Analysing Amazonian Forest Productivity Using a new Individual & Trait-Based Model (TFS v.1).


1. Ecology and Global Change, School of Geography, University of Leeds, UK
2. School of Geography, University of Exeter, UK
3. Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil
4. School of GeoSciences, University of Edinburgh, Scotland, UK
5. Instituto de Investigaciones para el Desarrollo, Forestal Facultad de Ciencias Forestales y Ambientales, Universidad de Los Andes, Venezuela
6. Department of Wildlife Conservation and Management, Universidad Estatal Amazónica, Puyo, Pastaza, Ecuador
7. Universidade Federal do Acre, Rio Branco, Brazil
8. Museu Paraense Emílio Goeldi, Belém, Brazil
9. Environmental Change Institute, School of Geography and the Environment, University of Oxford, UK
10. Centre for Tropical Environmental and Sustainability Science (TESS) and School of Marine and Tropical Biology, James Cook University, Australia
11. Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, UK.

* Currently at: Department of Ecology & Systematics, Faculty of Biology, University of Athens, Greece.

Abstract

Repeated long-term censuses have revealed large-scale spatial patterns in Amazon Basin forest structure and dynamism, with some forests in the west of the Basin having up to a twice as high rate of aboveground biomass production and tree recruitment as forests in the east. Possible causes for this variation could be the climatic and edaphic gradients across the Basin and/or the spatial distribution of tree species composition. To help understand causes of this variation a new individual-based model of tropical forest growth, designed to take full advantage of the forest census data available from the Amazonian Forest Inventory Network (RAINFOR), has been developed. The model allows for within-stand variations in tree size distribution and key functional traits and between-stand differences in climate and soil physical and chemical properties. It runs at the stand level with four functional traits, leaf dry mass per area (Mₐ), leaf nitrogen (Nₗ) and phosphorus (Pₗ) content and wood density (Dₚ) varying from tree to tree in a way that replicates the observed continua found within each stand. We first applied the model to validate canopy-level water fluxes at three eddy covariance flux measurement sites. For all three sites the canopy-level water fluxes were adequately simulated. We then applied the model at seven plots, where intensive measurements of carbon allocation are available. Tree-by-tree multi-annual growth rates generally agreed well with observations for small trees, but with deviations identified for larger trees. At the stand-level, simulations at 40 plots were used to explore the influence of climate and soil nutrient availability on the gross (Πₐ) and net (Πₙ) primary production rates as well as the carbon use efficiency (Cₑ). Simulated Πₐ, Πₙ and Cₑ were not associated with temperature. On the other hand, all three measures of stand level productivity were positively related to both mean annual precipitation and soil nutrient status. Sensitivity studies showed a clear importance of an
accurate parameterisation of within- and between-stand trait variability on the fidelity of model predictions. For example, when functional tree diversity was not included in the model (i.e., with just a single plant functional type with mean Basin-wide trait values) the predictive ability of the model was reduced. This was also the case when Basin-wide (as opposed to site-specific) trait distributions were applied within each stand. We conclude that models of tropical forest carbon, energy and water cycling should strive to accurately represent observed variations in functionally important traits across the range of relevant scales.

Keywords: Amazon Basin, tropical forest, individual-based model, functional traits, stomatal conductance, soil nutrient availability, gross primary productivity, net primary productivity, carbon use efficiency.

Author for correspondence: Nikolaos M. Fyllas, nfyllas@gmail.com
1. Introduction

The Amazon Basin, encompassing one of the planet’s largest forest areas and hosting one quarter of the Earth’s biodiversity, constitutes a large reservoir of living biomass (Malhi and Phillips, 2005). Amazon forests also have a substantial influence on regional and global climates (Shukla et al. 1990; Spracklen et al., 2012). These forests are, however, under strong human pressure through logging, forest to pasture conversion, and face at present a warming and more variable climate and changing atmospheric composition (Lewis et al., 2004; Gloor et al., 2013). Due to the enormous area of forest within the Amazon Basin, these factors have the potential to modify global atmospheric greenhouse concentrations, regional and global climate, and the overall biodiversity of the planet (Cramer et al., 2004).

Traditionally, two approaches have been followed to understand current and future state of the Amazon forests. First, Dynamic Global Vegetation Models (DGVM) have been used to simulate vegetation patterns and carbon fluxes across Amazonia (Moorcroft et al., 2001; Galbraith et al., 2010) with some predicting substantial carbon losses under scenarios of global change (White et al., 1999; Cox et al., 2004) but with others less so (Cramer et al., 2004), or even gains (Huntingford et al., 2013). A second approach to understand Amazonian forests dynamics is through the analysis of long-term field observations of patterns of tree growth and mortality as they relate to climatic and edaphic variations across the Basin (e.g. Phillips et al., 2004; Quesada et al., 2012).

Analyses of Amazon forest inventory data, and particularly those of the Amazon Forest Inventory Network (RAINFOR) (Malhi et al., 2002), have revealed large-scale temporal trends in biomass and species composition as well as intriguing spatial patterns in many stand properties (Phillips et al., 1998; Baker et al., 2004, Phillips et al., 2009). Specifically, there is systematic spatial variation in species composition, biomass, growth and turnover rates, with western forests exhibiting higher wood productivity, faster turnover time and lower stand wood density compared to eastern forests (Baker et al., 2004; Malhi et al., 2006). This macroecological variation may possibly be explained by the Basin-wide observed climate and soil nutrient availability gradients (ter Steege et al., 2006; Quesada et al., 2012). The climatic gradient comprises a southeast to northwest increase in annual precipitation and decrease in dry season length (Sombroek, 2001), with aboveground wood productivity positively related to precipitation (Malhi et al., 2004). On the other hand, a soil age/nutritional axis spans from the northeastern part of the basin to southwestern Amazonia, with generally younger and richer soils in the west and highly weathered nutrient poor soils in the east (Sombroek, 2000; Quesada et al., 2011), although at regional and local scales the patterns are often more complicated than this macro-gradient might imply (Higgins et al., 2011). Soil physical properties (such as rooting depth, drainage and water holding capacity and soil structure) are similarly related to soil age and parental material (Quesada et al., 2010). Poor physical (for example soil depth) conditions (less weathered soils) are often associated with higher soil nutrient availability (Walker and Syers, 1976; Vitousek and Farrington, 1997) leading to increased nutrient concentrations at the leaf level (Fyllas et al., 2009) and thus a potential for higher photosynthetic rates (Reich et al., 1994; Raaimakers et al., 1995). In addition, increased disturbance-associated mortality rates in soils of poor physical properties lend towards more dynamic stands where faster growing species dominate (Chao et al., 2009, Quesada et al., 2012). This positive feedback
mechanism could explain the higher aboveground productivity and turnover rates observed for western forests (Quesada et al., 2012).

The simplistic ways by which plant functional diversity is currently reflected in DGVMs is an important shortcoming in predicting ecosystem response to environmental gradients and their vulnerability to global change (Lavorel et al., 2007). Some of the widely applied DGVMs represent Amazonian plant diversity with only few plant functional types (PFT), for example the LPJ model uses only two tropical-oriented PFTs (Sitch et al., 2003) and the JULES model only one (Clark et al., 2011). The mean values of key model parameters like photosynthetic capacity, wood density and leaf turnover times are selected to describe an a priori PFT definition (Fyllas et al., 2012). This means that many processes are controlled by a set of fixed parameters that describe viable plant strategies within very limited boundaries. Such PFT implementation has important drawbacks. It is usually based on the average value of a plant trait recorded from different field studies and different species. But recent studies have shown that key traits present a wide variation, dependent upon species identity and site growing conditions (Sultan, 2000; Fyllas et al., 2009; Baraloto et al., 2010a). Thus any given species has the potential to exhibit site-dependent shifts in its trait value; this being in addition to the inter-specific trait variability expected at any given site. Ignoring this plasticity could potentially bias modelling through an underestimation of the PFT's resilience by projecting dramatic but artificial switches in vegetation state caused by the limited and discrete (step-wise) nature of PFT descriptions.

Such unaccounted variability could be particularly important when modelling Amazonian forest dynamics, where environmental heterogeneity and plant functional diversity comprise key components of the ecosystem (Townsend et al., 2008). For example, the variation in leaf mass per area (ML) recorded within Amazon forests covers an approximately similar range to the one identified in global datasets, ranging from 30 to 300 gm⁻² (Fyllas et al., 2009). Similarly, there are large contrasts in soil physical and chemical conditions (Quesada et al., 2010). These important ecosystem flux drivers have now been better quantified with Amazon-wide climate (Malhi and Wright, 2004), soil (Quesada et al., 2011) and functional trait datasets also having been obtained (Baker et al., 2009; Fyllas et al., 2009; Patiño et al., 2009; Patiño et al., 2012). This is in addition to continually expanding long-term forest inventory data in which tree growth, mortality and species composition data are regularly being recorded (Keeling et al., 2008; Chao et al., 2009).

We here introduce a vegetation dynamics model developed as a tool to better analyse these observed Amazonian large-scale productivity patterns. This is achieved through specific incorporations of observed environmental and the biotic variations into the model formulation. Specifically we focus a) on the architectural variability, expressed through the size-class distribution of a stand, and b) on the functional variability, expressed through simulated distributions of four important functional traits which are allowed to vary from tree to tree within individual plots. Following a continuum approach, we replace the use of a discrete number of PFTs, with distributions of a functional traits "quartet" the within-stand distributions of which also vary from plot to plot in accordance with observation.

Two axes of functional variation/strategy are represented in the model: the leaf economic and the tree architecture spectra. The four functional traits include leaf mass per area (ML), leaf nitrogen and phosphorous dry mass concentration (NLm and PLm respectively) and wood density (DLw). The first three
traits express one component of the leaf economic spectrum (Reich et al., 1997; Wright et al., 2004), i.e. a global photosynthetic tissue trade-off between inexpensive, short-lived and fast payback leaves vs. costly, long-lived and slow payback leaves; although we emphasise that other factors such as leaf cation concentrations may be important in this respect (Fyllas et al., 2012; Patiño et al., 2012). Low $M_a$ and high nutrient content leaves are associated with comparably short longevity and usually have high (mass-based) gas exchange rates (Reich et al., 1994; Raaimakers et al., 1995). Lately the role of $P_{lm}$ has been highlighted as it expresses alternative limitations of the photosynthetic efficiency of tropical tree species (Domingues et al., 2010). The fourth trait, $D_W$, is used to represent a tree architectural axis with denser wood species supporting an overall higher aboveground biomass and thus having a higher maintenance respiration (Chave et al., 2005; Mori et al., 2010, although see Larjavaara and Muller-Landau, 2012). These two dimensions capture essentially a growth vs. survival trade-off. There is mixed evidence for a coordination between leaf and stem traits, i.e. a correlation between slow return related leaf traits and denser wood (Chave et al., 2009), with Baraloto et al. (2010b) suggesting that these two axes are independent, but with Patiño et al. (2012) showing some important correlations with foliar traits such as $P_{lm}$. For the purpose of this study we consider leaf and stem dimensions as independent axes of tree functional variation, with no predefined interrelationship between the representative traits. However, the observed among-stand variability of these four characters is used to express how growing conditions control plant processes while the within-stand trait variation represents a range of ecological strategies found under the same growing conditions.

The model is initialised with site-specific tree diameter and functional traits data, and forced with daily climate data. We first test the ability of the model to estimate stand-level water fluxes at three eddy-flux tower sites. For a subset of seven RAINFOR plots where site-specific carbon allocation coefficients are known, a tree-level test of stem growth rates is applied. We further validate the ability of the model to simulate the spatial patterns of aboveground biomass productivity at 40 RAINFOR plots, and subsequently explore the variation of Gross Primary Productivity ($\Pi_G$), Net Primary Productivity ($\Pi_N$) and Carbon Use Efficiency ($\CUE$) along established Amazonian climatic and edaphic gradients.

2. Materials and Methods

2.1 Model Description

“Traits-based Forest Simulator” (TFS) is an individual-based forest model, i.e. it simulates water and carbon fluxes for each tree in a stand. In the current version of the model, stand structure is prescribed in terms of the number of trees and their diameter at breast height ($d$). This is thus a “snapshot” version of the model, which does not take into account tree recruitment and mortality. In this version of TFS, each individual is fully described through $d$, with allometric equations used to estimate other attributes of interest like tree height ($H$), crown area ($C_A$), total leaf area ($L_A$) and tree-level leaf area index ($L$). Whole tree biomass is then partitioned to leaf ($B_L$), stem ($B_S$), coarse root ($B_{CR}$) and fine root ($B_{FR}$) biomass using established allometric equations. Allocation of assimilated carbon to different plant components is static, i.e. it does not change with size or resource availability, but rather implements field-derived allocation coefficients (Aragão et al., 2009). The general architecture of the model is presented in Fig. 1.
Tree functional diversity is expressed through four traits \( \{M_a, N_L, P_L, D_W\} \), which are randomly assigned to each tree: these pseudo-data being generated from local observations using a random vector generation algorithm. Leaf photosynthesis is calculated using a modified version of the Farquhar biochemical model (Farquhar et al., 1980), that incorporates leaf chemical and soil moisture effects. The maximum photosynthetic rate is regulated by \( N_L \) or \( P_L \) through the co-limitation model of Domingues et al. (2010). In contrast to most ecosystem fluxes models, where photosynthetic rates are directly regulated by water availability (Scheiter and Higgins, 2009; Clark et al. 2011), we couple water ‘stress’ to reduction of canopy conductance by estimating a daily fractional available soil water content for each tree in the stand. Carbon fluxes are simulated on an hourly and water fluxes on a daily time-step.

Light competition is based on the assumption of a perfect canopy tessellation. The flat-top version of the perfect plasticity model (Purves et al., 2007) has been used in the current version of TFS to characterise canopy and sub-canopy trees, by assuming that all of a tree’s foliage is found at the top of its stem (S1, Canopy Architecture and Radiation Environment). A canopy height \( Z^* \) is estimated for each forest stand, defining canopy and sub-canopy trees. By summing up the crown area \( C_A \) of all trees in the stand, \( Z^* \) is estimated as the height of the last tree that enters to the sum before the cumulative crown area is equal to the plot area. Canopy trees are absorbing a mean daily amount of shortwave solar radiation equal to the sum of mean beam, diffuse and scattered daily radiation in correspondence to the sun-shade model of de Pury and Farquhar (1997). The direct and diffuse fraction of solar radiation is estimated using the Spitters et al. (1986) approximation. The functional configuration of a tree (i.e. the values of the traits quartet) does not affect its light competitive status, as tree height and crown area are not directly associated to any of the four traits. Future versions of the model will incorporate such effects.

Soil water balance is approximated through a simple bucket model, with soil water content affecting leaf conductance and thus photosynthetic rates. Competition for soil water is approximated through a size hierarchy, i.e. bigger trees, with a more extensive root system are assumed to have access to deeper water (S1, Water Balance Algorithm). By assuming that a tree with a higher leaf biomass \( B_L \) requires a higher fine root biomass \( B_{FR} \), we indirectly implement a \( M_a \) effect on water competition (S1, Definition, Allometry and Stoichiometry of Individual Trees in TFS). In particular, between two trees of the same size, the higher \( M_a \) tree will be more competitive in terms of acquiring soil water.

TFS is coded in Java and it is fully described in S1. The main effects of including functional diversity are realised through trait-driven effects on photosynthesis and respiration (Reich et al., 2008; Reich et al., 2009). Model components that are linked with any of the four base traits are described in following paragraphs. All statistical analyses and graphs were made with R (R Development Core Team, 2013).

### 2.1.1 Within-stand Functional Diversity

As noted above, TFS employs neither species nor PFT descriptions, but rather a different discrete combination of each the four key functional traits \( M_a, N_L, P_L, \) and \( D_W \) is assigned to each individual tree along with a diameter-based allometry. To achieve this, the four functional characters assigned are generated using a procedure based on the actual values recorded within each plot. This is
achieved using a random vector generation algorithm (Taylor and Thompson, 1986). This algorithm, appropriate for generating non-repeated pseudo-observations from a relatively small sample of observations, was originally developed to provide for a realistic probabilistic representation of shrapnel projectile distributions in military battlefield simulations in the face of only a limited amount of available data (due to the cost and difficulty of undertaking the appropriate experiments). This “ballistic method” is notable in that it was specifically designed to short-circuit the usual step of multivariate density in the generation a pseudorandom population with approximately the same moments as the original sample. The ballistic method is readily programmable as follows (with the underlying rationale as discussed in Taylor & Thompson (1986) and Thompson (1989)) and with the following description based on Visual Numerics (2014):

First take a vector $X$ with $n$ multivariate observations $(x_1, \ldots, x_n)$. To generate a pseudodataset $\xi$ from $X$, one observation ($x_i$) is first chosen at random and its nearest $m$ neighbours, $x_{i1}, x_{i2}, \ldots, x_{im}$ are then determined and with the mean $\bar{x}_j$ of those nearest neighbours subsequently calculated. Next, a random sample $u_1, u_2, \ldots, u_m$ is generated from a uniform distribution with lower bound $\frac{1}{m} - \sqrt{\frac{3(m-1)}{m^2}}$, and upper bound $\frac{1}{m} + \sqrt{\frac{3(m-1)}{m^2}}$. The random variate $z_j$ is then estimated as $\sum_{i=1}^{m} u_i (x_{ij} - \bar{x}_j) + \bar{x}_j$, and the process then repeated as required. Somewhat subjective here is the selection of the appropriate value of the number of nearest neighbours ($m$) although the nature of the simulations is not strongly dependent upon that value (Taylor & Thompson, 1986). Thus, following their recommendation and as in the Visual Numerics (2014) default, we have taken here $m = 5$.

In our case, applying this procedure resulted in a coordinated trait quartet for each tree in a stand being generated on the basis on the smaller observational trait quartets sampled from trees in the same stand (Baker et al., 2009; Fyllas et al., 2009; Patiño et al., 2012) and without any assumptions having to be made about their underlying statistical distributions. Thus no single functional trait “average stand” value is used (or even required). Further, between-stand differences in the traits distributions and their covariances are also intrinsically taken into account. This is because each stand is characterised by its own multivariate trait sample and size distribution. More fertile plots have an overall lower $M_{\alpha}$ and higher $N_{\alpha m}$ and $P_{\alpha m}$ compared to infertile plots (Fyllas et al., 2009), with this being reflected in the photosynthetic capacity of individual trees, as described in the next paragraph.

### 2.1.2 Photosynthesis

A tree-level leaf area index ($L$), estimated as the ratio of $L_A$ to $C_A$, is used to compute the energy, carbon and water fluxes for each tree in a stand. The net photosynthetic rate ($\mu$mol m$^{-2}$ s$^{-1}$) is given from:

$$A_{\mu} = g_s (C - C^*) \ (1)$$
with $C_a$ the atmospheric CO₂ mixing ratio ($\mu$mol mol$^{-1}$), $C_c$ the CO₂ mixing ratio inside the chloroplast and $g$ the CO₂ stomatal conductance (mol m$^{-2}$ s$^{-1}$) calculated from Medlyn et al. (2011) and modulated by a soil moisture term. The leaf-level photosynthetic rate $A_n$ is scaled-up to the tree-level by multiplying with the $C_A$ of the tree.

The co-limitation equation suggested by Domingues et al. (2010), where the leaf level photosynthetic capacity (area basis) is potentially limited by either nitrogen or phosphorus is used TFS to estimate the leaf maximum carboxylation and electron transport rates:

$$V_{\text{max}} = M_a \left( \min \{a_{\text{NV}} + v_{\text{NV}} N_{L_m}, a_{\text{PV}} + v_{\text{PV}} P_{L_m} \} \right)$$ (2)

$$J_{\text{max}} = M_a \left( \min \{a_{\text{NJ}} + v_{\text{NJ}} N_{L_m}, a_{\text{PJ}} + v_{\text{PJ}} P_{L_m} \} \right)$$ (3)

both in ($\mu$mol m$^{-2}$ s$^{-1}$), and $a_{\text{NV}}$, $a_{\text{NJ}}$, $a_{\text{PV}}$, $a_{\text{PJ}}$ in ($\mu$molg$^{-1}$ s$^{-1}$) and $v_{\text{NV}}$, $v_{\text{NJ}}$, $v_{\text{PV}}$, $v_{\text{PJ}}$ in ($\mu$mol mg$^{-1}$ s$^{-1}$) empirical coefficients (see S1). The canopy-level photosynthetic capacity $V_{\text{cmax}}$ ($\mu$mol m$^{-2}$ s$^{-1}$) is estimated using the tree-level leaf area index $L$, taking into account within canopy gradients in light and photosynthetic capacity based on Lloyd et al. (2010). Nutrient optimisation is approximated using equations in Lloyd et al. (2010), with $M_a$ also dependent on the height of each tree ($H$) and the mean canopy height ($\overline{H_s}$):

$$M_a^* = M_a \cdot \exp \left[ a_H \cdot (H - \overline{H_s}) \right]$$ (4), with $a_H$ an empirical coefficient.

### 2.1.2 Respiration

Tree respiration includes a growth and a maintenance component, both computed daily. Growth respiration is considered as a constant fraction (0.25) of daily photosynthesis (Cannell and Thornley, 2000). Three different maintenance respiration formulations are allowed in TFS (S1, Respiration), but in this study we use the one described below. Leaf maintenance respiration $R_{\text{ml}}$ is estimated as a fraction of $V_{\text{cmax}}$ (Scheiter and Higgins, 2009):

$$R_{\text{ml}} = 0.015 V_{\text{cmax}}$$ (5)

Stem maintenance respiration is estimated from the sapwood volume ($V_S$) of a tree:

$$R_{\text{ms}} = \delta V_S$$ (6), with $\delta=39.6$ ($\mu$mol m$^{-3}$ s$^{-1}$) as reported in Ryan et al. (1994) for tropical trees.

Sapwood volume is estimated by inverting the pipe model and assuming that the ratio of leaf area to sapwood area ($\Phi_{LS}$) increases with the height and the wood density for tropical trees following (Calvo-Alvarado et al., 2008; Meinzer et al., 2008):

$$\Phi_{LS} = 0.5 \times (\lambda_1 + \lambda_2 \cdot H + \delta_1 + \delta_2 D_w)$$ (7),

with $\lambda_1=0.066$ m$^2$ cm$^{-2}$, $\lambda_2=0.017$ m cm$^{-2}$, $\delta_1=-0.18$ m$^2$ cm$^{-2}$ and $\delta_2=1.6$ cm$^3$ g$^{-1}$. 
Sapwood area ($m^2$) and volume ($m^3$) are then calculated from:

\[ S_A = L_A / \Phi_{LS} \] (8), with $L_A$ the total leaf area of the tree ($m^2$) and

\[ S_v = S_A \cdot (H - C_D) \] (9), with $C_D$ the crown depth (m)

Coarse root maintenance respiration $R_{mCR}$ is estimated as in Scheiter and Higgins (2009):

\[ R_{mCR} = 0.218 \beta_R \frac{R_{CR}}{\Phi_{CN}} \] (10)

where $\Phi_{CN}$ is the root C:N ratio estimated on the basis of the simulated $N_R$ assuming a dry weight carbon fraction of 0.5.

Fine root maintenance respiration $R_{mFR}$ is assumed to be equal to leaf respiration.

All respiratory components are corrected with the temperature dependence function of Tjoelker et al. (2001). The total maintenance respiration $R_m$ is then:

\[ R_m = R_{ml} + R_{ms} + R_{mCR} + R_{mFR} \] (11)

### 2.1.3 Stomatal Conductance

Initially, a maximum (no water stress) stomatal conductance, $g_{s,max}$ is calculated following Medlyn et al. (2011, 2012):

\[ g_{s,max} = g_0 + 1.6 \cdot (1 + \frac{g_1}{D_C}) \times \frac{A_c}{C_a} \] (12)

with $g_0$ (mol m$^{-2}$ s$^{-1}$) the minimum stomatal conductance, $g_1$ an empirical coefficient that represents the water use efficiency of the plant, and $D_C$ the leaf-to-atmosphere vapour pressure difference. Values of $g_0$ and $g_1$ that lead to the best model performance were different between sites, as indicated by the model calibration procedure. For the basin-wide simulations constant values of $g_0=0.020$ (mol m$^{-2}$ s$^{-1}$) and $g_1=5.0$ (-) were used, close to the estimates of Domingues et al. (2014). In future versions of the model, we anticipate that $g_0$ and $g_1$ will be related to other functional traits. The maximum stomatal conductance is subsequently reduced to the actual $g_s$ by multiplying the second term of equation 8 with a water stress coefficient.

In contrast to most ecosystem fluxes model, where photosynthetic rates are directly regulated by water availability (Scheiter and Higgins, 2009; Clark et al. 2011), we couple soil water deficit to canopy conductance by estimating a daily fractional available soil water content $\mathcal{I}$, for each i tree in the stand ($S_1$, Water Balance and Soil Water Stress). This term is then used to estimate the water stress $\gamma$ that has a direct effect on stomatal conductance, as also described in Keenan et al. (2010).
2.2 Study Sites & Simulations Set-up

Three sets of site data were used to explore the behaviour of the model. These include a set of three eddy flux measurements (EFM) sites, seven plots with intensive carbon balance and allocation measurements (IM), and 40 permanent measurement plots (PM).

2.2.1 Eddy flux (EFM) sites

Daily climate and energy flux data from three EFM sites (Caxiuanã [1.72S, 51.46W], Manaus [2.61S, 60.21W] and Tapajós [2.86S, 54.96W]) were used to assess the ability of the model to estimate canopy-level water fluxes. Data were obtained from the Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) project ([http://daac.ornl.gov/LBA/lba.shtml](http://daac.ornl.gov/LBA/lba.shtml)). In particular mean daily climate parameters including incoming radiation, temperature, precipitation, relative humidity and wind speed were used to force the model. Latent heat flux ($\lambda E$ in Wm$^{-2}$) was used to estimate a daily mean canopy conductance defined as $G_C = \frac{\lambda E}{D_C}$. The EFM data cover a period from 2001 to 2008 for Caxiuanã, from 2000 to 2005 for Manaus and from 2002 to 2004 for Tapajós. $G_C$ was only estimated for days with a complete diurnal record of $\lambda E$. At each one of the EFM sites the mean daily $G_C$ (mol m$^{-2}$ s$^{-1}$) was compared between observations and simulations. The model was initialized with size-class distribution and functional traits data from RAINFOR permanent plots located near the eddy flux towers. Specifically, CAX-06 inventory data were used for Caxiuanã, BNT-04 for Manaus, and TAP-55 for Tapajós. We note that the EFM sites are mainly found at the eastern part of Amazonia (Fig. 2) growing on low nutrient status soils.

The model was initially calibrated to the site specific values for $g_0$ and $g_1$ of equation 8 that gave the best performance. A Standardised Major Axis (SMA) regression, forced through zero was used to verify the ability of the model to simulate $G_C$, with a regression slope close to one indicating a good model performance.

2.2.2 Intensive measurement (IM) sites

The ability of the model to realistically simulate carbon fluxes at the tree-level is evaluated using data from the seven intensive measurement plots (Aragão et al., 2009; Malhi et al., 2009). These sites are amongst the intensively surveyed plots within the RAINFOR network (Fig. 2), where measurements of all major components of the C cycle are recorded (Malhi et al., 2009). At these plots, a detailed assessment of the carbon stocks is applied, and $\Pi_N$ allocation coefficients to different plant components are estimated (Aragão et al., 2009; Malhi et al. 2011; Doughty et al., 2013). These site-specific coefficients are used to calculate the amount of simulated $\Pi_N$ that is allocated to stems $\Pi_{N,s}$ (kgC y$^{-1}$).

The IM sites of interest include two plots at Agua Pudre in Colombia (AGP-01 & AGP-02), one (ALP-30) at Allpahuayo/Peru, one (BNT-04) at Manaus/Brazil, one in Caxiuanã /Brazil (CAX-06), one in Tambopata/Peru (TAM-05) and one in (TAP-55) Tapajós /Brazil. Based on data from Quesada et al. (2011), AGP-01, AGP-02, TAM-05 can be considered to be located on fertile soils, with the other
four plots on infertile ones. Available soil depth data (Quesada et al., 2011) and functional traits data (Fyllas et al., 2009) were used for site specific simulations. For all seven sites we estimated the observed average multi-annual growth rate (2000-2006) of each tree from forest census data, in order to compare it with the simulated $\Pi_{N,s}$.

The daily climate was extracted from the Princeton Global Meteorological Forcing Dataset (Sheffield et al., 2006). These simulations are used to validate the ability of the model to accurately estimate tree-level stem growth, under a given stand structure, a given climatic and soil profile and functional traits configuration of the established trees. Average observed stem growth rate (per 10 cm $d$ bins), expressed in carbon units (i.e. kg C y$^{-1}$), is compared with simulated $\Pi_{N,s}$ using the York method of best straight line, which holds when both $x$ and $y$ observations are subject to correlated errors that vary from point to point (York et al., 2004).

2.2.3 Permanent measurement (PM) sites

Inventory data from 40 RAINFOR permanent measurement plots (Fig. 2), including tree diameter and multiannual growth for all trees greater than 10 cm curated/managed in ForestPlots.net (Lopez-Gonzalez et al., 2009; Lopez-Gonzalez et al., 2011), are used to a) validate the ability of the model to accurately simulate stand-level carbon fluxes and b) explore patterns of $\Pi_G$, $\Pi_N$ and $C_U$ along the Amazonian climatic and soil nutrient availability gradient. The size class distribution within each PM site is used to initialise the stand structure of the model and simulate patterns of productivity for the 2000-2006 period. Climate data for the same period were used here with the first year again used as a spin-up period (Sheffield et al., 2006). For those 40 PM plots, sample distributions of the traits quartet are available (Fyllas et al., 2009) as well as a description of soil chemical and physical properties (Quesada et al., 2011).

At the PM sites the simulated stand-level aboveground $\Pi_N$ was compared with observed rates of aboveground growth ($\Delta B_{ABG}$ (kg C m$^{-2}$ y$^{-1}$)) for trees that survived during the 2000-2006 time period using a SMA regression. A second step was to explore the way $\Pi_G$, $\Pi_N$ and $C_U$ vary across an Amazon climatic and soil nutrient availability gradient (Quesada et al., 2010). The site scores of a principal components analysis (PCA) on the soil properties of the 40 PM plots (see Fyllas et al., 2009) are used to categorise plots along a nutrient availability gradient ($\Phi_1$), while the key climatic variables used were the annual mean temperature $T_A$ and annual total precipitation $P_A$. A Kendall correlation coefficient ($\tau$) was used to identify potential relationships of $\Pi_G$, $\Pi_N$ and $C_U$ with $T_A$, $P_A$ and $\Phi_1$, as in most cases non-linear associations were observed.

2.2.4 Randomisation Exercise

In order to explore a) the importance of including trait variability and thus functional diversity in our simulations and b) the importance of including constrains that are known to control the large scale patterns of Amazonian forest dynamics, we conducted a randomisation exercise with the model being run under four alternative set-ups at the 40 permanent RAINFOR plots. The first set-up denoted as var-tr is the variable-trait simulation with trait initialization based on the observed stand-level trait distribution as described in the previous paragraphs (default set-up). The second set-up, denoted as fix-
is a fixed-trait simulation with all trees having the same (dataset mean) values for each trait: This thus representing a single PFTs case. The third set-up (rand-trt) is a variable-trait simulation with trait initialization based on random values of the traits quartet as recorded in any individual along the 40 permanent plots. This setup thus ignores any potential patterns of functional trait biogeography, i.e. traits are not related to the environmental or edaphic conditions under which a tree is growing. The fourth setup (rand-tr-N) is a variable trait simulation where the photosynthetic capacity of an individual is only defined by its leaf N content and thus the NP co-limitation constraint is removed. These alternative set-ups were compared by considering both the slope and the R² of SMA regressions between the predicted and the observed Πₙₕ.

3. Results

3.1 Canopy conductance simulations at the EFM sites

Values of best model performance for \( g₀ \) and \( g₁ \) were different between sites, with \( g₀=0.035 \) (mol m⁻² s⁻¹) and \( g₁=7.5 \) at Caxiuanã, \( g₀=0.035 \) and \( g₁=7.0 \) at Manaus with \( g₀=0.01 \) and \( g₁=2.5 \) these being somewhat lower than the estimates of Domingues et al. (2013) at Tapajós. Simulated \( Gc \) was underestimated for Caxiuanã (\( α=0.85±0.05 \)) and Manaus (\( α=0.90±0.02 \)), with the model overestimating \( Gc \) in Tapajós (\( α=1.28±0.04 \), but exhibiting an overall adequate performance (Fig. 3). For simulations at the IM and the PM sites, constant values of \( g₀=0.02 \) (mol m⁻² s⁻¹) and \( g₁=5 \) (-) were used, which are found within the range of values in the EFM sites and reported estimates (Medlyn et al. 2012; Domingues et al., 2013).

3.2 Stem growth rate simulations at the IM sites

The mean simulated stem growth rate \( Πₙₕ \) of each tree in the seven IM plots was compared with the observed aboveground biomass gains (\( ΔΒ_{ABG} \)) for the 2000-2006 period. An accurate simulation of \( Πₙₕ \) can be seen for small size classes, but with greater differences between the observed and the simulated multi-annual growth found for bigger trees (Fig. 4). At infertile ALP-30, the estimate slope of the York model indicated an overestimation of aboveground production (\( α=1.18±0.06 \)), driven mainly by an overestimation of the mid-size classes. At BNT-04 the model underestimated the overall growth (\( α=0.82±0.03 \)). Aboveground growth was overestimated in CAX-06 (1.11±0.07). At TAP-55 (\( α=1.44±0.15 \)) the model underestimated aboveground production (0.90±0.06). At fertile AGP-01 (\( α=1.36±0.08 \)) and AGP-02 (\( α=1.25±0.05 \)) an overestimation of aboveground productivity was observed although with simulations of most size classes falling within the observed ranges. At TAM-05 (\( α=0.79±0.07 \)) though, the simulated aboveground growth was underestimated with the overall slope driven by divergences in smaller size classes. The range and distribution of \( Πₙ \) allocation to stem growth is adequately captured by TFS as summarised in Fig B1.
3.3 GPP, NPP and CUE simulations at the PM sites

Simulated stand-level aboveground net primary productivity $\Pi_{\text{G}}$ was positively associated with observed changes in aboveground biomass of trees that survived in the PM plots over the 2000-2006 period $\Delta B_{\text{ABG}}$, with an $R^2=0.42$, suggesting an adequate model behaviour (Fig. 5). A summary of simulated stand-level $\Pi_{\text{G}}$, $\Pi_{\text{N}}$, and $C_{\text{U}}$ relationship to key environmental drivers is given in Table 1 (see also Fig B2). $\Pi_{\text{G}}$ and $\Pi_{\text{N}}$, and $C_{\text{U}}$ were not associated with temperature. However, all three measures of stand level productivity were positively related to annual precipitation and soil nutrient availability.

3.4 Randomisation Exercise Simulations

Results from the randomisation exercise (Fig. 6) found the fully constrained default set-up (var-tr) to have the best predictive performance ($R^2=0.42$ with a SMA slope $a=0.92$). This is as compared to the fixed trait simulations (fix-tr) single PFT parameterization with a decreased predictive ability of TFS ($R^2=0.29$, $a=0.82$) and an overall higher mean predicted aboveground productivity. Not accounting for the site specific distribution of the traits quartet, i.e. bypassing potential biogeographic patterns of functional diversity and/or environmental-trait interactions (rand-tr) also reduced the predictive ability of the model ($R^2=0.29$, $a=0.74$). Finally the random trait no NP co-limitation set-up (rand-tr-N) similarly lead to an inferior model performance ($R^2=0.33$, $a=0.88$) and with the highest mean simulated aboveground productivity.

4. Discussion

We report here on the core components of an individual-based model that has been developed in order to help better understand the patterns revealed by recent integrated measurements of climate, soils, functional diversity and stand dynamics for a wide range of forests across the Amazon Basin. In its current setup the model does not explicitly simulate regeneration and mortality dynamics but rather uses the observed size distribution of trees at the study sites, thus taking into account stand structure and functional trait variability as observed along the main climatic and edaphic axes of the Amazon Basin. With the current setup we were able to reproduce the tree- and stand- level $\Pi_{\text{N}}$ patterns found across Amazonia and to explore for potential environmental controls over stand-level $\Pi_{\text{G}}$, $\Pi_{\text{N}}$ and $C_{\text{U}}$.

4.1 Scientific Outcomes

Our simulations found no association of stand level gross primary productivity ($\Pi_{\text{G}}$) with temperature, probably due to the relatively small range of variation of temperature across our plots. $\Pi_{\text{G}}$ decreased until an annual temperature of approximately 26°C but remained relative constant above this point (Table 1, Fig. A.2.2). However, our simulations suggest that a strong association of $\Pi_{\text{G}}$ with the annual precipitation and soil nutrient availability of the plots. $\Pi_{\text{G}}$ was positively related to annual precipitation over the entire range observed in the 40 PM plots. The association of $\Pi_{\text{G}}$ with the nutrient availability axis is in agreement with fertilisation experiments showing an increase with nutrient supply (Giardina et al., 2003). In our Basin-wide examination of $\Pi_{\text{G}}$ the soil nutrient availability and
stand structure gradients are not, however, independent (Quesada et al., 2012), as in the RAINFOR network permanent plots it has been observed that bigger/older trees are more abundant on eastern infertile forests, where soil physical conditions can support a bigger tree size (Baker et al., 2009) with a lower risk of trees being uprooted (Chao et al., 2009). Bigger trees generally support a greater foliage area and thus could significantly contribute to the overall carbon assimilation of the stand. However, bigger trees on infertile plots are generally characterised by lower leaf nutrient concentrations (Fyllas et al., 2009) and thus slower assimilation rates (Reich et al., 1994; Domingues et al., 2010). On the other hand a higher abundance of smaller trees with higher gas exchange rates is observed on more dynamic, fertile plots. Ultimately this indicates that stand structure should be specifically taken into account when simulating $\Pi_G$ in tropical forests, and thus individual-based models could significantly contribute towards a deeper understanding of the functioning and sensitivity of these ecosystems.

In our simulations stand-level net primary productivity ($\Pi_N$) showed no significant association to annual temperature but increased with soil nutrient availability and annual precipitation (Table 1, Fig. A.2.2). Our $\Pi_N$ simulations are in agreement with field observations of increasing aboveground wood productivity with precipitation (Quesada et al., 2012). Based on TFS parameterisation, photosynthetic rates are expected to be higher at a greater soil nutrient availability due to associated higher leaf N and P concentration (Fyllas et al., 2009; Domingues et al., 2010). Using a similar parameterisation for a “sun and shade” big leaf model, Mercado et al. (2011) found an increase in net canopy assimilation rate with leaf P content in agreement with our positive association between $\Pi_N$ and soil nutrient availability. Their simulated $\Pi_G$ accounted for approximately 0.30 of the observed wood productivity in 33 study plots, and thus the $R^2$=0.42 between simulated $\Pi_N$ and aboveground growth found here suggest a marginally improved model behaviour. It should be noted that our definition of soil nutrient availability ($\Phi_1$), based on the PCA analysis in Quesada et al. (2010), directly relates to soil P content. As shown first in the analysis of Quesada et al. (2012), where data from almost 60 plots were considered, aboveground $\Pi_N$ is positively related to soil P content in lowland tropical forest. The increased $\Pi_N$ in fertile environments (apart from the higher $\Pi_G$) seems to be enhanced by the greater abundance of small trees there. As tree size increases maintenance respiration likely “consumes” an increasing proportion of assimilated carbon, and thus at large size classes the proportion of trees which have enough carbon to allocate to growth decreases (Givnish, 1988; Cavaleri et al., 2008). This is in line with the negative relationship between coarse wood production and maximum height documented for some Amazonian trees (Baker et al., 2009).

In our simulations carbon use efficiency ($C_U$) ranged from 0.43 to 0.54. Recent research suggests that the $C_U$ is not as constant as had been previously suggested (De Lucia et al., 2007; Zhang et al., 2009). For example the meta-analysis of De Lucia et al. (2007) found that $C_U$ varies from 0.23 to 0.83 in different forest types. Our average estimate of $C_U=0.51$ is, however, above the range of reported in Malhi (2012). Zhang and colleagues (2009) identified a negative trend of the $\Pi_N/\Pi_G$ ratio with temperature at the range of 20 to 30 °C, as also simulated here especially above 26°C (Fig A2.2). Simulated $C_U$ increased with soil nutrient availability, being marginally lower at infertile (0.48) compared to fertile (0.50) plots. This is attributable to smaller size class trees (with lower relative respiratory costs) constituting a greater proportion of the total stand biomass on higher nutrient status soils. One factor relating to soil nutrient availability but not included in the current version is an implicit consideration of the respiratory costs of plant nutrient uptake (Lambers et al., 2008) either directly, or through other
processes such as organic acid exudation (Jones et al., 2009) or the symbiotic associations (Duponnois et al., 2012.) One would expect these costs to be proportionally higher for stand of a low nutrient status, especially with regard to P (Quesada et al., 2012).

4.2 Practical implications

The modelling of tropical forest carbon fluxes and stand dynamics has traditionally involved approaches aimed at a balance between simplicity, computational economy, and complexity. On one hand, the enormous biological and biogeochemical heterogeneity of tropical forests (Townsend et al., 2008) places special importance on how modelers prioritise both the amount and the detail of processes that should be included to capture the main controls and feedbacks. On the other hand, the finding that Amazonia is dominated by just 227 tree species (ter Steege et al., 2013) implies that most biogeochemical cycling in the world’s largest tropical forest is performed by a tiny sliver of its diversity.

At one end of the complexity spectrum are individual-based models which are able to properly simulate population dynamics and thus lags due to demography. Individual-based models of tropical forests have traditionally focused on realistically representing the light environment (TROLL - Chave, 1999) or grouping tree species on the basis of their different responses to environmental resources as suggested by field observations (FORMIND - Kohler & Huth, 1998, LPJ-GUESS – Helly et al., 2006). At the other end of the complexity spectrum are DGVMs which simulate population dynamics more simplistically (but see Moorcroft et al., 2001; Scheiter & Higgins, 2009). Using a DGVM model Verheijen et al. (2013) allowed for within-PFT climate-driven trait variation to occur and achieved an improvement of the predicted vegetative biomass and PFT distribution patterns. A similar rationale was followed in Wang et al. (2012) where it was shown that the inclusion of multi-trait covariance in DGVM can be used to constrain model parameters and reduce uncertainties in simulated ecosystem productivity. Fisher et al. (2010) applied the individual-based Ecosystem Demography model (Moorcroft et al., 2001), and showed that by varying traits related to demographic processes, forest and biomass dynamics exhibited a wide range of responses to climate forcing.

Most of the above approaches have used discrete PFTs to represent tree species and functional diversity. These studies suggest that by allowing for within PFT trait variability a more plastic and realistic response to the relevant environmental drivers is observed. In contrast to the above, TFS replaces the use of PFTs with traits distributions, following a different model philosophy and architecture using the concept of multidimensional trait continua. In particular, considering functional diversity to be expressed by a multidimensional trait space, the use of PFTs selects a number of clusters where the central vector defines the average trait values of each PFT (Fyllas et al., 2012). Recent studies (Verheijen et al., 2013; Wang et al., 2012) allow for the average trait values to be shifted based on empirical climatic and/or trait inter-correlation functions. In contrast, the use of trait continua does not cluster the multidimensional trait space but rather allows any realistic trait combination (as suggested by the limited sampling of the actual population) to be simulated. Successful trait combinations under given environmental conditions may then be expected to emerge as a by-product of model dynamics (Higgins et al., 2014). A similar to TFS representation of functional diversity has been implemented in the aDGVM model (Scheiter & Higgins, 2009; Scheiter et al., 2013) where the importance of including trait-variability in simulations of vegetation dynamics has also been highlighted. In TFS, variable-trait ($R^2=0.42$) simulations led to a better model performance compared to fixed-trait ($R^2=0.29$) simulations.
Thus including functional diversity in simulations of vegetation dynamics is expected not only to suggest less vulnerable communities under changing climatic conditions (Fauset et al., 2012; Scheiter et al., 2013) but also, it seems, to better describe the current patterns of key ecosystem properties like aboveground productivity.

A few modelling studies that implement a similar traits continua approach have recently been published. Scheiter & Higgins (2009) were the first to develop an individual-based framework that eschews the use of PFTs and allows for plants to allocate carbon as a function of local environmental conditions. Falster and colleagues (2011) presented a model where they used leaf economic strategy, height, wood density and seed size to scale-up from individual scale processes to landscape predictions. Pavlick et al. (2013) applied an interesting approach where they used 15 traits to incorporate trait diversity within plant community in a DGVM. The rationale of the above models is that they allow different plant functional strategies to be available in a specific location with given environmental conditions (for example a grid cell), and that by setting up a set of functional trade-offs they “filter out” poorly adapted trait combinations from the community. This is effectively an implementation of ideas arising from the environmental filtering/community assembly theory to predict an optimum plant community at a given location (Keddy, 1992; Scheiter et al., 2013; Fortunel et al., 2014). By contrast, drawing on recent findings on the processes controlling Amazonian forest dynamics, we have here attempted to incorporate within TFS the relevant observed associations between functional trait diversity, stand-structure and soil physical and chemical properties (Fyllas et al., 2009, Quesada et al., 2012). Although there are similarities with some of the more recent models discussed above to our knowledge this is the first time all these linkages have been represented in a single modelling framework. Our approach has been made possible (and thus differs from others) because of the type and quantity of observational constraints used. For example in any given plot we do not force the model to select some “optimum” trait combination based on the prevailing environmental conditions, but we rather assume that the observed trait distribution reflects that of the evolutionary stable community structure occurring at each site. Similarly we don’t require the model to predict what the optimum tree-size class distribution would be. Rather, we initialize simulations with what is observed.

We have here employed this implementation as our primarily aim in this first instance has been to validate the predictive ability of the model at some extensively monitored Amazonian plots.

Even with these prescribed constraints, the trait randomization exercise yielded some interesting outputs regarding the importance of trait variability in simulations of forest dynamics. As already discussed the default variable-trait (var-tr) simulations gave the best TFS performance in terms of predicting patterns of aboveground production at the 40 permanent measurement plots with fixed trait (fix-tr) TFS simulations showing a lower predictive ability and an overall higher mean $\Pi_{AN}$. This pattern of trait variability reducing above-ground biomass is in contrast with a similar simulation from Scheiter et al. (2013), where variable trait simulations gave rise to a higher mean biomass because of an increased chance of selecting a trait combination allowing trees to grow larger. This difference arises from the photosynthesis NP co-limitation constraint hardwired into the current version of TFS as the use of the Amazon wide mean $N_L$ and $P_L$ values, leads inevitably to universally phosphorus limited estimates of $V_{\text{cmax}}$ and $J_{\text{max}}$ that reduce the overall predictive ability of the model. Indeed, when the NP co-limitation is removed, the variable trait simulations (rand-tr-N) do actually yield the highest $\Pi_{AN}$ estimates. Finally the random variable trait setup (rand-tr) resulted again in a similarly poor TFS
behavior ($R^2=0.29$), emphasizing the importance of potential environment – trait interactions in accounting for between-stand structural differences. In other words, in the modelling tropical forest dynamics it is clear that trait distributions cannot be used without a consideration of how they may be shifted by the local growing conditions.

The four functional traits used in the current version of TFS, i.e. leaf dry mass per area, leaf nitrogen and phosphorous concentrations and wood density, are directly related to the rates of tree photosynthesis and respiration. For that reason they provide a stable basis that should allow alternative ecological strategies based on well known trade-offs such as the “growth vs. survival” to be implemented in trait-based vegetation dynamics models. These four traits have been extensively studied around 70 plots in the Amazon and their patterns of variation and inter-correlation have been analysed (Baker et al., 2009; Fyllas et al., 2009; Patiño et al., 2009; Patiño et al., 2012). For the purpose of this study, it is important to know the within stand variation of the functional traits used, i.e. the trait values at the individual level across different plots. Additional functional traits that were considered but not included as base traits in this version of TFS were the seed size and the leaf area to sapwood area ratio. Seed size is an important functional trait that expresses a tolerance vs. fecundity trade-off, with seed size trading-off with seed number and with larger seed species being more tolerant at more stressful places (Muller-Landau, 2010). However, data on seed size are usually available at the species level, i.e. intraspecific variation is not usually recorded, and thus this kind of data cannot be included in the current version of TFS. The leaf area to sapwood area ratio, $\Phi_{LS}$, is an important trait that can be used to constrain the hydraulic architecture of trees (Meinzer et al., 2008). Here $\Phi_{LS}$ is expressed as a function of $D_w$ and $H$ (equation 7) and it is not used as an independent (base) trait. Future version of TFS will include this aspect of functional variability, but for this first study we have selected just a small set of key traits in order to maintain a relative simple model structure.

Like most modelling efforts, TFS represents work in progress. We identify three particularly promising avenues for future improvements. Firstly, discrepancies between the observed and simulated stem level growth rates, particularly in larger size classes, could result from the allometric equations used to estimate aboveground biomass and growth not being species or size specific. The allometric equations used here express a generic height ($H$) vs. $d$ relationship for Amazonia, without taking into account habitat and species differences, so a more accurate representation of tree architecture would probably result in better biomass growth estimation. Indeed, $H-d$ relationships do vary significantly among species (King, 1996; Poorter et al., 2006) and across regions (Nogueira et al., 2008; Feldpausch et al., 2011; Goodman et al., 2013). An additional source of bias when estimating stem-level growth rates could be related to the uniform (static) allocation coefficient used in this study. For example, Litton et al. (2007) showed that allocation to aboveground tree biomass components increases with age and the availability of resources. Furthermore, Castanho et al. (2013) improved the predictions of a DGVM by adjusting allocation coefficients based on soil texture. Such ontogenetic and/or resource based shifts in patterns of carbon allocation could be potentially modelled through the use of dynamic allocation schemes (Friedlingstein et al., 1999; Franklin et al., 2012).

The importance of realistically representing autotrophic respiration processes in models of vegetation dynamics is also highlighted here. Modelling respiration has proven to be a difficult task (Cannell and Thornley, 2000), and accurately representations of this component is of great importance.
for understanding the global C cycle (Valentini et al., 2000). For example the way respiration is represented in DGVMs could have a substantial control over the way the dynamics of Amazonian forest under scenarios of climatic change are simulated (Huntingford et al., 2004; Galbraith et al., 2010). Nitrogen content of plant tissue has been proven a good predictor of respiration rates (Reich et al., 2008). However, Mori et al. (2010) suggested a mixed-power equation where the exponent varies from 1 to 3/4 as size increases. Both the Reich and Mori models are implemented in TFS, but we found that a third method, combining the size and nitrogen control, performed better. Thus we suggest that an amalgamation of those two approaches could provide a better way to estimate respiration fluxes in the new generation of dynamic vegetation models. In addition leaf phosphorous content seems to constrain respiration rates stronger than nitrogen content in some tropical forests (Meir et al., 2001; Meir and Grace, 2002), and thus inclusion of a phosphorus constraint in future equations of leaf respiration could increase their realism.

Finally, discrepancies in the observed versus the simulated canopy conductance $G_c$ could result from the parameterisation of the stomata conductance model of Medlyn et al. (2011). The estimates for $g_0$ and $g_1$ used in the 40 PM plots simulations were taken as constant. However, Medlyn et al. (2011) suggested that $g_0$ and $g_1$ could vary with functional group. Thus the Amazon wide parameterisation used here should be replaced with local level estimates when appropriate gas exchange data are available, and ultimately with estimates based on linked functional traits as evidenced through recently documented associations between structural characteristics such as wood density and leaf area: sapwood ratio with leaf physiological traits such as $M_a$ and leaf $^{13}$C/$^{12}$C ratio (Patiño et al., 2012), although we also note that the extent of such structural/physiological linkages remains the subject of debate (Baraloto et al., 2010b). Alternative stomatal closure equations as a function of soil water availability (Harris et al. 2004) should also be tested along with the conductance model in future versions of the model.

5. Conclusions

We set out to develop a modelling framework for tropical forests that is relatively simple yet adequately complex to capture the main ecological gradients in the world’s most extensive tropical forest. Our study places special emphasis on processes highlighted by recent field studies to strongly influence Amazonian forest dynamics, especially functional trait diversity and its association with multiple soil properties (Fyllas et al., 2009). In summary TFS is characterised by a relatively simple setup, which is capable to reproduce water and carbon fluxes as observed at both daily and multi-annual time scales. TFS represents an important link between inventory data, and large scale models with the incorporation of the continuum of plant strategies, through the inclusion of trait distributions providing a step towards better representing diversity in vegetation modelling (Lavorel et al., 2007), representing important processes and trait variation that cannot be adequately accounted for by a DGVM approach to vegetation modelling. Since TFS is based heavily on measured data, the model is well suited to testing hypotheses related to the present day Amazon biogeography and biogeochemical fluxes.
6. Code availability

The JAVA source code can be obtained upon request. Contact: nfyllas@gmail.com

Acknowledgments

This research was supported by a Marie Curie Intra-European Fellowship within the 7th European Community Framework Programme to NM Fyllas. Manuel Gloor and Lina Mercado were funded by the AMAZONICA NERC consortium grant. OP is supported by an ERC Advanced Grant and (as for JL) a Royal Society Wolfson Research Merit Award. The work of RAINFOR between 2000 and 2006 was supported primarily by the UK Natural Environment Research Council and the European Community Framework Programme through grants to OP, YM, and JL. Natalia Restrepo Coupe provided the gap filled eddy flux tower data. Gabriela Lopez-Gonzalez and Sophie Fauzet helped with data preparation. The Missouri Botanical Garden, Rodolfo Vasquez, Abel Monteagudo, Nigel Pitman, Adriana Prieto, Agustin Rudas, Natalino Silva, Gerardo Aymard, Chiqui Arroyo and Alejandro Araujo Murakami contributed permanent plot data to the RAINFOR network.
References


21


Table 1. Kendall correlation coefficients (τ) and associated significance levels (p) between simulated gross primary productivity (\(\Pi_G\)), net primary productivity (\(\Pi_N\)), carbon use efficiency (\(C_U\)) and key environmental factors.

<table>
<thead>
<tr>
<th>Mean Annual Temperature - (T_A) (°C)</th>
<th>(\Pi_G) (kgC m(^{-2}) y(^{-1}))</th>
<th>(\Pi_N) (kgC m(^{-2}) y(^{-1}))</th>
<th>(C_U) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau = -0.17)</td>
<td>(\tau = -0.21)</td>
<td>(\tau = -0.11)</td>
<td></td>
</tr>
<tr>
<td>(p = 0.131)</td>
<td>(p = 0.065)</td>
<td>(p = 0.33)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Annual Precipitation (P_A) (mm)</th>
<th>(\tau = 0.54)</th>
<th>(\tau = 0.60)</th>
<th>(\tau = 0.36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p &lt; 0.001)</td>
<td>(p &lt; 0.001)</td>
<td>(p = 0.002)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Soil nutrient availability (\Phi_1) (PCA Axis 1)</th>
<th>(\tau = 0.48)</th>
<th>(\tau = 0.50)</th>
<th>(\tau = 0.39)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p &lt; 0.001)</td>
<td>(p &lt; 0.001)</td>
<td>(p &lt; 0.001)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: The five basic components of the model and information flow among them. Tree by tree traits and size initialisation takes place at the beginning of each simulation. Carbon and water fluxes, as well as gross and net primary productivity are estimated daily.
Figure 2. Geographic distribution of study sites. Dark grey triangles indicate the three eddy flux tower sites (with local names), light gray circles indicate the seven intensive measurement plots (with plot codes), and crosses indicate the coordinates of the 40 RAINFOR permanent measurement plots.
Figure 3: Simulated against observed mean daily canopy conductance $G_C$ for the three sites with eddy flux data. The broken line represents an 1:1 relationship and the continuous line illustrates a standardised major axis (SMA) regression.
Figure 4: Simulated stem growth rate $\Pi_{N,s}$ against observed aboveground biomass change $\Delta B_{ABG}$ for different size classes for the 2000-2006 period. Upper panel: infertile plots. Lower panel: fertile plots. The broken line represents an 1:1 relationship continuous. The continuous line illustrates the straight line fit using the York method (see text for details).
Figure 5: Simulated stand-level aboveground net primary productivity (Π_{AN}) against observed stand-level aboveground biomass growth (ΔB_{ABG}) of surviving trees, at the 40 PM plots. The line illustrates a SMA regression of α=0.92 (0.72...1.18) and R²=0.42. Red dots indicate high nutrient availability and blue dots indicate low nutrient availability plots.
Figure 6: Summary of the randomisation exercise simulations. a) Simulated stand-level aboveground net primary productivity ($\Pi_{AN}$) against observed stand-level aboveground biomass growth ($\Delta B_{ABG}$) for the four different set-ups. The slope of the SMA (a) and the adjusted $R^2$ are given in the parentheses for each set-up. Different colours indicate different setups b) Simulated Amazon-wide aboveground net primary productivity ($\Pi_{AN}$) for the four different set-ups.