Supplementary Information

Methods

1) Calculating the covariance of point observations to derive \( \sigma_x^2 \)

The basis for estimating \( \sigma_x^2 \) is a modified version of the formula that computes the variance of a given dataset, here dataset \( X \) of point variables \( x_i \) and \( x_j \) at location \( i \) and \( j \), (Zhang and Lei Cheng, 2012):

\[
\sigma_x^2 = \text{cov}(X, X) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{2} (x_i - x_j)^2
\]

(Eq.S1)

where \( N \) is the number of observations. The advantage of equation (SI 1) is that it is based on pairs of \( x \) and does not require the mean of \( X \). By calculating the average half-square difference of pairs of \( x_i \) and \( x_j \) at location \( i \) and \( j \) for a specified distance that corresponds to the size of a grid cell, we obtain an estimate of the average within-grid-cell variance of point observations \( \sigma_x^2 \). This mathematical representation is conceptually equivalent to the nugget effect in a semivariogram. This nugget effect is interpreted as the sum of variance caused by small-scale variability and observation errors (Cressie, 1993) and is thus fully consistent with our interpretation of \( \sigma_x^2 \).

2) Calculating the corrected correlation coefficient \( r_{corr} \) and the maximum achievable correlation coefficient \( r_{max} \)

The overall difference between two datasets \( X \) and \( Y \) of observed \( x \) and point values \( y \) at spatial location \( i \) is captured by the mean squared deviation \( E^2 \):

\[
E^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2
\]

(Eq.S2)

However, the convenience of expressing differences in a single figure comes at the cost of not being able to distinguish in which aspects (the mean, the “amplitude” meaning the variability or roughness of the pattern, or the “shape” of the spatial pattern) the two datasets differ. For this purpose, \( E^2 \) can be decomposed into different components. In the first step, the overall bias in the mean can be isolated:

\[
E^2 = E^{i2} + (\bar{x} - \bar{y})^2
\]

(Eq.S3)

where \( \bar{x} \) and \( \bar{y} \) are the arithmetic means of \( X \) and \( Y \), respectively, and \( E^{i2} \) is the mean squared difference of the centred patterns defined as:

\[
E^{i2} = \frac{1}{N} \sum_{i=1}^{N} ((x_i - \bar{x}) - (y_i - \bar{y})^2
\]

(Eq.S4)

\( E^{i2} \) is related to the correlation coefficient \( r \) and the variances \( \sigma_x^2 \) and \( \sigma_y^2 \) of \( X \) and \( Y \), respectively:

\[
E^{i2} = \sigma_x^2 + \sigma_y^2 - 2\sigma_x\sigma_y r
\]

(Eq.S5)
Eq. S5 is the underlying principle of the Taylor diagram (Taylor, 2001), in which the root-mean-square difference $E'$, the correlation coefficient $r$, and differences between standard deviations $\sigma_x$ and $\sigma_y$ are visualized in a concise and easily recognizable way. However, the above described relationship between $E'$ and the difference of $\sigma_x$ and $\sigma_y$ is ambiguous as an underestimation of standard deviation can lead to a higher or lower $E'$, depending on the value of $r$. Therefore, in the following, we will focus on the variance $\sigma_x^2$ and $\sigma_y^2$, and on the correlation coefficient $r$ and will show the use of the Taylor diagrams below.

Similar to the variance, errors also add quadratically so that a corrected estimate of the mean squared difference $E''^2$ is also obtained by subtracting $\sigma_x^2$:

$$E''^2_{\text{corr}} = E''^2 - \sigma_x^2$$  

(Eq. S6)

Thus, the root-mean-square difference $E''_{\text{corr}}$ is always smaller than $E''$. Due to the quadratic relationship, the increment $E'' - E''_{\text{corr}}$ is inversely related to $E''$, i.e. the correction is stronger for smaller $E''$.

The quantification of the effect of $\sigma_x^2$ on $r$ is less obvious but can be derived by rearranging equation Eq. S5 for $r$:

$$r = \frac{\sigma_x^2 + \sigma_y^2 - E''^2}{2\sigma_x\sigma_y}$$  

(Eq.S7)

By replacing $\sigma_x^2$ with $\sigma_{x,\text{corr}}^2$ and $E''^2$ with $E''^2_{\text{corr}}$, we can obtain an equation for the calculation of $r_{\text{corr}}$:

$$r_{\text{corr}} = \frac{\sigma_{x,\text{corr}}^2 + \sigma_y^2 - E''^2_{\text{corr}}}{2\sigma_{x,\text{corr}}\sigma_y}$$  

(Eq.S8)

By combining Eq. S7 and Eq. S8 we find that:

$$\frac{r_{\text{corr}}}{r} = \frac{\sigma_x}{\sigma_{x,\text{corr}}}$$  

(Eq.S9)

This means that $r$ is corrected by the same factor $f$ as $\sigma_y/\sigma_x$, but the implications are different. The correlation coefficient varies between +1 and -1, with +1 indicating total positive correlation, 0 indicating no linear correlation, and -1 indicating total negative correlation. Since the correction for $\sigma_x^2$ is equivalent to multiplying with a factor $f > 1$, $r_{\text{corr}}$ is always larger than $r$ if $r$ is positive and smaller if $r$ is negative. This means that the $\sigma_x^2$ equally affects positive and negative correlation and has no effect if $X$ and $Y$ are totally uncorrelated.

However, the correction does not have the potential to change the sign of $r$. There is an upper bound for positive $r$ and a lower bound for negative $r$ that can be achieved for a given $\sigma_x^2$. For the case $\sigma_{x,\text{corr}}^2 = \sigma_x^2 / 2$, these limits are $1/\sqrt{2}$ and $-1/\sqrt{2}$, respectively.

In this way, we can also calculate a “maximum achievable correlation coefficient” $r_{\text{max}}$. Assume, we correlate the observed dataset vs. the observed dataset we would get a correlation coefficient $r = 1$. If we now correct one of these datasets by $\sigma_x$, its correlation reduces by factor $\frac{\sigma_x}{\sigma_{x,\text{corr}}}$, which gives us $r_{\text{max}}$ similar to Eq. SI 9 with $r=1$.

This means that $r$ is corrected by the same factor $f$ as $\frac{\sigma_x}{\sigma_{x,\text{corr}}}$. Assuming that $r_{\text{corr}} = r \times f$ and $\sigma_{x,\text{corr}} = \sigma_x \times f$

then $= \frac{\sigma_{x,\text{corr}}}{\sigma_x}$. 

2
With $r=1$, it becomes

\[ r_{\text{corr}} = 1 \times \frac{\sigma_{x,\text{corr}}}{\sigma_x} = r_{\text{max}} \]  

(Eq. S10)
3) Description of the dynamic global vegetation models (DGVMs)

Below, we provide a short description for each of the four DGVMs applied in our study, which is taken from the Appendix S2 from Johnson et al. (2016):

ORCHIDEE

The ORCHIDEE model (Krinner et al., 2005) consists of a DGVM coupled to the SECHIBA land-surface model (Ducoudré et al., 1993). ORCHIDEE has been previously evaluated against data from flux tower sites (Verbeeck et al. 2011) and forest plot data (Delbart et al., 2010).

Photosynthesis in ORCHIDEE is simulated following the formulations of Farquhar et al. (1980) and Collatz et al. (1992), while stomatal conductance is computed via the technique of Ball et al. (1987). Maintenance respiration of plant pools in ORCHIDEE is calculated using PFT-specific functions of (a) temperature and biomass and (b) nitrogen/carbon ratios (see Ruimy et al., 1996). Soil layering characteristics are site dependent, with rooting distributions determined by availability of water, light and nitrogen. By definition, vegetation phenology is prognostic and is based on PFT-specific temperature and moisture constraints (Krinner et al. 2005). With respect to biomass pools, the model consists of four separate carbon pools, plus total soil carbon (Verbeeck et al. 2011). Representation of vegetation dynamics and disturbance follows the approach described in the LPJ model (Sitch et al., 2003). For the simulations in this study an 11 layer soil hydrology scheme was used (Guimberteau et al., 2012).

LPJmL

In LPJmL (Bondeau et al., 2007;Gerten et al., 2004;Sitch et al., 2003), most physiological and hydrological processes are simulated at daily time steps, whereas vegetation dynamics and PFT composition are updated annually. Natural vegetation is represented by nine plant functional types (PFTs) which describe the main characteristics of plants within the different biomes across the globe. Over Amazonia, the dominant PFTs are tropical evergreen trees and tropical raingreen trees. Photosynthesis is based on the Farquhar model approach (Farquhar and von Caemmerer, 1982;Farquhar et al., 1980) with air temperature and radiation controlling photosynthetic activity at the leaf level. Transpiration and photosynthesis are coupled through stomatal conductance of the leaves, where increasing transpirational losses or carbon starvation due to closed stomata can reduce NPP under drought conditions or high temperatures. With continued drought depleting soil water storage, tropical raingreen trees shed their leaves during the dry season to avoid carbon loss and mortality. Tropical evergreen broadleaf trees keep their leaves and are thus usually outcompeted in a seasonal dry tropical climate.

Carbon gained is allocated annually to the living carbon pools where basic allometric relations between crown area, tree height and stem diameter are met (Sitch et al. 2003). The pipe model ensures that each unit of leaf area is supported by a corresponding area of transport tissue, i.e. the sapwood cross-sectional area. Canopy closure is assumed but no crown overlap is permitted. Furthermore, plants can invest more carbon to fine roots under water-limited conditions to reduce drought risks This term is parameterized for each PFT.

Tree mortality results from heat stress, fire and light competition. The latter can occur due to low growth efficiency or thinning effects. Mortality from heat stress occurs when a PFT-specific temperature is crossed (Sitch et al. 2003), and individuals lost through fire are quantified by a PFT-specific parameter describing fire intensity and severity (Thonicke et al., 2001).
JULES

The Joint UK Land Environment Simulator (JULES) is the UK community land surface model (Best et al., 2011; Clark et al., 2011) and the land surface scheme for the Hadley Centre climate model. It is closely based on the MOSES-TRIFFID land surface scheme (Cox, 2001), which was used in some of the first studies that predicted ‘die-back’ of the Amazon region. This study utilized version 2.1 of JULES. JULES simulates five PFTs: broadleaf trees, needleleaf trees, shrubs, C3 grasses and C4 grasses, which compete with each other following Lotka-Volterra dynamics (Cox 2001). Over Amazonia, broadleaf trees are the dominant plant functional type. In our simulations, a four-layer soil model is simulated with a total depth of 10 m, although individual plant functional types differ in their rooting depth. Net leaf photosynthesis is calculated based on Collatz et al. (Collatz et al., 1991; Collatz et al., 1992). Leaf photosynthesis is coupled to stomatal conductance through the leaf internal CO$_2$ concentration, calculated using the approach of Jacobs (1994). Leaf photosynthesis is scaled to canopy level using a multi-layer approach which adopts the 2-stream approximation of radiation interception from Sellers (1985). JULES simulates 3 vegetation pools (foliage, roots and wood), with maintenance respiration for each pool calculated dependent on tissue temperature and nitrogen content. Carbon fluxes from JULES are accumulated and passed to the TRIFFID vegetation dynamics model every 10 days. NPP is partitioned into a fraction used for growth of existing vegetation and a fraction for ‘spreading’ (Clark et al. 2011), based on the leaf area index. Tree mortality is not explicitly considered in the model. Biomass losses occur via turnover of carbon pools, each with specific turnover times, and prescribed large-scale disturbance rates.

INLAND

The Integrated Model of Land Surface Processes (INLAND) is the land surface module currently under development for the Brazilian Earth System Model, within the Brazilian scientific community (Costa et al. in prep.). It is originally based on IBIS model (Foley et al., 1996; Kucharik et al., 2000), and further adapted with special focus on the representation of tropical ecosystems of South America. INLAND simulates 12 different PFTs competing for available resources within the grid cell and the relative success of each PFT determines its fractional coverage. The model allows trees and herbaceous plants or grasses to experience different light and water availability: while trees in the upper canopy have priority to capture available light (thus shading the shrubs and grasses in the lower part of the canopy), the herbaceous plants are able to capture soil water first when it infiltrates the ground (Foley et al. 1996). INLAND uses the mechanistic treatment of canopy photosynthesis proposed by Farquhar et al. (1980) and the semi-mechanistic Ball-Berry approach to estimate stomatal conductance (Ball et al., 1987; Collatz et al., 1991; Collatz et al., 1992), computing gross photosynthesis, maintenance respiration and growth respiration to yield the annual carbon balance for each PFT. The vegetation dynamics module simulates biomass changes for each PFT on a yearly time step. Net primary productivity (NPP) is allocated to individual biomass pools (leaves, roots, wood) according to fixed allocation coefficients. Mortality is not explicitly modelled. Instead, biomass losses occur via turnover of the existing carbon pool, according to fixed turnover rates as well as via large-scale disturbance caused by fire or land use change.

4) Modelling protocol for the DGVM simulation runs

Model runs were performed based on the standardized Moore Foundation Andes-Amazon Initiative (AAI) modelling protocol (Zhang et al. 2015). To derive pre-industrial equilibrium of carbon pools and vegetation
distribution, the models were run through a 500 year spin-up period by repeating 39 years of bias-corrected NCEP climatology (Sheffield, Goteti & Wood 2006) with pre-industrial atmospheric CO$_2$ concentration of 278 ppm and for the transient runs with increasing CO$_2$ concentrations from Zhang et al. (2015).

5) Further information on inventory data

Mitchard et al. (2014) provide AGB values (in Mg ha$^{-1}$) calculated with six different biomass equations (see Table S1) provided by Chave et al. (2006;2014) taking into account two (diameter-at-breast-height (DBH) and tree height) or three (DBH, height, wood density) parameters by using continental or regional tree-height models and stem-level wood density (mean or derived from species data). From the efforts of the TEAM, RAINFOR and ATDM projects, measurements from 413 plots across the Amazon region are available (Lopez-Gonzalez et al., 2011;Lopez-Gonzalez et al., 2014). Plot measurements were taken from 1956 to 2013.

Additionally, we use data from Brienen et al. (2015) who provide values for WP and woody loss (WL) in Mg ha$^{-1}$ from 321 plots.

In order to compare measured and modelled biomass, both variables need to match in terms of units. We convert the measured dry biomass of AGB and WP (in Mg dry matter (DM) ha$^{-1}$ and Mg DM ha$^{-1}$ yr$^{-1}$, respectively) to carbon mass (in MgC ha$^{-1}$ and MgC ha$^{-1}$ yr$^{-1}$). For this conversion, Martin and Thomas (2011) suggest a value of 0.474 gC/gDM, which we apply within our study. Martin and Thomas (2011) give a standard error of 0.025 gC/gDM or about 5.2% for their estimate of carbon content in dry matter.

6) Description of allometric equations

Table S1: Description of the six allometric equations applied to calculate AGB of the inventory data used here, according to Mitchard et al. (2014) and Lopez-Gonzalez et al. (2014). To derive biomass from the measurements, the “Chave Moist equation” (Chave et al., 2006;Chave et al., 2014) for moist forests was used, which can have either three or two parameters, the latter one excluding tree height.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Height model</th>
<th>Wood density</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdhr</td>
<td>Regional</td>
<td>Species-level</td>
<td>Above-ground biomass calculated using the 3-parameter Chave Moist equation, using regional height models and stem-level wood density (derived from species data)</td>
</tr>
<tr>
<td>Kdh</td>
<td>Regional</td>
<td>Mean</td>
<td>Above-ground biomass calculated using the 3-parameter Chave Moist equation, using regional height models and the dataset mean wood density for every stem</td>
</tr>
<tr>
<td>Kdr</td>
<td>Continental</td>
<td>Species-level</td>
<td>Above-ground biomass calculated using the 3-parameter Chave Moist equation, using a continental height model and stem-level wood density (derived from species data)</td>
</tr>
<tr>
<td>Kd</td>
<td>Continental</td>
<td>Mean</td>
<td>Above-ground biomass calculated using the 3-parameter Chave Moist equation, using a continental height model and the dataset mean wood density for every stem</td>
</tr>
</tbody>
</table>
| Kd2p          | -            | Species-level| Above-ground biomass calculated using the 2-parameter Chave Moist equation (which excludes
| Kd2p  | -     | Mean | Above-ground biomass calculated using the 2-parameter Chave Moist equation (which excludes height) and the dataset mean wood density for every stem |

Results

1) AGB maps for the remaining five allometric models

Figure S1: Maps of mean aboveground biomass (AGB) from inventory data within the simulated pixels derived from six allometric equations. The map for the first allometric equation (Kdh) is shown in the main text.
2) Regional variability of within-pixel variability

Figure S2: Regional variability of within-pixel variability $\sigma_e$ according to the Feldpausch regions (Feldpausch et al., 2011) for the different allometric models (see Table S1). AMZ denotes the variability across the whole Amazon basin (and is used in our calculations). WA = Western Amazonia, BS = Brazilian Shield, GS = Guiana Shield, CEA = Central Eastern Amazon.
3) Mean aboveground biomass (AGB) for different allometric models

Table S2: Values for observed mean aboveground biomass (AGB) for six different allometric models.

<table>
<thead>
<tr>
<th>Observed AGB Allometric models</th>
<th>Mean $\bar{x}$ (MgC/ha)</th>
<th>Mean global variability $\sigma_x$ (MgC/ha)</th>
<th>Within-pixel variability $\sigma_x$ (MgC/ha)</th>
<th>Corrected global variability $\sigma_{x,corr}$ (MgC/ha)</th>
<th>Max. achievable correlation $r_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdhr</td>
<td>135.96</td>
<td>49.80</td>
<td>31.14</td>
<td>38.86</td>
<td>0.78</td>
</tr>
<tr>
<td>Kdh</td>
<td>134.76</td>
<td>41.04</td>
<td>28.20</td>
<td>29.82</td>
<td>0.72</td>
</tr>
<tr>
<td>Kdr</td>
<td>134.66</td>
<td>42.45</td>
<td>30.73</td>
<td>29.29</td>
<td>0.69</td>
</tr>
<tr>
<td>Kd</td>
<td>134.19</td>
<td>35.94</td>
<td>28.44</td>
<td>21.97</td>
<td>0.61</td>
</tr>
<tr>
<td>Kdr2p</td>
<td>153.14</td>
<td>49.25</td>
<td>35.84</td>
<td>33.78</td>
<td>0.68</td>
</tr>
<tr>
<td>Kd2p</td>
<td>152.62</td>
<td>41.95</td>
<td>33.25</td>
<td>25.58</td>
<td>0.61</td>
</tr>
</tbody>
</table>
Table S3: Comparison metrics for AGB.

<table>
<thead>
<tr>
<th>AGB</th>
<th>INLAND</th>
<th>JULES</th>
<th>ORCHIDEE</th>
<th>LPJmL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean $\bar{y}$  (MgC/ha)</td>
<td>114.36</td>
<td>151.33</td>
<td>217.60</td>
<td>169.92</td>
</tr>
<tr>
<td>Mean global variability $\sigma_y$ (MgC/ha)</td>
<td>32.78</td>
<td>13.39</td>
<td>61.96</td>
<td>54.00</td>
</tr>
<tr>
<td>Mean bias ($\bar{y}/\bar{x}$)</td>
<td>Kdhr: 0.84</td>
<td>1.12</td>
<td>1.60</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Kdh: 0.85</td>
<td>1.12</td>
<td>1.61</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Kdr: 0.85</td>
<td>1.12</td>
<td>1.62</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Kd: 0.85</td>
<td>1.12</td>
<td>1.62</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Kdr2p: 0.75</td>
<td>0.99</td>
<td>1.42</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>Kd2p: 0.75</td>
<td>0.99</td>
<td>1.42</td>
<td>1.11</td>
</tr>
<tr>
<td>Pattern amplitude ($\sigma_y/\sigma_{x_{corr}}$)</td>
<td>Kdhr: 0.86</td>
<td>0.35</td>
<td>1.62</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>Kdh: 1.12</td>
<td>0.45</td>
<td>2.11</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>Kdr: 1.14</td>
<td>0.46</td>
<td>2.15</td>
<td>1.88</td>
</tr>
<tr>
<td></td>
<td>Kd: 1.55</td>
<td>0.62</td>
<td>2.91</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td>Kdr2p: 0.99</td>
<td>0.40</td>
<td>1.87</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>Kd2p: 1.33</td>
<td>0.54</td>
<td>2.50</td>
<td>2.19</td>
</tr>
<tr>
<td>Similarity of pattern ($r_{corr}$)</td>
<td>Kdhr: 0.46</td>
<td>0.42</td>
<td>0.53</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Kdh: 0.55</td>
<td>0.49</td>
<td>0.54</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Kdr: 0.48</td>
<td>0.43</td>
<td>0.56</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Kd: 0.59</td>
<td>0.53</td>
<td>0.56</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Kdr2p: 0.47</td>
<td>0.43</td>
<td>0.56</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Kd2p: 0.58</td>
<td>0.52</td>
<td>0.55</td>
<td>0.34</td>
</tr>
</tbody>
</table>
Table S4: Mean simulated values for woody productivity (WP) and woody residence time (τ) calculated only for pixels that contain observational plots.

<table>
<thead>
<tr>
<th>WP</th>
<th>INLAND</th>
<th>JULES</th>
<th>ORCHIDEE</th>
<th>LPJmL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (\bar{y}) (MgC/ha/yr)</td>
<td>8.01</td>
<td>5.18</td>
<td>9.14</td>
<td>4.47</td>
</tr>
<tr>
<td>Mean variability (\sigma_\bar{y}) (MgC/ha/yr)</td>
<td>1.09</td>
<td>0.72</td>
<td>1.60</td>
<td>0.75</td>
</tr>
<tr>
<td>τ (years)</td>
<td>INLAND</td>
<td>JULES</td>
<td>ORCHIDEE</td>
<td>LPJmL</td>
</tr>
<tr>
<td>Mean (\bar{y})</td>
<td>14.73</td>
<td>31.03</td>
<td>25.79</td>
<td>34.68</td>
</tr>
<tr>
<td>Mean variability (\sigma_\bar{y}) (years)</td>
<td>3.15</td>
<td>5.66</td>
<td>5.79</td>
<td>8.45</td>
</tr>
</tbody>
</table>

Supplementary Literature


Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., Henry, M., Martinez-Yzirzar, A., Mugasha, W. A.,


