LETTER

Demographic controls of aboveground forest biomass across North America

Mark C. Vanderwel,[1,2]*
Hongcheng Zeng,[3] John P.
Caspersen,[3] Georges Kunstler[4,5]
and Jeremy W. Lichstein[2]

Abstract
Ecologists have limited understanding of how geographic variation in forest biomass arises from differences in growth and mortality at continental to global scales. Using forest inventories from across North America, we partitioned continental-scale variation in biomass growth and mortality rates of 49 tree species groups into (1) species-independent spatial effects and (2) inherent differences in demographic performance among species. Spatial factors that were separable from species composition explained 83% and 51% of the respective variation in growth and mortality. Moderate additional variation in mortality (26%) was attributable to differences in species composition. Age-dependent biomass models showed that variation in forest biomass can be explained primarily by spatial gradients in growth that were unrelated to species composition. Species-dependent patterns of mortality explained additional variation in biomass, with forests supporting less biomass when dominated by species that are highly susceptible to competition (e.g. *Populus* spp.) or to biotic disturbances (e.g. *Abies balsamea*).

Keywords
biomass, carbon, growth, mortality, productivity, trees.

Ecology Letters (2016) **19**: 414–423

INTRODUCTION

Forests contain 45% of global terrestrial carbon stores (Bonan 2008), and each year absorb about 25% of the carbon emitted to the atmosphere through the burning of fossil fuels (Pan *et al.* 2011a). Despite the present and future importance of forest carbon, changes in the global terrestrial carbon sink remain the largest uncertainty in predictive models of the Earth's coupled carbon–climate system (Friedlingstein *et al.* 2006). The biomass of living trees represents a large share of all vegetation carbon, and is also the precursor to a substantial fraction of soil carbon. Processes that regulate forest biomass at broad spatial and temporal scales therefore have a major influence on global terrestrial carbon cycling.

Annual biomass accumulation (or loss) is the net result of forest growth (defined here as aboveground woody biomass production per unit area) and mortality. Global variation in forest productivity has been mapped with both ground-based and remote sensing data, and ecologists have successfully developed predictive models that relate productivity to underlying environmental drivers (Leith 1975; Yuan *et al.* 2007). Our knowledge of broad-scale patterns in forest mortality is less well developed, but is advancing. In recent years, ecologists have quantified continental and global patterns in mortality and disturbance (Stephenson & van Mantgem 2005; Lines *et al.* 2010), and have begun defining the causal relationships underlying spatial variation in mortality (Stephenson *et al.* 2011). At present though, the available data present an

unclear picture of how climate-related variation in mortality affects global terrestrial carbon balance (Smith *et al.* 2013).

Spatial variation in forest biomass can provide valuable insights into carbon storage under global change. A key question concerns the degree to which forest biomass is regulated by differences in growth vs. mortality across space. Both of these processes can be affected by spatial variation in environmental conditions, such as climate and soil, as well as by the species composition of a particular area through interspecific variation in demographic performance. Species turnover across space and time could be particularly important if shifts in tree species composition influence ecosystem responses to environmental change. For example, climate warming is projected to induce physiological increases in the productivity of New Zealand forests, but many of these growth increases could be offset if slow-growing lowland species track their climate niche and migrate into cooler montane forests (Coomes *et al.* 2014). A multi-decadal drought in Ghana has produced shifts in species composition that have in turn increased forest biomass, in strong contrast to the biomass losses that often follow drought events in the short term (Fauset *et al.* 2012). These examples suggest that tree species turnover can have important regional effects on forest biomass.

To quantify demographic controls over forest biomass at broad scales, we must disentangle the influence of species composition on growth and mortality from that of inherent spatial and environmental variation, then evaluate the degree to which each source of variation affects forest biomass.

[1]*Department of Biology, University of Regina, 3737 Wascana Pkwy, Regina, SK S4S 0A2, Canada*
[2]*Department of Biology, University of Florida, Gainesville, FL 32611, USA*
[3]*Faculty of Forestry, University of Toronto, 33 Willcocks St., Toronto, ON M5S 3B3, Canada*

[4]*Irstea, UR EMGR, 2 rue de la Papeterie-BP 76, St-Martin-d'Hères, F-38402, France*
[5]*Univ. Grenoble Alpes, Grenoble, F-38402, France*
Correspondence: E-mail: mark.vanderwel@uregina.ca

Addressing this issue empirically requires demographic data across suitably broad spatial and compositional gradients, as well as analytical methods that can isolate the effects of different sources of demographic variation on forest biomass. Here, we integrate data from forest inventory plot networks spanning the United States and Canada to estimate continental variation in demographic performance from repeat censuses of over 4.9 million individual trees. These data cover most of the spatial extent and species composition of temperate and boreal forests across North America. We statistically partition continental-scale variation in both growth and mortality of 49 tree species groups into three components: (1) spatial effects, which describe common gradients in demographic performance across space for all species; (2) species effects, which describe consistent differences in performance among species groups; and (3) residual effects, which capture idiosyncratic variation in species group performance in different locations. Finally, we quantify how geographic variation in forest biomass across North America depends on the above three types of demographic effects for both growth and mortality.

METHODS

Forest inventory data

We estimated continental variation in forest growth, mortality and biomass from networks of forest inventory plots across the United States and Canada. Data for the United States were obtained from 121 521 plots in the Forest Inventory and Analysis (FIA) program that had been re-measured after a 5.3 ± 1.8 (mean ± SD) year interval. These plots collectively included 3 479 132 trees > 2.54 cm DBH. Plots spanned 46 of the 48 conterminous states (excluding New Mexico and Wyoming, where plots had not yet been re-measured), as well as Alaska, Puerto Rico and the US Virgin Islands. Data for Canada were obtained from individual government agencies in 8 of the 10 provinces (excluding Alberta and Prince Edward Island). Across Canada, 18 973 plots collectively included 1 431 906 individual trees. Plot size, minimum DBH (range 1–9 cm) and re-measurement intervals (8.5 ± 3.5 years) all varied among provinces.

We divided the inventory plots into four regions by political jurisdiction (conterminous states plus US territories, Canada plus Alaska) and longitude (east or west of 97 °W). We then applied a k-means clustering procedure within each region to aggregate plots into a set of 150 total groups, hereafter referred to as clusters (Fig. 1). The plot clusters averaged 54×10^3 km^2 in size, and were created so that each contained a suitably large sample of trees for estimating mean demographic performance across a broad landscape. We chose this number of clusters such that each cluster represented a relatively distinct set of environmental conditions, there was high beta diversity among clusters and cluster size approximated the resolution of global vegetation models (Appendix D provides a detailed rationale for the number of clusters, as well as a sensitivity analysis that shows our main results were robust when plots were aggregated into more or fewer clusters). All statistical modelling was subsequently performed at the cluster level rather than for individual plots or trees.

For each cluster (c), we calculated the mean biomass ($B_{c,s}^{<obs>}$; Mg·ha^{-1}), relative abundance by biomass ($A_{c,s}$; unitless), growth ($G_{c,s}^{<obs>}$; Mg ha^{-1} year^{-1}) and biomass mortality ($M_{c,s}^{<obs>}$; year^{-1}) for each of 24 softwood and 25 hardwood species groups (s) defined by the FIA program (O'Connell et al. 2013; see Appendix A for detailed data processing procedures). Most (26 of 49) of the species groups effectively represented a single species (> 90% of individuals in the group being one species), but others comprised multiple species that were geographically or taxonomically related and shared similar functional characteristics. To ease readability, we hereafter use the term 'species' in referring to these 49 groups (e.g. species composition, species effects).

We also estimated an annual harvest rate (H_{SPAT}; year^{-1}) for each cluster. For clusters in the eastern United States, we used the FIA data to calculate harvest rate as the biomass of trees that were harvested between censuses divided by the biomass of all trees initially alive and by the number of years between censuses. Plots in the western United States and in Canada did not generally contain tree-level data on harvesting. For these regions, we estimated the annual harvest rate from reported state- or province-level timber yields relative to the estimated growing stock (Canadian Council of Forest Ministers 2005, Smith et al. 2009).

Partitioning variation in demographic rates

Variation in community-level growth or mortality can arise through different combinations of species and spatial effects on demographic performance. We distinguish four cases, depicted conceptually in Fig. 2. In the first case, we expect that species performance does not change across clusters, and that community-level growth or mortality is determined only by the relative abundance of species with different demographic rates (Fig 2a). Conversely, species might have nearly equal performance in clusters where they co-occur, with their demographic rates increasing in concert along latent environmental gradients (Fig. 2b). Even without identifying the particular environmental factors that define these gradients, this pattern can be discerned by estimating spatial effects that describe consistent changes in the performance of all species present in a given cluster. Combining the previous two cases, species may respond not only to environmental gradients in a common manner, but also have intrinsic differences in performance (Fig. 2c). Finally, species may respond to environmental conditions individualistically (Fig. 2d). Species' performances would be only weakly correlated with one another in this last case, implying strong species × environment interactions in their demographic rates.

None of these cases imply anything about how species composition is itself controlled by environmental conditions. We are not attempting to explain why particular species are found in a given location, or how they might perform outside of their current range. Rather, we seek to tease out the relative importance of environmental conditions and species composition in determining variation in forest growth and mortality (and hence biomass) across North America.

To discriminate between the four cases described above, we constructed hierarchical Bayesian models that predict species
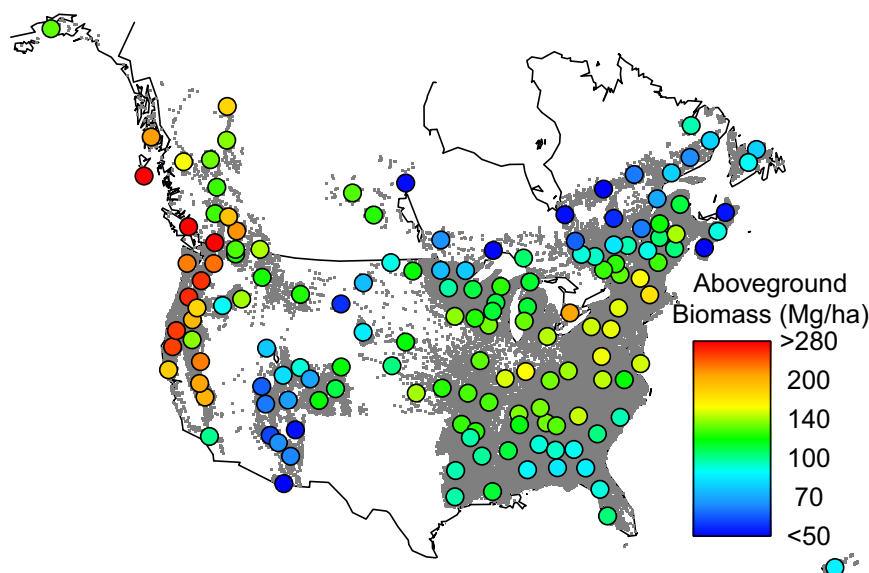
Figure 1 Map of the geographic distribution of individual forest plots (grey) and 150 plot clusters (coloured circles) across North America. Plots were aggregated into clusters using a k-means clustering algorithm that minimised the sum of the distance between each plot and the centre of the nearest cluster. We then calculated mean cluster-level biomass, growth and mortality for each of the 49 species groups defined by the Forest Inventory and Analysis program. We used a combination of plot data and state- or province-level reporting to estimate an annual harvest rate for each cluster.

growth and mortality in each cluster from a combination of species and spatial effects. We predicted the growth ($G_{c,s}^{<pred>}$; Mg ha^{-1} year^{-1}) and mortality ($M_{c,s}^{<pred>}$; year^{-1}) rates of species s in cluster c as:

$$G_{c,s}^{<pred>} = A_{c,s} \cdot exp(g_0 + g_c + g_s + g_{c \cdot s})$$
$$M_{c,s}^{<pred>} = logit^{-1}(m_0 + m_c + m_s + m_{c \cdot s}) \tag{1}$$

where $A_{c,s}$, the relative abundance of the species in that cluster, scales growth for differences in species' abundance among clusters (i.e. the biomass growth of a species is proportional to its share of the occupied growing space). The parameters of the models included: (1) global constants (g_0, m_0) that represent the performance of a typical (or mean) species in a typical environment (cluster); (2) spatial effects (g_c, m_c for each cluster c) that represent how the performance of a typical species would vary among clusters with different environmental conditions; (3) species effects (g_s, m_s for each species s) that represent how the performance of each species would vary from others in a typical cluster and (4) residual effects ($g_{c \cdot s}$, $m_{c \cdot s}$ for each species × cluster combination) that capture individualistic variation in species performance in each cluster. For both the growth and mortality models, spatial effects, species effects and residual effects for each species were all modelled as normally distributed random effects. We designed these demographic models to closely fit raw observations so that the models could effectively partition observed variation among different effects. The relative magnitudes of species, spatial and residual effects thus determined which of the cases in Fig. 2 best describe species-level performance: (1) large species effects, zero spatial or residual effects; (2) small species effects, large spatial effects, zero residual effects; (3) large species and spatial effects, zero residual effects; (4) large species, spatial, and residual effects.

We estimated the posterior distribution of all model parameters using Bayesian methods. Observed growth and mortality for each species × cluster combination were modelled as normal and beta-distributed random variables, respectively, with expected values given by the predicted rates in eqn 1. Further technical details on model specification and fitting can be found in Appendix B. Species that were absent from a given cluster (i.e. $A_{c,s} = 0$) had no effect on model fitting nor on the estimation of cluster-level performance described in the following section.

We implemented the demographic models using the Stan modelling language and software (Stan Development Team 2014). We obtained posterior samples from six MCMC chains and confirmed that these chains all converged to a well-mixed distribution of posterior parameter values. We verified that model parameters were identifiable by generating new pseudo-data using the hierarchical parameter distributions, then confirming that our estimation procedure could successfully recover known underlying parameter values (see Appendix C). All parameters were summarised by their posterior mean for subsequent analyses.

How demographic variation controls forest biomass

We next used the statistical models of demographic performance to isolate how species and spatial effects on growth and mortality each correlated with cluster-level variation in forest biomass. We did this by using the parameters for species-level performance to construct a series of reduced models that estimated growth and mortality at the level of each individual cluster (Table 1); we then evaluated the goodness of fit of simple regressions between each model of cluster-level performance and forest biomass. The cluster-level demographic
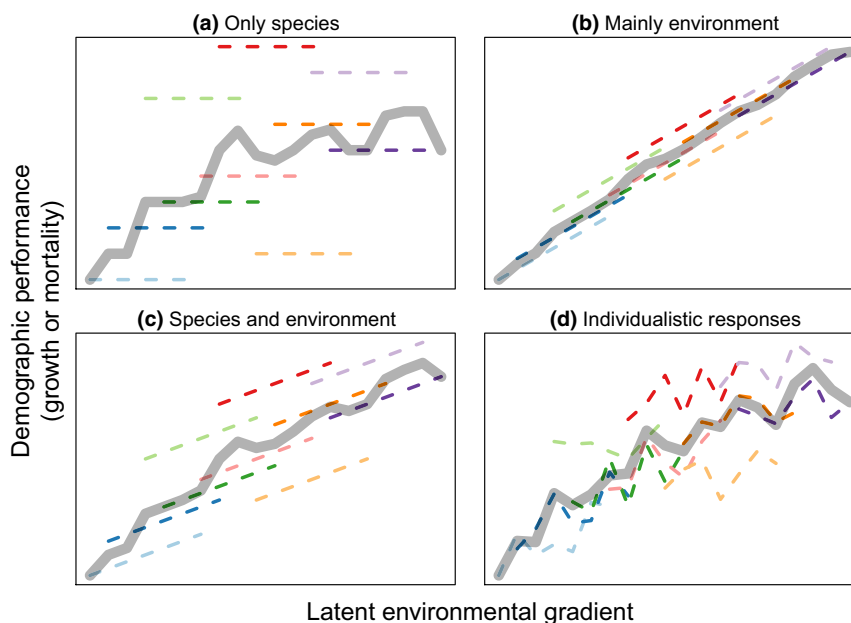
Figure 2 Hypothetical scenarios for how species composition and/or environmental conditions may explain variation in forest growth or mortality across space. The *x*-axis represents a latent environmental gradient (i.e. related to unmeasured physical factors) affecting species-level demographic performance. The dashed coloured lines represent the performance of individual species across this gradient, and the thick solid lines represent community-level performance averaged across the species present in a given location. (a) Performance varies among species, but individual species (where they occur) are unaffected by environmental conditions; community-level performance is determined only by species composition. (b) All species have approximately equal performance in a given environment; community-level performance is largely determined by the environment and does not strongly depend on species composition. (c) Performance varies among species, but each species responds to environmental conditions in a common manner; community-level performance depends on both species composition and the environment. (d) Species responses to environmental conditions are only weakly correlated with one another (due to species × environment interactions); community-level performance depends on how individual species perform in a particular environment. The figure illustrates hypothetical scenarios based on the artificial data that were generated from a simple additive model where the ratios of the standard deviations for species, environment and interaction terms were (a) 1 : 0 : 0; (b) 1 : 4 : 0; (c) 1 : 1 : 0; (d) 2 : 2 : 1.

models all used the same set of parameter estimates defined above, but each reduced model omitted the influence of one or more sources of variability by setting the relevant random-effect parameters to zero.

Fig. 2 provides a convenient way to visualise this procedure. Our full demographic models generated predictions that resembled Fig. 2d, and could nearly perfectly explain variation in cluster-level growth and mortality. By replacing the estimated residual effects with zeros, the full models are reduced to the case shown in Fig. 2c (i.e. perfect correlation in species performance across space). If species × environment interactions in demographic rates are unimportant, then we would expect the cluster-level demographic rates predicted in Fig. 2c to explain just as much variation in forest biomass as those in Fig. 2d. Following similar logic, we can test the importance of species or spatial effects to variation in biomass by setting the appropriate terms to zero to obtain the patterns in Fig. 2a or 2b. We can then assess how much our ability to predict cluster-level biomass from demographic performance is degraded by each simplification of the full demographic model.

It is important to recall that spatial and species effects were estimated only once, so that the simpler models were each obtained from the full models by dropping specific terms. For example, our predictions for mortality with spatial effects only (M_{SPAT}, depicted in Fig. 2b) were obtained using parameter estimates from the full model but with the non-spatial terms set to zero (i.e. m_0 and m_c estimated using eqn 1, and $m_s = m_{c \cdot s} = 0$). If instead one were re-estimate a model with spatial effects (m_0 and m_c for the M_{SPAT} model) while ignoring species composition, the resulting model would implicitly include differences in species composition among clusters, rather than controlling for species composition as our analysis does.

Modelling forest biomass from demography

We sought to estimate the relative importance of different demographic effects (e.g. spatial and species effects on growth and mortality) for explaining variation in forest biomass. To do this, we incorporated the cluster-level demographic models in Table 1 into a simple model that predicted biomass (B; Mg ha^{-1}) as a function of growth (G; Mg ha^{-1} year^{-1}), mortality (M; year^{-1}), harvest removals (H; year^{-1}) and mean forest age (T, year):

$$B = \frac{G}{M + H} \cdot \left(1 - e^{-(M+H) \cdot T}\right) \qquad (2)$$

Equation 2 specifies that, within each cluster, biomass increases monotonically with mean forest age towards an asymptote calculated as the ratio of annual increment to loss rate ($G/(M + H)$). The loss rate determines how quickly

Table 1 Formulae and interpretations for simplified models of cluster-level forest growth and mortality

Model labels	Random effects set to zero	Formulae	Interpretation
G_{CONST}/M_{CONST}	All	$G_{NULL} = \exp(g_0)$ $M_{NULL} = \text{logit}^{-1}(m_0)$	There is no spatial variation in growth/mortality. This is a trivial case, but provides a basis of comparison for the other models
G_{SPEC}/M_{SPEC}	Spatial, residuals	$G_{SPEC} = \sum_s (A_s \cdot \exp(g_0 + g_s))$ $M_{SPEC} = \sum_s (A_s \cdot \text{logit}^{-1}(m_0 + m_s))$	There is variation in growth/mortality among species, but species performance is constant across clusters (Fig. 2a)
G_{SPAT}/M_{SPAT}	Species, residuals	$G_{SPAT} = \exp(g_0 + g_c)$ $M_{SPAT} = \text{logit}^{-1}(m_0 + m_c)$	There is variation in growth/mortality among spatial clusters, but not among co-occurring species (Fig. 2b)
G_{BOTH}/M_{BOTH}	Residuals	$G_{BOTH} = \sum_s (A_s \cdot \exp(g_0 + g_c + g_s))$ $M_{BOTH} = \sum_s (A_s \cdot \text{logit}^{-1}(m_0 + m_c + m_s))$	There is variation among both clusters and species. Species share a common response to latent environmental factors (Fig. 2c)
G_{FULL}/M_{FULL}	None	$G_{FULL} = \sum_s (A_s \cdot \exp(g_0 + g_c + g_s + g_{c \cdot s}))$ $M_{FULL} = \sum_s (A_s \cdot \text{logit}^{-1}(m_0 + m_c + m_s + m_{c \cdot s}))$	Saturated model that approximates the observed growth/mortality data. Species can respond to latent environmental factors in an individualistic manner (Fig. 2d)

biomass accumulates, with a cluster supporting half of its maximum biomass at a mean age of $T = \ln(2)/(M + H)$. By using the different formulations for growth and mortality in Table 1 with this model, we can evaluate how well each demographic effect explains variation in biomass while controlling for differences in age and harvest rate among clusters.

Similar model formulations have been used for modelling vegetation biomass based on the difference between input and loss rates elsewhere (Olson 1963; Kucharik *et al.* 2000; Fisher *et al.* 2008; Smith *et al.* 2013). The differential equation form of this model only strictly applies at the stand level, and so eqn 2 is not appropriate for predicting dynamic changes in cluster biomass. We do not expect this limitation to affect comparisons of different growth and mortality effects, and we use eqn 2 in a purely heuristic way to predict forest biomass from cluster-level estimates of G, M, H and T. (To verify that results were robust, Appendix E presents an alternative, multiple regression approach for predicting biomass from different growth, mortality, harvest and age effects; this modelling approach produced very similar conclusions to those derived from eqn 2.)

We used Pan *et al.*'s (2011b) forest age map to estimate the mean age of each cluster (spatially averaged across the minimum convex polygon enclosing its plots). We then used eqn 2 to predict cluster-level biomass with each combination of the growth and mortality models in Table 1. By comparing the fit of models with spatial, species and residual effects incorporated into G and M, we quantified the contribution of each type of effect to predicting observed biomass. We also considered three formulations for harvest removals: a version where harvest rates were estimated for each cluster separately (H_{SPAT}), a version where the global mean harvest rate was used ($H_{CONST} = 1.03\%$ per year)

and a version that did not explicitly account for harvest losses ($H_{ZERO} = 0\%$ per year). Together, these yielded a set of 5 growth × 5 mortality × 3 harvesting = 75 biomass models. We fit a simple linear regression between log-observed and log-predicted biomass for each of these models to correct for potential biases in the biomass estimates from eqn 2, and compared the resulting R^2 and AIC values across models. Each of the 75 biomass models had three fitted parameters (regression slope and intercept, plus residual error) regardless of which effects were included in G, M, and H.

The importance of spatial, species and residual effects on growth and mortality, as well as spatial variation in harvest rate, was estimated from the sum of the Akaike weights ($\Sigma\omega_i$) associated with each variant of G, M and H. These weights measure the evidence for each particular demographic effect (*CONST, SPEC, SPAT, BOTH* and *FULL* versions of G and M; *ZERO, CONST* and *SPAT* versions of H) belonging in the best model within the set (Burnham & Anderson 2002).

Environmental correlates of spatial variation

To aid in interpreting spatial gradients in growth and mortality, we performed simple regressions between our estimated spatial effects models (G_{SPAT}, M_{SPAT}) and a basic set of climatic and edaphic variables. Climatic variables included mean annual temperature and precipitation from the World-Clim database (Hijmans *et al.* 2005). Edaphic variables included the sand, silt and clay content of the topsoil layer from the Unified North American Soil Map (Liu *et al.* 2013). All environmental variables were spatially averaged across the minimum convex polygons that enclosed each cluster's plots.

RESULTS

Partitioning components of demographic variation

Growth of different species varied in concert across space (Fig. 3), and this species-independent growth variation was strongly correlated with forest biomass. Specifically, spatial effects accounted for 83% of the variation in forest growth across North America (Table 2), and explained considerable variation in forest biomass in each of four sub-continental regions ($R^2 = 48$–78%; Fig. 4c). By contrast, spatial mortality effects explained 51% of all variation in mortality across North America, but were an inconsistent predictor of forest biomass ($R^2 = 0.01$–0.48 within regions; $R^2 = 0.18$ continent wide; Fig. 4d).

We detected little variation in growth among different species when controlling for spatial effects (Fig. 3). Systematic differences in the growth of different species (i.e. species effects) explained little variation in forest biomass on their own, and produced no improvement when added to the biomass model that included spatial growth effects (Fig. 4a,e). Conversely, species variation in mortality appeared to be an important component of forest demography (Fig. 3). When species mortality effects were considered alongside spatial effects, they explained 26% more variation in mortality (Table 2) and yielded stronger and more consistent relationships with forest biomass than spatial mortality effects alone ($R^2 = 0.36$ vs. 0.18 continent wide; $R^2 = 0.15$–0.38 within regions; Fig. 4f).

Idiosyncratic patterns of variation in species performance (i.e. residual variation after accounting for the main spatial and species effects) explained 14% of variation in growth and 17% of variation in mortality (Table 2). However, these full demographic models did not improve overall relationships with observed biomass across North America (continent wide $R^2 = 0.49$ for growth, 0.32 for mortality; Fig. 3g,h).

Demographic controls of forest biomass variation

Our age-dependent model (eqn 2) showed that spatial variation in growth was the most important factor for predicting

Table 2 Comparison of the goodness of fit and model selection evidence for individual components of growth and mortality, and differences in harvest intensity, in models predicting variation in forest biomass across North America

Process	Effects included	R^2 for demographic rate*	$\overline{\Delta R^2}$ for biomass†	Sum of Akaike weights‡
Growth	None (G_{CONST})	–	–	< 0.001
	Species (G_{SPEC})	0.01	−0.03	< 0.001
	Spatial (G_{SPAT})	0.83	0.62	0.997
	Species and spatial (G_{BOTH})	0.86	0.58	< 0.001
	All (G_{FULL})	1.00	0.60	0.003
Mortality	None (M_{CONST})	–	–	< 0.001
	Species (M_{SPEC})	0.10	0.04	0.906
	Spatial (M_{SPAT})	0.51	0.01	< 0.001
	Species and spatial (M_{BOTH})	0.77	0.07	0.069
	All (M_{FULL})	0.94	0.07	0.025
Harvesting	None (H_{ZERO})	–	–	0.953
	Mean (H_{CONST})	–	−0.01	0.047
	Spatial (H_{SPAT})	–	−0.04	< 0.001

*Variance in growth or mortality that is explained by a given formulation of the demographic models.

†Mean difference in log-biomass variance explained compared to the 'None' version of each demographic model. The mean is calculated across all combinations of models for the other demographic processes.

‡Sum of Akaike weights ($\Sigma\omega_i$) for biomass models that include the given demographic model. This metric reflects the weight of evidence for each formulation being included in the best biomass model within the set, and is interpreted as a measure of relative importance within each demographic process.

biomass, but that species variation in mortality rates contributed to differences in biomass as well. Six biomass models had a $\Delta_{AIC} < 10$, with each explaining 72–73% of the variance in log biomass. The growth model that included spatial but not species variation (G_{SPAT}) was in each of the six top-performing models, and was unambiguously supported over other growth formulations ($\Sigma\omega_i = 0.997$; Table 2). Across
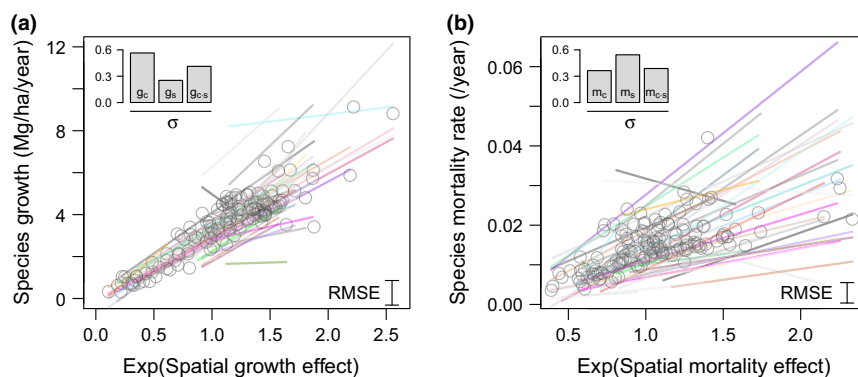
Figure 3 Trends in species- and cluster-level performance across gradients in growth and mortality defined by estimated spatial effects. (a) The growth trends of most species (coloured lines) converged around a common log-linear relationship with estimated spatial growth effects, suggesting that growth variation among species was relatively weak and that species tended to respond to environmental conditions in a common manner. This pattern reflected greater variation in estimated spatial effects than in species or residual effects (barplot inset). Because species effects on growth were relatively weak, cluster-level growth (circles) was strongly correlated with spatial effects. (b) Species-level mortality rates tended to increase with estimated spatial mortality effects, but unlike growth, there was large variation among species (i.e. different elevations for species-level trend lines). For mortality, variation in species effects was larger than that in spatial or residual effects (inset), producing a weaker correlation between spatial effects and cluster-level mortality (circles). In each panel, the error bar at the bottom right shows the average root mean square error (RMSE) around the species-level trends.
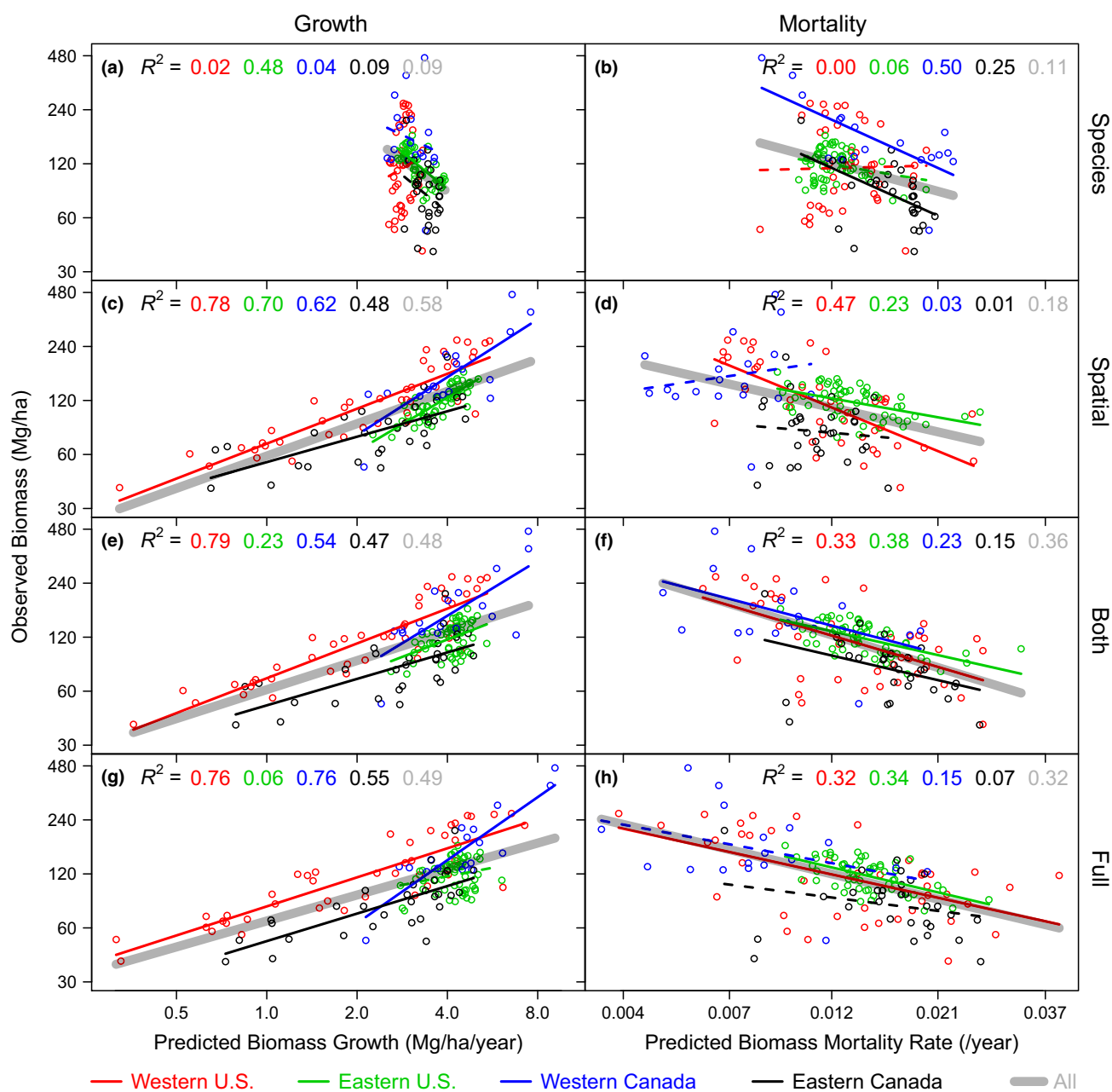
Figure 4 Relationships between forest biomass, growth and mortality across North America. Circles represent individual plot clusters. Solid lines represent linear relationships with non-zero slopes (based on 95% credible interval) for individual regions and for all regions together (grey); dashed lines represent linear relationships whose true slope may be zero (within 95% credible interval). Spatial effects on growth were strongly related to biomass in all four regions (c), but neither the additional consideration of species variation (e) nor the inclusion of residual effects (g) improved these relationships further. Relationships between mortality and biomass were strongest when both spatial and species effects were accounted for (f). Residual effects failed to improve mortality–biomass relationships (h), while ignoring species variation (d) produced weaker relationships everywhere but in the western United States.

the full set of models, spatial variation in growth increased the average variance explained by $\Delta R^2 = 0.62$ compared to models with a constant growth term. Three different formulations of mortality were represented in the top models (M_{SPEC}, M_{BOTH}, M_{FULL}, with $\Sigma\omega_i = 0.906$ for M_{SPEC}; Table 2), all of which included differences among species. There was only a small chance that spatial or residual mortality effects also helped in predicting biomass. The average increase in variance explained by these mortality terms ranged from $\Delta R^2 = 0.04$ to 0.07 compared to a constant mortality rate. Constant and

zero harvest rates were each included in three of the top six models ($\Sigma\omega_i = 0.953$ for H_{ZERO}). Spatial variation in harvest intensity marginally weakened the fit to biomass (Table 2). The age-dependent biomass model whose raw predictions best fit the biomass data contained the terms G_{SPAT}, M_{SPEC} and H_{CONST}, and had a root mean square error (RMSE) of 30.0 Mg/ha (Appendix F).

Supplementary analyses showed that the amount of variance explained by individual effects could vary somewhat when plots were aggregated into different numbers of clusters

(Appendix D). For instance, spatial growth effects showed the largest sensitivity to cluster size, with the average increase in log-biomass variance explained ranging from 0.68 to 0.51 when 50 or 500 clusters were created respectively. Despite this variability, the main patterns described above were robust across a wide range of cluster sizes.

Environmental predictors of spatial demographic effects

In both western and eastern Canada (plus Alaska), growth was most strongly related to mean annual temperature (R^2 = 0.50, 0.34 respectively; Appendix G), likely reflecting the effect of shorter growing seasons. Mortality rates in Canada were related to the silt content of the topsoil (R^2 = 0.32 in west, 0.29 in east), but these spatial mortality effects did not have clear relationships with biomass, as noted previously (Fig. 4d). Both spatial growth and mortality effects in the western United States were related to mean annual precipitation (R^2 = 0.54, 0.28 respectively), and in turn had important effects on forest biomass (Fig. 4c,d). In the eastern United States, spatial effects on growth were related to the silt content of the topsoil (R^2 = 0.49), which was highest in the fertile valleys of the Ohio and Mississippi rivers. Mortality effects in this region were positively related to mean annual temperature (R^2 = 0.49), reflecting an apparent trend for trees to exhibit higher mortality rates towards their southern range limits.

DISCUSSION

Global and continental-scale studies have documented that forest biomass varies along gradients in temperature and precipitation (Keith *et al.* 2009; Larjavaara & Muller-Landau 2012; Liu *et al.* 2014), being particularly high in areas with a cool and moist climate such as the Pacific Northwest. However, we have a weaker understanding of how this broad-scale variation in forest biomass arises from underlying demographic processes (Pan *et al.* 2013). Using demographic observations from continental networks of forest plots, we have estimated the contributions of both spatial- and species-related variation in growth and mortality to forest biomass. Our results show that, at a broad scale, forest biomass across North America appears to be related primarily to environmental controls over growth, and secondarily to variation in mortality rates arising from both local environmental conditions and species composition. These findings provide novel insights for understanding the demographic processes that determine continental-scale variation in forest biomass.

Forest biomass is known to be strongly related to productivity in boreal and temperate forests (Keeling & Phillips 2007). Spatial variation (independent of changes in species composition) accounted for the overwhelming share of variation in forest growth across North America, and explained more than half of the variation in forest biomass. This spatial growth variation mapped well onto environmental factors that limit growth within different parts of North America: temperature in Canadian forests where growing seasons are short, precipitation in the semi-arid western United States and soil texture in the eastern United States, where neither precipitation nor temperature are a strong limiting factor. These results support empirical and mechanistic models that have been used to predict vegetation productivity based on Liebig's law of the minimum (Leith 1975; Yuan *et al.* 2007). Moreover, our findings indicate that inherent spatial variation in growth is a major control over forest biomass across North America.

We detected remarkably little variation in growth among different species after accounting for spatial effects, however. While biomass growth certainly differs among forest types in North America (Gower *et al.* 1997; Jenkins *et al.* 2001), at broad scales species composition appears not to be an important predictor of growth after controlling for intrinsic differences in environmental conditions. This can be explained by at least two considerations. First, interspecific variation in productivity in a given area tends to be smaller than intraspecific variation across environmental gradients (Reich *et al.* 1997; Hudiburg *et al.* 2009). This was indeed the case, as our hierarchical models showed that spatial effects on growth were stronger than the corresponding species effects (Fig. 3a). Second, species tend to be more abundant on soil types or topographic positions on which they exhibit their best performance (Canham *et al.* 2006; Zhang *et al.* 2014). Interspecific differences in growth are therefore expected to be lower across a heterogeneous landscape than they are at individual sites because species occur more often in sites where they grow well. Averaging performance across many individual plots that make up a cluster likely masks site-level differences among species that produce variation in species composition within a given landscape (Whittaker 1956). As such, our findings concerning the demographic controls over forest biomass at broad scales (among landscapes) do not imply anything about their importance at finer scales (among sites within a landscape, or in different microenvironments).

A combination of spatial and species differences in mortality explained significant additional variation in forest biomass. Spatial environmental effects on mortality were most important in the western and eastern United States (Fig. 3d), where mortality was related to climatic gradients in precipitation and temperature respectively. Water deficit acts as a major control over forest mortality and biomass in the western United States, both at present and in projections of future climate change (Van Mantgem *et al.* 2009; Jiang *et al.* 2013). In the eastern United States, geographic mortality patterns have been previously related to mean annual temperature (Lines *et al.* 2010). These temperature-related patterns appear to determine the southern range limits of several different forest types (Vanderwel *et al.* 2013).

Relationships between mortality and forest biomass were strengthened when species variation was considered in addition to spatial effects. For example, in eastern Canada, mortality rates were positively related to the abundance of spruce (*Picea* spp.) and balsam fir (*Abies balsamea*) – not because these species tend to occur in climatically unfavourable northern forests, but rather because they have high mortality rates as hosts of the spruce budworm (*Choristoneura fumiferana*), a major biotic disturbance agent in eastern boreal forests. In western forests of both Canada and the United States, mortality increased with the relative abundance of cottonwood and aspen (*Populus* spp.), and decreased with that of Douglas fir (*Pseudotsuga menziesii*). Although both are considered to be early successional species groups, Douglas fir has adaptations

for crown maintenance and fire resistance that allow it to survive for over a thousand years (Ishii & Ford 2002). By contrast, western aspen is much more susceptible to competition and tends to accumulate pathogens over time, leading to rapid deterioration as stands reach about 120 years in age (Mueggler 1989). A more general understanding of species mortality effects proved elusive, however, as they appeared to be unrelated to either wood density or shade tolerance (Appendix H).

Compared with growth, tree mortality is less directly related to resource acquisition and its environmental controls, and more strongly influenced by species-specific strategies for coping with pests, pathogens and natural disturbances. As a result (and in contrast to our findings for growth), species differences explained appreciable variance in mortality rates. Much of this variation among species arises from the multitude of factors that may threaten a tree's survival, including competition, windstorms, drought, fire, insects and disease. Tree species exhibit a range of adaptations to such processes, and in turn these can lead to striking differences in the mortality responses of co-occurring species (Canham *et al.* 2001; Mueller *et al.* 2005). Variation in mortality determines the composition and development of forests disturbed by wind or fire (Papaik & Canham 2006; Seidl *et al.* 2014), and has been identified as the dominant process governing long-term forest succession in the US Lake States region (Purves *et al.* 2008). Large-scale mortality events can have dramatic impacts ecosystem dynamics, as evidenced by recent bark beetle outbreaks in western North America (Edburg *et al.* 2012), by drought-induced shifts from forest to pinyon-juniper woodland in the southwestern United States (Allen & Breshears 1998), and by the loss of a foundation species (eastern Hemlock, *Tsuga canadensis*) from Appalachian forests (Ford & Vose 2007). Our results further demonstrate that these interspecific differences in mortality rates are an important factor regulating variation in forest biomass across North America.

CONCLUSIONS

Our results provide valuable insights into how dynamic vegetation models (DVMs) may be refined to more accurately predict the live carbon balance of terrestrial ecosystems. Current DVMs typically represent global plant functional diversity with ~ 10 broadly defined plant functional types (PFTs). However, roughly the same amount of functional trait variation occurs among species within these PFTs as occurs among PFTs (Kattge *et al.* 2011). We show not only that there is considerable interspecific variation in mortality rates that cannot be explained by spatial environmental gradients, but also that this species-level variation has significant consequences for forest biomass, an important store of terrestrial carbon. Across current DVMs, uncertainty in future changes in carbon residence time (primarily a function of tree mortality rates) is the largest source of uncertainty in predicted carbon storage (Friend *et al.* 2013). While there has been recent progress in understanding drought-induced mortality (McDowell *et al.* 2013), mechanisms underlying species variation in tree mortality remain poorly understood in general. New modelling frameworks that appropriately capture species differences in mortality and biomass turnover could help in predicting future changes in terrestrial carbon storage.

AUTHORSHIP

MV and JL designed the study; MV, HZ, JP and GK compiled the data; MV performed modelling work and wrote the manuscript; all authors contributed substantially to revisions.

REFERENCES

Allen, C.D. & Breshears, D.D. (1998). Drought-induced shift of a forest-woodland ecotone: rapid landscape response to climate variation. *Proc. Natl Acad. Sci. USA*, 95, 14839–14842.

Bonan, G.B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science*, 320, 1444–1449.

Burnham, K.P. & Anderson, D.R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn.. Springer, New York, p. 488.

Canadian Council of Forest Ministers (2005). Wood Supply in Canada, 2005 report. Canadian Forest Service. Available at: https://cfs.nrcan.gc.ca/publications?id=25854. Last accessed 9 July 2014.

Canham, C.D., Papaik, M.J. & Latty, E.F. (2001). Interspecific variation in susceptibility to windthrow as a function of tree size and storm severity for northern temperate tree species. *Can. J. For. Res.*, 31, 1–10.

Canham, C.D. , Papaik, M.J., Uriarte, M., McWilliams, W.H., Jenkins, J.C. & Twery, M.J. (2006). Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. *Ecol. Appl.*, 16, 540–554.

Coomes, D.A., Flores, O., Holdaway, R., Jucker, T., Lines, E.R. & Vanderwel, M.C. (2014). Wood production response to climate change will depend critically on forest composition and structure. *Glob. Chang. Biol.*, 20, 3632–3645.

Edburg, S.L., Hicke, J.A., Brooks, P.D., Pendall, E.G., Ewers, B.E., Norton, U. *et al.* (2012). Cascading impacts of bark beetle-caused tree mortality on coupled biogeophysical and biogeochemical processes. *Front. Ecol. Environ.*, 10, 416–424.

Fauset, S., Baker, T.R., Lewis, S.L., Feldpausch, T.R., Affum-Baffoe, K., Foli, E.G. *et al.* (2012). Drought-induced shifts in the floristic and functional composition of tropical forests in Ghana. *Ecol. Lett.*, 15, 1120–1129.

Fisher, J.I., Hurtt, G.C., Thomas, R.Q. & Chambers, J.Q. (2008). Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecol. Lett.*, 11, 554–563.

Ford, C.R. & Vose, J.M. (2007). *Tsuga candensis* (L.) Carr. mortality will impact hydrologic processes in southern Appalachian forest ecosystems. *Ecol. Appl.*, 17, 1156–1167.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V. *et al.* (2006). Climate-carbon cycle feedback analysis: results from the (CMIP)-M-4 model intercomparison. *J. Climate*, 19, 3337–3353.

Friend, A.D., Lucht, W., Rademacher, T.T., Keribin, R., Betts, R., Cadule, P. *et al.* (2013). Carbon residence time dominates uncertainty

in terrestrial vegetation responses to future climate and atmospheric CO_2. *Proc. Natl Acad. Sci. USA*, 111, 3280–3285.

Gower, S.T., Vogel, J.G., Norman, J.M., Kucharik, C.J., Steele, S.J. & Stow, T.K. (1997). Carbon distribution and aboveground net primary production in aspen, jack pine, and black spruce stands in Saskatchewan and Manitoba, Canada. *J. Geophys. Res.*, 102, 29029–29041.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, 25, 1965–1978.

Hudiburg, T., Law, B., Turner, D.P., Campbell, J., Donato, D. & Duane, M. (2009). Carbon dynamics of Oregon and Northern California forests and potential land-based carbon storage. *Ecol. Appl.*, 19, 163–180.

Ishii, H. & Ford, E.D. (2002). Persistence of *Pseudotsuga menziesii* (Douglas-fir) in temperate coniferous forests of the Pacific Northwest coast, USA. *Folia Geobot.*, 37, 63–69.

Jenkins, J.C., Birdsey, R.A. & Pan, Y. (2001). Biomass and NPP estimation for the mid-Atlantic region (USA) using plot-level forest inventory data. *Ecol. Appl.*, 11, 1174–1193.

Jiang, X., Rauscher, S.A., Ringler, T.D., Lawrence, D.M., Williams, A.P. & Allen, C.D. (2013). Projected future changes in vegetation in western North America in the twenty-first century. *J. Climate*, 26, 3671–3687.

Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bonisch, G. et al. (2011). TRY - a global database of plant traits. *Glob. Chang. Biol.*, 17, 2905–2935.

Keeling, H.C. & Phillips, O.L. (2007). The global relationship between forest productivity and biomass. *Glob. Ecol. Biogeogr.*, 16, 618–631.

Keith, H., Mackey, B.G. & Lindenmayer, D.B. (2009). Re-evaluation of forest biomass carbon stocks and lessons from the world's most carbon-dense forests. *Proc. Natl Acad. Sci. USA*, 106, 11635–11640.

Kucharik, C.J., Foley, J.A., Delire, C., Fisher, V.A., Coe, M.T., Lenters, J.D. et al. (2000). Testing the performance of a dynamic global vegetation model: water balance, carbon balance, and vegetation structure. *Global Biogeochem. Cy.*, 14, 795–825.

Larjavaara, M. & Muller-Landau, H.C. (2012). Temperature explains global variation in biomass among humid old-growth forests. *Glob. Ecol. Biogeogr.*, 21, 998–1006.

Leith, H. (1975). Modelling the primary production of the world. In: *Primary Productivity of the Biosphere* (eds Leith, H. & Whittaker, R.H.). Springer, Berlin, pp 237–263.

Lines, E.R., Coomes, D.A. & Purves, D.W. (2010). Influences of forest structure, climate and species composition on tree mortality across the eastern US. *PLoS ONE*, 5, e13212.

Liu, S., Wei, Y., Post, W.M., Cook, R.B., Schaefer, K. & Thornton, M.M. (2013). The unified North American soil map and its implication on the soil organic carbon stock in North America. *Biogeosciences*, 10, 2915–2930.

Liu, Y., Yu, G., Wang, Q. & Zhang, Y. (2014). How temperature, precipitation and stand age control the biomass carbon density of global mature forests. *Glob. Ecol. Biogeogr.*, 23, 323–333.

McDowell, N.G., Fisher, R.A., Xu, C., Domec, J.C., Holtta, T., Mackay, D.S. et al. (2013). Evaluating theories of drought-induced vegetation mortality using a multimodel-experiment framework. *New Phytol.*, 200, 304–321.

Mueggler, W.F. (1989). Age distribution and reproduction of Intermountain aspen stands. *West. J. Appl. For.*, 4, 41–45.

Mueller, R.C., Scudder, C.M., Porter, M.E., Trotter, R.T. III, Gehring, C.A. & Whitham, T.G. (2005). Differential tree mortality in response to severe drought: evidence for long-term vegetation shifts. *J. Ecol.*, 93, 1085–1093.

O'Connell, B.M., LaPoint, E.B., Turner, J.A., Ridley, T., Boyer, D., Wilson, A.M. et al. (2013). The Forest Inventory and Analysis database: data description and user guide for phase 2 (version 6.0.1). Available at: http://www.fia.fs.fed.us/library/database-documentation/current/ver6.0/FIADB%20User%20Guide%20P2_6-0-1_final.pdf. Last accessed 30 April 2015.

Olson, J.S. (1963). Energy storage and the balance of producers and decomposers in ecological systems. *Ecology*, 44, 322–331.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A. et al. (2011a). A large and persistent carbon sink in the world's forests. *Science*, 333, 988–993.

Pan, Y., Chen, J.M., Birdsey, R., McCullough, K., He, L. & Deng, F. (2011b). Age structure and disturbance legacy of North American forests. *Biogeosciences*, 8, 715–732.

Pan, Y., Birdsey, R.A., Phillips, O.L. & Jackson, R.B. (2013). The structure, distribution, and biomass of the world's forests. *Ann. Rev. Ecol. Evol. Syst.*, 44, 593–622.

Papaik, M.J. & Canham, C.D. (2006). Species resistance and community response to wind disturbance regimes in northern temperate forests. *J. Ecol.*, 94, 1011–1026.

Purves, D.W., Lichstein, J.W., Strigul, N. & Pacala, S.W. (2008). Predicting and understanding forest succession using a simple tractable model. *Proc. Natl Acad. Sci. USA*, 105, 17018–17022.

Reich, P.B., Grigal, D.F., Aber, J.D. & Gower, S.T. (1997). Nitrogen mineralization and productivity in 50 hardwood and conifer stands on diverse soils. *Ecology*, 78, 335–347.

Seidl, R., Rammer, W. & Spies, T.A. (2014). Disturbance legacies increase the resilience of forest ecosystem structure, composition, and functioning. *Ecol. Appl.*, 24, 2063–2077.

Smith, W.B., Miles, P.D., Perry, C.H. & Pugh, S.A. (2009). Forest resources of the United States, 2007. General Technical Report WO-78. US Department of Agriculture Forest Service, Washington DC, USA.

Smith, M.J., Purves, D.W., Vanderwel, M.C., Lyutsarev, V. & Emmott, S. (2013). The climate dependence of the terrestrial carbon cycle, including parameter and structural uncertainties. *Biogeosciences*, 10, 583–606.

Stan Development Team (2014). Stan: a C++ library for probability and sampling, version 2.3. Available at: http://mc-stan.org. Last accessed 30 Apr 2015.

Stephenson, N.L. & van Mantgem, P.J. (2005). Forest turnover rates follow global and regional patterns of productivity. *Ecol. Lett.*, 8, 524–531.

Stephenson, N.L., vanMantgem, P.J., Bunn, A.G., Bruner, H., Harmon, M.E., O'Connell, K.B. et al. (2011). Causes and implications of the correlation between forest productivity and tree mortality rates. *Ecol. Monogr.*, 81, 527–555.

Van Mantgem, P.J., Stephenson, N.L., Byrne, J.C., Daniels, L.D., Franklin, J.F., Fule, P.Z. et al. (2009). Widespread increase of tree mortality rates in the western United States. *Science*, 323, 521–524.

Vanderwel, M.C., Lyutsarev, V.S. & Purves, D.W. (2013). Climate-related variation in mortality and recruitment determine regional forest-type distributions. *Glob. Ecol. Biogeogr.*, 22, 1192–1203.

Whittaker, R.H. (1956). Vegetation of the Great Smoky Mountains. *Ecol. Monogr.*, 26, 1–80.

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L.L., Baldocchi, D. et al. (2007). Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agric. For. Meteorol.*, 143, 189–207.

Zhang, T., Lichstein, J.W. & Birdsey, R.A. (2014). Spatial and temporal heterogeneity in the dynamics of eastern US forests: implications for developing broad-scale forest dynamics models. *Ecol. Model.*, 279, 89–99.

SUPPORTING INFORMATION