

Field-based tree mortality observations constrain model-projected forest carbon sinks across continents

Kailiang Yu (✉ ky9hc@virginia.edu)

Laboratoire des Sciences du Climat et de l'Environnement

Philippe Ciais

Laboratoire des Sciences du Climat et de l'Environnement <https://orcid.org/0000-0001-8560-4943>

Sonia Seneviratne

Swiss Federal Institute of Technology in Zurich <https://orcid.org/0000-0001-9528-2917>

Zhihua Liu

University of Montana

Han Chen

Lakehead University <https://orcid.org/0000-0001-9477-5541>

Jonathan Barichivich

Laboratoire des Sciences du Climat et de l'Environnement

Craig Allen

University of New Mexico

Hui Yang

Laboratoire des Sciences du Climat et de l'Environnement

Yuanyuan Huang

CSIRO Ocean and Atmosphere

Ashley Ballantyne

Laboratoire des Sciences du Climat et de l'Environnement

Article

Keywords: Atmospheric CO₂ Sequestration Capacity, Biomass Loss by Mortality, Net Primary Productivity, Heterotrophic Respirations, Net Carbon Sink

Posted Date: May 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-544120/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on April 19th, 2022. See the published version at <https://doi.org/10.1038/s41467-022-29619-4>.

Field-based tree mortality observations constrain model-projected forest carbon sinks across continents

Kailiang Yu^{1,2*}, Philippe Ciais¹, Sonia Seneviratne³, Zhihua Liu², Han Y.H. Chen⁴, Jonathan Barichivich^{1,5}, Craig D. Allen⁶, Hui Yang^{1,7}, Yuanyuan Huang^{1,8}, Ashley P. Ballantyne^{1,2}

¹Le Laboratoire des Sciences du Climat et de l'Environnement, IPSL-LSCECEA/CNRS/UVSQ
Saclay, Gif-sur-Yvette, France

²Department of Ecosystem and Conservation Sciences, University of Montana, Missoula, USA

³Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland

⁴Faculty of Natural Resources Management, Lakehead University, Thunder Bay, Ontario,
Canada

⁵Instituto de Geografía, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

⁶Department of Geography and Environmental Studies, University of New Mexico,
Albuquerque, NM, USA

⁷Max Planck Institute for Biogeochemistry, Jena, Germany

⁸CSIRO Oceans and Atmosphere, Aspendale, Australia

Abstract Considerable uncertainty and debate exist in projecting the future capacity of forests to sequester atmospheric CO₂. Here we generate spatially explicit patterns of biomass loss by mortality (LOSS) from a dataset (n = 2676) of long-term (1951 to 2018), largely unmanaged forest plots to constrain projected (2015-2099) net primary productivity (NPP), heterotrophic respiration (HR) and net carbon sink in six dynamic global vegetation models (DGVMs) across North and South America, Africa, Asia, and Australia. This approach was motivated by the higher accuracy of field-derived LOSS estimates and the strong relationship among LOSS, NPP, and HR at continental or biome scales. The field observations showed higher values of LOSS in tropical regions (0.53 Kg m⁻² y⁻¹) than in North America (0.22 Kg m⁻² y⁻¹). The upscaled gridded LOSS map show hotspots in Southern Asia & Australia, Northwestern South America, and the western coast of North America. The DGVMs overestimated LOSS, particularly in tropical regions and eastern North America by as much as 0.5 Kg m⁻² y⁻¹. The spread of DGVM-projected NPP and HR uncertainties was substantially reduced after constraining the simulations with the observed LOSS patterns. The observation-constrained models show a decrease in the tropical forest carbon sink by the end of the century, particularly across South America (from 2 to 1.3 PgC y⁻¹), and an increase in the sink in North America (from 0.75 to 0.97 PgC y⁻¹). These results suggest that forest demographic data can be used to constrain land carbon sink projections.

Main text

Forests are major drivers of biophysical land-atmosphere feedbacks, the global carbon and water cycles, and thus overall planetary climate regulation¹ – yet large uncertainties and substantial debate persist regarding ongoing and potential changes in the capacity of forests carbon sinks across continents or biomes (i.e., tropical vs boreal) under anticipated future climates^{2,3}. It is often thought that increased net primary productivity (NPP) across spatial resource gradients or under elevated CO₂ will increase carbon sinks^{4,5}. The loss of CO₂ to the atmosphere through ecosystem heterotrophic respiration (HR), however, has been observed to increase worldwide⁶ and could lead to a trade-off or even outweigh the positive effects of NPP (i.e., at regional scales)⁷, thus potentially amplifying future climate change via enhanced rates of carbon emission to atmosphere^{8,9}. Reducing uncertainties about the future forest carbon sink thus requires an integrated understanding of NPP and HR across space or time.

Recent studies have suggested that faster forest growth (NPP) leads to higher tree mortality at local and regional scales, particularly in tropical forests^{10,11}. This has been recently demonstrated across spatial scales in boreal forests using tree-ring datasets¹² and across forest biomes in Earth system models (ESMs)¹³. While the fraction of biomass subjected to tree mortality is often treated as a constant proportion of standing stocks in ESM simulations¹³, this steady state representation of mortality and growth is not necessarily observed in forests responding to global change. Ongoing climate change can result in disproportionate mortality relative to growth, from direct physiological mortality from more extreme drought and heat events, insect outbreaks, windthrow, lightning, and wildfire^{14,15}. In contrast, growth may be exceeding mortality in some forests due to increased atmospheric CO₂ reducing drought stress¹⁶ and/or nitrogen deposition enhancing growth¹⁷. At long-term and broad spatial scales, however,

NPP, biomass loss from tree mortality (hereafter LOSS) and HR could be coupled with their strong positive relationships because of the linkage of aboveground and belowground processes^{18,19}, while the instantaneous rates of NPP, LOSS and HR could be decoupled over short term or local scales^{20,21}. Thus, data on LOSS which can be directly measured in ground-based forest plots with high certainty¹³ and linked or correlated with NPP and HR, may provide an unique constrain on the forest carbon sink in a future changing climate.

Uncertainty in projections of forest carbon cycling could be reduced through an emergent constraint (EC) approach by identifying heuristic relationships between a multi-model ensemble and an observational estimate. This approach has been used to constrain the projected land carbon storage²², or photosynthesis/GPP^{23–25}, with the assumption that the processes driving the long-term response are also driving the historical (short-term) patterns^{22,23}. Interpreting the results of emergent constraints needs caution in the confirmation of the verified mechanisms^{26,27} and the mismatch of spatial and temporal scales between data and models^{25,26}. Earlier studies usually used atmospheric variables (i.e., CO₂ concentrations) to constrain the carbon storage and photosynthesis at regional to global scales^{22,23}. Using spatially explicit observational products (i.e., GPP, evapotranspiration and leaf area index), studies have emergently constrained the future terrestrial carbon cycling projections at grid scale and then spatially aggregated to broader spatial scales^{24,28}. Recent studies highlight the potential of integrating emergent constraints at lumped broad spatial scales and machine learning to generate the spatially explicit constraint of projected gross primary production (GPP) by accounting for the non-linear relationships between GPP and environmental drivers²⁹. However, to date it remains unclear whether the increased availability of ground-based datasets (forest plot data) could better constrain the projected forest carbon sink at broad spatial scales in a future climate. Because observational uncertainty has

more influence than model ensemble uncertainty in EC²⁶, a LOSS constraint based on forest plot datasets is urgently needed and expected to largely reduce the uncertainties in projections of future forest carbon sinks and associated feedbacks to climate.

Here we generate spatially explicit patterns of LOSS from long-term (1951 to 2018) forest plot data of largely unmanaged forest plots ($n = 2676$; Supplementary Fig. 1) to constrain projected (2015-2099) NPP, HR and net ecosystem exchange ($NEE = NPP - HR$) in six DGVMs across continents - North America, South America, Africa and Asia & Australia. This approach is motivated by positive relationships (Supplementary Figs. 2,3) between historical LOSS and projected NPP and HR carbon fluxes in DGVMs, in line with the known pattern of faster growth and higher mortality¹⁰⁻¹³ and the couplings of growth, mortality and respiration at long-term and broad spatial scales^{18,19}. We used LOSS because: 1) it can be directly measured in forest inventory datasets¹³ with higher accuracy than input fluxes such as growth or woody NPP (see Methods). Indeed, quantification of growth includes components of recruitment of new trees and growth of surviving trees¹³, and differing thresholds of diameter at breast height as a criterion to determine the recruitment of new trees could lead to large uncertainties in estimates of growth across forest plots³⁰; and 2) the accuracy of constraining the projected carbon cycle in DGVMs largely depends on the observational uncertainty²⁶. We first used a random forest to upscale spatial variations of LOSS with 58 environmental variables (see Supplemental Table 1) to generate a spatially explicit map of LOSS at 0.25-degree resolution across continents (see Methods). Second, we compared the observed spatial patterns of LOSS with DGVM³¹ patterns in which tree mortality was explicitly reported (see Methods). Finally, two complementary approaches – a conventional emergent constraint and a novel machine learning constraint – were used to constrain the projected NPP and HR at continental scale in DGVMs (see Methods,

Supplementary Fig. 4). The emergent constraint (EC) approach was applied by identifying a linear relationship between historical LOSS, averaged at local forest-plot scale or continental scale, and projected NPP and HR, aggregated as sums at continental scale across the DGVM ensemble. By comparison, the machine learning (ML) approach was used to examine the non-linear relationships by training a ML (random forest) model between historical simulated LOSS and projected NPP and HR for each DGVM at grid scale.

The original forest plot data aggregated at 0.25 degree show large spatial variations (Fig. 1a) across the continents, with the greatest LOSS in Asia & Australia (mean \pm 1 SE; 0.65 ± 0.05 Kg m⁻² y⁻¹) > South America (0.49 ± 0.02 Kg m⁻² y⁻¹) and Africa (0.46 ± 0.02 Kg m⁻² y⁻¹) > North America (0.23 ± 0.01 Kg m⁻² y⁻¹ in boreal and 0.2 ± 0.01 Kg m⁻² y⁻¹ in temperate)³² (Fig. 1b; Supplementary Fig. 5a). This pattern was robust to bootstrapping (1000 iterations) to randomly select 90% of plots for estimating the probability distribution of the mean continental values (Supplementary Fig. 5b). The upscaled gridded LOSS maps generated by our random forest algorithm (see Methods) over the spatial domain of our full datasets shows hotspots of high LOSS in Southern Asia & Australia (> 0.6 Kg m⁻² y⁻¹), Northwestern South America (Amazon) (> 0.5 Kg m⁻² y⁻¹), and the western coast of North America (> 0.3 Kg m⁻² y⁻¹)^{10,32-34} (Supplementary Fig. 6a), where climate is warm and dry – at least at seasonal scales. These patterns were robust to two bootstrapping approaches – based on the sampled biomes of each point feature and also randomly sampling 90% data with replacement (see Methods) (Fig. 2a; Supplementary Fig. 6b). The uncertainty (coefficient of variance - CV; %mean) was generally low (<10%) across continents, with the exception of temperate forests in North America (CV > 10%), despite the larger sample size (n > 500 at 0.25 degree) (Fig. 2b; Supplementary Fig. 6c), likely because of potential effects of forest recovery or regrowth following past disturbance³⁵.

Mean annual temperature (MAT), aridity index (the ratio of precipitation to potential evapotranspiration), and precipitation seasonality were identified as the dominant predictors of LOSS across continents (Supplementary Fig. 7a), with positive relationships with LOSS (Fig. 3a)^{10,32}. In contrast to local-scale studies^{36,37}, wood density, forest stand density, and soil conditions were poor predictors of LOSS when all data were used. These relationships were largely driven by the spatial pattern of LOSS and climate gradients, whereby LOSS and MAT, aridity index, and precipitation seasonality were high in wet tropical forests (Supplementary Fig. 8). This motivated us to examine the drivers of LOSS in tropical vs non-tropical biomes (Supplementary Fig. 7b, c; Fig. 3b, c, d). With a smaller gradient in climate within wet tropical forests, soil properties such as nutrient content and cation exchange capacity (CEC) were significant predictors of LOSS (Supplementary Fig. 7b; Fig. 3b)³⁸. In wet tropical forests, the relationships between soil nutrient content and CEC and LOSS were positive (Fig. 3b) and thus appeared to support the framework of higher mortality in more productive tropical forests in locations of nutrient enriched soils^{38,39}. In non-tropical regions, basal area or a competition index based on the degree of crowding within stocked areas⁴⁰ (see Methods) were the dominant predictors of LOSS in extra-tropical North America with their positive relationships (Supplementary Fig. 7c; Fig. 3c, d), thereby highlighting the role of competition in driving the spatial patterns of LOSS^{40,41}. This pattern also support the existence of a spatial tradeoff between faster growth and higher mortality because of resource limitations or younger death, whereby competition plays the fundamental role^{13,41}. In contrast to other studies^{15,42}, forest age (available in boreal and temperate forests in North America) was not a good predictor of LOSS (Supplementary Fig. 9), likely because of our focus on mature and old-growth forests (i.e., age > 80 and 100 years in boreal and temperate forests, respectively).

We then compared the observed patterns of LOSS with those simulated by six state-of-the-art DGVMs in which tree mortality and LOSS were explicitly simulated³¹. The results show divergent predictions of LOSS among DGVMs with four models (ORCHIDEE, JULES, LPJmL, and SEIB-DGVM) overestimating LOSS compared to our observation-based estimate, particularly in tropical forests and temperate eastern North America, while CABLE-POP and LPJ-GUESS underestimate LOSS across continents (Supplementary Fig. 10). This led to a model ensemble mean overestimation of LOSS, particularly in tropical forests (historical $\Delta\text{LOSS} > 0.5 \text{ Kg m}^{-2} \text{ y}^{-1}$) and eastern North America ($\Delta\text{LOSS} > 0.4 \text{ Kg m}^{-2} \text{ y}^{-1}$) (Figs. 2c, 2e), while the spread of LOSS prediction between models was greatest ($\text{CV} > 130\%$) in western boreal forests in North America (Fig. 2d).

With the emerging positive and linear relationships between LOSS and projected NPP and HR across grids within each continent (Supplementary Fig. 2), we used the conventional emergent constraint approach²² to constrain future (2015-2099) NPP and HR aggregated across continents by using the historical LOSS averaged at forest-plot scale (derived from original plot data of LOSS) or continental scale (derived from the map of LOSS) (Supplementary Fig. 4; see Methods). We found that the emergent constraint approach worked well in North America, where the relationship between historical LOSS and projected NPP and HR was significant (the scenario of using original plot data of LOSS: $R^2 = 0.68$ and $P = 0.04$ for grid-level NPP; $R^2 = 0.97$ and $P = 0.0001$ for grid-level HR; the scenario of using map of LOSS at continent scale: $R^2 = 0.7$ and $P = 0.04$ for grid-level NPP; $R^2 = 0.95$ and $P = 0.0008$ for grid-level HR) (Supplementary Fig. 11a; Supplementary Fig. 12a). This emergent constraint approach was less effective for other continents, where tropical forests are predominant (all $P > 0.05$; Supplementary Fig. 11b, c, d; Supplementary Fig. 12b, c, d). These results suggest a weak linear

relationships when observations are lumped at broad continental scales, thus highlighting the importance of spatial scale in this emergent constraint²⁶. We interpret the result that this LOSS emergent constraint works better in North America than in the tropical forests, by a better representation of forest plot distribution and couplings of LOSS and NPP and HR across space in North America.

To overcome this limitation, we trained a machine learning algorithm²⁹ to reproduce the emerging relationship between historical LOSS and projected NPP and HR at grid level in each DGVM by incorporating all grid values without or with climate predictors, expressed as NPP_{pro} or $HR_{pro} = f(LOSS_{his})$ or $f(LOSS_{his}, MAT_{pro}, MAP_{pro})$, respectively, where pro refers to projected variables, his refers to historical variables, and MAT and MAP is mean annual temperature precipitation, respectively (see Methods). The results show consistently positive non-linear relationships between $LOSS_{his}$ and NPP_{pro} or HR_{pro} across DGVMs (Supplementary Fig. 3) and our machine learning algorithms can surrogate well the results of process-based models between the historical LOSS and the projected NPP and HR ($R > 0.65$ and $R > 0.9$ in both scenarios without climate effects and with climate effects, respectively; see Methods) (Supplementary Fig. 13). After including the $LOSS_{his}$ (derived from LOSS) in the machine learning algorithm, we were able to generate spatially-explicit constrained estimates²⁹ of projected NPP and HR, and then compare them with the scenario without the constraint (Supplementary Fig. 14; Supplementary Fig. 15). These patterns essentially show a lower NPP_{pro} or HR_{pro} in locations of overestimated $LOSS_{his}$ in DGVMs, consistent with the positive relationship between $LOSS_{his}$ and NPP_{pro} or HR_{pro} (Supplementary Fig. 3).

Our results show that DGVMs overestimate tree mortality, particularly in tropical regions (Figs. 2c, 2e). Thus, if modeled mortality is over-estimated, we expect that NPP is over-

estimated as well. Our constraint approach by LOSS reduces this common bias of models and decreases projected NPP down to 7.4, 2.2, 1.9 Pg C y^{-1} in South America, Africa and Asia & Australia, compared to original NPP values of 9, 2.4, 2.3 Pg C y^{-1} (Fig. 4a). The reason for this is that NPP or growth is strongly positively correlated with LOSS across space in both inventory data and DGVMs (Supplementary Figs. 2,3; Supplementary Fig. 16). The constant mortality used in most models may be too large if modelers have tuned this parameter to obtain reasonable biomass stocks, thus compensating for an overestimate of NPP in absence of modeled competition between individuals and nutrients (e.g. phosphorus) limitations in tropical forests¹³. Likewise, HR_{pro} showed similar patterns with NPP_{pro} because of coupling of HR and NPP and LOSS at broad spatial and long term scales^{18,19}, despite the likely decoupling of the instantaneous rate of HR and NPP and LOSS at local and short-term scales^{20,21}. Thus, we also constrained a decrease in projected grid-level HR with values of 6.2, 1.8, 1.6 Pg C y^{-1} in South America, Africa and Asia & Australia compared to 7, 1.9, 1.8 Pg C y^{-1} in the original model ensemble (Fig. 4b). Taken together, our results constrain a weaker future tropical forest carbon sink from observation-based LOSS estimates down to 1.3, 0.4, 0.3 Pg C y^{-1} in South America, Africa and Asia & Australia as compared to 2, 0.5, 0.5 Pg C y^{-1} in the original models. The sink is reduced in particular over the Amazon basin, while North America showed an enhanced future carbon sink (0.97 and 0.75 Pg C y^{-1} after and before constraint, respectively). The constraint by the machine learning approach substantially similarly reduced the model spread in grid-level NPP_{pro} and HR_{pro} across continents (Fig. 4; Table 1).

Our results were robust to the inclusion of projected climate (temperature and precipitation) across space in the machine learning algorithm (Supplementary Fig. 17), while we note that our approach does not account for the slow varying effects of temporal climate change

and atmospheric CO₂ concentrations^{22,23,25}. Indeed, our study focuses on the carbon flux – NPP and HR and carbon sink averaged over the long-term projected future (2015-2099) in mature forests across continents. The mechanistic basis underlying this approach is the observed pattern of faster growth and higher mortality and thus coupling of growth, mortality, and respiration averaged over long-term and broad spatial scales, which have been demonstrated in forest inventory datasets^{11,43}, tree rings¹², eddy flux towers¹⁸ and Earth System models¹³ and DGVMs here. Our approach assumes that this coupling holds both in the short-term historical period and will hold in the future, a space for time approximation allowing us to use observed historical LOSS to constrain the future projected NPP and HR carbon fluxes. Models typically have a simple representation of tree mortality processes, as a fraction of standing stock¹³, but increasingly more detailed global datasets of tree biomass and demography will also make it possible to test more realistic simulations of competition to access to light, water and nutrients in the next generation of models^{31,44}. The finite resources and vegetation carrying capacity govern the tradeoffs between growth and mortality and respiration across space or time. Our results indicate that the projected increase in tropical forest productivity may not be as great as previously thought, especially over the Amazon, which may offset projected increases in boreal and temperate forest productivity, thereby reducing the carbon sink potential of global forests.

Methods

Forest plot datasets

Forest plot data used in this study met the following criteria¹³: (1) all plots had at least three consecutive censuses and long-term (>9 years) observations between the first and last census so that LOSS averaged over all censuses was representative of the historic forest status; (2) plots were natural, unmanaged forest stands that have not been disturbed by fires, harvesting, and

other human activities; (3) the plots were largely mature or old-growth forests and were screened by criteria such as forest age or forest gymnosperm fraction (see Supplementary Information for details); (4) the plots were in a quasi-steady state and were screened by criteria that plots with growth more than 3 times of LOSS or LOSS more than 3 times of growth were excluded – such large differences (i.e., more than 3 times) between LOSS and NPP are likely due to disturbances such as fires that are not well-represented yet in DGVMs. Thus this criterion is thus used to further select natural mature and old-growth forest stands; and (5) the plots with low values of biomass (i.e., $< 3 \text{ Kg m}^{-2}$) were excluded because they are not fully stocked and thus not likely to be mature forests. Ultimately, we compiled a broad-scale ($n = 2676$) and long-term (1951 to 2018) dataset of largely unmanaged forest plots in a quasi-steady state, distributed in Canada, USA, Amazon, Africa and Asia & Australia (Supplementary Fig. 1). To be comparable with DGVMs, the aboveground woody biomass loss was converted to total woody biomass loss including belowground roots using the root-shoot biomass ratio product⁴⁵. More details for the criterion of plots selected, plot establishment, and measurements are described in Supplementary Methods. We stress the advantage of using LOSS rather than growth or woody NPP to constrain the projected carbon sink because of different thresholds of a diameter at breast height to record the new recruitments to quantify growth across forest plots^{13,30}. For instance, the criteria of 12.7 cm used in the Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service is much higher than Canada forests and thus underestimated the growth because of because of limited records of recruitments.

DGVMs

Simulated NPP, mortality flux, and HR of six DGVMs (ORCHIDEE, CABLE-POP, JULES, LPJ-GUESS, LPJmL, and SEIB-DGVM) were publicly available from the carbon turnover inter

comparison project <https://zenodo.org/communities/veg-c-turnover-comp/?page=1&size=20>³¹.

We used these six DGVMs because they explicitly reported tree mortality. To be comparable with forest plot data, we used the components of tree mortality flux without disturbance, such as fires (Supplementary Table 2). The historic mortality flux was averaged over 1961-2014 forced by CRU-NCEP v5 climate data and the projected NPP and HR were averaged over 2015-2099 forced by bias-corrected IPSL-CM5A-LR RCP 8.5 simulated climate data. Because the output of HR is reported at grid scale and it is not able to be decomposed into tree-level HR and non-tree level HR, we used grid scale NPP to estimate forest carbon sink, quantified as the difference of grid-level NPP and grid-level HR. To be consistent with forest inventory data, outputs of historical mortality at tree-level were used to compare with observations. All outputs were resampled and analyzed at 0.5 degree.

Geospatial modeling and environmental drivers

We used the Random Forest machine learning algorithm⁴⁶ (see Supplementary Methods) with the derived 58 environmental covariates including climate, vegetation and soil conditions (Supplementary Table 1) to extrapolate these relationships between LOSS and environmental conditions across continents and generate the first spatially-explicit and quantitative map of LOSS at the continental scale. The 58 environmental covariates were derived based on the georeferenced coordinates of forest plot data ($n = 814$) aggregated at 0.25 degree. 10-fold cross validation was used to evaluate the strength of prediction and the best model having high coefficient of determination and low standard deviation in the mean cross-validation were used to generate the map of LOSS. The standard error sharply decreased with increasing sample size across all vegetation biomes and the analysis showed that an efficient prediction required a large sample size ($n > 400$) (Supplementary Fig. 18a). Random Forest was able to predict the variation in LOSS with

high predictive accuracy ($R^2 = 0.48$ in 10-fold cross validation; $R^2 = 0.93$ in final model; Supplementary Fig. 18b). Two types of bootstrapping were used to evaluate the uncertainty (standard deviation as a fraction of the mean predicted value) in the map of LOSS. One was based on a stratified bootstrapping (100 iterations) procedure⁴⁷, which was the sampled biomes of each point feature (LOSS) with the total number collection of LOSS points to avoid biases. The second bootstrap was based on randomly sampling 90% with replacement (10 iterations) to account for the biases from an unbalanced sample distribution. In both of the two types of bootstrapping, ensemble mean and 95% confidence intervals of LOSS were computed by grid. Rasters of tree cover, human footprint index, percentage of annual burn area and managed land cover (see Supplementary Methods) were used as mask to define the natural forest areas across continents.

To examine the environmental controls of LOSS, we chose the top drivers which include climate conditions (mean annual temperature, MAT; aridity index; precipitation seasonality), vegetation properties (tree density, basal area, competition index, wood density and biome type) and soil properties (soil organic carbon, SOC; soil N; soil hydraulic conductivity, Ks; cation exchange capacity; clay content; and pH) (see Supplementary Methods). Competition index was defined as the degree of crowding within stocked areas and quantified as a function of tree density and the quadratic mean diameter in centimeters⁴⁰. These variables were examined to avoid multicollinearity using a matrix of pairwise correlations to remove any variable with high correlations ($R > 0.7$) with other predictor variables⁴⁸ and variation inflation factor ($VIF < 4$). The Random Forest machine learning algorithm was then used to determine the importance of each predictor variable⁴⁹. Mean decrease in accuracy (%IncMSE) were reported and the variables with greater values of %IncMSE are more important in influencing LOSS. To account for each plot as

a random effect, we also used a linear mixed model to examine the dominant factors on LOSS across continents and in tropical regions vs non-tropical regions (see Supplementary Methods).

Constraining projected forest carbon sinks

We first attempted to use the classic emergent constraint approach to constrain the projected (2015-2099) NPP, HR and forest carbon sink. This was achieved by least-squares linear regressions²² between historical (1961-2014) LOSS and projected grid-level NPP and grid-level HR across DGVMs:

$$NPP_{pro} \text{ or } HR_{pro} (i) = a \times LOSS_{his}(i) + b \quad (1)$$

where *pro* is ‘projected’, *his* is ‘historic’, *i* is the index of model, and *a* and *b* are coefficients.

NPP and HR were aggregated as sum within each continent and LOSS were aggregated as average across forest plot sites or within each continent (Supplementary Fig. 4). In details, we used LOSS at forest-plot local scale to constrain the projected NPP and HR at continental scale in DGVMs and found its limited feasibility in the classic emergent constraint approach because of spatial mismatch between data and model²⁶. This motivated us to generate a spatially explicit map of LOSS derived from the machine learning algorithm (see Methods) to constrain the projected NPP and HR at continental scale in DGVMs.

Alternatively, we used the historical (1961-2014) LOSS as the predictor to train a random forest model and constrain the projected (2015-2099) NPP, HR and forest carbon sink:

$$NPP_{pro} \text{ or } HR_{pro} (i) = f (LOSS_{his}(i)) \quad (2)$$

$$NPP_{pro} \text{ or } HR_{pro} (i) = f (LOSS_{his}(i), MAT_{pro}, MAP_{pro}) \quad (3)$$

where equation 2 is the scenario without the projected climate effects and equation 3 is the scenario with the projected climate effects. Climate effects were incorporated to account for their potential influence on the couplings or relationships between historical LOSS and projected NPP

and HR across environmental (climate) gradients. CO₂ was not included in eqn (3) because of the minimal spatial heterogeneity on annual timescales. We clarified that all of these variables were averages over the historical or future periods and our study thus focused on the spatial patterns of variables of interests in the quasi-steady state.

We stress that the random forest model allowed for inclusion of all grid values of variables of interests and consideration of non-linearity. This was achieved with 1000 trees and 10 maximum tree depth and 80% of the data for training purpose and the rest 20% for validation. With this approach, we surrogated the predictive relationship between the historical (1961-2014) LOSS and the projected (2015-2099) NPP and HR from a complex but physically driven DGVMs with an empirical machine learning model (random forest). Then, we feed the observed maps of LOSS derived from forest plots datasets into the trained random forest model to assess the impacts of historic LOSS on the projected NPP and HR. In this way, we were able to generate the spatially explicit²⁹ constraint of the projected NPP and HR, which was subsequently aggregated as sum within each continent to assess the forest carbon sink at the continental scale. The mechanistic underpinning justifying the linear (emergent) and non-linear (machine learning) constrain was the theory of faster growth and higher mortality^{12,13,43} and thus coupling of growth, mortality and respiration averaged over long term and broad spatial scales^{18,19}.

Table 1 Modeled projected (2015-2099, Pg C y⁻¹) grid-level NPP and grid-level HR before and after constrain across continent in six DGVMs

Continent	Before constrain		After constrain	
	Mean	SD	Mean	SD
North America	6.1	3.2	6.2	2.9
South America	9.0	3.2	8.0	2.7
Africa	2.4	0.9	2.3	0.8
Asia & Australia	2.3	0.9	2.0	0.7
North America	6.2	3.2	6.2	2.9
South America	9.0	3.2	8.0	2.7
Africa	2.4	0.9	2.3	0.8
Asia & Australia	2.3	0.9	2.0	0.7

Note: the scenario is without accounting for the effects projected climate – precipitation and temperature

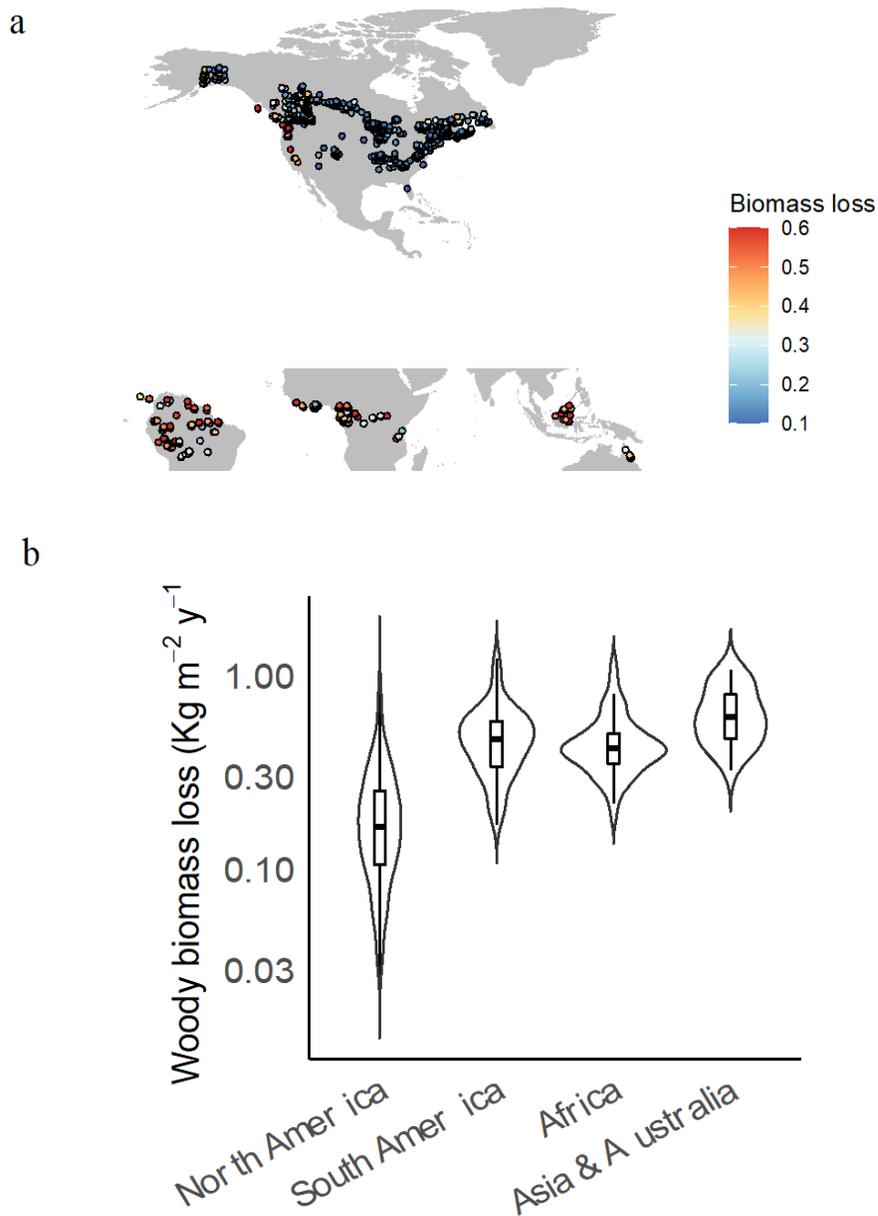


Figure 1. Map of sample locations and biomass loss to mortality (LOSS) data. a, Sampling sites. A total of 2676 samples were collected and aggregated into 814 grids at 0.25 degree that were used for geospatial modelling. b, The median and interquartile range of LOSS across continents – North America, South America, Africa, and Asia & Australia.

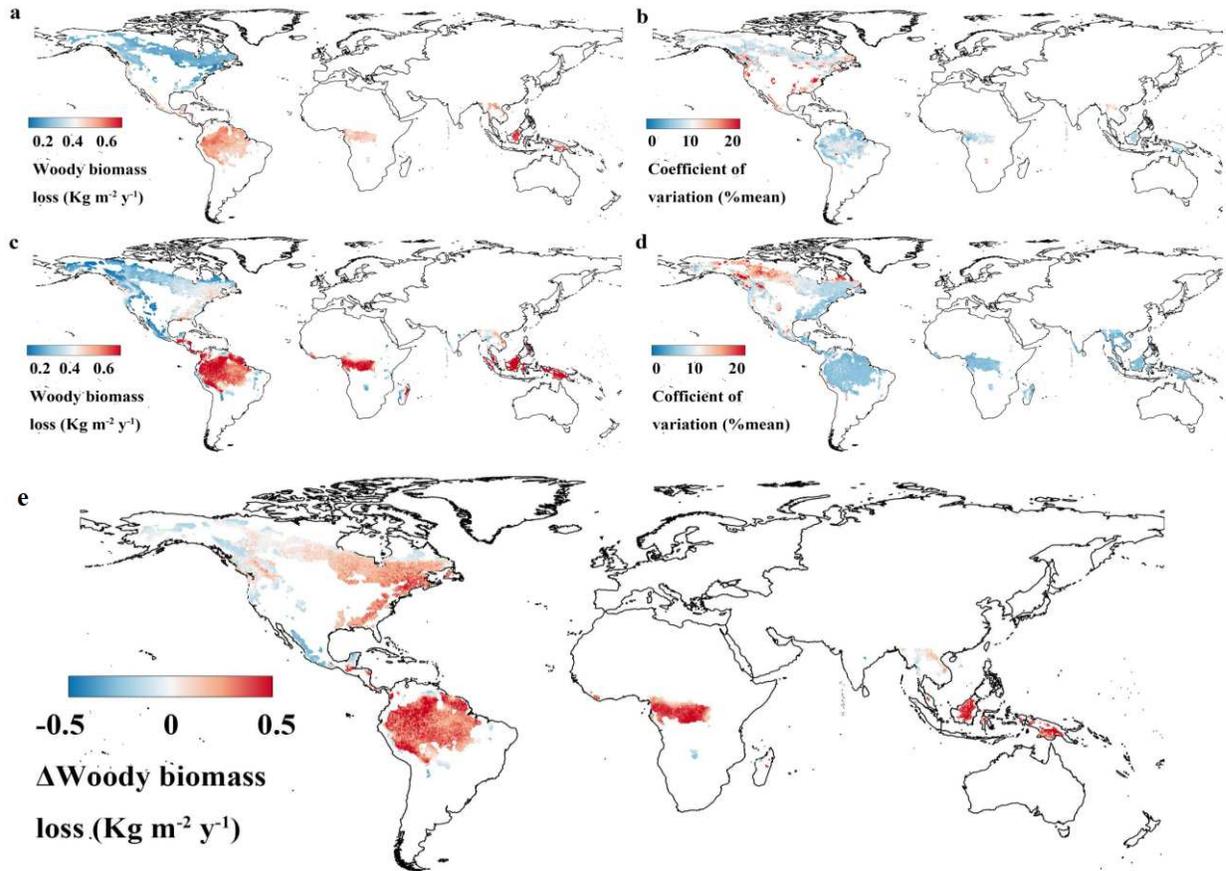


Figure 2. Map of biomass loss to mortality (LOSS) and its uncertainty across continents. (a, b) Ensemble mean of LOSS (a) and its uncertainty (coefficient of variation, b) across continents at 0.25 degree derived from forest plot data using the bootstrapped (10 iterations) approach by randomly sampling 90% plots with replacement. (c, d) Ensemble mean of LOSS (c) and its uncertainty (coefficient of variation, d) across continents at 0.5 degree derived from six dynamic vegetation models (DGVMs, ORCHIDEE, CABLE-POP, JULES, LPJ-GUESS, LPJmL, and SEIB-DGVM). Coefficient of variation was quantified as standard deviation divided by the mean predicted value as a measure of prediction accuracy. (e) The difference of LOSS between ensemble mean of DGVMs and ensemble mean of LOSS derived from forest plots data across continents at 0.5 degree, quantified as difference between Fig. 2c and Fig. 2a, whereby LOSS in Fig. 2a is resampled at 0.5 degree.

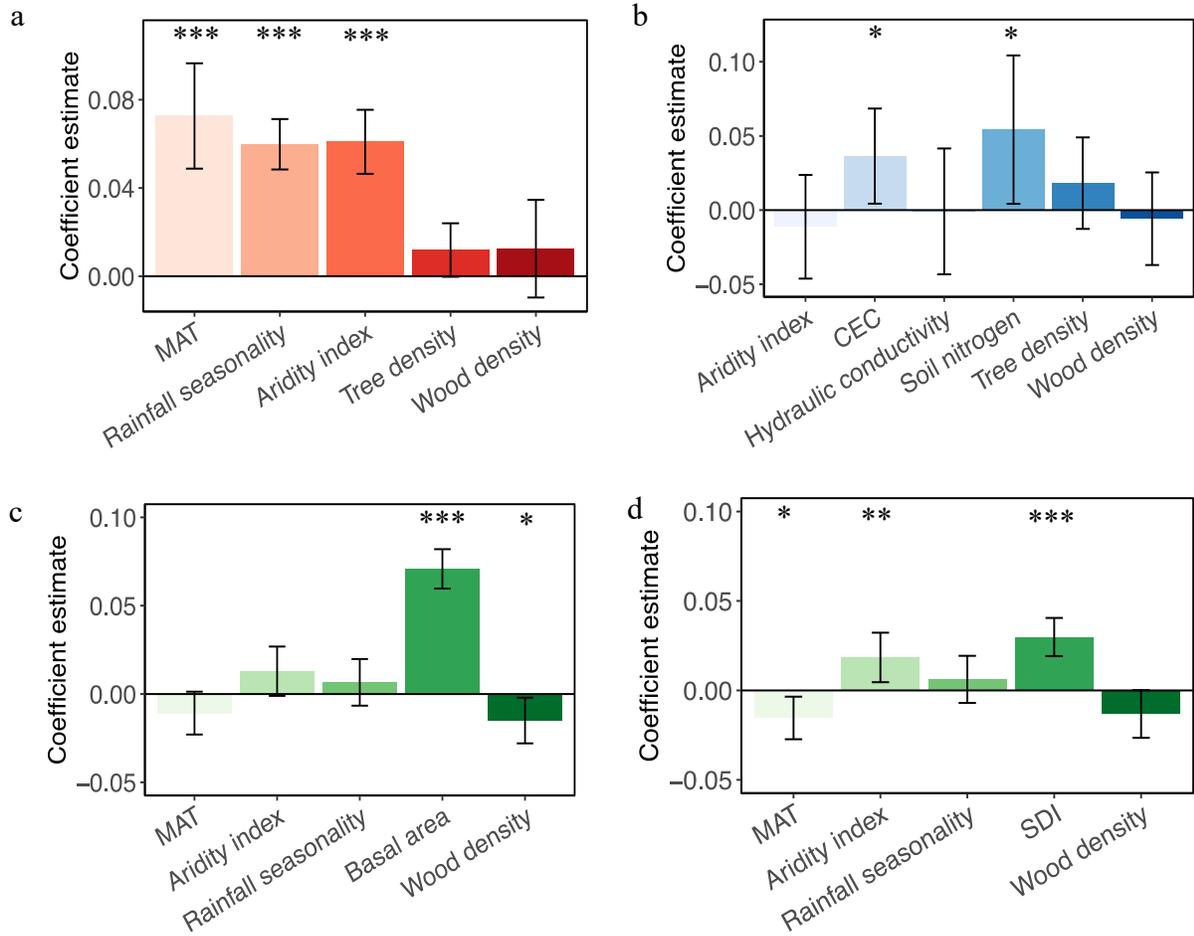


Figure 3. Standardized response coefficients (mean \pm 95% CIs) between LOSS and dominant environmental drivers across continents (a), tropical regions (b) vs non-tropical regions (c, d) quantified by linear mixed model which account for each plot as a random effect. Competition index (SDI) was defined as the degree of crowding within stocked areas and quantified as a function of tree density and the quadratic mean diameter in centimeters. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

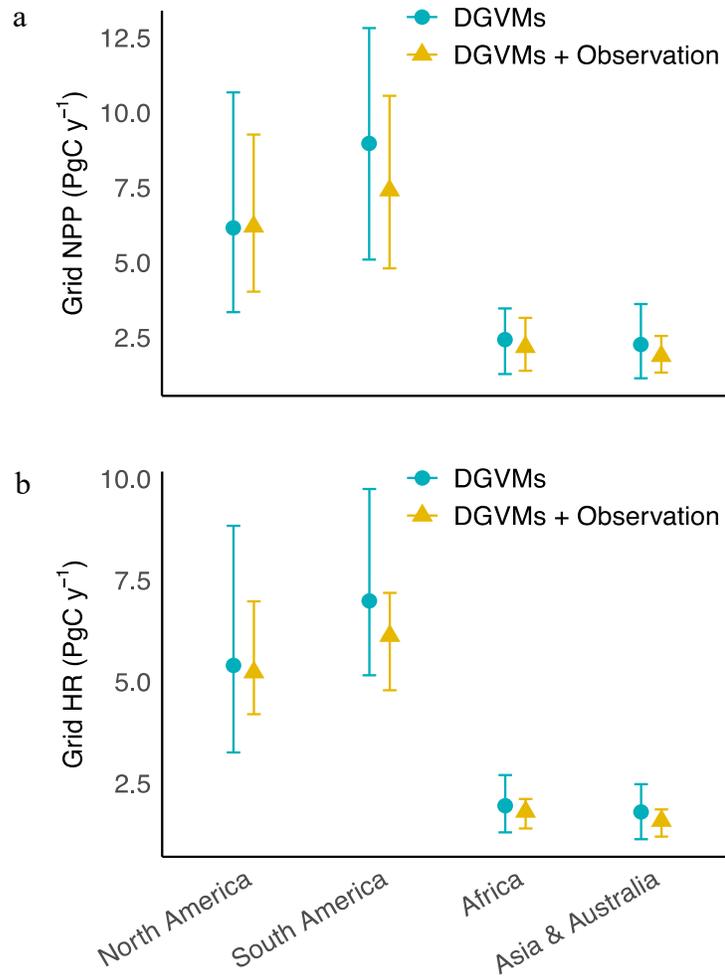


Figure 4. Projected grid-level NPP and grid heterotrophic respiration (HR) across continents. (a, b) Projected (2015-2099) grid-level NPP (a) and grid-level HR (b) across continents quantified by six dynamic vegetation models - DGVMs (ORCHIDEE, CABLE-POP, JULES, LPJ-GUESS, LPJmL, and SEIB-DGVM). The y axes are the minimum, mean, and maximum values in six DGVMs. DGVMs refers to the scenario before constrain and DGVMs refers to the scenario after constrain without climate predictors. The constrain was achieved by using the observational map of biomass loss to mortality (LOSS) derived from forest plots data to feed into the trained ML (random forest) model to assess the impacts of historic LOSS on the projected NPP and HR in each DGVM.

References

1. Bonan, G. B. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* (80-.). **320**, 1444–1449 (2008).
2. Tagesson, T. *et al.* Recent divergence in the contributions of tropical and boreal forests to the terrestrial carbon sink. *Nat. Ecol. Evol.* (2020) doi:10.1038/s41559-019-1090-0.
3. Schimel, D., Stephens, B. B. & Fisher, J. B. Effect of increasing CO₂ on the terrestrial carbon cycle. *Proc. Natl. Acad. Sci. U. S. A.* (2015) doi:10.1073/pnas.1407302112.
4. Norby, R. J. *et al.* Forest response to elevated CO₂ is conserved across a broad range of productivity. *Proc. Natl. Acad. Sci. U. S. A.* **102**, (2005).
5. Piao, S. *et al.* Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO₂ trends. *Glob. Chang. Biol.* **19**, (2013).
6. Bond-Lamberty, B., Bailey, V. L., Chen, M., Gough, C. M. & Vargas, R. Globally rising soil heterotrophic respiration over recent decades. *Nature* (2018) doi:10.1038/s41586-018-0358-x.
7. Naidu, D. G. T. & Bagchi, S. Greening of the earth does not compensate for rising soil heterotrophic respiration under climate change. *Glob. Chang. Biol.* (2021) doi:10.1111/gcb.15531.
8. Friedlingstein, P. *et al.* Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. *J. Clim.* **19**, 3337–3353 (2006).
9. Heimann, M. & Reichstein, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* **451**, 289–292 (2008).
10. Brienen, R. J. W. *et al.* Long-term decline of the Amazon carbon sink. *Nature* **519**, 344–348 (2015).

11. Korner, C. Through enhanced tree dynamics carbon dioxide enrichment may cause tropical forests to lose carbon. *Philos. Trans. R. Soc. B Biol. Sci.* **359**, 493–498 (2004).
12. Büntgen, U. *et al.* Limited capacity of tree growth to mitigate the global greenhouse effect under predicted warming. *Nat. Commun.* **10**, 2171 (2019).
13. Yu, K. *et al.* Pervasive decreases in living vegetation carbon turnover time across forest climate zones. *Proc. Natl. Acad. Sci. U. S. A.* (2019) doi:10.1073/pnas.1821387116.
14. Anderegg, W. R. L., Kane, J. M. & Anderegg, L. D. L. Consequences of widespread tree mortality triggered by drought and temperature stress. *Nature Climate Change* vol. 3 30–36 (2013).
15. McDowell, N. G. *et al.* Pervasive shifts in forest dynamics in a changing world. *Science* (80-.). **368**, (2020).
16. Schimel, D., Stephens, B. B. & Fisher, J. B. Effect of increasing CO₂ on the terrestrial carbon cycle. *Proc. Natl. Acad. Sci.* **112**, 436–441 (2015).
17. Thomas, R. Q., Canham, C. D., Weathers, K. C. & Goodale, C. L. Increased tree carbon storage in response to nitrogen deposition in the US. *Nat. Geosci.* **3**, (2010).
18. Baldocchi, D. TURNER REVIEW No. 15. ‘Breathing’ of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems. *Australian Journal of Botany* vol. 56 (2008).
19. Ciais, P., Sabine, C. & Bala, G. Chapter 6: Carbon and other biogeochemical cycles. *Clim. Chang. 2013 Phys. Sci. Basis* (2014).
20. Moore, D. J. P. *et al.* Persistent reduced ecosystem respiration after insect disturbance in high elevation forests. *Ecol. Lett.* **16**, (2013).
21. Yang, H. *et al.* Coarse woody debris are buffering mortality-induced carbon losses to the

- atmosphere in tropical forests. *Environmental Research Letters* vol. 16 (2021).
22. Cox, P. M. *et al.* Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature* **494**, (2013).
 23. Wenzel, S., Cox, P. M., Eyring, V. & Friedlingstein, P. Projected land photosynthesis constrained by changes in the seasonal cycle of atmospheric CO₂. *Nature* **538**, (2016).
 24. Winkler, A. J., Myneni, R. B., Alexandrov, G. A. & Brovkin, V. Earth system models underestimate carbon fixation by plants in the high latitudes. *Nat. Commun.* **10**, (2019).
 25. Liu, Y. *et al.* Field-experiment constraints on the enhancement of the terrestrial carbon sink by CO₂ fertilization. *Nat. Geosci.* **12**, (2019).
 26. Winkler, A. J., Myneni, R. B. & Brovkin, V. Investigating the applicability of emergent constraints. *Earth Syst. Dyn.* **10**, (2019).
 27. Hall, A., Cox, P., Huntingford, C. & Klein, S. Progressing emergent constraints on future climate change. *Nature Climate Change* vol. 9 (2019).
 28. Mystakidis, S., Davin, E. L., Gruber, N. & Seneviratne, S. I. Constraining future terrestrial carbon cycle projections using observation-based water and carbon flux estimates. *Glob. Chang. Biol.* **22**, (2016).
 29. Schlund, M. *et al.* Constraining Uncertainty in Projected Gross Primary Production With Machine Learning. *J. Geophys. Res. Biogeosciences* **125**, (2020).
 30. Zhu, K., Zhang, J., Niu, S., Chu, C. & Luo, Y. Limits to growth of forest biomass carbon sink under climate change. *Nature Communications* (2018).
 31. Pugh, T. A. M. *et al.* Understanding the uncertainty in global forest carbon turnover. *Biogeosciences* **17**, (2020).
 32. Sullivan, M. J. P. *et al.* Long-term thermal sensitivity of earth's tropical forests. *Science*

- (80-). **368**, (2020).
33. Peng, C. *et al.* A drought-induced pervasive increase in tree mortality across Canada's boreal forests. *Nat. Clim. Chang.* **1**, 467–471 (2011).
 34. Van Mantgem, P. J. *et al.* Widespread increase of tree mortality rates in the Western United States. *Science (80-)*. **323**, 521–524 (2009).
 35. Pugh, T. A. M. *et al.* Role of forest regrowth in global carbon sink dynamics. *Proc. Natl. Acad. Sci.* **116**, 4382–4387 (2019).
 36. Greenwood, D. L. & Weisberg, P. J. Density-dependent tree mortality in pinyon-juniper woodlands. *For. Ecol. Manage.* **255**, (2008).
 37. Chao, K. J. *et al.* Growth and wood density predict tree mortality in Amazon forests. *J. Ecol.* **96**, (2008).
 38. Rozendaal, D. M. A. *et al.* Competition influences tree growth, but not mortality, across environmental gradients in Amazonia and tropical Africa. *Ecology* **101**, (2020).
 39. Phillips, O. L. & Gentry, A. H. Increasing Turnover Through Time in Tropical Forests. *Science (80-)*. **263**, 954–958 (1994).
 40. Zhang, J., Huang, S. & He, F. Half-century evidence from western Canada shows forest dynamics are primarily driven by competition followed by climate. *Proc. Natl. Acad. Sci.* **112**, 4009–4014 (2015).
 41. Luo, Y. & Chen, H. Y. H. Climate change-associated tree mortality increases without decreasing water availability. *Ecology Letters* vol. 18 1207–1215 (2015).
 42. Neumann, M., Mues, V., Moreno, A., Hasenauer, H. & Seidl, R. Climate variability drives recent tree mortality in Europe. *Glob. Chang. Biol.* **23**, 4788–4797 (2017).
 43. Körner, C. A matter of tree longevity. *Science (80-)*. **355**, 130–131 (2017).

44. Fisher, R. A. *et al.* Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology* vol. 24 (2018).
45. Huang, Y. *et al.* A global map of root biomass across the world's forests. *bioRxiv* (2020) doi:10.1101/2020.01.14.906883.
46. Crowther, T. W. *et al.* Mapping tree density at a global scale. *Nature* **525**, 201–205 (2015).
47. van den Hoogen, J. *et al.* Soil nematode abundance and functional group composition at a global scale. *Nature* (2019) doi:10.1038/s41586-019-1418-6.
48. Anderegg, L. D. L., Anderegg, W. R. L., Abatzoglou, J., Hausladen, A. M. & Berry, J. A. Drought characteristics' role in widespread aspen forest mortality across Colorado, USA. *Glob. Chang. Biol.* **19**, 1526–1537 (2013).
49. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixFieldbasedtreemortalityobservationsconstrainmodelprojectedforestcarbonsinksacrosscontinents.pdf](#)
- [SupplementaryTable1.xlsx](#)