Goldewijk, K. K., Van Der Hoek, K. W., Beusen, A. H., Van Vuuren, D. P., Willems, J., Rufino, M. C., and Stehfest, E.: Exploring global changes in nitrogen and phosphorus cycles in agriculture induced by livestock production over the 1900–2050 period, Proc. Natl. Acad. Sci. U. S. A., 110, 20882-20887, http://doi.org/10.1073/pnas.1012878108, 2013b.

Bünemann, E. K.: Assessment of gross and net mineralization rates of soil organic phosphorus–A review, Soil Biol. Biochem., 89, 82-98, http://doi.org/10.1016/j.soilbio.2015.06.026, 2015.

Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini, B. E. L., Cornelissen, J. H. C., Craine, J. M., Craven, D., de Vries, F. T., Díaz, S., Domingues, T. F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W. N., Hickler, T., Jansen, S., Kramer, K., Kraft, N. J. B., Kurokawa, H., Laughlin, D. C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N. A., Spasojevic, M. J., Sosinski, E., Thornton, P. E., Valladares, F., van Bodegom, P. M., Williams, M., Wirth, C., and Reich, P. B.: Mapping local and global variability in plant trait distributions, Proc. Natl. Acad. Sci. U. S. A., 114, E10937-E10946, http://doi.org/10.1073/pnas.1708984114, 2017.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., Chédin, A., and Ciais, P.: Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res.-Atmos., 110, http://doi.org/10.1029/2005JD006390, 2005.

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 biogeochem. cycles, 13(2), 623-645, https://doi.org/10.1029/1999GB900014,1999.

Conant, R. T., Berdanier, A. B., and Grace, P. R.: Patterns and trends in nitrogen use and nitrogen recovery efficiency in world agriculture, Global Biogeochem. Cycles, 27, 558-566, http://doi.org/10.1002/gbc.20053, 2013.

Du, E., Terrer, C., Pellegrini, A. F., Ahlström, A., van Lissa, C. J., Zhao, X., Xia, N., Wu, X., and Jackson, R. B.: Global patterns of terrestrial nitrogen and phosphorus limitation, Nat. Geosci., 13, 221-226, http://doi.org/10.1038/s41561-019-0530-4, 2020.

Fardeau, J., Morel, C., and Boniface, R.: Phosphate ion transfer from soil to soil solution: kinetic parameters [available P, phosphate flow], Agronomie, http://doi.org/10.1051/agro:19910909, 1991.

Frossard, E., Achat, D. L., Bernasconi, S. M., Bünemann, E. K., Fardeau, J.-C., Jansa, J., Morel, C., Rabeharisoa, L., Randriamanantsoa, L., Sinaj, S., Tamburini, F., and Oberson, A.: The use of tracers to investigate phosphate cycling in soil–plant systems, in: Phosphorus in Action, Springer, 59-91, 2011.

Gobron, N., Pinty, B., Aussedat, O., Chen, J., Cohen, W. B., Fensholt, R., Gond, V., Huemmrich, K. F., Lavergne, T., Mélin, F., Privette, J. L., Sandholt, I., Taberner, M., Turner, D. P., Verstraete, M. M., and Widlowski, J.: Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: Methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations, J. Geophys. Res.-Atmos, 111, D13110, http://doi.org/10.1029/2005JD006511, 2006a.

Gobron, N., Pinty, B., Taberner, M., Mélin, F., Verstraete, M., and Widlowski, J.-L.: Monitoring the photosynthetic activity of vegetation from remote sensing data, Adv. Space Res., 38, 2196-2202, http://doi.org/10.1016/j.asr.2003.07.079, 2006b.

Goll, D. S., Brovkin, V., Parida, B., Reick, C. H., Kattge, J., Reich, P. B., Van Bodegom, P., and Niinemets, Ü.: Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling, Biogeosciences, 9, 3547-3569, http://doi.org/10.5194/bg-9-3547-2012, 2012.

Goll, D., Vuichard, N., Maignan, F., Jornet-Puig, A., Sardans, J., Violette, A., Peng, S., Sun, Y., Kvakic, M., and Guimberteau, M.: A representation of the phosphorus cycle for ORCHIDEE (revision 4520), Geosci. Model Dev., 10, 3745-3770, http://doi.org/10.5194/gmd-10-3745-2017, 2017a.

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., Holland, E. A. Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., Vöosmarty, C. J.: Nitrogen cycles: past, present, and future, Biogeochemistry, 70, 153-226, http://doi.org/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), Geosci. Model Dev., 10, 2009-2030, http://doi.org/10.5194/gmd-10-2009-2017, 2017b.

Goll, D. S., Joetzjer, E., Huang, M., and Ciais, P.: Low phosphorus availability decreases susceptibility of tropical primary productivity to droughts, Geophys. Res. Lett., 45, 8231-8240, http://doi.org/10.1029/2018GL077736, 2018.

Helfenstein, J., Jegminat, J., McLaren, T. I., and Frossard, E.: Soil solution phosphorus turnover: derivation, interpretation, and insights from a global compilation of isotope exchange kinetic studies, Biogeosciences, 15, 105-114, http://doi.org/10.5194/bg-15-105-2018, 2018.

Helfenstein, J., Pistocchi, C., Oberson, A., Tamburini, F., Goll, D. S., and Frossard, E.: Estimates of mean residence times of phosphorus in commonly considered inorganic soil phosphorus pools, Biogeosciences, 17, 441-454, http://doi.org/10.5194/bg-17-441-2020, 2020.

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G., Walsh, M. G. Gonzalez, M. R.: SoilGrids1km—global soil information based on automated mapping, PLoS One, 9, http://doi.org/10.1371/journal.pone.0105992, 2014.

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, 327-330, http://doi.org/10.1038/nature07028, 2008.

Jiang, M., Caldararu, S., Zaehle, S., Ellsworth, D. S., and Medlyn, B. E.: Towards a more physiological representation of vegetation phosphorus processes in land surface models, New Phytol., 222, 1223-1229, http://doi.org/10.1111/nph.15688, 2019.

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

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Damiano, G., 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, la- tent heat, and sensible heat derived from eddy covariance, satel- lite, and meteorological observations, J. Geophys. Res.-Biogeo., 116, http://doi.org/10.1029/2010JG001566, 2011.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Glob. Biogeochem. Cycle, 19, http://doi.org/10.1029/2003GB002199, 2005.

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., and Garnier, J.: 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland, Environ. Res. Lett., 9, 105011, http://doi.org/10.1088/1748-9326/9/10/105011, 2014.

Liu, W., Yang, H., Ciais, P., Stamm, C., Zhao, X., Williams, J. R., Abbaspour, K. C., and Schulin, R.: Integrative Crop‐Soil‐Management Modeling to Assess Global Phosphorus Losses from Major Crop Cultivations, Glob. Biogeochem. Cycle, 32, 1074-1086, http://doi.org/10.1029/2017GB005849, 2018.

Lu, C. C., and Tian, H.: Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hot spots and nutrient imbalance, Earth Syst. Sci. Data, 9, 181, http://doi.org/10.5194/essd-9-181-2017, 2017.

Lun, F., Liu, J., Ciais, P., Nesme, T., Chang, J., Wang, R., Goll, D., Sardans, J., Peñuelas, J., and Obersteiner, M.: Global and regional phosphorus budgets in agricultural systems and their implications for phosphorus-use efficiency, Earth Syst. Sci. Data, 10, 1-18, http://doi.org/10.5194/essd-10-1-2018, 2018.

Margalef, O., Sardans, J., Fernández-Martínez, M., Molowny-Horas, R., Janssens, I., Ciais, P., Goll, D., Richter, A., Obersteiner, M., Asensio, D., and Peñuelas, J.: Global patterns of phosphatase activity in natural soils, Sci. Rep., 7, 1-13, http://doi.org/10.1038/s41598-017-01418-8, 2017.

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

Mekonnen, M. M., and Hoekstra, A. Y.: Global anthropogenic phosphorus loads to freshwater and associated grey water footprints and water pollution levels: A high-resolution global study, Water Resour. Res., 54, 345-358, http://doi.org/10.1002/2017WR020448, 2018.

McGill, W., and Cole, C.: Comparative aspects of cycling of organic C, N, S and P through soil organic matter, Geoderma, 26, 267-286, http://doi.org/10.1016/0016-7061(81)90024-0, 1981.

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

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

Peng, J., Wang, Y. P., Houlton, B. Z., Dan, L., Pak, B., and Tang, X.: Global carbon sequestration is highly sensitive to model-based formulations of nitrogen fixation, Glob. Biogeochem. Cycle, http://doi.org/10.1029/2019GB006296, 2019.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., Ma, Y., Myneni, R. B., Poulter, B., Sun, Z., Wang, T., Viovy, N., Zaehle, S., and Zeng, N.: Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends, Global Change Biol., 19, 2117– 2132, http://doi.org/10.1111/gcb.12187, 2013.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO2 flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919-1964, 2003.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H.: A continuous satellite-derived measure of global terrestrial primary production, BioScience, 54, 547-560, http://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2, 2004.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, Proc. Natl. Acad. Sci. U. S. A., 108, 9899-9904, http://doi.org/10.1073/pnas.1019576108, 2011.

Santoro, M., Eriksson, L. E., and Fransson, J. E.: Reviewing ALOS PALSAR backscatter observations for stem volume retrieval in Swedish forest, Remote Sens., 7, 4290-4317, http://doi.org/10.3390/rs70404290, 2015.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., Yuan, H.: A global soil data set for earth system modeling, J. Adv. Model Earth Syst., 6.1, 249-263, https://doi.org/10.1002/2013MS000293, 2014.

Smith, W. K., Reed, S. C., Cleveland, C. C., Ballantyne, A. P., Anderegg, W. R., Wieder, W. R., Liu, Y. Y., and Running, S. W.: Large divergence of satellite and Earth system model estimates of global terrestrial CO2 fertilization, Nat. Clim. Chang., 6, 306-310, http://doi.org/10.1038/NCLIMATE2879, 2016.

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Pe- ters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langen- felds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak northern and strong tropical land carbon uptake from ver- tical profiles of atmospheric CO2, Science, 316, 1732–1735, , http://doi.org/10.1126/science.1137004, 2007.

Sullivan, B. W., Smith, W. K., Townsend, A. R., Nasto, M. K., Reed, S. C., Chazdon, R. L., and Cleveland, C. C.: Spatially robust estimates of biological nitrogen (N) fixation imply substantial human alteration of the tropical N cycle, Proc. Natl. Acad. Sci. U. S. A., 111, 8101-8106, http://doi.org/10.1073/pnas.1320646111, 2014.

Sun, Y., Goll, D. S., Ciais, P., Peng, S., Margalef, O., Asensio, D., Sardans, J., and Peñuelas, J.: Spatial pattern and environmental drivers of acid phosphatase activity in Europe, Front. Big Data, http://doi.org/10.3389/fdata.2019.00051, 2020.

Tum, M., Zeidler, J., Günther, K. P., and Esch, T.: Global NPP and straw bioenergy trends for 2000–2014, Biomass Bioenerg., 90, 230-236, http://doi.org/10.1016/j.biombioe.2016.03.040, 2016.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., and Ahl, D.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sens. Environ., 102, 282-292, http://doi.org/10.1016/j.rse.2006.02.017, 2006.

van der Laan-Luijkx IT, van der Velde IR, van der Veen E, Tsuruta A, Stanislawska K, Babenhauserheide A, Zhang HF, Liu Y, He W, Chen H, Masarie KA, Krol MC, and Peters W: The CarbonTracker Data Assimilation Shell (CTDAS) v1. 0: implementation and global carbon balance 2001-2015, Geosci. Model Dev., 10, 2785-2800, http://doi.org/10.5194/gmd-10-2785-2017, 2017.

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

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

Wanek, W., Zezula, D., Wasner, D., Mooshammer, M., and Prommer, J.: A novel isotope pool dilution approach to quantify gross rates of key abiotic and biological processes in the soil phosphorus cycle, Biogeosciences, 16, 3047-3068, http://doi.org/10.5194/bg-16-3047-2019, 2019.

Wang, Y., Law, R., and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, Biogeosciences, 7, http://doi.org/10.5194/bg-7-2261-2010, 2010.

Wang, Y. P., Houlton, B., and Field, C.: A model of biogeochemical cycles of carbon, nitrogen, and phosphorus including symbiotic nitrogen fixation and phosphatase production, Glob. Biogeochem. Cycle, 21, http://doi.org/10.1029/2006GB002797, 2007.

Wang, Y. P., and Houlton, B. Z.: Nitrogen constraints on terrestrial carbon uptake: Implications for the global carbon-climate feedback, Geophys. Res. Lett., 36, http://doi.org/10.1029/2009GL041009, 2009.

Wißkirchen, K., Tum, M., Günther, K., Niklaus, M., Eisfelder, C., and Knorr, W.: Quantifying the carbon uptake by vegetation for Europe on a 1km 2 resolution using a remote sensing driven vegetation model, Geosci. Model Dev., 6, http://doi.org/10.5194/gmd-6-1623-2013, 2013.

Yang, X., Thornton, P., Ricciuto, D., and Post, W.: The role of phosphorus dynamics in tropical forests--a modeling study using CLM-CNP, Biogeosciences, 11, http://doi.org/10.5194/bg-11-1667-2014, 2014.

Zaehle, S., and Friend, A.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Glob. Biogeochem. Cycle, 24, http://doi.org/10.1029/2009GB003521, 2010.

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., and Shen, Y.: Managing nitrogen for sustainable development, Nature, 528, 51-59, http://doi.org/10.1038/nature15743, 2015.

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

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

Table S1 Plant functional type (PFT) specific leaf N:P ratio (gN g^{-1}P) for ORCHIDEE v1.1 and v1.2: tropical evergreen broadleaf forest (TREBF), tropical raingreen broadleaf forest (TRDBF), temperate evergreen needleleaf forest (TEENF), temperate evergreen broadleaf forest (TEEBF), temperate summergreen broadleaf forest (TEDBF), boreal evergreen needleleaf forest (BOENF), boreal summergreen broadleaf forest (BODBF), boreal summergreen needleleaf forest (BODNF), C3 grassland (C3G), C4 grassland (C4G), C3 cropland (C3C) and C4 cropland (C4C).

	TREBF	TRDBF	TRENF	TREBF	TRDBF	BOENF	BODNF	BODBF	C3G	C4G	C3C	C4C	Sources
$np_{leaf,min}$ for v1.1	16.68	16.68	8.34	10.84	10.84	8.34	10.84	10.84	10.84	10.84	-	-	McGrodd et al. (200
$np_{leaf,max}$ for v1.1	22.57	22.57	11.29	14.67	14.67	11.29	14.67	14.67	14.67	14.67	-	-	McGrodd et al. (200
$np_{leaf,min}$ for v1.2	5	5	5	5	5	5	5	5	5	5	5	5	
$np_{leaf,max}$ for v1.2	30	30	30	30	30	30	30	30	30	30	30	30	

Table S2 Total GPP and NPP (PgC yr-1) simulated by ORCHIDEE-CNP, ORCHIDEE and from data-driven products averaged over 2001-2010.

	GPP	Sources		NPP	Sources
ORCHIDEE-CNP	119	This study	ORCHIDEE-CNP	48	This study
ORCHIDEE	140	Krinner et al., 2005	ORCHIDEE	60	Krinner et al., 2005
MODIS	110±23	Turner et al., 2006	MODIS	55±11	Turner et al., 2006
MTE	120±6	Jung et al., 2009	BETHY	60±3	Tum et al., 2016
BESS	122±25	Jiang and Ryu, 2016	GIMMS	33~49	Smith et al., 2016

Table S3 Global P inputs and outputs estimated by ORCHIDEE-CNP and previous studies. The estimations by Lun et al. (2018) and Mekonnen et al. (2018) are based on the balance of P inputs and removal of P.

		ORCHIDEE-CNP	Lun et al., (2018)	Mekonnen and Hoekstra (2018)
Study period		2002-2010	2002-2010	2002-2010
Study biomes		Both natural and managed biomes	Cropland and pasture	Only cropland
P inputs	P deposition input	3.8	1.4	-
	Sludge P input	-	1.4	1.2
	Chemical fertilizer P input	20.1	16.8	17
	Manure P input	6.7	19.8	5.9
P outputs	P leaching	3.8	-	-
	P leaching from natural land	0.38	-	-
	P leaching from managed land	3.3	-	-
	P erosion loss	-	-	-
	P runoff and leaching loss	3.8	5.4	-
	P erosion, leaching and runoff loss	-		0.6
P erosion-runoff-leaching to freshwater from diffuse sources			12.5%	2.8%

Table S4 Averaged N and P load via drainage and runoff (Tg yr^{-1}) over 2002-2010 simulated by ORCHIDEE-CNP and GOLUM-CNP.

	ORCHIDEE-CNP			GOLUM-CNP
	ALL	Natural biome	Managed biome	
N leaching	55	35	20	30
P leaching	3.5	0.38	3.2	0.39

Table S5 Inorganic sorption dynamic parameters for Oxisols, Molisols and other soils in ORCHIDEE-CNP.

	Oxisols	Molisols	Other soils	References
Freundlich Isotherm coefficient (LH_2O kg^{-1}soil)	348.9	185.5	72.3	Kvacik et al. (in prep)

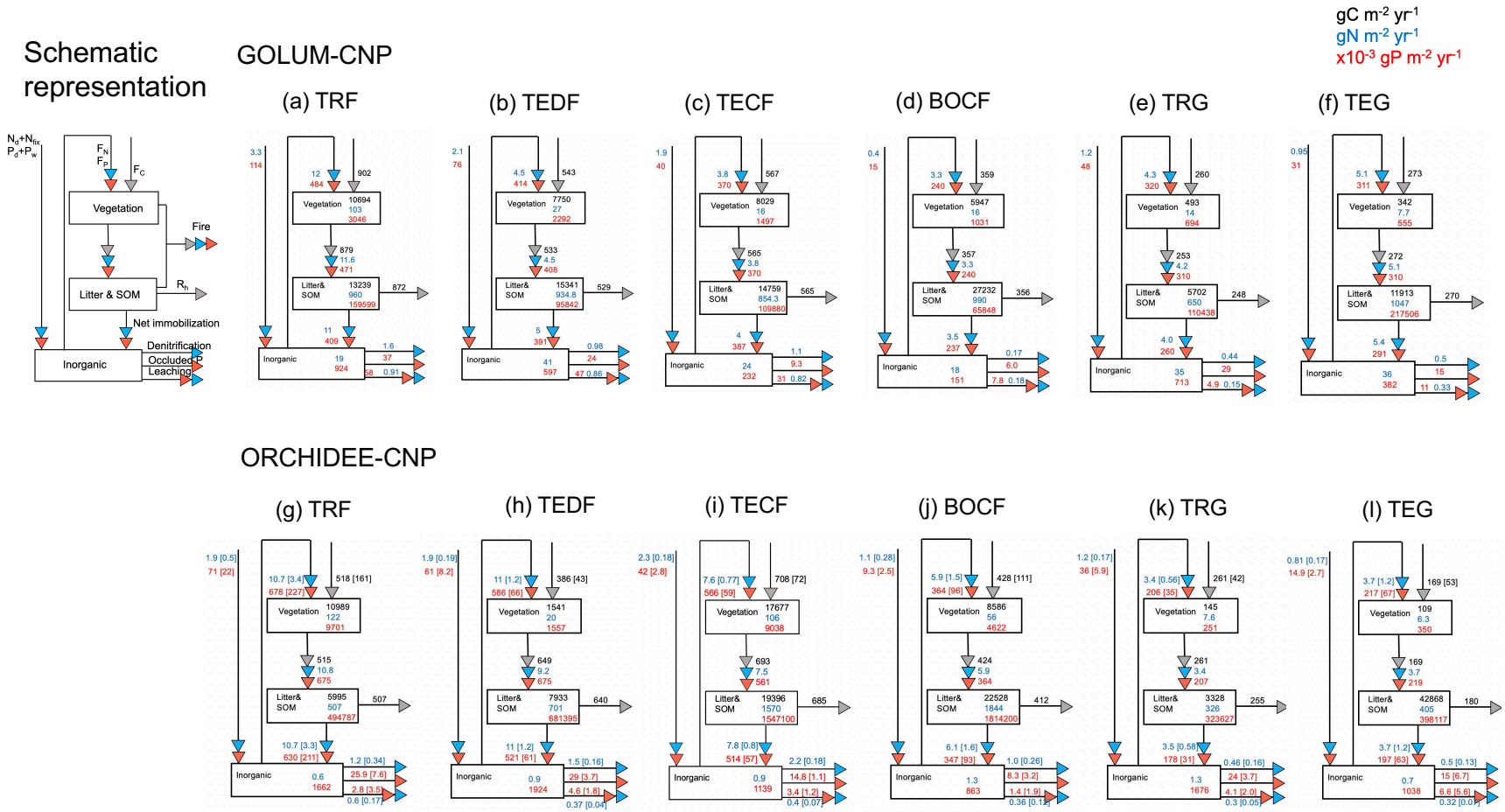


Figure S1 Fluxes and pool sizes of C, N and P for 6 biomes by ORCHIDEE-CNP and GOLUM-CNP. The targeted biomes are tropical rainforests (TRF, a, g), temperate deciduous forests (TEDF, b, h), temperate coniferous forests (TECF, c, i), boreal coniferous forests (BOCF, d, j), tropical/C4 grasslands (TRG, e, k), temperate/C3 grasslands (TEG, f, l). Numbers in square brackets indicate the standard deviations for accounting the spatial spread of C, N and P fluxes.

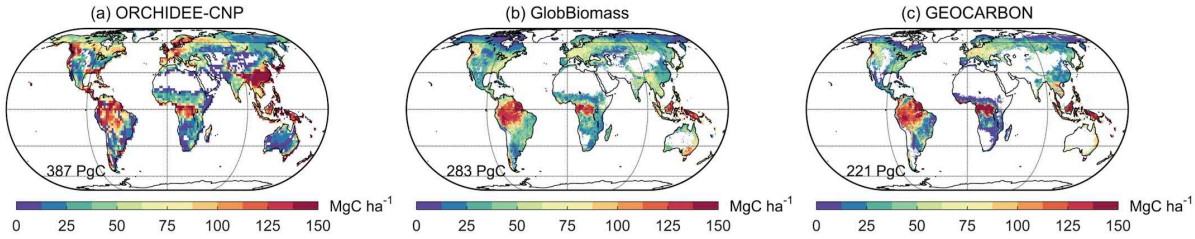


Figure S2 Global pattern of aboveground forest biomass carbon density from ORCHIDEE-CNP and two data-driven map of GlobBiomass and GERCARBON.

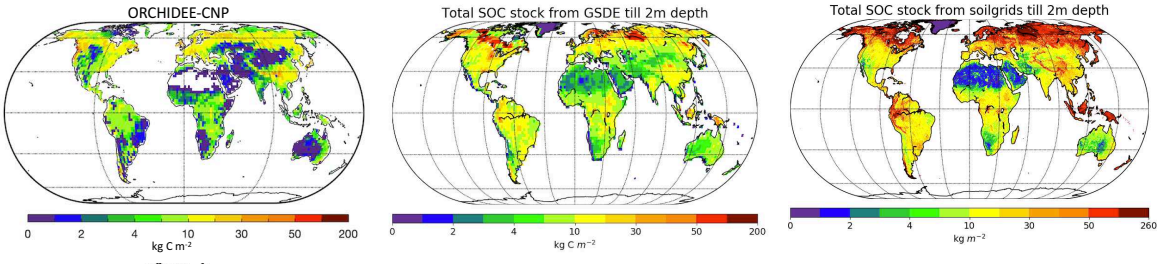


Figure S3 Global pattern of soil organic carbon of 0~2 meters depth from GSDE and Soilgrids and derived from ORCHIDEE-CNP.

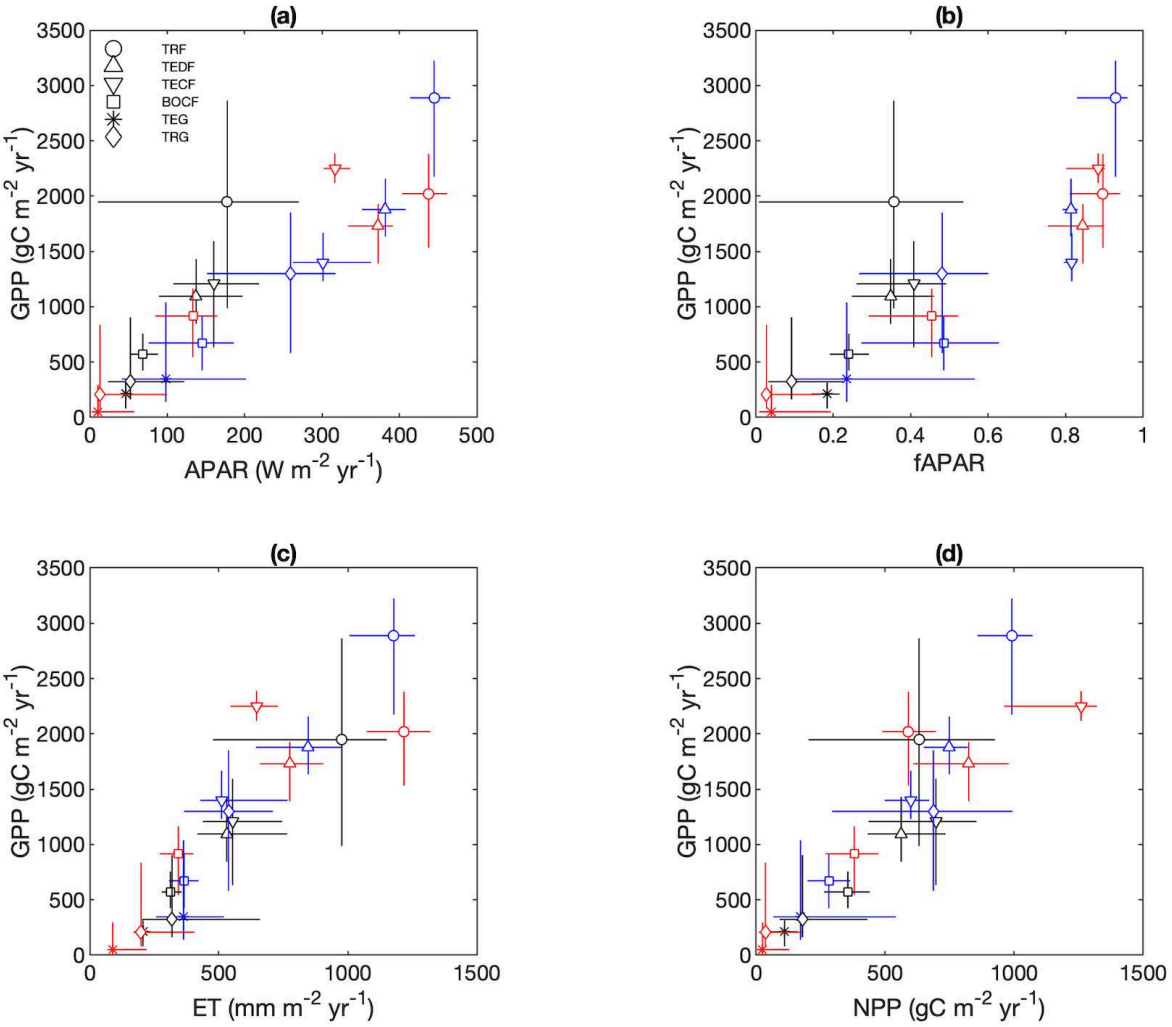


Figure S4 Comparisons between GPP and absorbed light by canopy (APAR; a), fraction of absorbed light by canopy (fAPAR; b) and ecosystem evapotranspiration (ET; c) used by 6 biomes: tropical rainforests (TRF), temperate deciduous forests (TEDF), temperate coniferous forests (TECF), boreal coniferous forests (BOCF), tropical/C4 grasslands (TRG) and temperate/C3 grasslands (TEG). (d) shows the relations between GPP and NPP on biome scale. Reference data (same with Figure 14), ORCHIDEE-CNP and ORCHIDEE are discriminated by color black, red and blue. Error bars indicate the standard deviation of GPP, NPP and resources used by biome.

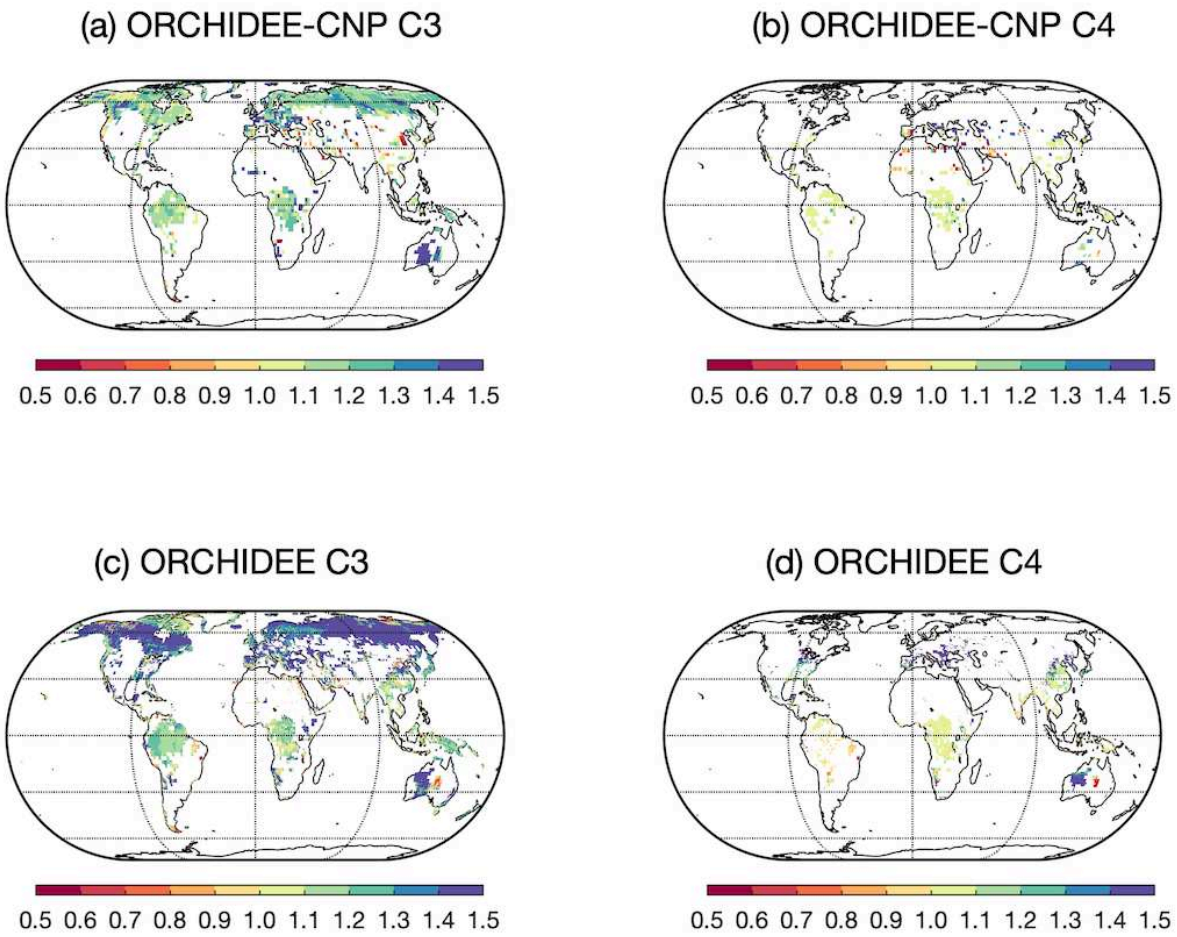


Figure S5 Global pattern of CO_2 fertilization effect (E_{CO2}), defined as the ratio of current GPP with atmospheric CO_2 concentration of 396 ppm (GPP_{396}) and GPP under pre-industrial atmospheric CO_2 concentration of 296 ppm (GPP_{296}), for natural C3 plants (a, c) and natural C4 plants (b, d) by ORCHIDEE-CNP (a, b) and ORCHIDEE (c, d).

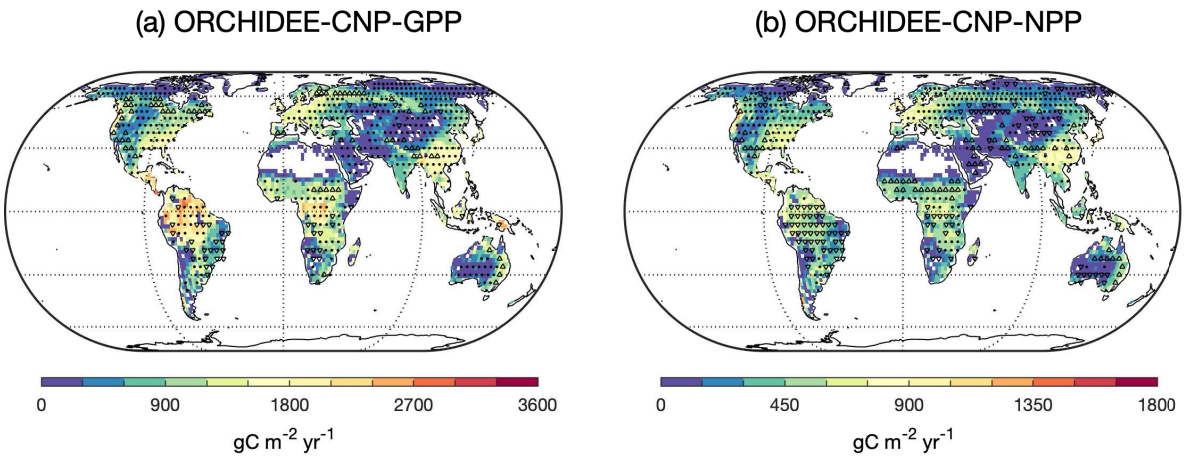


Figure S6 Comparisons of global pattern for plant primary productivity between ORCHIDEE-CNP and data-driven products (MODIS-GPP, MTE-GPP, BESS-GPP, MODIS-NPP, BETHY-NPP, GIMMS-NPP). Black point, '⊗' and '▽' indicates the GPP or NPP simulated by ORCHIDEE-CNP lie within the ranges / higher than the upper limits / lower than the lower limits of estimations by data-driven products, respectively. The ranges of data-driven GPP and NPP involves uncertainties within each product and among all products.

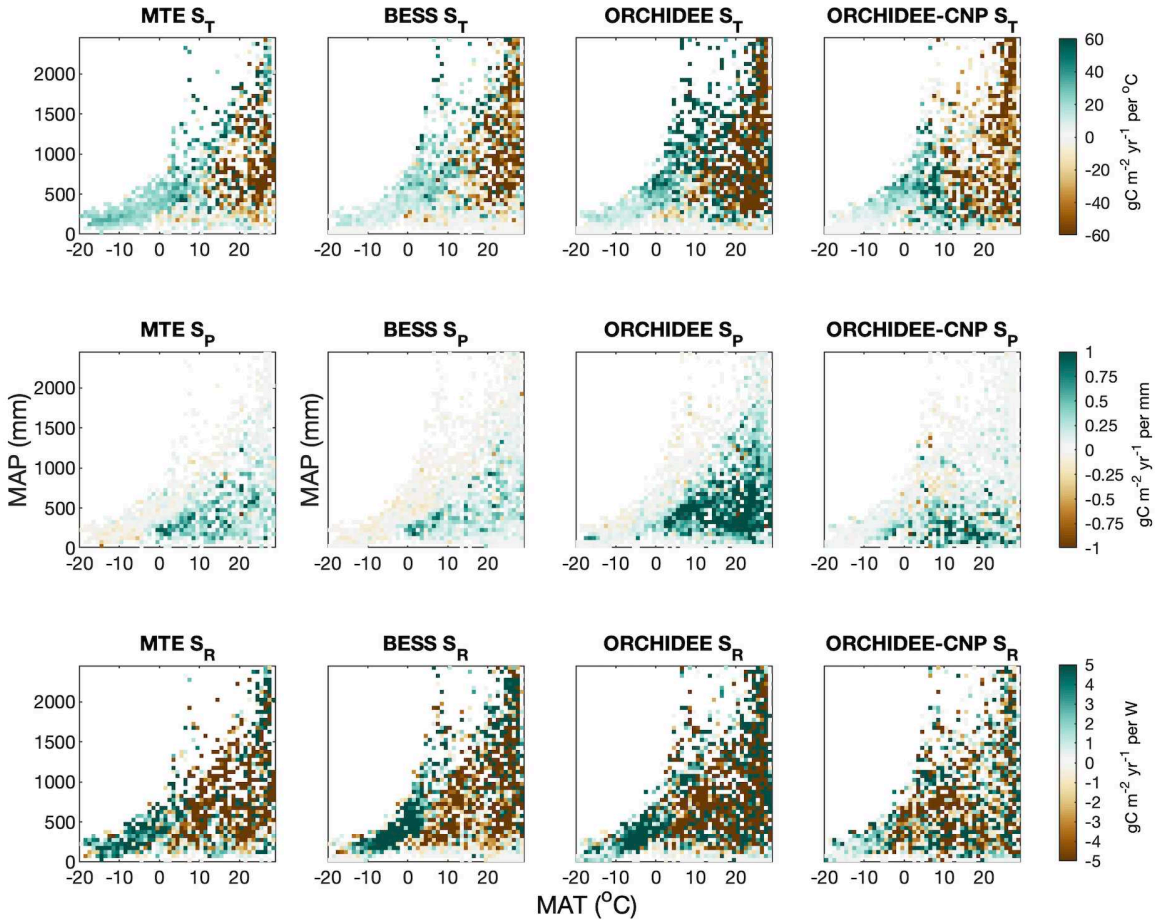


Figure S7 Sensitivity of GPP anomalies to anomalies of mean annual temperature (S_T), annual precipitation (S_P) and shortwave radiation (S_R) plotted in a mean annual temperature (MAT) – mean annual precipitation (MAP) space for MTE-GPP, BESS-GPP, ORCHIDEE and ORCHIDEE-CNP.

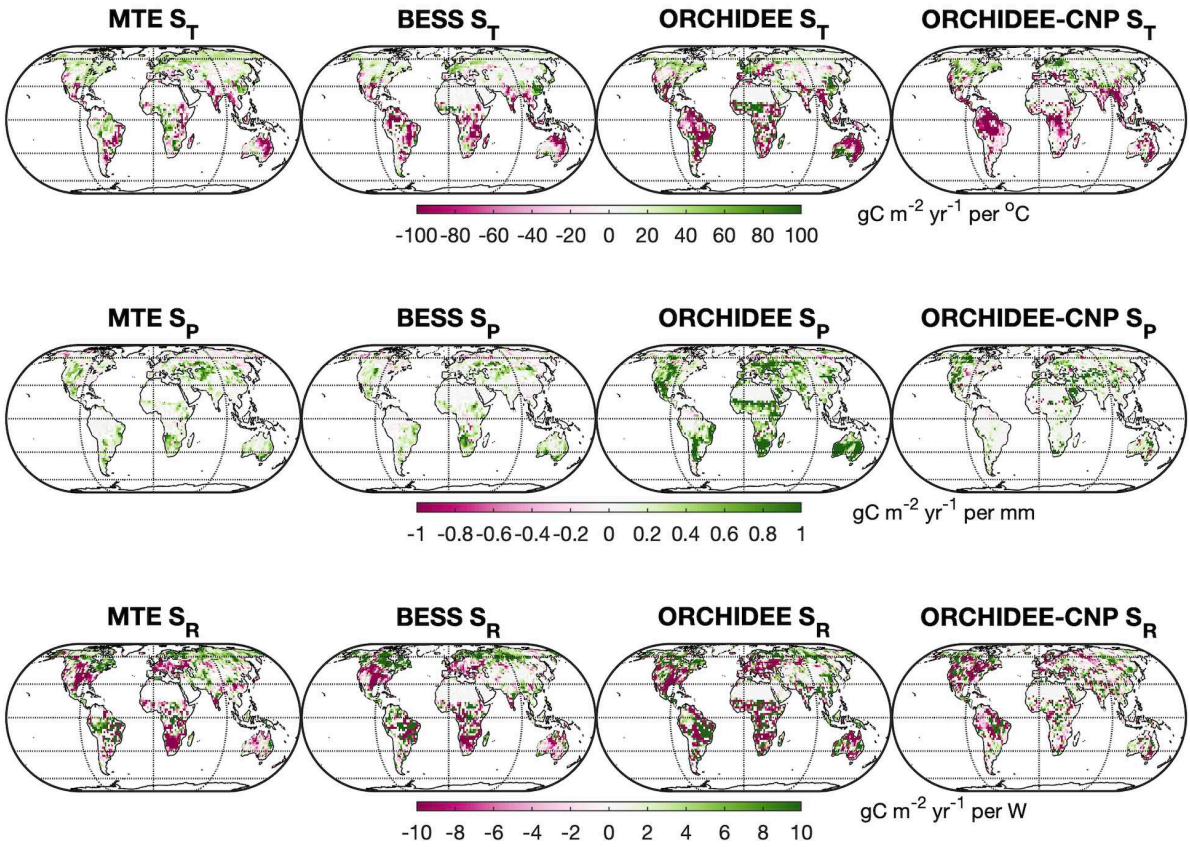


Figure S8 Global patterns of sensitivities of GPP anomalies to anomalies of mean annual temperature (S_T), annual precipitation (S_P) and shortwave radiation (S_R) derived from MTE-GPP, BESS-GPP, ORCHIDEE and ORCHIDEE-CNP.

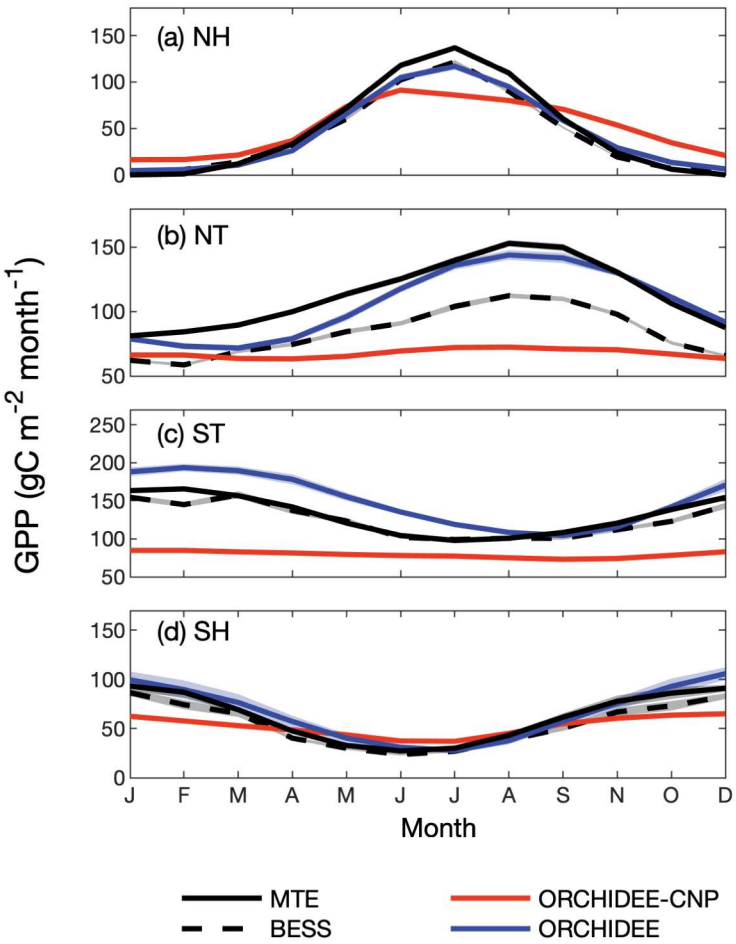


Figure S9 Seasonal cycles of GPP by BESS-GPP, MTE-GPP, ORCHIDEE and ORCHIDEE-CNP for the northern hemisphere (30°N-90 °N; a), north tropical (0°-30°N; b), south tropical (0°-30 °S; c) and the southern hemisphere (30°S-90 °S; d).

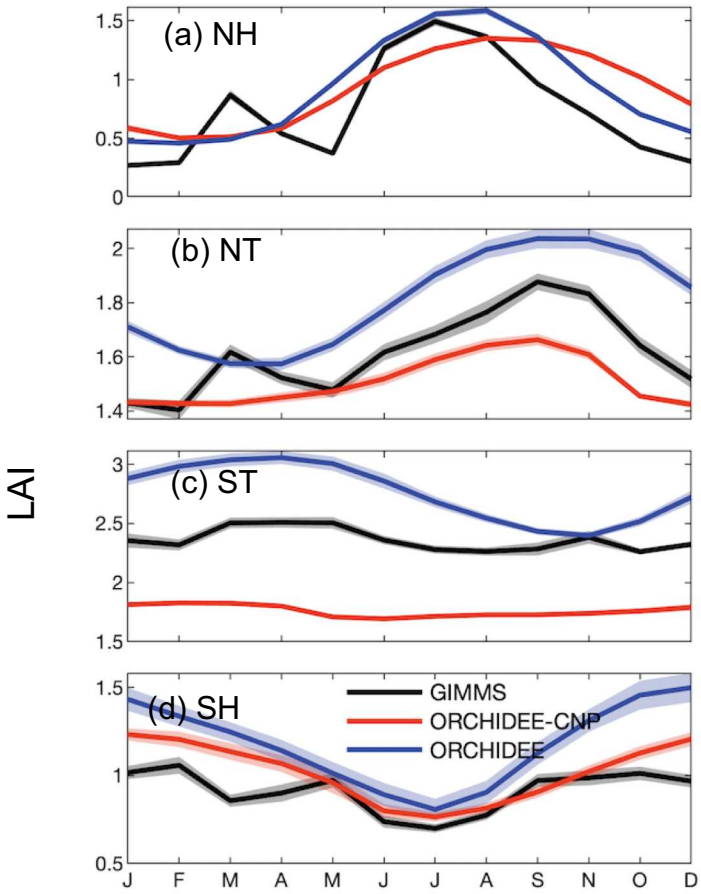


Figure S10 Seasonal cycle of LAI derived from GIMMS, ORCHIDEE and ORCHIDEE-CNP for the northern hemisphere (30°N-90 °N; a), the north tropical (0°-30°N; b), the south tropical (0°-30 °S; c) and the southern hemisphere (30°S-90 °S; d).

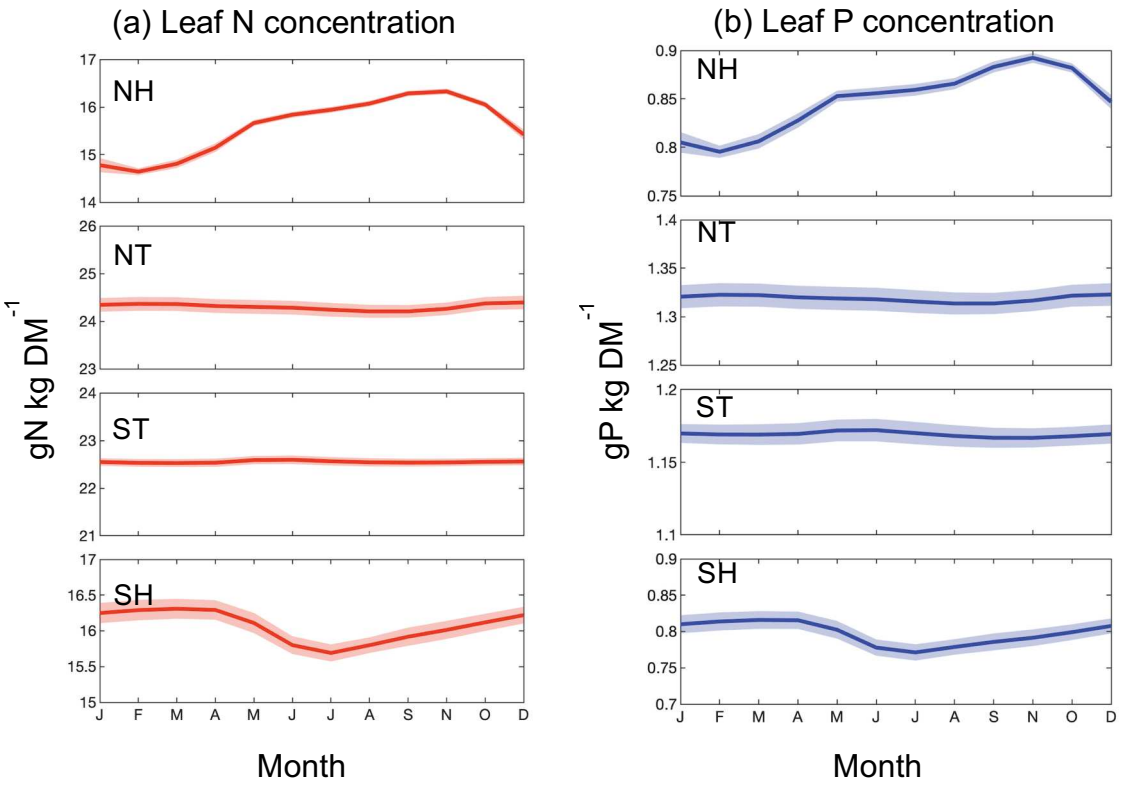


Figure S11 Seasonal cycle of leaf N (a) and P concentration (b) simulated by ORCHIDEE-CNP with excluding managed lands (crop and pastures) for north hemisphere (NH; 30°N-90 °N), north tropical (NT; 0°-30°N), south tropical (ST; 0°-30 °S) and south hemisphere (SH; 30°S-90 °S).

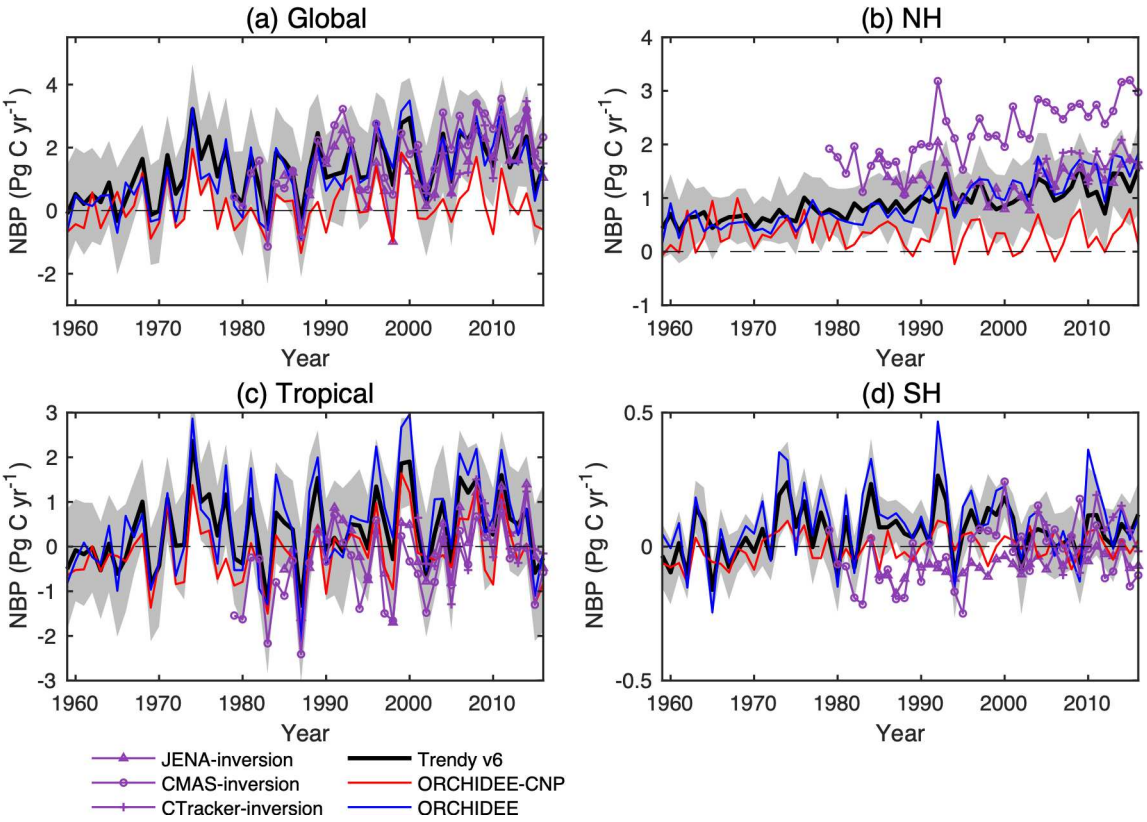


Figure S12 Comparisons of net biome productivity (NBP) between ORCHIDEE-CNP, ORCHIDEE, GCP and inversion dataset for (a) globe (b) the northern hemisphere (30°N-90 °N), (c) tropical (30°S-30 °N), and (d) the southern hemisphere (30°S-90°S). Grey shading indicates ±1σ uncertainty of DGVM results for Trendy v6.

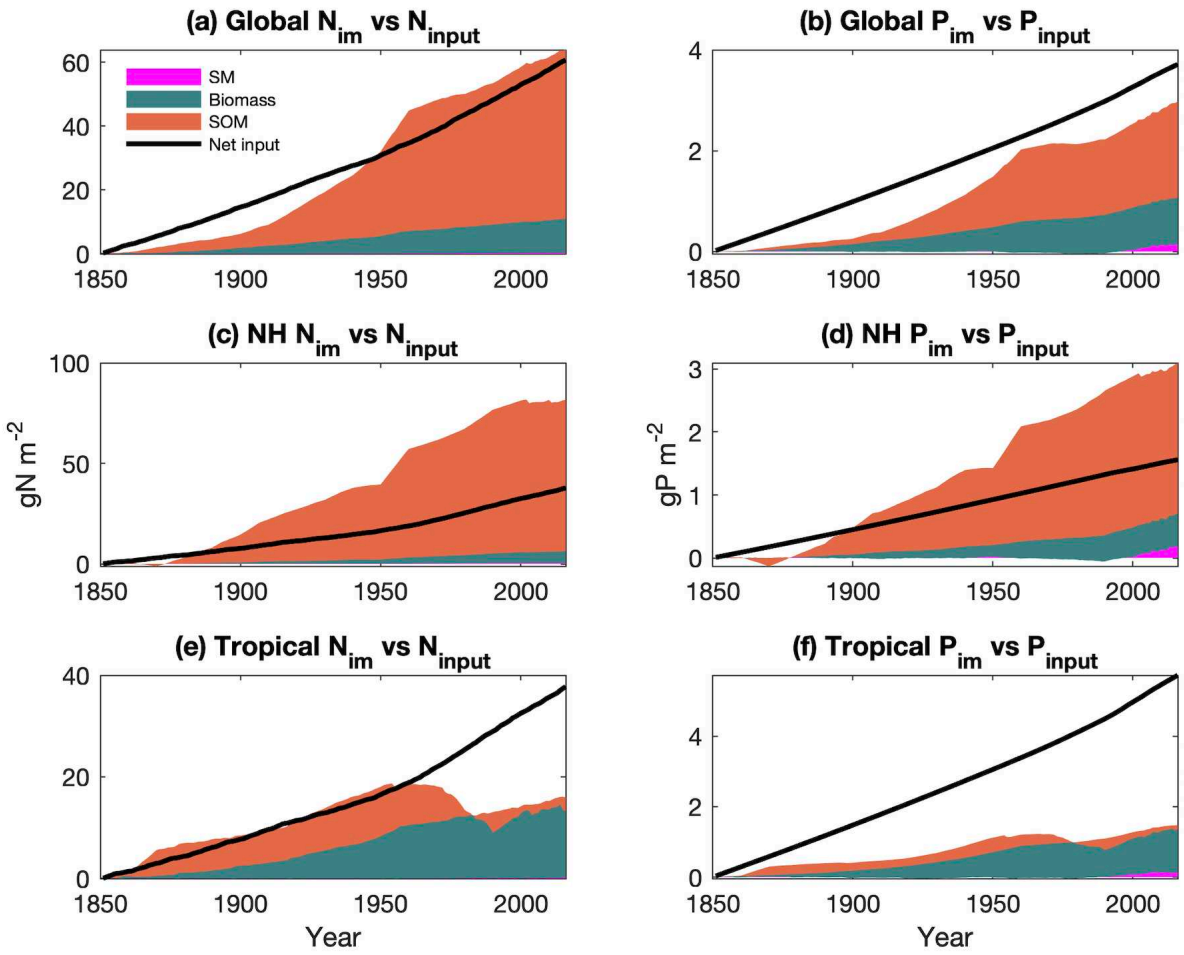


Figure S13 Changes of accumulated immobilized nutrient and net nutrient inputs (area-weighted average) for natural biomes during 1850-2016 on global, north hemisphere (NH, 30°N~90°N) and tropical region (30°S~30°N).

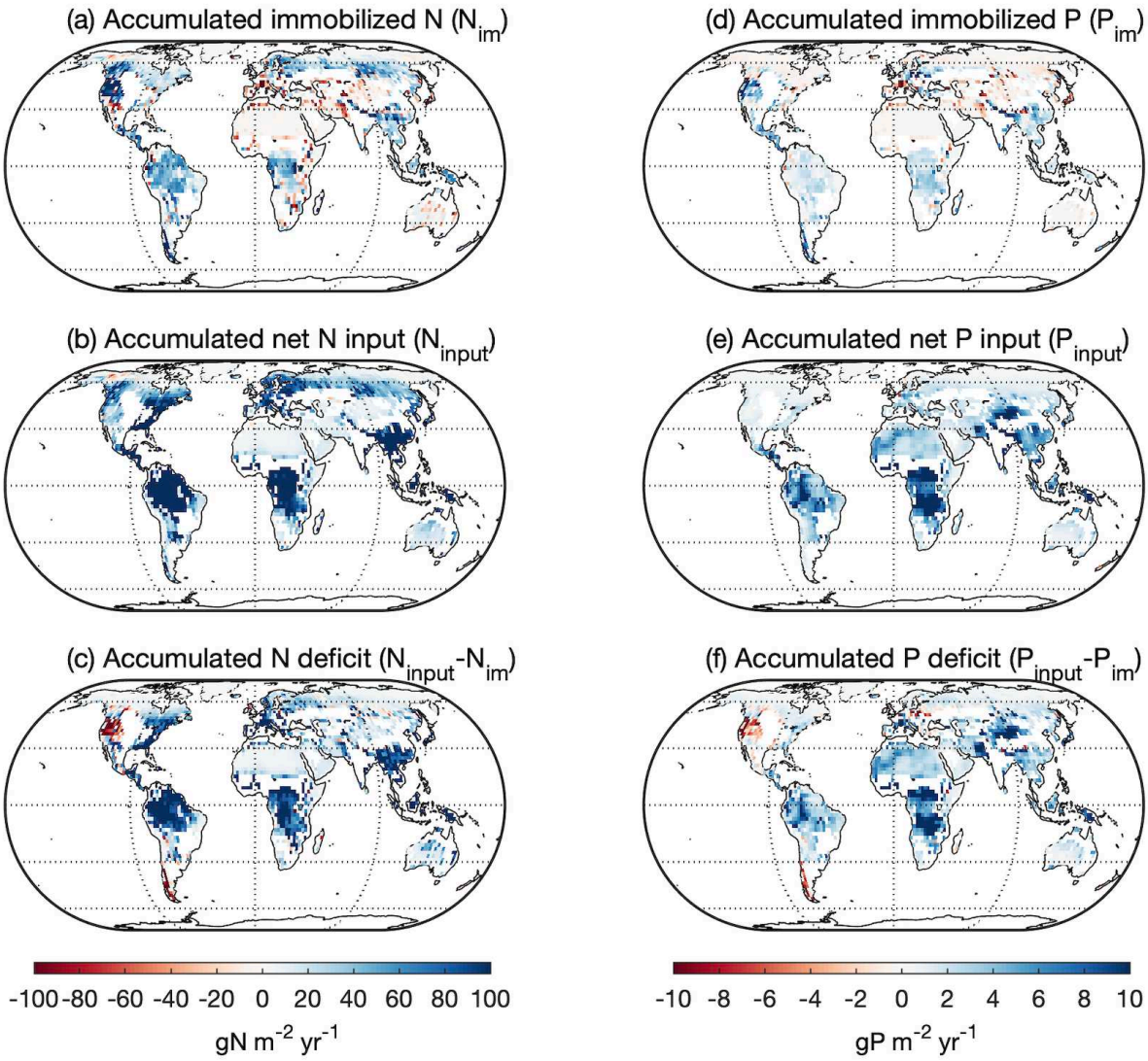


Figure S14 Accumulated immobilized nutrient, net nutrient inputs and deficit during 1850-2016 for N (a, b, c) and P (d, e, f). Grids dominated with managed biomes (cropland and pasture) (area fraction>50% in 2016) are masked.

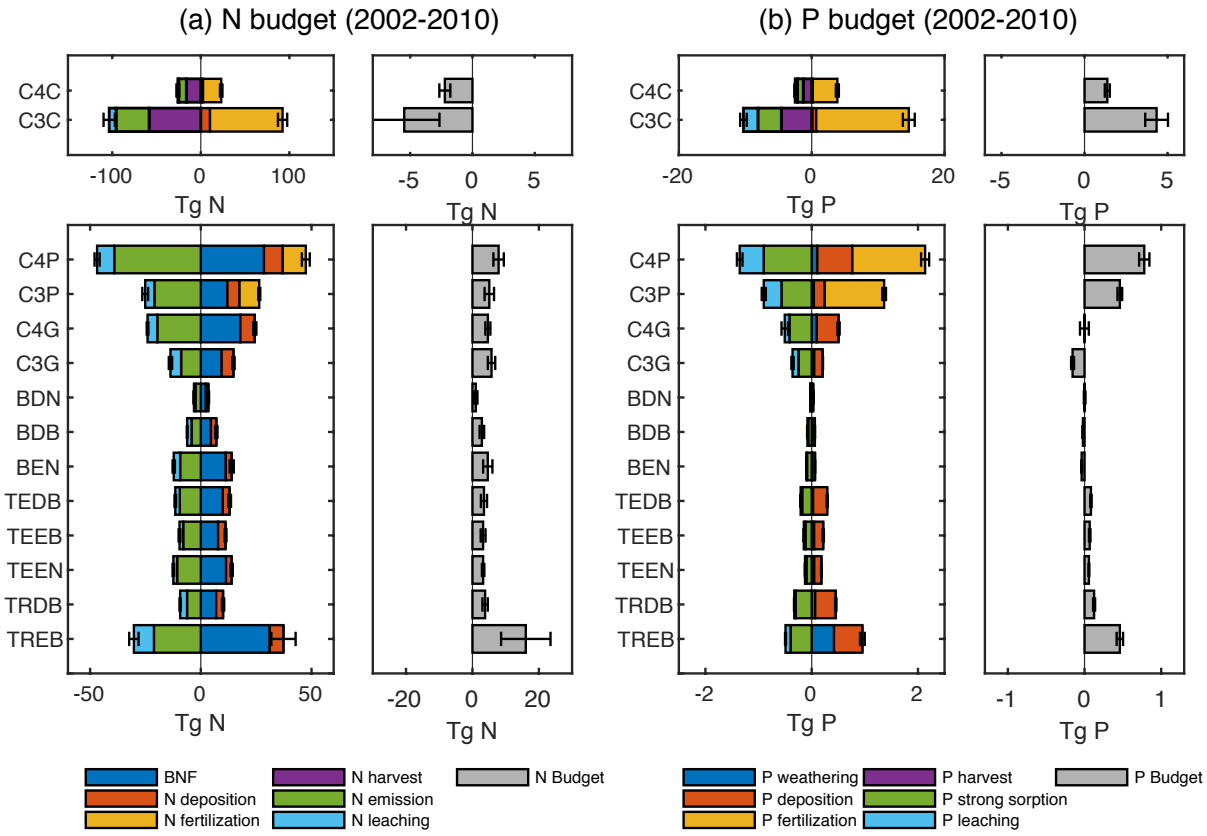


Figure S15 Total N and P budgets for each biome. Error bars indicate the standard deviation over 2002-2010. TREB: tropical evergreen broadleaf forest; TRDB: tropical deciduous broadleaf forest; TEEN: temperate evergreen conifer forest; TEEB: temperate evergreen broadleaf forest; TEDB: temperate deciduous broadleaf forest; BEN: boreal evergreen conifer forest; BDN: boreal deciduous conifer forest; C3G: C3 grassland; C4G: C4 grassland; C3P: C3 pasture; C4P: C4 pasture; C3C: C3 cropland; C4C: C4 cropland.

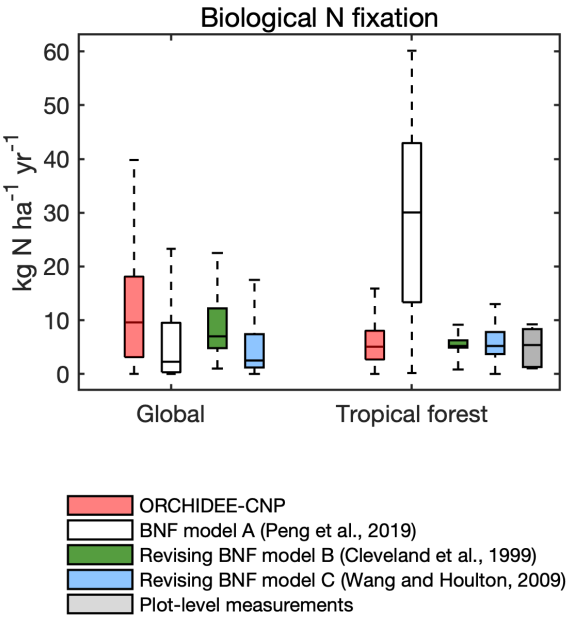


Figure S16 Global BNF and tropical forest BNF simulated by ORCHIDEE-CNP, BNF model A by Peng et al. (2019), and revising BNF model B (Cleveland et al., 1999) and C (Wang and Houlton, 2009) based on plot-level measurements in tropical forests by Sullivan et al. (2014) (grey box).

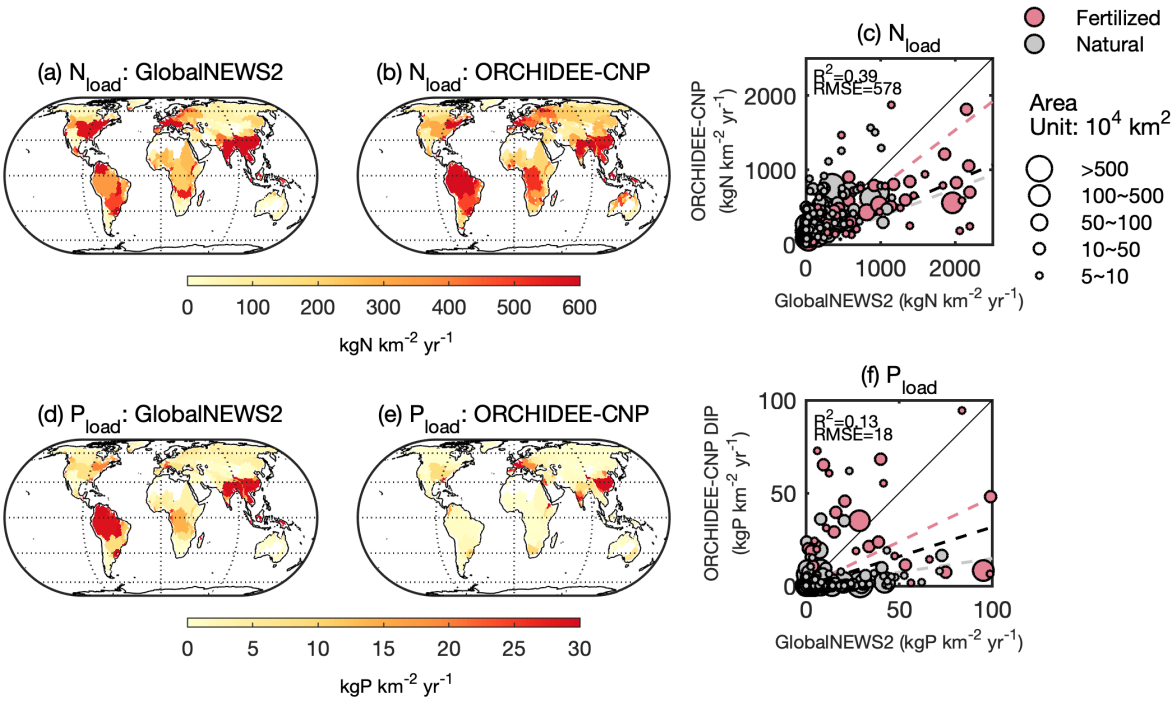


Figure S17 Global pattern of N (N_{load}, a-c) and P loads (P_{load}, d-f) from land to rivers for basins with area larger than 50,000 km^2. (c) and (f) show the comparisons of N and P leaching between ORCHIDEE-CNP and GlobalNEWS2 (Mayorga et al., 2010). Black line indicates 1:1 line. Pink dashed line indicates the linear regression line for basins with fertilized basin (N fertilization higher 1 $gNm^{-2}yr^{-1}$ or P fertilization higher than 0.5 g $Pm^{-2}yr^{-1}$), while grey dashed line indicates the linear regression line for basins with natural basin (N fertilization lower 1 $gNm^{-2}yr^{-1}$ or P fertilization lower than 0.5 g $Pm^{-2}yr^{-1}$). Black dashed lines indicate the linear regression line for all basins. R^2 and RMSE refer to the coefficient of determination and root-mean-square error between estimations of ORCHIDEE-CNP and estimates from GlobalNEWS2 (Mayorga et al., 2010) for all basins.

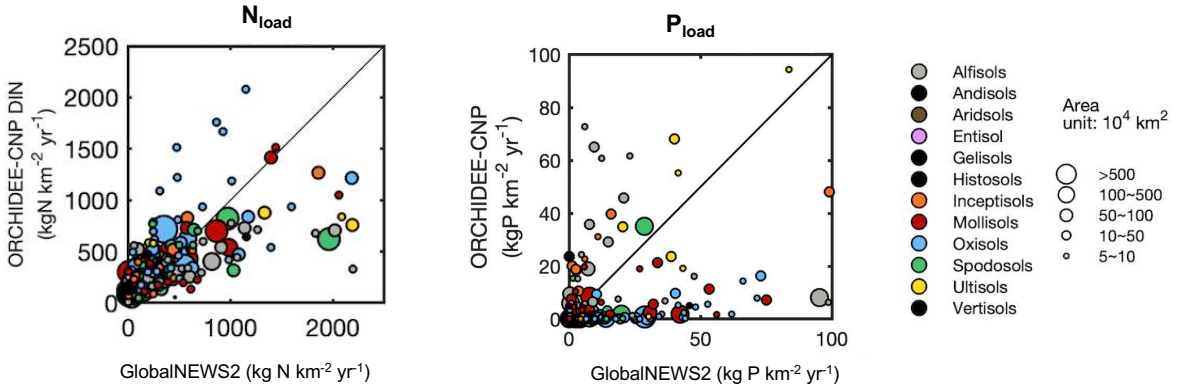


Figure S18 Global pattern of N and P load rates from land to river (N_{load} and P_{load}) on catchment scale for ORCHIDEE-CNP and GlobalNEWS2 model for basins with area larger than 50,000 km^2. (c) and (f) show the comparisons of N_{load} and P_{load} between ORCHIDEE-CNP and GlobalNEWS2. Black line indicates 1:1 line.

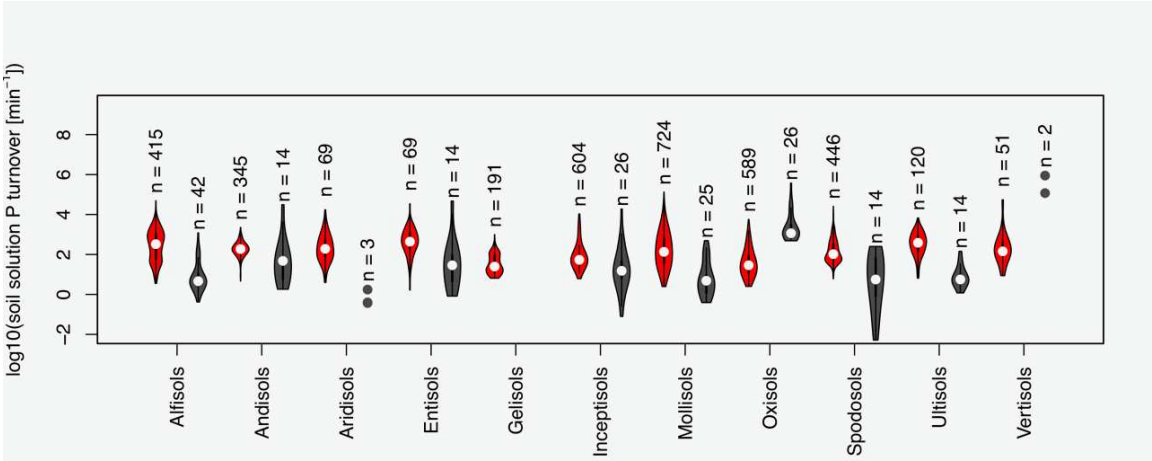


Figure S19 Violin plots of soil solution inorganic P turnover rates by soil order between ORCHIDEE-CNP (red) and measurements (black) (Helfenstein et al., 2018). The number of measurements and grid cells for each soil suborder are also shown. Open circles are medians of all grid cells within each biome, with balloons representing the probability density distribution of each value.

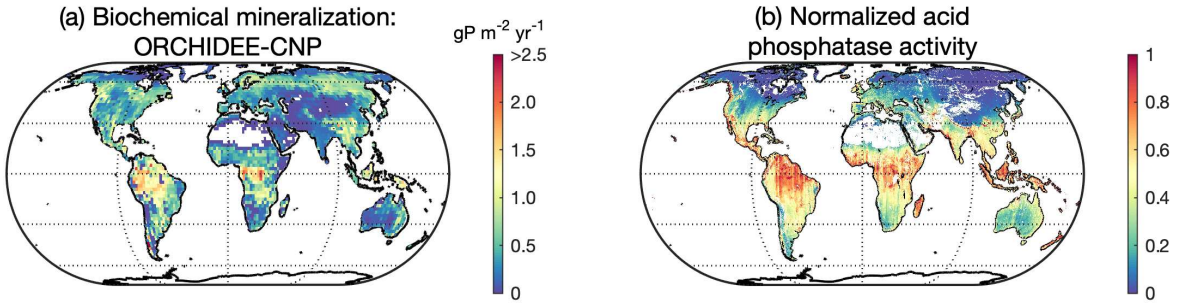


Figure S20 Global pattern of biochemical mineralization by ORCHIDEE-CNP (a) and normalized acid phosphatase activity (b).

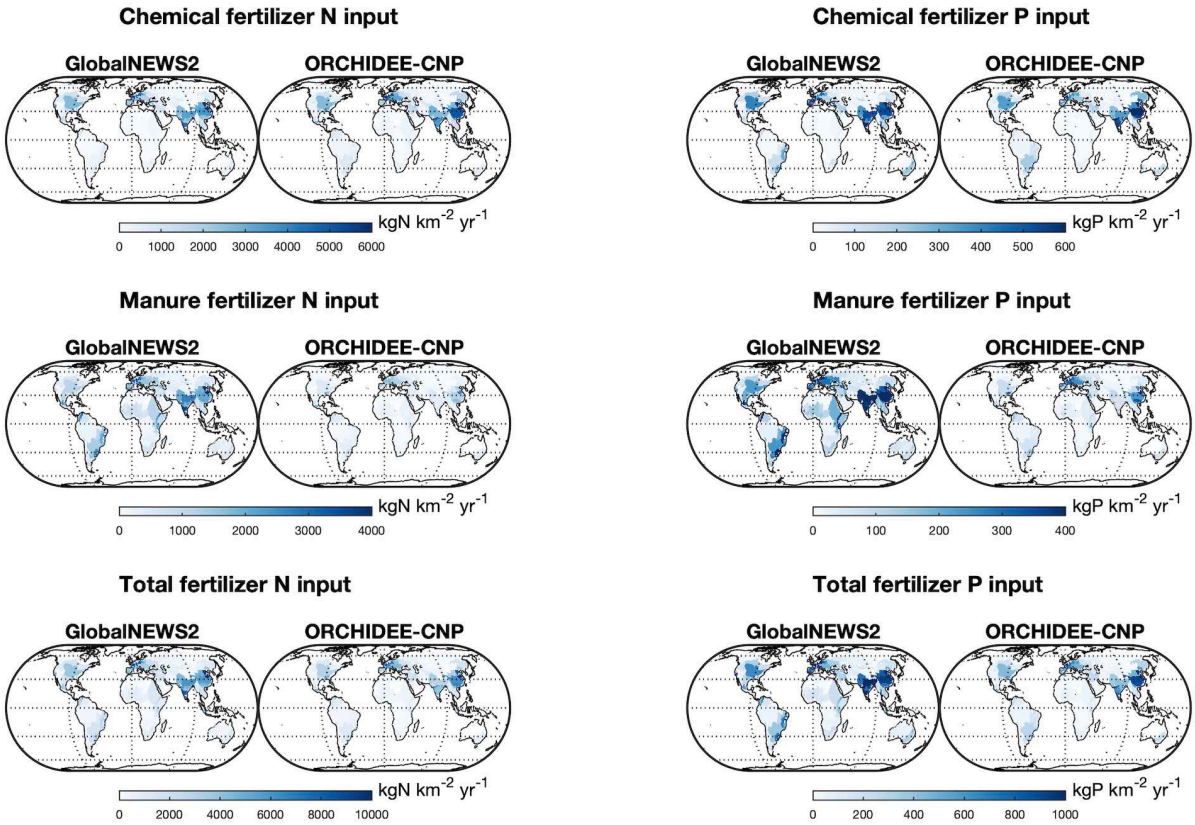


Figure S21 Chemical and manure N and P input on catchment scale by GlobalNEWS2 and ORCHIDEE-CNP.

SI Text 1: Global evaluation of nutrient enabled version land surface model ORCHIDEE-CNP v1.2 (r5986)

Sun et al.

Correspondence to: Daniel S. Goll (dsgoll123@gmail.com)

Abstract. Here, model modifications for the global version of ORCHIDEE-CNP (revision 5986) since the last published revision (r4630) are described. For complete description of nutrient cycles see Goll et al. (2017, 2018), for general model description see Krinner et al. (2005), and for plant allocation see Zaehle and Friend (2010), Naudts et al. (2015). For code access see main manuscript.

1 Modifications

The representation of the nitrogen and phosphorus cycles is described in detail in Goll et al. (2017) with some modification related to soils in Goll et al. (2018). Here we report only model modifications related to the nutrient cycles since the last published revision (r4630). If not stated otherwise, we use the parameterization of ORCHIDEE trunk version revision 4695. We use a similar nomenclature as in earlier publications (Naudts et al., 2015; Goll et al., 2017, 2018).

1.1 Photosynthetic capacity

In ORCHIDEE-CNP (r4630) (Goll et al., 2017) the photosynthetical capacity on leaf level is a function of leaf nitrogen concentration using an empirical relationship (Kattge et al., 2009). There is evidence that leaf phosphorus concentration affects photosynthesis characteristics (Walker et al., 2015). Therefore, we exchanged the original empirical relationship with a new one which links leaf nitrogen and phosphorus concentration (P^*_{leaf}, N^*_{leaf}; $\times 10^{-3}$ $\mathrm{g(N,P)g^{-1}(Dryweight)}$]) and maximum carboxylation capacity (V_{cmax}; $\mu\mathrm{mol(CO_2)m^{-2}s^{-1}}$) and maximum rate of electron transport (J_{max}; $\mu\mathrm{mol(CO_2)m^{-2}s^{-1}}$) of photosynthesis based on a compilation of measuremens on 451 plant species (Ellsworth et al. in preparation). The coefficicent relating maintenance respiration to tissue nitrogen concentration ($c_{r,maint}$) had to be re-calibrated to the relationship between new leaf nitrogen and V_{cmax} (Table 1).

1.2 Stoichiometry

In ORCHIDEE-CNP (r4630) (Goll et al., 2017), narrow plant functional type (PFT)-specific ranges for leaf phosphorus to nitrogen concentration ($pn_{leaf,min}$ - $pn_{leaf,max}$) are used. Here, we use a wide range for leaf phosphorus to nitrogen concentration which is common to all PFTs (Table 1). A globally uniform range facilitates the evaluation of the model as the predicted gradient in leaf phosphorus to nitrogen concentration along latitudes is independent of the prescribed land cover. The range of

leaf phosphorus to nitrogen concentration chosen here is in line with the validity of the leaf nutrient - photosynthetic capacity relationship (Ellsworth et al. in preparation).

As a consequence of the wide range in leaf phosphorus to nitrogen concentration, the coefficicents in the function for the plasticity of the leaf phosphorus to nitrogen concentration had to be recalibrated (replacing eq.11 in Goll et al. (2017)):

$$D_{leaf,np} = \begin{cases} D_{max}(1 - exp[-(1.3\frac{1/pn_{leaf} - 1/pn_{leaf,min}}{1/pn_{leaf,max} - 1/pn_{leaf,min}})^{7.1}]) & \text{for } P_{labile} < G_P \\ D_{max} exp[-(1.3\frac{1/pn_{leaf} - 1/pn_{leaf,min}}{1/pn_{leaf,max} - 1/pn_{leaf,min}})^{7.1}] & \text{otherwise} \end{cases} \quad (1)$$

where D_{max} [ ] is the maximum change in phosphorus to nitrogen ratio of new biomass relative to the stoichiometry of existing biomass (Goll et al., 2017), P_{labile} [$\mathrm{g(P)m^{-2}}$] the size of plant labile phosphorus pool, G_P [$\mathrm{g(P)m^{-2}t^{-1}}$] posphorus allocated to growth.

Further, we disabled the stoichiometric flexibility of sapwood and heartwood biomass which was shown to improve the realism of the simulated response of vegetation to artifically elevated carbon dioxide concentration at two forest sites (Meyerholt and Zaehle, 2015) based on an earlier nitrogen enabled version of ORCHIDEE (Zaehle and Friend, 2010).

1.3 Temperature control on biochemical mineralisation and plant nutrient uptake

The empirical function (f_{temp}; eq.5 in Goll et al. (2017)) to scale biochemical mineralisation and plant nutrient (nitrogen and phosphorus) uptake according to soil temperature was removed. f_{temp} was originally introduced in the nitrogen enabled version of ORCHIDEE (Zaehle and Friend, 2010) to avoid the accumulation of nitrogen within plants when temperatures are low. This function has proven to be not needed due to the control of nutrient uptake by the respective other nutrient (Goll et al., 2017). It was thus removed.

1.4 Phenology of grasses

In ORCHIDEE, grass senescence is controlled by cold and drought-stress (Krinner et al., 2005). The original module assumes that the turnover of all plant tissue is controlled by leaf senescence with the exception of carbohydrate reserve (i.e., synchronous senescence of leaf, root and stem as an annual grass). This is not in line with observational evidence. A review by Gill and Jackson (2000) estimated that annual root turnover is 53% (100% indicating a synchronous leaf root senescence) in perennial grasses. Belowground parts of grasses can enter a dormant state undergoing significant metabolic changes including decreased respiration (e.g., Shane et al. (2009). This facilitates the regrowth of a plant following dormancy from the belowground carbon and nutrient reserves. In fact, perennial grasslands with root survival during winter are widespread

Here, we uncouple root senescence from aboveground parts (e.g. leaves) during periods of cold and drought-stress. Instead, root turnover is solely based on root age (Krinner et al., 2005). We implement root dormancy (after leaf senescence completion) by reducing its maintenance respiration by 90% following Table 1 of Shane et al. (2009), and allow root acquisition of soil nutrients as long as root biomass exists (Malyshev and Henry, 2012). It should be noted that complete root senescence (i.e. total loss of root biomass) can still happen for extremely long droughts when maintenance respiration depletes carbohydrate reserves.

1.5 Fire

When biomass is consumed by fire, a part of its nutrients are released into the atmosphere as emissions (Krinner et al., 2005), the rest remains within the biomass residues. In ORCHIDEE-CNP, we apply the emission factors (e_N, e_P) from Akagi et al. (2011) to estimate N and P emitted from different types of biomass during burning, whereas the remaining ($1-e_N$, $1-e_P$) nutrient in biomass are transfered to the litter pools. In ORCHIDEE we use PFT-specific emission factors distinguishing between tropical forest, temperate forest, boreal forest, and C3 (pasture maintenance) and C4 grassland (savanna) (Table 1).

1.6 Mineral fertilizer and manure

Manure applications are management practices in sustaining agricultural production since thousands of years. More recently, mineral fertilizer application (both N and P) has been the the major causes of the dramatic increase in agricultural productivity in the 20^{th} century. In ORCHIDEE-CNP, both organic (manure) and inorganic (mineral) fertilization which enrich the soil nutrient pools are accounted for.

N and P in mineral fertilizer goes directly into soil mineral pools (i.e. dissolved labile P, ammonia, and nitrate). Mineral N fertilizer is treated as ammonium nitrate for simplicity with half of N as ammonia and half as nitrate. For manure applied to cropland and pasture, we assume a typical slurry application with major part as ammonium contained in the liquid slurry. We assume that 90% of N in manure is in the liquid part of the slurry (like urine), which goes into soil ammonia pool. For the solid part of the slurry, we assume it goes into litter pool with a C:N ratio of 10:1 as suggested by Soussana and Lemaire (2014), and a P:N ratio of 0.2 (a value between the ratio in ruminant manure (0.15-0.18) and monogastric manure (0.24-0.28); see Table SI3 of Lun et al. (2018) for detail).

1.7 Recalibration

Several parameters in ORCHIDEE had to be re-calibrated to new model formulations (see above) or are corrected in case of turnover of sapwood for TrEBF and TrDBF to achieve realistic wood growth rates (not shown). We further adjusted the recycling efficiency of nutrients from root ($f^N_{trans,root}$,$f^P_{trans,root}$) and leaf ($f^N_{trans,leaf}$,$f^P_{trans,leaf}$) according to global data compilations (Freschet et al., 2010; Vergutz et al., 2012). The new values of vegetation parameter are given in Table 1.

Table 1. Plant functional type (PFT) specific parameters: tropical evergreen broadleaf forest (TrEBF), tropical raingreen broadleaf forest (TrDBF), temperate evergreen needleleaf forest (TeENF), temperate evergreen broadleaf forest (TeEBF), temperate summergreen broadleaf forest (TeDBF), boreal evergreen needleleaf forest (BoENF), boreal summergreen broadleaf forest (BoDBF), boreal summergreen needleleaf forest (BoDNF), C3 grassland (C3grass), C4 grassland (C4grass), C3 cropland (C3crop), and C4 cropland (C4crop). the maximum and minimum ratio between leaf area and sapwood area ($k_{ls,max}$ $k_{ls,min}$), the turnover of tree (τ_{tree}), the coefficient relating maintenance respiration rate to tissue nitrogen concentration ($c_{r,maint}$).

	TrEBF	TrDBF	TeENF	TeEBF	TeDBF	BoENF	BoDBF	BoDNF	C3grass	C4grass	C3crop	C4crop	Source
$cn_{leaf,min}$ [g(C)g^{-1}(N)]	12.5	12.5	28.	16.	16.	28.	16.	16.	16.	16.			Zaehle and Friend (2010)
$cn_{leaf,max}$ [g(C)g^{-1}(N)]	60.	60.	75.	45.	45.	75.	45.	45.	45.	45.			Zaehle and Friend (2010)
$np_{leaf,min}$ [g(N)g^{-1}(P)]	5.	5.	5.	5.	5.	5.	5.	5.	5.	5.			(Ellsworth in review)
$np_{leaf,max}$ [g(N)g^{-1}(P)]	30.	30.	30.	30.	30.	30.	30.	30.	30.	30.			(Ellsworth in review)
$c_{r,maint}$ [$\times 10^{-2}$]	2.5	2.5	3.84	3.84	3.84	13.80	13.80	13.80	4.90	13.80	3.84	3.84	calibrated
$k_{ls,max}$ [mm $-$ 2]	4000.	4000.	1600.	1800.	2100.	1500.	3100.	2500.	2.0	1.0	2.5	3.0	calibrated
$k_{ls,min}$ [mm $-$ 2]	1475.	1475.	417.	675.	1600.	400.	2400.	1500.	1.5	.5	1.75	1.5	calibrated
τ_{sapw} [[yr-1]]	0.018	0.018	0.024	0.031	0.042	0.031	0.031	0.042	5.2	1.3	1.3	1.3	this study, Naudts et al. (2015)
τ_{tree} [yr $-$ 1]	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	–	–	–	–	calibrated
$f^N_{trans,root}$ [ ]	.275	.275	.275	.275	.275	.275	.275	.275	.275	.275	.275	.275	Freschet et al. (2010)
$f^N_{trans,leaf}$ [ ]	0.561	0.612	0.627	0.561	0.612	0.627	0.612	0.627	0.746	0.746	0.746	0.746	Vergutz et al. (2012)
$f^P_{trans,root}$ [ ]	.57	.57	.57	.57	.57	.57	.57	.57	.57	.57	.57	.57	Freschet et al. (2010)
$f^P_{trans,leaf}$ [ ]	.65	.65	.65	.65	.65	.65	.65	.65	.65	.65	.65	.65	Vergutz et al. (2012)
e_N [g(N) g(C)^{01}]	0.0067	0.0067	0.0045	0.0045	0.0045	0.0086	0.0086	0.0086	0.0056	0.0056	0.0074	0.0074	Akagi et al. (2011)
e_P [g(P) g(C)^{01}]	3.69e-06	3.69e-06	0.0	0.0	0.0	0.0	0.0	0.0	3.01e-06	3.01e-06	0.0	0.0	Akagi et al. (2011)

References

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., and Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric models, Atmospheric Chemistry and Physics, 11, 4039–4072, doi:10.5194/acp-11-4039-2011, 2011.

Freschet, G. T., Cornelissen, J. H. C., van Logtestijn, R. S. P., and Aerts, R.: Substantial nutrient resorption from leaves, stems and roots in a subarctic flora: What is the link with other resource economics traits?, New Phytologist, 186, 879–889, doi:10.1111/j.1469-8137.2010.03228.x, 2010.

Gill, R. A. and Jackson, R. B.: Global patterns of root turnover for terrestrial ecosystems, New Phytologist, 147, 13–31, doi:10.1046/j.1469-8137.2000.00681.x, 2000.

Goll, D., Winkler, A., Raddatz, T., Dong, N., Colin Prentice, I., Ciais, P., and Brovkin, V.: Carbon-nitrogen interactions in idealized simulations with JSBACH (version 3.10), Geoscientific Model Development, 10, doi:10.5194/gmd-10-2009-2017, 2017.

Goll, D. S., Joetzjer, E., Huang, M., and Ciais, P.: Low Phosphorus Availability Decreases Susceptibility of Tropical Primary Productivity to Droughts, Geophysical Research Letters, pp. 1–10, doi:10.1029/2018GL077736, http://doi.wiley.com/10.1029/2018GL077736, 2018.

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

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, 1–33, doi:10.1029/2003GB002199, http://doi.wiley.com/10.1029/2003GB002199, 2005.

Lun, F., Liu, J., Ciais, P., Nesme, T., Chang, J., Wang, R., Goll, D., Sardans, J., Peñuelas, J., and Obersteiner, M.: Global and regional phosphorus budgets in agricultural systems and their implications for phosphorus-use efficiency, Earth System Science Data, 10, doi:10.5194/essd-10-1-2018, 2018.

Malyshev, A. V. and Henry, H. A. L.: Frost damage and winter nitrogen uptake by the grass Poa pratensis L .: consequences for vegetative versus reproductive growth, PLANT ECOLOGY, 213, 1739–1747, doi:10.1007/s11258-012-0127-0, 2012.

Meyerholt, J. and Zaehle, S.: The role of stoichiometric flexibility in modelling forest ecosystem responses to nitrogen fertilization, New Phytologist, 208, 1042–1055, doi:10.1111/nph.13547, 2015.

Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., B??nisch, G., Campioli, M., Ghattas, J., De Groote, T., Haverd, V., Kattge, J., MacBean, N., Maignan, F., Merila??, P., Penuelas, J., Peylin, P., Pinty, B., Pretzsch, H., Schulze, E. D., Solyga, D., Vuichard, N., Yan, Y., and Luyssaert, S.: A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes, Geoscientific Model Development, 8, 2035–2065, doi:10.5194/gmd-8-2035-2015, 2015.

Shane, M. W., Mccully, M. E., Canny, M. J., Pate, J. S., Ngo, H., Mathesius, U., Cawthray, G. R., and Lambers, H.: Summer dormancy and winter growth : root survival strategy in a perennial monocotyledon, NEW PHYTOLOGIST, 183, 1085–1096, doi:10.1111/j.1469-8137.2009.02875.x, 2009.

Soussana, J. F. and Lemaire, G.: Coupling carbon and nitrogen cycles for environmentally sustainable intensification of grasslands and crop-livestock systems, Agriculture, Ecosystems and Environment, 190, 9–17, doi:10.1016/j.agee.2013.10.012, 2014.

Vergutz, L., Manzoni, S., Porporato, A., Novais, R. F., and Jackson, R. B.: Global resorption efficiencies and concentrations of carbon and nutrients in leaves of terrestrial plants, Ecological Monographs, 82, 205–220, doi:10.1890/11-0416.1, 2012.

Walker, A. P., Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Asao, S., Hickler, T., 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, 5, 1–20, doi:10.1002/2014GB004995.Received, 2015.

Zaehle, S. and Friend, A. D.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, GLOBAL BIOGEOCHEMICAL CYCLES, 24, 1–13, doi:10.1029/2009GB003521, 2010.

Chapter 4 A new method to reduce the very long time to reach equilibrium state for nutrient-carbon models based on machine learning

Summary

The integration of observational data into land surface models (LSMs) based on parameter optimizations require a high computational efficiency of the LSMs. Because a CNP model has parameters that determine the equilibrium state of fluxes and pools, each time a parameter is varied, the equilibrium has to be re-simulated. This looks trivial but time scales needed to reach a balance between input and output carbon (C), nitrogen (N) and phosphorus (P) fluxes and a steady state equilibrium (i.e. spin-up) of their pools are of several millennia. Even when land models are numerically simple, or work on coarse time steps, re-computing their spin-up and equilibrium state each time in the context of varying dozens of parameters is already a huge numerical challenge. For ORCHIDEE-CNP that solves the surface energy budget with a full physics each 30 minutes' time step with a coupling with C, N, P biogeochemistry, we found that 1) a spin up of 10 thousand years is needed to achieve steady state; and 2) this spin-up at 2° global resolution takes 6 weeks. This is for one set of parameters. If a modeler wants to change a parameter value or an equation, everything has to be redone before doing any scientific study. This is clearly a limitation of the research that must be overcome.

In this chapter, I designed and tested a general Machine Learning (ML)-based procedure (Bagging ensemble) for acceleration of the equilibration of CNP cycles in LSMs which are responsible for the low computational efficiency like ORCHIDEE-CNP.

This ML-based acceleration approach (MLA) requires to spin-up only a small subset of model pixels (14.1%) from which the equilibrated state of the remaining pixels is estimated by ML, thus decreasing the computational demand by about one order of magnitude. MLA predicts the equilibrium state of soil, biomass and litter C, N and P on both PFT and global scale well (R^2>0.85) for ORCHIDEE-CNP. Biases in global and regional land carbon balance for the recent decades due to the use of MLA are marginal.

The reduction in the computational consumption opens the opportunity of data assimilation using the ever-growing observation datasets. Different from previous acceleration approaches with precondition of linear process and fixed C-N ratios, this approach breaks those limits and can be widely used in LSMs, no matter if the model processes are linear or nonlinear.

This chapter will be submitted as '*Sun Y.*, *Goll DS., Ciais P., Huang Y., Wang YP., Wang Y., Equilibration of coupled biogeochemical cycles in land surface model based on machine learning drastically increases the computational efficiency'*.

4.1 Introduction

Nitrogen (N) and phosphorus (P) are essential nutrients for photosynthesis, respiration and growth of plants and metabolism of soil microbes, thereby are critically important for primary production and carbon accumulation by land biosphere (Elser et al., 2007; Norby et al., 2010; Cleveland et al., 2013; Hou et al., 2019; Gärdenäs et al., 2011; Melillo et al., 2011). Furthermore, limitation of N and P on land carbon uptake was projected to exacerbate under climate change and rising atmospheric CO_2 (Wieder et al., 2015; Sun et al., 2017; Fleischer et al., 2019).

As a result, N and P cycles have been implemented into an increasing number of land surface models (LSMs) (e.g. Zaehle et al., 2014; Wang et al., 2010; Yang et al., 2014). Simulations of those LSMs commonly use steady states of all biogeochemical pools as the initial conditions for most applications, e.g. such as the simulations of the recent past (Thornton & Rosenbloom, 2005). The steady states are typically obtained by running the model with constant or periodic boundary conditions for hundreds to thousands of years (i.e. spin-up) (Bondeau et al, 2007; Randerson et al., 2009; Xia et al., 2012). Models including P cycles generally require the longest duration to reach steady state (~10k+ years, Sun et al, submitted) due to the long residence time of P compared to C and N. Such long spin-up simulations may lead to a high burden of computation resources, in particular as the duration of the spin-up simulation greatly exceeds the actual model experiment which typically spans decades to few centuries.

For LSMs with weak non-linear feedback from belowground SOM state to aboveground processes, the slow equilibration of soil organic matter (SOM) can be shortcut by analytically solving differential equations governing SOM dynamics by relating input of litter to soil carbon pool size (Xia et al., 2012). As an example, in the ORCHIDEE LSM, a 200yr long simulation is performed to equilibrate the aboveground C cycle including litter soil carbon inputs, followed by analytical spin-up for SOM using results from the aboveground spin-up (Krinner et al., 2005). LSMs in which vegetation processes depend on soil biogeochemistry, as in case of accounting for plant nutrient, require alternative solutions. Several acceleration approaches for CN models have been proposed such as accelerated decomposition rates for soil carbon pools (Thornton and Rosenbloom, 2005), decelerated bulk denitrification and leaching method (Shi et al., 2013), exact solution of a linearized system (Qu et al., 2018), semi-analytical solution for soil organic C and N pools (Xia et al., 2012) and gradient projection method (Fang et al., 2014).

However, those approaches are based on linear or weak nonlinear differential equations. Therefore, they are not applicable to LSMs which include non-linear soil nutrient processes (Wang et al., 2012, Goll et al., 2017) such as inorganic soil P transformation. As a consequence, an efficient spin-up procedure suitable for new generation of LSMs including complex C, N and P interactions among multiple time-scales is missing. This hampers the application of model-data integration systems and uncertainty assessments as they rely on the realization of large numbers of (spin-up) simulations.

Machine learning (ML) approaches build mathematical models based on 'training data' for extrapolation, which have been widely used in ecology and earth science fields. Reichstein et al. (2019) suggested that the hybrid between physical process models and data-driven ML for a priority direction in earth system model development, and recommended that ML can surrogate sub-models or emulations of process models thanks to the advantages in computation efficiency.

In this study, we designed a general ML-based procedure for acceleration of equilibration of coupled C, N and P cycles for LSMs with strong nonlinear processes (MLA) (Figure 1). First, we simulate the biogeochemical equilibrium of a subset of representative model pixels using a LSM. Second, we use that data to train a ML based model to predict the steady state of the C, N, P of the remaining pixels. As the computational efficiency of LSM scales linearly with the number of pixels, this approach reduces the computation time as only a subset of pixels is simulated while the remaining is derived from ML which has a much higher computation efficiency than a LSM. We tested the

applicability of this procedure and the accuracy of this procedure for one nutrient enabled LSMs ORCHIDEE-CNP.

4.2 Methods

4.2.1 Overview: a machine learning-based procedure for accelerated spin-up

Fraction of pixels (training pixels) representing the climates, edaphic environments, and N and P deposition features were selected (section 4.2.2) to perform ordinary spin-up simulations with the land surface model until all biogeochemical pools reach to their steady-states (Figure 4.1). The steady-state criterion for carbon fluxes is the 100-year mean change in total ecosystem carbon stock must be below 1 g m^{-2} yr^{-1}. Using bagging ensemble regression trees method, we built a model that relates the biogeochemical pools at equilibrium (*Response Y*) and the model forcing data (*Predictor X*) for the training pixels. That model was then used to calculate equilibrium biogeochemical pool sizes for all other pixels within the model domain (section 4.2.3). In developing the model based on machine learning, we found that including information on net primary productivity and leaf area index from an additional short global spin-up simulation (NPP_g, LAI_g) into *Predictor X* strongly improved the predictive skills of the ML models (section 4.2.4). Thus, two sets of land surface model simulations are needed to build the ML model: a set of long (10k year) site-level simulations for the training pixels (hereinafter *site-run*), and a short (300 or 450yr) global simulation (hereinafter *pre-run*).

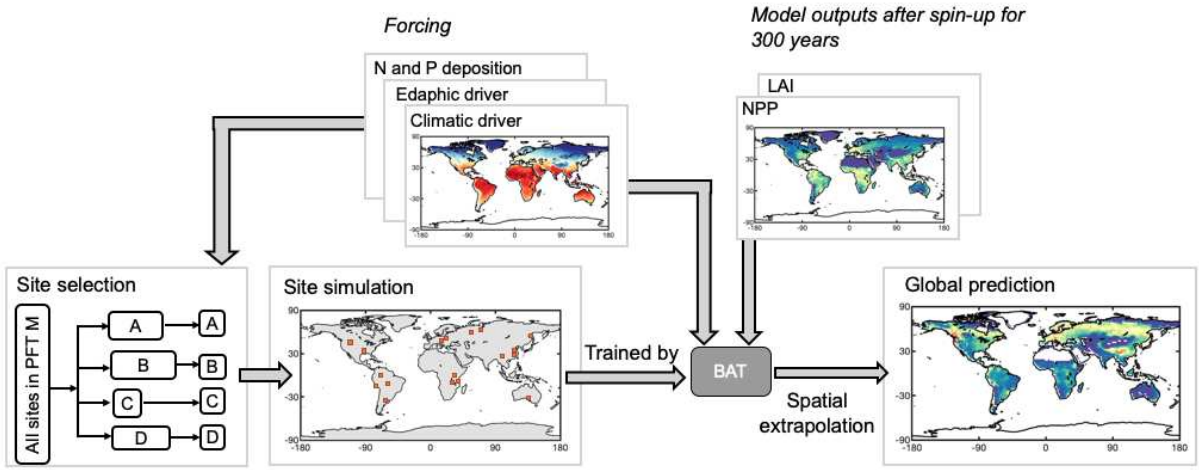


Figure 4.1 Schematic of acceleration procedure for spin-up of ORCHIDEE-CNP.

We evaluated the ML model performance on training sites using leave-one-out cross validation, which can help test the reproductively of ML prediction model on independent data (Cawley, 2006), before applying ML model to all pixels using the global forcing datasets (Table 4.1). Finally, we evaluated the ML model performance at plant functional type (PFT) and global scales (Figure 4.2) using results from a long (10k yr) global simulation. If successful, the ML procedure will lead to a substantial decrease in computational time. We tested the procedure on ORCHIDEE-CNP with coupled C, N and P cycles. The model information can be found below.

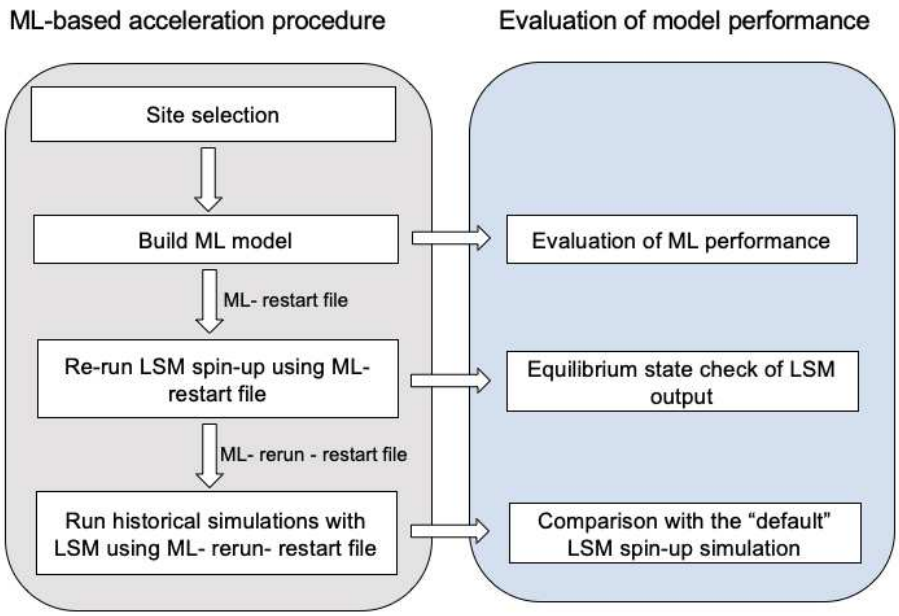


Figure 4.2 Flow chart for the build-up of prediction model and model evaluation.

4.2.2 Land surface model ORCHIDEE-CNP

ORCHIDEE-CNP v1.2 (r5986) (Goll et al., 2017, 2018; Chapter 3) were simulated on a spatial resolution of 2^o x 2^o and each pixel contains information on biogeochemical cycles of multiple plant functional types (PFTs) using a tiling approach. As steady-state conditions, the model requires 467 variables at each pixel and tile: 172 variables are state variables of C, N and P cycles (i.e. C, N and P in soil, biomass and litter, NPP, GPP and LAI); 200 variables are auxiliary variables which equilibrate with a few decades or shorter (e.g. soil temperature, leaf temperature); the remaining 95 are variables related to time-invariant forcing data. Only the state variables need to be predicted by ML, while the auxiliary variables and the invariant forcing variables are taken from the short global spin-up simulation (Figure S4.1; see below).

The model was driven by gridded information of climatic, edaphic, land cover and atmospheric nutrient deposition variables as well as globally uniform variables like atmospheric CO_2 concentration (Chapter 3). The ORCHIDEE-CNP spin-up used the cyclic climate forcing from the period 1901-1920 from CRU-JRA-55 meteorological dataset (Harris et al., 2014; Kobayashi et al., 2015). Edaphic forcing variables used for ORCHIDEE-CNP spin-up include soil texture (Zobler, 1986), pH (Global Soil Data Task Group, 2000), bulk density (HWSD, 30 FAO/IIASA/ISRIC/ISSCAS/JRC, 2012) and soil types (following the USDA Soil Taxonomy, Goll et al., 2017). 15 Plant Functional Type (PFT) area fractions were used in ORCHIDEE-CNP. The ORCHIDEE-CNP spin-up used the constant PFT map for year 1700 from HYDE land use data set (HYDE v3.2; Klein Goldewijk et al., 2017a, b). Gridded maps of atmospheric N and P deposition rates of the year 1850 were used for the spin-up. Detailed information for the forcing datasets can be found in Chapter 3.

4.2.3 Machine Learning substitute models

4.2.3.1 Bagging ensemble regression tree

Bagging (known as bootstrap aggregation) is an ensemble technique proposed by Breiman (1996), that combines multiple predictions from 'weak learners' built using randomly generated training sets to form one final prediction. This 'strong learner' improves the stability and accuracy of machine learning algorithms. Bagging ensemble is fast and can efficiently handle unbalanced and large databases with thousands of features (Breiman, 1996).

In this study, we grow 100 trees (i.e. weak learners) in each ensemble for predicting the

equilibrium steady-states of NPP, LAI and C, N and P storage for each soil, biomass and litter compartment. 90% of total training pixels were randomly selected to train each tree. The minimum number of samples of every tree leaf is set as 5. We increase the weights for samples falling out of the 10_{th} ~90_{th} quantiles as 2~6 times of other samples to reduce the overestimation in the high Y region (Table S1).

4.2.3.2 Training site selection

The use of a representative training set is important for accurate predictions by machine learning models. The representativeness is related to the quality and size of the training data (Kavzoglu, 2009). A high quality of the training data refers to evenly distributed samples across different classes (i.e. balanced principle), which guarantees a high prediction accuracy in both major and minor class on test datasets (e.g. Buda et al., 2018; Rendón et al., 2020). To satisfy this balance principle in training data for ML approaches, we designed a procedure for training site selection, which includes a K-mean cluster method aiming for balanced *Predictor X* (Farquad & Bose, 2012) and Synthetic Minority Oversampling TEchnique (SMOTE) aiming for balanced *Response Y* (Chawla et a., 2002) (SI Text S1). Increasing the size of sampling sites will improve the performance of machine learning models. To determine the suitable size of training sets, we performed the same machine learning method using different number of training sites (Figure S4.4).

4.2.4 Predictors

Original climatic variables used to force land surface simulations with a time step of 6 hours were aggregated into monthly values. From the monthly climatic variables, we derived 16 climatic predictors for ML models, which include 5 temperature-related variables (amplitude, maximum, minimum, mean and standard deviation of monthly temperature), 2 precipitation-related variables (mean annual precipitation and standard deviation of monthly precipitation), 3 growing season (GS) -related variables (monthly mean temperature > -4^{o}C) (accumulated monthly temperature and precipitation during GS and GS length), 2 interaction variables between temperature and precipitation (the product of mean temperature and precipitation, and the product for GS temperature and precipitation) and mean values for specific humidity, surface pressure and radiation (Table 4.1). Those climatic predictors strongly control the C, N and P inputs into soils and the mineralization rates of SOM (Wiesmeier et al., 2019), and thus affect the size of soil C, N and P pools.

In addition to the climate, soil priority and nutrients state are also regarded to affect the accumulation of SOC (Wiesmeier et al., 2019; Rovai et al., 2019). Nine edaphic and deposition predictors are directly extracted from the land surface model forcing datasets, which include soil pH, clay and silt fraction, soil bulk density, shield factor and N (NHx and NOy) and P deposition (Table 4.1).

Vegetation productivity drives the C, N and P inputs to soils which cannot just be "inferred" by a ML model driven by climate predictors only. Thus, we included predictor NPP_g and LAI_g which substantially improve the ML model performance (Figure 4.1; Figure S4.5). Those two variables are from the *pre-run* of ORCHIDEE-CNP for the 300th year. The length of the spin-up simulation is a tradeoff between ML model performance and computation costs of running LSM.

Table 4.1 Predictors used in prediction model for ORCHIDEE-CNP. The climatic predictors, nutrient deposition predictors and LAIg and NPPg are yearly variables.

Abbreviations	Variables	Units
Tamp	Amplitude of monthly temperature	
Tmax	Maximum monthly temperature	
Tmin	Minimum monthly temperature	
Tmean	Mean monthly temperature	^{o}C
Tstd	Standard deviation of monthly temperature	
AT_gs	Accumulated temperature during growing season (monthly temperature > -4^{o}C)	
Rainf	Mean annual precipitation	
Rainf_std	Standard deviation of monthly precipitation	kg m^{-2} yr^{-1}
Pre_gs	Precipitation during growing season (monthly T > -4^{o}C)	
Qair	Near surface specific humidity	kg kg^{-1}
Psurf	Surface pressure	Pa
SWdown	Shortwave down radiation	W s^{-1}
LWdown	Longwave down radiation	
INT_Pre_Tmean	Rainf · Tmean	^{o}C kg m^{-2} yr^{-1}
INT_Pre_AT_gs	Pre_gs · AT_gs	
GSL	Growing season length	month
Clay	Clay fraction	%
Slit	Slit fraction	%
Soilph	Soil pH	-
Ndep_noy	Nitrogen deposition (NOy)	gN m-2 yr-1
Ndep_nox	Nitrogen deposition (NHx)	
Pdep	Phosphorus deposition	
Soilbulk	Soil bulk density	g soil cm^{-3}
Soilshield	Soil shield factor	-
Soilsuborder	Soil suborders (categorical variable)	-
NPP_g	Annual net primary productivity after 300 years of spin-up	gC m2 yr-1
LAI_g	Annual leaf area index after 300 years of spin-up	-

4.2.5 Re-run the spin-up till equilibrium state using ML predictions

Using ML approaches to predict equilibrium state of pools of those unselected pixels will inevitably yield errors. Thus, ML-based predictions need to be adjusted with additional fully coupled spin-up simulation, which can help reduce the bias induced by the imperfect ML initialization. We re-run the spin-up for 1000 years starting from the previous fully coupled simulation (i.e. *pre-run*) but replacing the state variables (i.e. C, N and P pools, NPP and LAI) using the predicted initial pool sizes by ML (hereinafter *re-run*) (Figure 4.2).

4.2.6 Evaluation of ML-based spin-up

To evaluate the applicability of ML-based spin-up, we assessed the bias in C, N and P storages

induced by the ML models (4.2.6.1) and by the whole ML approach (hereinafter MLA) which includes the ML prediction and *re-run* (4.2.6.2) (Figure 4.2).

4.2.6.1. Bias in C, N and P storages induced by ML compared to the "*true equilibrium*" from the coupled simulations

In order to test the accuracy of MLA, we performed a traditional fully coupled spin-up simulation for 10k years to have the reference "*true equilibrium*" states. We compared the predicted C, N and P pool sizes by ML against the "true equilibrium" values at PFT-level and global scales. We used four metrics to assess the model performance: coefficient of determination (R^2), relative bias, relative mean-square deviation (*reMSE*), and the regression slope between the results from "*true equilibrium*" and ML (slope). Relative bias is defined as:

$$rs = \frac{I_{ML} - I_{equi}}{I_{equi}} \times 100\% \qquad \text{(Eq. 4.1).}$$

, while *reMSE* is defined as:

$$reMSE = \frac{\sum_{j=1}^{n} \left(X_{ML,j} - X_{equi,j}\right)^2}{\sum_{j=1}^{n} \left(X_{equi,j} - \hat{X}_{equi,j}\right)^2} \text{ (Eq. 4.2).}$$

X_{ML} and X_{equi} are values from MLn and Equi respectively, $\hat{X}_{equi,j}$ is the mean value across all sites.

4.2.6.2 Impact of errors induced by MLA on simulated historical C balance

To assess the impact of bias of MLA on historical simulations, we performed two historical simulations with ORCHIDEE-CNP for the period 1700-2017 starting from (1) the *"true equilibrium"* state (*Simulation EQ*) and from (2) the MLA generated state (*Simulation ML*). Both of those two simulations were forced by historical climate forcing, land cover changes and management (i.e. mineral fertilizer application, crop harvest; Chapter 3), and atmospheric CO_2 concentrations (S3 type simulation). Then we evaluated the biases in global spatiotemporal pattern of net biome productivity (NBP) from *Simulation ML* against *Simulation EQ* to test the applicability of ML approach on historical simulation. Other variables related to C, N and P cycles were also evaluated using the metrics mentioned in 4.2.6.1.

4.3 Result

4.3.1 Evaluation of ML models

The reproducibility of ML models on independent data is sufficiently high on PFT level, as indicated by the good performance ($R^2 > 0.8$) in leave-one-out cross validation (LOOCV) on training sets for passive SOC and the aboveground heartwood carbon (HeartAB C) pools (Table S4.4). We extrapolated the ML models to global pixels and found that the ML predictions of soil, vegetation and litter C, N and P pool on PFT level compares well to the *"true equilibrium"* pools simulated by ORCHIDEE-CNP. The ML models can generally well represent the true equilibrium values of slow changing C pools, including all of the soil C pools and 4 woody vegetation (aboveground sapwood: SapAB; belowground sapwood: SapBE; aboveground heartwood: HeartAB; belowground heartwood: Heart BE) and woody litter C pools, with high R^2 (>0.75), low reMSE (<25%) and comparable regression slope that close to 1 (0.95~1.14) (Figure 4.3). However, ML models show deficiencies in predicting fast-changing C pools, such as leaf, root and vegetation labile C pools (R^2=0.23~0.97, *reMSE*=7%~95%, slope=0.5~1.07). Relatively large biases in the fast and small biases in slow-changing pools will be corrected by re-running spin-up (section 4.2.5 and 4.3.3). Performance of ML for N and P pools is comparable to that for C pools with generally higher R^2 (>0.8) and lower *reMSE* (<15%) except for soil P pools (Figure 4.4 & 4.5). The somewhat different performance for soil P pools could be related to stoichiometric decoupling of soil organic P, while C and N are coupled in ORCHIDEE-CNP (Goll et al., 2017).

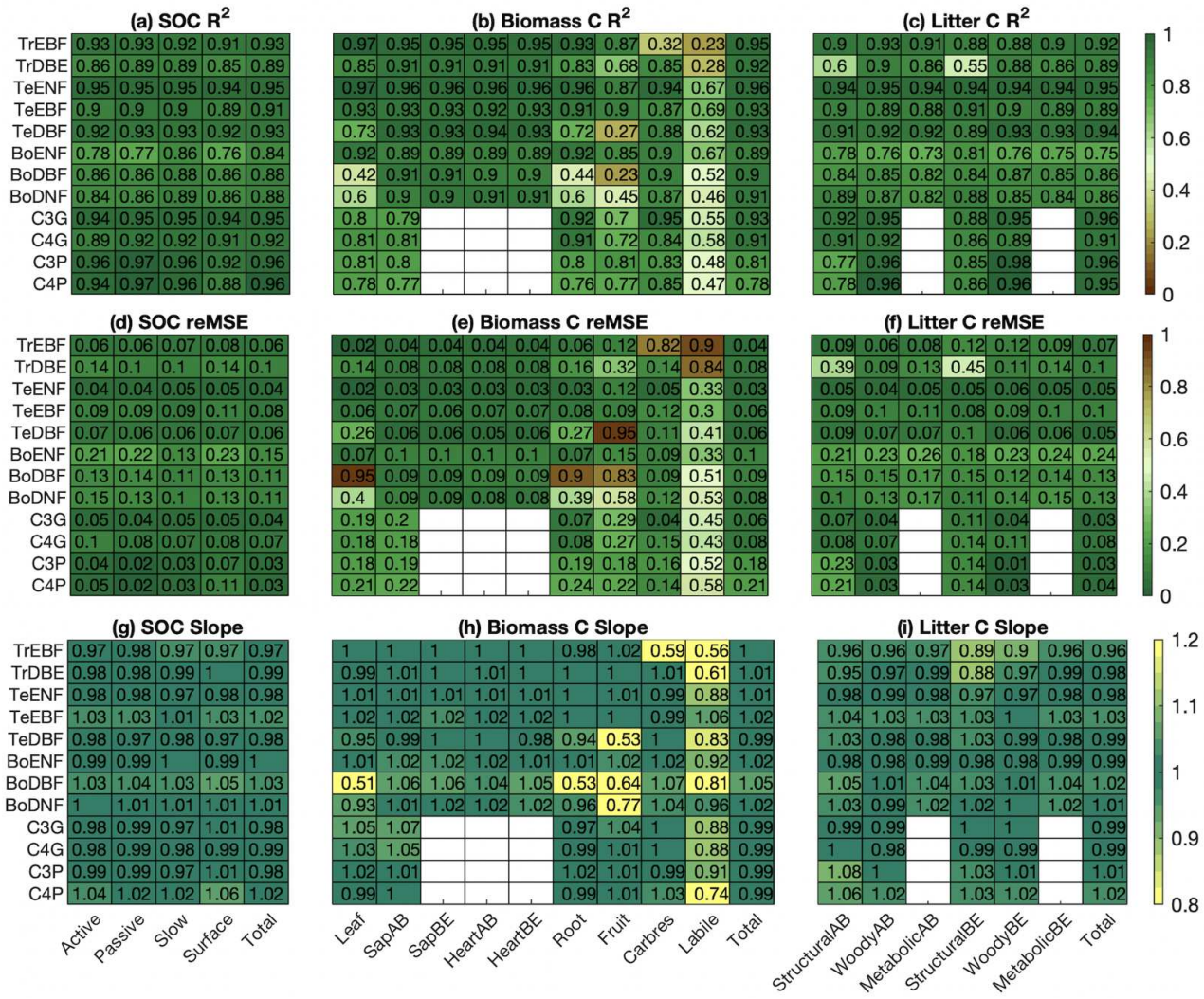


Figure 4.3 The performances of machine learning models (ML) on soil, biomass and litter C storages for each PFT compared to the 'true equilibrium' state simulated by ordinary spin-up (Equi). Three statistics represent the model performance: coefficient of determination (R^2), relative mean-square deviation (*reMSE*), and the regression slope between the results from and ML and Equi (slope).

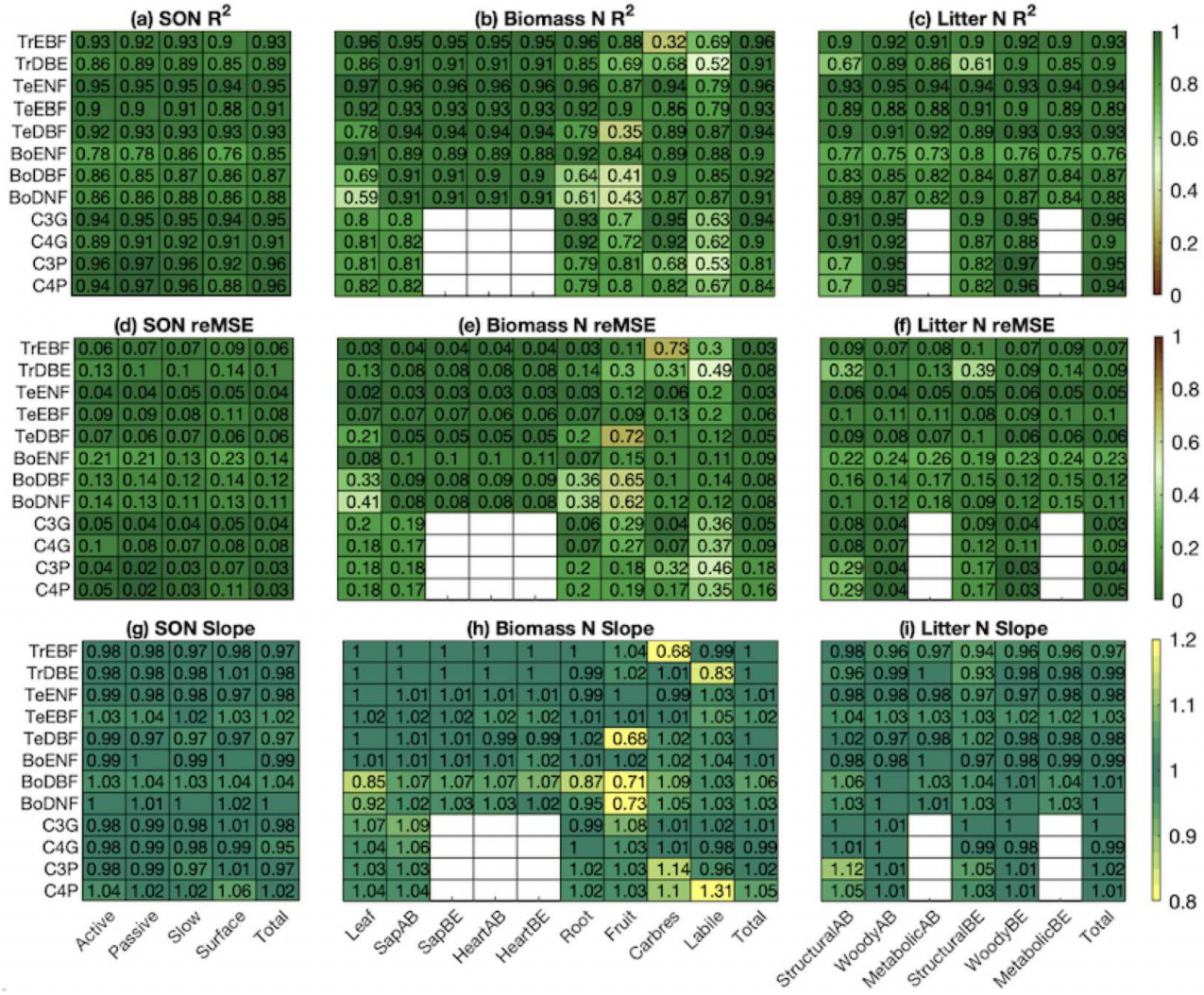


Figure 4.4 The performances of machine learning models on soil, biomass and litter N storages for each PFT compared to the 'true equilibrium' state simulated by ordinary spin-up. The statistics are the same with Figure 4.3.

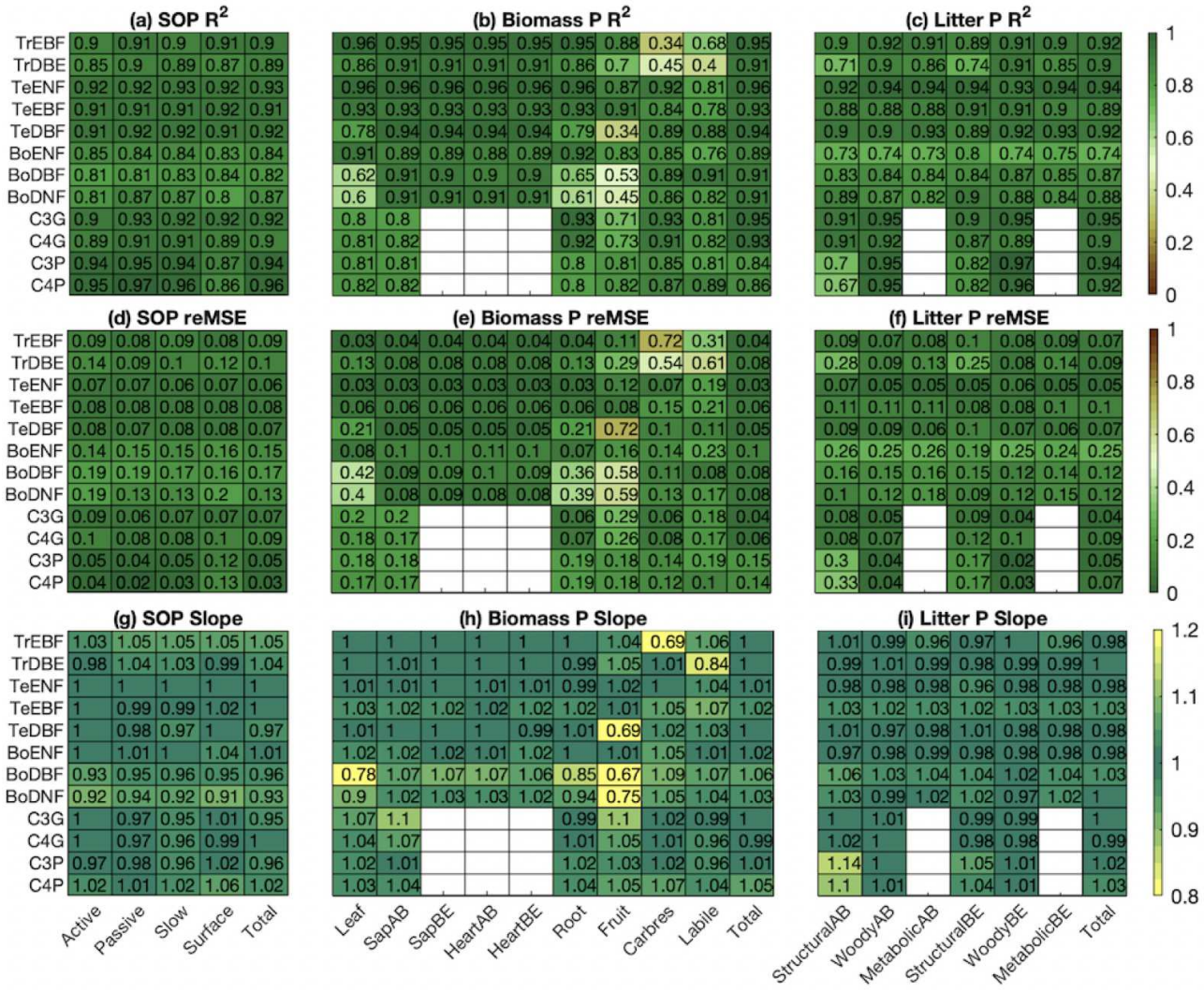


Figure 4.5 The performances of machine learning models on soil, biomass and litter P storages for each PFT compared to the 'true equilibrium' state simulated by ordinary spin-up. The statistics are the same with Figure 4.3.

Across different regions, ML generally produced similar global patterns of each pool to the "*true equilibrium*" (Figure S4.6, S4.7). The high positive bias in total soil C mainly occurs in North and South America, while negative biases are mainly located in high latitude regions of Eurasia for SOC and SON and the whole Eurasia for SOP (Figure S4.6c, f, i). ML performs better for slow-changing biomass pools than for soil pools. For example, the fractions of pixels with a low relative bias of 20% is higher in total biomass C (81%) than in total SOC (69%) (excluding pixels with values lower than their 10th quantiles; Figure S4.6, S4.7). This is due to the faster equilibrium rates for slow-changing biomass pools.

4.3.2 The dependence of ML performance on number of training pixels

The performance of ML models depends on the number of training sites (Figure S4.4). For example, the ML performance on Passive SOC and biomass HeartAB C pools improves with increasing training sites, as supported by the increasing R^2, decreasing *reMSE*, and converging regression slope (≈1)**.** The maximum potential performance of ML, with 100% of global pixels were used for training, can be with R^2 >0.92 and *reMSE* < 0.3% for Passive SOC and biomass HeartAB C pools on PFT level (Table S4.5).

4.3.3 Re-run of land surface model to approach further to the equilibrium state

In order to reduce the errors produced by ML predictions (section 4.3.1), we re-run the spin-up for additional 1000 years to approach further to the equilibrium state (section 4.2.5). The number of years of *re-run* used in practice will be determined via comparing the R^2, *reMSE* and slope between values from "*true equilibrium*" and *re-run* on PFT scale. We found that the bias in fast-changing variables only require 10-100 years of *re-run* to reach the lowest bias with R^2>0.7 and *reMSE*<0.3 (Figure S4.8). But the slow-changing pools need >100 years of *re-run* to reach the state near to the "*true equilibrium*" (Figure S4.9). For vegetation HeartAB C pool, the R^2 firstly increased and then remained constant or decreased slightly with the increasing number of simulation years, while the *reMSE* firstly decreased, then remained constant or slightly increased (Figure S4.9a, c). The regression slope of vegetation HeartAB C pool between *Simulation ML* and *EQ* converges to =1 except for BoDNF (Figure S4.9e). Compared to vegetation HeartAB C pool, fitness indexes for passive SOC evolve much slower with increasing simulation years (Figure S4.9b, d, f), indicating that biases in passive SOC by ML predictions need more than 1000 years to adjust. Even though the R^2 reaches up to >0.85, *reMSE* decreases to <0.2, and the slope converges to near to 1 for all PFTs except for BoENF with re-run 300 simulation years.

We also assess the effect of duration of *re-run* on historical simulation (see section 4.2.5) with respect to NBP and total soil and vegetation C, N and P pools (section 4.3.4). We determined that 300 year of *re-run* is the best to trade-off the high efficiency and low biases, as no significant improvement with longer *re-run* simulation (Figure S4.10; Figure S4.11b-c; Figure S4.12b-c).

4.3.4 Evaluation of historical simulations using MLA spin-up

ML will inevitably introduce errors in the initial pool sizes at steady states, which will affect the model simulations. The question is whether errors in the initial pool sizes will have a significant impact on model applications. In the following, we assess the biases in the results in the historical simulation introduced by the inaccuracy of the MLA based equilibrium state. We focus on the spatiotemporal pattern of NBP, and other C, N and P fluxes and storages.

4.3.4.1 NBP

Net biome productivity (NBP) is defined as the net C exchange between the atmosphere and the terrestrial biosphere. It is the sum of net primary productivity, heterotrophic respiration and emissions due to disturbances. Positive NBP values denote a land carbon sink. Annual total NBP from *Simulation ML* is highly comparable with that from *Simulation EQ* on both global and regional scale during the period of 1959-2016 (Figure 4.6). For period of 1990-2016, the global total annual NBP by *Simulation ML* (0.27 ± 0.88 PgC yr^{-1}) is slightly lower than that from *Simulation EQ* (0.30 ± 0.86 PgC yr^{-1}) (mean value and the inter-annual variability using 1 sigma standard deviation). The accumulated global total NBP by *Simulation ML* (-23.6 PgC yr^{-1}) for period of 1900-2016 is lower than that from *Simulation EQ* (-18.4 PgC yr^{-1}).

Global pixel-to-pixel comparisons in NBP for single years (1990, 2000 and 2010) between *Simulation ML* and *Simulation EQ* show high spatial R^2 (0.89-0.97) and low RMSE (7.45-8.75 gC m^{-2} yr^{-1}) (Figure S4.13). *Simulation ML* produced NBP pattern comparable with the one from *Simulation EQ* (Figure S4.14): ~86% global pixels have an absolute bias lower than 10 gC m^{-2} yr^{-1}. The pattern of NBP biases by *Simulation ML* shows no indication of a concentration of high biases in certain regions (Figure S4.14c, f, i), indicating that there are no systematic errors in NBP.

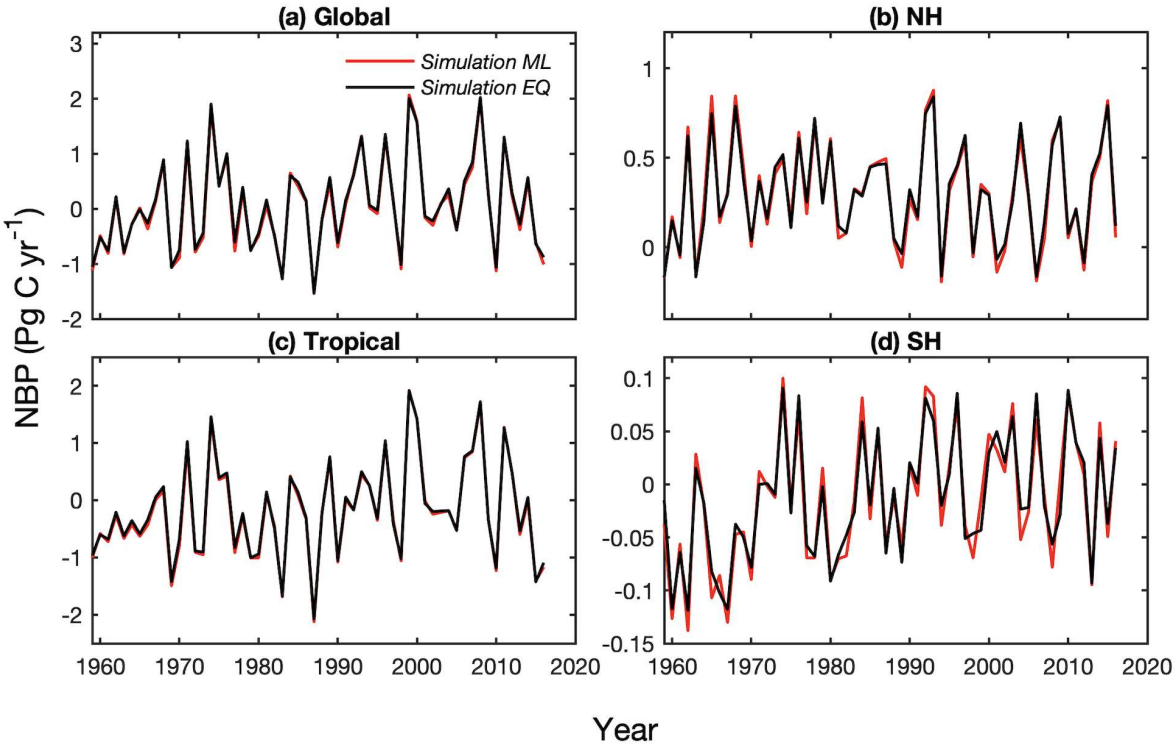


Figure 4.6 Changes in total yearly net biome productivity (NBP) during 1959-2016 from *Simulation ML* and *Simulation EQ* for global (a), North Hemisphere (NH; 30°N - 90°N; b), Tropical (30°S - 30°N; c), and South Hemisphere (SH; 30°S - 90°S; d).

4.3.4.2 GPP, NPP and LAI

The biases in GPP, NPP and LAI in *Simulation ML* are also important checks. *Simulation ML* produced similar global pattern of annual GPP and NPP averaged over the period 2001-2010 compared to *Simulation EQ* (R^2=0.985 and RMSE=75 gC m^{-2} yr^{-1} for GPP; R^2=0.975 and RMSE=40 gC m^{-2} yr^{-1} for NPP; Figure S16). 94 % global pixels have absolute relative bias lower than 15%. The NPP and GPP by *Simulation ML* are systematically overestimated by 12%~87% (25_{th}~75_{th} quantiles) in high latitudes of North America, but are underestimated by 4%~22% (25_{th}~75_{th} quantiles) in some regions of Russia, India and central Africa. Those biases in GPP and NPP are consistent with the biases in LAI. Overall, the biases in C fluxes and LAI by ML seems acceptable.

4.3.4.3 N and P fluxes and storages

Simulation ML produced annual N and P fluxes averaged over the period 2001-2010 which are comparable to the ones from *Simulation EQ*, with R^2 >0.9 and *reMSE* <0.1, for most of PFTs except for C3 and C4 cropland (Figure S17). The regression slope between those fluxes for *Simulation ML* versus *EQ* are all near to 1 except for C3 and C4 croplands (Figure S17). The global pattern of average annual N fluxes and P fluxes over 2001-2010 by *Simulation ML* is highly consistent with that from *Simulation EQ* (Figure S19), while the relative bias in some N and P fluxes such as N2O emission and P leaching are very high due to the low absolute flux values (>30%; Figure S20).

Simulation ML produced N and P storages averaged over the period 2001-2010 which are comparable to the ones from *Simulation EQ* for most of PFTs except for the fast-changing pools (e.g. soil labile N), with R^2 >0.85, *reMSE* <0.1 and regression slope approaching to 1 (Figure S18). Total N and P in soil organic matter and biomass averaged over the period 2001-2010 by *Simulation ML* are systematically overestimated in high latitude of North America, but are generally underestimated in some regions of Eurasia (Figure S11c; Figure S12c). The sign of biases in soil C, N and P storages by *Simulation ML* are generally same with that of ML predictions (Figure S11a; Figure S12a), but their magnitude of absolute biases is smaller than that of ML predictions (Figure S11; Figure S12).

4.3.5 Savings in computational consumption

This ML-based acceleration approach requires as input only 14.1% of total pixels being equilibrated in a fully coupled site simulations, and two 300 year-long global simulations (*pre-spin* and *re-run*). In total, ML saved 80.5 % of computational time compared to equilibrating all pixels in fully coupled spin-up simulation (Figure 4.7). The site simulation takes the largest proportion (68%) of total time consuming for ML-based spin-up.

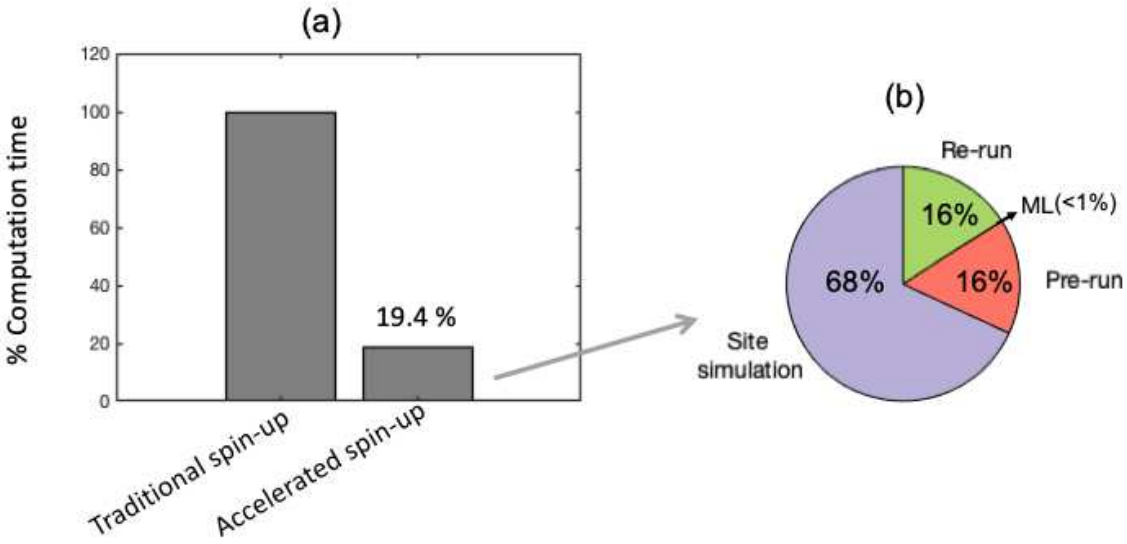


Figure 4.7 Time savings of ML-based spin-up compared to traditional spin-up (a) and the time-consuming proportion of each step for ML-based spin-up: site simulation, pre-run, re-run and ML (b).

4.4 Discussion

4.4.1 Bias in historical NBP due to the application of ML

The bias in global yearly NBP by *Simulation ML* during 1990-2016 (0.03 PgC yr^{-1}) is much smaller than ranges by GCP models (± 0.12 PgC yr^{-1}; 1-sigma standard deviation; Figure S15; Le Quéré et al., 2019), indicating that ML-based spin-up successfully substitute traditional spin-up by giving almost accurate yearly NBP. The C sink in North Hemisphere (NH; 30°N-90°N) is the most questionable issue in current ORCHIDEE-CNP version (r5986) due to the lack of forest management process and/or insufficient model parameterizations (Sun et al., submitted). Due to the very low bias in NH NBP induced from ML-based spin-up (2.8%, 0.009 PgC yr^{-1}), this approach can be safely applied into future model processes refinement and parameter calibration with accuracy guaranteed.

4.4.2 Factors influencing the ML performances

4.4.2.1 Predictor importance

Considering the biases in SOC equilibrium stocks predicted by ML (section 4.3.1; Figure 4.3), it is necessary to understand the factors influencing the ML performances. Ranking the predictor importance is a useful approach to assess the key predictors for predicting SOC equilibrium stocks in ML. For ORCHIDEE-CNP, the size of passive SOC pool is mainly affected by productivity (LAI_g and NPP_g) and climates for most tropical and temperate forests (Figure S4.21; Figure S4.22). Although this ranking result from the ML varies with different training sets, the ranking for top 5 key predictors are generally robust (~15% vs 100% pixels for ML training; Figure S4.21; Figure S4.22), indicating a representative training site selection for tropical and temperate forests. However, for boreal forests and grassland, the ranking of predictor importance is contradictory between different training sets. When using larger size of training set (i.e. 100%), edaphic variables (soil suborder type, soil bulk density and clay content) instead of climates or NPP greatly influence the passive SOC on those PFTs (Figure S4.21; Figure S4.22). The low representativeness of training sites on those PFTs

inevitably induce high biases in predicted SOC.

Particularly, passive SOC of BoENF is the most biased prediction from ML, which was robustly detected to be mostly controlled by soil suborder type (Figure S4.21; Figure S4.22). Those high biases in BoENF's passive SOC can be partly attributed to the low representativeness of rare soil suborder type (such as Andisols, Entisols, Gelisols, and Oxisols) by training sites for ML (Figure S4.23).

4.4.2.2 Partial dependence of SOC to key predictors

Excluding sites with rare soil suborder types, which show the largest biases, there is still -17%~31% relative biases for BoENF sites' passive SOC (25_{th} and 75_{th} quantiles on soil suborder scale; Figure S4.23). This is because the representativeness of training sites on other predictors (such as climate variables) are still insufficient. With increasing number of training sites from 13.4% to 100% of total pixels for BoENF, the predicted passive SOC by ML in Tmax-Tgs space decreases by 3%~18% in regions with Tgs >80 ^{o}C and Tmax >17 ^{o}C (Figure S4.24c). This bias in BoENF passive SOC due to the insufficient training sites for Tgs and Tmax are generally consistent with the sign of prediction bias in passive SOC compared to "true equilibrium" value. This indicates increasing the number of training sites in high T region can actually improve the ML performance in passive SOC.

4.4.3 Trade-off between ML accuracy and time saving

Increasing the training sites for particular PFT (as demonstrated for BoENF) can improve the ML accuracy (see 4.4.2), but also increases the computation cost: In ORCHIDEE-CNP, doubling the number of training sites (from 13.4 % to 26.4 % of total pixels) leads to an increase in R^2 of passive SOC from 0.78 to 0.86 and reduced *reMSE* from 0.22 to 0.13. It also leads to a reduced the computational time saving from 80.6% to 73.4%. Depending on the needed the accuracy, the time savings will vary. We expect that LSM runs at higher spatial resolution than examined here, a small increase in pixels' fraction for training (i.e. small increase in computation cost) will involve large amounts of pixels into the training pool, and thus may lead to larger increase in accuracy.

4.4.4 Implications and Conclusions

This study is the first to involve machine learning methods to substitute the traditional fully coupled spin-up. In this study, the application of ML in ORCHIDEE-CNP saved 80.5% of computational time for global spin-up. Different from previous acceleration approaches with precondition of linear process and fixed C-N ratios, this approach breaks those limits and can be widely used in LSMs, no matter if the model processes are linear or nonlinear. The reduction in the computational consumption to about one order of magnitude less opens the opportunity of data assimilation using the ever-growing observation datasets as well as increasing the models spatial resolution.

References

Breiman, L. (1996). Bagging predictors. Machine learning, 24(2), 123-140.

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., ... & Smith, B. (2007). Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology, 13(3), 679-706. https://doi.org/10.1111/j.1365-2486.2006.01305.x

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. Neural Networks, 106, 249-259.

Cawley, G. C. (2006, July). Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. In The 2006 ieee international joint conference on neural network proceedings (pp. 1661-1668). IEEE.

Cleveland, C. C., Houlton, B. Z., Smith, W. K., Marklein, A. R., Reed, S. C., Parton, W., ... & Running, S. W. (2013). Patterns of new versus recycled primary production in the terrestrial biosphere. *Proceedings of the National Academy of Sciences*, *110*(31), 12733-12737. http://doi.org/10.1073/pnas.1302768110, 2013

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.

Desgraupes, B. (2017). Clustering indices. University of Paris Ouest-Lab Modal'X, 1, 34.

Elser, J. J., Bracken, M. E., Cleland, E. E., Gruner, D. S., Harpole, W. S., Hillebrand, H., ... & Smith, J. E. (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology letters*, *10*(12), 1135-1142. https://doi.org/10.1111/j.1461-0248.2007.01113.x

Farquad, M. A. H., & Bose, I. (2012). Preprocessing unbalanced data using support vector machine. Decision Support Systems, 53(1), 226-233.

Fleischer, K., Rammig, A., De Kauwe, M. G., Walker, A. P., Domingues, T. F., Fuchslueger, L., ... & Haverd, V. (2019). Amazon forest response to CO 2 fertilization dependent on plant phosphorus acquisition. Nature Geoscience, 12(9), 736-741. https://doi.org/10.1038/s41561-019-0404-9

Gärdenäs, A. I., Ågren, G. I., Bird, J. A., Clarholm, M., Hallin, S., Ineson, P., ... & Ogle, S. (2011). Knowledge gaps in soil carbon and nitrogen interactions–from molecular to global scale. *Soil Biology and Biochemistry*, *43*(4), 702-717. https://doi.org/10.1016/j.soilbio.2010.04.006

Goll, D., Vuichard, N., Maignan, F., Jornet-Puig, A., Sardans, J., Violette, A., ... & Guenet, B. (2017). A representation of the phosphorus cycle for ORCHIDEE (revision 4520). Geoscientific Model Development Discussions, 10(10), 3745-3770. http://doi.org/10.5194/gmd-10-3745-2017, 2017.

Goll, D. S., Joetzjer, E., Huang, M., and Ciais, P.: Low phosphorus availability decreases susceptibility of tropical primary productivity to droughts, Geophys. Res. Lett., 45, 8231-8240, http://doi.org/10.1029/2018GL077736, 2018.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset, Int. J. Climatol., 34, 623-642, http://doi.org/10.1002/joc.3711, 2014.

Hou, E., Luo, Y., Kuang, Y., Chen, C., Lu, X., Jiang, L., ... & Wen, D. (2020). Global meta-analysis shows pervasive phosphorus limitation of aboveground plant production in natural terrestrial ecosystems. *Nature communications*, *11*(1), 1-9. https://doi.org/10.1038/s41467-020-14492-w

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene – HYDE 3.2, Earth Syst. Sci. Data, 9, 927-953, http://doi.org/10.5194/essd-9-927-2017, 2017a.

Klein Goldewijk, K., Dekker, S. C., and van Zanden, J. L.: Per-capita estimations of long-term historical land use and the consequences for global change research, J. Land Use Sci., 12, 313-337,

http://doi.org/10.1080/1747423X.2017.1354938, 2017b.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., ... & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. Global Biogeochemical Cycles, 19(1). https://doi.org/10.1029/2003GB002199

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., and Endo, H.: The JRA-55 reanalysis: General specifications and basic characteristics, J. Meteorol. Soc. Jpn., 93, 5-48, http://doi.org/10.2151/jmsj.2015-001, 2015.

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., ... & Swenson, S. C. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. Biogeosciences, 10(11), 7109.

Melillo, J. M., Butler, S., Johnson, J., Mohan, J., Steudler, P., Lux, H., ... & Vario, C. (2011). Soil warming, carbon–nitrogen interactions, and forest carbon budgets. Proceedings of the National Academy of Sciences, 108(23), 9508-9512. https://doi.org/10.1073/pnas.1018189108

Norby, R. J., De Kauwe, M. G., Domingues, T. F., Duursma, R. A., Ellsworth, D. S., Goll, D. S., ... & Pavlick, R. (2016). Model–data synthesis for the next generation of forest free-air CO 2 enrichment (FACE) experiments. *New Phytologist*, *209*(1), 17-28. https://doi.org/10.1111/nph.13593

Qu, Y., Maksyutov, S., & Zhuang, Q. (2018). An efficient method for accelerating the spin-up process for process-based biogeochemistry models. *Biogeosciences, 15(13)*.

Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowald, N. M., Lindsay, K., LEE, Y. H., ... & Covey, C. (2009). Systematic assessment of terrestrial biogeochemistry in coupled climate–carbon models. Global Change Biology, 15(10), 2462-2484. https://doi.org/10.1111/j.1365-2486.2009.01912.x

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195-204. https://doi.org/10.1038/s41586-019-0912-1

Rendón, E., Alejo, R., Castorena, C., Isidro-Ortega, F. J., & Granda-Gutiérrez, E. E. (2020). Data Sampling Methods to Deal With the Big Data Multi-Class Imbalance Problem. Applied Sciences, 10(4), 1276.

Rovai, A. S., Twilley, R. R., Castañeda-Moya, E., Riul, P., Cifuentes-Jara, M., Manrow-Villalobos, M., ... & Pagliosa, P. R. (2018). Global controls on carbon storage in mangrove soils. *Nature Climate Change, 8(6)*, 534-538.

Sun, Y., Peng, S., Goll, D. S., Ciais, P., Guenet, B., Guimberteau, M., ... & Poulter, B. (2017). Diagnosing phosphorus limitations in natural terrestrial ecosystems in carbon cycle models. Earth's future, 5(7), 730-749. https://doi.org/10.1002/2016EF000472

Thornton, P. E., & Rosenbloom, N. A. (2005). Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. Ecological Modelling, 189(1-2), 25-48. https://doi.org/10.1016/j.ecolmodel.2005.04.008

Wang, Y. P., Law, R. M., & Pak, B. (2010). A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences, 7(7).

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

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., ... & Wollschläger, U. (2019). Soil organic carbon storage as a key function of soils-a review of drivers and indicators at various scales. *Geoderma*, 333, 149-162.

Xia, J. Y., Luo, Y. Q., Wang, Y. P., Weng, E. S., & Hararuk, O. (2012). A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. Geoscientific Model

Development, 5(5), 1259-1271. https://doi.org/10.5194/gmd-5-1259-2012.

Yang, X., Thornton, P. E., Ricciuto, D. M., & Post, W. M. (2014). The role of phosphorus dynamics in tropical forests--a modeling study using CLM-CNP. Biogeosciences, 11(6).

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., ... & Jain, A. (2014). Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two temperate F ree-A ir CO 2 E nrichment studies. New Phytologist, 202(3), 803-822. https://doi.org/10.1111/nph.12697

Zobler, L.: A world soil file global climate modeling, 32, 1986.

Supplementary

Text S1 K-means cluster method and Synthetic Minority Oversampling Technique

To achieve a balanced *Predictor X*, we applied K-means cluster to classify the global multi-variables X space into K clusters (Figure 4.1) according to 'europ' distance. Highly skewed predictors (i.e. N and P deposition) were log-transformed before entering into the cluster procedure. Soil suborder type, as the categorical predictor, was excluded from *Predictor X* for the K-mean cluster. Different optimal values of K are determined if using different criterions (Table S4.2; Desgraupes, 2017). Considering that very high K (K>5) cannot help improve the partition but create some clusters containing very few samples, we applied K=4 for each PFT in this study (Figure S4.2). The same number of sites were selected from each cluster (N_c) as the training set for each PFT (Figure S4.3). A set of suitable N_c for each PFT used for post- ML model building and analysis are shown in Table S4.3. In total, 14.1% of total pixels were selected for the fully coupled site simulation. The choice of number of N_c influences the ML performance and computation cost, which were shown in section 4.3.3 (Figure S4.4).

Irrespectively of a balanced *Predictor X*, the *Response Y* can be unbalanced. E.g. insufficient sampling of training sites in the high Y ranges will lead to underestimation of Y in high Y regions. Vice versa for low Y ranges.

Synthetic Minority Oversampling TEchnique (SMOTE) is an over-sampling approach which is mainly used to create new synthetic examples in a minority class of the training data set (Chawla et al., 2002). Those 'synthetic' examples for minor class are generated as follows: obtain the K-nearest neighbors of sample M in feature space, select *n* samples randomly and record them as M_i. Finally, the new sample M_{new} is defined by interpolation as follows:

$$M_{new} = M_{origin} + rand \times \left(M_i - M_{origin} \right),\ i = 1, 2, \ldots, n \quad \text{Eq. (S1)}$$

where *rand* is a random number uniformly distributed within the range (0,1).

In this study, we applied SMOTE to generate 4 duplicated samples for each site with response Y falling out > 90^{th} quantiles of its distribution.

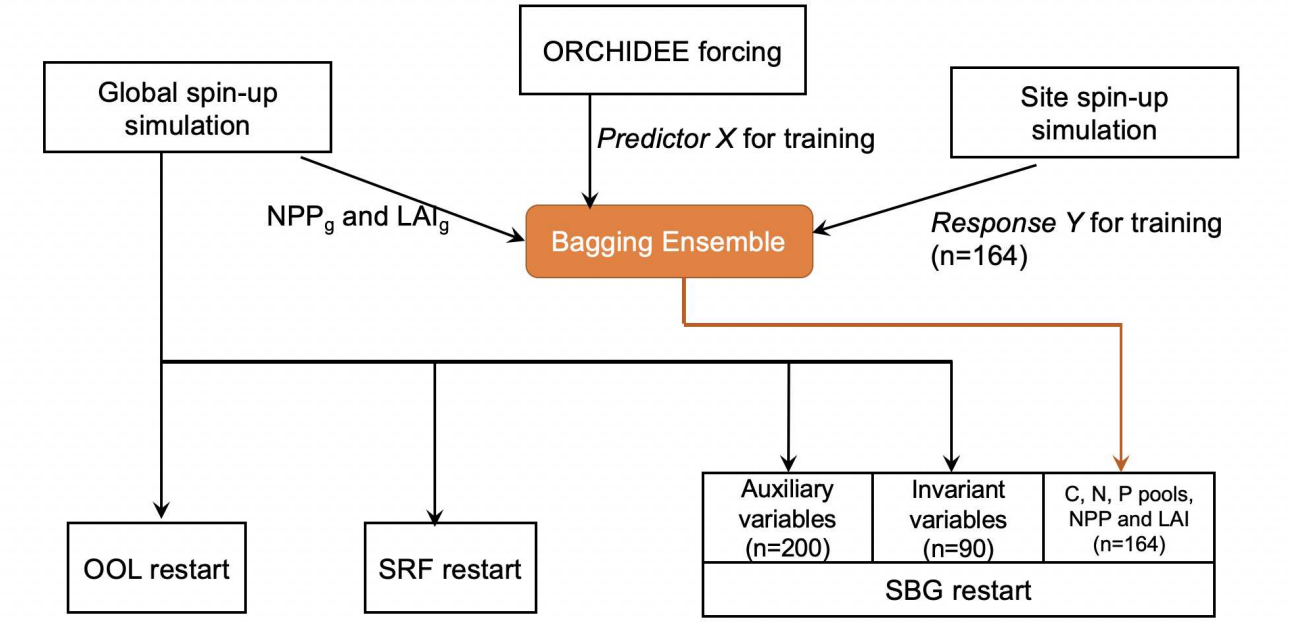


Figure S4.1 Flow chart for dealing with state variables, auxiliary variables and invariant forcing variables in the restart file of ORCHIDEE-CNP.

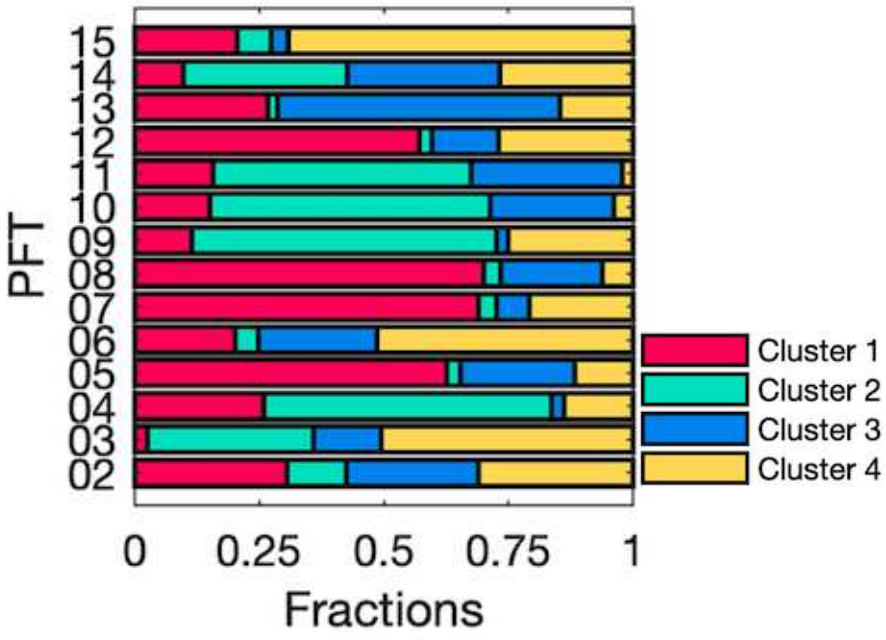


Figure S4.2 Fractions of sites in each cluster using K-mean cluster (K=4) on Predictor X space approach for each PFT.

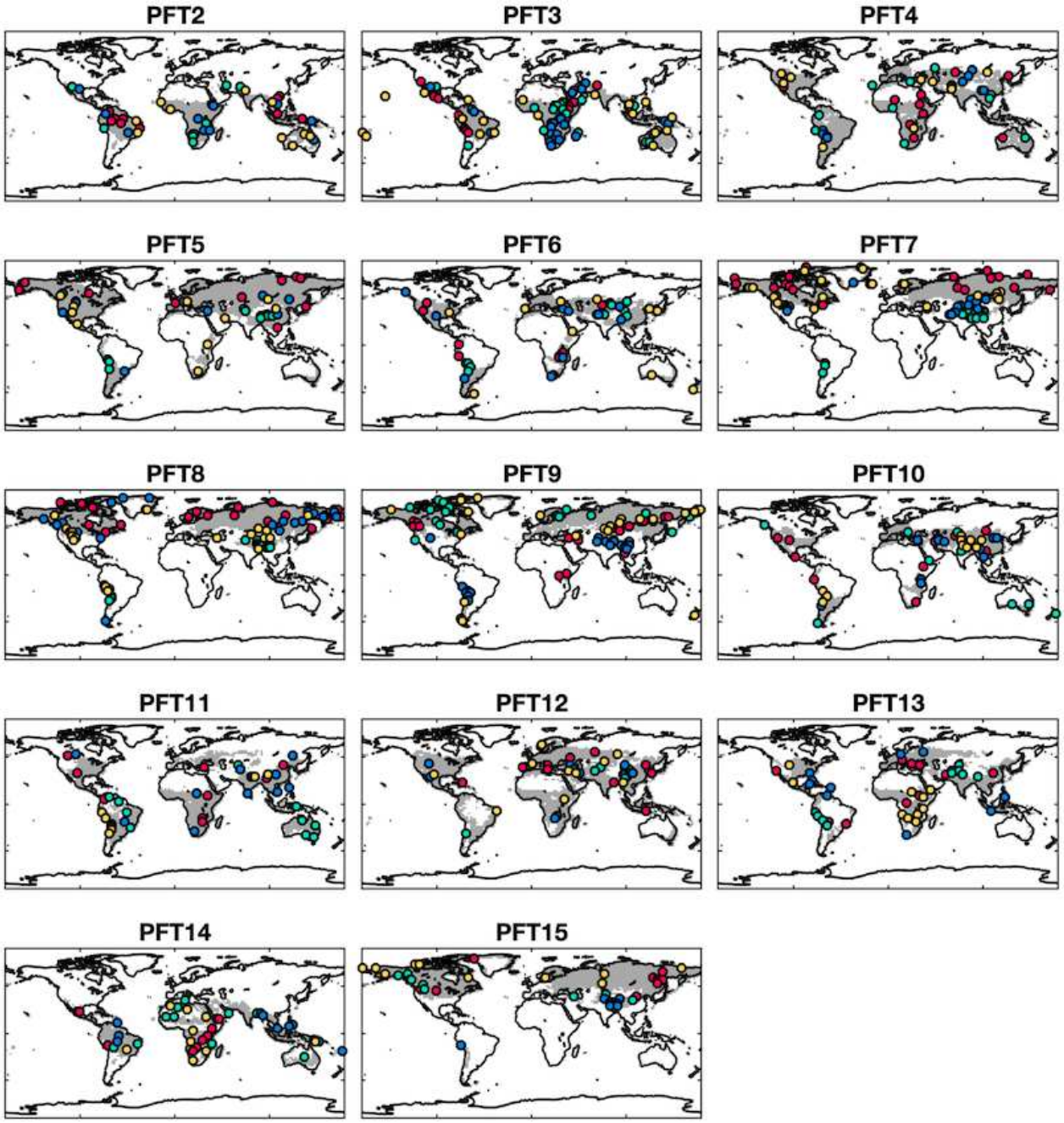


Figure S4.3 Distribution of selected training sites for each PFT. The colors of scatters are corresponding to the 4 different clusters in Figure S4.2.

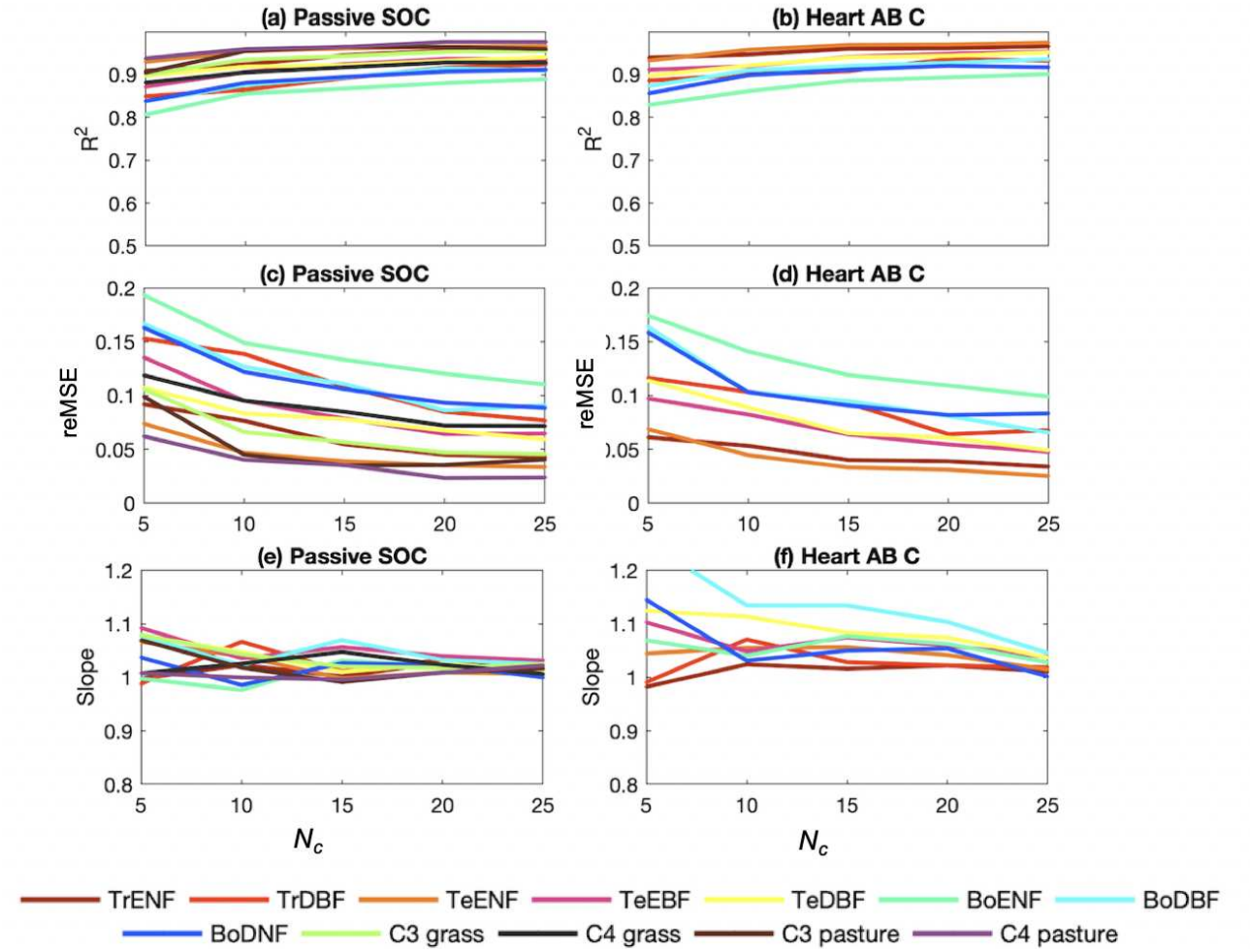


Figure S4.4 Performances of machine learning models (ML) on passive soil organic carbon (Passive SOC) and aboveground heartwood C (HeartAB C) for each PFT using different numbers for training sites for each cluster (Nc). Three statistics represent the model performance: coefficient of determination (R^2), relative mean-square deviation (reMSE), and the regression slope between the results from and ML and Equi (slope).

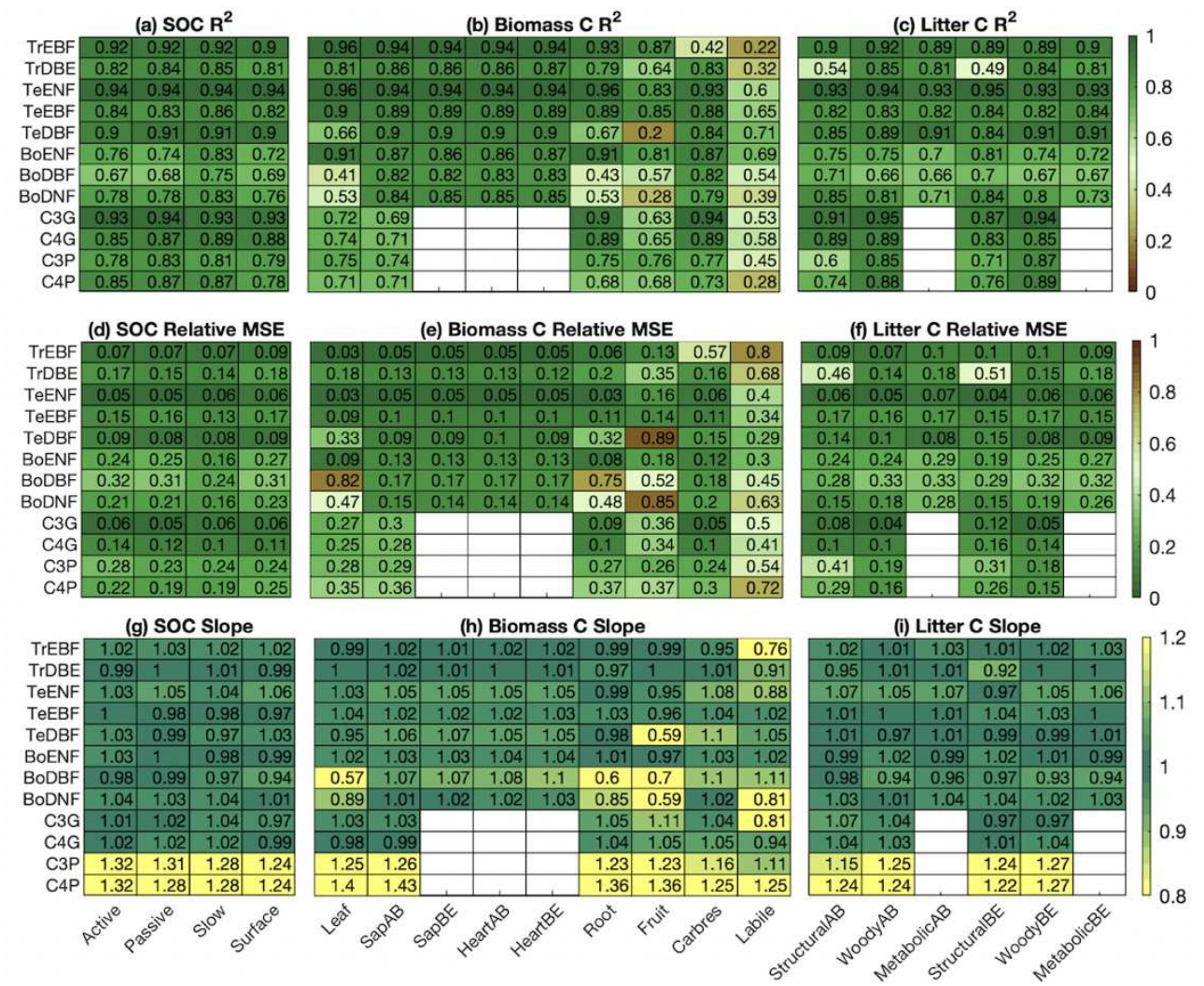


Figure S4.5 Performances of machine learning models (ML) on soil, biomass and litter C storages for each PFT excluding NPPg and LAIg from predictor X for ORCHIDEE-CNP. The statistics are the same with Figure 4.3-4.5.

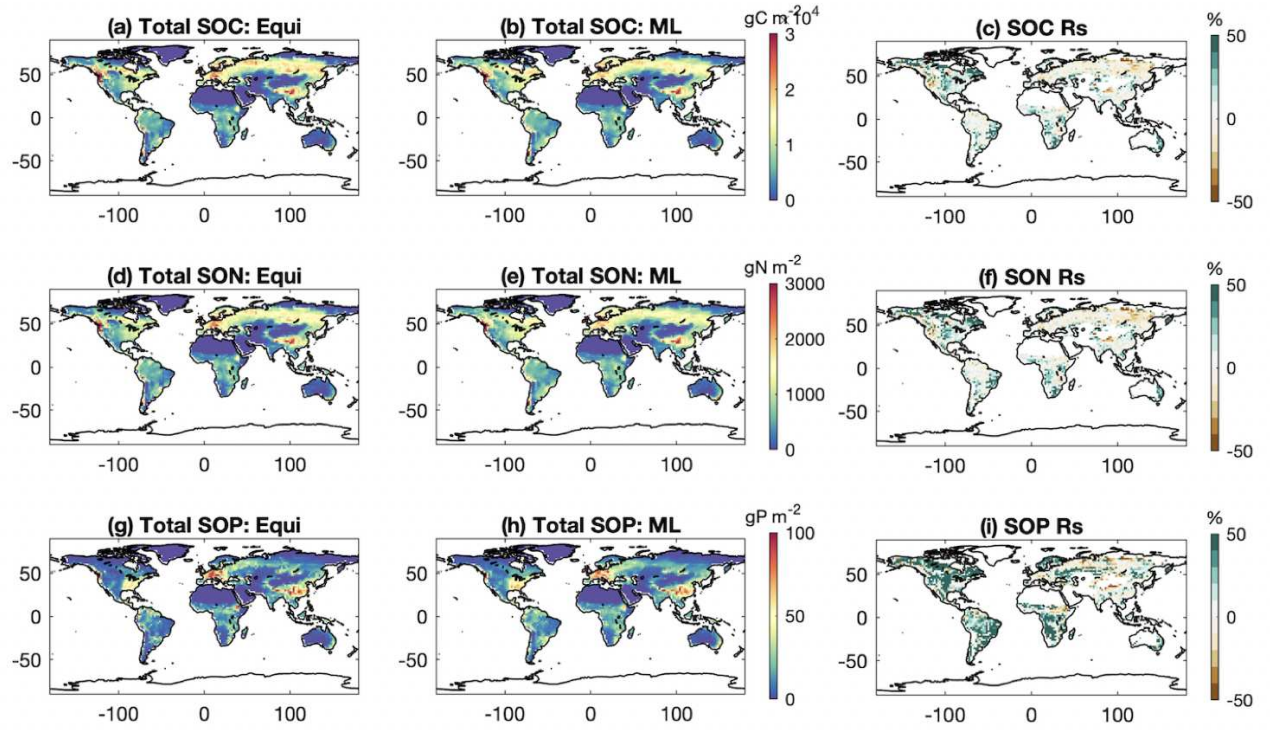


Figure S4.6 Global pattern of total soil C, N and P storages from machine learning models' predictions (ML) and "true equilibrium" (Equi). Relative differences (Rs) for each pool are also shown.

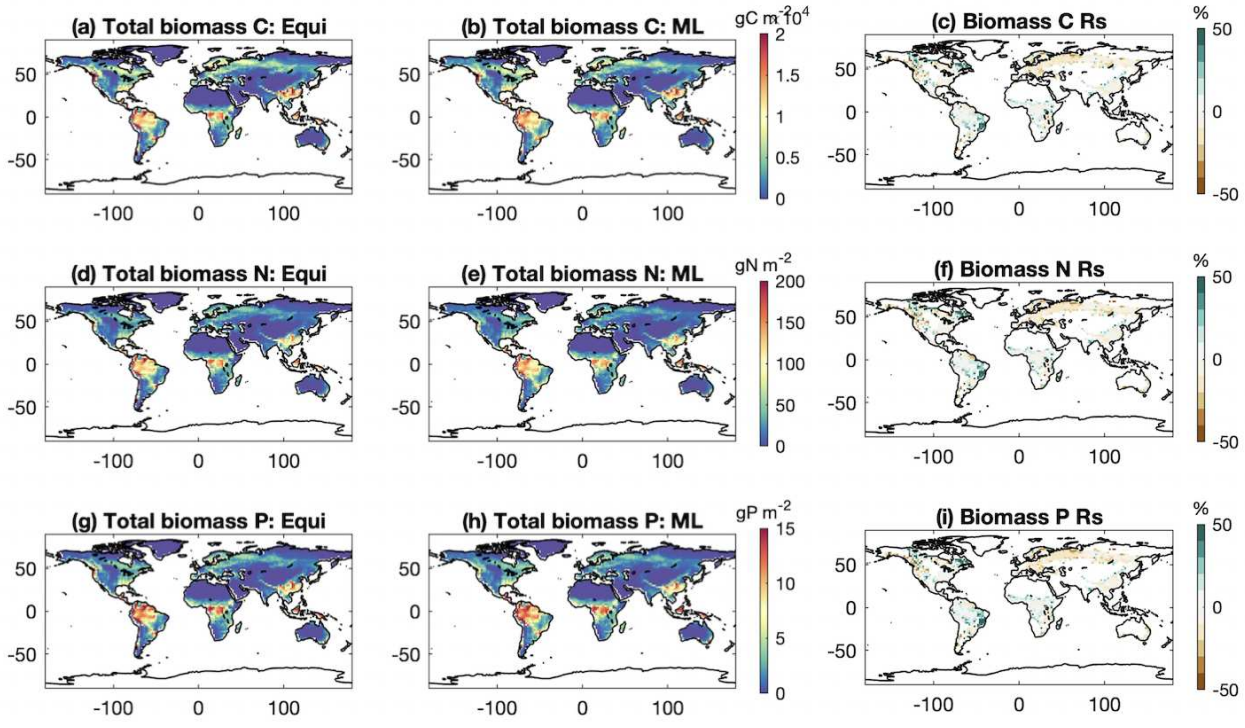


Figure S4.7 Global pattern of total biomass C, N and P storages from machine learning models' predictions (ML) and "true equilibrium" (Equi). Relative differences (Rs) for each pool are also shown.

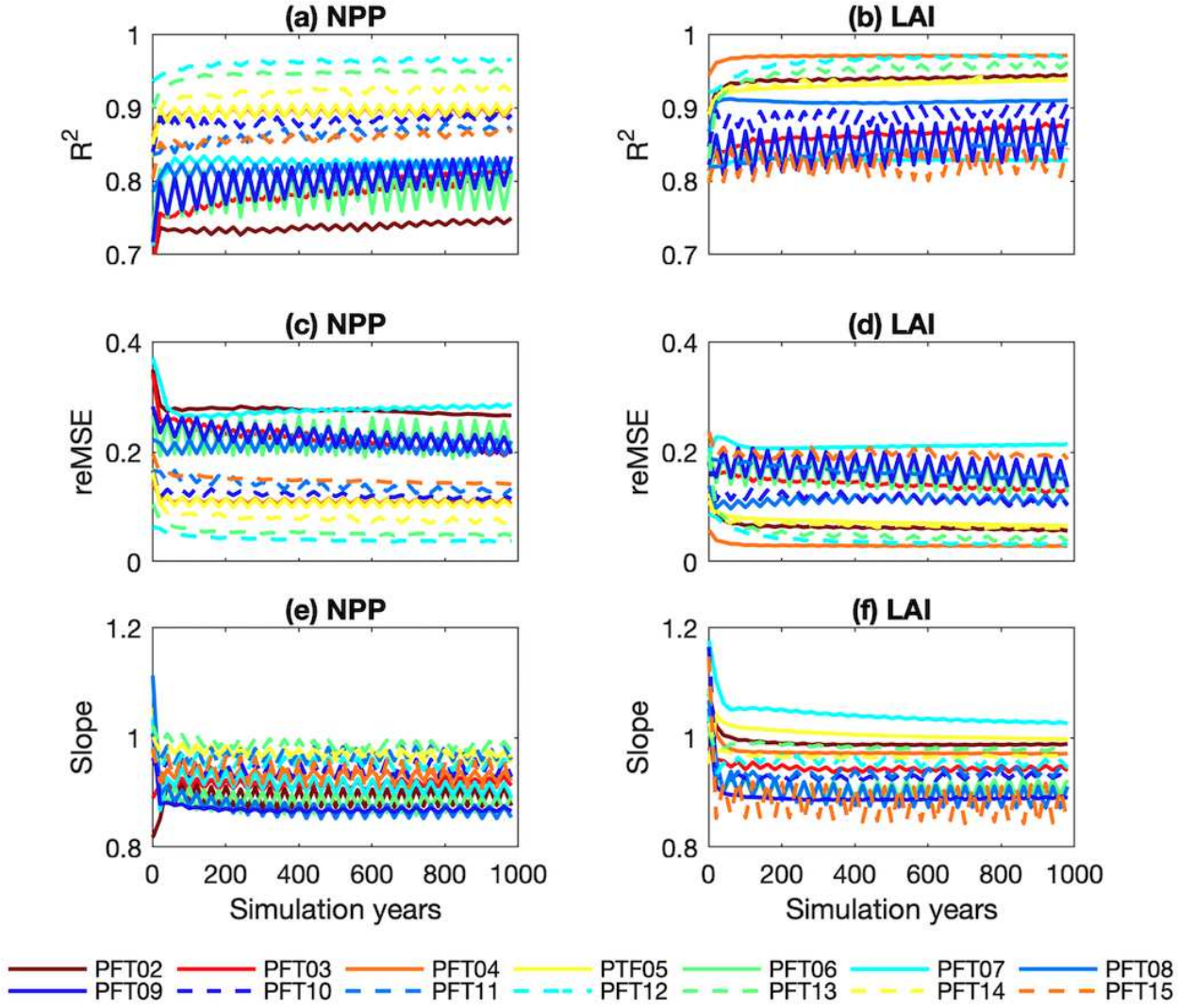


Figure S4.8 Evolution of the fitness of NPP and LAI from ML-based re-run to the "true equilibrium" state with increasing simulation years of re-run. Three statistics, coefficient of determination (R2), relative mean-square deviation (reMSE), and the regression slope (slope), were used for the evaluation.

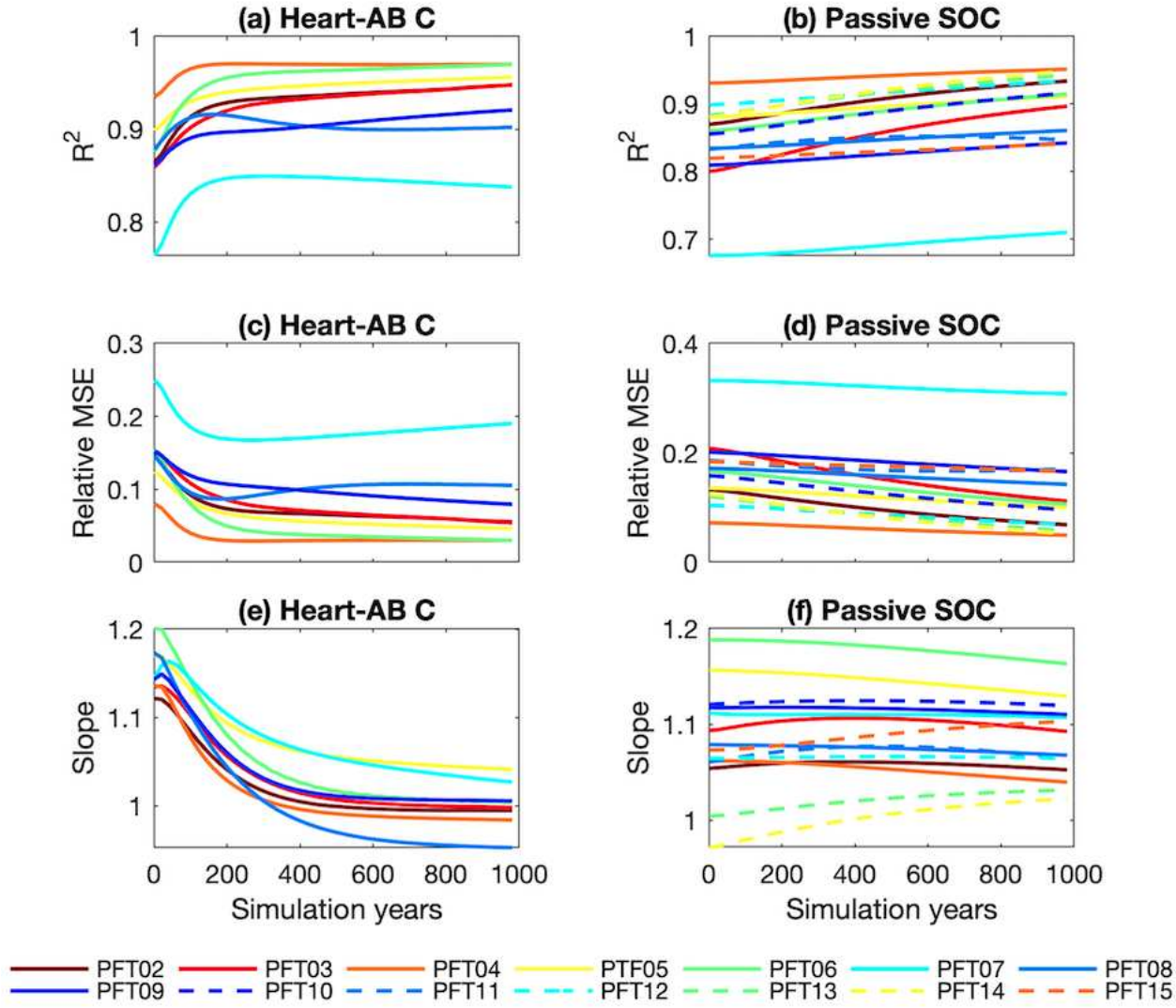


Figure S4.9 Evolution of the fitness of pool sizes from ML-based re-run to the "true equilibrium" state with increasing simulation years of re-run. Two pools, passive soil organic carbon (Passive SOC) and aboveground heartwood C (HeartAB C), were evaluated using three statistics that are the same with Figure S8.

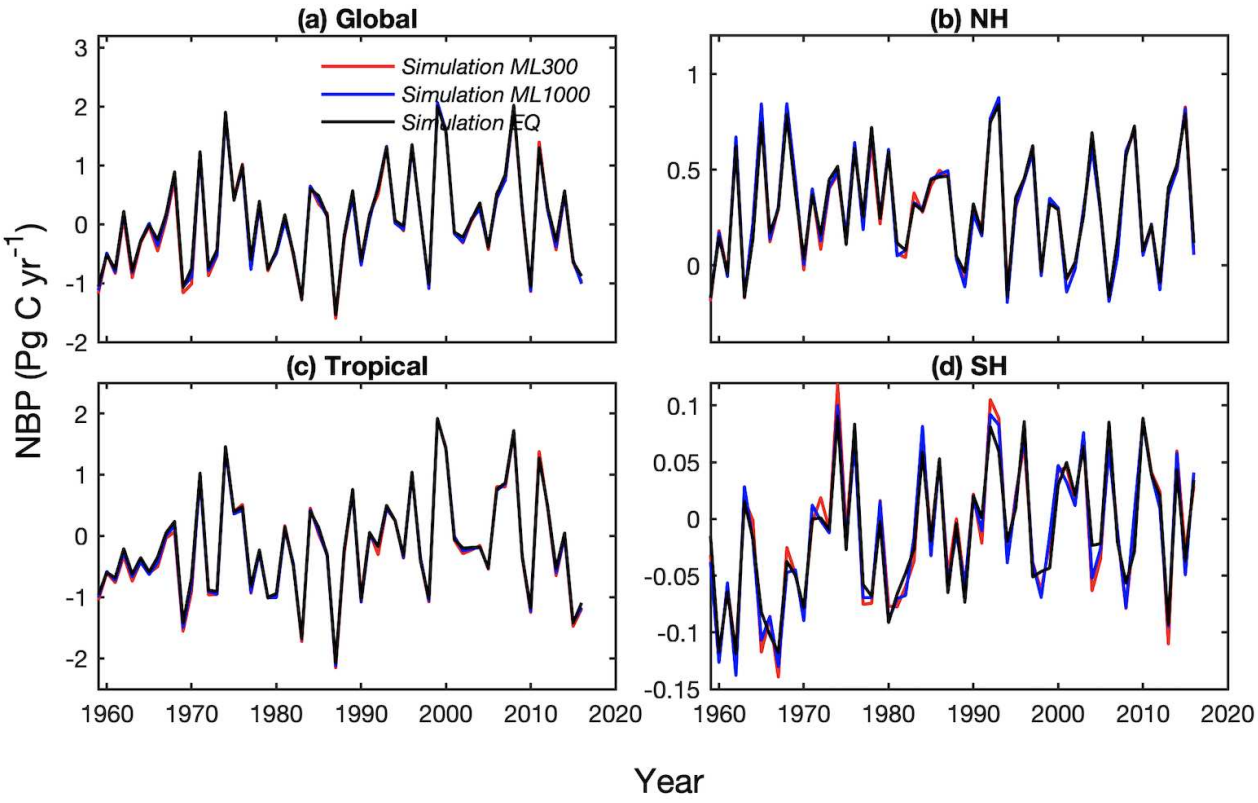


Figure S4.10 Changes in total yearly net biome productivity (NBP) during 1959-2016 from Simulation EQ and two Simulation ML with different years of re-run (300 vs. 1000) for global (a), North Hemisphere (NH; 30oN - 90oN; b), Tropical (30oS - 30oN; c), and South Hemisphere (SH; 30oS - 90oS; d).

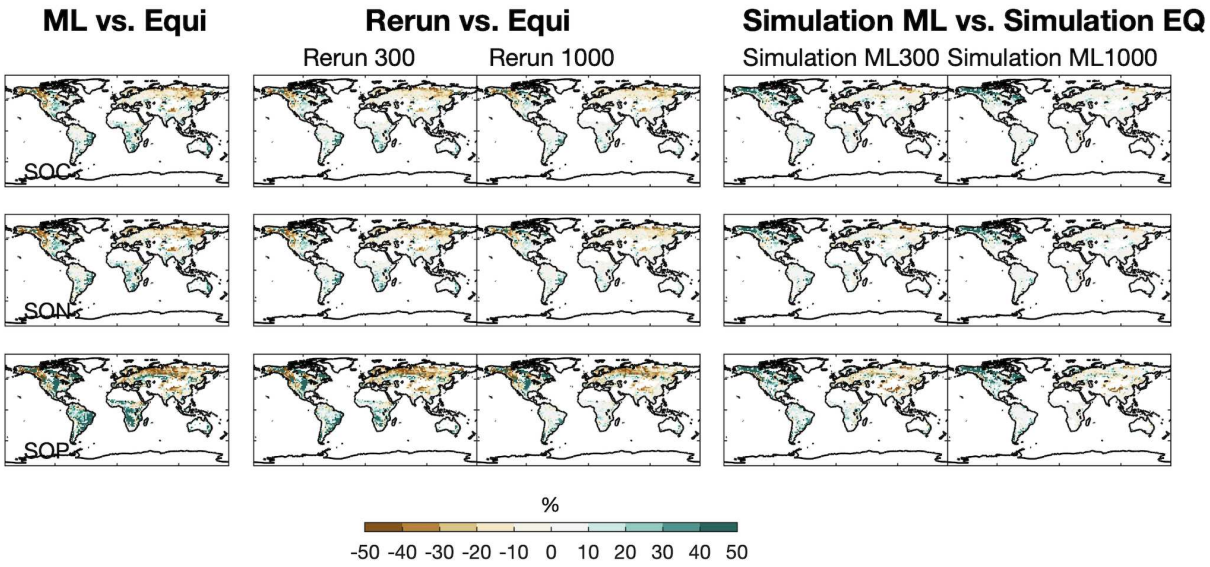


Figure S4.11 Global pattern of the relative bias in total soil organic C, N and P storages (SOC, SON, SOP) in each step of ML-based spin-up compared to the traditional spin-up. (a) shows the relative bias in SOC, SON, SOP due to ML compared to "true equilibrium" state (Equi); (b) shows the relative bias after re-run the spin-up using ML-based restart file for 300 and 1000 years (i.e. Rerun 300 and Rerun 1000); and (c) shows the relative bias in mean total SOC, SON and SOP across 2000-2010 from two Simulation ML with different years of re-run (300 vs. 1000) compare to Simulation EQ. Pixels with the low total SOC (<1kgC m-2 yr-1) for Simulation EQ were masked.

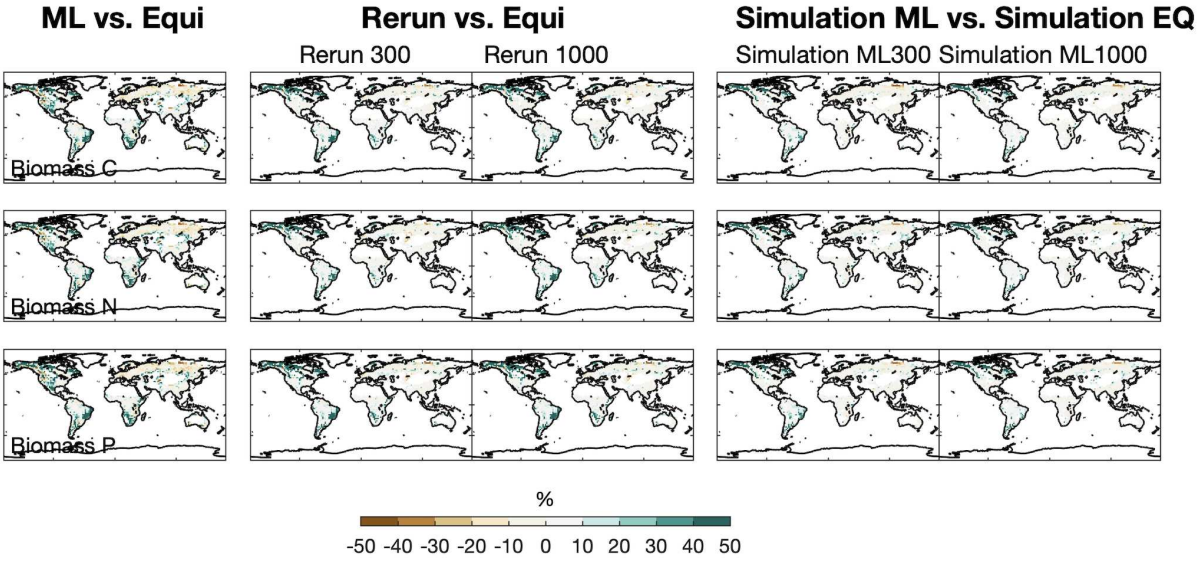


Figure S4.12 Global pattern of the relative bias in total biomass organic C, N and P storages in each step of ML-based spin-up compared to the traditional spin-up. (a) shows the relative bias in biomass C, N and P due to ML compared to "true equilibrium" state (Equi); (b) shows the relative bias after re-run the spin-up using ML-based restart file for 300 and 1000 years (i.e. Rerun 300 and Rerun 1000); and (c) shows the relative bias in mean total biomass C, N and P across 2000-2010 from two Simulation ML with different years of re-run (300 vs. 1000) compare to Simulation EQ. Pixels with the low total biomass C (<500 gC m-2 yr-1) for Simulation EQ were masked.

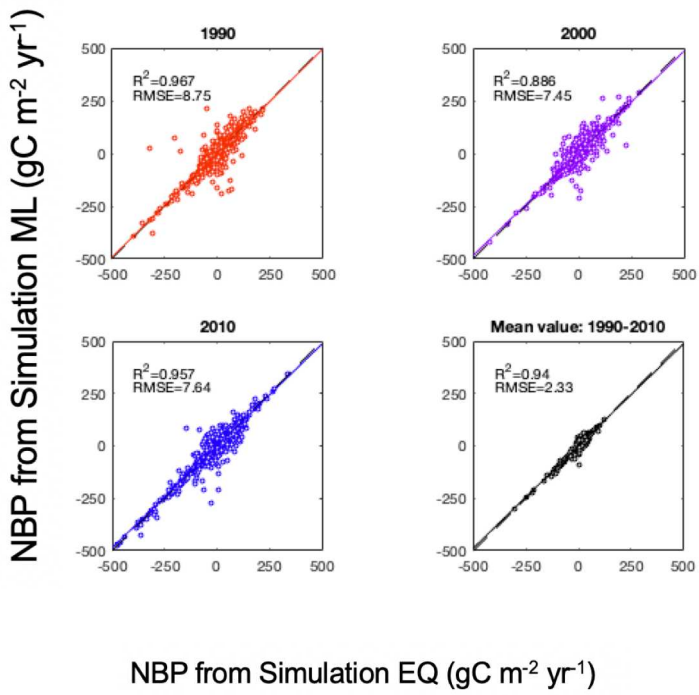


Figure S4.13 Global pixel-to-pixel comparisons in yearly net biome productivity (NBP) for 1990, 2000, 2010 and across 1990-2010 between Simulation ML and Simulation EQ.

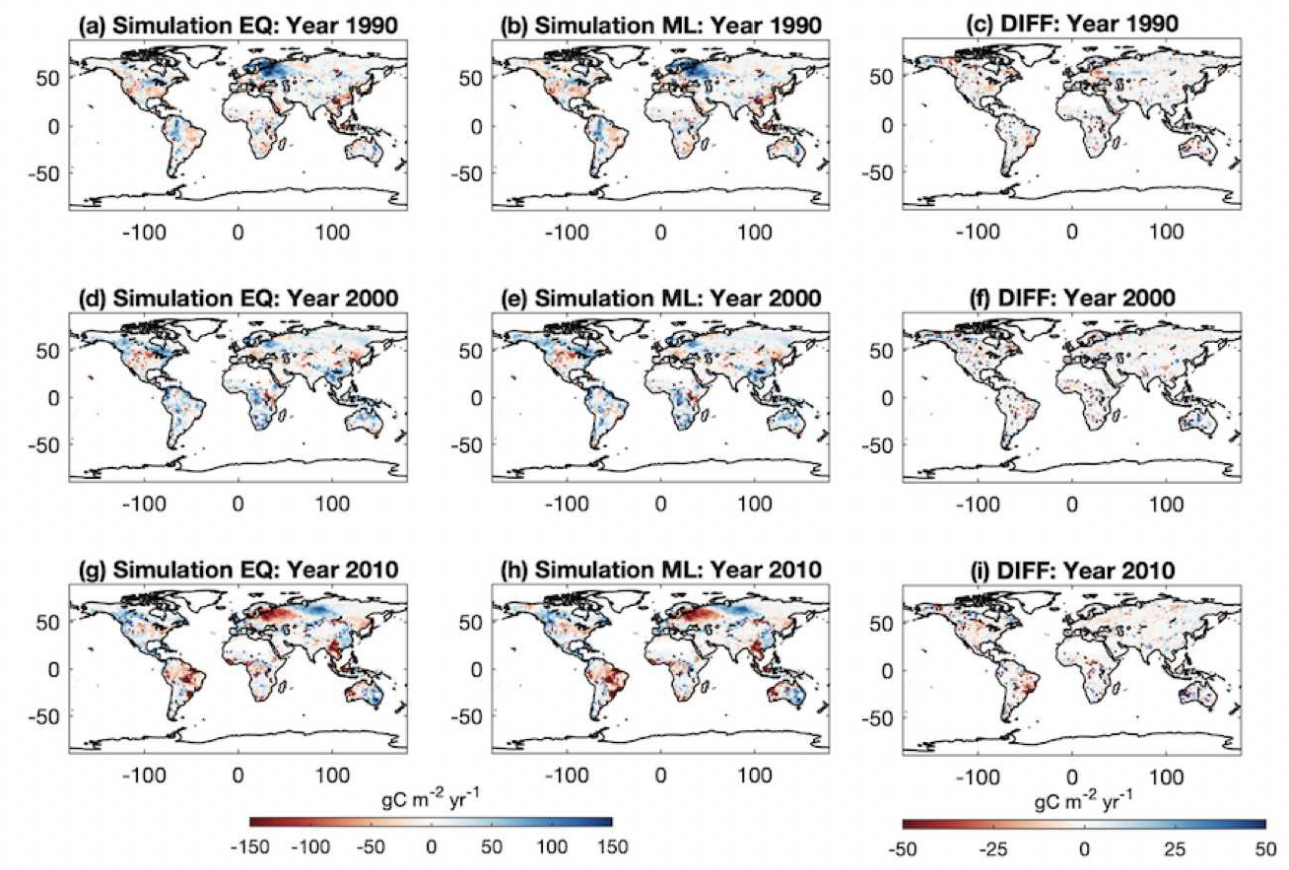


Figure S4.14 Global pattern of yearly net biome productivity (NBP) for 1990, 2000 and 2010 from Simulation ML and Simulation EQ. Bias in NBP (DIFF; Simulation ML minus Simulation EQ) are also shown (c, f, i).

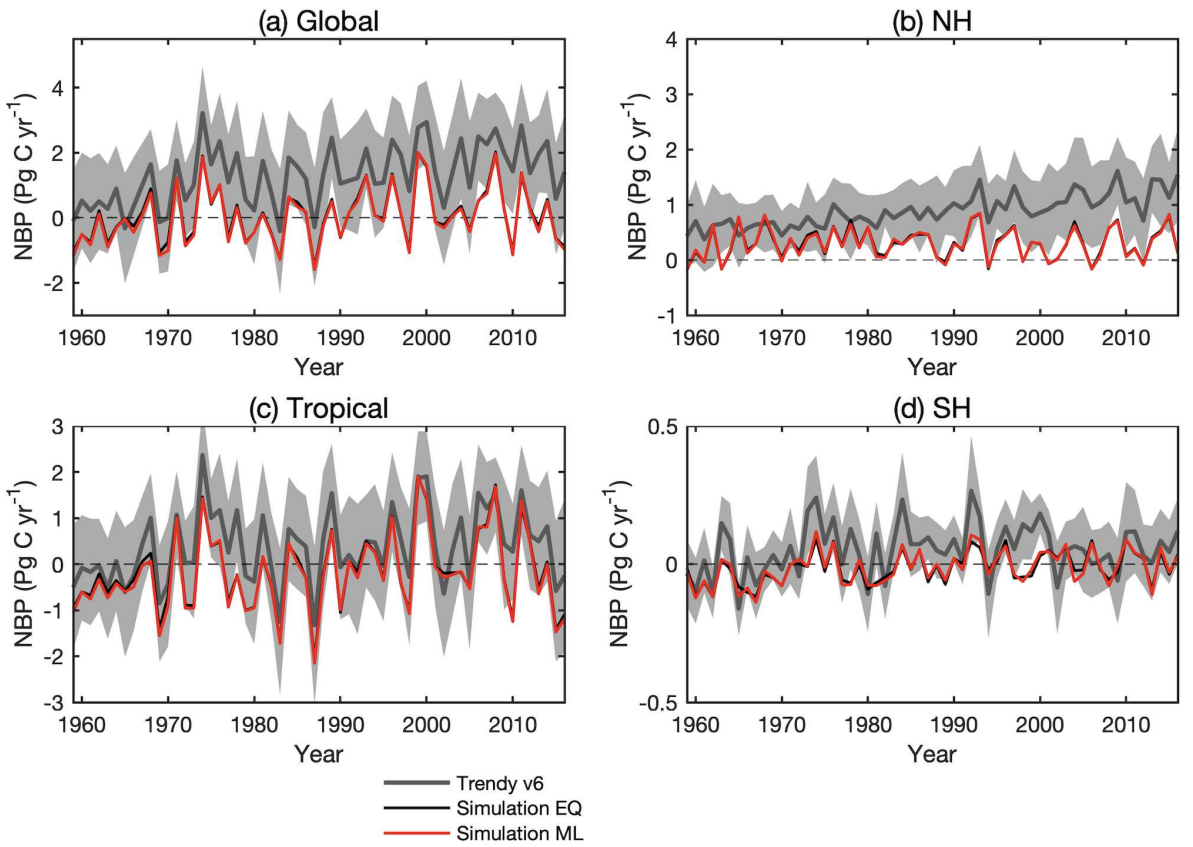


Figure S4.15 Changes in total yearly net biome productivity (NBP) during 1959-2016 from Simulation EQ and two Simulation ML and Trendy v6 model for global (a), North Hemisphere (NH; 30oN - 90oN; b), Tropical (30oS - 30oN; c), and South Hemisphere (SH; 30oS - 90oS; d). Grey shading indicates ±1σ uncertainty of DGVM results for Trendy v6.

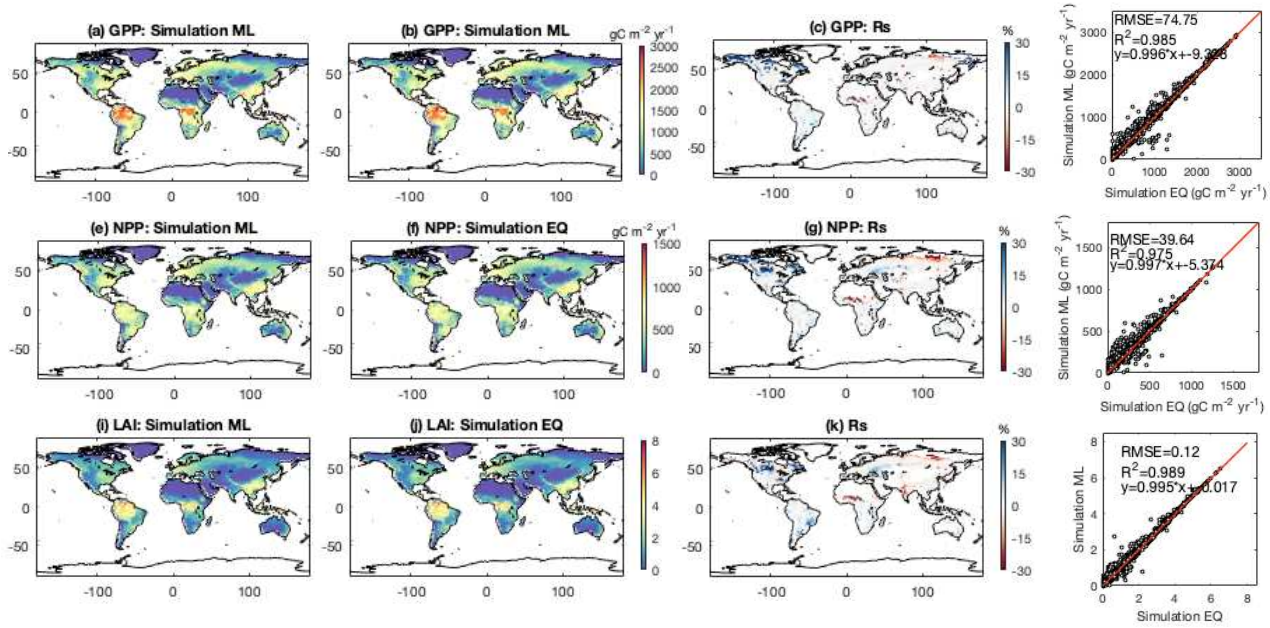


Figure S4.16 Global pattern of historical mean annual GPP, NPP, LAI across 2000-2010 from Simulation ML and Simulation EQ. Relative differences (Rs; c, g, k) and global pixel-to-pixel comparisons (d, h, l) for each item are also shown. Pixels with the low LAI (LAI<0.1) for Simulation EQ were masked.

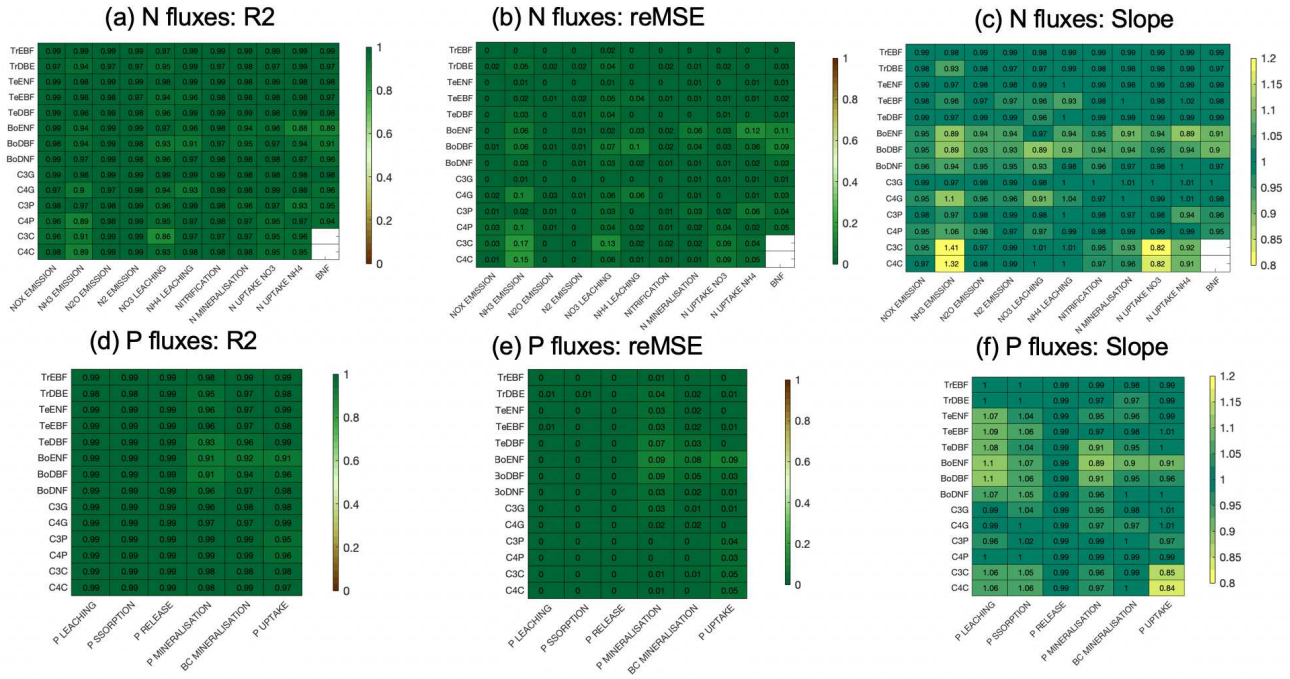


Figure S4.17 Performances of machine learning models (ML)-based historical simulation (i.e. Simulation ML) on mean yearly N and P fluxes across 2000-2010, compared to Simulation EQ. The statistics are the same with Figure 3-5.

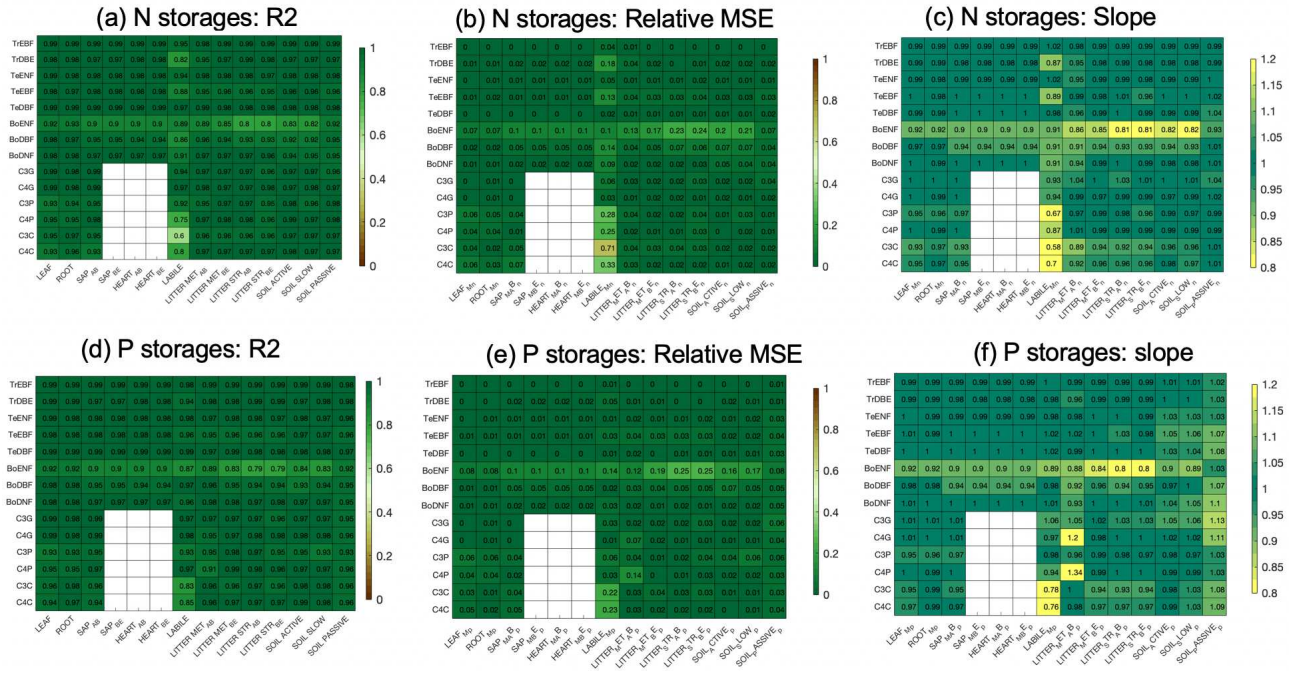


Figure S4.18 Performances of machine learning models (ML)-based historical simulation (i.e. Simulation ML) on mean N and P storages across 2000-2010, compared to Simulation EQ. The statistics are the same with Figure 3-5.

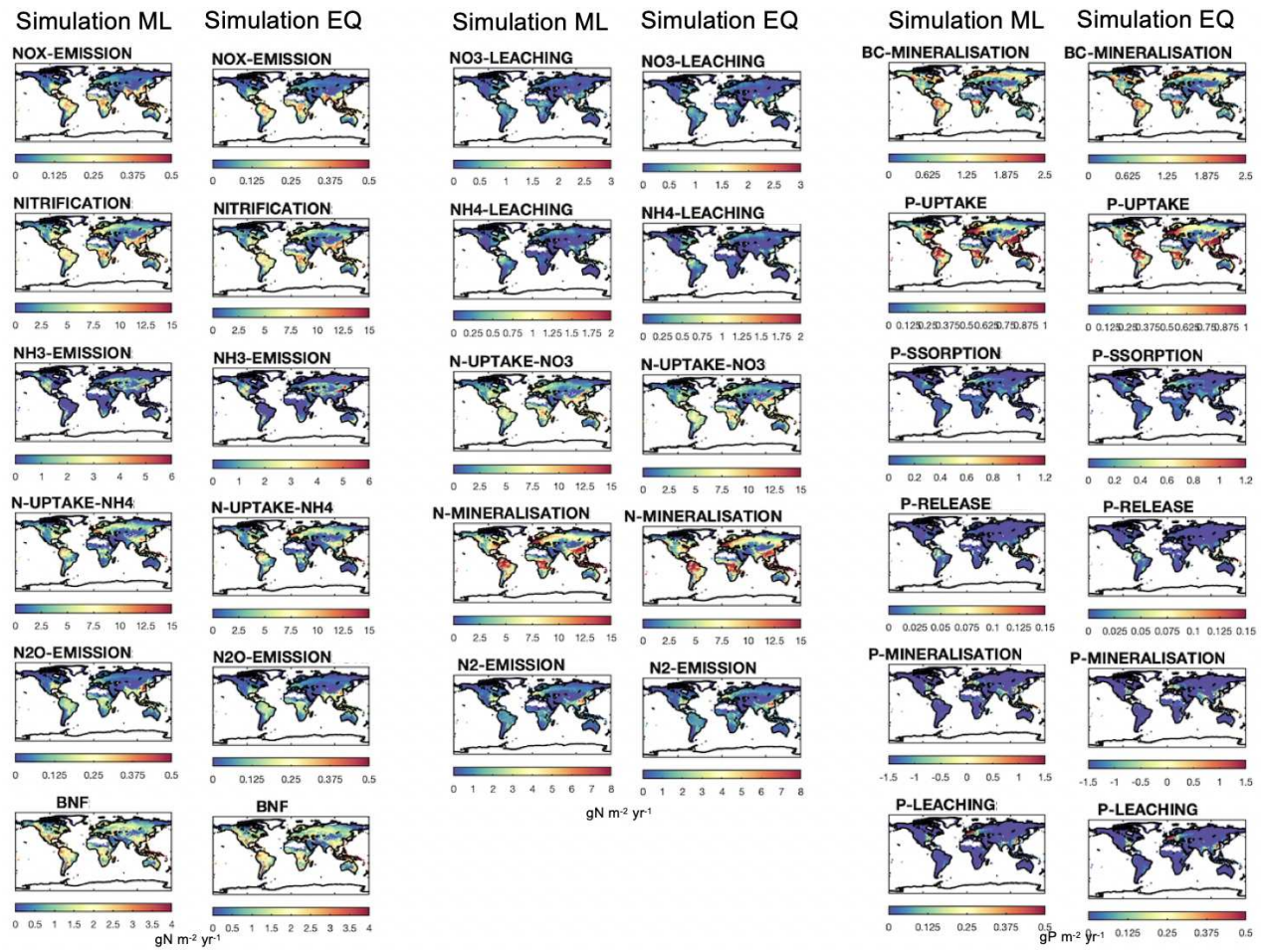


Figure S4.19 Global pattern of mean annual N and P flux components across 2000-2010 by Simulation ML and Simulation EQ.

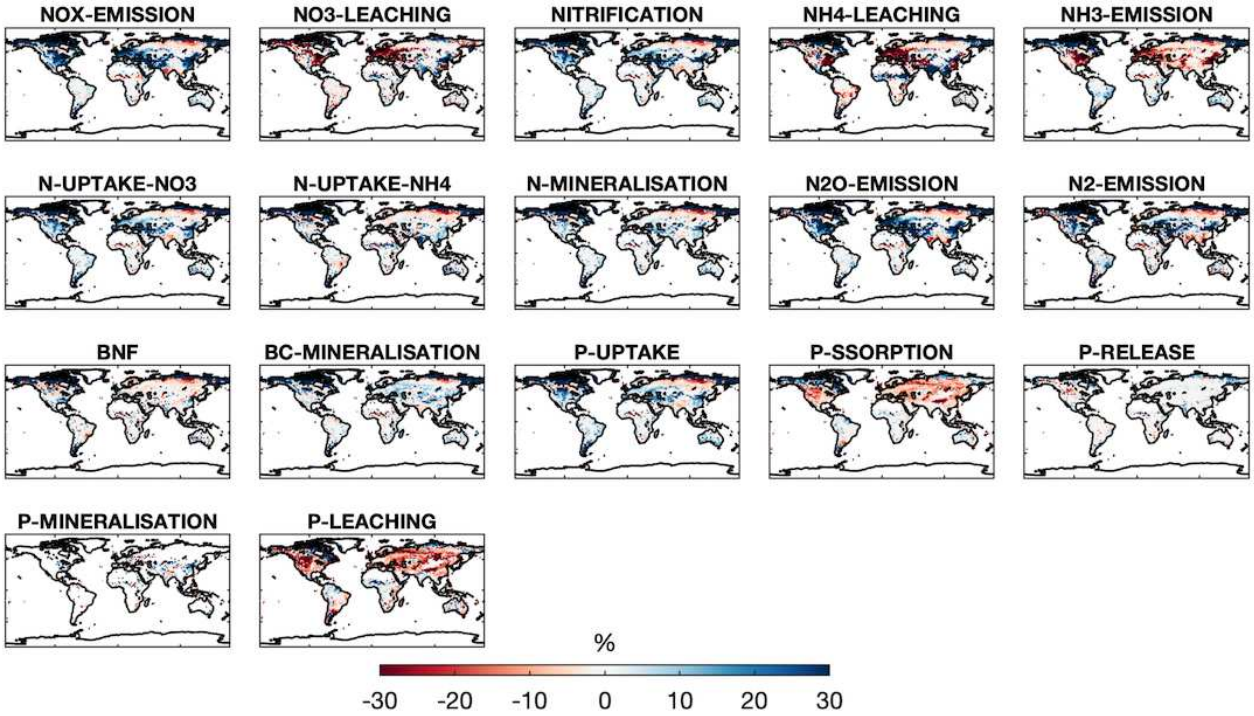


Figure S4.20 Relative bias in mean N and P flux components across 2000-2010 by Simulation ML compared to Simulation EQ. Pixels with the flux value lower than their 25th quantiles for Simulation EQ were masked.

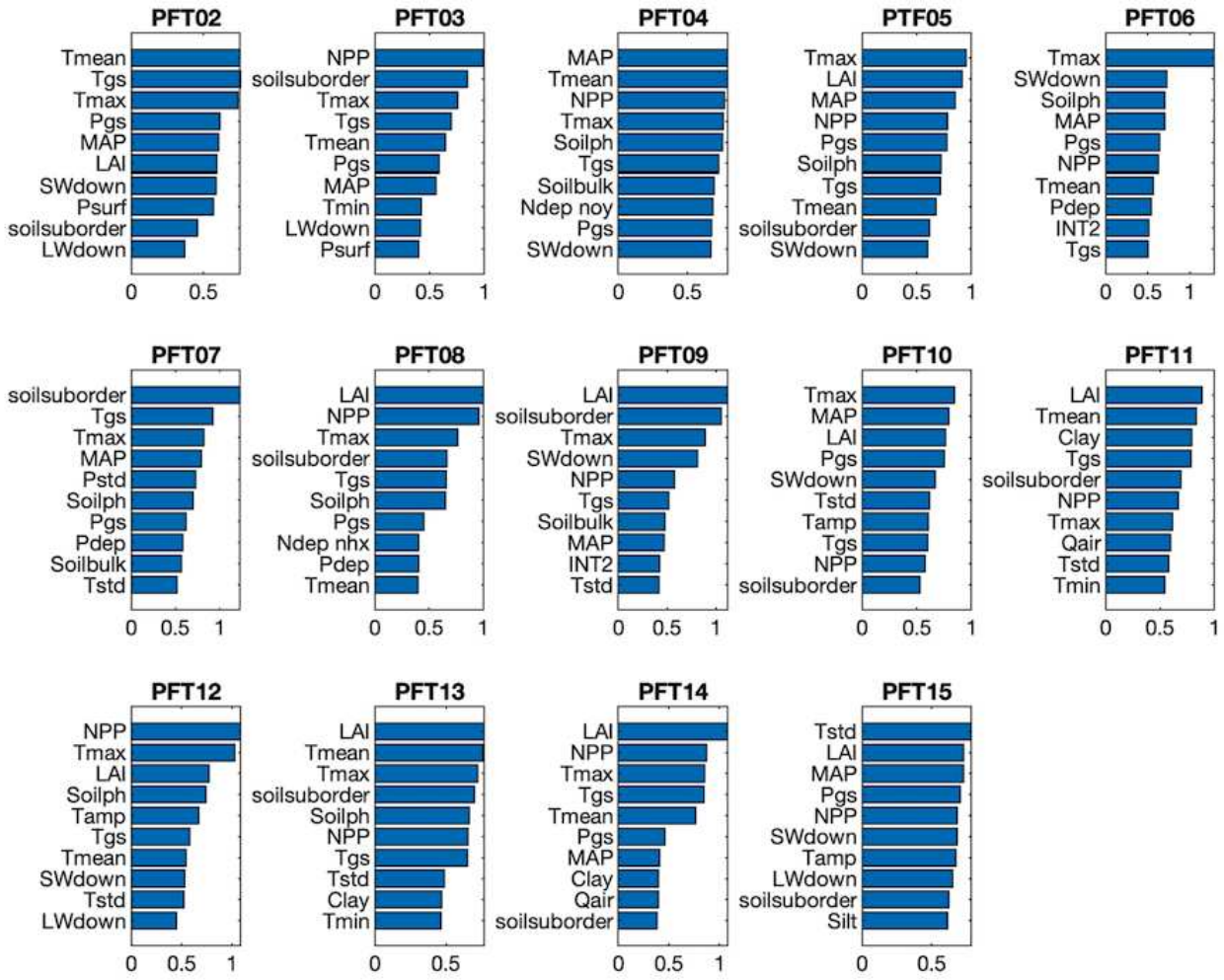


Figure S4.21 Predictor importance for predicting passive SOC of each PFT by ML using small fractions of pixels (Table S4.3) as the training sites.

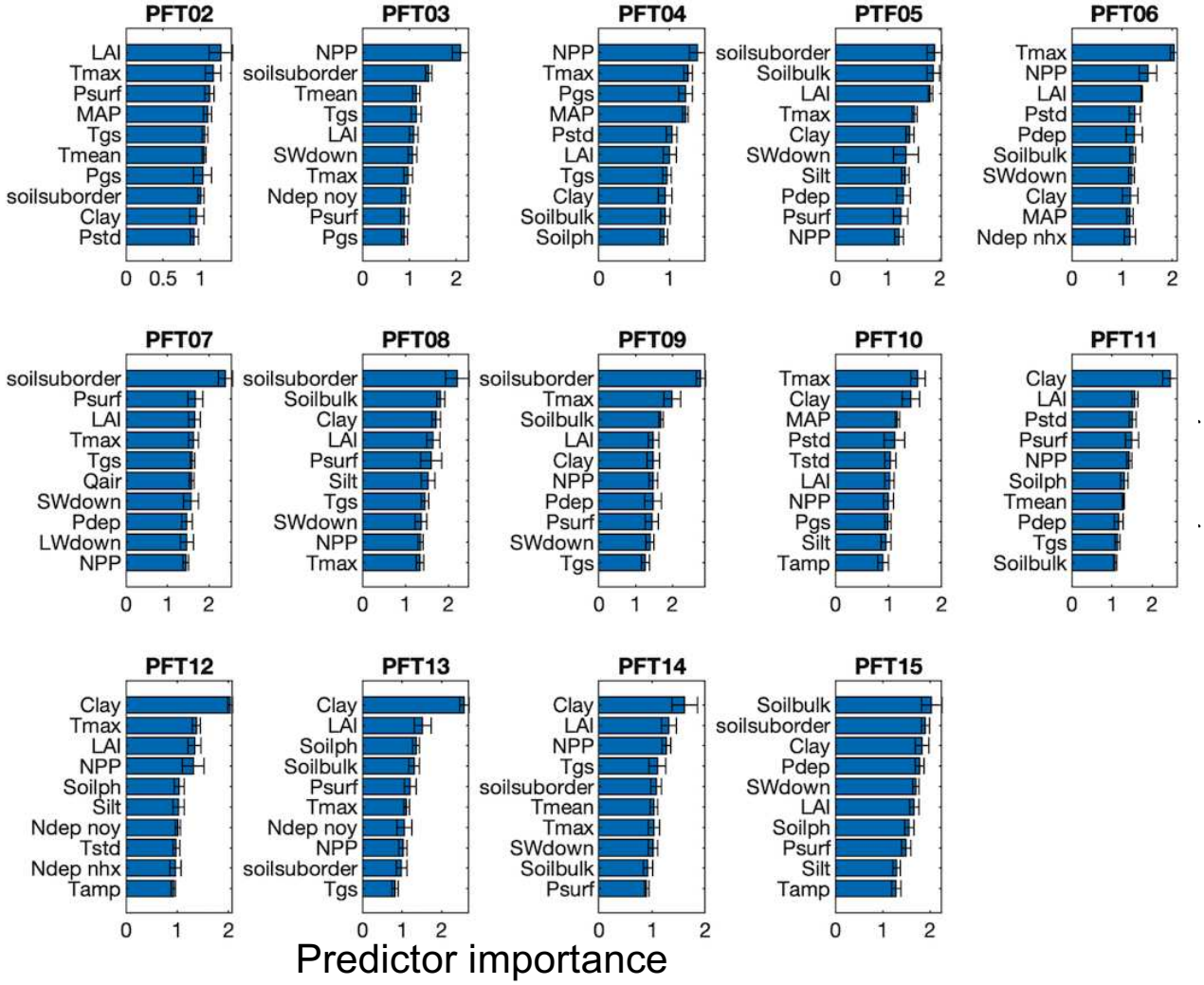


Figure S4.22 Predictor importance for predicting passive SOC of each PFT by ML using 100% pixels as the training sites.

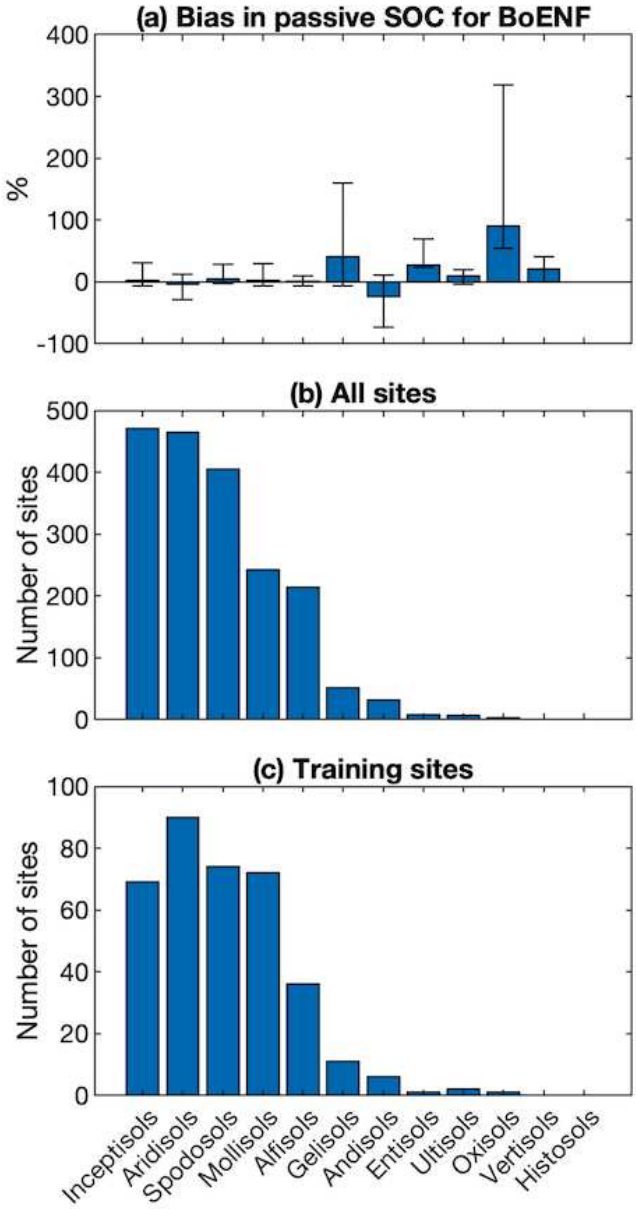


Figure S4.23 Relative Bias in passive SOC by ML (vs. Equi) for boreal evergreen needle-leaf forest (BoENF) on each soil suborder type. Corresponding numbers of all sites and selected training sites for each soil suborder type are also shown (b, c).

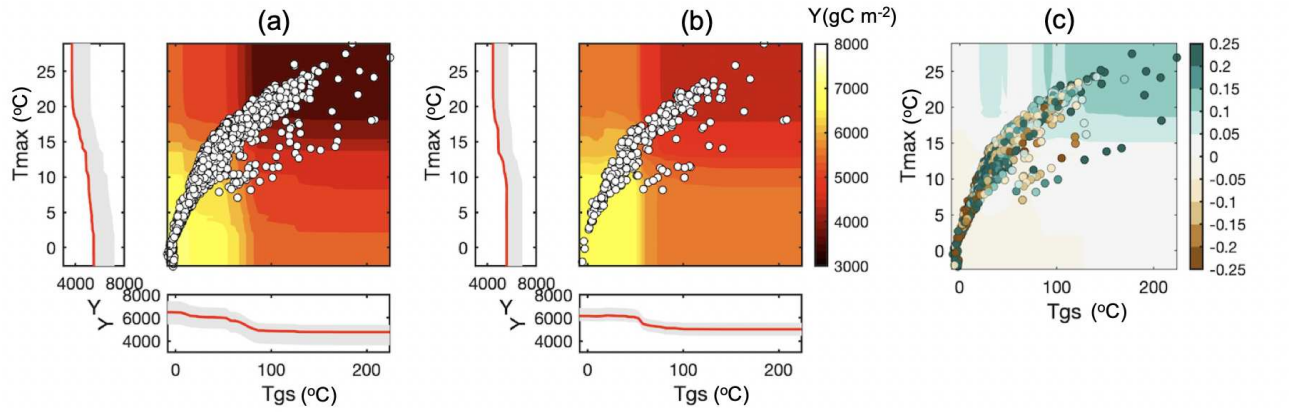


Figure S4.24 Two-way partial dependence plot between passive SOC of boreal evergreen needle-leaf forest (BoENF) and two predictors for ML, Tmax (maximum monthly temperature) and Tgs (accumulated monthly temperature during growing season with monthly temperature >-4 ^{o}C), using all of the sites (a) and sites for the ML training (b). Scatters in (a) and (b) show the distribution of sites in Tmax-Tgs space. The background color in (a) and (b) show the joint value of passive SOC (Y) by Tmax-Tgs using ML. Leaf and bottom panels show the variation of Y with Tmax and Tgs, respectively. (c) show the relative biases in in Y in LAI-Tgs space due to the low number of training sites (background color) and the distribution of sites with high relative bias by ML (>15%; vs. *"true equilibrium"* value; scatter color) in Tmax-Tgs space.

Table S4.1 Quantiles-specific weights of *response Y* for ML training.

	Quantiles-specific weights
Soil organic C, N and P storages	6 for >99.9%; 4 for <2%; 2 for 2~10% and >99%~99.5%; 1 for 10%~90%
Biomass and litter C, N and P storages	4 for >99%; 2 for >98%~99%; 1 for <98%

Table S4.2 Minimum K values should be considered in the K-mean cluster using different criterions. Note that higher K value can always reduce the sum of within-cluster distances (Figure S4.22), but may create some clusters containing very few samples.

PFT	PFT Name	Calinski-Harabasz	Davies-Bouldin	Silhouette
PFT2	Tropical Evergreen Broadleaf Forest (TrEBF)	2	2	2
PFT3	Tropical Deciduous Broadleaf Forest (TrEDF)	2	2	2
PFT4	Temperate Evergreen Needleleaf Forest (TeENF)	3	2	2
PFT5	Temperate Evergreen Broadleaf Forest (TeEBF)	3	2	2
PFT6	Temperate Deciduous Broadleaf Forest (TeDBF)	2	2	2
PFT7	Boreal Evergreen Needleleaf Forest (BoENF)	9	2	3
PFT8	Boreal Deciduous Broadleaf Forest (BoDBF)	9	2	3
PFT9	Boreal Deciduous Needleaf Forest (BoDNF)	3	3	2
PFT10	C3 grass	4	2	2
PFT11	C4 grass	3	2	3
PFT12	C3 pasture	3	2	3
PFT13	C4 pasture	2	2	2
PFT14	C3 grass (C3 crop in historical simulation)	2	2	2
PFT15	C3 grass (C4 crop in historical simulation)	3	3	3

Table S4.3 Fractions of selected training sites used for ML training for each PFT and for all pixels. Number of training sites for each cluster (N_c), number of training sites for ML (N_{ML}), and total number of pixels (N_{total}) for each PFT are also shown.

PFT	PFT Name	N_c	N_{ML}	N_{total}	Fraction (%)
PFT2	Tropical Evergreen Broadleaf Forest (TrEBF)	10	264	1516	17.4
PFT3	Tropical Deciduous Broadleaf Forest (TrEDF)	20	279	1627	17.1
PFT4	Temperate Evergreen Needleleaf Forest (TeENF)	10	478	2687	17.8
PFT5	Temperate Evergreen Broadleaf Forest (TeEBF)	10	442	2951	15.0
PFT6	Temperate Deciduous Broadleaf Forest (TeDBF)	10	300	1531	19.6
PFT7	Boreal Evergreen Needleleaf Forest (BoENF)	20	297	2213	13.4
PFT8	Boreal Deciduous Broadleaf Forest (BoDBF)	20	297	2219	13.4
PFT9	Boreal Deciduous Needleaf Forest (BoDNF)	20	326	2338	13.9
PFT10	C3 grass	10	280	1461	19.2
PFT11	C4 grass	10	341	1852	18.4
PFT12	C3 pasture	10	463	2719	17.0
PFT13	C4 pasture	10	366	1912	19.1
PFT14	C3 grass (C3 crop in historical simulation)	10	268	1526	17.6
PFT15	C3 grass (C4 crop in historical simulation)	10	305	2312	13.2
All			674	4770	14.1

Table S4.4 The R^2 for leave-one-out cross validation on training sets for passive soil organic carbon (Passive SOC), slow soil organic carbon (Slow SOC), aboveground biomass heartwood C (HeartAB C) and belowground biomass heartwood C (HeartBE C).

PFT	Passive SOC	Slow SOC	HeartAB C	HeartBE C
PFT2	0.87	0.83	0.94	0.94
PFT3	0.82	0.83	0.96	0.96
PFT4	0.92	0.93	0.98	0.98
PFT5	0.88	0.89	0.97	0.97
PFT6	0.9	0.9	0.97	0.97
PFT7	0.73	0.82	0.94	0.94
PFT8	0.85	0.86	0.96	0.96
PFT9	0.82	0.85	0.95	0.95
PFT10	0.92	0.82		
PFT11	0.89	0.9		
PFT12	0.92	0.91		
PFT13	0.9	0.89		
PFT14	0.91	0.92		
PFT15	0.88	0.89		

Table S4.5 Maximum potential performance of ML on Passive SOC and biomass HeartAB C pools when used 100% of global pixels for training.

	R^2				*reMSE*			
PFT	Passive SOC	Slow SOC	HeartABC	HeartBEC	Passive SOC	Slow SOC	HeartABC	HeartBEC
PFT2	0.99	0.997	0.999	0.999	0.01	0.003	<0.001	<0.001
PFT3	0.98	0.999	0.999	0.999	0.02	0.002	<0.001	<0.001
PFT4	0.99	0.999	0.999	0.999	0.01	<0.001	<0.001	<0.001
PFT5	0.98	0.999	0.999	0.999	0.02	<0.001	<0.001	<0.001
PFT6	0.99	0.999	0.999	0.999	0.02	0.002	0.001	0.001
PFT7	0.97	0.999	0.999	0.999	0.03	0.001	0.001	0.001
PFT8	0.97	0.999	0.998	0.998	0.03	0.003	0.002	0.002
PFT9	0.87	0.999	0.99	0.99	0.01	0.001	0.008	0.008
PFT10	0.92	0.999			0.02	<0.001		
PFT11	0.99	0.999			0.01	0.001		
PFT12	0.98	0.999			0.01	<0001		
PFT13	0.98	0.999			0.01	<0.001		
PFT14	0.99	0.999			0.02	0.001		
PFT15	0.99	0.999			0.02	0.001		

Chapter 5 Conclusions and perspectives

In summary, the starting point was that current land models (LSMs) with P biogeochemistry have large uncertainties in simulating C, N and P dynamics and interactions, which hinders a reliable simulation of the constraints of P on land C cycle. This thesis combined machine learning models, observations on various temporal and spatial scales, and land surface models to provide measures to reduce those uncertainties. Through combining site observations and machine learning methods, I generated a novel benchmark dataset for LSMs (chapter 2), which served, in combination with additional data, for model evaluation (chapter 3). Through using a comprehensive evaluation strategy tailored for LSMs with C, N and P interactions (chapter 3), I identified biases in a LSM in simulating coupled cycles of C, N and P and proposed ways to address them. Through developing a general approach to enhanced computational efficiency of P-enabled LSMs, I reduced the computational demand of spin-up for a LSM by about one order of magnitude, which facilitates data assimilation of the ever-growing observation datasets (chapter 4).

In answer to the first question(s): "Is it possible to produce a spatially explicit dataset of soil phosphatase activity from current scarce field data and what are major drivers of spatial variation?", I found that data of phosphatase is yet too sparse to produce reliable global maps, but sufficient for Europe. I designed an effective method explaining 58% of the data spatial variance in Europe. Phosphatase potential activity is mainly positive related to temperature and rainfall, and total soil nitrogen. This result suggests that multiple trade-offs between the biological demand for new inorganic P uptake (mainly associated to the overall ecosystem production determined by climate and soil conditions) and the capacity to synthetize phosphatase are the main causes underlying current soil phosphatase concentrations and activities.

In answer to the second question(s): "How well does the LSM ORCHIDEE reproduce observed nutrient-related properties of the land biosphere? Is it possible to identify processes which should be preferentially refined in future model development given? What are these processes?", I found that ORCHIDEE-CNP has overall sufficient skills in reproducing observational data and I was able to detect specific improvements and deteriorations due the inclusion of nutrient cycles. A clear identification of model processes responsible for model biases was only partly possible given the modelling setup and available observational data. Nonetheless, I was able to propose a focus on a set of processes (e.g. soil organic P mineralization and soil inorganic P transformation) to improve ORCHIDEE-CNP in next model versions.

In answer to the third question: "Is it possible to develop a general acceleration approach for the spin-up?", I found it is possible using machine learning methods to save ~80% of computational time for global spin-up compared to the conventional approach. Different from previous acceleration approaches with precondition of linear process and fixed ecosystem stoichiometries, this approach breaks those limits and can be widely used in LSMs, no matter if the model processes are linear or nonlinear. The reduction in the computational consumption by about one order of magnitude opens the opportunity of data assimilation using the ever-growing observation datasets as well as increasing the models spatial resolution.

The work presented in this thesis helps to prioritize areas of future model developments and provides a technical solution to facilitate model-data assimilation. It further identified promising areas of application of ORCHIDEE-CNP taking into account the lessons learnt from my PhD work.

(1) Model - data assimilation for ORCHIDEE-CNP:

One important uncertainty source of ORCHIDEE-CNP is related to the insufficient parameterization of processes, which often relies on (compilation of) scattered site data not representative for the diverse ecosystems on land. Data assimilation using the ever-growing observations (e.g. IMBALANCE-P project) and satellite datasets could help reduce the model-data misfit, and thus

reduce the uncertainty in simulating the nutrient constraints on C cycle. Besides, satellite datasets enable the model-data assimilation on large patterns, which help exploiting the spatial heterogeneity of key parameters (Reichstein & Carvalhais, 2019). For example, it has been shown that the spatial heterogeneity of parameters controlling plant allcoation in LSM can be derived from global MODIS LAI, biomass map and soil C datasets by using model-data fusion framework - revealing global emergent relationships of C allocation (Bloom et al., 2016). Current model-data assimilation tools available for ORCHIDEE (e.g. ORCHIDAS; Kuppel et al., 2012; Bacour et al., 2015; MacBean et al., 2015; Peylin et al., 2016; https://orchidas.lsce. ipsl.fr/) cannot be applied to optimize parameters which affect the simulated initial state due to the requirements of re-running the spin-up procedure each time changing such parameters. Due to the strong two-way coupling of processes operating on contrasting time scales in P-enabled LSM, e.g. coupling of 'fast' vegetation and 'slow' soil processes, the conventional approaches developed for C-only models are not suited as they rely on decoupling of process, e.g. optimization of vegetation parameter independently of the soil. Using the MLA-based spin-up procedure I have developed, it is becoming possible to realize ensemble of model simulations with ORCHIDEE-CNP which are ~10 larger than before. Nonetheless, still only a small subset of ORCHIDEE-CNP's parameters can be assimilated. Based on my analysis in chapter 3, parameters constraining the seasonality and inter-annual variability of C cycles should be a first priority. Specifically, parameters related to soil nutrient mineralization and nutrient uptake by plants, phenology dynamics and canopy light absorption need to be better calibrated.

(2) Large ensemble of model simulations for the past equilibrium land soil C storage

The evolution of climate since the Holocene was found to have substantially affected the current state of soil C stocks in terrestrial ecosystems (Delgado-Baquerizo et al., 2017). LSMs have been used to explore the footprint of paleoclimate in SOC accumulation, and found that the size of SOC of present-day is strongly correlated (R^2=0.99) with the initial soil carbon stock (e.g. Ding et al., 2019). Thus, it is critically important for projecting changes of SOC to investigate past states of land soil C storage. Previous studies showed that the size of SOC on ecosystem scale is controlled primarily by climate, vegetation productivity, biome type, texture (e.g. Wiesmeier et al., 2019), and soil total nitrogen and phosphorus (Rovai et al., 2019). However, there is a lack of understanding how SOC stocks have evolved during the Holocene, and what role nutrients have played in it. Soil P form and fractions have undergone changes along with the pedogenesis, vegetation succession, vegetation shifts, and extinction of large herbivores during the Holocene period (Eger et al., 2011), indicating a large difference of Holocene soil P availability from the present-day. Benefit from the high efficiency of MLA-based spin-up procedure designed in this thesis (Chapter 4), we can run large ensemble of model simulations with varied key parameters (e.g. Vcmax), nutrient availability, different climate forcing, and soil priorities, to diagnose the key factors in determining the equilibrium land soil C storage.

(3) Application of ORCHIDEE-CNP in simulating C cycle in tropical forests:

Intact tropical forests act as a substantial aboveground C sink having increased biomass stocks over time (Brienen et al., 2015; Phillips et al.,2017). The evolution of the aboveground C sink as well as potential belowground C sinks is potentially constrained by emerging P limitation (e.g. Fleischer et al., 2019). Datasets from free-air concentration enrichment (FACE) experiments and/or open-top chambers (e.g. Amazon FACE; Hofhansl et al., 2016), and nutrient addition (e.g. IMBALANCE-P; Van Langenhove et al., 2020) in ecosystems on P-poor soils and site observations for soil and vegetation (e.g. Soong et al., 2020) are emerging which allow to explore the role of P in the C cycle response to increasing CO_2 concentration. The current version of ORCHIDEE-CNP has a good performance in tropical regions (chapter 3), and thus could be used for studying the mechanism of plant adaptation and plasticity to P deficient in tropical regions.

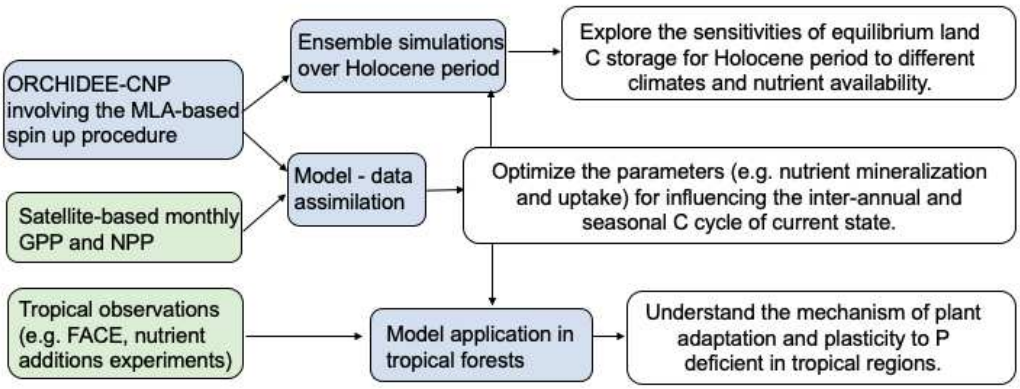


Figure 5. Possible model applications of ORCHIDEE-CNP.

References

Bacour, C., Peylin, P., MacBean, N., Rayner, P., Delage, F., Chevallier, F., ... Berveiller, F. (2015). Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model. *Journal of Geophysical Research: Biogeosciences*, *120*, 1839–1857.

Brienen, R. J. W. et al. Long-term decline of the Amazon carbon sink. *Nature* **519**, 344–348 (2015).

Bloom, A. A., Exbrayat, J. F., Van Der Velde, I. R., Feng, L., & Williams, M. (2016). The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences*, *113*(5), 1285-1290.

Delgado-Baquerizo, M., Eldridge, D. J., Maestre, F. T., Karunaratne, S. B., Trivedi, P., Reich, P. B., & Singh, B. K. (2017). Climate legacies drive global soil carbon stocks in terrestrial ecosystems. *Science Advances, 3(4)*, e1602008.

Ding, J., Wang, T., Piao, S., Smith, P., Zhang, G., Yan, Z., ... & Dai, F. (2019). The paleoclimatic footprint in the soil carbon stock of the Tibetan permafrost region. *Nature communications, 10(1)*, 1-9.

Eger, A., Almond, P. C., & Condron, L. M. (2010). Quantifying the soil-and ecosystem-rejuvenating effects of loess in a high leaching environment, West Coast, New Zealand. *In 19th World Congress of Soil Science*.

Fleischer, K., Rammig, A., De Kauwe, M. G., Walker, A. P., Domingues, T. F., Fuchslueger, L., ... & Haverd, V. (2019). Amazon forest response to CO 2 fertilization dependent on plant phosphorus acquisition. *Nature Geoscience*, *12*(9), 736-741.

Hartmann, J., Moosdorf, N., Lauerwald, R., Hinderer, M., & West, A. J. (2014). Global chemical weathering and associated P-release—The role of lithology, temperature and soil properties. *Chemical Geology*, *363*, 145-163.

Hofhansl, F., Andersen, K. M., Fleischer, K., Fuchslueger, L., Rammig, A., Schaap, K. J., ... & Lapola, D. M. (2016). Amazon forest ecosystem responses to elevated atmospheric CO_2 and alterations in nutrient availability: filling the gaps with model-experiment integration. *Frontiers in Earth Science*, *4*, 19.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., & Richardson, A. (2012). Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, *9*, 3757–3776.

Maire, V., Wright, I. J., Prentice, I. C., Batjes, N. H., Bhaskar, R., Bodegom, P. M., ... Reich, P. B. (2015). Global effects of soil and climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography*, *24*, 706–717.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P. J., Kuppel, S., ... Prunet,P.(2016).Anewstep-wiseCarbonCycleDataAssimilationSystem using multiple data streams to constrain the simulated land surface carbon cycle. *Geoscientific Model Development Discussions*, *9*, 3321–3356.

Penuelas, J., Fernández-Martínez, M., Vallicrosa, H., Maspons, J., Zuccarini, P., Carnicer, J., ...

Reichstein, M., & Carvalhais, N. (2019). Aspects of forest biomass in the earth system: Its role and major unknowns. *Surveys in Geophysics*, *40*(4), 693-707.

Phillips, O. L. & Brienen, R. J. W. Carbon uptake by mature Amazon forests has mitigated Amazon nations' carbon emissions. *Carbon Balance Manag.* **12**, 1 (2017).

Rovai, A. S., Twilley, R. R., Castañeda-Moya, E., Riul, P., Cifuentes-Jara, M., Manrow-Villalobos, M., ... & Pagliosa, P. R. (2018). Global controls on carbon storage in mangrove soils. *Nature Climate Change, 8(6)*, 534-538.

Soong, J. L., Janssens, I. A., Grau, O., Margalef, O., Stahl, C., Van Langenhove, L., ... & Freycon, V.

(2020). Soil properties explain tree growth and mortality, but not biomass, across phosphorus-depleted tropical forests. *Scientific reports*, *10*(1), 1-13

Van Langenhove, L., Janssens, I. A., Verryckt, L., Brechet, L., Hartley, I. P., Stahl, C., ... & Peguero, G. (2020). Rapid root assimilation of added phosphorus in a lowland tropical rainforest of French Guiana. *Soil Biology and Biochemistry*, *140*, 107646.

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., ... & Wollschläger, U. (2019). Soil organic carbon storage as a key function of soils-a review of drivers and indicators at various scales. *Geoderma*, 333, 149-162.