# Quantifying the effect of forest age in annual net forest carbon balance

Simon Besnard, Nuno Carvalhais, M Altaf Arain, Andrew Black, Sytze de Bruin, Nina Buchmann, Alessandro Cescatti, Jiquan Chen, Jan G P W Clevers, Ankur R Desai, et al.

## ▶ To cite this version:

Simon Besnard, Nuno Carvalhais, M Altaf Arain, Andrew Black, Sytze de Bruin, et al.. Quantifying the effect of forest age in annual net forest carbon balance. Environmental Research Letters, 2018, 13 (12), 10.1088/1748-9326/aaeaeb . hal-02621050

## HAL Id: hal-02621050
## https://hal.inrae.fr/hal-02621050v1

Submitted on 26 May 2020

Environmental Research Letters

LETTER • OPEN ACCESS

# Quantifying the effect of forest age in annual net forest carbon balance

To cite this article: Simon Besnard *et al* 2018 *Environ. Res. Lett.* **13** 124018

IOP Publishing Environ. Res. Lett. 13 (2018) 124018 https://doi.org/10.1088/1748-9326/aaeaeb

Environmental Research Letters

**OPEN ACCESS**

RECEIVED
6 July 2018

REVISED
21 October 2018

ACCEPTED FOR PUBLICATION
24 October 2018

PUBLISHED
7 December 2018

LETTER

# Quantifying the effect of forest age in annual net forest carbon balance

**Simon Besnard[1,2], Nuno Carvalhais[1,3], M Altaf Arain[4], Andrew Black[5], Sytze de Bruin[2], Nina Buchmann[6], Alessandro Cescatti[7], Jiquan Chen[8], Jan G P W Clevers[2], Ankur R Desai[9], Christopher M Gough[10], Katerina Havrankova[11], Martin Herold[2], Lukas Hörtnagl[6], Martin Jung[1], Alexander Knohl[12], Bart Kruijt[13], Lenka Krupkova[11], Beverly E Law[14], Anders Lindroth[15], Asko Noormets[16,17], Olivier Roupsard[18,19], Rainer Steinbrecher[20], Andrej Varlagin[21], Caroline Vincke[22] and Markus Reichstein[1]**

[1] Department for Biogeochemical Integration, Max-Planck-Institute for Biogeochemistry, Hans-Knöll-Str. 10, D-07745 Jena, Germany
[2] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University and Research, PO Box 47, 6700 AA Wageningen, The Netherlands
[3] CENSE, Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, Portugal
[4] School of Geography and Earth Sciences and McMaster Center For Climate Change, McMaster University, Hamilton, Ontario L8S 4K1, Canada
[5] Faculty of Land and Food Systems, 2357 Main Mall, University of British Columbia, Vancouver V6T 1Z4, Canada
[6] ETH Zurich, Department of Environmental Systems Sciences, LFW C56, Universitätstrasse 2, 8092 Zurich, Switzerland
[7] European Commission, Joint Research Centre, Directorate for Sustainable Resources, Ispra, Italy
[8] Department of Geography, Environment and Spacial Sciences, Center for Global Change and Earth Observations, Michigan State University, East Lansing, MI 48823, United States of America
[9] Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, 1225 W Dayton St, Madison, WI 53711, United States of America
[10] Department of Biology, University of Virginia Commonwealth, 1000 W. Cary Street, Richmond, VA 23284-2012, United States of America
[11] Department of Matters and Energy Fluxes, Global Change Research Institute CAS, Belidla 4a, 60300 Brno, Czech Republic
[12] Bioclimatology, University of Göttingen, Büsgenweg 2, D-37077 Göttingen, Germany
[13] Wageningen Environmental Research (Alterra), PO Box 47, 6700 AA Wageningen, The Netherlands
[14] 328 Richardson Hall, College of Forestry, Oregon State University, Corvallis, OR 97331-5752, United States of America
[15] Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, SE-22362 Lund, Sweden
[16] Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, United States of America
[17] Department of Ecosystem Science and Management, Texas A&M University, College Station, TX 77843, United States of America
[18] CIRAD, UMR Eco&Sols, LMI IESOL, B.P. 1386 CP 18524, Dakar, Senegal
[19] Eco&Sols, University Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, Montpellier, France
[20] Department of Atmospheric Environmental Research, Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Kreuzeckbahnstr. 19, D-82467 Garmisch-Partenkirchen, Germany
[21] A.N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, 119071, Leninsky pr.33, Moscow, Russia
[22] Université catholique de Louvain, Earth and Life Institute, Environmental Sciences, Croix du sud 2 L7.05.09, B-1348 Louvain-la-Neuve, Belgium

**E-mail:** sbesnard@bgc-jena.mpg.de

**Keywords:** carbon cycle, eddy covariance, net ecosystem production, empirical modelling, forest age, climate, soil properties

Supplementary material for this article is available online

## Abstract

Forests dominate carbon (C) exchanges between the terrestrial biosphere and the atmosphere on land. In the long term, the net carbon flux between forests and the atmosphere has been significantly impacted by changes in forest cover area and structure due to ecological disturbances and management activities. Current empirical approaches for estimating net ecosystem productivity (NEP) rarely consider forest age as a predictor, which represents variation in physiological processes that can respond differently to environmental drivers, and regrowth following disturbance. Here, we conduct an observational synthesis to empirically determine to what extent climate, soil properties, nitrogen deposition, forest age and management influence the spatial and interannual variability of forest NEP across 126 forest eddy-covariance flux sites worldwide. The empirical models explained up to 62% and 71% of spatio-temporal and across-site variability of annual NEP, respectively. An investigation of model structures revealed that forest age was a dominant factor of NEP spatio-temporal variability in both space and time at the global scale as compared to abiotic factors, such as nutrient availability, soil characteristics and climate. These findings emphasize the importance of forest age in quantifying spatio-temporal variation in NEP using empirical approaches.

## 1. Introduction

Forests cover about 30% of the Earth's terrestrial surface and store around 90% of terrestrial vegetation carbon (C) (Canadell *et al* 2000, Gower 2003, Le Quéré *et al* 2018), indicating their fundamental role in terrestrial C dynamics (Bonan 2008, Beer *et al* 2010, Pan *et al* 2011, Hicke *et al* 2012, Carvalhais *et al* 2014). However, the functioning of forest ecosystems is likely to be altered by changing climate (Ciais *et al* 2005, Xiao *et al* 2009, Zhao and Running 2010, Reichstein *et al* 2013), ecological disturbances (Chambers *et al* 2007, Bowman *et al* 2009, Amiro *et al* 2010) and management (Noormets *et al* 2015, Naudts *et al* 2016). Therefore, it is important to characterize current and future forest net ecosystem production (NEP) for regional to country-level assessments, and to evaluate mitigation strategies that minimize carbon dioxide ($CO_2$) emissions to the atmosphere (Becknell *et al* 2015, Trumbore *et al* 2015, Law *et al* 2018).

The overall NEP dynamic at a given site emerges from combined responses to factors that control both gross primary productivity (GPP) and ecosystem respiration (ER) (NEP = GPP-ER). At the ecosystem level, the forest NEP patterns following stand-replacing disturbance are mostly controlled by the time-varying dominance between autotrophic and catabolic processes. After disturbance, heterotrophic respiration (Rh) generally tends to increase because of an aboveground biomass transfer to the litter and soil organic matter C pools (Law *et al* 2003, Kurz *et al* 2008, Harmon *et al* 2011, Noormets *et al* 2012, Lindauer *et al* 2014, Paul-Limoges *et al* 2015), while GPP collapses due to an instantaneous reduction in leaf area, resulting in a net release of $CO_2$ to the atmosphere. On the one hand, a shift from C source to C sink occurs as canopy development supports GPP and net C accumulation in plants increase. On the other hand, Rh and ER decline due to a reduction in litterfall and substrate availability through decomposition. The resulting imbalance between GPP and ER persists until vegetation and soil C pools increase up to the point when ER comes into equilibrium with GPP (Schwarz *et al* 2004, Lindroth *et al* 2008, Luyssaert *et al* 2008, Tang *et al* 2014).

Several approaches used for assessing forest NEP include micro-meteorological and biometric techniques, process-based models, and/or satellite data. However, annual regional C stock assessments that account for age-related physiology, regrowth, and soil processes following disturbance are challenging due to lack of information in disturbance history or management practices (Zscheischler *et al* 2017). Flux sites provide the annual net uptake of $CO_2$ from the atmosphere (i.e. NEP) that can be used to calibrate empirical models for mapping annual NEP at regional scales. However, current empirical uspcaling exercises (Jung *et al* 2011, Tramontana *et al* 2016) do not directly include proxies that allow the dynamics of C fluxes with age to emerge, therefore it is not clear how well the aforementioned data-driven models captured such dynamics. Thereby, empirical estimation of annual NEP that explicitly accounts for disturbance and forest age effects are of relevance for regional C stock studies (Ciais *et al* 2014).

Despite the recognized effects of forest age in controlling spatial and interannual variability of NEP, there is still debate about the quantitative role of forest age in the empirical annual forest C estimates. In fact, the most recent observation-based synthesis studies tackling NEP spatio-temporal variability and its drivers reached diverging conclusions on the importance of forest age. While some authors have shown that forest age is a key factor controlling forest C balance (Chen *et al* 2002, 2003, Coursolle *et al* 2012, Yu *et al* 2014, Gao *et al* 2016), others have indicated that spatial and interannual variability of NEP is mainly controlled by nutrients availability and soil properties (Bhatti *et al* 2002, Janssens *et al* 2010, Vicca *et al* 2012, Fernández-Martínez *et al* 2014) or climate conditions (Thornton *et al* 2002, Amiro *et al* 2006, Coursolle *et al* 2006), although several authors report that the C budget in forest ecosystems is less sensitive to climatic conditions than expected in certain regions (Law and Falge 2002, Reichstein *et al* 2007, Yi *et al* 2010).

Given the fundamental understanding of the role of forest age in NEP and the contrasting results from previous meta-analyses, we revisited the importance of forest age to the spatial and temporal variability in NEP based on a more up-to-date, larger, and higher quality eddy-covariance (EC) dataset including 126 forest ecosystem sites. We further expanded previous observation-based syntheses by exploring nonlinear empirical model formulations to incorporate forest developmental stage and environmental factors for calculating realistic NEP spatio-temporal variability. Such a model can eventually be used to estimate NEP at a global scale and infer likely limits to NEP variation and the future forest C sink as forests age.

## 2. Methods

### 2.1. Datasets

We used a global dataset of 126 EC forest sites ranging from 0 to ⩾300 year-old stands (table S1 and figure S1 is available online at stacks.iop.org/ERL/13/124018/mmedia). The sites were part of both version 2 of the LaThuile FLUXNET and the FLUXNET 2015 datasets (https://fluxnet.fluxdata.org) of the FLUXNET network (Baldocchi *et al* 2001, Baldocchi 2008). Five vegetation types were considered in the study, including 76 evergreen forests, 27 deciduous forests, 11 mixed forests, seven shrublands, three savannas and two wetlands.

We aggregated daily NEP, GPP, ER and the associated uncertainties into annual sums (i.e. site-years) and computed an annual-average from all available years per site (i.e. site-average) (see supplementary information for details on EC data processing). One relevant aspect to consider is that the observation-derived GPP is determined via the measured night-time NEP (Reichstein *et al* 2005). This challenges the statistical independence of both variables, therefore risking a spurious correlation between GPP and NEP at annual scales (see supplementary materials). To avoid any spurious relationship between NEP and GPP, we used a proxy for GPP, i.e. GPP′, which was determined as the ratio between latent heat flux (LE) and the square root of vapor pressure deficit $\left(\mathrm{GPP}' = \frac{\mathrm{LE}}{\sqrt{\mathrm{VPD}}}\right)$, (see supplementary materials for more details and figure S3) as based on physiological principles of the coupling between the C and water cycles at the leaf level (Chen *et al* 2002, Katul *et al* 2009, 2009). While LAI and FPAR could have been used as proxies for GPP, GPP′ is the integrated response of phenology and physiology, therefore a direct metric of primary productivity, while the former are mostly phenology driven.

We considered forest age as the time since forest establishment or as the time since the occurrence of the last stand-replacing disturbance (see supplementary information for more details on the definition of forest age). Sites were selected based on the availability of information about forest age, disturbance history and management that would allow for an appropriate definition of a meaningful site stand age. These included a range of young and old growth forest sites (figure S2) that were established after complete, nearly complete or substantial removal of forest vegetation (e.g. harvest, fire, wind-throw, insect outbreaks), followed by reforestation/succession/afforestation activities within the flux tower footprint. We did not consider sites with ambiguous historical information or those that had experienced only low to partial intensity disturbances, which would not allow the determination of whole stand forest age. For uneven-aged stands, we followed Spies and Franklin (1991) and estimated the age of a stand as the age of the largest 10% of trees. Undisturbed old-growth forests where age information was available were also included. The information for each site was obtained from the literature, provided by the site principal investigators or from the Biological, Ancillary, Disturbance and Metadata database (table S1). In general, the wide span in stand age among sites and the multi-year record of observations per site permit evaluating the effect of age on both the mean and the interannual variability in NEP.

In addition to the C and water fluxes, we also obtained the following variables as statistical covariates for model development for each EC site:

(i) local microclimatic variables from *in situ* observations (i.e. mean annual air temperature (MAT), total annual precipitation (P), and global radiation (Rg));

(ii) information on nutrient availability (NA) divided into three classes: low NA ($n = 67$ sites); medium NA ($n = 41$ sites); and high NA ($n = 18$ sites) (based on (Fernández-Martínez *et al* 2014) study and/or expert knowledge, tables S1);

(iii) additional information on soil texture up to 2 m depth from the SoilGrids1km dataset (Hengl *et al* 2014);

(iv) information on forest management based on Campioli *et al* (2015), Luyssaert *et al* (2008) datasets, and indications from the PIs (managed forests $n = 44$ sites; and unmanaged forests $n = 81$ sites). Managed sites were dominated by human activity while unmanaged sites were undisturbed or experienced low human impacts;

(v) gridded monthly temperature and precipitation observations from the Climate Research Unit (CRU, http://cru.uea.ac.uk) (Harris *et al* 2014) to determine long-term linear climate trends and anomalies; and

(vi) local total atmospheric nitrogen deposition ($\mathrm{N}_{deposition}$) from *in situ* observations. We collected $\mathrm{N}_{deposition}$ estimates from the gridded emissions dataset (Wang *et al* 2017) at 0.5° × 0.5° resolution when they were not available at site level.

### 2.2. NEP model development

The development of an NEP statistical model principally aimed to provide a data-driven estimate of the several factors that control the temporal and spatial variability in NEP, and further to estimate the relative contribution of the different predictive variables—especially age and GPP′—to variation in NEP. To do so, we used the aforementioned statistical covariates (i.e. GPP′, microclimatic data, nutrient availability, soil texture, $\mathrm{N}_{deposition}$, forest age, and forest management) to train and evaluate the ability of a random forest (RF; (Breiman 2001)) algorithm to explain NEP variability.

GPP′ and ER are co-determinants of NEP (figure S3), therefore both were initially compared to represent site level NEP. However, given the higher correlation of NEP with GPP′ compared to the relationship with ER, and the aforementioned statistical dependence between

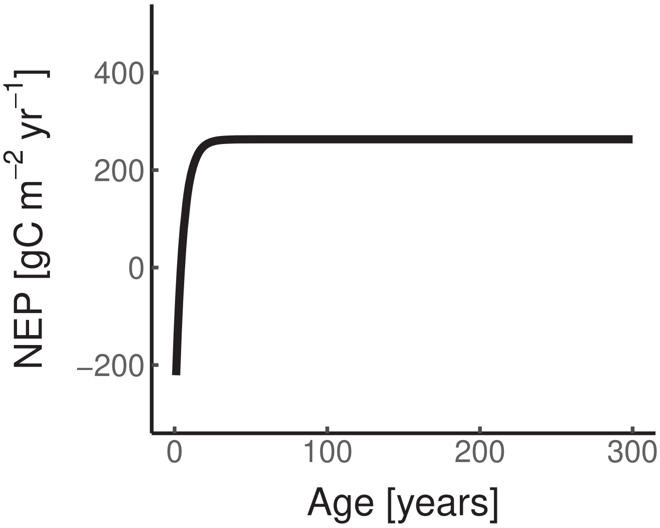

**Figure 1.** NEP-age dynamics for the Amiro *et al* (2010) stand age model. $a = 587.74$ gC m$^{-2}$ y$^{-1}$, b $= -0.19$, and $e = -324.54$ gC m$^{-2}$ y$^{-1}$.

NEP and the gross C fluxes (i.e. ER and GPP), we chose to discard ER in the statistical analysis.

The role of forest age as an explanatory variable of NEP was additionally evaluated with a published nonlinear model (from now on identified as f(age)) to represent the NEP-age relationship (equation (1) and figure 1)

$$\mathrm{NEP} = e + a \cdot (1 - \exp^{b \cdot Age}) \quad (1)$$

adapted from Amiro *et al* 2010, where the model parameters *a*, *b* and *e* (offset) were estimated.

Although the NEP-age model (equation (1)) was originally developed to represent the temporal patterns of annual GPP-to-ER ratio in forest chronosequences, here we used it to describe the dependency of spatio-temporal NEP variability on forest age. The selection of the Amiro *et al* (2010) model to describe NEP spatial temporal dynamics assumes that the age-related patterns in GPP/ER are qualitatively similar to NEP, and is supported by a comparison to two other different empirical models presented in the literature (Coursolle *et al* 2012, Tang *et al* 2014). These two models were also tested but showed poorer model performance than the Amiro *et al* (2010) model in the multivariate analysis (table S2).

### 2.3. Experiment design

*Estimation of forest age model parameters*: The model parameters of the forest-age model (equation (1)) were first estimated in a leave-one-site-out cross-validation (CV) mode for the entire dataset based on a generalized nonlinear least squares (gnls) model using R software (Team 2015). To account for temporal autocorrelation of the observations, we combined the gnls model with an auto-regression moving average model. We minimized the sum of squared residuals weighted for the uncertainty of the observations (Richardson and Hollinger 2005). The standard errors of model parameters were estimated using a bootstrapping algorithm ($N = 500$ random resamplings). The model output (i.e. f(Age)) was further included as a covariate in the training of the RF algorithms as a nonlinear formulation of age effects on spatio-temporal variability of NEP.

*RF algorithm and variables selection*: We tested and assessed a RF algorithm using the caret R package (Kuhn *et al* 2008). This is a non-parametric technique, i.e. it makes no assumption about the residuals of the model. A priori, we used a regression algorithm (i.e. the Boruta algorithm (Kursa and Rudnicki *et al* 2010)), to determine the best set of predictive variables for NEP among the aforementioned variables. The Boruta method relies on an RF method and determines relevance of each variable by comparing the relevance of the original predictors to that of the randomized variables. It iteratively removes the features that are shown by a statistical test to be less relevant than random probes. We decided to keep the five best variables to improve the accuracy of predictions.

*Model performance and model sensitivities*: The performance of the statistical model was evaluated by directly comparing model estimates with observed values of NEP for each site-average or site-year in a leave-one-site-out CV mode. In other words, we excluded one site at a time in every training set to predict the mean NEP (site-average) or the annual variations in NEP (site-year) at the excluded site. The statistics used to analyze the results included the coefficient of determination ($R^2$), Nash–Sutcliffe model efficiency coefficient (NSE), root mean squared error (RMSE), and mean absolute error (MAE) (Omlin and Reichert 1999). To quantify the importance of each predictive variable, we performed a model sensitivity analysis by removing a predictor from each regression analysis, then refitted and re-assessed the model

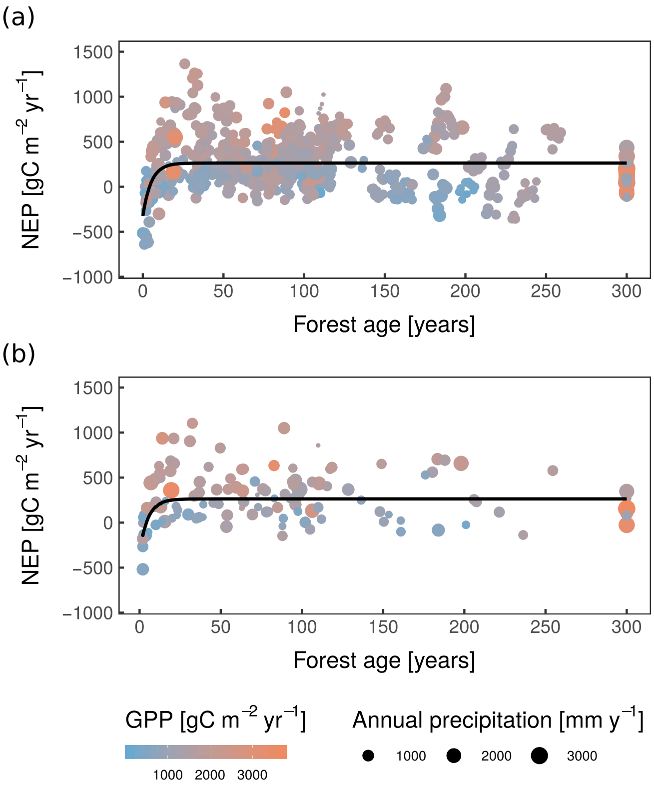

**Figure 2.** Relationship between NEP and forest age for (figure 2(a)) site-years and (figure 2(b)) site-average. The dots indicate the measurements. The lines indicate modeled NEP using the median of the parameter estimates from Amiro *et al* (2010) function.

**Table 1.** Model parameter estimates of univariate relationships between NEP and forest age. The standard errors of the parameter estimates are shown in brackets. Statistics of the forest age model are also shown. $R^2$ = coefficient of determination; MAE = mean absolute error; AIC = Akaike Information Criterion; total $n$ = 716 for all years per site and $n$ = 126 for average site.

| | Parameters | | | Model performance | | |
|---|---|---|---|---|---|---|
| | Offset [gC m$^{-2}$ y$^{-1}$] | a [gC m$^{-2}$ y$^{-1}$] | $b$ | $R^2$ | MAE | AIC |
| All years per site | −324.7 (106.8)$^{***}$ | 587.7 (106.9)$^{***}$ | −0.2 (0.07)$^{***}$ | 0.09 | 224.2 | 9 226.4 |
| Average per site | −482.6 (366.5) | 760.8 (362.7)$^{***}$ | −0.3 (0.2) | 0.2 | 209.7 | 1 753.2 |

without the variable left out. For the site-average analysis, the statistical model was trained using the site-years dataset. The same predictions were further averaged per site and compared to the site-average observations.

## 3. Results

### 3.1. Age-dependent forest carbon dynamics

The statistical dependence of NEP on forest age supports the nonlinear NEP-age relationship (figure 2 and equation (1)) in that NEP increased rapidly with age followed by stabilization with forests aging (figure 2). This finding reflects expected age-related change in the size and the dynamics of the C pools (Noormets *et al* 2015, Woodall *et al* 2015, Zhang *et al* 2015, Gray *et al* 2016). However, NEP-age dynamics appeared to depend of climatic conditions, because biomass accumulation rates of regrowing forests vary with climate (Chazdon *et al* 2016, Anderson-Teixeira *et al* 2013). This covariation partly explains the differences of the timing when a maximum NEP is reached and then gradually decreases as forest ages among different environmental conditions and the substantial scatter of observations around the model response due to inter-site variability.

The low correlation coefficient between NEP and age (table 1) ($R^2$ = 0.09 and 0.2 for site-years and site-average, respectively) could be attributed to the substantial contribution of other environmental factors to the spatial and temporal variability of NEP (figure 2). A model based on forest age alone is unable to capture such dependencies and warrants the need to include additional factors in a regression analysis. Although the regression does not show a substantial correlation, the fitted function showed a strong statistical significance, mostly because of the initial curve inflection attributed to the large effect of disturbances on the NEP fluxes in

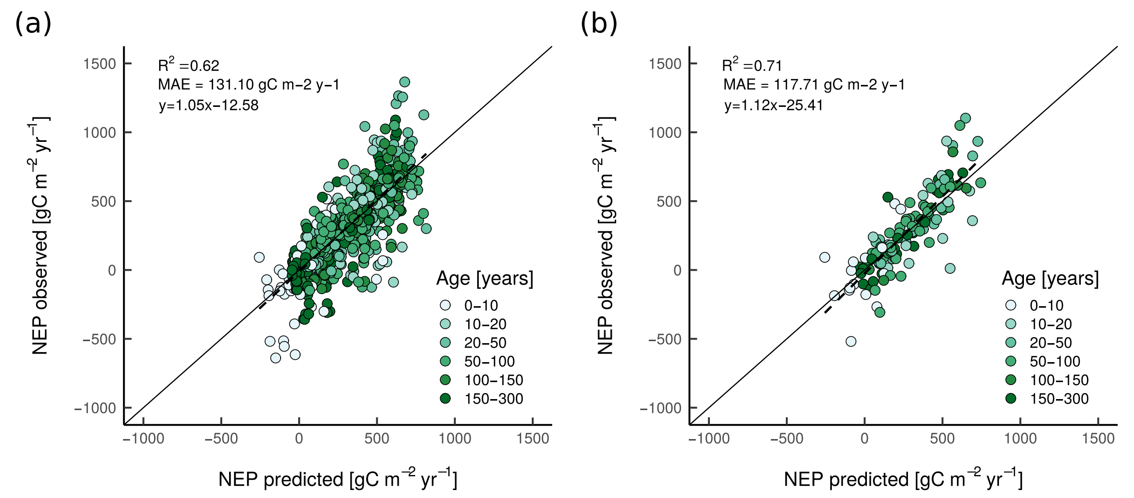

**Figure 3.** Cross-validated comparison of predicted versus observed NEP estimates using the best model set-up for (figure 3(a)) site-years and (figure 3(b)) site-average. The scatterplots of the observed versus modelled annual sums of NEP were grouped by age classes. NEP = f(age, f(Age), GPP′, MAT, $N_{deposition}$).

the first years of a recovery process (Amiro *et al* 2010, Coursolle *et al* 2012, Tang *et al* 2014).

### 3.2. Random forest algorithm performance and model sensitivity

Based on the aforementioned feature selection criterion (table S3), the RF algorithm accounted for the effects of forest age (i.e. age and f(Age)), GPP′, MAT, and $N_{deposition}$. Both site-years and site-average variability were well captured by the different RF models (NSE = 0.62 for site-years and NSE = 0.71 for site-average) (figure 3), suggesting that the structure of the models is suitable for reproducing the spatio-temporal patterns of annual NEP. In addition, for both scenarios (i.e. site-years and site-average), we found that a model including only forest age and GPP′ as predictive variables (i.e. NEP = f(age, f(Age), GPP′)) had a good predictive capacity for both site-years and site-average (NSE = 0.60 for site-years and NSE = 0.67 for site-average). Although we depicted some high values in the residuals across-site (maximum = 454.4 gC $m^{-2}$ $y^{-1}$; minimum = −537.4 gC $m^{-2}$ $y^{-1}$), we found no significant patterns of residuals against covariates (figure S4).

Model sensitivity tests whereby predictors were sequentially removed (table 2) supported the importance of forest age for explaining NEP variability. Whenever we removed forest age from the RF models, model performance decreased, while there were only small changes in model performance when removing either GPP′, MAT or $N_{deposition}$.

**Table 2.** Changes in model performance caused by removing predictors from the best model set-up and then refitting the model without the left out variable(s). These results were computed by leave-one-site-out cross-validation. The (-) symbol means that the predictive variable(s) were removed from the Random Forest models. $R^2$ = coefficient of determination; NSE = Nash–Sutcliffe model efficiency coefficient; RMSE = root mean squared error; MAE = mean absolute error; total $n$ = 716 for site-years and $n$ = 126 for site-average. NEP = f(age, f(Age), GPP′, MAT, $N_{deposition}$).

| | Δ $R^2$ | Δ NSE | Δ RMSE (gC $m^{-2}$ $y^{-1}$) | Δ MAE (gC $m^{-2}$ $y^{-1}$) |
|---|---|---|---|---|
| **Site-years** | | | | |
| Full model | 0.62 | 0.62 | 180.14 | 131.09 |
| (-) [Age + f(Age)] | −0.42 | −0.43 | +84.10 | +69.70 |
| (-) GPP′ | −0.01 | −0.02 | +4.13 | +1.54 |
| (-) MAT | −0.01 | −0.01 | +0.52 | −1.06 |
| (-) [GPP′ + MAT] | −0.05 | −0.06 | +14.54 | +9.81 |
| (-) $N_{deposition}$ | −0.01 | −0.01 | +3.40 | +0.38 |
| **Site-average** | | | | |
| Full model | 0.71 | 0.71 | 162.38 | 117.71 |
| (-) [Age + f(Age)] | −0.50 | −0.51 | +98.10 | +85.85 |
| (-) GPP′ | −0.04 | −0.04 | +5.98 | +3.10 |
| (-) MAT | −0.04 | −0.04 | +5.94 | +0.61 |
| (-) [GPP′ + MAT] | −0.12 | −0.13 | +25.96 | +18.93 |
| (-) $N_{deposition}$ | −0.02 | −0.02 | +0.59 | −3.22 |

## 4. Discussion

### 4.1. Forest age as a key driver of spatial and interannual variability in NEP

Based on theoretical principles of the C cycle at ecosystem scale, forest age is expected to play a significant role in NEP. Consistent with the early forest dynamics theory on net primary productivity (NPP) trajectories with forest age (Odum 1969), we empirically found strong support for a nonlinear relationship between NEP and forest age, although an age effect was not evident when looking at a univariate relationship, due to spatial variability of other local covariates (figure 2). Hence, we followed a multi-variate approach (figure 3) that accounts for the co-variarying

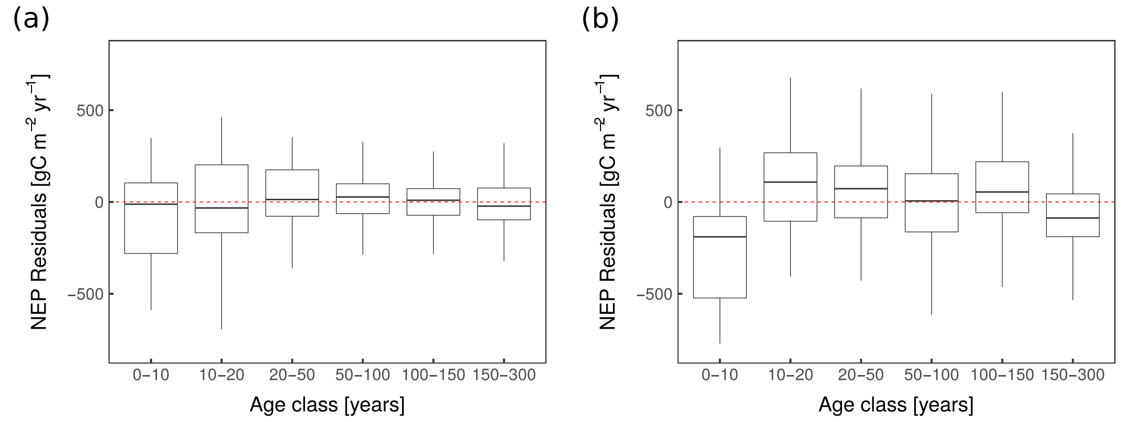

**Figure 4.** Model residuals per age class of (figure 4(a)) a RF model including forest age as a predictive variable, i.e. NEP = f(age, f(Age), GPP′, MAT, $N_{deposition}$) and (figure 4(b)) a RF model excluding forest age as a predictive variable, i.e. NEP = f(GPP′, MAT, $N_{deposition}$).

effects of other factors that change in space and time in order to assess the role of age in explaining NEP spatio-temporal variability. Furthermore, RF models have no prior assumption on the functional response between dependent and independent variables, therefore the relevance of forest age (i.e. age + f(age)) was also addressed by contrasting the model performance when removing variables (table 2) and by looking at the model residuals across age class (figure 4). Forest age emerges as the variable that explains most of the spatial and temporal variability in NEP, despite including information on climate and environmental conditions. Photosynthesis and respiration processes drive the link between NEP and forest age, therefore having long term time series of all component fluxes—enabling to establish individual curves per site—and observationally independent estimates of GPP/ER/NEP, could help disentangling whether the NEP-age dynamics are driven by the links between GPP and age or between ER and age.

Previously, the effect of stand age on the temporal variability of NEP has been demonstrated via the control of age on NPP using a global dataset (Pregitzer and Euskirchen 2004, Tang *et al* 2014). Similarly, forest age plays a dominant role in explaining spatial variability in NEP in the East Asian monsoon region (Yu *et al* 2014, Gao *et al* 2016). Unlike previous studies, we tested the effect of several drivers on both site-average and site-years, allowing us to evaluate both spatial and interannual NEP variability. Both analyses (i.e. site-average, site-years) showed that forest age was one of the main drivers of NEP variability (table 2). However, some factors are temporally invariant at yearly scales (e.g. soil texture) or do not change over time due to data limitations (e.g. nutrient availability), while others could change (e.g. forest age, GPP′, MAT). Therefore, the lack of temporal variability in these factors could reduce their contribution to the NEP in the site-years analysis.

GPP has been suggested as one of the main drivers of NEP spatial variability (Fernández-Martínez *et al* 2014). We found that excluding GPP′ from the RF algorithm decreased the model efficiency by 0.02 and 0.04 for site-years and site-average, respectively (table 2). GPP′ emerged as superior predictor compared to climate and soil properties (table S3), likely because it is more closely coupled with NEP, whereas climate and soil properties had variable effects depending on site characteristics. The statistical relation of GPP′ to NEP appeared to be significantly stronger for stands younger than 20 years, than for intermediate-aged/old-growth forests (⩾20 years) (figure S5). In the initial successional pathway, most of the year-to-year variability in NEP is explained by the changes in GPP and climate. However, as forest ecosystems mature and the autotrophic and respiratory processes start to balance each other out, the variations in NEP become more a function of forest age, or time since disturbance, rather than of individual variations in GPP or ER. Having the full representation of stand development stages is important for representing forest spatio-temporal C dynamics after stand-replacing disturbances more realistically (figure 4). This means that the controls of GPP on NEP are strongly dependent on the distribution of forest age, which emphasizes the relevance of age class distributions for understanding the dynamics of biosphere-atmosphere fluxes. The interactions between forest age and local conditions (e.g. GPP) suggest to move beyond stand age in reflecting changes in plant and soil pools, but also in appropriately parameterizing forest age-related changes in ecophysiological mechanisms both at plant and soil levels. Still, we have limited knowledge on the disturbance effects on detrital pools (and thus heterotrophic respiration, Rh), the type of transition between previous land cover/use (Carvalhais *et al* 2010) followed by different regeneration types (e.g. regrowth, plantation on pasture, former agricultural lands, and afforestation), and site history. These ecosystem conversions may strongly influence ecosystem C balance (Kutsch and Kolari 2015) and could explain the current bias present for the young forests (<20 years) (figure 4(a)).

### 4.2. Climate and soil properties controls on spatial annual NEP variability

While several environmental factors exert controls on NEP, we found that their statistical effect was minor in comparison to forest age. In many cases, sensitivities and even the sign of the relationship between environmental factors and NEP differ among case studies in the literature. For instance, NEP can be positively correlated with MAT in space (Fernández-Martínez *et al* 2014), whereas other studies find only very weak relationships (Law and Falge 2002, Reichstein *et al* 2007, Piao *et al* 2008). In boreal regions, air and soil temperature are the main factors affecting interannual NEP variability in old stands, while climatic conditions could not explain temporal patterns of NEP of young stands (Coursolle *et al* 2012). Here, MAT had a modest contribution to explaining NEP variability in the final model (table 2).

Rather than mean temperature, temperature changes in the recent past significantly influence current spatial variability of the forest C sink (Piao *et al* 2009). We tested both annual climate anomalies and climate trends (i.e. from 1960 to 2012 based on the CRU dataset) in the final models, but found that they had limited effect on explaining NEP variability (table S3). Nevertheless, future increases in temperature, changes in precipitation patterns and more extreme events will likely have significant effects on the C budgets of forest ecosystems (Thuiller *et al* 2011, Trumbore *et al* 2015).

Soil characteristics and fertility may play an important role in the spatial and interannual variability of NEP (Oren *et al* 2001, Janssens *et al* 2010, Fernández-Martínez *et al* 2014). Ecosystem C exchanges are generally limited by nutrient availability (often N) that may increase following disturbance (e.g. stand replacement, harvest). N mineralization increases available N while N uptake decreases N availability (Thornton *et al* 2002). However, we found that nutrient availability and clay content were not considered statistically strong drivers in explaining NEP variability. This was in contrast to previous studies concluding that nutrient availability is even more important than forest age in explaining across-site forest NEP variability (Fernández-Martínez *et al* 2014). This apparent contradiction emerged from the fact that earlier studies used a linear relationship, while we also included a nonlinear relationship. We showed that removing a nonlinear relationship between NEP and stand age (i.e. f(age)) results in a substantial loss in overall model performance and a significant reduction in the apparent importance of forest age for NEP (table S4).

Nevertheless, the apparent low contribution of climate, nutrient status, and soil properties to explaining NEP variability can be explained by the fact that their information is already embedded either in forest age or GPP′. The latter is clearly climatically driven, whereas the GPP-NEP dynamic is strongly controlled by forest nutrient availability (Fernández-Martínez *et al* 2014). Forest age is likely a superior predictor of spatio-temporal variation in NEP because it integrates relevant ecological information not captured by other single variables. In fact, forest age is rather a composite measure of numerous drivers that are more directly mechanistically coupled with C cycling processes.

## 5. Implications

While GPP, climate and soil properties are significant factors influencing the variability of NEP across space and time, we conclusively demonstrate that forest age performs as a strong indicator of spatio-temporal variability in NEP and is a useful integrated proxy for ecological changes that constrain NEP at the global scale.

Many global ecosystem models rely on simple representations of forest age dynamics and few consider the role of successional changes in C cycling processes (Anderson-Teixeira *et al* 2013), which introduces uncertainties into long term simulations of forest C dynamics (Friend *et al* 2014, Friedlingstein *et al* 2014). Additionally, given the statistical power of the proposed model in comparison to other state-of-the art approaches (Jung *et al* 2011, Tramontana *et al* 2016), this study points out new directions towards further developments in bottom-up upscaling exercises based on EC data. Regardless of the modeling strategy, reliable annual maps of forest age distribution and/or disturbance history will be required in order to make accurate predictions of NEP in space and time. These will further support the design of sustainable forest management and climate change mitigation strategies that depend on the effect of forest aging and age class distribution (Pan *et al* 2011, Thuiller *et al* 2011, Trumbore *et al* 2015).

This study emphasizes the need for increased focus on forest demography, which may amplify or exceed the importance of climate sensitivity for predicting the future of the terrestrial C cycle.

## Acknowledgments

This work used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DEFG0204ER63917 and DEFG0204ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, FluxnetCanada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOSSiberia and

USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAOGTOSTCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval, Environment Canada, US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California, Berkeley and University of Virginia. This research was funded by both the BACI project and the Independent Monitoring Project. We would like to thank Sönke Zaehle for providing the $N_{deposition}$ data. We thank Mirco Migliavacca and Ulrich Weber for the assistance with the processing of the EC and the CRU data and Andrew Durso for proofreading the manuscript. Finally, we would like to thank Tomislav Hengl for providing the SoilGrids1km dataset and Sara Vicca for helping us in determining the nutrient availability classification of some sites used in this study. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme I (NPU I), grant number LO1415. This research was funded by both the BACI project and the Independent Monitoring Project. We would also like to acknowledge NOVA grant UID/AMB/04085/2013 and the GlobBiomass project.

## ORCID iDs

Simon Besnard https://orcid.org/0000-0002-1137-103X
Sytze de Bruin https://orcid.org/0000-0002-6884-2832
Jiquan Chen https://orcid.org/0000-0003-0761-9458

## References

Amiro B *et al* 2006 Carbon, energy and water fluxes at mature and disturbed forest sites, Saskatchewan, Canada *Agric. Forest Meteorol.* **136** 237–51

Amiro B D *et al* 2010 Ecosystem carbon dioxide fluxes after disturbance in forests of North America *J. Geophys. Res.: Biogeosci.* **115**

Anderson-Teixeira K J, Miller A D, Mohan J E, Hudiburg T W, Duval B D and DeLucia E H 2013 Altered dynamics of forest recovery under a changing climate *Glob. Change Biol.* **19** 2001–21

Baldocchi D 2008 'breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems *Aust. J. Bot.* **56** 1–26

Baldocchi D *et al* 2001 Fluxnet: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities *Bull. Am. Meteorol. Soc.* **82** 2415–34

Becknell J M *et al* 2015 Assessing interactions among changing climate, management, and disturbance in forests: a macrosystems approach *BioScience* **65** 263–74

Beer C *et al* 2010 Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate *Science* **329** 834–8

Bhatti J, Apps M and Jiang H 2002 Influence of nutrients, disturbances and site conditions on carbon stocks along a boreal forest transect in central canada *Plant Soil* **242** 1–14

Bonan G B 2008 Forests and climate change: forcings, feedbacks, and the climate benefits of forests *Science* **320** 1444–9

Bowman D M *et al* 2009 Fire in the Earth system *Science* **324** 481–4

Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32

Campioli M *et al* 2015 Biomass production effciency controlled by management in temperate and boreal ecosystems *Nat. Geosci.* **8** 843–6

Canadell J *et al* 2000 Commentary: carbon metabolism of the terrestrial biosphere: a multitechnique approach for improved understanding *Ecosystems* **3** 115–30

Carvalhais N *et al* 2014 Global covariation of carbon turnover times with climate in terrestrial ecosystems *Nature* **514** 213–7

Carvalhais N, Reichstein M, Ciais P, Collatz G J, Mahecha M D, Montagnani L, Papale D, Rambal S and Seixas J 2010 Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints *Glob. Change Biol.* **16** 2813–29

Chambers J Q, Fisher J I, Zeng H, Chapman E L, Baker D B and Hurtt G C 2007 Hurricane katrina's carbon footprint on us gulf coast forests *Science* **318** 1107–1107

Chazdon R L *et al* 2016 Carbon sequestration potential of second-growth forest regeneration in the latin american tropics *Sci. Adv.* **2** e1501639

Chen J, Falk M, Euskirchen E, Paw U, Suchanek K T, Ustin T H, Bond S L, Brosofske B J, Phillips K D and Bi N R 2002 Biophysical controls of carbon flows in three successional douglas-fir stands based on eddy-covariance measurements *Tree Physiol.* **22** 169–77

Chen J M, Ju W, Cihlar J, Price D, Liu J, Chen W, Pan J, Black A and Barr A 2003 Spatial distribution of carbon sources and sinks in canada's forests *Tellus* B **55** 622–41

Ciais P *et al* 2014 Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system *Biogeosciences* **11** 3547–602

Ciais P *et al* 2005 Europe-wide reduction in primary productivity caused by the heat and drought in 2003 *Nature* **437** 529

Coursolle C *et al* 2012 Influence of stand age on the magnitude and seasonality of carbon fluxes in canadian forests *Agric. Forest Meteorol.* **165** 136–48

Coursolle C *et al* 2006 Late-summer carbon fluxes from canadian forests and peatlands along an east west continental transect *Can. J. Forest Res.* **36** 783–800

Fernández-Martínez M *et al* 2014 Nutrient availability as the key regulator of global forest carbon balance *Nat. Clim. Change* **4** 471–6

Friedlingstein P *et al* 2014 Persistent growth of co2 emissions and implications for reaching climate targets *Nat. Geosci.* **7** 709

Friend A D *et al* 2014 Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric co2 *Proc. Natl Acad. Sci.* **111** 3280–5

Gao S, Zhou T, Zhao X, Wu D, Li Z, Wu H, Du L and Luo H 2016 Age and climate contribution to observed forest carbon sinks in east asia *Environ. Res. Lett.* **11** 034021

Gower S T 2003 Patterns and mechanisms of the forest carbon cycle *Annu. Rev. Environ. Res.* **28** 169–204

Gray A N, Whittier T R and Harmon M E 2016 Carbon stocks and accumulation rates in pacific northwest forests: role of stand age, plant community, and productivity *Ecosphere* **7** e01224

Harmon M E, Bond-Lamberty B, Tang J and Vargas R 2011 Heterotrophic respiration in disturbed forests: a review with examples from north america *J. Geophys. Res.: Biogeosci.* **116**

Harris I, Jones P, Osborn T and Lister D 2014 Updated high-resolution grids of monthly climatic observations-the cru ts3.10 dataset *Int. J. Climatol.* **34** 623–42

Hengl T *et al* 2014 Soilgrids1kmglobal soil information based on automated mapping *PLoS One* **9** e105992

Hicke J A *et al* 2012 Effects of biotic disturbances on forest carbon cycling in the united states and canada *Glob. Change Biol.* **18** 7–34

Janssens I *et al* 2010 Reduction of forest soil respiration in response to nitrogen deposition *Nat. Geosci.* **3** 315

Jung M *et al* 2011 Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations *J. Geophys. Res.: Biogeosci.* **116**

Katul G, Manzoni S, Palmroth S and Oren R 2009 A stomatal optimization theory to describe the effects of atmospheric co$_2$ on leaf photosynthesis and transpiration *Ann. Bot.* **105** 431–42

Katul G G, Palmroth S and Oren R 2009 Leaf stomatal responses to vapour pressure deficit under current and co$_2$-enriched atmosphere explained by the economics of gas exchange *Plant, Cell Environ.* **32** 968–79

Kuhn M *et al* 2008 Caret package *J. Stat. Softw.* **28** 1–26

Kursa M B *et al* 2010 Feature selection with the boruta package *J. Stat. Softw.* **36** 1–13

Kurz W A, Dymond C, Stinson G, Rampley G, Neilson E, Carroll A, Ebata T and Safranyik L 2008 Mountain pine beetle and forest carbon feedback to climate change *Nature* **452** 987–90

Kutsch W L and Kolari P 2015 Data quality and the role of nutrients in forest carbon-use efficiency *Nat. Clim. Change* **5** 959

Law B *et al* 2002 Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation *Agric. Forest Meteorol.* **113** 97–120

Law B E, Hudiburg T W, Berner L T, Kent J J, Buotte P C and Harmon M E 2018 Land use strategies to mitigate climate change in carbon dense temperate forests *Proc. Natl Acad. Sci.* 201720064

Law B E, Sun O, Campbell J, Van Tuyl S and Thornton P 2003 Changes in carbon storage and fluxes in a chronosequence of ponderosa pine *Glob. Change Biol.* **9** 510–24

Le Quéré C *et al* 2018 Global carbon budget 2017 *Earth Syst. Sci. Data* **10** 405

Lindauer M, Schmid H, Grote R, Mauder M, Steinbrecher R and Wolpert B 2014 Net ecosystem exchange over a non-cleared wind-throw-disturbed upland spruce forestmeasurements and simulations *Agric. Forest Meteorol.* **197** 219–34

Lindroth A *et al* 2008 Leaf area index is the principal scaling parameter for both gross photosynthesis and ecosystem respiration of northern deciduous and coniferous forests *Tellus* B **60** 129–42

Luyssaert S, Schulze E-D, Börner A, Knohl A, Hessenmöller D, Law B E, Ciais P and Grace J 2008 Old-growth forests as global carbon sinks *Nature* **455** 213

Naudts K, Chen Y, McGrath M J, Ryder J, Valade A, Otto J and Luyssaert S 2016 Europeas forest management did not mitigate climate warming *Science* **351** 597–600

Noormets A, Epron D, Domec J, McNulty S, Fox T, Sun G and King J 2015 Effects of forest management on productivity and carbon sequestration: a review and hypothesis *Forest Ecol. Manage.* **355** 124–40

Noormets A, McNulty S G, Domec J-C, Gavazzi M, Sun G and King J S 2012 The role of harvest residue in rotation cycle carbon balance in loblolly pine plantations. respiration partitioning approach *Glob. Change Biol.* **18** 3186–201

Odum E P 1969 The strategy of ecosystem development *Science* **164**

Omlin M and Reichert P 1999 A comparison of techniques for the estimation of model prediction uncertainty *Ecol. Modelling* **115** 45–59

Oren R *et al* 2001 Soil fertility limits carbon sequestration by forest ecosystems in a co2-enriched atmosphere *Nature* **411** 469

Pan Y, Chen J M, Birdsey R, McCullough K, He L and Deng F 2011 Age structure and disturbance legacy of north american forests *Biogeosciences* **8** 715

Paul-Limoges E, Black T A, Christen A, Nesic Z and Jassal R S 2015 Effect of clearcut harvesting on the carbon balance of a douglas-fir forest *Agric. Forest Meteorol.* **203** 30–42

Piao S *et al* 2008 Net carbon dioxide losses of northern ecosystems in response to autumn warming *Nature* **451** 49

Piao S, Friedlingstein P, Ciais P, Peylin P, Zhu B and Reichstein M 2009 Footprint of temperature changes in the temperate and boreal forest carbon balance *Geophys. Res. Lett.* **36**

Pregitzer K S and Euskirchen E S 2004 Carbon cycling and storage in world forests: biome patterns related to forest age *Glob. Change Biol.* **10** 2052–77

Reichstein M *et al* 2013 Climate extremes and the carbon cycle *Nature* **500** 287

Reichstein M *et al* 2005 On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm *Glob. Change Biol.* **11** 1424–39

Reichstein M *et al* 2007 Determinants of terrestrial ecosystem carbon balance inferred from european eddy covariance flux sites *Geophys. Res. Lett.* **34**

Richardson A D and Hollinger D Y 2005 Statistical modeling of ecosystem respiration using eddy covariance data: maximum likelihood parameter estimation, and monte carlo simulation of model and parameter uncertainty, applied to three simple models *Agric. Forest Meteorol.* **131** 191–208

Schwarz P A, Law B, Williams M, Irvine J, Kurpius M and Moore D 2004 Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems *Glob. Biogeochem. Cycles* **18**

Spies T A 1991 The structure of natural young, mature, and old-growth douglas-fir forests in oregon and washington *Wildlife and Vegetation of Unmanaged Douglas-Fir Forests* (Portland, OR: US Department of Agriculture) 91–109

Tang J, Luyssaert S, Richardson A D, Kutsch W and Janssens I A 2014 Steeper declines in forest photosynthesis than respiration explain age-driven decreases in forest growth *Proc. Natl Acad. Sci.* **111** 8856–60

R Development Core Team 2015 R: a language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing) (www.r-project.org)

Thornton P E *et al* 2002 Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests *Agric. Forest Meteorol.* **113** 185–222

Thuiller W, Lavergne S, Roquet C, Boulangeat I, Lafourcade B and Araujo M B 2011 Consequences of climate change on the tree of life in europe *Nature* **470** 531

Tramontana G *et al* 2016 Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms *Biogeosciences* **13** 4291–313

Trumbore S, Brando P and Hartmann H 2015 Forest health and global change *Science* **349** 814–8

Vicca S *et al* 2012 Fertile forests produce biomass more efficiently *Ecol. Lett.* **15** 520–6

Wang R *et al* 2017 Global forest carbon uptake due to nitrogen and phosphorus deposition from 1850 to 2100 *Glob. Change Biol.* **23** 4854–72

Woodall C, Russell M, Walters B, D'Amato A, Fraver S and Domke G 2015 Net carbon flux of dead wood in forests of the eastern us *Oecologia* **177** 861–74

Xiao J, Zhuang Q, Liang E, Shao X, McGuire A D, Moody A, Kicklighter D W and Melillo J M 2009 Twentieth-century droughts and their impacts on terrestrial carbon cycling in china *Earth Interact.* **13** 1–31

Yi C *et al* 2010 Climate control of terrestrial carbon exchange across biomes and continents *Environ. Res. Lett.* **5** 034007

Yu G, Chen Z, Piao S, Peng C, Ciais P, Wang Q, Li X and Zhu X 2014 High carbon dioxide uptake by subtropical forest ecosystems in the east asian monsoon region *Proc. Natl Acad. Sci.* **111** 4910–5

Zhang B, Zhou X, Zhou L and Ju R 2015 A global synthesis of below-ground carbon responses to biotic disturbance: a meta-analysis *Glob. Ecol. Biogeogr.* **24** 126–38

Zhao M and Running S W 2010 Drought-induced reduction in global terrestrial net primary production from 2000 through 2009 *Science* **329** 940–3

Zscheischler J *et al* 2017 Reviews and syntheses: an empirical spatiotemporal description of the global surface-atmosphere carbon fluxes: opportunities and data limitations *Biogeosciences* **14** 3685–703