Modelling carbon and water balance of Eucalyptus plantations at regional scale: Effect of climate, soil and genotypes

Ahmed Attia, Yann Nouvellon, Santiago Cuadra, Osvaldo Cabral, Jean-Paul Laclau, Joannès Guillemot, Otavio Campoe, Jose-Luiz Stape, Marcelo Galdos, Rubens Lamparelli, et al.

Modelling carbon and water balance of *Eucalyptus* plantations at regional scale: effect of climate, soil and genotypes

Ahmed Attia [1], Yann Nouvellon [2,3,4], Santiago Cuadra [5], Osvaldo Cabral [6], Jean-Paul Laclau [2,3,4], Joannès Guillemot [2,3,4], Otavio Campoe [7,8], José-Luiz Stape [8], Marcelo Galdos [9], Rubens Lamparelli[1], Guerric le Maire[1,2,3]*

(1) UNICAMP, Campinas, SP, CEP: 13083-860, Brazil,

(2) CIRAD, UMR Eco&Sols, F-34398 Montpellier, France,

(3) Eco&Sols, Univ Montpellier, CIRAD, INRA, IRD, Montpellier, SupAgro, Montpellier, France,

(4) ESALQ, Universidade de São Paulo, Piracicaba, SP, CEP 13418-900, Brazil,

(5) EMBRAPA CNPTIA, Campinas, SP, CEP: 13083-886, Brazil,

(6) EMBRAPA Meio Ambiente, Jaguariuna,SP, CEP: 13820-000, Brazil,

(7) Federal University of Lavras – UFLA, Lavras, MG, CEP: 37.200-000, Brazil

(8) UNESP-FCA, Botucatu, SP, CEP 18.610-300, Brazil

(9) Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, LS2 9JT Leeds, UK

* corresponding author

Abstract

Carbon and water budgets of forest plantations are spatially and temporally variable and hardly empirically predictable. We applied G'DAY, a process-based ecophysiological model, to simulate carbon and water budgets and stem biomass production of *Eucalyptus* plantations in São Paulo State, Brazil. Our main objective was to assess the drivers of spatial variability in plantation production at regional scale. We followed a multi-site calibration approach: the model was first parameterized using a detailed experimental dataset. Then a subset of the parameters were re-calibrated on two independent experimental datasets. An additional genotype-specific calibration of a subset of parameters was performed. Model predictions of key carbon-related variables (e.g., gross primary production, leaf area index and stem biomass) and key water-related variables (e.g., plant available water and evapotranspiration) agreed closely with measurements. Application of the model across *ca.* 27,500 ha of forests planted with different genotypes of *Eucalyptus* indicated that the model was able to capture 89% of stem biomass variability measured at different ages. Several factors controlling *Eucalyptus* production variability in time and space were grouped in three categories: soil, climate, and the planted genotype. Modelling analysis showed that calibrating the model for genotypic differences was critical for stem biomass prediction at regional scale, but that taking into account climate and soil variability significantly improved the results. We conclude that application of process-based models at regional scale can be used for accurate predictions of *Eucalyptus* production, provided that an accurate calibration of the model for key genotype-specific parameters is conducted.

Keywords: *Eucalyptus* plantations, ecophysiological model, G'DAY, optimization, productivity

1. Introduction

Forest vegetation plays a major role in determining the state of the global climate system and carbon cycle, both of which are undergoing significant anthropogenic perturbations. Among forest vegetation types, *Eucalyptus* is the most widely planted tropical hardwood genus, covering more than 20 million ha worldwide (Albaugh *et al.*, 2013). In Brazil, forest plantations cover 7.6 million ha, of which 72% are planted with high-productivity *Eucalyptus* clones (average annual increment of 40 m^3 ha^{-1} yr^{-1} of roundwood, ranging from 25 to 60 m^3 ha^{-1} yr^{-1} (Gonçalves *et al.*, 2013)). *Eucalyptus* plantations in Brazil are generally managed in 6-7 years rotation, with canopy closure occurring within 2-3 years after planting. The relatively low susceptibility to pests and diseases, rapid growth and high productivity, adaptability to varying soil and climate, and adequate fiber quality for the industry explain the expansion of commercial *Eucalyptus* tropical plantations worldwide.

Water deficit, nutrient deficiency, soil type and compaction are the main drivers of *Eucalyptus* plantation functioning in Southern Brazil (Ryan *et al.*, 2010; Stape *et al.*, 2010). However, predicting how forests grow in response to soil and climate constraints, and determining their carbon storage capacities remains a key challenge for modelers. Forest productivity is driven by complex interactions and feedbacks among biological mechanisms. Furthermore, spatial variability in resources supplies, management, and characteristics of the genetic tree material critically influence forest productivity.

Ecophysiological process-based models that simulate water and carbon fluxes in forests proved to be useful tools to formalize biophysical hypotheses on forest functioning and to test for the importance of environmental drivers on productivity. Over the last two decades, a range of process-based models was developed, varying in resolution, complexity, generality, and applicability (Mäkelä *et al.*, 2000; Battaglia *et al.*, 2004; Corbeels *et al.*, 2005a; Dufrêne *et al.*, 2005; Marsden *et al.*, 2013). Some of these models were developed for research

purposes to understand and quantify carbon and water cycling at fine time-scale (Dufrêne *et al.*, 2005), others were constructed as management tools, in close collaboration with intended end-users (Landsberg and Waring, 1997; Sands *et al.*, 2000; Almeida *et al.*, 2004b). Importantly, wood productivity relates on both the net amount of carbon (C) sequestered by trees (net C balance) and the way this C is allocated among tree organs.

The modelling of *Eucalyptus* plantations at regional scale is of critical economic importance and has been the focus of a growing body of studies (e.g., Almeida *et al.* (2004a). Marsden *et al.* (2013) modified the Generic Decomposition And Yield model (G'DAY) to simulate the productivity of 16 *Eucalyptus* plantation stands of the same genotype in São Paulo region, Brazil. It was shown that soil water holding capacity explains a large part of spatial variability in tree growth rate and biomass productivity in this area. This study also highlighted two important limitations that currently hinder regional modelling using G'DAY. First, the impact of water availability on the C allocation toward root growth, leaf growth and litterfall production is not correctly represented in the model. Because leaf and root areas are critical C and water exchange surfaces, further improvements of the G'DAY C allocation scheme is required for accurate simulations of carbon and water fluxes and productivity along soil water gradients. Second, the diversity of species and hybrid *Eucalyptus* genotypes used in Brazil's plantations was shown to be challenging for large-scale modelling (Almeida *et al.* (2010). *Eucalyptus* materials may indeed strongly differ in the control of several key processes, such as photosynthesis and light use efficiency (Warrier and Venkataramanan, 2010; le Maire *et al.*, accepted) and carbon allocation (Ngugi *et al.*, 2003). However, a detailed parameterization of process-based models for each genotype from field measurements is today out of reach, because physiological and C partitioning measurements along plantation rotation are missing for most of the numerous cultivated genotypes. Previous studies therefore generally used a unique model parameterization over a whole region, either

by using generic parameterization at the level of plant functional type (e.g., Almeida *et al.*, 2004a), or by calibrating the model based on some genotype-specific measurements (Almeida *et al.*, 2004a; Gonzalez-Benecke *et al.*, 2016). Calibration can be conducted by using independent measurements to parameterize the equations of the model, or by inverting the model to constrain the values of these parameters (Gonzalez-Benecke *et al.*, 2016). This last option is considered more reliable for application purposes, since the model simulations are constrained within the range of a plausible domain, even though equifinality of the simulations (i.e., the possibility of 'getting the right answers for the wrong reason') may lead to a wrong interpretation of the results. Although using a unique parameterization for regional modelling proved useful to capture the response of productivity to large environmental gradient, it creates an additional uncertainty that need to be quantified if the simulations are to be used for management purposes.

The overall objectives of the present study was 1) to predict temporal and spatial variations of stem biomass production in *Eucalyptus* plantations in the São Paulo state of Brazil using the G'DAY model 2) to use the validated model to test for the effect of genotype, soil and climate on plantation wood productivity at regional scale. Efforts were done to improve the ability of the model in capturing the responses of forest functioning and C allocation to water stress. We used a multi-site calibration approach to provide an end-to-end calibration and application scheme from local to regional scale. G'DAY was first optimized using detailed measurements from one experimental site. A subset of "genotype-specific" parameters was then calibrated at two other sites where physiological data were also available, and presenting differences in genotype and location. Finally a calibration and application of the model at large spatio-temporal scale was performed using a network of 1472 stands. Constrained simulations were then conducted using the final model to highlight the drivers of productivity at regional scale.

2. Materials and Methods

2.1. The four *Eucalyptus* datasets

The site of the first experimental dataset (DATASET 1) was a commercial plantation of *Eucalyptus grandis*, located at 22°58'04''S, 48°43'40''W, 750 m.a.s.l, planted in November 2009 at a 3x2 m spacing, and monitored continuously since then as part of the EUCFLUX project (http://www.ipef.br/eucflux/en/, (Nouvellon *et al.*, 2010; Nouvellon *et al.*, 2018)). The average annual precipitation was 1540 mm from 2008 to 2016, with an average temperature of 19.3°C, and a wet hot summer from October to May and dry cold winter from June to September. A detailed description of the site is given in (Campoe *et al.*, 2013) and (Christina *et al.*, 2016; Christina *et al.*, 2017)Measurements included volumetric soil moisture content, tree height, leaf area index (LAI), H_2O and CO_2 gas exchanges between the ecosystem and the atmosphere, leaf, bark and branch litterfall, biomass of all tree compartments, leaf photosynthesis. Daily net ecosystem exchange (NEE) and evapotranspiration (ET) were obtained from a flux-tower, using the Eddy-covariance method (Christina *et al.*, 2017; Nouvellon *et al.*, 2018; Vezy *et al.*, 2018). Gross primary productivity (GPP) was estimated from NEE and meteorological data using the standard computation from Reichstein *et al.* (2005). Soil water content was measured using calibrated CS616 probes (Campbell Scientific Inc., Logan, UT, USA) inserted at 0.15, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 m soil depths. Soil is sandy loam with an average maximum plant available water (PAWmax) of 92 mm m^{-1} (Christina *et al.*, 2017). Leaf area index was measured by combining destructive measurements and dendrometric inventories, and leaf and bark litterfall was collected monthly using 48 litterfall traps of 0.52 by 0.52 m placed in the field following a voronoï scheme to sample every distance to trees. Branch litterfall was collected monthly on four 6.6

142 m^2 area. Litter samples were dried, weighted, and values of dry mass by unit area were
143 computed (gDM/m^2).

144 The site of the second experimental dataset (DATASET 2), located at Santa Rita
145 municipality (21°35' 48" S, 47°36' 0" W, 761 m.a.s.l.) was planted on April 2004 with a clone
146 of *E. grandis* × *E. urophylla*, at a 3 x 2.4 m spacing. The average annual precipitation is 1500
147 mm, and mean temperature was 22.1°C. Data collection started two years after planting on
148 March 2006 until March 2008. Collected data included measurements of soil water content,
149 LAI, biomass of leaves and stem, fully described in Cabral *et al.* (2010) and Cabral *et al.*
150 (2011). Soil moisture content was measured at 0.1, 0.3, 0.5, 0.7, and 1 m soil depth. The soil
151 was sandy, with PAWmax of 50 mm m^{-1} (Cabral *et al.*, 2011). Daily LAI values was
152 computed from the reflection coefficient of photosynthetically active radiation with a linear
153 relationship previously calibrated against destructive LAI measurements. NEE and ET were
154 measured through the Eddy covariance method by using a flux tower at a height of 27 m,
155 together with meteorological variables, including photosynthetically active radiation
156 measurements (Cabral *et al.*, 2010; Cabral *et al.*, 2011).

157 The third experimental stand (DATASET 3) was located at Itatinga (23°02'28'' S,
158 48°37'33 W, 850 m.a.s.l) and planted with *Eucalyptus grandis* seedlings on April 2004. Data
159 collection started at planting and continued until the end of the rotation, which included soil
160 water, biomass of leaves, stem, roots, leaf area index, leaf, branch and bark litterfall. Soil
161 water content was measured using TDR probes installed at various soil depth down to 10 m
162 depth (Laclau *et al.*, 2010). The soil was sandy loam, with PAWmax of 110 mm m^{-1}
163 (Maquere, 2008; Marsden *et al.*, 2013). This site was used in a previous application of
164 G'DAY model (Marsden *et al.*, 2013).

165 The fourth dataset (DATASET 4) was made of measurements of 1472 *Eucalyptus* species
166 stand polygons comprised in an area ranging from 22°33' S to 20°50' S and 48°14' W to

46°49'W (i.e., spread on an 183 × 151 km area). The 1472 polygons have an average area of 29 ha, and were planted with various genotypes of *Eucalyptus* (there is almost no coppicing practice in this dataset). Soil was highly variable in sand and clay content and therefore showed a large range of PAWmax (see section 2.2). Plant height and stem biomass were obtained from field inventories performed at two to three inventory dates (ages) for each polygon, between 2000 and 2012. The dataset and biomass calculations are presented in le Maire *et al.* (2011a), Baghdadi *et al.* (2014) and Baghdadi *et al.* (2015).

2.2. Regional scale soil and meteorological datasets

Application of the model at regional scale required reliable sources of weather and soil data. We obtained the gridded weather data from the open-access dataset for daily meteorological variables in Brazil (1980-2013) (Xavier *et al.*, 2016). This dataset provides high resolution grids (0.25° by 0.25°) of daily precipitation, evapotranspiration, maximum and minimum temperature, solar radiation, relative humidity, and wind speed developed by the CLIMA research team using ground-based weather stations in Brazil, operated by federal (INMET, ANA) and state (DAEE for São Paulo) agencies. This gridded weather dataset was compared with the three DATASET 1, 2 and 3 for solar radiation, maximum and minimum temperature, and precipitation. Gridded solar radiation (Supplementary Figs. 1A, 1B, and 1C) and maximum and minimum temperature recorded at the three sites (Supplementary Figs 1D, 1E, and 1F and Figs. 2A, 2B, and 2C) matched very well the local measurement data. For precipitations, the errors were larger, but the order of magnitude was respected (Supplementary Figs. 3 and 4). Overall, the gridded weather data was considered a reliable source for the regional simulations.

Several sources of soil data were tested to determine the dataset that best captured the variability of soil texture, and possibly the water retention properties subsequently used in the

model. The digitized 1:250.000 scale soil map of Sao Paulo state from Rossi (2017) was the most detailed one and included a soil texture attribute (Fig. 1). We overlay this map with about 600 soil profiles measured at various locations across São Paulo state obtained from the forest company, the Brazilian Agricultural Research Corporation (Embrapa) and the Geological Service of Brazil (CPRM). These soil profiles generally were to 2 m deep and included soil texture properties, soil organic matter content and bulk density. This information was used within Tomasella's pedotransfer function (level 3) to estimate the water retention parameters (Tomasella *et al.*, 2000), which enables the computation of the maximum plant available water content (PAWmax) on each soil profile, with a per meter depth unit ($mm_{water}\ m_{soil}^{-1}$). The average PAWmax value per soil textural class extracted from Sao Paulo soil map were computed, together with their standard deviation, and further extrapolated to the entire map (Supplementary Fig. 5).

2.3. G'DAY model

In the present study, the G'DAY model (Comins and McMurtrie, 1993; Corbeels *et al.*, 2005a; Corbeels *et al.*, 2005b; Marsden *et al.*, 2013) was used to simulate water and carbon budgets of *Eucalyptus* plantations at a number of experimental sites and commercial stands in the state of São Paulo, Brazil. The G'DAY model is an ecophysiological process-based model, functioning at a daily time-step, which uses minimum and maximum air temperatures, precipitation, vapor pressure deficit and solar radiation as daily weather inputs. G'DAY simulates the water and carbon fluxes between the environment and a number of soil and tree biomass pools. The ecophysiological processes in the model are represented by several sub-models of plant production, soil water balance, and litter decomposition. Marsden *et al.* (2013) modified the soil water balance in order to reflect the dynamics of moisture content in three layers: litter, top soil, and rooting zone. The maximum PAW in the litter layer was

modified to increase linearly as the C mass of the forest floor increased, whereas in the rooting zone the total maximum PAW increased with rooting depth during the rotation. Here, we used the version presented in details in Marsden *et al.* (2013), but added a new modification of the C allocation scheme to better consider the impact of environmental constraints such as water availability on the partitioning of C between shoots and roots and to enhance the model capability of capturing seasonal leaf area index (LAI) variations. Allocation fraction of the net primary production (NPP) to the different organs, is obtained following a "goal-seeking" scheme, where a constant allocation value was modulated in function of a target compartment biomass (see Corbeels *et al.* (2005a) and Marsden *et al.* (2013)), and mortality is computed as turnover rates. We slightly modified this C allocation scheme as following:

i) The proportion of carbon allocated to fine roots compared to foliage production was based on an assumption of a higher allocation to root when the soil is dry (Landsberg and Sands, 2010);

ii) The leaf turnover was set to increase when the soil was dry (severe and prolonged stress episodes) and was positively correlated to the production of new leaves, as observed in litterfall measurements (Pook *et al.*, 1997);

iii) The "target" value of leaf area, which came from a height-dependent target ratio between leaf area and sapwood cross-sectional area (Corbeels *et al.*, 2005a; Corbeels *et al.*, 2005b; Marsden *et al.*, 2013), was previously reached by modulating the carbon allocation to leaves. Within the new allocation scheme, we kept the same principle but modulated the leaf turnover ratio in order to match this algometrical constraint, while C allocation to new leaves was modulated by the equation described in i)

241 Use of water from very deep layers was shown to sustain transpiration of *Eucalyptus*
242 plantation during extreme drought events (Christina *et al.*, 2017), permitted by the fast soil
243 exploration of *Eucalyptus* trees along the rotation and the ability for the root to access soil
244 water table. A minimum access to water (i.e, a minimum value for soil PAW), was therefore
245 included in the model in order to avoid a total transpiration break and reproduce the effect of
246 deep soil water use on tree functioning observed in the field. Such minimum soil PAW was
247 however very low, and correspond to high water stress effects on trees (reduced transpiration
248 and photosynthesis).

249

250 ## 2.4. Model Calibration

251 ### *2.4.1 Calibration on the detailed DATASET 1*

252 The G'Day model has more than 200 parameters, among which physical constants,
253 parameters linked to the nitrogen cycle that were not considered in the present study, and a
254 subset of plant and soil-specific parameters. The entire set of these 57 plant and soil-specific
255 parameters, described in Supplementary Table 1, were optimized against DATASET 1 using
256 the Hooke-Jeeves algorithm for derivative-free optimization in the R package *dfoptim*
257 (Varadhan, 2018). This algorithm minimizes the residual sum of squares between observed
258 and predicted data while allowing bond constraints on the parameters. These bonds, given in
259 Supplementary Table 1, were set based on the literature and previous work on the model
260 (Marsden *et al.*, 2013). The calibration was done using all the following measurements, after
261 normalization of their residual sum of squares: plant height (m), LAI, C mass of leaves (kg C
262 m^{-2}), leaves, branches, and bark litter fall (kg C m^{-2}), stem biomass (kg C m^{-2}), PAW (mm),
263 evapotranspiration (ET, mm day^{-1}), net ecosystem exchange (NEE, g C m^{-2} day^{-1}), and gross
264 primary photosynthesis (GPP, g C m^{-2} day^{-1}).

265

266 *2.4.2 Calibration on the DATASET 2 and 3*

267

268 Only a subset of the parameters were re-calibrated (Supplementary Table 1) using DATASET
269 2 and 3. A subset of 43 parameters among the 57 was selected based on expert knowledge on
270 their possible variations among genotypes: parameters which are known to be stables among
271 genotypes, or showing very low sensitivity of stem biomass evolution through time, were
272 kept at the values obtained after calibration on DATASET 1. The 43 parameters are mainly
273 linked to carbon assimilation and allocation, allometry, wood density properties, water
274 balance submodels. Parameters obtained after calibration on DATASET 1 were used as initial
275 values of those re-calibrated parameters. The parameters re-calibrated locally were
276 constrained by new bounds values, based on the calibrated parameters obtained on
277 DATASET 1. Indeed, these parameters are not expected to vary much among genotypes or
278 with management practices, and therefore the values obtained on DATASET 1 were
279 considered reliable priors. Bounds of +/- 10 %, 20% or 30 % around this value was chosen in
280 function of the current knowledge on physiological differences among clones (Supplementary
281 Table 1). When little information was available on literature, larger bounds of 30% was
282 given. On these 2 datasets, local climate information from meteorological station and fixed
283 soil parameters were used. The recalibration was done independently for DATASET 2 and 3.
284 For DATASET 2, the calibration was done on LAI, C mass of stem, NEE, and ET, and for
285 DATASET 3 on LAI, C mass of leaves and stem, cumulative leaves, branches, bark litterfall,
286 and PAW.

287 *2.4.3 Calibration on DATASET 4*

Genotypes having fewer than 30 stands on DATASET 4, and therefore with little representativity for model calibration, were not used. Eleven genotypes were kept, which included many common genotypes used in fast growing *Eucalyptus* plantation in this region of Brazil at that time, including *E. grandis* and *E urophylla* x *grandis* materials. The dataset was then partitioned into two equal halves for each genotype, one for calibrating the parameters of the model and the other for evaluating the model (2*736 stands). Among the 57 parameters of G'DAY calibrated on DATASET 1, only 22 were calibrated for each genotype (Supplementary Table 1) using DATASET 4, with the same bounds defined for DATASET 2 and 3. Indeed, 21 parameters re-calibrated on DATASET 2 and 3 out of the 43 genotype-specific parameters were found to be hardly variable between the 3 first datasets, or to have no influence on simulated biomass. Values of the parameters obtained on DATASET 1 were used as generic values for parameters not recalibrated. For DATASET 4, the calibration was performed on two measured variables: plant height (m) and stem biomass (kg C m^{-2}) measured at two to three inventory dates for each polygon, but on the high range of soil and climate conditions found in this area. In these simulations, the planting date was prescribed as its observed value recorded in DATASET 4. The climate and soil information from the gridded inputs (see section 2.2) was used for each polygon. Model performance was evaluated by R^2, RMSE, relative RMSE, and Nash-Sutcliffe efficiency (NSE) statistics using *modeval* function in the R package *sirad* (Bojanowski, 2016). Because some stands were measured two or three times during their growth, some of the observations within the evaluation set were not statistically independent. To avoid this eventual issue, the statistics described above were computed on a subset where only one measurement per polygon was randomly selected. This random selection was repeated a thousand time, to generate a distribution of the model performance statistics, from which the average and standard deviation was computed.

313

314 2.5. Analysis of the model results at regional scale

315 Simulation of biomass at regional scale by G'DAY model was compared to a subset of
316 736 polygons set apart during the parameter calibration step. We quantified the importance of
317 spatial variation of climate, soil properties, and genotype specific parameters sets, and their
318 combinations, for the model accuracy. Eight different simulations were run on all 736
319 validation polygons (Table 2). The first set was a "base simulation", representing a scenario
320 where no spatial variation of climate, soil or genotype was taken into account: one grid point
321 of the meteorological gridded dataset, chosen at the center of DATASET 4, was selected for
322 all polygons. Similarly, one value of PAWmax was used, based on the most frequent value in
323 the total area. In base simulation, the set of parameters obtained on DATASET 1 after
324 optimization was used for all stands, to eliminate the spatial variability due to change in
325 genotypes (Table 1). The second, third, and fourth sets of simulations were performed with
326 spatial variation of climate only, genotype specific parameters only, and soil data only,
327 respectively. The fifth, sixth, and seventh sets of simulations included combinations of two of
328 these spatial variability. The last set of simulations had all drivers varying spatially, which
329 correspond to the more precise parameterization and input data in this study. The abilities of
330 the eight sets of simulations to predicting stem biomass were compared using Tukey's HSD.
331 The absolute difference between stem biomass predictions and observations for each
332 individual point (inventory date) in the 736 polygons were computed. The dependency of
333 these absolute differences (a measure of the model performance) to Climate, Genotype and
334 Soil was quantified using one way ANOVA. Climate, Genotype and Soil were considered as
335 binary dummy variables, with 0 value when no variation was accounted for, and 1 when
336 spatial variation of that characteristics was accounted for. All interactions among predictors
337 were included. This ANOVA therefore allowed to quantify the importance of the spatial

variability in climate, genotype and soil for the model performance in simulating wood productivity at regional scale.

3. Results

3.1. Model calibration on DATASET 1, 2 and 3

After calibration using DATASET 1, G'DAY model was able to simulate correctly the seasonal and inter-annual dynamics of many measured variables related to the carbon, water and energy cycling, from daily to yearly time-steps. Note that parameters for which local field measurements were available (e.g. leaf photosynthetical parameters from Christina et al., 2016) were more constrained in the inversion procedure. Calibrated values are given in Supplementary Table 1.

After calibration, the simulated LAI on DATASET 1 were fairly similar to observed LAI, except little underestimation in the third and fourth years of the rotation when observed LAI was measured at approximately 5.7 m^2_{leaf}/m^2_{soil} while simulated LAI was little less than 5 m^2_{leaf}/m^2_{soil} (Table 1 and Fig. 2B). The LAI curve followed the typical trend observed in commercial rotation of *Eucalyptus* trees, increasing to a maximum value three years after planting and then gradually decrease until harvest, with seasonal variations. Similar observations were reported for the prediction of C mass of leaves when the model slightly underestimated this C variable in the second and third years of the rotation, but matched the observed data in the following years (Table 1 and Fig. 2C).

Leaf, branch, and bark litterfalls were also correctly predicted by the model for cumulative sum, with R^2 of 0.99 and NSE close to 1 (Table 1 and Fig. 2D). Plant height was correctly predicted throughout the rotation reaching maximum height of 28 m by the end of the rotation in 2016 (Fig. 2A).The model also simulated correctly the final stem biomass that reached 8 kg C m^{-2} at the end of the rotation (Table 1 and Fig. 2E).

364 The calibrated model correctly simulated PAW throughout the rotation, except for a little
365 underestimation in early 2016 (Table 1). Figure 2F shows observed and predicted PAW and
366 the green line in the graph represents PAWmax as used in the model. This value do not
367 corresponds to the real PAWmax experienced by trees (that should be integrated down to the
368 real depth of roots) but reflects a "functional" PAWmax. Indeed, PAW is used in many
369 processes of G'DAY, and increasing PAWmax for very deep roots would results in very low
370 fraction of available water during most of the rotation, even on rainy seasons. As explained in
371 2.3, we however included a minimum PAW in the model, which represents access to deep
372 soil water. Such access is important for sustaining transpiration throughout the year, but
373 represents a small amount of transpired water as shown in Christina *et al.* (2017) and
374 confirmed in these simulations.

375 Simulation of ET correctly matched the observed data throughout the rotation, except a
376 small underestimation in the second half of 2013 and early 2014 (Fig. 3A). The goodness of
377 fit statistics of that variable indicated acceptable level of agreement between observed and
378 simulated data (Table 1). The model reasonably simulated the NEE with NSE of 0.75 and R^2
379 of 0.75 (Fig. 3B), but most importantly showed good seasonality. For GPP, simulated data
380 followed a temporal dynamic similar to observed data, which show peaks every summer
381 season along the rotation (Fig. 3C). There was an underestimation of GPP by the model in
382 2012-2013 which could be attributed to the underestimation of LAI during this period. Other
383 research is ongoing to provide more confidence on these estimates. Nevertheless, observed
384 and simulated daily GPP had R^2 of 0.92 with NSE of 0.84 and RRMSE of 16.5% (Table 1),
385 and well simulated seasonality.

386 A subset of parameters was re-calibrated using DATASET 2 and 3. The model
387 reasonably simulated LAI and stem biomass of DATASET 2 with average LAI of 3
388 m^2l_{eaf}/m^2_{soil} and C mass of stem of 3 kg C m^{-2} in the beginning of 2008 (Figs. 4A and 4B).

389 Note that LAI at this site was not measured using destructive sampling but was an estimation
390 based on reflected light, which may result in higher uncertainty. This site has a sandy soil
391 with average PAWmax of 50 mm m^{-1} of soil which could explain the lower values of LAI
392 and biomass yield at this site compared to DATASET 1. In addition, low precipitation in
393 March to November of 2006 and 2007 resulted in reduced LAI (Supplementary Fig. 3D and
394 E). The model was able to capture the seasonal variations of NEE (Fig. 4C) and ET (Fig. 4D),
395 except for underestimation of NEE at the beginning of 2007. The high ET values in summer
396 2007 were most likely due to the high precipitation (about 1000 mm) observed at that time
397 (Supplementary Fig. 3D). The model predicted ET with R^2 of 0.76 and RMSE of 1.08 (mm
398 day^{-1}) and NEE with R^2 of 0.28 and RMSE of 1.7 (g C m^{-2} day^{-1}) (Table 1). To reach this
399 result, only some of the re-calibrated parameters were significantly changed such as those
400 controlling the relationship between stem biomass and height (Ht0 and Htpower), turnover
401 rate (Bfall), stomatal conductance (Fs1 and Fs2) and photosynthesis (Jref) (Supplementary
402 Table 1).

403 Results at DATASET 3 included the model simulations of LAI, carbon mass of leaves
404 and stem, litterfall, and PAW (Fig. 5). Overall, after calibration, the model correctly
405 simulated LAI and C mass of leaves throughout the 6-year rotation (Table 1), except for a
406 little underestimation of LAI in 2006 and overestimation of C mass of leaves in 2007 (Fig.
407 5A and B). Other plant compartments of leaves, branches, and bark litterfall were also well
408 simulated with R^2 ranging from 0.99 to 0.80 and NSE close to 1 (Fig. 5C) as well as C mass
409 of stem with RMSE of 0.33 (kg C m^{-2}), RRMSE of 14.7% and NSE close to 1 (Fig. 5D).
410 Simulated and observed PAW were comparable with R^2 of 0.71 and NSE of 0.69, apart from
411 some underestimation in the first half of 2006 which could be related to the underestimation
412 of LAI in this period (Figs. 5A and G). Similarly to DATASET 2, to reach these results, only
413 some of the re-calibrated parameters were significantly changed such as parameters of growth

414 efficiency (GrowthEff), relationship between stomatal conductance and VPD (g_{max} and Fs2),
415 relationship between stem biomass and height (Ht0 and Htpower), and turnover rates (Bdecay
416 and Bfall) (Supplementary Table 1).
417 Finally, based on the results from DATASET 1, 2 and 3, we concluded that: 1) the model
418 is flexible and generic enough to represent the main processes controlling carbon and water
419 balance of these eucalypt plantations; 2) the use of large dataset including many field
420 measured variables, and the choice of adequate bounds for parameters based on field
421 measurements and literature, allow to constrain enough the model in a calibration procedure;
422 3) only a small subset of 22 parameters is necessary to be modified in the model to be able to
423 simulate very different plantations, in terms of climate conditions, soils and genotypes. The
424 procedure was therefore considered suitable, and was extrapolated to DATASET 4, which
425 have large range of spatial and genotype variation. In that case, the model was calibrated at
426 genotype-scale on biomass and stand height measurements.
427
428
429 3.1. Model calibration and validation at regional scale (DATASET 4)
430 The large scale spatio-temporal DATASET 4 included measurements of plant height and
431 stem biomass on commercial plantations. Calibrated values from DATASET 1 were used for
432 the regional simulation along with genotype specific parameters that were calibrated for each
433 genotype in the regional simulations (Supplementary Table 1). Plant height was fairly well
434 simulated with some differences in the level of precision among genotypes. For instance,
435 prediction of plant height on independent stands of genotypes C, D, F, and K had greater
436 correlation with observed data and lower RMSE values than genotypes B, H, and J where the
437 level of correlation was lower and the RMSE was greater (Supplementary Figs 6, 7, and 8).

Across all genotypes, simulated and observed height were correlated by 93% with RMSE of 1.83 m and RRMSE of 10.47% and NSE 0.86 (Fig. 6).

Similar results was obtained for biomass yield as the model performances were higher for genotypes C, F, and G with R^2 of 0.91-0.93 and RMSE of 0.63-0.73 kg C m^{-2}, while for some other genotypes like B the R^2 was *ca*. 0.80 and RMSE was 0.84 kg C m^{-2} (Supplementary Figs. 6, 7, and 8). Nonetheless, the overall correlation between observed and simulated biomass was 89% and RMSE was 0.75 kg C m^{-2} (Fig. 6, All genotypes), on stands independent from model calibration.

3.2. Impact of climate, genotype specific parameters, soil data, and their combinations on model error estimation

There was significant differences among simulation scenarios for stem biomass prediction (Figure 7 and Table 3). Prediction of stem biomass and plant height without taking into account the spatial variation of climate, soils and genotypes had the greatest error. Improvement of the model to simulate stemwood biomass was drastically improved when including differences of Genotype, then Climate and then Soil. Effect of Genotype was even further improved when climate or soil was considered together (Figure 7). The results of the ANOVA similarly showed significant individual and two-way interaction effects of climate, soil, and genotype on the accuracy of stem biomass prediction (Table 3). Fisher values confirmed the dominant effect of Genotype calibration, followed by Climate and Soil. Interactions reached a high level, in particular the Soil x Genotype interactions.

4. Discussion

4.1. Parameterization and evaluation of the model

We used a slightly modified version of the G'DAY allocation model based on the "balanced growth hypothesis" that postulates that plants invest more carbon to roots when the limiting

463 factor for growth is water or nutrients, and reduce the foliage size as drought avoiding
464 mechanism (Shipley and Meziane, 2002). In addition, the influence of water stress on
465 litterfall production and LAI dynamics was also taken into consideration, as observed on long
466 term measurements (Pook *et al.*, 1997). Evaluation of the model showed that the
467 incorporation of the modified carbon scheme enhanced the model's ability to capture the
468 decline of LAI as a result of soil drought, and consequently improved the simulation of leaf
469 biomass.

470 In the DATASET 1 calibration, the final values of some parameters were restrained
471 by the lower or upper bounds of prior distributions (Supplementary Table 1). In the present
472 work the objective was mainly to get plausible prior probability distribution of parameters to
473 be further included on larger scale applications. These parameter values, even when
474 constrained on bound limits, are still plausible and gave good results at this site with respect
475 to carbon and water measurements. On other datasets, parameters also were constrained by
476 bounds, but this was necessary because of the fewer measurements available for calibration.
477 The overall picture and the good results obtained on independent data on DATASET 4
478 showed that these limits were sufficient for a first trial. More spatial information on
479 vegetation characteristics, such as variables estimated from remote sensing (le Maire *et al.*,
480 2011b; Baghdadi *et al.*, 2014; Baghdadi *et al.*, 2015), and to more datasets could allow to
481 improve data assimilation, with a refinement of the distribution of calibrated parameters
482 probability distribution.

483 Widely distributed *Eucalyptus* species show high levels of genetic diversity in light
484 absorption, gross primary production and differences in C allocation which determine
485 together genotype productivity (Aspinwall *et al.*, 2018; le Maire *et al.*, accepted). In the
486 current study, we calibrated 22 parameters to address differences among genotype.
487 Parameters related to leaf and branch mortality and turnover rate, C allocation, and stomatal

conductance in response to climatic factors showed the greatest variations among genotypes (Supplementary Table 1). Previous research on similar Brazilian *Eucalyptus* plantations has reported variations in photosynthetic capacity between E. *grandis* × *urophylla* hybrids in comparison with E. *grandis* impacting productivity (Almeida *et al.*, 2004a). Leaf photosynthetic parameters of *Eucalyptus* plantations can also vary considerably within clones of the same species (Shem *et al.*, 2009; Warrier and Venkataramanan, 2010; le Maire *et al.*, accepted). Other reports suggested that C allocation to roots and other C sinks may be dominant drivers of genotypic variation in productivity in response to environmental factors (Resco de Dios *et al.*, 2016). This results in large differences in light absorption and light use efficiencies and growth along rotations (le Maire *et al.*, accepted).

4.2. Drivers of modelled spatial variability of stand biomass

Several factors influence biomass production in *Eucalyptus* plantations among which the climate, soil type, genetic material, and their interactions. One of the main advantages of process-based models, following proper parameterization, is the possibility to identify, quantify and disentangle the influence of spatial variation in soil, climate, and genotypes on productivity. Among these factors, genotype specific parameters were the most important for the accuracy of stem biomass predictions (Fig. 7 and Table 3). Genotype-specific calibration of ecophysiological models is critical for accurate characterization of stem biomass production among stands planted with different hybrids of *Eucalyptus*. In the present study, genotype-specific parameters were obtained by constrained optimization on only two variables, trunk biomass and height. In the future, other measurements could help improving the estimation of these parameters, such as data issued from remote sensing (le Maire, 2018).

The spatial variability of climate was the second factor in importance for prediction accuracy of stem biomass (Fig. 7 and Table 3). G'DAY was sensitive to varying climate scenarios: a reduced error in stem biomass and plant height predictions of about 20% was

513 obtained when taking into account the local climate compared with holding the climate
514 scenario constant for all stands (from a gridded dataset). While this result is expected on large
515 climate gradient ("regional scale"), we showed here that even on reduced area ("regional
516 scale") of 183 × 151 km the climate has a major importance. Rainfall and drought occurrence
517 was the most important climatic factor constraining the growth of *Eucalyptus* in this area, as
518 observed in other studies (Mummery and Battaglia, 2004; Stape *et al.*, 2004; Whitehead and
519 Beadle, 2004). In addition, research conducted in tropical Brazilian forest, near the Atlantic
520 Coast, reported 70 to 110% annual variations in production in response to soil and air
521 humidity (Almeida *et al.*, 2010). Accuracy of the input rainfall dataset could be ameliorated
522 using denser networks of pluviometers.
523 Among the three drivers of modelled spatial variability of stand biomass, varying the soil
524 type parameter was the least important driver, reducing the error in prediction by only 15%
525 (Fig. 7 and Table 3). Nevertheless, the soil x climate or soil x genotype interaction effects
526 reduced the error in stem biomass prediction in a larger way than climate alone or genotype
527 alone, indicating an improvement of the simulations when considering variations in soil water
528 holding capacity, in the studied range. A previous study aiming at simulating the spatial
529 variability of *Eucalyptus* plantation in São Paulo state using the G'DAY model reported
530 significant improvement in stem biomass prediction when using stand-specific PAW values
531 compared with holding it constant for all stands (Marsden *et al.*, 2013). The authors attributed
532 the higher performances of simulations using stand-specific PAW values to the linkage of
533 maximum PAW with tree height. The current version of the model has shown improvements
534 in simulation of stem biomass using stand-specific PAW values as well; however, it did not
535 show much improvements when comparing all drivers variable (scenario 8) with varying
536 climate and genotype (scenario 5) (Figure 7). This indicates that further improvements in the
537 model responses to the climate x genotype x soil interactions is needed. Improvement should

also be done on spatial parameterization of the maximum plant available water (PAWmax). In the current study, this gridded map has several limitations: 1) the texture class associated to each soil type of the soil map was not always precise, in particular for some complex soils; 2) associating a unique soil textural class in function of soil types do not allow to represent eventual variability within a soil type; 3) the association of a texture class to a single PAWmax value does not reflect the variability visible on Supplementary Figure 5; 4) differences in soil depth is not taken into account: here all stands are supposed to have access to the first ~3 meters of soil, and keep having minimal water access along their rotation. However, even with these limitations, the use of these gridded estimates of soil PAWmax did improved the results compared to using a single value at all location.

5. Conclusions

A modified version of the G'DAY model was able to simulate seasonal variations in growth and the exchange of key C and H_2O variables between the ecosystem and the atmosphere along complete commercial rotations of *Eucalyptus* plantations. Application of the model at the regional scale showed reasonable level of accuracy in the simulation of stem biomass and plant height. The main drivers of spatial variability in simulated stem productivity was the genetic differences among genotypes followed by climatic and soil variables. This work will benefit in the future from other data sources, such as remote-sensing, allowing to further constrain the mechanisms embodied in ecophysiological models at various temporal and spatial scales.

6. Acknowledgements

This project was funded by the FAPESP-Microsoft Research the project "Characterizing and Predicting Biomass Production in Sugarcane and *Eucalyptus* Plantations in Brazil" (SEMP) of the Sao

Paulo Research Foundation (FAPESP-Microsoft Research n° 2014/50715-9). Ahmed Attia was supported by a post-doctoral grant 2017/00886-0 from Sao Paulo Research Foundation (FAPESP). The contribution of MG was supported by grant 2012/06933-6 from Sao Paulo Research Foundation (FAPESP). This project was funded by the EUCFLUX project (cooperative program with Arcelor Mittal, Cenibra, Bahia Specialty Cellulose, Duratex, Fibria, International Paper, Klabin, Suzano, and Vallourec Florestal), coordinated by the Forestry Science and research Institute - IPEF (https://www.ipef.br/). The experiment also partially benefited of fundings from Agence Nationale de la Recherche (MACACC project ANR-13-AGRO-0005, Viabilité et Adaptation des Ecosystèmes Productifs, Territoires et Ressources face aux Changements Globaux AGROBIOSPHERE 2013 program). The experimental site belongs to the SOERE F-ORE-T, which is supported annually by Ecofor, Allenvi and the French National Research Infrastructure ANAEE-F (http://www.anaee-france.fr/fr/). We acknowledge support from the IN-SYLVA network. We are grateful to the staff at the Itatinga Experimental Station, in particular Rildo Moreira e Moreira (Esalq, USP) and Eder Araujo da Silva (http://www.floragroapoio.com.br) for their technical support.

7. References

Albaugh, J.M., Dye, P.J., King, J.S., 2013. Eucalyptus and Water Use in South Africa. International Journal of Forestry Research 2013, 11.

Almeida, A.C., Landsberg, J.J., Sands, P.J., 2004a. Parameterisation of 3-PG model for fast-growing Eucalyptus grandis plantations. For Ecol Manag 193, 179-195.

Almeida, A.C., Landsberg, J.J., Sands, P.J., Ambrogi, M.S., Fonseca, S., Barddal, S.M., Bertolucci, F.L., 2004b. Needs and opportunities for using a process-based productivity model as a practical tool in Eucalyptus plantations. For Ecol Manag 193, 167-177.

Almeida, A.C., Siggins, A., Batista, T.R., Beadle, C., Fonseca, S., Loos, R., 2010. Mapping the effect of spatial and temporal variation in climate and soils on Eucalyptus plantation production with 3-PG, a process-based growth model. For Ecol Manag 259, 1730-1740.

Aspinwall, M.J., Blackman, C.J., de Dios, V.R., Busch, F.A., Rymer, P.D., Loik, M.E., Drake, J.E., Pfautsch, S., Smith, R.A., Tjoelker, M.G., 2018. Photosynthesis and carbon allocation are both important predictors of genotype productivity responses to elevated CO 2 in Eucalyptus camaldulensis. Tree Physiol 38, 1286-1301.

Baghdadi, N., Le Maire, G., Bailly, J., Ose, K., Nouvellon, Y., Zribi, M., Lemos, C., Hakamada, R., 2015. Evaluation of ALOS/PALSAR L-Band Data for the Estimation of Eucalyptus Plantations Aboveground Biomass in Brazil. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 8, 3802-3811.

Baghdadi, N., Le Maire, G., Fayad, I., Bailly, J.S., Nouvellon, Y., Lemos, C., Hakamada, R., 2014. Testing different methods of forest height and aboveground biomass estimations from ICESat/GLAS data in eucalyptus plantations in Brazil. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 290-299.

Battaglia, M., Sands, P., White, D., Mummery, D., 2004. CABALA: a linked carbon, water and nitrogen model of forest growth for silvicultural decision support. For Ecol Manag 193, 251-282.

Bojanowski, J.S., 2016. Package 'sirad'. In: https://cran.r-project.org/web/packages/sirad/sirad.pdf (Ed.).

Cabral, O.M.R., Gash, J.H.C., Rocha, H.R., Marsden, C., Ligo, M.A.V., Freitas, H.C., Tatsch, J.D., Gomes, E., 2011. Fluxes of CO2 above a plantation of Eucalyptus in southeast Brazil. Agric For Meteorol 151, 49-59.

611 Cabral, O.M.R., Rocha, H.R., Gash, J.H.C., Ligo, M.A.V., Freitas, H.C., Tatsch, J.D., 2010. The energy
612 and water balance of a Eucalyptus plantation in southeast Brazil. Journal of Hydrology 388, 208-
613 216.
614 Campoe, O.C., Stape, J.L., Nouvellon, Y., Laclau, J.-P., Bauerle, W.L., Binkley, D., Le Maire, G., 2013.
615 Stem production, light absorption and light use efficiency between dominant and non-dominant
616 trees of Eucalyptus grandis across a productivity gradient in Brazil. For Ecol Manag 288, 14-20.
617 Christina, M., Nouvellon, Y., Laclau, J.-P., Stape, J.L., Bouillet, J.-P., Lambais, G.R., le Maire, G., 2017.
618 Importance of deep water uptake in tropical eucalypt forest. Funct Ecol 31, 509-519.
619 Christina, M., Nouvellon, Y., Laclau, J.P., Stape, J.L., Campoe, O.C., le Maire, G., 2016. Sensitivity and
620 uncertainty analysis of the carbon and water fluxes at the tree scale in Eucalyptus plantations
621 using a metamodeling approach. Can J For Res 46, 297-309.
622 Comins, H.N., McMurtrie, R.E., 1993. Long-Term Response of Nutrient-Limited Forests to CO"2
623 Enrichment; Equilibrium Behavior of Plant-Soil Models. Ecol Appl 3, 666-681.
624 Corbeels, M., McMurtrie, R.E., Pepper, D.A., O'Connell, A.M., 2005a. A process-based model of
625 nitrogen cycling in forest plantations: Part I. Structure, calibration and analysis of the
626 decomposition model. Ecol Model 187, 426-448.
627 Corbeels, M., McMurtrie, R.E., Pepper, D.A., O'Connell, A.M., 2005b. A process-based model of
628 nitrogen cycling in forest plantations: Part II. Simulating growth and nitrogen mineralisation of
629 Eucalyptus globulus plantations in south-western Australia. Ecol Model 187, 449-474.
630 Dufrêne, E., Davi, H., François, C., le Maire, G., Le Dantec, V., Granier, A., 2005. Modelling carbon and
631 water cycles in a beech forest: Part I: Model description and uncertainty analysis on modelled
632 NEE. Ecol Model 185, 407-436.
633 Gonçalves, J.L.d.M., Alvares, C.A., Higa, A.R., Silva, L.D., Alfenas, A.C., Stahl, J., Ferraz, S.F.d.B., Lima,
634 W.d.P., Brancalion, P.H.S., Hubner, A., Bouillet, J.-P.D., Laclau, J.-P., Nouvellon, Y., Epron, D., 2013.
635 Integrating genetic and silvicultural strategies to minimize abiotic and biotic constraints in
636 Brazilian eucalypt plantations. For Ecol Manag 301, 6-27.
637 Gonzalez-Benecke, C.A., Teskey, R.O., Martin, T.A., Jokela, E.J., Fox, T.R., Kane, M.B., Noormets, A.,
638 2016. Regional validation and improved parameterization of the 3-PG model for Pinus taeda
639 stands. For Ecol Manag 361, 237-256.
640 Laclau, J.-P., Levillain, J., Deleporte, P., Nzila, J.d.D., Bouillet, J.-P., Saint André, L., Versini, A.,
641 Mareschal, L., Nouvellon, Y., Thongo M'Bou, A., Ranger, J., 2010. Organic residue mass at planting
642 is an excellent predictor of tree growth in Eucalyptus plantations established on a sandy tropical
643 soil. For Ecol Manag 260, 2148-2159.
644 Landsberg, J.J., Sands, P., 2010. Physiological ecology of forest production: principles, processes and
645 models. Academic Press.
646 Landsberg, J.J., Waring, R.H., 1997. A generalised model of forest productivity using simplified
647 concepts of radiation-use efficiency, carbon balance and partitioning. For Ecol Manag 95, 209-
648 228.
649 le Maire, G., 2018. Advances in remote sensing: potential to gain insight into the ecosystem services
650 of eucalypt plantations. In, Eucalyptus 2018. CIRAD, Montpellier, France, p. 8.
651 le Maire, G., Guillemot, J., Campoe, O.C., Stape, J.-L., Laclau, J.-P., Nouvellon, Y., accepted. Light
652 absorption, light use efficiency and productivity of 16 contrasted genotypes of several Eucalyptus
653 species along a 6-year rotation in Brazil. For Ecol Manag.
654 le Maire, G., Marsden, C., Nouvellon, Y., Grinand, C., Hakamada, R., Stape, J.-L., Laclau, J.-P., 2011a.
655 MODIS NDVI time-series allow the monitoring of Eucalyptus plantation biomass. Remote Sens
656 Environ 115, 2613–2625.
657 le Maire, G., Marsden, C., Verhoef, W., Ponzoni, F.J., Lo Seen, D., Bégué, A., Stape, J.-L., Nouvellon,
658 Y., 2011b. Leaf area index estimation with MODIS reflectance time series and model inversion
659 during full rotations of Eucalyptus plantations. Remote Sens Environ 115, 586-599.

660 Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen,
661 P., 2000. Process-based models for forest ecosystem management: current state of the art and
662 challenges for practical implementation. Tree Physiol 20, 289-298.
663 Maquere, V., 2008. Dynamics of Mineral Elements under a Fast-growing Eucalyptus Plantation in
664 Brazil. Implications for Soil Sustainability. In, Ph.D. Thesis. AgroParisTech.
665 Marsden, C., Nouvellon, Y., Laclau, J.-P., Corbeels, M., McMurtrie, R.E., Stape, J.L., Epron, D., le
666 Maire, G., 2013. Modifying the G'DAY process-based model to simulate the spatial variability of
667 *Eucalyptus* plantation growth on deep tropical soils. For Ecol Manag 301, 112-128.
668 Mummery, D., Battaglia, M., 2004. Significance of rainfall distribution in predicting eucalypt
669 plantation growth, management options, and risk assessment using the process-based model
670 CABALA. For Ecol Manag 193, 283-296.
671 Ngugi, M.R., Hunt, M.A., Doley, D., Ryan, P., Dart, P., 2003. Dry matter production and allocation in
672 Eucalyptus cloeziana and Eucalyptus argophloia seedlings in response to soil water deficits. New
673 For 26, 187-200.
674 Nouvellon, Y., Stape, J.L., Laclau, J.P., Bonnefond, J.M., Da Rocha, H.R., Campoe, O.C., Marsden, C.,
675 Bouillet, J.P., Loos, R.A., Kinana, A., Le Maire, G., Saint Andre, L., Roupsard , O., 2010. Water and
676 energy fluxes above an Eucalyptus plantation in Brazil: environmental control and comparison
677 with two eucalypt plantations in Congo. In, Sir Mark Oliphant Canopy Processes in a Changing
678 Climate Conference (formally the IUFRO Canopy Processes Meeting), Falls Creek, Victoria and
679 Tarraleah, Tasmania.
680 Nouvellon, Y., Stape, J.L., Le Maire, G., Bonnefond, J.-M., Guillemot, J., Christina, M., Bouillet, J.-P.,
681 Camargo Campoe, O., Laclau, J.-P., 2018. Full-rotation carbon, water and energy fluxes in a
682 tropical eucalypt plantation. In, Eucalyptus 2018. CIRAD, Montpellier, France, pp. 102-103, Oral
683 presentation.
684 Pook, E.W., Gill, A.M., Moore, P.H.R., 1997. Long-term variation of litter fall, canopy leaf area and
685 flowering in a Eucalyptus maculata forest on the south coast of New South Wales. Aust J Bot 45,
686 737-755.
687 Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C.,
688 Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D.,
689 Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.M.,
690 Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T.,
691 Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and
692 ecosystem respiration: review and improved algorithm. Global Change Biology 11, 1424-1439.
693 Resco de Dios, V., Mereed, T.E., Ferrio, J.P., Tissue, D.T., Voltas, J., 2016. Intraspecific variation in
694 juvenile tree growth under elevated CO2 alone and with O3: a meta-analysis. Tree Physiol 36,
695 682-693.
696 Rossi, M., 2017. Mapa pedológico do Estado de São Paulo: revisado e ampliado. São Paulo: Instituto
697 Florestal, 2017. V.1. 118p. (inclui Mapas). Available online @
698 http://iflorestal.sp.gov.br/files/2017/11/Livro_Solos1.pdf.
699 Ryan, M.G., Stape, J.L., Binkley, D., Fonseca, S., Loos, R.A., Takahashi, E.N., Silva, C.R., Silva, S.R.,
700 Hakamada, R.E., Ferreira, J.M., Lima, A.M.N., Gava, J.L., Leite, F.P., Andrade, H.B., Alves, J.M.,
701 Silva, G.G.C., 2010. Factors controlling Eucalyptus productivity: How water availability and stand
702 structure alter production and carbon allocation. For Ecol Manag 259, 1695-1703.
703 Sands, P.J., Battaglia, M., Mummery, D., 2000. Application of process-based models to forest
704 management: experience with PROMOD, a simple plantation productivity model. Tree Physiol 20,
705 383-392.
706 Shem, K., Catherine, M., Ong, C., 2009. Gas exchange responses of Eucalyptus, C. africana and G.
707 robusta to varying soil moisture content in semi-arid (Thika) Kenya. Agrofor Syst 75, 239-249.
708 Shipley, B., Meziane, D., 2002. The balanced-growth hypothesis and the allometry of leaf and root
709 biomass allocation. Funct Ecol 16, 326-331.

Stape, J.L., Binkley, D., Ryan, M.G., 2004. Eucalyptus production and the supply, use and efficiency of use of water, light and nitrogen across a geographic gradient in Brazil. For Ecol Manag 193, 17-31.
Stape, J.L., Binkley, D., Ryan, M.G., Fonseca, S., Loos, R.A., Takahashi, E.N., Silva, C.R., Silva, S.R., Hakamada, R.E., Ferreira, J.M.d.A., Lima, A.M.N., Gava, J.L., Leite, F.P., Andrade, H.B., Alves, J.M., Silva, G.G.C., Azevedo, M.R., 2010. The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production. For Ecol Manag 259, 1684-1694.
Tomasella, J., Hodnett, M.G., Rossato, L., 2000. Pedotransfer functions for the estimation of soil water retention in Brazilian soils. Soil Sci Soc Am J.
Varadhan, R., Borchers, H. W, 2018. dfoptim: Derivative-Free Optimization. R package version 2018.2-1. https://CRAN.R- project.org/package=dfoptim.
Vezy, R., Christina, M., Roupsard, O., Nouvellon, Y., Duursma, R., Medlyn, B., Soma, M., Charbonnier, F., Blitz-Frayret, C., Stape, J.-L., Laclau, J.-P., de Melo Virginio Filho, E., Bonnefond, J.-M., Rapidel, B., Do, F.C., Rocheteau, A., Picart, D., Borgonovo, C., Loustau, D., le Maire, G., 2018. Measuring and modelling energy partitioning in canopies of varying complexity using MAESPA model. Agric For Meteorol 253-254, 203-217.
Warrier, K.C., Venkataramanan, K., 2010. Gas exchange characteristics in Eucalyptus clones. Indian J Plant Physiol 15, 226.
Whitehead, D., Beadle, C.L., 2004. Physiological regulation of productivity and water use in Eucalyptus: a review. For Ecol Manag 193, 113-140.
Xavier, A.C., King, C.W., Scanlon, B.R., 2016. Daily gridded meteorological variables in Brazil (1980–2013). International Journal of Climatology 36, 2644-2659.

Figure 1. Soil map of São Paulo state (Rossi, 2017) showing the location of the four datasets.

Figure 2. Comparison of measured (OBS, black line/symbol) and simulated (SIM, red line/symbol) variables after on calibration DATASET1: (A) plant height (m), (B) leaf area index (m^2 m^{-2}), (C) C mass of green leaves (kg C m^{-2}), (D) cumulative dead leave, branch, and bark litterfall (kg C m^{-2}), (E) C mass of stem (kg C m^{-2}), and (F) plant available water (mm). The green line represents simulated PAWmax.

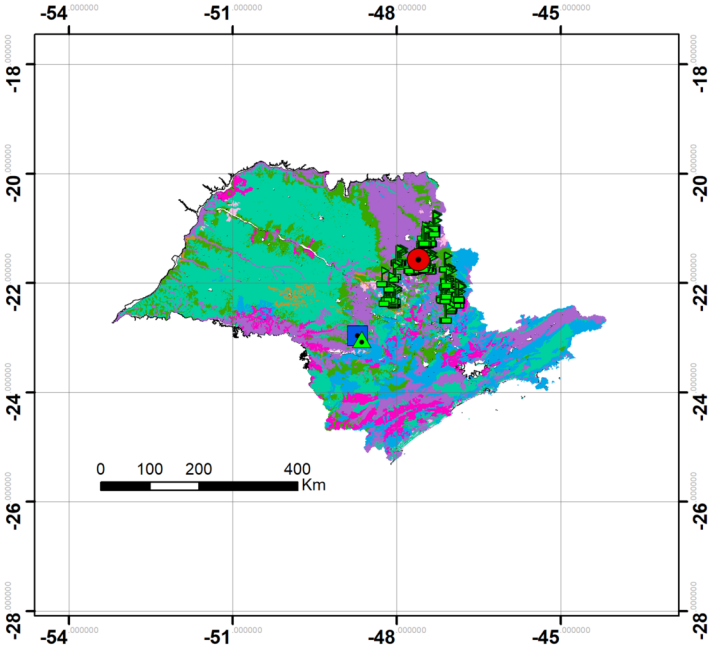
Figure 3. Comparison of measured and simulated variables on DATASET 1: (A) evapotranspiration (mm day^{-1}), (B) net ecosystem exchange (g C m^{-2} day^{-1}), and gross primary production (g C m^{-2} day^{-1}).

Figure 4. Comparison of measured and simulated variables on DATASET 2: (A) leaf area index (m^2 m^{-2}), (B) C mass of stem (kg C m^{-2}), (C) net ecosystem exchange (g C m^{-2} day^{-1}), and (D) evapotranspiration (mm day^{-1}).

Figure 5. Comparison of measured and simulated variables on DATASET 3: (A) leaf area index (m^2 m^{-2}), (B) C mass of leaves (kg C m^{-2}), (C) cumulative leave, branch, and bark litterfall (kg C m^{-2}), and (D) C mass of stem (kg C m^{-2}), and plant available water (mm). The green line represents simulated maximum PAW.

Figure 6. Measured vs. simulated plant height (m) and stem biomass (g C m^{-2}) of all *Eucalyptus* genotypes on commercial stand (DATASET 4). Model performance statistics were described in the section 2.4.3.

Figure 7. Quantifying the effects of including the spatial variation of climate, soil properties, and genotype specific parameters sets, and their combinations, on the model error estimation. Eight different simulations were run on all 736 validation polygons (see description of the scenarios in Table 2). Scenarios with the same letter are not significantly different.



Experimental sites

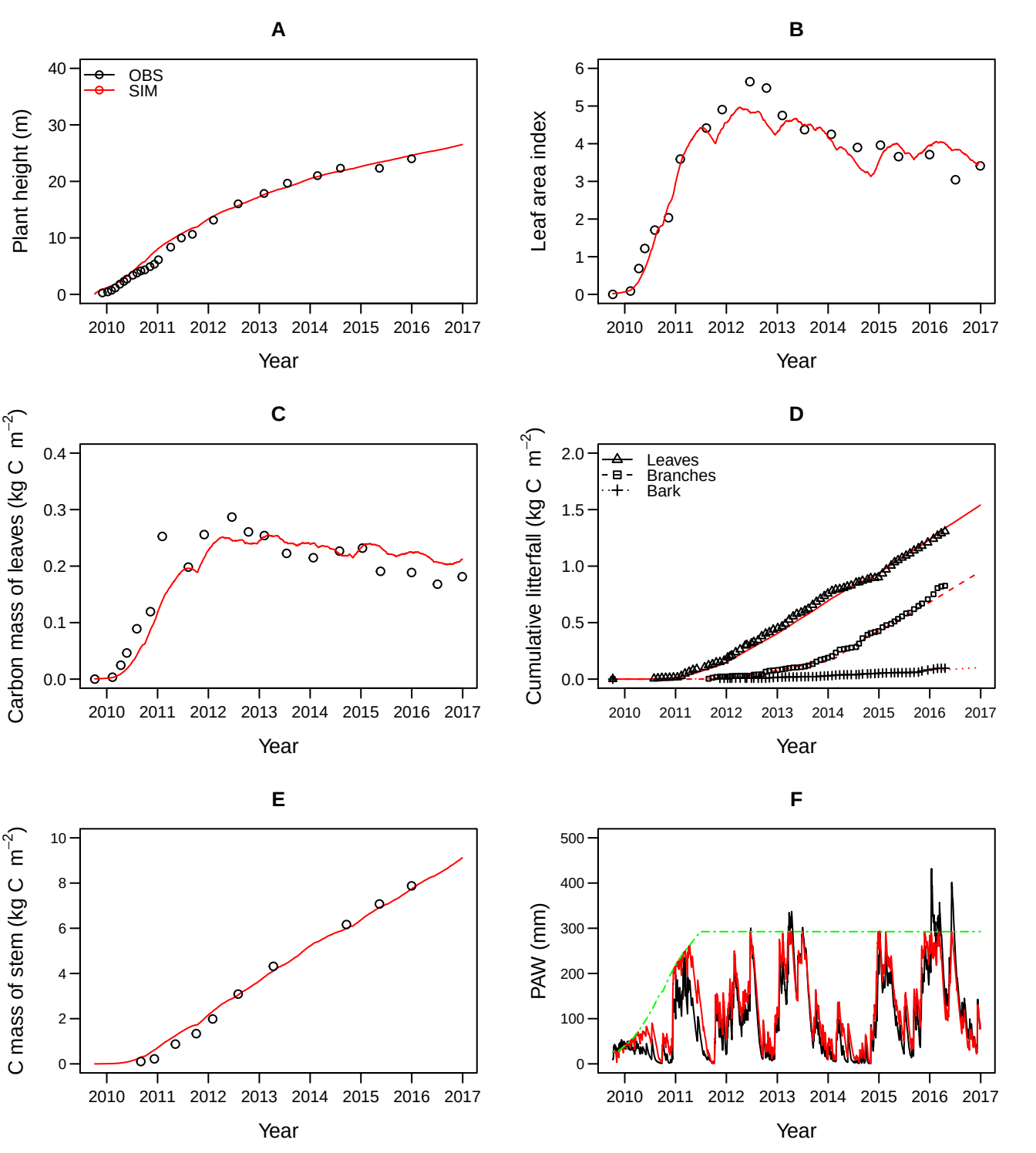
- DATASET1
- DATASET2
- DATASET3

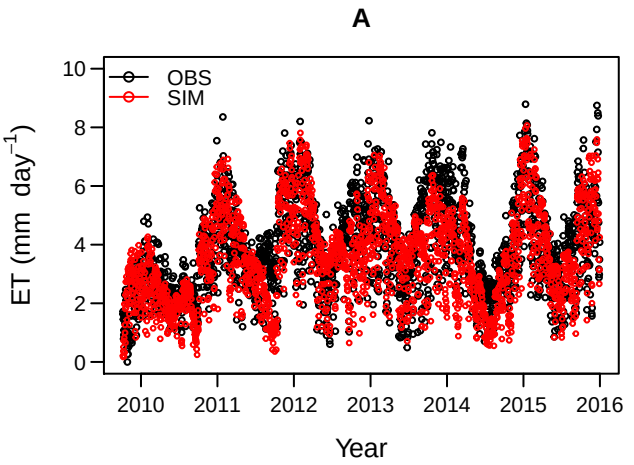
Commercial sites

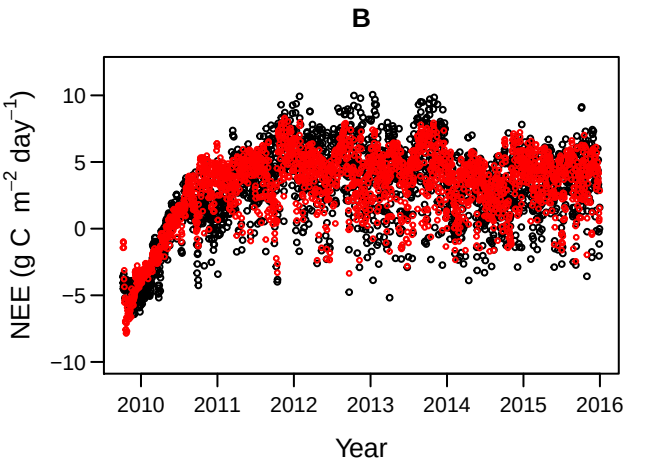
- Large scale dataset

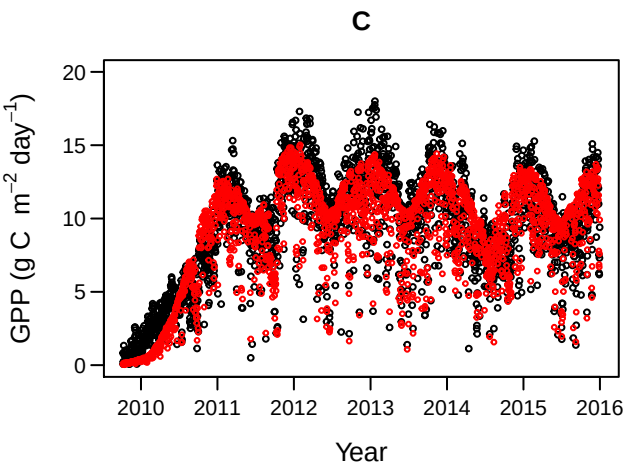
Soil texture

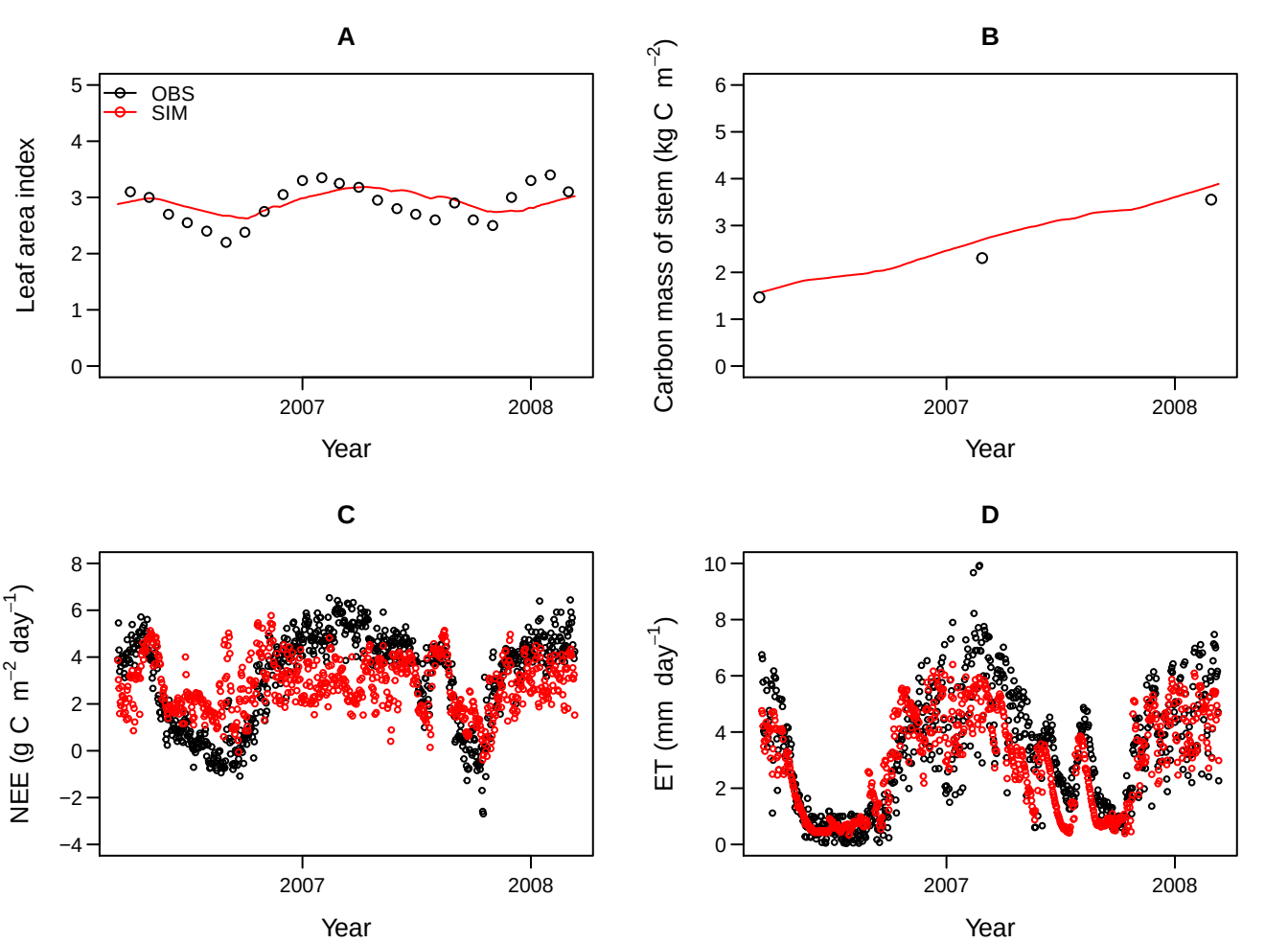
- Urban/water area
- Clay
- Clay loam
- High clay
- Loam
- Sandy
- Sandy clay loam
- Sandy loam

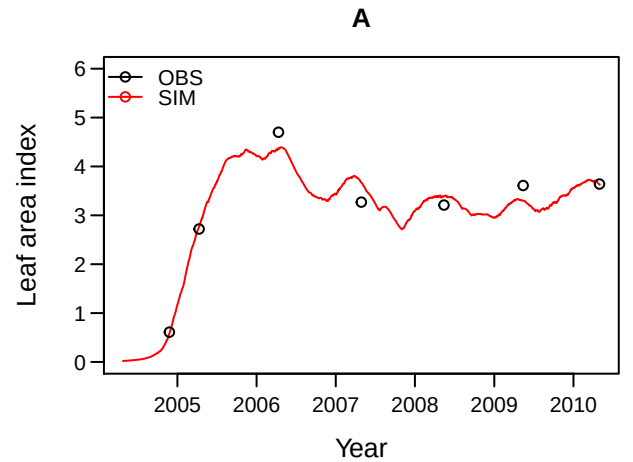

A
Plant height (m)
OBS
SIM
Year
B
Leaf area index
Year
C
Carbon mass of leaves (kg C m^{-2})
Year
D
Cumulative litterfall (kg C m^{-2})
Leaves
Branches
Bark
Year
E
C mass of stem (kg C m^{-2})
Year
F
PAW (mm)
Year

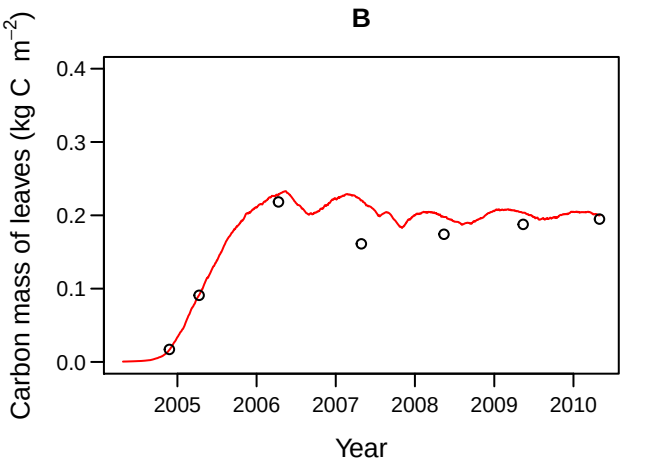

A
OBS
SIM
ET (mm day^{-1})
0
2
4
6
8
10
2010
2011
2012
2013
2014
2015
2016
Year

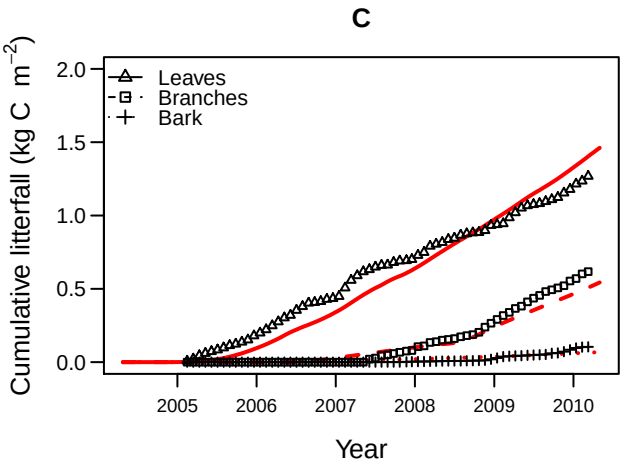

B
NEE (g C m^{-2} day^{-1})
−10
−5
0
5
10
2010
2011
2012
2013
2014
2015
2016
Year

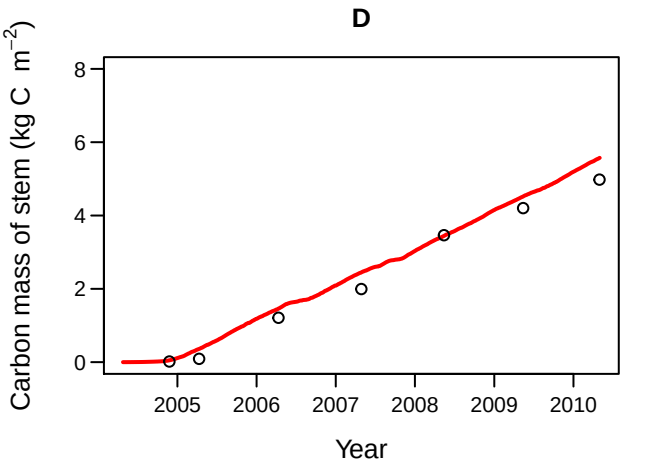

C
GPP (g C m^{-2} day^{-1})
0
5
10
15
20
2010
2011
2012
2013
2014
2015
2016
Year

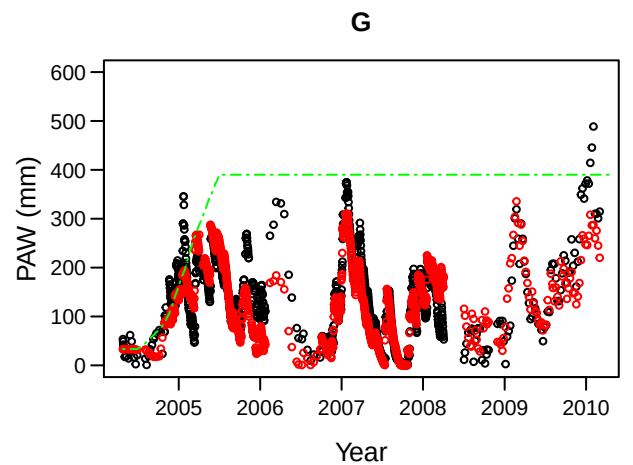

A
OBS
SIM
Leaf area index
0
1
2
3
4
5
2007
2008
Year
B
Carbon mass of stem (kg C m−2)
0
1
2
3
4
5
6
2007
2008
Year
C
NEE (g C m−2 day−1)
−4
−2
0
2
4
6
8
2007
2008
Year
D
ET (mm day−1)
0
2
4
6
8
10
2007
2008
Year

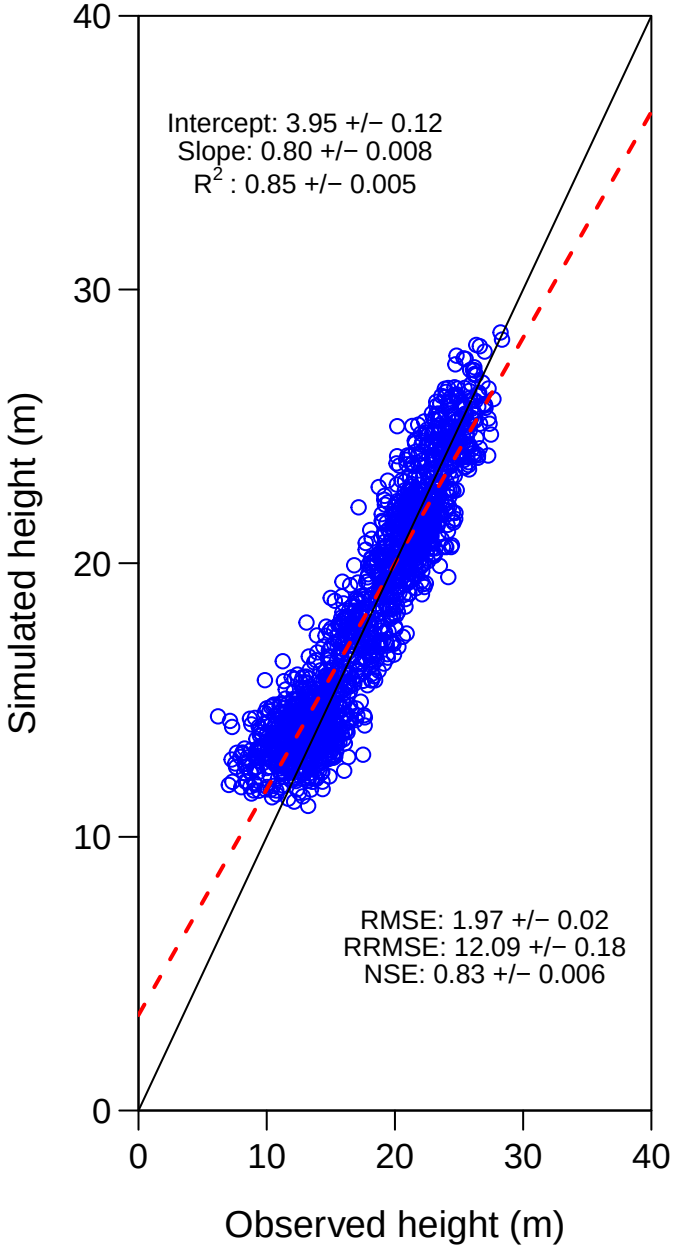

A
OBS
SIM
Leaf area index
Year

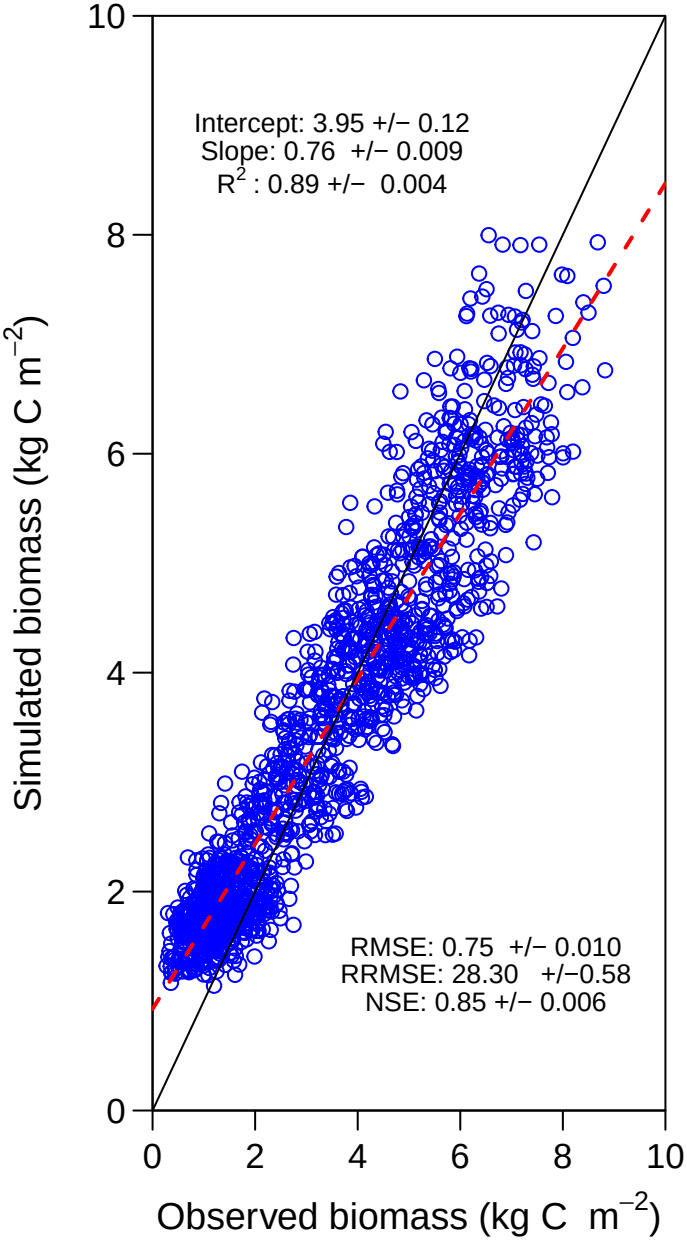

B
Carbon mass of leaves (kg C m-2)
Year


C
Leaves
Branches
Bark
Cumulative litterfall (kg C m-2)
Year


D
Carbon mass of stem (kg C m-2)
Year


G
PAW (mm)
Year


Intercept: 3.95 +/− 0.12
Slope: 0.80 +/− 0.008
R2 : 0.85 +/− 0.005
RMSE: 1.97 +/− 0.02
RRMSE: 12.09 +/− 0.18
NSE: 0.83 +/− 0.006
Simulated height (m)
Observed height (m)
0
10
20
30
40


Intercept: 3.95 +/− 0.12
Slope: 0.76 +/− 0.009
R2 : 0.89 +/− 0.004
RMSE: 0.75 +/− 0.010
RRMSE: 28.30 +/−0.58
NSE: 0.85 +/− 0.006
Simulated biomass (kg C m−2)
Observed biomass (kg C m−2)
0
2
4
6
8
10

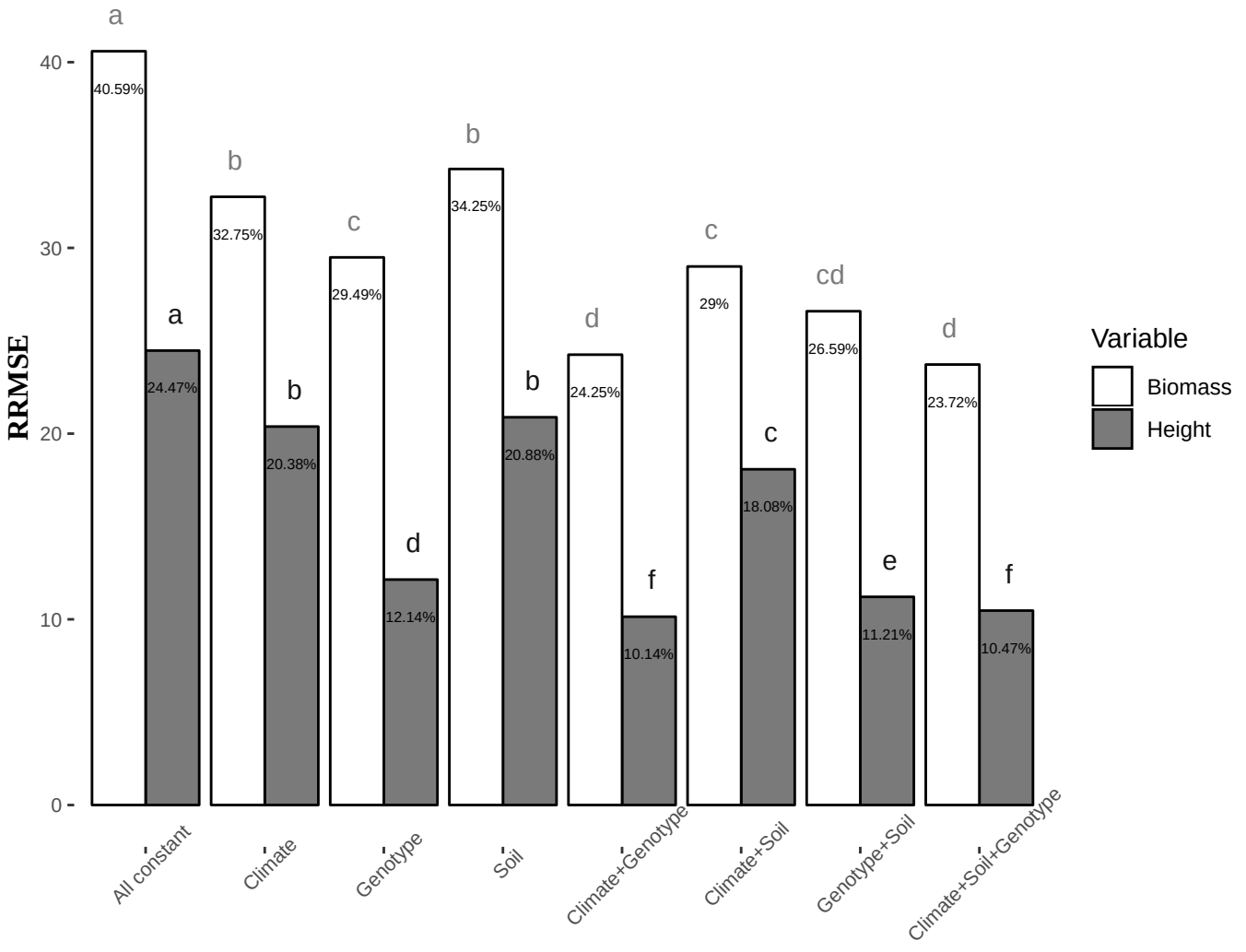

RRMSE
Scenario
Variable
Biomass
Height
0
10
20
30
40
All constant
Climate
Genotype
Soil
Climate+Genotype
Climate+Soil
Genotype+Soil
Climate+Soil+Genotype
40.59%
24.47%
32.75%
20.38%
29.49%
12.14%
34.25%
20.88%
24.25%
10.14%
29%
18.08%
26.59%
11.21%
23.72%
10.47%
a
a
b
b
c
d
b
b
d
f
c
c
cd
e
d
f

Table 1. Model performance statistics after model calibration on three experimental stands (DATASET 1, 2, and 3) : coefficient of determination (R^2), root mean square error (RMSE), relative RMSE (RRMSE), and Nash-Sutcliffe efficiency (NSE).

Variables	Corresponding Figure	R^2	RMSE	RRMSE	NSE
DATASET 1					
Plant height	Fig. 2A	0.98	1.03 (m)	11.4%	0.98
Leaf area index	Fig. 2B	0.94	0.42	13.1%	0.93
C mass of leaves	Fig. 2C	0.85	0.03 (kg C m^{-2})	21.7%	0.82
Leaves litterfall	Fig. 2D	0.99	0.03 (kg C m^{-2})	6.2%	0.99
Branches litterfall	Fig. 2D	0.99	0.02 (kg C m^{-2})	9.5%	0.98
Bark litterfall	Fig. 2D	0.97	0.004 (kg C m^{-2})	13.3%	0.96
C mass of stem	Fig. 2E	0.99	0.33 (kg C m^{-2})	10.2%	0.98
Plant available water	Fig. 2F	0.86	38.6 (mm)	37.9%	0.81
Evapotranspiration	Fig. 3A	0.71	0.91 (mm day^{-1})	24.1%	0.66
Net ecosystem exchange	Fig. 3B	0.75	1.64 (g C m^{-2} day^{-1})	57.9%	0.74
Gross primary production	Fig. 3C	0.85	1.5 (g C m^{-2} day^{-1})	16.5%	0.84
DATASET 2					
Leaf area index	Fig. 4A	0.31	0.28	9.8%	0.29
C mass of stem	Fig. 4B	0.99	0.25 (kg C m^{-2})	13.4%	0.96
Net ecosystem exchange	Fig. 4C	0.28	1.7 (g C m^{-2} day^{-1})	55.2%	0.24
Evapotranspiration	Fig. 4D	0.76	1.08 (mm day^{-1})	32.9%	0.72
DATASET 3					
Leaf area index	Fig. 5A	0.95	0.23	7.6%	0.95
C mass of leaves	Fig. 5B	0.93	0.02 (kg C m^{-2})	16.9%	0.85
Leaves litterfall	Fig. 5C	0.97	0.09 (kg C m^{-2})	15.5%	0.92
Branches litterfall	Fig. 5C	0.99	0.03 (kg C m^{-2})	29.8%	0.94
Bark litterfall	Fig. 5C	0.80	0.01 (kg C m^{-2})	87.5%	0.73
C mass of stem	Fig. 5D	0.99	0.33 (kg C m^{-2})	14.7%	0.96
Plant available water	Fig. 5E	0.71	46.6 (mm)	33.4%	0.69

Table 2: The height different simulations scenarios at regional scale, taking into account or not the spatial variation in climate, *Eucalyptus* genotype and soil. When no climate variation were input, the meteorology from the central grid point was used. When no soil variation were input, the more frequent soil type was used. When no genotype variation were input, the genotype from DATASET1 was used.

Simulation number	**Simulation name**	**Spatial variation of Climate ?**	**Spatial variation of Genotypes ?**	**Spatial variation of Soil ?**
1	All constant	No	No	No
2	Climate	Yes	No	No
3	Genotype	No	Yes	No
4	Soil	No	No	Yes
5	Climate+Genotype	Yes	Yes	No
6	Climate+Soil	Yes	No	Yes
7	Genotype+Soil	No	Yes	Yes
8	Climate+Soil+Genotype	Yes	Yes	Yes

Table 3. Analysis of variance of simulation scenarios described in section 2.5. The variance of the residuals (absolute values of measured and simulated stand height or stem biomass) is analysed in function of the use of spatially variables information on Climate, Genotype and Soil and their interactions. Analysis was performed on stem biomass and plant height. *, **, and *** refer to p values lower than 0.05, 0.01, and 0.001, respectively. ns: not significant

Source	Plant height	Stem biomass
	F value	*F value*
Climate	480.3***	106.41***
Genotype	3292.39***	260.77***
Soil	20.12***	48.87***
Climate × Genotype	0.085ns	15.75***
Climate × Soil	6.29*	11.77***
Genotype × Soil	305.67***	27.23***
Climate×Soil×Genotype	69.99***	1.29 ns