Supporting Information for;

“Does climate directly influence NPP globally?”

Chengjin Chu, Megan Bartlett, Youshi Wang, Fangliang He, Jacob Weiner, Jérôme Chave, Lawren Sack

Table S1. (Excel spreadsheet). Models for predicting annual net primary productivity (NPP) for ecosystems worldwide from stand biomass ($M_{\text{tot}}$), stand age ($a$), and climate variables representing temperature ($T$), precipitation ($P$) and growing season length ($l_{\text{gs}}$). The models were run in four configurations: A, using annual mean temperature and precipitation ($T_{\text{ann}}$ and $P_{\text{ann}}$); B using growing season mean temperature and precipitation ($T_{\text{gs}}$ and $P_{gs}$); and C and D, testing NPP previously divided by $l_{gs}$, as done by Michaletz et al. to control out the influence of $l_{gs}$. For each model we provide the slope and $R^2$ of observed plotted against predicted values, the number of parameters, and the Akaike Information Criterion corrected for small sample sizes (AICc), and its difference relative to the model number 8 or 16 which did not include any climate variables. Models in bold face were selected with the lowest AICc values by at least 2. Note that models including climate are strongly selected over models without climate variables. Parameter values with standard errors are provided in columns to the right. To the right are the values predicted for NPP by models including stand age and biomass and $T$, $P$ or $l_{gs}$ alone, with other variables assigned their mean values in the dataset, and $T$, $P$ or $l_{gs}$ assigned the highest and lowest values in the dataset; we thus estimated the shift in NPP driven by these climate variables all else being equal, according to these models. Note the substantial response of NPP to shifts in individual climate variables. Because the climate variables were correlated with each other, applying the other models with multiple climate variables to assess the influence of climate variables would require simultaneous shifting of all input climate variables; doing that also indicates a strong influence of climate on NPP. Columns far to the right contain the results of Independent Effects Analysis (IEA) applied to selected models of NPP, including the full model (eqn M3), and models removing terms (i.e., removing stand age, or stand biomass, or climate variables). Unlike partial correlations, the IEA robustly partitions variance in the dependent variable among correlated predictor variables. Models 1-8 include the Michaletz et al. model and more parsimonious versions based on removing climate variables; models 9-16 are the same, but removing the stand age ($a$) term, which resulted in circularity in the model, predicting NPP from ABC (see text); model 17 is eqn M3 without including $M_{\text{tot}}$ or $a$. The analysis was applied for models using annual or growing season means for temperature and precipitation, and also for models predicting NPP divided by the length of the growing season.
Table S2. Structural equations modeling of net primary productivity (NPP) across 1247 ecosystems worldwide in the dataset of Michaletz et al. (2014) analyzed by stand biomass and climate variables, as shown in Fig. 7. Stand structure and climate variables include stand biomass ($M_{\text{tot}}$), length of growing season ($l_{gs}$), and mean annual or mean growing season temperature and precipitation ($T_{\text{ann}}$, $T_{gs}$, $P_{\text{ann}}$ and $P_{gs}$ respectively). For each term, direct effects on NPP are given in bold-face (with % of total effect that is direct), and indirect effects. Model A was based on the hypothesis of Michaletz et al. (eqn M3) that climate variables would have direct effects on NPP or indirect effects via $M_{\text{tot}}$ and stand age ($a$); that model was not supported, likely due to the circularity of including $M_{\text{tot}}$ and $a$ (and therefore stand-age averaged biomass change, ABC) as a predictor of NPP. Model B was the same model without including the $a$ term, and Model C had additional influence of $l_{gs}$ on $T$ and $P$ (as explained in the text). Each model was tested using annual or growing season mean temperature and precipitation. Support for the models was determined using the Chi-square test; when a model was not supported, we did not determine the coefficients. The best supported model by AIC was Model B using growing season mean climate variables (underlined); 94% of the effects of climate on NPP were direct. In all models all variables but temperature were log-transformed, as by Michaletz et al. Note that the results are sensitive to transformation of factors, but alternative transformation did not change the overall conclusion that climate had a strong total influence on NPP, equal to or greater than $M_{\text{tot}}$, and the direct climate effects were equal to or greater than any indirect effects via $M_{\text{tot}}$. *All of the positive effects on NPP were direct, given that the indirect effect was negative. All effects were significant at $P < 0.05$ except † $P \leq 0.075$ and ‡ $P = 0.33$.

<table>
<thead>
<tr>
<th>Model</th>
<th>annual (ann)</th>
<th>Support or growing season (gs)</th>
<th>AIC</th>
<th>$R^2$</th>
<th>direct effects on NPP in bold face (with % of total effect that is direct), and indirect effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>climate used</td>
<td></td>
<td></td>
<td>$M_{\text{tot}}$</td>
</tr>
<tr>
<td>A. eqn M3</td>
<td>a</td>
<td>N</td>
<td>na</td>
<td></td>
<td>na</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gs</td>
<td>N</td>
<td></td>
<td>na</td>
</tr>
<tr>
<td>B. eqn M3 without the $a$ term</td>
<td>a</td>
<td>Y</td>
<td>13100</td>
<td>0.51</td>
<td>0.493 0.19* -0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gs</td>
<td>Y</td>
<td>13045</td>
<td>0.53 0.546 0.31 (86%), 0.051</td>
</tr>
<tr>
<td>C. eqn M3 without the $a$ term and with addition $l_{gs}$ influence on $T$ and $P$</td>
<td>a</td>
<td>Y</td>
<td>13108</td>
<td>0.51</td>
<td>0.493 0.19 (40%), 0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gs</td>
<td>N</td>
<td></td>
<td>na</td>
</tr>
</tbody>
</table>
Appendix S1. Statistical methods

Correlation analyses
For testing correlations, we report the strongest relationship for untransformed data or log-transformed data, given that some relationships are apparently approximately linear, and others approximated by power-laws. We report $R^2$ values corresponding to the correlation coefficients (i.e., $r$ values) because one objective is to quantify how much variation in NPP is explained by individual drivers, and this is summarized by $R^2$.

Maximum likelihood model selection
We compared likelihood support for models with and without climate variables to evaluate the influence of climate on NPP. We used the optim package in R (v.3.1.0) to fit 17 models (Table A1). Models 1-8 are the Michaletz et al. model and more parsimonious versions based on removing climate variables, models 9-16 are the same but without the stand age ($a$) term to avoid the circularity inherent in predicting NPP from ABC (see text); and model 17 includes the three climate predictors but not stand biomass ($M_{tot}$) or $a$. The best-fit models were identified as those with Akaike’s Information Criterion (AIC) values, corrected for small sample size (AICc; $\text{AIC} + \frac{2K(K+1)}{n-K-1}$), within 2 units of the minimum AICc. ($K$ is the number of fitted parameters for each model and $n = 1247$ forests) (Burnham & Anderson, 2002). We compared the AICc value for model 8, which included only $M_{tot}$ and $a$ as predictors, with values for models 1-7, which included $M_{tot}$, $a$, and the climate variables; and repeated this comparison for model 16 and models 9-15, which are the same as models 1-9 with the $a$ term removed. We conducted these model comparisons in four ways to consider the range of treatments of climate variables used by Michaletz et al., by testing both NPP and NPP normalized by growing season length (NPP/l_{gs}) against mean average values and growing season averages for temperature and climate.

Independent Effects Analysis
To further test for direct impacts of climate on NPP, independently of $M_{tot}$ and $a$, we re-analyzed the effect of each variable in the Michaletz et al. eqn M3 on NPP using independent effects analysis instead of partial correlations. Independent effects analysis, which calculates the average contribution of each predictor variable to variance in the dependent variable across all possible models that represent a subset of the predictor variables (Murray & Conner 2009), is more appropriate and effective to separate out the independent relationships of correlated variables than partial correlation analysis, which only permits very limited inference. Michaletz et al. stated that “partial correlation plots show the correct strength of the relationship between the dependent variable and each independent variable while controlling for the influence of all other independent variables included in the model”. They further justified this partial correlation analysis because they performed a test for collinearity and found the variance inflation factors (VIFs), or the
increases in variance in estimated regression parameters due to correlations among predictor variables, were below critical thresholds for collinearity (Ryan, 1997). However, these argument are invalid for justifying the use of partial correlations to disentangle correlated drivers: as shown by the theoretical and simulation studies by Murray & Conner (2009), “partial correlations are not designed to partition the variance shared between multiple correlated predictors and the dependent variables.” The VIF test is typically used to justify including partially correlated variables within a multiple regression analysis, because each variable contains some independent information (Ryan, 1997), and not to justify partitioning the importance of factors using a partial correlation analysis, which cannot separate their independent effects. Independent effects analysis accurately partitions variance among correlated predictors, and thus can be used to determine variable importance, rather than partial correlation analysis (Murray & Conner, 2009; see section 3).

We calculated independent effects for each variable in the Michaletz et al. model using the hier.part package in R. We repeated this analysis for the model without $M_{tot}$, and then without $a$, to evaluate the impact of climate on NPP when the circular ABC term is removed. We calculated independent effects for these three models in four ways, using both NPP and NPP normalized by growing season length (NPP/$l_{gs}$) as dependent variables, and mean average values and growing season averages for temperature and climate as predictors.

Structural Equations Modeling

In SEM, researchers usually firstly propose a hypothetical model based on the a priori knowledge, then evaluate the model based on empirical data. We proposed three hypothetical models as following (Fig. S1): Model A was based on the hypothesis of Michaletz et al. that climate variables would have direct effects on NPP or indirect effects via $M_{tot}$ and stand age ($a$). Model B was the same model without including the $a$ term to avoid the circular term ABC ($=M_{tot} / a$; see main text). Model C had additional influence of $l_{gs}$ on $T$ and $P$, as expected from eqns 1a and 2a in the main text.
Figure S1. Structural equation models (SEM) for net annual productivity (NPP) determined by climate variables, stand biomass ($M_{\text{tot}}$) and stand age ($a$) applied to the 1247 ecosystems worldwide in the dataset of Michaletz et al. (2014); climate variables include length of growing season ($l_{\text{gs}}$), and mean annual or mean growing season temperature and precipitation ($T_{\text{ann}}$, $T_{\text{gs}}$, $P_{\text{ann}}$ and $P_{\text{gs}}$ respectively).

We tested models using $T_{\text{ann}}$ and $P_{\text{ann}}$, or using $T_{\text{gs}}$ and $P_{\text{gs}}$. We generally tested versions with temperature data untransformed and all other data log-transformed, for consistency with eqn M3, but our tests using additional transformations did not affect the overall findings for the importance of the direct role of climate on NPP, independently of $M_{\text{tot}}$. Support for the models were determined using the Chi-square test. The SEM analyses were conducted using the lavaan package (Rosseel, 2012) in R (v. 3.1.0).
Appendix S2. Supplementary Results: On estimating the influence of stand age on NPP

Michaletz et al. tested the correlation of NPP against selected climate variables, and stand biomass (M_{tot}). However, rather than directly testing the correlation of NPP with stand age (a), Michaletz et al. presented a more involved analysis, described below. When we performed a direct correlation test, we found that NPP was very weakly related to a and that this relationship arose only due to the higher NPP for stands < 50 years of age (Fig. 3a, b), much younger than would be suffering from stand age related decline in most natural woody ecosystems according to previous studies (Gower et al., 1996, Weiner & Thomas, 2001, Pregitzer & Euskirchen, 2004, He et al., 2012). Indeed, we found in an analysis of variance that while NPP was lower for stands < 50 years old than for the stands of older categories, it did not differ among the other stand age categories (Table S3). We found that the higher NPP for stands < 50 years of age in this database was likely to be related to their disproportionate sampling in moist, warm environments. Indeed, in this database, the stands 0-50 years in age differed significantly in climate from the others, with a l_{gs} of 1-2 months longer, and a mean T_{ann} of several degrees C higher than the three other age classes, and P_{ann} of > 100 mm greater than two of the three other age classes (Table S3). Indeed, even among the stands 0-50 years in age, a was negatively correlated with T_{ann}, P_{ann} and l_{gs}. We hypothesize that this bias arose because in the data available to Michaletz et al. for ecosystem NPP, more information was available for young stands fully developed enough to merit NPP measurement, which would be those with higher NPP, especially occurring in warm, moist climates.

Table S3. Mean values ± standard error NPP and climate variables for stands categorized by ages as conducted by Michaletz et al. (2014). Additionally, correlation coefficients are reported for stand age (a) versus each climate variable for ecosystems within each age category. The stand categories differed in their NPP and all climate variables (one-way ANOVAs; P < 0.001), and those aged 0-50 years had significantly higher T_{ann} and longer l_{gs} than all three other categories, and significantly higher P_{ann} than two of the three other age categories. Only for the youngest category of stands was a negatively related to all climate variables. This difference in climate would result in higher NPP for the youngest age category, and for the relationship of NPP to a across all ecosystems, not found for stands > 50 years old, in which age-related decline would be expected. Correlation coefficients are the strongest for untransformed or log-transformed data. *P<0.05; **P<0.01; ***P<0.001; ns>P>0.05. Same subscript letters in a given column indicate no significant difference in Tukey post-hoc comparisons at P < 0.05.

<table>
<thead>
<tr>
<th>Stand age (a)</th>
<th>NPP</th>
<th>T_{ann}</th>
<th>P_{ann}</th>
<th>l_{gs}</th>
<th>M_{tot}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-50 (570)</td>
<td>1175±26a</td>
<td>10.7±0.25a; -0.25*** 921±22a; -0.13**</td>
<td>8.6±0.12a; -0.26*** 11885±288a; 0.15***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-100 (415)</td>
<td>992±26b</td>
<td>8.0±0.40 b; -0.059** 811±19c; 0.25***</td>
<td>7.3±0.12b; -0.075ns 16429±521b; 0.51***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101-200 (235)</td>
<td>909±34b</td>
<td>5.75±1.51c; -0.092ns 801±29bc; 0.023ns</td>
<td>6.1±0.17bc; -0.065ns 27192±852c; 0.033ns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>201- (27)</td>
<td>993±134ab</td>
<td>8.67±0.26abc; 0.16ns 1038±130abc; 0.42* 7.0±0.60c; -0.17ns</td>
<td>42286±5588d; 0.64***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What led Michaletz et al. to their conclusion that a was a strong driver of NPP decline across ecosystems of all ages despite the very weak correlation of NPP to a, and that correlation evidently arising due to a confounding of a with climate in the youngest stand category? Michaletz et al. were apparently misled by their more involved analysis. They plotted NPP against M_{tot} for the ecosystems sorted into four age categories, and showed that the fitted lines for younger stand
ages were higher than those for older stand ages (reproduced in Fig. 2d of the main text of this paper). That analysis indicated that younger stands had higher NPP, at a given $M_{\text{tot}}$. However, this analysis had an automatic outcome, which in our view does not permit inference of an association between $a$ and NPP. Because $M_{\text{tot}}/a$ is equivalent to the stand age-averaged biomass change (ABC), which is mechanistically dependent on NPP and statistically correlated with NPP (Figs 1 and 4), plotting NPP against $M_{\text{tot}}$ will tend to produce higher NPP values for the younger stand categories, at a given $M_{\text{tot}}$. Such an analysis does not imply that $a$ influences NPP in this global dataset; these variables were actually very weakly related in the direct correlation analysis.

For illustration of the misleading nature of the analysis consider a simple example. One might collect data for vehicles on their distances $d$ and times $t$ travelled during the day, and calculate their average velocity ($v = d/t$). In a simulated dataset for 10 vehicles, the average $v$ was unrelated to $t$ (Fig. S2a). However, if one plots $v$ against $d$, one might at first sight conclude that the five vehicles driven for a smaller $t$ appear to have higher $v$ than those driven for a longer $t$ (Fig. S2b). That would be a spurious conclusion, because the higher velocities of the vehicles driven for a smaller $t$ arise only at a given $d$, an automatic result, because $v = d/t$. Note that the same result would arise even if $v$ were not directly calculated from $d$ and $t$, but a correlated measurement was used, e.g., data were collected from radar measurements. (In the same way, Michaletz et al. plotted NPP against $M_{\text{tot}}$ and considered shifts among categories of $a$ within this plot; these shifts were automatic, because ABC, a close correlate of NPP, is equal to $M_{\text{tot}}/a$).

![Figure S2](image.png)

**Figure S2.** Plots to illustrate the spurious analysis of Michaletz et al. (2014) from which they concluded that NPP was driven by stand age in their database. Simulated data were plotted for average velocity against driving time and driving distance for 10 hypothetical vehicles, with the open symbols representing those driven for < 5 hours per day, and the closed symbols those for > 5 hours per day. Note that despite no relationship of velocity to driving time (panel (a)), when velocity is plotted against distance, the vehicles driven for a shorter period of time have a higher mean velocity, at a given driving distance (shown by the separate fitted lines). This is an automatic result, which is spurious if used to infer that velocity is influenced by driving time for these vehicles; similarly, this analysis was spurious when used to infer that NPP was driven by stand age in the database of Michaletz et al. (2014).
Appendix S3: Derivation of eqn M3a

\[ \ln(\text{NPP}) = \alpha \ln(M_{\text{tot}}) + \alpha_a \ln(a) - \frac{E}{kT} + \alpha_p \ln(P) + \alpha_{lgs} \ln(lgs) + \beta_{0,1} \quad \text{eqn M3} \]

\[ \ln(\text{NPP}) = \ln(M_{\text{tot}}^\alpha \times a^{\alpha_a}) - \frac{E}{kT} + \alpha_p \ln(P) + \alpha_{lgs} \ln(lgs) + \beta_{0,1} \]

\[ \ln(\text{NPP}) = \ln\left(\frac{M_{\text{tot}}^\alpha}{a^{-\alpha_a}} \times M_{\text{tot}}^{\alpha + a_a}\right) - \frac{E}{kT} + \alpha_p \ln(P) + \alpha_{lgs} \ln(lgs) + \beta_{0,1} \]

\[ \ln(\text{NPP}) = -\alpha_a \ln\left(\frac{M_{\text{tot}}}{a}\right) + (\alpha + \alpha_a) \ln(M_{\text{tot}}) - \frac{E}{kT} + \alpha_p \ln(P) + \alpha_{lgs} \ln(lgs) + \beta_{0,1} \]

\[ \ln(\text{NPP}) = -\alpha_a \ln(ABC) + (\alpha + \alpha_a) \ln(M_{\text{tot}}) - \frac{E}{kT} + \alpha_p \ln(P) + \alpha_{lgs} \ln(lgs) + \beta_{0,1} \quad \text{eqn M3a} \]
Appendix S4. Supplementary Discussion: Relationship of NPP and ABC

NPP and ABC (\(= M_{\text{tot}}/a\)) are mechanistically related (Fig. 1), and statistically inter-correlated (Fig 4). Indeed, ABC is the stand-age average of net biomass change (NBC), and the relationship of NBC to NPP is well known. Notably, total ecosystem NBC is rarely determined, and it is more convenient to estimate the woody above-ground version, as aboveground woody NPP minus aboveground woody biomass loss (Malhi et al., 2015). A model used for estimating NBC (and especially that of the above-ground woody pool) is,

\[ \text{NBC} = \text{NPP} - M_{\text{tot}}/\tau \]

where \(\tau\) is the residence time of the biomass pool in years (Galbraith et al., 2013).

Notably, in addition to the correlation of NBC with NBC due to their mechanistic linkage (Fig. 1), the two may be additionally correlated in practice due to field methodology. In many studies, supposed measurements of NPP are in fact estimates of NBC; this occurs when woody NPP is estimated as the change in standing biomass between censuses, without correction. Such an approach in fact measures NBC, and underestimates NPP by up to approximately 25%; the loss of biomass due to tree mortality and decomposition should be added for the correct estimation of woody NPP (Clark et al., 2001a; Fig. 2 in that paper).

Despite being mechanistically linked (Fig. 1), one might expect that NPP and ABC could be decoupled for three reasons. First, ABC is averaged over the age of a stand, and thus while it might be expected to relate to stand-age averaged NPP, would it be related to recent annual NPP? This would indeed be expected, because a stand with higher average NPP should have higher NPP in a typical year, given that a higher NPP is driven by more productive climate, soil type and/or species (Chapin et al., 2012); many studies show that across ecosystems, NPP values are correlated across years and decades, and when averaged over different time scales (Mohamed et al., 2004, Sang & Su, 2009, Gong et al., 2012, Zhu & Southworth, 2013). A second reason for potential decoupling of NPP and ABC is if heterotrophic respiration \(R_h\) varies strongly across ecosystems independently of NPP. However, studies have shown that \(R_h\) and NPP are positively correlated across ecosystems globally, as are NPP and biomass turnover rates (i.e., mortality and recruitment), thinning, and disturbance rates (Stephenson & van Mantgem, 2005, Keeling & Phillips, 2007, Chapin et al., 2012, Cole et al., 2014; Brienen et al., 2015). Finally, and most importantly, the relationships of NPP, NBC and ABC might be expected to hold only for stands of a limited range of age. For example, one would expect ABC to correlate with NPP especially strongly during intervals of succession after a large disturbance, when a stand is still increasing in mass. However, for old-growth stands in “equilibrium” (i.e., if NPP = heterotrophic respiration and thus NBC is approximately zero and \(M_{\text{tot}}\) remains relatively stable) or in decline (i.e., if NPP < heterotrophic respiration, and thus NBC < 0 and \(M_{\text{tot}}\) is declining) one might imagine that ABC might become decoupled from NPP; NPP could remain the same while \(M_{\text{tot}}\) could remain the same or decline, and thus ABC=\(M_{\text{tot}}/a\) inevitably declines as the stand becomes older. Still, it would take many years of equilibrium or stand decline, relative to the years of biomass accumulation, for the signal
of a high NPP and a high NBC to disappear from ABC completely across such stands. Thus, we found in the Michaletz et al. database that the correlation of ABC with NPP (log-transformed data) declined across the four age classes: it was highest for stands 0-50 years old ($R^2 = 0.85; n = 569; P < 0.001$), even higher than for all the data combined across age categories (Fig. 4), but remained strong for stands 51-100 and 101-200 years old (respectively $R^2 = 0.73, n = 414; P < 0.001$ and $R^2 = 0.41, n = 234; P < 0.001$) and even for stands 201+ years old ($R^2 = 0.32; n = 26; P = 0.002$).

Notably, NPP and ABC showed a curvilinear relationship (Fig. 4). We propose three potential mechanisms, related to the fact that stands with higher $M_{tot}$ tend to have higher NPP and ABC (Fig. 2). First, as explained above, in older stands with higher NPP, there is a greater likelihood of equilibrium or stand decline, weakening the relationship of ABC to NPP. Second, stands with larger $M_{tot}$ and longer stand history will tend to have experienced more disturbance or thinning episodes, such that the stand age averaged NPP (and thus ABC) will be more variable and lower relative to recently measured NPP. Third, amplifying this effect, stands of higher $M_{tot}$ tend to have a greater fraction of biomass stored in large trees (Slik et al., 2013), and given the disproportionate biomass loss that would arise from their death and decomposition, stands with higher $M_{tot}$ and NPP would at least sometimes have very high average $R_h$, leading to a weaker relationship of ABC and NPP at high NPP.

Regardless of the shape of the relationship between NPP and ABC and its specific explanation, their tight mechanistic and statistical correlation (Figs 1 and 4) mean that using ABC to predict NPP creates a strong relationship, and one which impedes the resolution of other predictive factors, including climate. Models for NPP including $M_{tot}$ and $a$ as multiplicative factors, and thus using ABC to predict NPP, include a circular term that produces a “spurious correlation”. The term “spurious correlation” has been applied previously to the situation when a ratio variable $Y/X$ is plotted against the numerator or denominator, i.e., against $Y$ or $X$, and a high correlation results automatically, regardless of the true underlying relationships among variables (Brett, 2004). The apparent predictive power is trivial, and if other predictor variables are included in such a model, their explanatory power will appear to be low. In this case of predicting NPP from a model that includes $M_{tot}$ and $a$ as multiplicative predictors (=ABC) one is predicting a ratio using a proxy for both its numerator and denominator. This circularity leads to the high apparent predictive power of the model ($R^2 = 0.77$), given the high degree of correlation of NPP to ABC $R^2 = 0.71$; see main text). The fit of such a model does not constitute evidence that the input variables drive the output variables, and such a model does not permit resolution of the factors with influence let alone their relative strengths.

By analogy, one might construct a multiplicative model for the velocity of vehicles ($v$), and include as multiplicative predictor variables, distance travelled ($d$), time ($t$), maximum engine capacity, driver experience, weather conditions, and road traffic. One would inevitably find a strong prediction of $v$, arising simply because $v = d / t$. The same is true if the $v$ one predicts was not measured as $d / t$, but using another method, e.g., radar measurements at random times. One would also find that that $d$ and $t$ are sufficient to explain $v$, with no apparent additional influence
of engine speed, driver experience, weather conditions, or traffic—even if those factors in fact had real, direct influence on \( v \) for the sample of vehicles. The model has a circularity, and one cannot reasonably conclude that the other factors have no direct influence, or that the other factors only indirectly influence velocity, via \( d \) or \( t \). This is an illustration of the error that led Michaletz et al. to conclude from their multiple regression analysis of eqn M3 that climate variables had a negligible direct influence on NPP, and that NPP was effectively only a function of stand mass and age. We require modeling that does not include hidden circular terms when attempting to elucidate driving variables.
Appendix S5: Improvements needed for datasets and analyses of drivers of NPP globally

For disentangling the drivers of NPP, improvements are needed both in the quality of the dataset and in the analyses. First, the determination of NPP is highly complex, and efforts are needed to account for the effects and potential biases of different designs and field methodologies. In particular, ongoing attention is needed to improve the estimation of below-ground NPP (Clark et al., 2001b, Malhi et al., 2015). Notably, stand age is not necessarily a meaningful variable, not just due to the difficulty of accurate estimation, or what point to take as the origin of the stand, but also because of the varied history of most natural stands, and the difficulty or impossibility of accounting for disturbances and partial harvesting. Michaletz et al. employed very coarse estimates; e.g., for some tropical forests, estimating stand age from biomass and a mean carbon residence time value taken from a published study for a different stand, a study which had concluded that carbon residence time should not be considered as constant across species or stands (Vieira et al., 2005).

A larger problem with the database of Michaletz et al. is the opportunistic and thus haphazard sampling of ecosystems with respect to location, age and climate. The geographical representativeness of the database is questionable. For example, 89% of data points represented China, and only six data points of 1247 represented tropical rainforests (their Fig. 1a). This is far from a good global sample of species and functional trait diversity, or of substrates and climate, and we advocate for the development of a truly representative global database. Further, one must avoid sampling biases that act to confound factors; in this database, younger stands were disproportionately representative of warmer, moister climates (Appendix S2).

How can we move toward an ideal analysis? First, fitting a range of structural relationships for the effects of climate would be more appropriate than treating these as independent multiplicative power laws, given that climate influences have been shown to be linear, or a range of nonlinear shapes, including sigmoidal, or exponential (e.g., Luo et al., 2004). More importantly, we need to move beyond analyses of NPP and its drivers across global ecosystems that conflate multiple processes across scales. Clearly, individual ecosystems (and their component species) have distinctive responses of NPP to climate variables, stand biomass and age (Coomes et al., 2014, Shi et al., 2014). Pooling data for ecosystems haphazardly sampled across climates and ages, and determining an overall, general response of NPP to climate and stand variables does not permit a clear analysis of the mechanistic drivers of the overall relationships or of the individual ecosystem responses that underlie these relationships. Indeed, this is a similar problem of design as that faced by studies analyzing general allometries that combine intra-individual, intra-specific and/or inter-specific data; such pooling of data across scales does not allow high resolution of the mechanistic drivers of the relationships, which could be different at any of these scales (Weiner & Thomas, 1992). To truly resolve the drivers of NPP, a sophisticated approach would sample each ecosystem (or even its component species) across the range of climates and ages that exist. In such a case, the variation among ecosystems arising from differences in climate variables and age could be resolved reliably with a hierarchical model (Lichstein et al., 2014).
Appendix S6. Supplementary References

Brett MT (2004) When is a correlation between non-independent variables "spurious"? Oikos, 105, 647-656.


