**Supporting Information**

**Accounting for trait variability in vegetation models enhances future projections of Amazon forest plant functional diversity and carbon storage**

**Authors:** Bianca Fazio Rius, João Paulo Darela Filho, Katrin Fleischer, Florian Hofhansl, Anja Rammig, Tomas Domingues, David Montenegro Lapola

Supporting information SI.1 - **Description of CAETÊ (Carbon and Ecosystem functional Trait Evaluation Model)**

Here we describe some additional information of the model CAETÊ, including all the ecophysiological equations. The symbols used for each variable, its description, unit and associated equations can be found in Table SI.5. The constant parameters with its respective values are shown in Table SI.6. Hereafter, the symbols , and correspond, respectively, to a PFT/PLS, a grid cell and a plant compartment [leaves, aboveground woody tissues (hereafter ABGW) and fine roots].

**SI.1.1. Model structure**

1. Creating and sampling plant life strategies (PLSs)

In order to create the hypervolume that contains all the possible combinations of functional traits and, hence, the PLSs, each trait owns a range of values that determines the minimum and the maximum value that will be sampled (Table SI.3). From this initial range , trait values are randomly and uniformly sampled. All the sampled traits are combined to create the potential functional space that contains all the possible combinations of values. From this potential space some combinations are excluded following a viability standard: first, the combination of values for allocation functional traits, necessarily, must sum 100%, otherwise this combination is excluded from the functional space. After that, all the combinations are submitted to a spinup simulation in order to verify if it presents carbon content on fine roots and on leaves that are higher than 0.001. Then, from the combinations that respect this premise, 3000 PLSs are sampled and seeded in each grid cell. The number of sampled PLSs was defined through a sensitivity test (see Supporting Information SI.2.)

1. Grid cell occupation and survivorship

In the model CAETÊ the abundance of a strategy (a PLS or a PFT) is the fraction of the grid cell that it occupies based on the carbon stock contribution of this strategy () in relation to the total carbon stock in a grid cell (). Therefore, the relative abundance () of a PFT/PLS in a grid cell is given by:

|  |  | (1) |
| --- | --- | --- |
|  |  | (2) |

where is the number of PFT/PLS in that grid cell.

This procedure has support on the biomass-ratio hypothesis (Grime, 1998) that predicts that the immediate effects of a species is proportional to its relative contribution to the total biomass of the community. Then, from the relative abundances it is possible to aggregate the biogeochemical variables from a PFT/PLS scale to the grid cell scale. For example, the net primary production () of a grid cell is equal to the sum of the of all PFTs/PLSs that survived () in that grid cell, weighted by their respective relative abundance ():

|  |  | (3) |
| --- | --- | --- |

This approach is also necessary because the competition scheme in CAETÊ model is still not mechanistic. However, the differential capture of carbon, water and light through the differential combination of functional traits is capable of leading to difference of abundances (Grime, 1998) between the PLSs (for the trait-based modeling approach) and between the PFTs (for the PFT modeling approach) in a grid cell. In that sense, the competition is indirectly considered by deriving the relative abundances (Grime, 1998). A PFT/PLS is considered as a survivor if it presents a positive carbon balance and also if it presents, concomitantly, at least 100 g of carbon in fine roots and in leaves; otherwise, the carbon in all the compartments is directed to the litter and to the decomposition process (see Lapola, Oyama & Nobre , 2009).

1. Spinup

Before the model initialization a spinup simulation was performed to determine the initial carbon content on plant compartments. This spinup was run until the stability attainment of the total carbon stock (i.e. the sum of carbon in all plant compartments) in all the grid cells, with a sensibility of 10% of difference between and . Thus, the initial amount of carbon () in each plant compartment in an specific PFT or PLS in a given time :

|  |  | (4) |
| --- | --- | --- |

where (Table SI.2 and SI.3) is the fraction of allocated to a plant compartment in each PFT/PLS and (Table SI.2 and SI.3) represents the carbon residence time in a compartment. Here, is considered as potential (), calculated from the model CPTEC Potential Vegetation Model 2 (CPTEC-PVM2; Lapola, Oyama, & Nobre, 2009)⁠, a precursor model of CAETÊ, using the same climatology applied to CAETÊ running.  
 **SI.1.2. Input Data**

The following climatic data were used as inputs for the model: air surface temperature, precipitation, shortwave radiation, relative humidity and atmospheric CO2 concentration. The first four variables were obtained from mean monthly data for the period between 1980 to 2010. which are available in the Inter-Sectoral Impact Model Intercomparison Project 2 (ISI-MIP2; Warszawski et al., 2014)⁠. Atmospheric CO2 concentration was obtained from the CO2.Earth website (Mc. Gee, 2020), employing the mean value of [CO2] for the same period used for the other climatic data.

**SI.1.3. Photosynthesis**

The photosynthesis equation () and the ones associated with it are based on Farquhar, Caemmerer, & Berry (1980) formulation, which takes into account three limiting factors: rubisco carboxylation ( ), light () and electron transport ( ). We also included water stress limitation ().

|  |  | (5) |
| --- | --- | --- |

where is the leaf level gross photosynthesis and is a function for upscaling the leaf level photosynthesis to the canopy level.

In the following equations are constants summarized in Table SI.5. The is calculated as the smallest root between the three limiting rates: , and :

|  |  | (6) |
| --- | --- | --- |

where is the minimum between and :

|  |  | (7) |
| --- | --- | --- |

is the photosynthesis rate limited by the Rubisco carboxylation capacity:

|  |  | (8) |
| --- | --- | --- |

where is the rate of Rubisco carboxylation, is the partial CO2 pressure at leaf interior, is the photorespiration compensation point, is the Michaelis-Menten constant for CO2 and the Michaelis-Menten constant for O2.

|  | ] | (9) |
| --- | --- | --- |

|  | ] | (10) |
| --- | --- | --- |

|  |  | (11) |
| --- | --- | --- |

where is the leaf level moisture deficit and is the atmospheric CO2 concentration (input).

|  |  | (12) |
| --- | --- | --- |

is obtained by the actual mixing ratio on leaf level and the saturated mixing ratio ():

|  |  | (13) |
| --- | --- | --- |

is a function of partial pressure of water vapor () and the surface pressure ():

|  |  | (14) |
| --- | --- | --- |

|  |  | (15) |
| --- | --- | --- |

is the photosynthetic rate limited by light as a function of the incident photosynthetically active radiation [; here we considered as 50% of the shortwave radiation (input)]:

|  |  | (16) |
| --- | --- | --- |

The photosynthetic rate limited by the electron transport is given by:

|  |  | (17) |
| --- | --- | --- |
|  |  | (18) |

where (Table SI.6) is the maximum rate of Rubisco carboxylation and (input) is temperature.

The function is used in the canopy scaling of photosynthesis. Since we assume a canopy division in sun and shade parts this function is subdivided in and . The sun part aims to represent the canopy portion in which solar radiation reaches it directly in a 90º angle, while the shade part receives diffuse radiation in a 20º angle.

|  |  | (19) |
| --- | --- | --- |
|  |  | (20) |

where is the leaf area index also splitted in and in order to follow the canopy subdivision. Then, following Beer-Lambert’s Law (see De Pury & Farquhar, 1997):

|  |  | (21) |
| --- | --- | --- |

The calculation of (specific leaf area) is based on the approach used by Pavlick et al. (2013):

|  |  | (22) |
| --- | --- | --- |

where (Table SI.1 and SI.2) is the carbon residence time on leaves for a PFT/PLS. This equation uses an empirical relationship derived from Reich, Walters, & Ellsworth (1997) and aims to compute the so-called leaf economic spectrum (Wright et al., 2004).

|  |  | (23) |
| --- | --- | --- |
|  |  | (24) |

## **SI.1.4. Respiration**

The autotrophic respiration () is splitted in growth respiration () and maintenance respiration () following Ryan (1991a; 1991b).

|  |  | (25) |
| --- | --- | --- |

## *SI.1.4.1. Growth Respiration*

The formulation is based on Ryan et al. (1991a; 1991b), which postulates that a reasonable growth respiration estimate can be done assuming that its metabolics costs in each tissue consumes an amount of carbon equal to a quarter of the carbon incorporated in the new tissue (i.e. the difference between the carbon content in a compartment in a time () and the carbon content in the same compartment in a time (). Total growth respiration is the sum of growth respiration in all the three plant compartments considered.

|  |  | (26) |
| --- | --- | --- |

*SI.1.4.2. Maintenance Respiration*

The maintenance respiration is calculated according to the nitrogen and carbon content in each compartment. The nitrogen content on plant tissues is intimately connected to the maintenance respiration since about 60% of it is used to protein repair and substitution (Ryan, 1991a, 1991b). Yet, the nitrogen supriment is not here considered as a limiting factor.

|  |  | (27) |
| --- | --- | --- |

where represents the N:C ratio for each plant compartment and (ºC) is the mean annual temperature. A considerable amount of the woody tissues do not respire (the so called heartwood; (Ryan, 1991b), and then we consider only 5% of the aboveground woody tissues carbon content on the maintenance respiration for this compartment. This 5% aims to represent the sapwood (Pavlick et al., 2013). The N:C ratio is equal to 0.034, 0.003 and 0.034 for leaves, sapwood and fine roots, respectively (Levis et al., 2004; Sitch et al., 2003). The parameter 0.07 is a sensibility factor of the maintenance respiration to temperature. Because in tropical forests the soil temperature is lower than the air temperature, then for calculating the fine roots maintenance respiration we used the soil temperature (; see Oyama & Nobre, 2004).

## **SI.1.5. Net primary productivity**

The net primary productivity () is the carbon available from photosynthesis () discounting the respiratory costs ().

|  |  | (28) |
| --- | --- | --- |

**SI.1.6. Carbon allocation and stock**

From NPP, carbon is allocated to the three different carbon compartiments (leaves, aboveground woody tissues and fine root) through allocation traits (Table SI.2 and SI.3) that determines the percentage of distributed to each compartment . Residence time traits (Table SI.2 and SI.3) represents how long the carbon stays in each compartment. Together, these functional traits ultimately define the stock of carbon () in each compartment of a PLS or PFT in a certain time step , and then, the total plant carbon of that PLS or PFT ().

|  |  | (29) |
| --- | --- | --- |
|  |  | (30) |

**SI.1.7. Stomatal conductance and canopy resistance**

Stomatal conductance () and canopy resistance () link the carbon cycle (through ) with the water balance sub-model (see SI.1.8.):

|  |  | (31) |
| --- | --- | --- |
|  |  | (32) |

is the stomatal conductance with a fix value of 0.001, is the conductance sensibility to the carbon assimilation (Medlyn et al., 2011) and has a constant value of 3.77. Lastly, is the deficit of vapor pressure on the leaf surface:

|  |  | (33) |
| --- | --- | --- |

where (input) is the relative humidity and is the evapotranspiration (see SI.1.8.).

**SI.1.8. Light capture for distinct PFTs/PLSs**

The carbon stored on aboveground wood tissues are highly correlated to light competition (Reich, 2014). However, the model CAETÊ is still not able to represent the competition ecological process mechanistically, since there is no representation of shading of PFTs/PLSs that resemble understory plants by the ones that resemble trees with higher height and higher canopy area, for example. The majority of models represents the plants competitive ability for light associated with woody tissues through variables that describe height, wood density and canopy ratio [e.g. aDGVM (Scheiter, Langan, & Higgins, 2013) and LPJ (Sitch et al., 2003). These variables are estimated via allometric relationships that were not used by CAETÊ in this study. So in order to implement a trade-off for the traits (allocation and residence time) that determines the carbon content on ABGW (), we used a simplified approach for the sake of representing the differential light capture between the different PFTs/PLSs in a grid cell. We considered that 5% of the PFTs/PLSs that present the higher relative carbon content on ABGW in a grid cell can capture 100% of the incident while the other ones can capture 80% of the IPAR .

**SI.1.9. Water balance sub-model**

The CAETÊ uses a sub-model to calculate the water balance in a grid cell scale. This sub-model is based on Oyama & Nobre (2004) and is evaluated over an homogeneous soil layer with a unique water maximum holding capacity () equal to 500 mm. In the present study different types of soil are not considered. Developments on CAETÊ have being done in order to increase the amount of soil layers and to implement different soil types with different water holding capacity and different nutrients availability.

*SI.1.9.1. Soil water content and saturation*

The model here employed calculates soil water and snow budgets separately, however given the climatic conditions of the study area we do not present here the soil snow budget formulation [see Oyama & Nobre (2004) for accessing it]. The actual soil water content () in a grid cell in a certain time step depends on the precipitation (), the evapotranspiration (; see Oyama & Nobre, 2004) and the runoff ( ; see Oyama & Nobre, 2004) on this grid cell:

|  |  | (34) |
| --- | --- | --- |

Then, the degree of water soil saturation in the grid cell (; Eq. 35) can be calculated:

|  |  | (35) |
| --- | --- | --- |

*SI.1.9.2. Water stress*

In order to include the limitation of photosynthesis by water availability and to comprise a trade-off for the fine roots traits investment, we added to the photosynthesis () equation a water stress factor (). This factor, is calculated in function of the ratio between the potential water supply for transpiration () and the atmospheric demand for transpiration (). This approach is based on Pavlick et al. (2013).

|  |  | (36) |
| --- | --- | --- |

is a function of the carbon stock on fine roots compartment () multiplied by a constant water uptake capacity () with a value of 0.0005 mmH2OkgC-1day-1, and by the degree of water soil saturation in the grid cell (). It is important to highlight that the value for is correspondent to that of the previous day.

|  |  | (37) |
| --- | --- | --- |

Following Gerten et al. (2004), represents the condition of a “unstressed transpiration” which occurs when stomatal opening is not limited by reduced water potential in the plant:

|  |  | (38) |
| --- | --- | --- |

where is the potential evapotranspiration (see Oyama & Nobre, 2004); is a coefficient with fixed value equal to 1.391, is the canopy scaling stomatal conductance and presents a value equal to 3.26 mms-1.is the canopy potential conductance when there is not water limitation and is calculated through minimum stomatal resistance () with a value equal to 100:

|  |  | (39) |
| --- | --- | --- |

**SI.1.10. Functional traits and trade-offs**

Each functional trait used in this study is related to at least one trade-off (the trade-offs associated to each functional trait are described in Table SI.3), which is a cost and benefit relation that impacts the ecophysiological process of a plant life strategy (i.e., a specific combination of functional traits), and, therefore, after all, determines its performance (its relative abundance in a grid cell; see section SI.1.1.b), and whether this strategy will be able to deal with the environmental conditions (Pavlick et al., 2013; Reu et al., 2011).

Intrinsic trade-off emerges between the allocation traits: since there is a limited amount of carbon to be directed to each plant compartment, the allocation in one of them limits the carbon availability for the other. Together with the allocation traits, the carbon residence time in the plant compartments determines the amount of carbon in each one of them, in other words, determines investment. Although higher values of residence time provide, in total, more carbon accumulation in a plant tissue, it also implies costs, since all the tissues are involved in respiratory costs, both for their maintenance and growth (Eq. 25-27). Then, for example, investing carbon in leaves can increase photosynthesis rate by increasing solar radiation absorption (Eq. 21 and 22), however, such investment is at the expense of investing in fine roots, for example, which is responsible for water uptake, also a limiting resource for photosynthesis (Eq. 36). Therefore, in a given environmental condition, some traits combinations will be favored from a balance between the costs and benefits of investing carbon in each compartment, maximizing survivorship and performance.

# Supporting information SI.2 - **Sensitivity test for the number of sampled PLSs**

# After creating the potential functional space with more than 105 a problem emerges: how many PLSs are enough to be sampled in order to make meaningful representation of all? To answer this question a set of 6 ensembles with different numbers of sampled PLSs (number of PLSs = 50, 100, 200, 500, 1000, 3000) was run to test the model sensitivity. Each ensemble contains 10 runs, this is required because of the random nature of the initial sampling process. The only difference between the ensembles is the number of initialized PLS.

To explore the sensitivity of the model for the number of sampled PLSs we assessed the mean values and the variance in each of the 10 runs for the 6 ensembles for total plant carbon storage and for all the six functional traits (Fig. SI.6). With this strategy we seeked to understand how the mean and the variances change between the runs for each ensemble of PLS numbers. Our idea is that since the sampling of functional trait values is random, we expect that the higher the number of PLSs the lower the difference in mean and variance between the runs. It would ensure that the results that emerge from the model, both for biogeochemical and functional diversity variables, are not a product of a “forged” diversity derived from the simple difference between the runs.

As expected, it is notable that the number of PLSs sampled from all the possible combinations presents a meaningful influence on estimating the mean value and variance between the runs, both for the total plant carbon stock and the functional traits. We observed that as more PLSs are sampled, the greater the convergence in the mean values (orange ticks on Fig. SI.6) and in the variance between the 10 runs. These results emerge from the fact that the increase in the number of sampled PLSs improves the ability in representing the total traits combinations in the potential functional space.

Supporting information SI.3 - **Building hypervolumes**

The method to analyze functional diversity from a multi-trait perspective is the hypervolume metric from Blonder et al. (2014). To use this method, the components entering the analysis need to be in comparable units (e.g. centred and scaled), uncorrelated and the number of variables should not exceed 5-8 (Barros et al., 2016; Blonder et al., 2014). In this study we used six variable functional traits that present distinct units and are correlated (mainly the allocation ones). In that sense, in order to overcome these limitations we performed, as recommended by Barros et al. (2016), a scaled principal component analysis (PCA; Fig. SI.7) and selected to perform the hypervolumes the factor scores of the first three orthogonal principal components PCs which retained a cumulative variance explanation > 95% (Table SI.7; Barros et al., 2016). Since our interest was to assess differences in the hypervolumes with the applied low precipitation scenario, the PCA was calculated using regular climate and low precipitation datasets together (Barros et al., 2016). After that, separate hypervolumes were calculated from the factor scores corresponding to each dataset (Fig. 5 and SI.5).

**References**

Barros, C., Thuiller, W., Georges, D., Boulangeat, I., & Münkemüller, T. (2016). N-dimensional hypervolumes to study stability of complex ecosystems. *Ecology Letters*, *19*(7), 729–742. https://doi.org/10.1111/ele.12617

Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, *23*(5), 595–609. https://doi.org/10.1111/geb.12146

De Pury, D. G. G., & Farquhar, G. D. (1997). Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. *Plant, Cell and Environment*, *20*(5), 537–557. https://doi.org/10.1111/j.1365-3040.1997.00094.x

Farquhar, G. D., Caemmerer, S. Von, & Berry, J. a. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta*, *90*, 78–90. https://doi.org/10.1007/BF00386231

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., & Sitch, S. (2004). Terrestrial vegetation and water balance - Hydrological evaluation of a dynamic global vegetation model. *Journal of Hydrology*, *286*(1–4), 249–270. https://doi.org/10.1016/j.jhydrol.2003.09.029

Grime, J. P. (1998). Benefits of plant diversity to ecosystems: Immediate, filter and founder effects. *Journal of Ecology*, *86*(6), 902–910. https://doi.org/10.1046/j.1365-2745.1998.00306.x

Lapola, D. M., Oyama, M. D., & Nobre, C. A. (2009). Exploring the range of climate biome projections for tropical South America: The role of CO2 fertilization and seasonality. *Global Biogeochemical Cycles*, *23*(3), 1–22. https://doi.org/10.1029/2008GB003357

Levis, S., Bonan, G. B., Vertenstein, M., & Oleson, K. W. (2004). The Community Land Model’s Dynamic Global Vegetation Model (CLM-DGVM). *Ncar/Tn-459+Ia*, 50. <https://doi.org/10.5065/D6P26W36>

McGee, M. (2020, March 2). *Monthly CO2.* CO2.Earth. https://www.co2.earth/monthly-co2

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., … Wingate, L. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134–2144. https://doi.org/10.1111/j.1365-2486.2010.02375.x

Oyama, M. D., & Nobre, C. a. (2004). A simple potential vegetation model for coupling with the Simple Biosphere Model (SiB). *Rev. Bras. Meteorol.*, *19*, 203–216.

Pavlick, R., Drewry, D. T., Bohn, K., Reu, B., & Kleidon, A. (2013). The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences*, *10*(6), 4137–4177. https://doi.org/10.5194/bg-10-4137-2013

Reich, P B, Walters, M. B., & Ellsworth, D. S. (1997). From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(25), 13730–13734. https://doi.org/10.1073/pnas.94.25.13730

Reich, Peter B. (2014). The world-wide “fast-slow” plant economics spectrum: A traits manifesto. *Journal of Ecology*, *102*(2), 275–301. https://doi.org/10.1111/1365-2745.12211

Reu, B., Proulx, R., Bohn, K., Dyke, J. G., Kleidon, A., Pavlick, R., & Schmidtlein, S. (2011). The role of climate and plant functional trade-offs in shaping global biome and biodiversity patterns. *Global Ecology and Biogeography*, *20*(4), 570–581. https://doi.org/10.1111/j.1466-8238.2010.00621.x

Ryan, M. G. (1991a). A simple method for estimating gross carbon budgets for vegetation in forest ecosystems. *Tree Physiology*, *9*(1\_2), 255–266. https://doi.org/10.1093/treephys/9.1-2.255

Ryan, M. G. (1991b). Effects of Climate Change on Plant Respiration, *1*(2), 157–167.

Scheiter, S., Langan, L., & Higgins, S. I. (2013). Next-generation dynamic global vegetation models: learning from community ecology. *The New Phytologist*. https://doi.org/10.1111/nph.12210

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., … Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2014). The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3228–3232. https://doi.org/10.1073/pnas.1312330110

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., … Gulias, J. (2004). The worldwide leaf economics spectrum. *Nature*, *428*, 821–827.