# DynACof: A process-based model to study growth, yield and ecosystem services of coffee agroforestry systems

Rémi Vezy, Guerric Le Maire, Mathias Christina, Selena Georgiou, Pablo Imbach, Hugo Hidalgo, Eric Alfaro, Céline Blitz-Frayret, Fabien Charbonnier, Peter Lehner, et al.

## Research Paper

# DynACof: a process-based model to study growth, yield and ecosystem services of coffee agroforestry systems

Rémi Vezy[a,b,c,d*], Guerric le Maire[a,b], Mathias Christina[a,b,e], Selena Georgiou[f], Pablo Imbach[f], Hugo G. Hidalgo[g,h], Eric J. Alfaro[g,h], Céline Blitz-Frayret[a,b], Fabien Charbonnier[a,b,i], Peter Lehner[j], Denis Loustau[c], Olivier Roupsard[a,f,k]

[a]CIRAD, UMR Eco&Sols, F-34398 Montpellier, France.

[b]Eco&Sols, Univ Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, Montpellier, France

[c]INRA, UMR 1391 ISPA, F-33140 Villenave d'Ornon, France

[d]CIRAD, UMR AMAP, F-34398 Montpellier, France

[e]CIRAD, UR 115, AIDA, 34398 Montpellier, France

[f]CATIE, Centro Agronómico Tropical de Investigación y Enseñanza, Turrialba 30501, Costa Rica

[g]Escuela de Física, University of Costa Rica, 2060-Ciudad Universitaria Rodrigo Facio San Pedro, San José, CR.

[h]Center for Geophysical Research, University of Costa Rica, 2060-Ciudad Universitaria Rodrigo Facio San Pedro.

[i]El Colegio de la Frontera Sur, CONACyT research fellow, San Cristóbal de las Casas, 29290 Chiapas, México

[j]Cafetalera Aquiares S.A., PO Box 362-7150, Turrialba, 7150, Costa Rica

[k]CIRAD, UMR Eco&Sols: LMI IESOL, B.P. 1386 CP 18524, Dakar, Senegal

**Corresponding author. Email address: remi.vezy@cirad.fr (Rémi Vezy).*

**Abstract**

The DynACof model was designed to model coffee agroforestry systems and study the trade-offs to e.g. optimize the system facing climate changes. The model simulates net primary productivity (NPP), growth, yield, mortality, energy and water balance of coffee agroforestry systems according to shade tree species and management. Several plot-scale ecosystem services are simulated by the model, such as production, canopy cooling effect, or potential C sequestration. DynACof uses metamodels derived from a detailed 3D process-based model (MAESPA) to account for complex spatial effects, while running fast. It also includes a coffee flower bud and fruit cohort module to better distribute fruit carbon demand over the year, a key feature to obtain a realistic competition between sinks.

The model was parameterized and evaluated using a highly comprehensive database on a coffee agroforestry experimental site in Costa Rica. The fluxes simulated by the model were close to the measurements over a 5-year period (nRMSE= 26.27 for gross primary productivity; 28.22 for actual evapo-transpiration, 53.91 for sensible heat flux and 15.26 for net radiation), and DynACof satisfactorily simulated the yield, NPP, mortality and carbon stock for each coffee organ type over a 35-year rotation.

Keywords: Crop model; *Coffea arabica*; MAESPA; GPP; *Erythrina poeppigiana*; plant-to-plot scale

## 1. Software and data availability

The DynACof model was developed as an R package (R Core Team, 2019), and as a Julia package (Bezanson et al., 2017). Full documentation is available on its dedicated website (https://vezy.github.io/DynACof), the code is open-source (GNU GPLv3 license) and available on Github repositories (R: https://github.com/VEZY/DynACof, Julia: https://github.com/VEZY/DynACof.jl) and archived on Zenodo (https://doi.org/10.5281/zenodo.1256816). The input data used for this study is available as example data from this repository, and the data for model evaluation is available from the FLUXNET website (http://www.europe-fluxdata.eu/home/site-details?id=CR-AqC).

## 2. Introduction

The key role of crop models is to help understand and predict the links between crop development and climate, soil, management, facilitation and competition between species. Crop models can provide insights into the main emerging agricultural challenges such as food security, sustainability, how to enhance ecosystem services, and how to cope with the possible negative effects of climate changes (Spiertz, 2012). There is an increasing need to address these issues at global scale to identify the different solutions available (Makowski et al., 2014), especially when the products are exchanged on the global market, like wheat, maize, soybean, cocoa or coffee. Perennial plantations are difficult to study, because their relatively long growing cycle extends the period necessary for data acquisition, and because the heterogeneity of the canopy sometimes significantly increases the intra-plot light and micro-meteorological anisotropy, such as for temperature, vapor pressure or aerodynamic conditions (Luedeling et al., 2014; Luedeling et al., 2016). Agroforestry systems (AFS) are probably the most complex perennial agroecosystems (Malézieux et al., 2009), because they have the most heterogeneous vertical and/or horizontal canopies, and these affect all ecosystem fluxes (Charbonnier et al., 2013; Vezy et al., 2018). Yet, AFS have the potential to enhance ecosystem services (Jose, 2009; Lin, 2010; Taugourdeau et al., 2014) such as carbon sequestration (Jose and Bardhan, 2012; Oelbermann et al., 2004), and to mitigate climate pressure on crops (Lin, 2007).

In Costa Rica, *Coffee arabica* is mostly grown under AFS management because it is assumed to improve coffee bean quality and to expand the cropping area to sub-optimal low altitude and warmer areas (Muschler, 2001), but such assumptions depend mostly on altitude, local climate and postharvest processing (Worku et al., 2018). Modeling the energy, water and carbon balance of these agroecosystems could provide insights into their functioning and allow stakeholders to test the *in-silico* trends of new management practices (e.g. density, pruning, thinning date or intensity) or species arrangement on given outputs such as yield or other ecosystem services. However, several factors make these systems challenging to model. First, there are many options for shade management with highly heterogeneous canopies, ranging from free growing, low density shade trees like *Cordia alliodora* to high density, heavily managed low trees or plants such as banana or pollarded *Erythrina poeppigiana* trees (van Oijen et al., 2010a). Second, the coffee reproductive phenology is a complex process that lasts for about two years (Camargo and Camargo, 2001), with competition between reproductive and

vegetative compartments (Charbonnier et al., 2017) and bienniality at the plant scale (Schnabel et al., 2018). Blossoming is mostly synchronized in sub-tropical regions but can also be highly asynchronous in equatorial regions, which impacts the distribution of fruit carbon demand and in turn, carbon allocation to other organs (Rodríguez et al., 2011). Third, coffee plants are often pruned every five to six years to sustain high levels of production on young resprouts. It is also assumed that the reserve compartment plays a major role in bean production, with biennial sprout dynamics (Cannell, 1985). Fourth, there are very few comprehensive datasets that can be used to calibrate and test multi-objective models for ecosystem services, energy, carbon, and water balance, aboveground and belowground biomass, NPP, fruit yield and more.

Model development implies identifying, prioritizing and balancing the most important processes and the scale at which the model should simulate them. In coffee systems, we assume that (i) absorbed light, (ii) light use efficiency (LUE), (iii) within canopy temperature, (iv) water, and (v) nutrient uptake are among the most important primary processes because they regulate carbon assimilation, respiration, evapotranspiration, vegetative growth, flowering, and fruit development. The next most important processes may be (vi) shade tree and coffee leaf phenology that regulate light absorption, canopy temperature and transpiration, (vii) carbon partitioning to compute net primary productivity (NPP), mortality (litterfall) and organ biomass, and (viii) a detailed phenology of the reproductive organs comprising all stages, from the appearance of cohorts of buds, flowers, and fruits, until harvest or the overripe stage.

Several models have already been developed to simulate coffee, grown in full sun or agroforestry systems:

- Rodríguez et al. (2011) proposed a model to simulate coffee in monoculture only, from branch to whole-plant scales. The model was calibrated from planting to five years old. The strength of this model lies in the fine phenology and physiological processes of the modeled coffee plant using branch-level cohorts of flowers and fruits over the entire two-year reproductive cycle. Indeed, cohorts are required to realistically distribute the demand for carbon of the fruits over the course of the season, and not all at once. This model was successfully calibrated for Colombian and Brazilian sites, two regions with contrasting climate and flower phenology (subtropical and equatorial). However, this model was not designed for large plots, long rotations or agroforestry: coffee light absorption is computed using the Beer-Lambert law using a constant coefficient of extinction, absorbed light is converted into photosynthesis using constant light use efficiency, and coffee pruning, shade trees, canopy temperature, water and energy balance are not implemented in the model.
- Another model was developed by Van Oijen et al. (2010b). This is a 1D plant average plot-scale model for coffee grown in agroforestry systems, simplifying the intra-plot microclimate into either below shade or in full sun. One clear advantage of this model is its ability to compute several ecosystem services and to incorporate various types of shade tree management and species, and the thorough Bayesian parameterization approach that was used. The model is simple, fast and can be run under changing climates. It was recently applied in East Africa under climate change scenarios by Rahn et al. (2018). The main limitations of the model are (i) its light transmission module does not consider light distribution as a continuum under shade trees, as described in Charbonnier et al. (2013), (ii) its

formalism of LUE which is not influenced by the intra-plot light variability even though Charbonnier et al. (2017) found it to be greatly impacted, (iii) it lacks of a reserve compartment and of a cohort module and again (iv) the absence of energy balance and temperature of the canopy to drive the plant respiration reproductive development.

- Two other models have also been applied to coffee in an agroforestry system using 3D light interception modules: in Dauzat et al. (2001), where only a sample of a few coffee plants were simulated, and using the MAESPA model to simulate the whole system. Since MAESPA was recently demonstrated to accurately predict light distribution, canopy temperature and water and energy balance in such systems (Charbonnier et al., 2013; Charbonnier et al., 2017; Vezy et al., 2018), the model can readily compute all variables that are potentially influenced by the complex canopy structure. However, its relatively high computation time still limits its application for full rotations of coffee under AFS, and the model does not simulate growth and yield.

We argue that a proper combination of the inherent strengths of the above-described models could provide significant improvements and extend application domains. It would involve combining cohorts and reserves at the plant scale, variable canopy temperature and intra-plot microclimate and LUE, while allowing a reasonable level of abstraction to insure rapid simulations (multiple plots, crop rotations, management, etc.). In this study, we built surrogate models (i.e. metamodels) of MAESPA for the spatial-dependent variables, and integrated them into a simpler growth and yield model to avoid expensive computation and development time. These metamodels are simple instantaneous equations that efficiently compute a given output of a complex model. In other words, it is a reduction of a complex model intended to emulate the behavior of complex interactions between variables (e.g. spatial heterogeneity) into one empirical equation, becoming an input for the next crop model. Metamodels are generally used to better understand the processes at stake in a model and to assess model sensitivity and uncertainty (Christina et al., 2016; Faivre et al., 2013), for the purpose of optimization (Razavi et al., 2012), or to make faster and reasonably accurate predictions for a given variable that is usually computed by a time-consuming model, but with fewer simulation errors compared to simpler models (Marie et al., 2014). Metamodels are often used as an efficient and simple tool to combine models at different time and/or space scales without running the finer-scale model iteratively.

Consequently, we designed DynACof to incorporate a plant-scale reproductive phenology formalism inspired by Rodríguez et al. (2011) but dependent on canopy temperature, with different sub-modules to adapt coffee and shade tree management, density and tree species, as in Van Oijen et al. (2010b), and metamodels calibrated from MAESPA simulations for spatially-dependent variables, such as diffuse and direct light extinction coefficients, light use efficiency, leaf water potential, transpiration, and sensible fluxes (Vezy et al., 2018).

Regarding model parameterization, calibration, and evaluation for coffee modelling, several strategies are proposed in the literature, most of which depend on the availability of field data. Rodríguez et al. (2011) assumed that the main factor that influences yield variability is latitude, because of its impact on phenology, so they tested their model on an equatorial site and on a sub-tropical site, with two distinct set of parameters for each situation. The former model was evaluated against field data on biomass (leaves, branches, berries, stem and

roots) and total plant nitrogen content gathered at three distinct sites. Van Oijen et al. (2010b), and later Rahn et al. (2018), who further developed the CAF2007 model, proposed a calibration based on an extensive screening of the literature and a sensitivity analysis, but no model evaluation against field data at this stage. Alternatively, we propose a multiple-objective strategy of evaluation in this study, relying on a large range of state and flux variables measured at the same time by the end of the crop rotation, including eddy-covariance fluxes, coffee and shade tree biomass measured at organ scale, necromass, yield, NPP, water balance and energy balance, and finally farm registers to describe management during a complete rotation (Charbonnier et al., 2013; Charbonnier et al., 2017; Defrenet et al., 2016; Gómez-Delgado et al., 2011; Taugourdeau et al., 2014; Vezy et al., 2018). Consequently, the aims of the present study are to:

(i) develop metamodels for spatially-dependent variables based on MAESPA, which has already been calibrated and validated on coffee agroforestry systems;

(ii) develop a new plot-scale 1D, 5-layer (shade tree, coffee, and 3 soil layers), average plant, ecophysiological process-based model for coffee crops grown under agroforestry or in full sun, while combining the advantages and strengths of three previously published models. Nutrients are considered non-limiting in this first version, which is realistic for many field conditions.

(iii) parameterize and evaluate the model using a multi-objective approach applied to an extensive dataset from a long-term observatory including the energy and water balance, GPP, yield, NPP and carbon mass per organ.

## 3. Materials and methods

### 3.1. Site description

The research site is located on the Aquiares Coffee Farm (6.6 $km^2$) between 9° 56’ 8” and 9° 56’ 35” N and 83° 44’ 39” and 83° 43’ 35” W, itself located in the central-Caribbean area of Costa Rica. The climate is tropical humid with no marked dry season, 3 014 mm mean annual precipitation and 19.5 °C mean annual temperature during the 1979 to 2009 period. The elevation of the research site ranges between 1 020 and 1 280 m a.s.l. and the mean slope is 11.31°. The vegetation consists of coffee plants (*Coffea arabica* L., var *Caturra*) planted below *Erythrina poeppigiana*, a leguminous shade tree. Both were planted in 1979. Shade trees were originally planted at a density of approximately 250 trees $ha^{-1}$ and pollarded twice a year to optimize the light transmitted to the coffee layer. They were thinned to about 7.4 trees $ha^{-1}$ in 2000 and left growing freely until the end of the study. The shade trees do not follow any particular planting pattern. The Aquiares farm is intensively managed with several applications of fertilizer per year (214 ± 44 kg N $ha^{-1}$ $yr^{-1}$) and a regular selective pruning of the coffee shoots, a practice often used by farmers to avoid a drop in production due to exhaustion. The farm complies with the Rainforest Alliance™ for its pest and weed management (weeds are scarce). Between 1995 and 2016, the Aquiares farm reported average yields of green coffee around 1333 ± 336 kg green coffee $ha^{-1}$ $yr^{-1}$.

### 3.2. Comprehensive database of measurements

"Coffee-Flux" is a collaborative research observatory (https://www.umr-ecosols.fr/recherche/projets/53-coffee-flux) monitored continuously since 2009 and located on the Aquiares coffee farm. This research site has been intensively studied and described in detail, notably for hydrology and eddy-covariance by Gómez-Delgado et al. (2011); LAI by Taugourdeau et al. (2014); light budget by Charbonnier et al. (2013); belowground biomass and NPP by Defrenet et al. (2016), ecosystem biomass, NPP and LUE by Charbonnier et al. (2017); and energy balance, water balance and surface temperature by Vezy et al. (2018).

The measurements used in this study to evaluate the model included: half-hourly $NEE$ (Net Ecosystem Exchange), $H$ (sensible heat flux), $LE$ (latent heat flux), and $Rn$ (net radiation) measured using an eddy-covariance tower. The online ReddyProc tool (Wutzler et al., 2018) was used to assess gross primary productivity (GPP), following the Lasslop et al. (2010) model option, based on daily hyperbolic curves to estimate photosynthesis. This model option was chosen because the Coffee-Flux site is mountainous, prone to night-time advection. The flux measurements were integrated at a daily time scale for comparison with DynACof outputs. The LAI of the shade tree ($LAI_{Tree}$) was also measured using a LAI2000 (LI-COR, NE, USA), and the coffee LAI ($LAI_{coffee}$) using a normalized difference vegetation index (NDVI) sensor positioned 25 m above the ground, at an angle of 15° to the vertical with 45° view angle, and converted to LAI according to Charbonnier et al. (2013). The light transmittance by shade trees was also measured using the LAI2000. The carbon mass ($g_C$ $m^{-2}$) of the shade tree stem and branches, and the coffee resprout wood, fine roots, stump and coarse roots were all measured during two consecutive years (2011/2012 and 2012/2013) using dimension measurements and site specific allometric relationships. The NPP of each compartment of the shade tree and coffee plant was also computed from the biomass increment and from litter production. Further details about the methods used for measurement are available in Charbonnier et al. (2013) and Charbonnier et al. (2017).

### 3.3. MAESPA model and metamodel conception

#### 3.3.1. Description of MAESPA

MAESPA is a 3D explicit process-based model (Duursma and Medlyn, 2012; Medlyn, 2004; Wang and Jarvis, 1990) used to simulate forest energy, water, and carbon fluxes at the scale of the individual tree. The light interception module, canopy temperature, water and energy balance of the model have already been calibrated, used and validated on the same agroforestry system (Charbonnier et al., 2013; Vezy et al., 2018). MAESPA is particularly well suited to simulate agroforestry system fluxes because it describes the forest at voxel scale, which is a homogeneous representation of a sub-part of the tree crown. It can manage several tree species with their own position on the plot, their overall structure (crown height, width, etc.), and their physical and physiological parameters. Thus, MAESPA computes a fine estimation of the light interception, energy, water, and carbon fluxes of each plant in the forest and of the soil, while taking the spatial heterogeneity of the canopy into account. However, MAESPA lacks a carbon allocation module or growth process, and requires computationally-intensive simulations, e.g. a week of computation on the shared Montpellier Bioinformatic

Biodiversity (MBB) computing cluster platform to complete a distributed simulation of a 0.2 ha AFS plot, including 4176 *Coffea arabica* sprouts and 14 *Erythrina poeppigiana* shade trees over one year at 30-minute time-scale. Consequently, MAESPA was used here mainly to compute metamodels to simulate spatial-dependent variables.

### 3.3.2. Metamodels

The main process affected by canopy complexity is probably the photosynthetically active radiation absorbed ($APAR$) by the canopy (Charbonnier et al., 2013). Heterogeneous canopies like those of shade trees in AFS tend to violate the assumption of a constant value for diffuse ($KDif$) and direct ($KDir$) light extinction coefficients because of non-uniform spatial distribution of leaf area, and because the leaf area density and foliage aggregation can change over time (Sampson and Smith, 1993; Sinoquet et al., 2007). Furthermore, a comparison between coffee plants grown in monoculture and under an agroforestry system showed that canopy complexity also affected canopy temperature, water and energy partitioning, light interception, transpiration and stomatal conductance (Vezy et al., 2018).

Metamodels were computed from MAESPA simulation outputs for diffuse ($KDif$) and direct ($KDir$) shade tree light extinction coefficients, coffee and shade tree light use efficiency ($LUE, g_C\ MJ^{-1}$), transpiration ($Tr, mm\ d^{-1}$), plant sensible heat flux ($H, MJ\ m^{-2}d^{-1}$) and soil net radiation ($Rn_{Soil}, MJ\ m^{-2}d^{-1}$). Metamodels are used to reproduce as far as possible the link between a set of input variables and the desired output variable, as if it had been computed by the MAESPA model.

MAESPA simulations were also used to find the values of some parameters such as the coffee layer light extinction coefficients that had low variability throughout the simulation and hence were assumed constant in DynACof. The partitioning parameter ($Soil\_LE\_P$, see appendix A.2) used to compute soil sensible and latent heat flux from the soil net radiation in DynACof was also parameterized using MAESPA outputs.

Metamodels were fitted using multilinear regressions selected according to a trade-off between the number of explanatory variables, their genericity and range of application and their accuracy obtained using different statistics (EF: modelling efficiency, $R^2$: r squared, RMSE: root mean squared error). MAESPA was used to simulate one year of the coffee agroforestry system from the Coffee-Flux site in the Aquiares farm in order to calibrate the metamodels. The parameterization and description of these MAESPA simulations are reported in Vezy et al. (2018). The outputs considered are converted from half hourly values at shade tree and coffee resprout scale to daily plot scale values for each plant layer during the same year (Figure 1). The metamodels were trained on 80% of MAESPA simulation outputs aggregated at daily time-scale and evaluated over the 20% remaining validation data to compute out-of-sample statistics. Both training and testing partitions were sampled based on the dependent variable subgroup percentiles using the "createDataPartition" function in the caret R package (Kuhn, 2019).

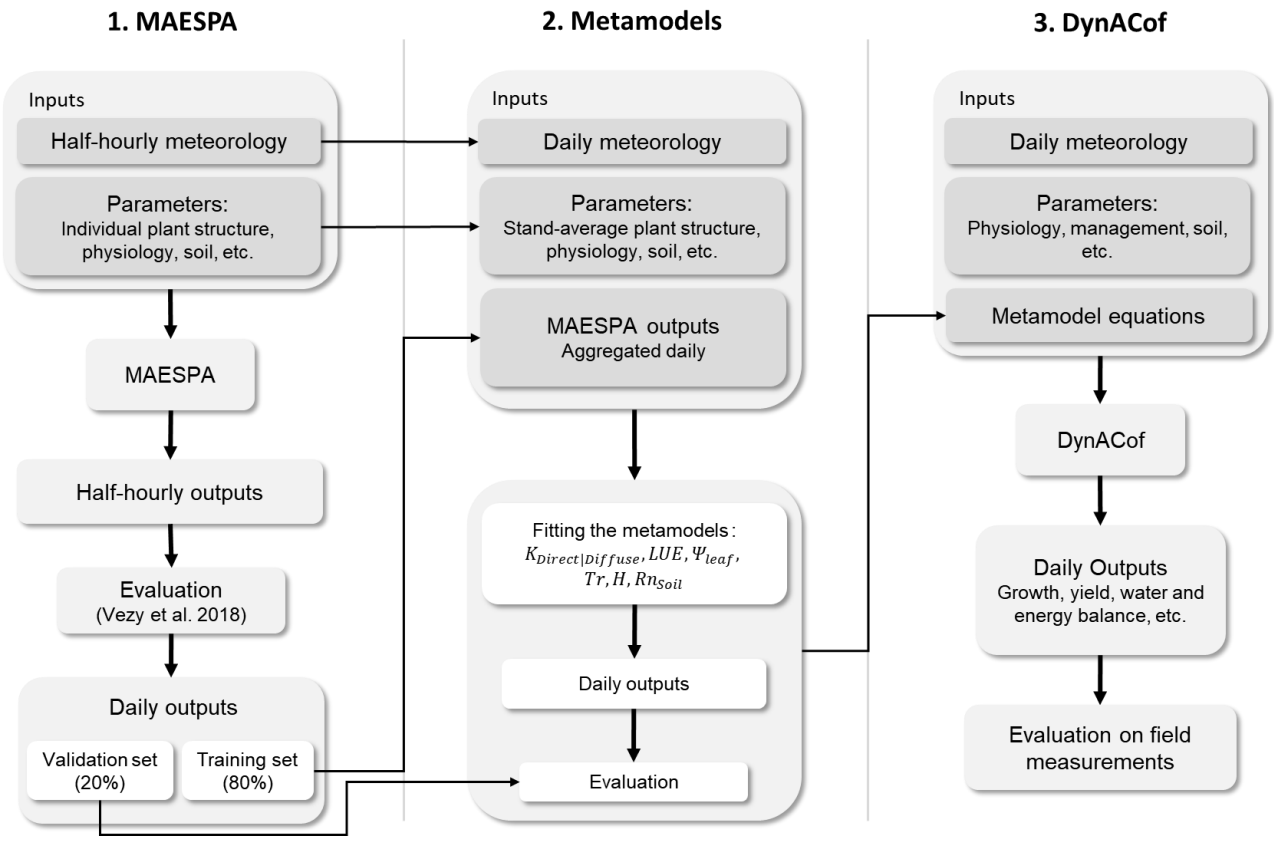


**Figure 1. Flowchart of the three-step metamodeling approach. 1/ The resolution for stand simulations in MAESPA was the individual tree and the half-hourly time step. 2/ The metamodels were fitted using inputs (meteorology and parameters) at daily and stand scale used as predictors to model the MAESPA outputs integrated at daily and stand-scale (training set, 80% of the data). The metamodels were evaluated on the validation set (20% of the data). 3/ The metamodel equations were injected into DynACof.**

The input variables of MAESPA used as predictors for metamodels were either related to climate or to plot-average plant structure. Climate variables included daily air temperature, vapor pressure deficit, PAR, diffuse and direct light fractions, wind speed, and air pressure. The atmospheric $CO_2$ concentration can also be used for the $LUE$ metamodel but were not useful for the present study. Input variables concerning shade tree structure for the metamodels were leaf area ($m^2_{leaves}\ tree^{-1}$), crown height ($m$), trunk height ($m$), crown radius ($m$), trunk diameter ($m$) and all derivatives such as the leaf area index (LAI, $m^2_{leaves}\ m^{-2}_{soil}$), leaf area density (LAD, $m^2_{leaves}\ m^{-3}_{crown}$), tree density ($trees\ ha^{-1}$) or crown projection ($m^2_{crown}\ m^2_{soil}$), all being plot averages. The LAI was measured continuously in the plot, while the other structural data were measured only once during the 2011 campaign. The variability and the interactions between the predictors were checked to insure the range of values was similar in the MAESPA simulation dataset and in the application dataset used for DynACof.

The metamodels are not mandatory in DynACof, and can be replaced by any equation or fixed value. Hence, DynACof was also run using standard equations instead of the metamodels to assess their relative contribution to the modeling performance for the plot-scale net radiation (Rn), latent (LE) and sensible (H) heat flux, and gross primary productivity (GPP). For this purpose, the FAO recommended plot scale equation from Allen et al. (1998) was chosen as a reference for comparison with the current computation for Rn in DynACof, LE was computed using the Penman-Monteith equation (Allen et al., 1998), H was computed as the difference between

Rn and LE, and GPP was computed using a constant light use efficiency ($LUE$) and constant light extinction coefficients ($KDif$ and $KDir$) for both the coffee and the shade tree layers. The average values computed from MAESPA simulations were used to compute the constant $LUE$ and $Ks$ coefficients on this modeling experiment.

### 3.4. Description of the DynACof model

#### 3.4.1. Introductory description

DynACof (which stands for Dyn*amic* A*groforestry* Cof*fee crop model*), is a daily plot scale crop model (Murthy, 2004) with two layers of vegetation (shade trees and coffee plants) and three soil layers, aimed at simulating the growth and yield of coffee plantations under various shade tree species and management options, considering the spatial heterogeneity of the shade tree canopy. The coffee layer can be simulated either in monoculture or in agroforestry systems. Each layer is simulated sequentially at a daily time step. Variables with high intra-plot variability, i.e. light absorption, LUE, transpiration, plant sensible heat flux and soil net radiation are all computed using metamodels from MAESPA. The model accounts for potential competition for light acquisition and water availability between plant and soil layers. Nutrients are considered non-limiting in this first version, which is realistic for many field conditions in Costa Rica. Water competition is simulated virtually from the day-to-day fluctuations in water content in each shared soil layer that can be reduced by drainage and evapotranspiration, or increased by precipitation through throughfall. This simple formalism should be sufficient given the absence of water limitations in the application concerned (Vezy et al., 2018), but it can also reproduce the competition between plants under more constrained conditions.

#### 3.4.2.Light interception and photosynthesis

The diffuse ($APARDif_i$, $MJ\ m^{-2}\ d^{-1}$) and direct ($APARDir_i$, $MJ\ m^{-2}\ d^{-1}$) daily absorbed photosynthetic active radiation of each plant layer are computed using the Beer-Lambert's law of light extinction:

$$APARDif_i = PARDif_i \cdot (1 - e^{-KDif \cdot LAI_i}) \quad (1)$$

$$APARDir_i = PARDir_i \cdot (1 - e^{-KDir \cdot LAI_i}) \quad (2)$$

with $PARDif_i$ and $PARDir_i$ the diffuse or direct daily photosynthetically active radiation ($MJ\ m^{-2}\ d^{-1}$) reaching the layer on day $i$, $KDif$ and $KDir$ the light extinction coefficient of the layer for the diffuse and direct light respectively and $LAI$ the leaf area index ($m^2_{leaf}\ m^{-2}_{soil}$) of the layer. Both stand scale light extinction coefficients of the shade tree layer are computed using metamodels from MAESPA, while the extinction coefficients of coffee were approximated to be constant after MAESPA simulations (only a slight variability could not be explained using stand scale factors). $PARDif$ and $PARDir$ are computed as the incoming $PAR$ minus the $PAR$ absorbed by the upper layer if any, neglecting the $PAR$ reflected back by the canopy.

The gross primary productivity ($GPP, g_C\ m^{-2}\ d^{-1}$) of each layer is then computed from the sum of diffuse and direct $APAR$ multiplied by the light use efficiency ($LUE$, $g_C\ MJ^{-1}$), which is derived from a MAESPA metamodel:

$$GPP_i = (APARDif_i + APARDir_i) \cdot LUE_i \quad (3)$$

### 3.4.3. Carbon supply

A whole plant daily gross carbohydrate budget (GCB) is computed from the daily $GPP$, available reserves, and plant maintenance respiration as follows:

$$GCB_i = GPP_i + kres \cdot CM_{RE,i-1} - Rm_i \quad (4)$$

The carbon available from reserves for day $i$ is computed as a fraction ($kres$) of the carbon mass of the reserves from the previous day ($CM_{RE,i-1}, g_C\ m^{-2}$). $Rm$ is the maintenance respiration ($g_C\ m^{-2} d^{-1}$, see Eq.(6)).

The plant daily carbon supply is then computed as:

$$Supply_i = \begin{cases} GCB_i & if\ GCB_i > 0 \\ 0 & if\ GCB_i \leq 0 \end{cases} \quad (5)$$

If $GCB_i$ is negative, it means that the carbon available from $GPP$ and reserves was not sufficient to support the maintenance respiration, so $Supply_i$ is equal to 0, and the missing carbon is considered as mortality imputed equally to all organs (see Eq. (23)).

### 3.4.4. Maintenance Respiration

The maintenance respiration requirement ($Rm$) is computed as the sum of maintenance respiration of all plant organs (Litton et al., 2007). The maintenance respiration of each organ ($Rm_j$) is computed using its carbon mass ($CM, g_C\ m^{-2}$) from the previous day, following a $Q_{10}$ formalism, as reported in Dufrêne et al. (2005):

$$Rm_{j,i} = pa_j \cdot CM_{j,i-1} \cdot NC_j \cdot MRN \cdot Q_{10_j}^{\frac{T_j - T_{MR}}{10}} \quad (6)$$

where $j$ is the organ, $i$ the day, $pa_j$ (0-1) the living fraction of the organ, $CM_{j,i-1}$ ($g_C\ m^{-2}$) the carbon mass of the organ from the previous day, $NC_j$ ($g_N\ g_C^{-1}$) the organ nitrogen content, $MRN$ ($g_C\ g_N^{-1}\ d^{-1}$) the respiration rate per nitrogen unit, $Q_{10_j}$ (dimensionless) the response of organ respiration to temperature, $T_j$ ($°C$) the temperature of the organ and $T_{MR}$ ($°C$) the base temperature for maintenance respiration. The leaf temperature of the plant layer ($Tc$) that is computed in Eq. (44) is used as a proxy for $T_j$, except for the roots where the soil temperature ($Tsoil$) is used instead.

### 3.4.5. Carbon allocation to organs

The allocation of carbohydrates to the organs is limited by either the $Supply_i$ or by the organ demand for carbon. The $Supply_i$ is distributed to the different organs following a hierarchical allocation scheme (Lacointe, 2000). For coffee, Charbonnier et al. (2017) reported that allocation to the woody compartments remained quite steady whatever the fruit load, so DynACof supplies this compartment with priority. In contrast, carbon allocation to fruits appeared to prevail over allocation to leaves (Charbonnier et al., 2017; Gutierrez et al., 1998), so DynACof supplies carbon assimilates to fruits just after woody compartments, and the remaining carbon is allocated to leaves and fine roots, and eventually reserves. The demand from the woody compartments is

considered to always be proportional to the $Supply_i$, the demand from fine roots and leaves is considered constant, whereas the fruits demand is computed (see section 3.4.11). If the demand for assimilates by the fruits is sufficiently high, the carbon allocated to fruits potentially represents all the remaining global carbon $Supply_i$ after allocation to the woody compartments, leaving nothing for the leaves and the fine roots.

In practice, the carbon $Supply_i$ is first partitioned between shoot wood ($CA_{shoot,i}$) and stump and coarse roots wood ($CA_{SCR,i}$) using constant coefficients $\lambda_{shoot}$ and $\lambda_{SCR}$:

$$CA_{shoot,i} = \lambda_{shoot} \cdot Supply_i \quad (7)$$

$$CA_{SCR,i} = \lambda_{SCR} \cdot Supply_i \quad (8)$$

$$\lambda_{shoot} + \lambda_{SCR} < 1$$

Carbon allocation to the coffee fruits is computed as the minimum between the total fruit demand (see Eq. (38)) and the remaining carbon supply.

$$CA_{fruit,i} = \min(DE_{fruit,i}\ ;\ Supply_i - CA_{shoot,i} - CA_{SCR,i}) \quad (9)$$

The remaining carbon supply, if any, is then shared between leaves and fine roots following a coefficient of allocation ($\lambda r_j$) for the remaining carbon:

$$Supply_{leaf,i} = \lambda r_{leaf} \cdot (Supply_i - CA_{shoot,i} - CA_{SCR,i} - CA_{fruit,i}) \quad (10)$$

$$Supply_{FRoot,i} = \lambda r_{FRoot} \cdot (Supply_i - CA_{shoot,i} - CA_{SCR,i} - CA_{fruit,i}) \quad (11)$$

$$\text{with } \lambda r_{leaf} + \lambda r_{FRoot} = 1 \quad (12)$$

And the allocation to both organs is the minimum between their respective supply and demand:

$$CA_{leaf,i} = \min(DE_{leaf,i}, Supply_{leaf,i}) \quad (13)$$

$$CA_{FRoot,i} = \min(DE_{FRoot}, Supply_{FRoot,i}) \quad (14)$$

Demand by the fine roots is a constant parameter, but demand by the leaves depends on the plant $LAI$ compared to the maximum observed $LAI$:

$$DE_{leaf,i} = DELM \cdot \frac{Stocking_i}{10{,}000} \quad (15)$$

where $DELM$ is the maximum leaf carbon demand ($g_C\ plant^{-1}\ d^{-1}$) per coffee plant and $Stocking_i$ is the coffee planting density on day $i$ ($plant\ ha^{-1}$).

Finally, if both leaf and fine root carbon demands are satisfied, the remaining carbon is stored in the reserves for future use:

$$CA_{RE,i} = Offer_i - CA_{shoot,i} - CA_{SCR,i} - CA_{fruit,i} - CA_{leaf,i} - CA_{FRoot,i} \quad (16)$$

### 3.4.6. Growth respiration

Growth respiration is the energy cost in carbohydrates to make new biomass from $CA$ on day $i$. It is computed using a construction cost coefficient ($\varepsilon_j, g_{C_{produced}}\ g_{C_{allocated}}^{-1}$) applied to the carbon allocated to the organ $j$:

$$Rg_{j,i} = CA_{j,i} \cdot \varepsilon_j \quad (17)$$

### 3.4.7. Net primary productivity

Maintenance respiration is already accounted for before any carbon allocation (see 3.4.3), therefore, the net primary productivity of each organ $j$ for day $i$ ($NPP_{j,i}, g_C\ m^{-2}\ d^{-1}$) is computed using the difference between the carbon allocated to the organ and the growth respiration:

$$NPP_{j,i} = CA_{j,i} - Rg_{j,i} \quad (18)$$

$$j = \{shoot|SCR|fruit|leaf|FRoot\} \quad (19)$$

The net primary productivity of the plant is the sum on $j$ of the $NPP_{j,i}$ of the layer considered, and the total net primary productivity of the stand is the sum of all $NPP_{j,i}$ in all the layers.

### 3.4.8. Mortality

Mortality is the sum of the natural mortality (turnover rate, $Mnat$), of pruning ($Mprun$), diseases and of the lack of carbon ($MC$) when the carbohydrate budget ($GCB_i$) is negative.

The natural mortality of an organ $j$ ($Mnat_{j,i}, g_C\ m^{-2} d^{-1}$) is proportional to its carbon mass from the previous day ($CM_{j,i-1}, g_C m^{-2}$), and is computed using a lifespan parameter ($lifespan_j, d$):

$$Mnat_{j,i} = \frac{CM_{j,i-1}}{lifespan_j} \quad (20)$$

The annual pruning of coffee plants affects the shoot wood and the leaves, and it is assumed that leaf loss is accompanied by fine root loss. Pruning also affects the branches of the shade trees. For the leaves and shoots, pruning is adjusted using a pruning intensity coefficient ($PI$) as follows. It is considered that two separate $PI$s are needed for the leaves and the shoots because the distribution of wood and leaf biomass may be heterogeneous between resprouts of different ages in the coffee plant as reported in Charbonnier et al. (2017).

$$Mprun_{j,i} = PI_j \cdot CM_{j,i-1} \quad (21)$$

Mortality due to pruning of the fine roots is related to the mortality of pruned leaves using a constant parameter ($m_{FRoot}$):

$$Mprun_{FRoot,i} = Mprun_{leaf,i} \cdot m_{FRoot} \quad (22)$$

Leaf disease mortality is only implemented for coffee, using a module to compute the American Leaf Spot (ALS) according to Avelino et al. (2007).

Mortality due to lack of carbohydrates to meet the maintenance respiration requirements ($MC_i$) is computed as:

$$MC_{j,i} = \begin{cases} -\dfrac{GCB_i \cdot CM_{j,i-1}}{CM_{tot,i-1}} & if\ GCB_i < 0 \\ 0 & if\ GCB_i \geq 0 \end{cases} \quad (23)$$

where $CM_{j,i-1}$ is the carbon mass of the organ $j$ from the day before $i$ (see Eq. (24)) and $CM_{tot,i-1}$ is the total carbon mass of the leaves, shoots, fine roots stump and coarse roots of the previous day.

Finally, the total mortality of each organ $M_{j,i}$ is computed as the sum of all mortalities: natural, pruning, leaf disease (for coffee only) and lack of carbohydrates.

### 3.4.9. Carbon and dry mass of organs

The carbon mass of an organ is incremented daily by adding its $NPP_{j,i}$ and removing its mortality:

$$CM_{j,i} = CM_{j,i-1} + NPP_{j,i} - M_{j,i} \quad (24)$$

The organ dry mass is computed using the carbon mass and the carbon content of each organ ($CC_j, g_C\ g_{DM}^{-1}$).

### 3.4.10. Branch nodes

The plagiotropic branches of a coffee plant present several nodes of decreasing age from the orthotropic branch to the tip:

- several unproductive old nodes, close to the orthotropic branch;
- productive nodes from the previous year that potentially bear flower buds, and then fruits;
- newly formed nodes that bear leaves.

The coffee reproductive phenology is based on a 2-year cycle (Camargo and Camargo, 2001). For a given year N, the new flush of nodes bearing new leaves will bear flower buds by the end of year N, while the leaves from year N-1 (born by the fruiting nodes N-1) are shed. Given that each new node bears two leaves, one expects some proportionality between node number and leaf area for the year N (Gutierrez et al., 1998). Such proportionality is used here to upscale the equations of Rodríguez et al. (2011) from the branch to the whole plant. The model computes the total number of newly formed nodes ($GN_i$) on the coffee plant for year N, i.e. the number of green wood nodes that will potentially bear flower buds by the end of the first year of the reproductive cycle:

$$GN_i = LAI_i \cdot RNL \cdot CN \quad (25)$$

where $RNL$ is the number of nodes per $LAI$ unit at 20 °C, assuming that this parameter is dependent on the growth temperature. $CN$ is an empirical temperature-dependant correction coefficient that is a function of the mean temperature during the vegetative growth period ($T_{gp}, °C$). This relationship is derived from Drinnan and Menzel (1995):

$$CN = 0.4194773 + 0.2631364 \cdot T_{gp} - 0.0226364 \cdot T_{gp}^2 + 0.0005455 \cdot T_{gp}^3 \quad (26)$$

Note that this coefficient is from one of the only datasets in the literature that links coffee reproductive and vegetative growth to the ambient temperature.

### 3.4.11. Fruit development

The reproduction module is mostly inspired by Rodríguez et al. (2011), but upscaled to the whole plant (e.g. in Eq. (25)). Two main development processes are computed in the model: the flower bud cohorts of year N, and the fruit cohorts of the following year. The bud itself has two stages of development, while the fruit has five. Buds can be initiated only during the bud initialization period (BIP) from day $D_{BIP1}$ to $D_{BIP2}$. The buds appear on branch nodes in daily cohorts every day within that time window. The $D_{BIP1}$ date is computed from the cumulative sum of degree days ($S_{ddTffb}$) after the date of end of vegetative development ($D_{VG2}$).

The degree days are computed as follows:

$$dd_i = \begin{cases} Tcan_i - T_{min} & if\ Tcan_i > Tmin \\ 0 & if\ Tcan_i \leq Tmin \end{cases} \quad (27)$$

where $dd_i$ are the degree days (°C) of day $i$, $Tcan_i$ the mean daily coffee canopy temperature, $Tmin$ the minimum (i.e. base temperature) and $Tmax$ the maximum temperature for physiological activity. One originality here is that the degree days are computed using the canopy temperature of each layer rather than measured air temperature because the microclimate close to the plant may differ depending on the local conditions (e.g. the cooling effect of shade trees).

$D_{BIP1}$is computed as:

$$\begin{cases} S_{ddTffb} = \sum_{i=D_{VG2}}^{D} dd_i & if\ S_{ddTffb} < F_{Tffb} \\ D_{BIP1} = D & if\ S_{ddTffb} = F_{Tffb} \end{cases} \quad (28)$$

where $F_{Tffb}$ is the threshold value in degree days that triggers the start of the bud development window. It is considered that bud cohorts stop initializing when the first bud cohort initiated on the plant enters the fruit stage. In other words, buds can be initiated every day as cohorts until the first fruit appears on the coffee plant. This day is denoted $D_{BIP2}$.

The number of buds initiated daily on a given cohort ($BudS1$) from $D_{BIP1}$ to $D_{BIP2}$ depends on several factors: the incoming radiation ($RAD, MJ\ m^{-2}\ d^{-1}$), the canopy degree day computed following Eq. (27) and the number of green nodes to support the buds ($GN_i$). This computation follows Eq. (12) from Rodríguez et al. (2011), here adapted to the whole plant:

$$BudS1_i = (a_{bud} - b_{bud} \cdot RAD_i) \cdot GN_i \cdot dd_i \quad (29)$$

where $a_{bud}$ and $b_{bud}$ are parameters.

Each bud initiated on day $i$ is considered to belong to the bud cohort of day $i$. There are as many cohorts as days between $D_{BIP1}$ and $D_{BIP2}$.

Once initiated, bud cohorts develop for $F_{bud_{S1}}$degree days until they become dormant on day $D_{bud_{S1}}$:

$$\begin{cases} S_{ddbS1} = \sum_{i=init}^{D} dd_i\ \ if\ \ S_{ddbS1} < F_{bud_{S1}} \\ D_{bud_{S1}} = D & if\ S_{ddbS1} = F_{bud_{S1}} \end{cases} \quad (30)$$

where $init$ is the day of initialization of the cohort.

To potentially break dormancy, the bud cohort has to experience a minimum amount of cumulated rainfall ($F_{rain}$) after $D_{bud_{S1}}$:

$$\begin{cases} S_{rain} = \sum_{i=D_{bud_{S1}}}^{D} rain_i & if\ S_{rain} < F_{rain} \\ D_{rain} = D & if\ S_{rain} = F_{rain} \end{cases} \quad (31)$$

However, the bud cohort can cumulate a maximum of $F_{bud_{S2}}$ degree days during its lifetime. If it does not break its dormancy before the day when this threshold is reached ($D_{bud_{S2}}$), it is considered desiccated. $D_{bud_{S2}}$ is computed as:

$$\begin{cases} S_{ddbS2} = \sum_{i=D_{bud_{S1}}}^{D} dd_i & if\ S_{ddbS2} < F_{bud_{S2}} \\ D_{bud_{S2}} = D & if\ S_{ddbS2} = F_{bud_{S2}} \end{cases} \quad (32)$$

The time window when buds from a cohort can potentially break their dormancy ($W_{fruit_{S1}}$) is then found by solving Eq. (31) and Eq. (32):

$$W_{fruit_{S1}} = [D_{rain}; D_{bud_{S2}}] \quad if\ S_{rain} = F_{rain}\ and\ D_{rain} < D_{bud_{S2}} \quad (33)$$

The time window is shared by all the buds in the same cohort, but all the buds will not necessarily break their dormancy on the same day. The number of buds in a given cohort that break their dormancy on a given day within $W_{fruit_{S1}}$ depends on the combination of two factors: the mean diurnal air temperature within the coffee canopy during bud growth ($CB, °C$) and the leaf water potential ($\Psi_{leaf}, MPa$) of the coffee plant. These conditions reflect the need for a drier period followed by an intense rainfall event for optimal dormancy break (Rodríguez et al., 2011). The number of buds breaking dormancy ($Budbreak$) on a given day is computed as:

$$Budbreak_i = BudS2_i \cdot Pbreak_i \cdot CB \quad (34)$$

where $BudS2_i$ is the number of stage 2 buds in the cohort that have not yet broken dormancy, $CB$ is a temperature-dependant correction coefficient, and $Pbreak_i$ is the rate at which they break dormancy, which is related to the leaf water potential ($\Psi_{leaf}, MPa$) using two parameters ($a_p$ and $b_p$) as follows:

$$Pbreak_i = \frac{1}{1 + e^{a_p + b_p \cdot \Psi_{leaf,i}}} \quad (35)$$

$CB$ is implemented to include the effect of the average canopy temperature on the number of buds initiated during growth of the coffee plant. It is computed from a monotone Hermite spline (Fritsch and Carlson, 1980) fitted on the rather unique Drinnan and Menzel (1995) dataset relating temperature and blossoming. More details are available on the help page of the eponym function of the R package of DynACof (https://vezy.github.io/DynACof/reference/CB.html).

When buds do break dormancy, they enter the fruit stage by forming a flower (stage 1 fruit). Then, the fruits develop and finally become mature (stage 4) then overripe (stage 5), when they fall onto the ground. Each bud breaking dormancy on day $i$ forms a new cohort of fruits, meaning that fruits forming a fruit cohort can originate from several different bud cohorts. The stage 1 fruits of the cohort then enter the carbon allocation scheme to undergo the successive stages of maturation. Under optimal conditions, the carbon mass of the fruit increases following a logistic growth over the growing period until it becomes overripe. This logistic growth during physiological development can be modeled as:

$$L = \left(1 + e^{\frac{F_{dd_{inf}} - S_{ddfruit}}{s}}\right)^{-1} \quad (36)$$

where $F_{dd_{inf}}$ is the inflexion value of the logistic growth (in degree days), $s$ is the steepness of the logistic growth, and $S_{ddfruit}$ is the number of degree days accumulated by the fruit since its flowering date.

The carbon demand of a cohort $c$ on day $i$ is computed as the sum of the carbon mass incremented by all fruits in the cohort under optimal conditions. It is computed using the derivative of the Eq. (36) as follows:

$$DE_c = \begin{cases} FruitS1 \cdot DE_{opt} \cdot L' & if\ S_{ddfruit} < F_{over} \\ 0 & if\ S_{ddfruit} \geq F_{over} \end{cases} \quad (37)$$

where $FruitS1$ is the number of fruits in the cohort, $DE_{opt}$ is the total carbon demand of a coffee fruit thoughout its development under optimal conditions including growth respiration, and $F_{over}$ (degree days) the physiological age for fruits to become overripe and no longer require carbohydrates because they fall to the ground. Thereby, the total plant scale fruit carbon demand is computed as the sum of the demand of all the cohorts growing on the coffee plant:

$$DE_{fruit,i} = \sum_{c=1}^{n} DE_{c,i} \quad (38)$$

The fruit demand for carbon can be considered as a genetic growth potential, with optimal fruit growth when there are no limitations to the supply. It is a sink strength that depends on the number of fruits and the degree days and is independent of the carbon supply. Consequently, the allocation of carbon to fruits is constrained either by fruit demand or by the carbon supply as in Eq. (9). The fruit mass is then computed as in Eq. (24). The fruits that become overripe (i.e. are not harvested) are removed from the coffee plant and considered as a mortality.

Coffee bean quality is also computed using the fruit sucrose content of each fruit cohort ($c$) based on the number of days after flowering, following the model of Pezzopane et al. (2012):

$$SM_i = \sum_{c=1}^{n} \frac{CM_{c,i}}{CC_{fruit}} \cdot \left( S_{y0} + \frac{S_a \cdot 100}{1 + \left(\frac{i}{S_{x0}}\right)^{S_b}} \right) \quad (39)$$

$$SMOpt_i = \sum_{c=1}^{n} \frac{CM_{c,i}}{CC_{fruit}} \cdot \left( S_{y0} + S_a \cdot 100 \right) \quad (40)$$

$$Mat_i = \frac{SM_i}{SMOpt_i} \quad (41)$$

where $SM_i$ is the sucrose mass of all the cohorts $c$ on day $i$, and $SMOpt_i$ is the optimal sucrose mass of the coffee bean, i.e. the sucrose mass of the bean if it was fully mature. $CM_{c,i}$ is the carbon mass of the $c^{th}$ cohort on day $i$, $CC_{fruit}$ is the carbon content of the fruit, $S_{y0}$ is the sucrose content at the beginning of the fruit development, $S_{x0}$ is the day at which the maturation is at the inflexion point, and $S_a$ and $S_b$ are two maturation parameters. Harvest maturity is then simply the global fruit maturity on the day of harvest.

Harvesting takes place once each growing season, it is triggered on the date when total cumulated fruit mortality has been higher than fruit growth for more than ten days. This simple method ensures that harvesting occurs when the fruit stock is at its maximum to optimize yield.

#### 3.4.12. Shade tree allometry

In addition to the common allocation scheme, each shade tree species has its own set of allometric relationships. Any kind of allometry can be implemented and can then be used as input for the metamodels, or as an informative model output. In this study, allometric relationships were used to compute the tree crown radius

and height from the tree branch dry mass and the tree density. The crown radius and height were then used to compute the LAD (leaf area density, $m^2_{leaves}\, m^{-3}_{crown}$) which is an input for the metamodels for the light extinction coefficients. Shade tree height is mandatory to compute the aerodynamic conductance and was computed in this study using an allometric relationships with the tree stem dry mass. The tree diameter at breast height was also computed from stem dry mass.

### 3.4.13. Temperature

The temperature of the foliage and the air inside the canopy is computed using the formalism proposed by Van de Griend and Van Boxel (1989) using the diffusivities for momentum transport and potential energy flow inside and in-between the inertial sublayer, the roughness sublayer, the air space of each canopy layers, the foliage and the soil. The temperature of the air inside the canopy is computed as:

$$TairCan_{i,l} = Tair_i + \frac{H_{i,l}}{\rho_i \cdot Cp \cdot gbulk_i} \tag{42}$$

with $i$ the day, $l$ the canopy layer, $Tair$ the air temperature measured above the canopy, $H$ the sensible heat flux ($MJ\ m^{-2}d^{-1}$), $\rho$ the air density ($kg\ m^{-3}$), $Cp$ the specific heat of air for constant pressure ($MJ\ K^{-1}kg^{-1}$) and $gbulk$ the aerodynamic conductance for heat above the canopy ($m\ s^{-1}$). If the coffee plants are grown in monoculture, Eq. (42) is applied directly to the coffee layer. If they are grown under shade trees, Eq. (42) is used for the shade tree, and the air temperature inside the coffee canopy is computed as follows:

$$TairCan_{i,cof} = TairCan_{i,tree} + \frac{H_{i,cof}}{\rho_i \cdot Cp \cdot ginterlay_i} \tag{43}$$

with $TairCan_{i,tree}$ the temperature of the air inside the shade tree canopy computed using Eq. (42), and $g_{interlay}$ the aerodynamic conductance at the interface between both canopy layers ($m\ s^{-1}$).

The leaf temperature ($Tcan$) of each layer is then computed as:

$$Tcan_{i,l} = TairCan_{i,l} + \frac{H_{i,l}}{\rho_i \cdot Cp \cdot gbh_i} \tag{44}$$

with $TairCan$ the air temperature inside the given canopy layer, and $gbh$ the leaf boundary layer conductance for heat ($m\ s^{-1}$).

The temperature at the soil surface is similarly computed as:

$$Tsoil_i = TairCan_{i,cof} + \frac{H_{i,soil}}{\rho_i \cdot Cp \cdot gsoilcan} \tag{45}$$

with $gsoilcan$ the canopy to soil aerodynamic conductance ($m\ s^{-1}$).

### 3.4.14. Soil, water, and energy

The soil water balance module was inspired by the BILJOU model (Granier et al., 1999), which has already been parameterized for this coffee agroforestry system (Gómez-Delgado et al., 2011) and is given in Appendix A. It has three layers, 0 to 1.25 m, 1.25 to 1.75 m and 1.75 to 3.75 m, respectively, to cover the whole root profile of the coffee plants (Defrenet et al., 2016).

The net radiation of the soil was computed using a MAESPA metamodel, and the partitioning between latent and sensible fluxes was parameterized using the average partitioning from the outputs of the MAESPA simulations (Vezy et al., 2018). The shade tree and coffee transpiration ($Tr$) and sensible heat ($H$) were simulated using MAESPA metamodels. The net radiation of the shade tree and the coffee layers was computed as the sum of the latent ($LE$) and the sensible ($H$) heat fluxes of each layer. The stand net radiation ($Rn$) was then computed by summing the net radiation of the shade tree, the coffee and the soil.

#### 3.4.15. Inputs and outputs

All parameters needed for a DynACof simulation are stored in specific input files for the shade tree, the coffee, the soil, the site and the meteorology. A set of example data is included into the package so any user can start a simulation without any data. This set of example file is also archived as a git repository on github.com (https://github.com/VEZY/DynACof_inputs) for convenience. The main input values used for this study are listed on the paragraph 3.5. The input variables for the meteorology file should provide at least the maximum and minimum air temperature of the day (°C), the RAD or PAR (MJ $m^{-2}$ $d^{-1}$), and the relative humidity (%) or vapor pressure deficit (hPa). The full list of mandatory and optional inputs the user can provide are available from the documentation (https://vezy.github.io/DynACof/reference/Meteorology.html).

The model has 245 output variables, mainly for energy balance, water balance, carbon assimilation and allocation, respiration, plant growth, yield and mortality for each organ type. The full list is available from the documentation (https://vezy.github.io/DynACof/reference/DynACof.html).

### 3.5. Model parameterization using a multi-objective approach

The MAESPA model was parameterized according to Vezy et al. (2018) and was run on a 0.2 ha sub-plot of 4176 coffee resprouts (3 resprouts per coffee plant in average) and 14 shade trees at a half-hourly time-step throughout the year 2011. The shade trees are not planted along a regular planting pattern and they present large crowns, increasing the heterogeneity of the light distribution in the plot. The simulated sub-plot was reduced to the minimum area possible to decrease computation time, while encompassing coffee plants that are always in full sun or only partly under shade in the morning or during the afternoon.

Metamodels were then built from daily plot scale aggregations of MAESPA outputs and integrated in DynACof (Figure 1). The metamodels were built using linear regressions with MAESPA input variables as predictors. The MAESPA dataset created from the simulations of the year 2011 was taken as a representative sample of most of the conditions of the growing cycle, with yearly climate variations, a highly variable shade tree LAI due to almost total leaf fall, and highly variable coffee plant structure, with resprouts ranging from 0 to 5 years old.

DynACof was run from January 1979 to the end of December 2016 at a daily time step. The climate inputs to the model came from the Coffee-Flux project between 2009 and 2016, and were computed between 1979 and 2008 using the method and data described in Hidalgo et al. (2017). The values and sources of the parameters used in DynACof are listed in Table 1 for climate and coffee, in Table 2 for the shade tree species, and in Table

3 for the soil. Given the large number of parameters and the scarce data on coffee and shade trees, some parameters were not measured in this study and were not found in the literature either, for example, the coffee carbon allocation coefficients (λ). Because allocation coefficients have multiple repercussions on other variables, notably on respiration through organ mass and on light interception through LAI, a multi-objective parameterization was needed. As computation time was a limiting factor, a manual multi-objective parameter tuning was preferred to algorithmic optimization. The values of the parameters were found by starting from expert *a priori*, and tuned manually to minimize the simulation error for NPP and biomass for each compartment using the first year of measurements reported in Charbonnier et al. (2017) at a stand age of 33 year after planting, always keeping the values within a plausible range. In any case, the outputs of the model were evaluated on the second year of measurements, which was not used for model parameterization.

**Table 1. Main parameters used in DynACof (full list available at https://github.com/VEZY/DynACof). NB: SCR= Stump and Coarse Roots, FRoot= Fine roots, Shoot= Resprout wood (i.e. orthotropic axis and branches).**

| Parameter | Unit | Value | Description | Source |
|---|---|---|---|---|
| *Site* | | | | |
| Latitude | *degree* | 9.93833 | Latitude | This study[1] |
| Longitude | *degree* | -83.728 | Longitude | This study[1] |
| Timezone | | 6 | Time zone | This study[1] |
| Elevation | $m$ | 1040 | Elevation | This study[1] |
| Height_Coffee | $m$ | 1.2 | Coffee height, for zht and z0 | This study[1] |
| ZHT | $m$ | 25 | Climate data meas. height | This study[1] |
| Stocking_Coffee | $Plant\ ha^{-1}$ | 5580 | Coffee initial planting density | This study[1] |
| *Coffee light interception* | | | | |
| KDif | *0-1* | 0.3906 | Diffuse light extinction coeff. | This study[1] (MAESPA) |
| KDir | *0-1* | 0.3410 | Direct light extinction coeff. | This study[1] (MAESPA) |
| *Vegetative development* | | | | |
| AgeCofMax | *Year* | 40 | Max. length of plantation cycle | This study[1] |
| AgePruning | *Year* | 5 | Age at first pruning | This study[1] |
| D_pruning | *DOY* | 74 | Day of year for pruning | This study[1] |
| SLA | $m^2_{leaf}\ kg^{-1}_{DM}$ | 10.97 | Specific leaf area | Charbonnier et al. (2017)[1] |
| RNL | $Node\ LAI^{-1}$ | 91.2 | Ref. # nodes per LAI unit at 20°C | Drinnan and Menzel (1995)[2] |
| DELM | $g_C\ cof^{-1} d^{-1}$ | 2.0 | Max. leaf carbon demand | This study[1] |
| DVG1 | *DOY* | 105 | Beginning of vegetative growth | This study[1] |
| DVG2 | *DOY* | 244 | End of vegetative growth | This study[1] |
| kres | *0-1* | 0.08 | Max. reserves used per day | Cambou (2012) |
| $\lambda$_Shoot | *0-1* | 0.12 | Allocation to resprout wood | Charbonnier et al. (2017)[3] |
| $\lambda$_SCR | *0-1* | 0.08 | Alloc. to perennial wood | Charbonnier et al. (2017)[3] |
| $\lambda r$_Leaf | *0-1* | 0.85 | Remaining carbon alloc. to leaves | Charbonnier et al. (2017)[3] |
| $\lambda r$_FRoot | *0-1* | 0.15 | Rem. carbon alloc. to fine roots | Charbonnier et al. (2017)[3] |
| lifespan_Leaf | *day* | 265 | Leaf lifespan | Charbonnier et al. (2017) |
| lifespan_Shoot | *day* | 7300 | Resprout lifespan | van Oijen et al. (2010a) |
| lifespan_SCR | *day* | 7300 | Perennial wood lifespan | van Oijen et al. (2010a) |
| lifespan_FRoot | *day* | 365 | Fine root lifespan | Defrenet et al. (2016) |
| m_FRoot | *0-1* | 0.05 | Fine root to leaf pruning effect | This study[3] |
| CC_Fruit | $g_C\ g^{-1}_{DM}$ | 0.4857 | Fruit dry mass carbon content | Cambou (2012) |
| CC_Leaf | $g_C\ g^{-1}_{DM}$ | 0.463 | Leaf dry mass carbon content | Cambou (2012) |
| CC_Shoot | $g_C\ g^{-1}_{DM}$ | 0.463 | Resprout wood dry mass C content | Cambou (2012) |
| CC_SCR | $g_C\ g^{-1}_{DM}$ | 0.475 | Perennial wood dry mass C content | Cambou (2012) |
| CC_Shoot | $g_C\ g^{-1}_{DM}$ | 0.463 | Resprout wood dry mass C content | Cambou (2012) |
| $\varepsilon$Fruit | $g_C\ g_C^{-1}$ | 1.6 | Fruit growth respiration cost | Poorter (1994)[2] |
| $\varepsilon$Leaf | $g_C\ g_C^{-1}$ | 1.279 | Leaf growth respiration cost | Dufrêne et al. (2005)[3] |
| $\varepsilon$FineRoot | $g_C\ g_C^{-1}$ | 1.279 | Fine root growth respiration cost | Dufrêne et al. (2005)[3] |
| $\varepsilon$Shoot | $g_C\ g_C^{-1}$ | 1.2 | Shoot wood growth resp. cost | Dufrêne et al. (2005)[3] |
| $\varepsilon$SCR | $g_C\ g_C^{-1}$ | 1.31 | Perennial wood growth resp. cost | Dufrêne et al. (2005)[3] |
| NC_Fruit | $mg_N\ g_C^{-1}$ | 11 | Fruit nitrogen content | van Oijen et al. (2010a) |
| NC_Leaf | $mg_N\ g_C^{-1}$ | 29.6 | Leaf nitrogen content | Ghini et al. (2015) |
| NC_Shoot | $mg_N\ g_C^{-1}$ | 4.1 | Resprout wood nitrogen content | Ghini et al. (2015) |
| NC_SCR | $mg_N\ g_C^{-1}$ | 5 | Perennial wood nitrogen content | Cambou (2012) |
| NC_FRoot | $mg_N\ g_C^{-1}$ | 18 | Fine root nitrogen content | van Praag et al. (1988) |
| Q10_Fruit | *1* | 2.4 | Temperature effect on Rm | Charbonnier (2013) |
| Q10_Leaf | *1* | 2.4 | | Charbonnier (2013) |
| Q10_Shoot | *1* | 2.4 | | Charbonnier (2013) |
| Q10_SCR | *1* | 1.65 | | van Oijen et al. (2010a) |
| Q10_FRoot | *1* | 1.65 | | van Oijen et al. (2010a) |
| MRN | $g_C\ g_N^{-1}\ d^{-1}$ | 0.06336 | Base maintenance respiration | Ryan (1991)[2] |
| pa_Fruit | *0-1* | 1 | Percentage of living cells in fruits | This study[3] |
| pa _Leaf | *0-1* | 1 | P. of living cells in leaves | This study[3] |
| pa_FRoot | *0-1* | 1 | P. of living cells in fine roots | This study[3] |
| pa_Shoot | *0-1* | 0.37 | P. of living cells in resprout wood | Dufrêne et al. (2005) |
| pa_SCR | *0-1* | 0.21 | P. of living cells in perennial wood | Dufrêne et al. (2005) |

**Table 1 (continued). Main parameters used in DynACof.**

| Parameter | Unit | Value | Description | Source |
|---|---|---|---|---|
| *Reproductive development* | | | | |
| a_bud | $Buds\ d^{-1}$ | 0.00287 | Number of buds initiated per day | Rodríguez et al. (2011) |
| b_bud | *1* | -4.1. $e^{-6}$ | Param. for bud initialization | Rodríguez et al. (2011) |
| F_Tffb | *Degree day* | 4000 | Time of first floral buds | Rodríguez et al. (2011) |
| a_p | *1* | 5.78 | Probability of bud dormancy break calculated from leaf water potential | Drinnan and Menzel (1995); Rodríguez et al. (2011) |
| b_p | *1* | 1.90 | | |
| F_rain | *mm* | 40 | Cumulative rain to break bud dormancy | Zacharias et al. (2008) |
| age_Maturity | *Year* | 3 | First age of flowering after planting | van Oijen et al. (2010a) |
| VF_Flowering | *Degree day* | 5500 | Very first flowering of coffee plant | Rodriguez et al., 2001 |
| F_buds1 | *Degree day* | 840 | Bud stage 1 | Meylan (2012); van Oijen et al. (2010a) |
| F_buds2 | *Degree day* | 2562 | Bud stage 2 | |
| F_over | *Degree day* | 3304 | From pinhead to overripe (stage 5) | Rodríguez et al. (2011) |
| s | *1* | 0.05 | Empirical coefficient for fruit growth | Rodríguez et al. (2011) |
| FtS | *0-1* | 0.63 | Fruit to seed dry mass ratio | Wintgens (2004) |
| *Sucrose accumulation* | | | | |
| S_a | *[sucrose]* | 5.3207 | Parameters used to model sucrose accumulation in coffee fruit | Pezzopane et al. (2012) |
| S_b | *1* | -28.556 | | Pezzopane et al. (2012) |
| S_x0 | *Degree day* | 190.972 | | This study[3] |
| S_y0 | *[sucrose]* | 3.4980 | | Pezzopane et al. (2012) |
| DE_opt | $g_{DM}$ | 0.164 | Optimum berry C demand | Wintgens (2004) |

(1) Parameters effectively measured or computed for this study.
(2) Parameters computed from another source (e.g. MRN from Ryan (1991) transformed to a daily time scale)
(3) Parameters tuned starting from source value to make the model outputs match the first year of measurements from Charbonnier et al. (2017). Expert *a priori* is used as start value when no value was found in the literature.

**Table 2. Parameters used in DynACof for the shade tree layer (*E.poeppigiana*). NB: Leaf life span was computed using the phenology routine. CR= Coarse Roots, FRoot= Fine roots.**

| Parameter | Unit | Value | Description | Source |
|---|---|---|---|---|
| *Vegetative development* | | | | |
| SLA | $m^2_{Leaf}\ kg_{DM}$ | 17.4 | Specific leaf area | van Oijen et al. (2010a) |
| DELM_Tree | $g_C\ tree^{-1} d^{-1}$ | 778.5 | Max. leaf carbon demand | Charbonnier et al. (2017)[2] |
| $\lambda$_Stem | *0-1* | 0.20 | Alloc. to stem | This study[3] |
| $\lambda$_Branch | *0-1* | 0.25 | Alloc. to branches | This study[3] |
| $\lambda$_CR | *0-1* | 0.10 | Alloc. to coarse roots | This study[3] |
| $\lambda$_Leaf | *0-1* | 0.26 | Alloc. to leaves | This study[3] |
| $\lambda$_*F*Root | *0-1* | 0.05 | Alloc. to fine roots | This study[3] |
| lifespan_Branch | *day* | 7300 | Branch life span | van Oijen et al. (2010a) |
| lifespan_FRoot | *day* | 90 | Fine root life span | van Oijen et al. (2010a) |
| lifespan_CR | *day* | 7300 | Coarse root life span | van Oijen et al. (2010a) |
| CC | $g_C\ g_{DM}^{-1}$ | 0.47 | Tree dry mass carbon content | van Oijen et al. (2010a) |
| $\varepsilon$_Branch | $g_C\ g_C^{-1}$ | 1.2 | Branch growth respiration cost | This study[3] |
| $\varepsilon$_Stem | $g_C\ g_C^{-1}$ | 1.2 | Stem growth respiration cost | This study[3] |
| $\varepsilon$_CR | $g_C\ g_C^{-1}$ | 1.33 | Coarse root growth resp. cost | Litton et al. (2007) |
| $\varepsilon$_Leaf | $g_C\ g_C^{-1}$ | 1.392 | Leaf growth respiration cost | Villar and Merino (2001) |
| $\varepsilon$_FRoot | $g_C\ g_C^{-1}$ | 1.392 | Fine root growth resp. cost | = $\varepsilon$Leaf |
| NC_Branch | $mg_N\ g_C^{-1}$ | 5.0 | Branch nitrogen content | This study[3] |
| NC_Stem | $mg_N\ g_C^{-1}$ | 5.0 | Stem nitrogen content | This study[3] |
| NC_CR | $mg_N\ g_C^{-1}$ | 8.4 | Coarse root nitrogen content | van Oijen et al. (2010a) |
| NC_Leaf | $mg_N\ g_C^{-1}$ | 35.9 | Leaf nitrogen content | van Oijen et al. (2010a) |
| NC_FRoot | $mg_N\ g_C^{-1}$ | 8.4 | Fine root nitrogen content | van Oijen et al. (2010a) |
| Q10_CR | *1* | 2.1 | Temperature effect on Rm | This study[3] |
| Q10_Leaf | *1* | 1.896 | Temperature effect on Rm | This study[3] |
| Q10_Branch | *1* | 2.1 | Temperature effect on Rm | This study[3] |
| Q10_Stem | *1* | 1.7 | Temperature effect on Rm | van Oijen et al. (2010a) |
| Q10_FRoot | *1* | 1.4 | Temperature effect on Rm | van Oijen et al. (2010a) |
| pa_Branch | *0-1* | 0.4:0.05 | Percentage of living cells in branch | This study[3] |
| pa_Stem | *0-1* | 0.3:0.05 | P. of living cells in stem | This study[3] |
| pa_CR | *0-1* | 0.21 | P. of living cells in coarse roots | Dufrêne et al. (2005) |
| pa_Leaf, FRoot | *0-1* | 1 | P. of liv. cells in leaves and fine roots | This study[3] |
| *Allometries* | | | | |
| LAD_max/min | $m^2_{Leaf}\ m^{-3}$ | 0.76/0.21 | Max/Min leaf area density | Charbonnier et al. (2013) |
| AgePruning | *year* | 1:21 | Ages at which trees are pruned | This study[1] |
| Stocking | $tree\ ha^{-1}$ | 250/7.38 | Tree density (before/after thinning) | Taugourdeau et al. (2014) |

(1) Parameters effectively measured or computed for this study.
(2) Parameters computed from another source
(3) Parameters set using an expert value (e.g. values from other species) or tuned starting from a source value to make the model outputs match the first year of measurements from Charbonnier et al. (2017).

**Table 3. BILJOU sub-module parameters**

| Parameter | Unit | Value | Description | Source |
|---|---|---|---|---|
| TotalDepth | *m* | 3.75 | Total simulated soil depth | This study |
| Wm1; Wm2; Wm3 | *mm* | 210;58;64 | Minimum water content, layers 1, 2, 3 | Gómez-Delgado et al. (2011) |
| Wf1; Wf2; Wf3 | *mm* | 290; 66; 69 | Field capacity, layer 1;2;3 | Gómez-Delgado et al. (2011) |
| IntercSlope | $mm\ LAI^{-1}$ | 0.2 | Rainfall interception by leaves | Gómez-Delgado et al. (2011) |
| WSurfResMax | mm | 120 | Max. water on the surface reservoir | Gómez-Delgado et al. (2011) |
| fc | $mm\ day^{-1}$ | 13.4 | Min. infiltration capacity | Gómez-Delgado et al. (2011) |
| alpha | *1* | 101.561 | Coef. for max. infilt. capacity | Gómez-Delgado et al. (2011) |
| kB | *day-1* | 0.038 | Discharge coeff. for surface runoff | Gómez-Delgado et al. (2011) |
| Soil_LE_p | *%* | 0.70 | Soil energy partitioning coefficient | This study (MAESPA) |

## 4. Results

### 4.1. Metamodels

The metamodels for shade tree $KDif$ and $KDir$ are presented in Table 4, and were computed using the shade tree $LAD$ ($LAD_{Tree}, m^2\ m^{-3}$) as the sole predictor. The metamodel for $LUE$ ($g_C\ MJ^{-1}$) was made using climate inputs because it was found to depend more on the environment than on the plant structure. The other metamodels for plant transpiration ($Tr, mm$), sensible heat fluxes ($H, MJ\ m^{-2}$) and soil net radiation are also presented in Table 4.

**Table 4. MAESPA metamodel equations. $LAD_{Tree}$ ($m^{-2}\ m^{-3}$) is the leaf area density of the shade tree, $PAR_l$ ($MJ\ m^{-2}\ day^{-1}$) the photosynthetically active radiation reaching the layer considered (e.g. PAR transmitted by the shade tree layer for the $LUE_{Cof}$ and transmitted by the coffee layer for $Rn_{Soil}$), $T_{air}$ ($°C$) and $VPD_{air}$ ($hPa$) are the air temperature and vapor pressure deficit measured above the canopy, $APAR_l$ ($MJ\ m^{-2}\ day^{-1}$) is the PAR absorbed a the given layer, and $Wind$ ($m\ s^{-1}$) is the wind speed. $R^2$: adjusted r-squared, nRMSE: normalized Root Mean Squared Error, EF: modelling efficiency. Bias, nRMSE and EF statistics were computed on out-of-sample data, $R^2$ on the sample used to train the model.**

| Metamodel | $R^2$ | Bias | nRMSE | EF |
|---|---|---|---|---|
| $KDif_{Tree} = 0.6146417 - 0.5321444 \cdot LAD_{Tree} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.02296613)$ | 0.95 | $1.2{\cdot}10^{-3}$ | 6.52 | 0.95 |
| $KDir_{Tree} = 0.4754740 - 0.4015379 \cdot LAD_{Tree} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.05439675)$ | 0.65 | $5.2{\cdot}10^{-3}$ | 20.89 | 0.60 |
| $LUE_{Cof} = 2.77258689 + 0.01034341 \cdot T_{air} - 0.71823829 \cdot \sqrt{PAR_{Cof}} + 0.01537693 \cdot VPD_{air} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.0921815)$ | 0.94 | $2.1{\cdot}10^{-3}$ | 7.81 | 0.94 |
| $LUE_{Tree} = 2.91332985 + 0.07195458 \cdot T_{air} - 0.03124228 \cdot VPD_{air} - 0.24092238 \cdot PAR_{tree} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.2664409)$ | 0.87 | $-4.9{\cdot}10^{-3}$ | 9.84 | 0.89 |
| $Tr_{Cof} = -0.82327839 + 0.03089959 \cdot T_{air} + 0.15777166 \cdot APAR_{Cof} + 0.07297392 \cdot VPD_{air} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.1481428)$ | 0.88 | $5.1{\cdot}10^{-3}$ | 19.98 | 0.87 |
| $Tr_{Tree} = -0.55231394 + 0.01798136 \cdot T_{air} + 0.02229318 \cdot VPD_{air} + 0.16053935 \cdot LAI_{Tree} + 0.48838455 \cdot APAR_{Tree} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.08087603)$ | 0.83 | $7.5{\cdot}10^{-3}$ | 24.11 | 0.81 |
| $H_{Cof} = -0.7678758 + 0.5461716 \cdot PAR_{Cof} - 0.3207445 \cdot VPD_{air} + 2.6837429 \cdot Tr_{Cof} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.4506168)$ | 0.96 | $6.6{\cdot}10^{-2}$ | 10.14 | 0.97 |
| $H_{Tree} = 0.19869391 + 0.70967904 \cdot APAR_{Tree} - 0.78428282 \cdot LAI_{Tree} - 0.68444250 \cdot Tr_{Tree} - 0.03157318 \cdot VPD_{air} + 0.06780685 \cdot Wind + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.08)$ | 0.77 | $4.3{\cdot}10^{-3}$ | 58.85 | 0.79 |
| $Rn_{Soil} = -1.050189 + 1.766872 \cdot PAR_{Soil} + \hat{\varepsilon_i}\ where\ \varepsilon \sim N(0,0.2257816)$ | 0.98 | $4.2{\cdot}10^{-2}$ | 8.80 | 0.98 |

Their performance was also assessed using four different statistics which showed that, despite being simple models, the metamodel predictions were in agreement with the validation sub-sample of the outputs of MAESPA in the simulated year 2011 (Table 4). Indeed, all metamodels gave high $R^2$ on the data on which they were trained, and high modelling efficiency, rather low nRMSE and low bias when applied to new data, except for $KDir$, which failed to capture the high day-to-day variability but still followed the overall trend. Although $H_{Tree}$ did present satisfactory statistics overall, its nRMSE was high mainly due to the low (and sometimes negative) sensible fluxes of the shade tree which yielded a low average value.

### 4.2. GPP and energy fluxes

The modeled GPP, water and energy balance from DynACof were compared to the whole period of measurements using data gathered during the long-term Coffee-Flux eddy-covariance monitoring.

Overall, the model outputs for these variables computed in 2011 through metamodels were close to those computed directly from MAESPA (nRMSE: GPP = 15.4, Rn= 7.2, ETR= 23.9). Indeed, plant transpiration, plant sensible heat fluxes and soil net radiation were computed using MAESPA metamodels, and the parameter for soil energy partitioning into sensible and latent heat (i.e. soil evaporation) was also determined using the average value from the MAESPA simulations.

The simulated net radiation was close to measured values, with relatively high modeling efficiency (0.88) (Figure 2, Rn) and a low bias of 0.24 MJ $m^{-2}$ $d^{-1}$ that was reflected in the cumulated energy in Figure 3.a. The high point density along the identity function indicated that modeled evapotranspiration (Figure 2, AET) was relatively close to the measurement in average, even though the simulations presented a relatively high error for a few measurements (in dark blue in the figure). Yet, the cumulated AET from DynACof (Figure 3.b) showed good consistency compared to cumulated measurements, indicating that simulated values converge towards the measured value asymptotically.

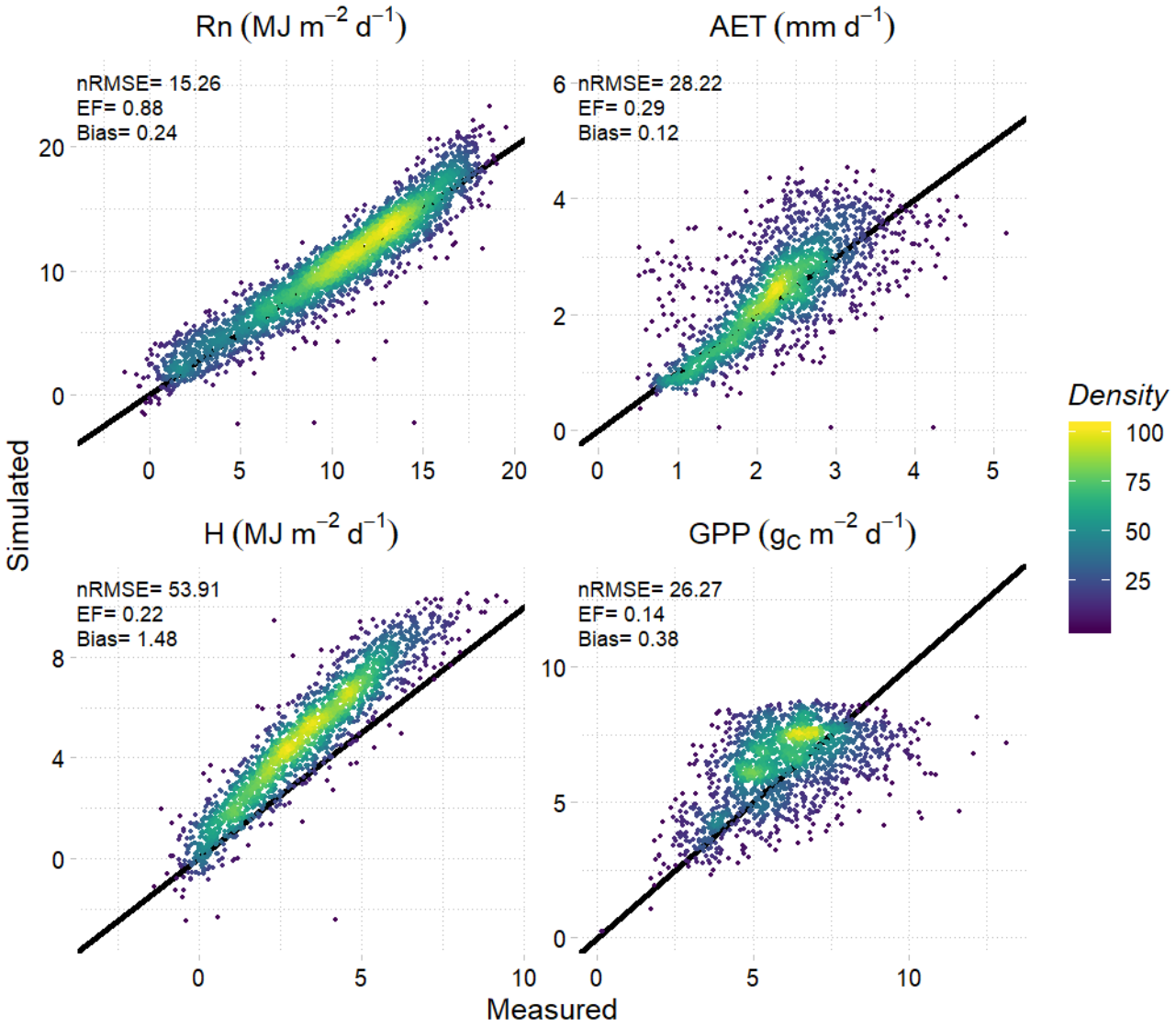


**Figure 2. Comparison between measured (x-axis) and modeled (y-axis, DynACof) net radiation (Rn, 2009-01-01 to 2016-11-28), actual evapotranspiration (AET, 2009-03-01 to 2013-12-31), sensible heat flux (H, 2009-03-01 to 2013-12-31) and gross primary productivity (GPP, 2009-01-01 to 2013-12-31) from the Coffee-Flux project in the Aquiares coffee agroforestry plantation at the scale of the whole plot (shade tree + coffee + soil layers) and at daily time scale. The color scale represents point density. nRMSE= normalized Root mean square error, EF= modeling efficiency and bias= modeling bias. One dot represents one day.**

The simulated sensible heat fluxes were in agreement with measured fluxes, except for a quasi-systematic bias of 1.48 MJ $m^{-2}$ $d^{-1}$ (Figure 2, H). The simulated GPP was close to the observed GPP in general (Figure 2, GPP), with a low positive bias of 0.38 $g_C$ $m^{-2}$ $d^{-1}$ and an nRMSE of 26.27. However, modelling efficiency remained low (0.14), mainly due to the high dispersion of the residuals, especially for high GPP values. Yet, when considering cumulated values of GPP over the whole period, the simulated GPP remained within the range of the measurement error (Figure 3.c). We also stress here that GPP remains a data-model product, derived from net ecosystem carbon flux measurements, but subject to several partitioning and modeling assumptions.

Overall, DynACof predictions of GPP and energy fluxes were in agreement with their respective measurements over the whole period (Figure 2 and Figure 3), confirming that the model performs reasonably well beyond the data on which the metamodels were calibrated (i.e. data from 2011).

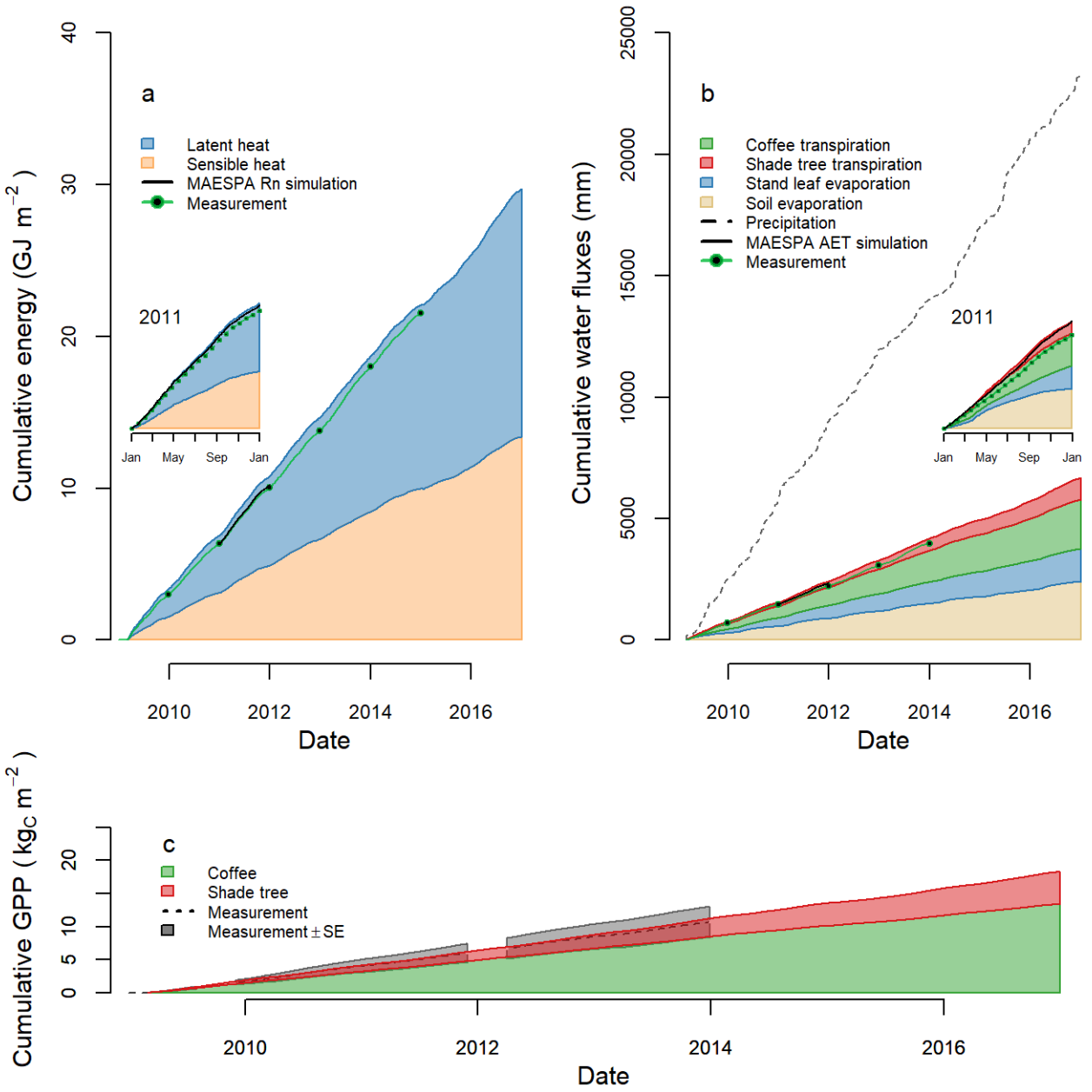


**Figure 3. DynACof simulation of cumulated partitioning for a\ energy, b\ evapotranspiration and c\ gross primary productivity over the 2009-2016 period at daily time-scale. MAESPA simulation of Rn (net radiation) and AET (actual evapotranspiration) in the year 2011 and measurements of the three variables are also presented for model assessment. Missing cumulative GPP values (i.e. gaps between measurement polygons) were gap filled using DynACof modelled GPP to present a longer measurement period. The inserts in figures a\ and b\ show the cumulative energy partitioning and evapotranspiration for the year 2011 for easier comparison with MAESPA simulations from Vezy et al. (2018).**

DynACof outputs were also compared to standard equations (Allen et al., 1998; Landsberg et al., 2001) that are widely used for crops or plantations with more homogeneous canopies (i.e. constant LUE and Ks parameters for GPP, Rn computed using an average albedo, LE computed using the Penman-Monteith equation). As expected, the error was higher when the net radiation, the sensible and latent heat flux and the GPP were simulated using this approach compared to a simulation using metamodels from MAESPA (Table 5 and Figure C.1). Hence, the choice of metamodels in DynACof was justified.

**Table 5. DynACof modelling performance with (*yes*) and without (*no*) metamodels from MAESPA.**

| Statistics | nRMSE | | EF | | Bias | |
|---|---|---|---|---|---|---|
| *Metamodel* | *yes* | *no* | *yes* | *no* | *yes* | *no* |
| Rn (MJ $m^{-2}$ $d^{-1}$) | 15.26 | 31.84 | 0.88 | 0.50 | 0.24 | 2.87 |
| LE (MJ $m^{-2}$ $d^{-1}$) | 28.22 | 42.46 | 0.28 | -0.68 | 0.29 | 1.44 |
| H (MJ $m^{-2}$ $d^{-1}$) | 53.92 | 130.07 | 0.22 | -3.56 | 1.48 | 4.03 |
| GPP ($g_C$ $m^{-2}$ $d^{-1}$) | 26.28 | 41.36 | 0.14 | -1.14 | 0.38 | 0.99 |

### 4.2.1. Growth and yield

Overall, the multi-objective calibration of DynACof yielded realistic results for most compartments that were documented by field measurements.

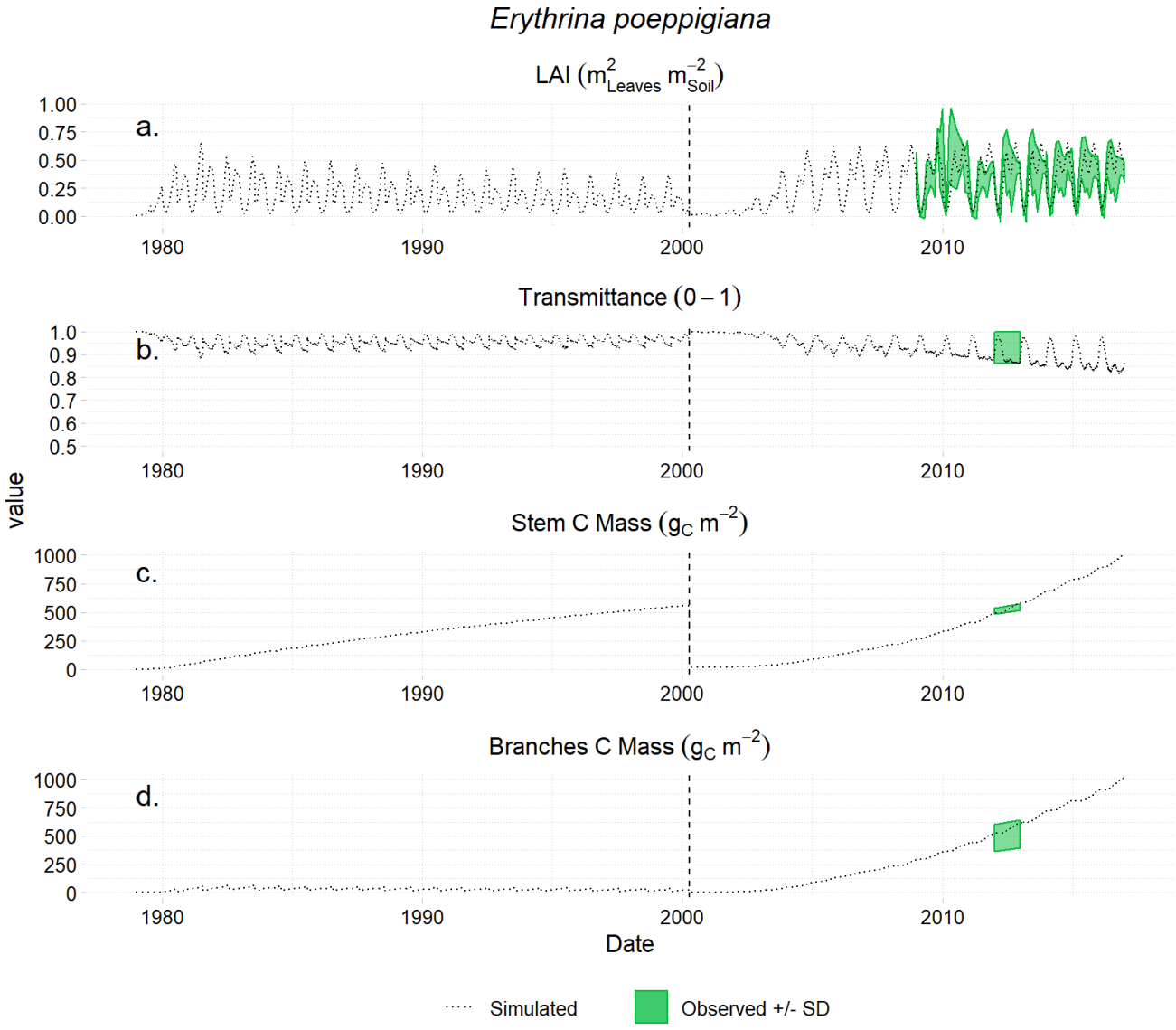


**Figure 4. Main outputs of the *Erythrina poeppigiana* shade tree simulated by DynACof over the whole rotation. Coffee plants were planted in 1979 together with Erythrina. Trees were pruned twice a year until thinning in 2000, and then allowed to grow freely. The thinning event is represented by the vertical dashed line. a\ LAI dynamics compared to measurements by LAI2000 (Taugourdeau et al., 2014) +/-SD ; b\ Light transmitted by the shade tree compared to Charbonnier et al. (2013) plot average; c\ Stem and d\ branch carbon mass compared to measured values reported in Charbonnier et al. (2017).**

Despite the high initial planting density, the LAI of the shade trees remained relatively low during the first period when the shade trees were pruned, i.e. from planting to thinning, when it dropped to a particularly low value of 0.02 $m^2$ $m^{-2}$ on average in the first year, then resumed during the second period. The LAI of the shade trees subsequently increased to reach a maximum of ca. 0.66 $m^2$ $m^{-2}$ in the last year of the simulation (Figure 4). The leaves of *E. poeppigiana* shed naturally between January and February in Aquiares and then recover rapidly until May (Taugourdeau et al., 2014). The observed phenology was matched by the model, with a simulated range and dynamic of LAI values close to the observations made in the same plot averaged over the whole measurement period (Taugourdeau et al., 2014). Despite a low density of 7.3 trees $ha^{-1}$ after thinning in 2000, the simulated shade trees intercepted up to 18.4% of the light at maximum LAI in 2016, which was consistent with the values measured in the same plot (Charbonnier et al., 2013). The simulated dry mass of tree stem and branches represented on average 47.2% and 4.4% of the total plot shade tree carbon mass while pruned, and 27.2% and 27.4% of the total carbon mass, respectively, at the end of the cycle when left to grow freely. The annual stem mass growth was close to linear under pruning management and became close to exponential when trees were not pruned after thinning, with the stem NPP increasing almost five times in the five last years of the simulation compared to the five years preceding thinning (Table 6). Branch mass grew much more slowly due to higher mortality when pruned but grew as fast as the stem when the tree was left to grow freely. As seen in Figure 4, both stem and branch mass simulations were in the range of the observations reported by Charbonnier et al. (2017) in the same plot.

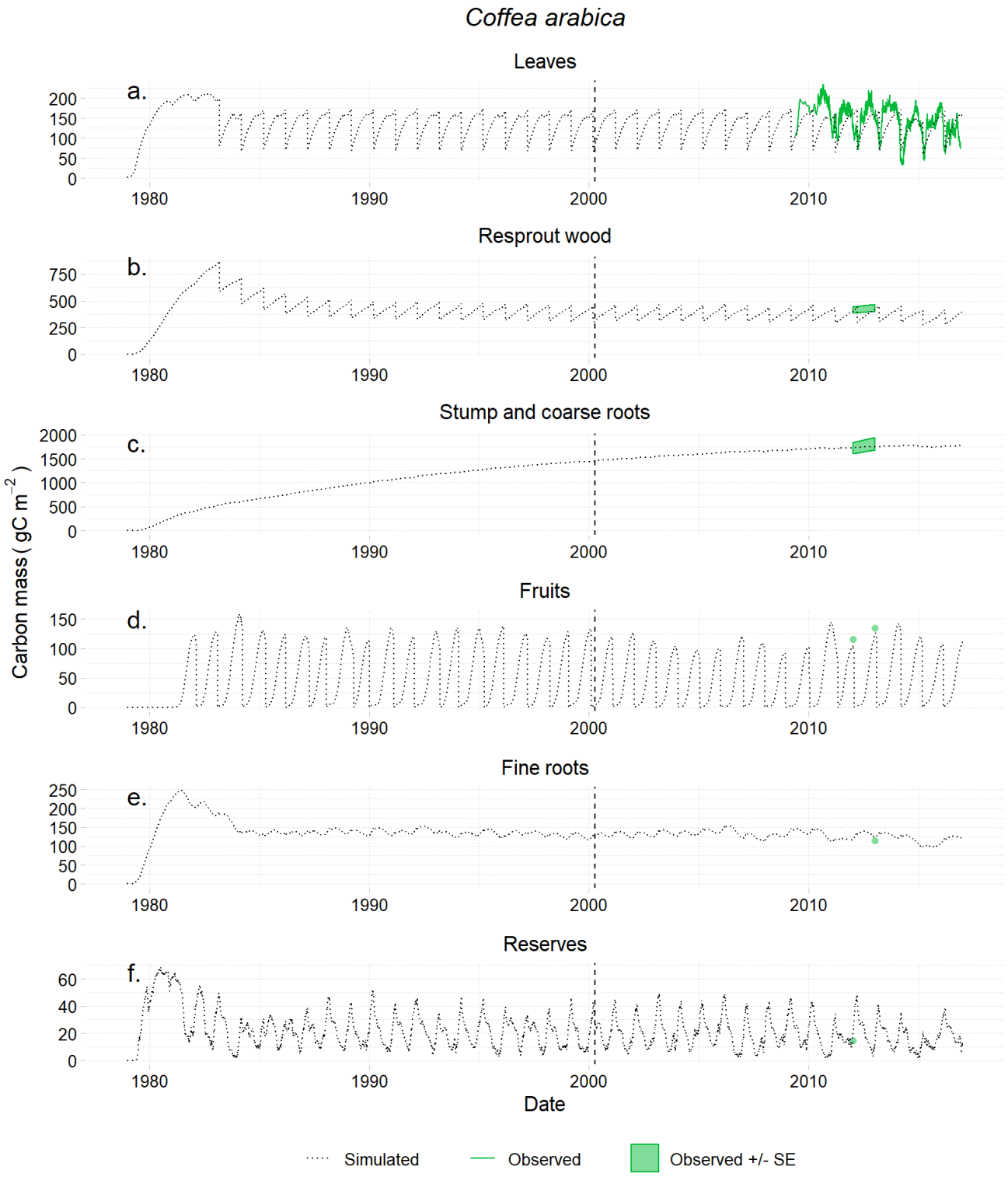


**Figure 5. Simulated coffee plant carbon biomass per organ (dotted lines) over a full plantation cycle (1979-2016), compared to measured biomass (green dots: one observation, green lines: continuous measurement and polygon: mean observation +/- SE). a/ Simulated leaf C mass compared to measured NDVI calibrated using LAI2000, and converted into dry mass using leaf SLA and dry mass to carbon mass ratio (Taugourdeau et al., 2014), see Table 1; b/ Simulated wood C mass from branches compared to average +/-SE measured by Charbonnier et al. (2017); c/ Simulated C mass in stump and coarse roots compared to measured stump C mass +/- SE in Charbonnier et al. (2017) + measured perennial root C mass from Defrenet et al. (2016); d/ Simulated fruit dry mass compared to values measured by Charbonnier et al. (2017) for 2011 and 2012 at harvest; e/ Simulated fine roots C mass compared to a 2011 measurement by Defrenet et al. (2016); and f/ Simulated reserves compared to reserves measured by Cambou (2012), assumed to represent the minimum value of the year.**

Dynamic simulations of the carbon mass of the coffee organs are plotted for a full growing cycle from 1979 to 2016 in Figure 5. Coffee leaf carbon mass (Figure 5.a) increased rapidly until reaching its maximum value of 211 $g_C$ $m^{-2}$ at three years old (onset of fruiting). Coffee pruning started 5 YAP, and leaf carbon mass then fluctuated between 58 and 174 $g_C$ $m^{-2}$ until the end of the simulation, corresponding to LAI values between 1.4 and 4.1 $m^2$ $m^{-2}$. The LAI dynamics expressed a yearly minimum by April after the drier season and just after pruning, and a second minimum in September, during bean filling, very similar to field observations reported by Taugourdeau et al. (2014). Interestingly, this realistic phenology was not prescribed into the model but appeared after the introduction of fruit cohorts, inducing a strong but progressive competition between sinks (here with the leaves) throughout the fruiting season. Resprout wood (Figure 5.b) grew rapidly from 0 to 4 YAP, before the onset of the pruning cycle which occurred every year for all resprouts aged more than 5. Resprout wood growth then decreased to reach a stable value of ca. 397 $g_C$ $m^{-2}$ at around 13 YAP, with only intra-annual fluctuations due to pruning. The perennial compartment represented by both the stump and the coarse roots grew progressively throughout the crop cycle because it was not subject to pruning and has a very long lifespan: it reached a maximum value of 1769 $g_C$ $m^{-2}$, or 37.3 $t_{DM}$ $ha^{-1}$ in the end of 2016 (Figure 5.c). The coffee fruit compartment (Figure 5.d) started to yield from the third YAP with a carbon mass of between 71 and 134 $g_C$ $m^{-2}$ at harvest and an average modeled green bean production of 1 382 ($\pm$205) kg $ha^{-1}$ $year^{-1}$. The simulated carbon mass of the fruits was close to values measured by Charbonnier et al. (2017) in 2012 and 2013. The fine roots (Figure 5.e) grew rapidly in the early growth stages, i.e. until the third YAP when the combined effects of pruning and natural mortality maintained their carbon mass at a relatively stable level from one year to the next with values of ca. 134 $g_C$ $m^{-2}$. The reserves compartment fluctuated seasonally, always in opposition with the fruit carbon growth. The simulated carbon reserve was close to the reserve measured by Cambou (2012) in the same plot (Figure 5.f) because the model was parameterized using this data.

The modelled coffee carbon allocation by organs showed that plant reserves represented the compartment with by far the highest carbon flow, capturing on average 40.4% of the yearly plant carbon supply during the five last years of the simulation, with a daily maximum allocation of 63.1% and a minimum of 0% during fruit production (Table 6). This compartment also had a high turnover rate because reserves were re-distributed back to the carbon supply pool, making it an effective carbon buffer for the coffee plant from one season to another. Leaves and branches were the organs with the highest NPP, representing 21.5% and 11.9% of the total yearly NPP respectively, during the last five years of the simulation. Fine roots represented 9.4%, stump and coarse roots 7.3%, and fruits 9.5% of total NPP. The high productivity of the branches was related to their high carbon demand that was often met by the supply, as well as for the stump and coarse roots. The simulated NPP was reasonably consistent with measurements for coffee (Table 5). For the shade trees, no direct measurements of the whole rooting system was available, therefore the comparison is not proposed here. Nevertheless, when cumulated into biomass, the results for both species were realistic (Figure 4, Figure 5).

**Table 6. DynACof simulation of NPP per organ type and plant layer separated according to shade tree management options, either with high density and pruned (until 2000) or free growing with low density after thinning in 2000. Average measurements for two years (2011/2012 – 2012/2013) from Charbonnier et al. (2017) are also provided for comparison.**

| Organ | Average NPP ($g_C\ m^{-2}\ year^{-1} \pm SD$) for the last 5 years | | |
|---|---|---|---|
| *Coffea arabica* | *Pruned Ep* | *Free growing Ep* | *Measured (free growing Ep)* |
| Leaves | 308.9 (3.0) | 303.3 (12.3) | 263.5 |
| Perennial wood: Stump + coarse roots | 104.2 (4.0) | 99.1 (12.2) | 161.5 |
| Branches | 170.7 (6.6) | 162.2 (20.0) | 210 |
| Fine roots | 134.5 (7.1) | 127.2 (18.4) | 148 |
| Fruits | 136.3 (7.1) | 133.0 (11.3) | 126 |
| Reserve balance | 580.0 (51.3) | 534.1 (119.1) | |

## 5. Discussion

### 5.1. Metamodels

Simulating complex processes in crop models using metamodels is a promising way to reduce computation intensity and avoid numerous equations that are often an important part of the development and application effort. Another advantage is that physiological data (e.g. photosynthetic parameters) are often sampled at a fine scale but can be used to parameterize a precise model (in this case, MAESPA) at a sub-hourly time step, and then be upscaled to field scale and to daily scale. In the present study, metamodels helped DynACof consider fine-scale processes explicitly such as spatial anisotropy instead of only using a parameter (e.g. a percentage of canopy cover), which significantly improved the model simulations. Replacing some metamodels by simple standard plot-scale models leads to a dramatic increase in simulation errors, as shown in the present study (Table 5), evidence that our metamodeling approach was appropriate.

The quality of the predictions of a metamodel relies first upon the ability of the original model to correctly simulate the processes involved in the system, and secondly on the ability of the metamodel to reproduce the outputs of the original model. The first point was investigated in a previous work (Vezy et al., 2018) where MAESPA successfully simulated the energy balance and the evapotranspiration in the same experimental plot. The second point was addressed by choosing metamodels according to the best possible balance between the model complexity, the number of explanatory variables, and out-of-sample prediction quality. Marie et al. (2014) found that despite being slower to compute, neural networks and multi-linear regressions with two or three level of interactions yielded higher $R^2$ than multi-linear regressions with no interactions, like those used in our study. However, four out of ten metamodels in our study gave $R^2$ higher than 0.90 with low nRMSE, which is considered highly accurate, three gave $R^2$ higher than 0.80, which is considered accurate (Villa-Vialaneix et al., 2012), and only one metamodel could be considered not sufficiently accurate with a $R^2$ of 0.65. For example, the LAD was used alone to predict the light extinction coefficient of the shade tree layer which was then used to compute its light absorption from its LAI. The diffuse extinction coefficient was predicted with high accuracy with an EF and a $R^2$ both equal to 0.95. The LAD is a useful proxy for within crown foliage aggregation, which was determined - along with the LAI - as the most important characteristic to model light penetration by Sampson and Smith (1993). However, other structural variables from the shade trees can be used as predictors to improve the metamodel accuracy. This was not the case in our study because shade trees structural data were measured only once for the year 2011, so only one value was given as input to MAESPA.

The metamodels for $LUE$ yielded $R^2$ and nRMSE similar to those found in Christina et al. (2016): $R^2$ of 0.87 compared to 0.87 and 0.94 for shade tree and coffee respectively in our study, and an RMSE of 0.20 $g_C$ $MJ^{-1}$ compared to 0.24 and 0.10 (shade tree and coffee resp.) in our study. Yet, despite being generally effective at reproducing a complex model output, applying metamodels to new conditions requires caution, as they can produce unexpected results outside their training values, especially if they use non-linear equations or when there is covariance between predictors. This was of particular concern in the present study because metamodels were fitted using a one year-long simulation of MAESPA only (2011). To overcome these problems, the 2011 data was checked to present a broad range of values for the target variables. Some conditions were still not met in the training sample, such as the period before the year 2000, when shade trees were pruned twice a year while the metamodels were trained on the system with free-growing trees. Yet, the 2011 training period included some days with very low *E. poeppigiana* LAI (i.e. 0.04 $m^2$ $m^{-2}$) because *E. poeppigiana* loses all its leaves once a year, which helped the metamodel simulate a plausible range of transmittance values under low LAI pruning conditions.

Likewise, the cumulated evapotranspiration and energy balance were satisfactorily predicted compared to measurements, even outside their training period, although both computations depended to a great extent upon metamodels. The metamodel for the coffee LUE predicted an increase in values with a reduction in incoming radiation on the plant layer, which is in agreement with previous results reported in Charbonnier et al. (2017). As a result, DynACof simulations of Rn, GPP and AET were close to those produced by MAESPA reported in Vezy et al. (2018), and more importantly, close to the eddy-covariance measurements from the long term Coffee-Flux monitoring from 2009 to the end of 2016 (Figure 2 and Figure 3). The agreement between simulated and measured values was particularly strong when cumulative fluxes were compared, which confirmed the high degree of consistency throughout the measurement period, mostly due to a low bias from the metamodels. Hence, metamodels proved to be powerful tools to overcome the long-term trade-off between speed, accuracy, genericity and fast development of growth and yield models able to simulate whole crop rotations.

### 5.2. Growth and yield model outputs from DynACof

The evaluation of a crop model is often challenging due to the lack of data for parameterization and validation, yet our model was subjected to multi-objective evaluation using numerous observations from the same experimental field or from the literature. Nevertheless, it should be noted that the model was mostly parameterized using values in the literature, and better results would be expected using measurements or an optimization algorithm (Van Oijen et al., 2005). However, the model satisfactorily predicted most processes at plot scale with little or no discrepancy. The total autotrophic respiration represented 55% of the GPP, which is close to the 57% reported by Litton et al. (2007) in a review of results from a wide range of forest ecosystems. The simulated LAI for coffee was in agreement with the measured LAI, and the mean simulated leaf dry mass from 2011 (= 132.2 $g_C\ m^{-2}$) was in agreement with the measured values reported in Charbonnier et al. (2017), Taugourdeau et al. (2014) and Siles et al. (2010), with values of 140.5, 143.7 and ranging from 102 to 176 $g_C\ m^{-2}$ respectively. The seasonal behavior of leaf biomass revealed a drop at the end of the "drier" season

corresponding to natural leaf shedding followed by pruning, then a rapid increase at the beginning of the rainy season with a secondary minimum when fruit dry mass was high. Interestingly, the simulations well mimicked the seasonal observations reported by Taugourdeau et al. (2014), and the simulation was close to measured values. The seasonality of the LAI was represented using two main drivers: pruning the leaves once a year, and the introduction of fruit cohorts. The first is a forced process, but the second is the result of successive computations that allows a smooth distribution of the demand for carbon required for grain filling during the period of reproductive development, depending on the developmental stage of each cohort. However, to date, the interannual variability in leaf area is barely perceivable in the simulations compared to the field conditions. We assume that some processes driving this variability still need to be incorporated in the model, for example, a dynamic leaf life span using cohorts of leaves, more leaf diseases or nitrogen effects. Indeed, American leaf spot (ALS) was already included in the model following Avelino et al. (2007), but coffee leaf rust is the predominant disease affecting coffee plants in this region, which is not yet included due to the absence of a published model linking disease severity and leaf loss.

Total NPP is the consequence of the carbon assimilation, its allocation and the respiration of each specific organ. Comparing the measured and simulated total NPP is an integrative evaluation of the model, and the total coffee NPP simulated by DynACof was in agreement with the NPP measured by Charbonnier et al. (2017), with an overestimation of ca. 3.6% for both years (2011-2012 and 2012-2013).

However, the most important but most challenging integrative process to simulate is fruit yield, because its allocation follows a complex scheme spread out over two years with numerous development stages (Camargo and Camargo, 2001), which was modeled in DynACof using a formalism inspired by Rodríguez et al. (2011). DynACof predicted a green bean yield of 1382 ($\pm$205) kg $ha^{-1}$ $year^{-1}$, which was within the range of values observed by Campanha et al. (2004), van der Vossen et al. (2015) and of the average yield in Central America (Söndahl et al., 2005). Furthermore, the Aquiares farm reported average yields of green beans of around 1333 ($\pm$336) kg $ha^{-1}$ $year^{-1}$ between 1995 and 2014 for fields close to the experimental plot, confirming that DynACof yield predictions were consistent. However, this comparison is only indicative and should be interpreted with caution because the conversion from whole fruit dry mass into processed green beans was made using a simple parameter (FtS, see Table 1) that may change depending on several factors, and because DynACof only simulates potential yield in the absence of fruit diseases or predators. Furthermore, the dataset used to fit the effect of temperature on the buds could overestimate the negative effect because Drinnan and Menzel (1995) used coffee plants grown in relatively small pots (10 L), which can restrict root growth and then negatively affect the physiology of the coffee plant (Ronchi et al., 2006). DaMatta et al. (2019) found similar results on unpublished data though, confirming this negative effect is not only coming from the restricted root growth.

Moreover, the model was also compared to *in situ* measurements reported in $g_C$ $m^{-2}$ by Charbonnier et al. (2017), and predictions were found to be close to measured values, even following the same pattern of yearly variability, which is particularly hard to achieve considering the number of formalisms in use, and the potential cumulated error from one process to another.

The shift in the shade tree management from pollarded to free-growing appeared to have little impact on fruit production or maturity at harvest. This apparent stability came from the low density of the shade trees, which still transmitted at least 81.6% of the light during the mature state according to DynACof, compared to 86% reported in Charbonnier et al. (2013). Charbonnier et al. (2017) reported that the higher LUE simulated by MAESPA for coffee plants under higher shade could compensate for most of the reduction in incident PAR, maintaining NPP at a nearly constant level as long as shade remains at low values. In DynACof, GPP was reduced only by 2.6% for a reduction of 7.9% in APAR, thanks to an increase of 5.8% in the LUE between the two periods of shade tree management, i.e. from low LAI in pruned shade trees to higher LAI in free-growing shade trees.

Another strength of the model is the prediction of canopy temperature under shade as a driver for plant biology, and water and energy balances, thanks to a full soil module inspired from the BILJOU model and to the MAESPA metamodels. Indeed, predictions of the cumulated AET and net radiation were very close to the nearly continuous measurements made from 2009 to 2015.

Given the model gave satisfactorily results for a wide range of processes using our multiple-objective strategy of parameterization and evaluation, i.e. different sets of variables measured in the same site at the same time, it can provide other information that cannot be discovered from the data only, and help researchers identify emergent properties in the system. For example, coffee LAI was strongly affected by pruning once a year and during the period in between by natural mortality and by the high fruit demand at the time of grain filling, which was also observed by Charbonnier et al. (2017). Another observation made using the model outputs is that except for stump and coarse roots, which are the only perennial compartments, biomass increased rapidly in the early stages of the plantation until it reached its maximum value for the whole rotation, after which biomass growth started to decrease with pruning, and found a new lower equilibrium between growth and mortality. The model also reproduced to some extent the biennial fruit production that was reported by Cannell (1985). Finally, although the measured variables were acquired mostly at the end of the rotation, the model ran for the full rotation and the simulated variables reached the values measured by the end of the rotation, thereby revealing dynamics that remained hidden during the unknown part of the rotation (e.g. biomass reaching a specific equilibrium, according to the pruning intensity). Overall, it is clear that the model is able to compute several long term ecosystem services and paves the way for the analysis of possible trade-offs between them, depending on different management options and climate.

Although the model now needs to be applied in different soil, climate and shade management conditions to evaluate its true genericity, the process-based equations implemented in most modules already imply relative genericity. For example, the flowering events and timing are based on the model from Rodríguez et al. (2011), which reproduced satisfactorily these processes on contrasted sites. However, the user should keep in mind that metamodels are not generic, and should be updated when applied to new conditions (e.g. different soil or planting design). The overall genericity of DynACof should allow its application to other climate conditions such as under climate changes, other locations with different soils and climates, and other shade management systems such as coffee grown in full sun or under *Cordia alliodora*, banana plants, *Eucalyptus sp.* or any species,

tree density and management such as pruning or thinning. In specific situations, it would be advisable to develop specific modules for coupling the nutrient cycles with the carbon and water cycles already available in the current version of DynACof. Therefore, the current version is assumed to simulate the potential outputs, only for situations without any nutrient stresses. Drought is already present in the model through leaf water potential and affects the reproductive phenology, but should be tested and refined for other vegetative limitations.

### 5.3. Model adaptation

DynACof can be easily adapted by users to other sites, other shade trees, other coffee varieties or managements by parameterizing the model accordingly. Adapting the model to other sites require mainly updating the site parameters (e.g. latitude, longitude, elevation), the soil properties (e.g. field capacity, pore fraction) and the meteorology. If the coffee cultivar is changed, the parameters linked to the coffee should be checked and updated if necessary (e.g. specific leaf area, allocation coefficients, base temperature). To change the management of the shade tree or the species, the user has to update the shade tree parameter file. The metamodels should be updated by the user whenever the plants structure or functioning is out of the range they were trained on. For example the LUE metamodel should be fitted using the atmospheric $CO_2$ concentration and air temperature to integrate their effect for simulations under climate changes. They also can be replaced by any plot-scale equations because they are written as input code.

The model can then be calibrated and evaluated using field data. LAI for both shade tree and coffee is a key variable to measure because it is the result, and has an impact on many processes. The biomass increment of each organ can be measured to evaluate the plant NPP. For the coffee, key variables from the reproductive development are useful to evaluate the simulated yield, e.g. time of initiation and number of buds and flowers, and of course yield itself. The soil water content can be measured to evaluate the water balance module, and the leaf and soil temperature for the energy balance module.

DynACof can also be modified easily by developers because it was designed as modular as possible, with each module called in sequence at daily time-step. Developers can replace any module while still leveraging the others easily. For example the coffee module could be replaced by a module simulating any other plant, annual or perennial (*e.g.* cocoa, wheat, sugar cane, rice...). New modules can also be added, such as for nitrogen cycle. This module could compute the nitrogen uptake, allocation and content for all organs from each plant species, and the mineralization processes from decaying organic matter (which already exist in the model) and microbial biomass in the soil, with eventual mineral fertilization input by management (e.g. urea, synthetic fertilizers, irrigation) or natural processes (e.g. precipitations).

## 6. Conclusion

DynACof (Dynamic Agroforestry Coffee Crop Model) was developed to simulate the effects of the environment, the soil, the species of shade tree and the management practices on coffee growth and yield. The shade management module can be set to any shade type and density, under full sun or agroforestry systems,

applying pruning or thinning at any age, if required. The model can be used for full rotations at a daily time step for any surface area, from plot to landscape, or even to a region if properly distributed, under current, past or future climate conditions as long as the metamodels, built from MAESPA model simulations, are updated to the target conditions. The model was parameterized and evaluated using a comprehensive and unique dataset for energy and water balance, biomass and NPP from an experimental site in the Aquiares farm in Costa Rica. A substantial advantage of DynACof being a tree-average plot model is the possibility to parameterize it using plot averages or totals, which are more frequently available from farms (e.g. yield, pruning intensity, coffee quality, etc.) because data remain scarce, especially under agroforestry management.

Two other important features of the model are the simulation of the canopy temperature (instead of air temperature) to control the plant growth according to the shade level, and the use of cohorts of flowers and fruits to consider grouped flowering in sub-tropical conditions and distributed flowering in equatorial climates. The model is implemented as an R and Julia packages for easy sharing and collaboration, and can be easily modified by adding new modules to compute pest attacks, nutrient cycling, soil organic matter decomposition or soil respiration. The methodology can be further generalized for any type of shade or climate by using different MAESPA simulation sets to train the metamodels. DynACof was built using a set of generic modules (e.g. functions for aerodynamic conductance) that can be used in other models to simulate any type of agroforestry systems or intercropped systems.

In conclusion, DynACof concentrates considerable ecophysiological knowledge on coffee, and is an efficient tool to evaluate and optimize coffee crop yield, ecosystem services and their trade-offs in response to climate conditions and management scenarios. It was also designed to predict the impacts of climate change on coffee yield and the potential of changes in management to mitigate such effects.

## Conflicts of interest

None.


## Acknowledgements

This project was funded by *Agence Nationale de la Recherche* (MACACC projet ANR-13-AGRO-0005, *Viabilité et Adaptation des Ecosystèmes Productifs, Territoires et Ressources face aux Changements Globaux* AGROBIOSPHERE 2013 programme), CIRAD (*Centre de Coopération Internationale en Recherche Agronomique pour le Développement*) and INRA (*Institut National de la Recherche Agronomique*). The authors are grateful for the support of the Aquiares farm (https://aquiares.com/) and for the long-term coffee agroforestry trial, 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/); the CIRAD-IRD-SAFSE project (France) and the PCP platform of CATIE. The Coffee-Flux observatory was supported and managed by CIRAD researchers. We are grateful to the staff in Costa Rica, in particular Alvaro Barquero, Alejandra Barquero, Jenny Barquero,

Alexis Perez, Guillermo Ramirez, Rafael Acuna, Manuel Jara, Alonso Barquero for their technical and field support.

The project analyzes largely benefited from the Montpellier Bioinformatics Biodiversity (MBB) computing cluster platform which is a joint initiative of laboratories grouped in the CeMEB LabEx "Mediterranean Center for Environment and Biodiversity", as part of the program "*Investissements d'avenir*" (ANR-10-LABX-0004).

## References

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M.J.F., Rome, 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. 300(9): D05109.

Avelino, J. et al., 2007. Topography and crop management are key factors for the development of American leaf spot epidemics on coffee in Costa Rica. Phytopathology, 97(12): 1532-1542.

Bezanson, J., Edelman, A., Karpinski, S. and Shah, V.B.J.S.r., 2017. Julia: A fresh approach to numerical computing. 59(1): 65-98.

Camargo, Â.P.D. and Camargo, M.B.P.D., 2001. Definition and outline for the phenological phases of arabic coffee under brazilian tropical conditions. Bragantia, 60(1): 65-68.

Cambou, A., 2012. Mesures des sucres lents et rapides d'organes de caféier par double approche VISNIR et Biochimique, Césure ENSAIA, Nancy, 30 pp.

Campanha, M.M. et al., 2004. Growth and yield of coffee plants in agroforestry and monoculture systems in Minas Gerais, Brazil. Agroforestry Systems, 63(1): 75-82.

Campbell, G.S.J.S.s., 1974. A simple method for determining unsaturated conductivity from moisture retention data. 117(6): 311-314.

Cannell, M.G.R., 1985. Physiology of the Coffee Crop. In: M.N. Clifford and K.C. Willson (Editors), Coffee: Botany, Biochemistry and Production of Beans and Beverage. Springer US, Boston, MA, pp. 108-134.

Charbonnier, F., 2013. Measuring and modeling light, water and carbon balance and net primary productivity in a coffee-based agroforestry system of Costa Rica, Université de Lorraine.

Charbonnier, F. et al., 2013. Competition for light in heterogeneous canopies: Application of MAESTRA to a coffee (Coffea arabica L.) agroforestry system. Agricultural and Forest Meteorology, 181: 152-169.

Charbonnier, F. et al., 2017. Increased light-use efficiency sustains net primary productivity of shaded coffee plants in agroforestry system. Plant Cell Environ, 40(8): 1592-1608.

Christina, M. et al., 2016. Sensitivity and uncertainty analysis of the carbon and water fluxes at the tree scale in Eucalyptus plantations using a metamodeling approach. Canadian Journal of Forest Research, 46(3): 297-309.

DaMatta, F.M., Rahn, E., Läderach, P., Ghini, R. and Ramalho, J.C.J.C.C., 2019. Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? , 152(1): 167-178.

Dauzat, J., Rapidel, B. and Berger, A., 2001. Simulation of leaf transpiration and sap flow in virtual plants: model description and application to a coffee plantation in Costa Rica. Agricultural and Forest Meteorology, 109(2): 143-160.

Defrenet, E. et al., 2016. Root biomass, turnover and net primary productivity of a coffee agroforestry system in Costa Rica: effects of soil depth, shade trees, distance to row and coffee age. Ann Bot.

Drinnan, J. and Menzel, C., 1995. Temperature affects vegetative growth and flowering of coffee (Coffea arabica L.). Journal of Horticultural Science, 70(1): 25-34.

Dufrêne, E. et al., 2005. Modelling carbon and water cycles in a beech forest: Part I: Model description and uncertainty analysis on modelled NEE. Ecological Modelling, 185(2-4): 407-436.

Duursma, R.A. and Medlyn, B.E., 2012. MAESPA: a model to study interactions between water limitation, environmental drivers and vegetation function at tree and stand levels, with an example application to [CO2] × drought interactions. Geoscientific Model Development, 5(4): 919-940.

Faivre, R., Iooss, B., Mahévas, S., Makowski, D. and Monod, H., 2013. Exploration par construction de métamodèles. Analyse de sensibilité et exploration de modèles. Applications aux modèles environnementaux. Quae: 159-194.

Fritsch, F.N. and Carlson, R.E., 1980. Monotone Piecewise Cubic Interpolation. 17(2): 238-246.

Ghini, R. et al., 2015. Coffee growth, pest and yield responses to free-air CO2 enrichment. Climatic Change: 1-14.

Gómez-Delgado, F. et al., 2011. Modelling the hydrological behaviour of a coffee agroforestry basin in Costa Rica. Hydrology and Earth System Sciences, 15(1): 369-392.

Granier, A., Bréda, N., Biron, P. and Villette, S., 1999. A lumped water balance model to evaluate duration and intensity of drought constraints in forest stands. Ecological modelling, 116(2-3): 269-283.

Gutierrez, A.P., Villacorta, A., Cure, J.R. and Ellis, C.K., 1998. Tritrophic analysis of the coffee (Coffea arabica) - coffee berry borer [Hypothenemus hampei (Ferrari)] - parasitoid system. Anais da Sociedade Entomológica do Brasil, 27: 357-385.

Hidalgo, H.G., Alfaro, E.J. and Quesada-Montano, B.J.C.C., 2017. Observed (1970–1999) climate variability in Central America using a high-resolution meteorological dataset with implication to climate change studies. 141(1): 13-28.

Jose, S., 2009. Agroforestry for ecosystem services and environmental benefits: an overview. Agroforestry Systems, 76(1): 1-10.

Jose, S. and Bardhan, S., 2012. Agroforestry for biomass production and carbon sequestration: an overview. Agroforestry Systems, 86(2).

Kuhn, M., 2019. caret: Classification and Regression Training. R Package version 6.0-82.

Lacointe, A., 2000. Carbon allocation among tree organs: A review of basic processes and representation in functional-structural tree models. Annals of Forest Science, 57(5): 521-533.

Landsberg, J.J., Johnsen, K.H., Albaugh, T.J., Allen, H.L. and McKeand, S.E., 2001. Applying 3-PG, a simple process-based model designed to produce practical results, to data from loblolly pine experiments. Forest Science, 47(1): 43-51.

Lasslop, G. et al., 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. 16(1): 187-208.

Lin, B.B., 2007. Agroforestry management as an adaptive strategy against potential microclimate extremes in coffee agriculture. Agricultural and Forest Meteorology, 144(1-2): 85-94.

Lin, B.B., 2010. The role of agroforestry in reducing water loss through soil evaporation and crop transpiration in coffee agroecosystems. Agricultural and Forest Meteorology, 150(4): 510-518.

Litton, C.M., Raich, J.W. and Ryan, M.G., 2007. Carbon allocation in forest ecosystems. Global Change Biology, 13(10): 2089-2109.

Luedeling, E., Kindt, R., Huth, N.I. and Koenig, K., 2014. Agroforestry systems in a changing climate — challenges in projecting future performance. Current Opinion in Environmental Sustainability, 6(0): 1-7.

Luedeling, E. et al., 2016. Field-scale modeling of tree–crop interactions: Challenges and development needs. Agricultural Systems, 142: 51-69.

Makowski, D., Nesme, T., Papy, F. and Doré, T., 2014. Global agronomy, a new field of research. A review. Agronomy for sustainable development, 34(2): 293-307.

Malézieux, E. et al., 2009. Mixing plant species in cropping systems: concepts, tools and models. A review. Agronomy for Sustainable Development, 29(1): 43-62.

Marie, G., Simioni, G. and Münkemüller, T., 2014. Extending the use of ecological models without sacrificing details: a generic and parsimonious meta-modelling approach. Methods in Ecology and Evolution, 5(9): 934-943.

Medlyn, B., 2004. A maestro retrospective. Forests at the land-atmosphere interface: 105-122.

Meylan, L., 2012. Design of cropping systems combining production and ecosystem services: developing a methodology combining numerical modeling and participation of farmers: application to coffee-based agroforestry in Costa Rica.

Murthy, V.R.K., 2004. Crop growth modeling and its applications in agricultural meteorology. Satellite remote sensing and GIS applications in agricultural meteorology: 235.

Muschler, R.G., 2001. Shade improves coffee quality in a sub-optimal coffee-zone of Costa Rica. Agroforestry systems, 51(2): 131-139.

Oelbermann, M., Paul Voroney, R. and Gordon, A.M., 2004. Carbon sequestration in tropical and temperate agroforestry systems: a review with examples from Costa Rica and southern Canada. Agriculture, Ecosystems & Environment, 104(3): 359-377.

Pezzopane, J., de Salva, T., de Lima, V. and Fazuoli, L., 2012. Agrometeorological parameters for prediction of the maturation period of Arabica coffee cultivars. International Journal of Biometeorology, 56(5): 843-851.

Poorter, H., 1994. Construction costs and payback time of biomass: a whole plant perspective. A whole plant perspective on carbon-nitrogen interactions: 111-127.

R Core Team, 2019. R: a language and environment for statistical computing. R Development Core Team, Vienna.

Rahn, E. et al., 2018. Exploring adaptation strategies of coffee production to climate change using a process-based model. Ecological Modelling, 371: 76-89.

Razavi, S., Tolson, B.A. and Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resources Research, 48(7).

Rodríguez, D., Cure, J., Cotes, J., Gutierrez, A. and Cantor, F., 2011. A coffee agroecosystem model: I. Growth and development of the coffee plant. Ecological Modelling, 222(19): 3626-3639.

Rojas-García, F., De Jong, B.H.J., Martínez-Zurimendí, P. and Paz-Pellat, F.J.A.o.F.S., 2015. Database of 478 allometric equations to estimate biomass for Mexican trees and forests. 72(6): 835-864.

Ronchi, C.P. et al., 2006. Growth and photosynthetic down-regulation in Coffea arabica in response to restricted root volume. 33(11): 1013-1023.

Ryan, M.G., 1991. Effects of climate change on plant respiration. Ecological Applications, 1(2): 157-167.
Sampson, D.A. and Smith, F.W., 1993. Influence of canopy architecture on light penetration in lodgepole pine (Pinus contorta var. latifolia) forests. Agricultural and Forest Meteorology, 64(1): 63-79.
Schnabel, F. et al., 2018. Shade trees: a determinant to the relative success of organic versus conventional coffee production. Agroforestry Systems, 92(6): 1535-1549.
Siles, P., Harmand, J.-M. and Vaast, P., 2010. Effects of Inga densiflora on the microclimate of coffee (Coffea arabica L.) and overall biomass under optimal growing conditions in Costa Rica. Agroforestry Systems, 78(3): 269-286.
Sinoquet, H., Stephan, J., Sonohat, G., Lauri, P.E. and Monney, P., 2007. Simple equations to estimate light interception by isolated trees from canopy structure features: assessment with three-dimensional digitized apple trees. New Phytol, 175(1): 94-106.
Söndahl, M., Van der Vossen, H., Piccin, A. and Anzueto, F., 2005. Espresso coffee: the chemistry of quality. In: A. Press (Editor).
Spiertz, H., 2012. Avenues to meet food security. The role of agronomy on solving complexity in food production and resource use. European Journal of Agronomy, 43(Supplement C): 1-8.
Taugourdeau, S. et al., 2014. Leaf area index as an indicator of ecosystem services and management practices: An application for coffee agroforestry. Agriculture, Ecosystems & Environment, 192: 19-37.
Van de Griend, A.A. and Van Boxel, J.H., 1989. Water and surface energy balance model with a multilayer canopy representation for remote sensing purposes. Water Resources Research, 25(5): 949-971.
van der Vossen, H., Bertrand, B. and Charrier, A., 2015. Next generation variety development for sustainable production of arabica coffee (Coffea arabica L.): a review. Euphytica, 204(2): 243-256.
van Oijen, M., Dauzat, J., Harmand, J.-M., Lawson, G. and Vaast, P., 2010a. Coffee agroforestry systems in Central America: I. A review of quantitative information on physiological and ecological processes. Agroforestry Systems, 80(3): 341-359.
Van Oijen, M., Dauzat, J., Harmand, J.-M., Lawson, G. and Vaast, P., 2010b. Coffee agroforestry systems in Central America: II. Development of a simple process-based model and preliminary results. Agroforestry Systems, 80(3): 361-378.
Van Oijen, M., Rougier, J. and Smith, R., 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiology, 25(7): 915-927.
van Praag, H.J., Sougnez-Remy, S., Weissen, F. and Carletti, G., 1988. Root turnover in a beech and a spruce stand of the Belgian Ardennes. Plant and Soil, 105(1): 87-103.
Vezy, R. et al., 2018. Measuring and modelling energy partitioning in canopies of varying complexity using MAESPA model. Agricultural and Forest Meteorology, 253–254: 203-217.
Villa-Vialaneix, N., Follador, M., Ratto, M. and Leip, A., 2012. A comparison of eight metamodeling techniques for the simulation of N2O fluxes and N leaching from corn crops. Environmental Modelling & Software, 34: 51-66.
Villar, R. and Merino, J., 2001. Comparison of leaf construction costs in woody species with differing leaf life-spans in contrasting ecosystems. New Phytologist, 151(1): 213-226.
Wang, Y.P. and Jarvis, P.G., 1990. Description and validation of an array model — MAESTRO. Agricultural and Forest Meteorology, 51(3–4): 257-280.
Wintgens, J.N., 2004. Coffee: growing, processing, sustainable production. A guidebook for growers, processors, traders, and researchers. WILEY-VCH Verlag GmbH & Co. KGaA.
Worku, M., de Meulenaer, B., Duchateau, L. and Boeckx, P., 2018. Effect of altitude on biochemical composition and quality of green arabica coffee beans can be affected by shade and postharvest processing method. Food Research International, 105: 278-285.
Wutzler, T. et al., 2018. Basic and extensible post-processing of eddy covariance flux data with REddyProc. Biogeosciences, 15(16): 5015-5030.
Zacharias, A.O., Camargo, M.B.P.d. and Fazuoli, L.C., 2008. Agrometeorological model for estimating the beginning of the flowering period for coffee crop (Coffea arabica L.). Bragantia, 67(1): 249-256.

## Appendix A: soil water balance model

The soil water balance model is largely derived from the BILJOU model (Granier et al., 1999). It is a "bucket" model, meaning that the soil is represented by the depth dimension only, itself divided into layers of given thickness.

### 1. Water balance

The water flow is managed sequentially as in any other bucket model.

#### 1.1. Interception and leaf evaporation

During a rainfall event, water can be either intercepted by the plants canopy, or directly reach the soil surface. The water intercepted by the plants is either stored on the canopy surface, or flow along the branches and trunks to finally reach the soil. The water that is stored on the canopy surface is then gradually evaporated back to the atmosphere.

The first step for computing the water balance in the model is to compute a maximum potential rainfall interception using the total stand LAI and an interception parameter as follows:

$$Interc_{max} = IntercSlope \cdot LAI_{plot} \quad \text{(A.1)}$$

Any water intercepted by the canopy when the canopy water retention is already at full capacity (*i.e.* at $Interc_{max}$, mm) is considered as throughfall water.

The daily evaporation of the water stored in the canopy bucket is computed using the Penman-Monteith equation as found in Allen et al. (1998) with an infinite stomatal conductance and a set of aerodynamic conductances.

#### 1.2. Surface runoff

Water reaching the soil surface enter the surface layer ($WSurfaceRes$, mm). Any water entering this layer when it is already full is considered as excess surface runoff, which is added to the superficial runoff itself computed using a parameter:

$$SuperficialRunoff = kB \cdot WSurfaceRes \quad \text{(A.2)}$$

#### 1.3. Infiltration

Water from the surface layer infiltrates the first soil layer up to the first layer infiltration capacity (W1), which is computed as follows:

$$InfilCapa = \begin{cases} fo & if\ W_1 \leq Wm_1 \\ fo - \frac{(W_1 - Wm_1)\cdot(fo-fc)}{Wf_1 - Wm_1} & if\ W_1 > Wm_1 \text{ and } W_1 \leq Wf_1 \\ fc & if\ W_1 > Wm_1 \text{ and } W_1 > Wf_1 \end{cases} \quad \text{(A.3)}$$

With $W_1$ the first layer water content (mm), $Wm_1$ the minimum water content of the first layer, $Wf_1$ the field capacity of the first layer and $fo$ and $fc$ the maximum and minimum infiltration capacity respectively (mm day$^{-1}$).

Then, $W_1$ is updated by adding the water that infiltrated from the surface bucket to its previous water content. If $W_1$ exceeds $Wf_1$ after this operation, the excess water drains into the second layer. This last operation is repeated for the two following layers, $W_2$ and $W_3$ with their own field capacity $Wf_2$ and $Wf_3$.

### 1.4. Soil surface evaporation

The soil surface evaporation (mm) is computed using the soil net radiation coming from a metamodel of MAESPA and a partitioning parameter ($Soil\_LE\_P$):

$$E_{soil} = \frac{Rn_{soil} \cdot Soil\_LE\_P}{\lambda} \quad \text{(A.4)}$$

with λ the latent heat of vaporization (MJ $kg_{H2O}^{-1}$).

### 1.5. Root water uptake

The water that is absorbed by the plants roots ($RootWaterExtract_l$) for a given layer $l$ is computed from the total stand transpiration ($T_{total}$), the total root fraction in the layer ($RootFraction_l$) and the extractable water from the layer ($EW_l$) as follows:

$$EW_l = W_l - Wm_l \quad \text{(A.5)}$$

$$RootWaterExtract_l = \min(T_{total} \cdot RootFraction_l, EW_l) \quad \text{(A.6)}$$

Each layer's water content is then updated by removing to it the root water uptake.

### 1.6. Water potential

The soil water potential ($\Psi_{soil,MPa}$) is computed using the equation from Campbell (1974) as follows:

$$\Psi_{soil} = \Psi_E \cdot \left(\frac{W_{tot} \cdot 1000}{\theta_s}\right)^B \quad \text{(A.7)}$$

with $\Psi_E$ (MPa) the air entry water potential, $W_{tot}$ the total soil water content (mm) and $\theta_s$ the saturated water content ($m^3\ m^{-3}$).

And the leaf water potential ($\Psi_{leaf}, MPa$) is computed as follows:

$$\Psi_{leaf,l} = \Psi_{soil} - \frac{T_l \cdot M_{H20}}{K_{tot}} \quad \text{(A.8)}$$

Where $T_l$ is the plant transpiration ($mm$), $M_{H20}$ the molar mass of water, and $K_{tot}$ the soil to leaf hydraulic conductance ($mol\ m^{-2}\ s^{-1}\ MPa^{-1}$)

## 2. Energy balance

The soil energy balance is computed using the soil net radiation from the metamodels of MAESPA and the partitioning coefficient as in Eq. (A.4) for both the latent and sensible fluxes. The soil heat storage is neglected because its variability is low at daily time-scale and because it tends to return at equilibrium after several days.

## Appendix B: shade tree allometric equations

Several allometric relationships are used in the shade tree growth module. Only the shade tree height is mandatory for the other parts of the model, because it is used to compute the canopy boundary layer conductance. Any other allometric equation can be added to the model via the tree parameter file, and can be used for user-custom metamodels or as informative output.

### 1. DBH

The diameter at breast height (DBH, m) of the *Erythrina poeppigiana* shade tree is computed following the equation from Rojas-García et al. (2015):

$$DBH = \frac{DM_{Stem}}{\left(CC_{wood} \cdot 10 \cdot \frac{Stocking}{0.5}\right)^{0.625}} \quad \text{(A.9)}$$

### 2. Height

Tree height (m) is computed following the equation from Van Oijen et al. (2010b):

$$Height = 0.46 \cdot \left(\frac{DM_{stem}}{1000 \cdot Stocking}\right)^{0.5} \quad \text{(A.10)}$$

If the shade tree is pruned, it breaks the relationship between the tree stem dry mass and its height as shown in Eq. (A.10). Instead, the tree height is computed using the trunk height as follows:

$$H_{trunk} = 3 \cdot (1 - e^{-0.2-Age}) \quad \text{(A.11)}$$

$$H_{tot} = H_{crown} + H_{trunk} \quad \text{(A.12)}$$

The underlying hypothesis behind Eq. (A.11) is that *E. poeppigiana* are pruned by stakeholders so the trunk height do not exceed 3 meters high.

### 3. Crown dimensions

The crown radius (m) is computed as follows:

$$R_{crown} = \sqrt{\frac{P_{crown}}{\pi}} \quad \text{(A.13)}$$

with $P_{crown}$ the crown projection, itself computed as:

$$P_{crown} = 8 \cdot \left(\frac{DM_{branch}}{1000 \cdot Stocking}\right)^{0.45} \quad \text{(A.14)}$$

with $DM_{branch}$the shade tree branch dry mass in $g\ m^{-2}$.

The crown height ($H_{crown}$) is taken as equal to the crown radius as shown in Table 2 from Charbonnier et al. (2013).

## 4. LAD

The leaf area density (LAD) is computed as follows:

$$LAD = \frac{LA_{tree}}{R_{crown}^2 \cdot \frac{H_{crown}}{2} \cdot \pi \cdot \frac{4}{3}} \quad \text{(A.15)}$$

with $LA_{tree}$ the shade tree leaf area.

## Appendix C: comparing DynACof with and without metamodels

To assess the importance of using metamodels from MAESPA to integrate fine scale processes to DynACof, we compared DynACof outputs for GPP, Rn, LE and H with and without metamodels. The GPP without metamodels was simulated using constant values for the light use efficiency and constant values for the light interception coefficients of the Beer-Lambert's equation for both the coffee and the shade tree layers. The parameterization was made using the average values from the MAESPA simulations. The net radiation without metamodels was simulated using the equation from Allen et al. (1998) at plot-scale using the albedo computed by MAESPA. The latent heat flux without metamodels was computed using the Penman-Monteith equation at stand-scale (PENMON function from DynACof: https://vezy.github.io/DynACof/reference/PENMON.html). The sensible heat flux without metamodels was computed as the difference between the net radiation and the latent heat flux, both computed without metamodels.

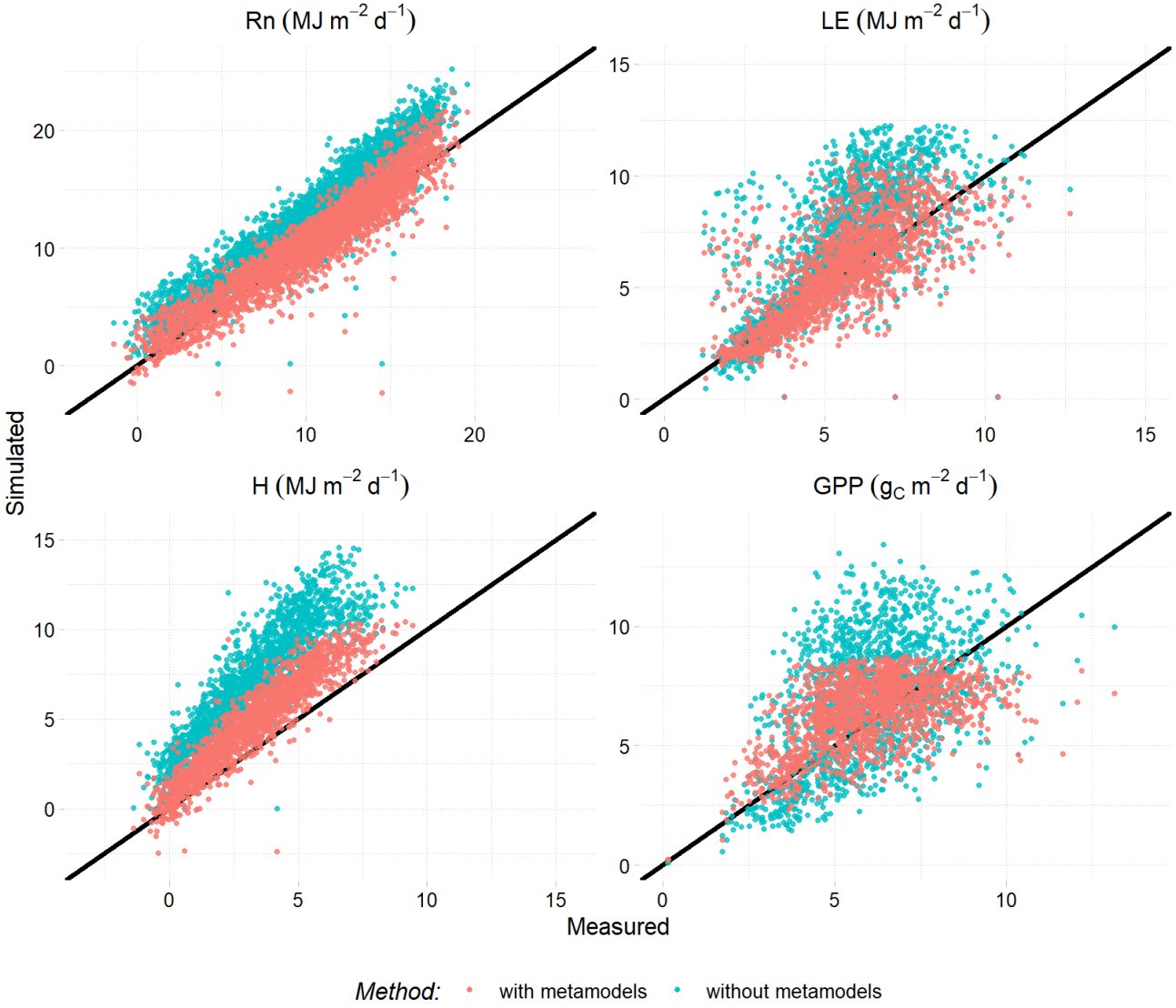


**Figure C.1. Net radiation (Rn), latent (LE) and sensible (H) heat flux, and gross primary productivity (GPP) simulations using DynACof with metamodels from MAESPA (red) or stand-scale equations (blue) compared to field measurements.**

Figure C.1 shows that GPP simulation was improved when using metamodels, and that net radiation was also improved using metamodels, but mainly because the stand-scale reference equation had a systematic positive bias compared to measurements. This bias could potentially be corrected knowing it a priori. The sensible heat flux were also systematically more biased without metamodels. The latent heat flux simulated without metamodel was close to the measurements for low values (< 5 MJ $m^{-2}$ $d^{-1}$), but was increasingly overestimated with higher values.

## Appendix D: further plots

### 1. Light interception, GPP and Ra

The contribution of the shade tree and the coffee plants to LAI, APAR, GPP and Ra are shown in Figure D.1. The highest coffee contribution to the plot-scale GPP can be explained by its higher LAI, and hence higher light interception compared to the shade tree. The shade tree thinning in 2000 has a high impact on these variables during the first three years.

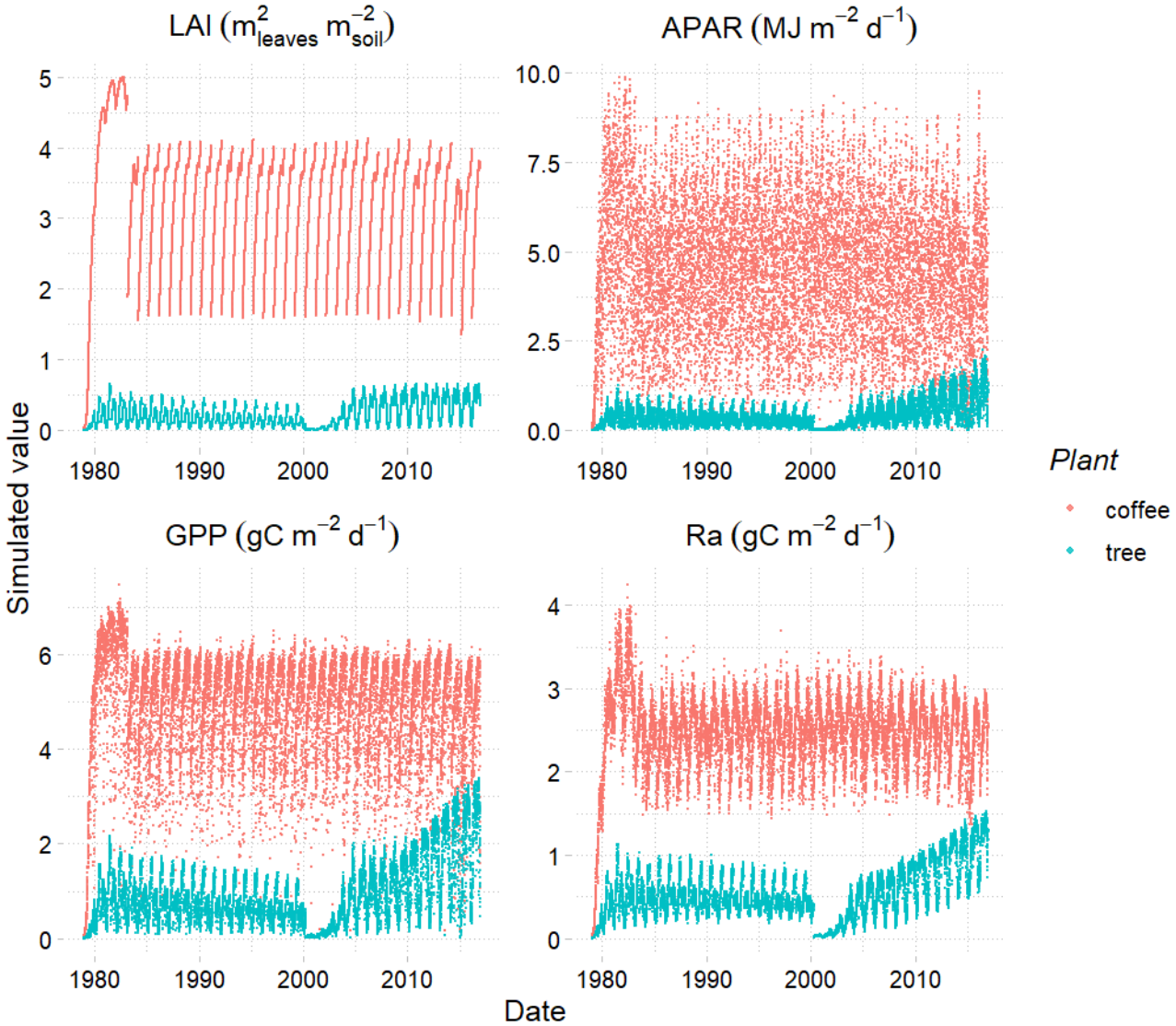


**Figure D.1. Comparison of dynamic outputs from DynACof between shade tree and coffee for leaf area index (LAI), absorbed photosynthetically active radiation (APAR), gross primary productivity (GPP) and autotrophic respiration (Ra) starting from 1979-01-01 until 2016-12-31.**

### 2. Water balance

DynACof simulates a broad range of informative variables related to the water balance of the system.

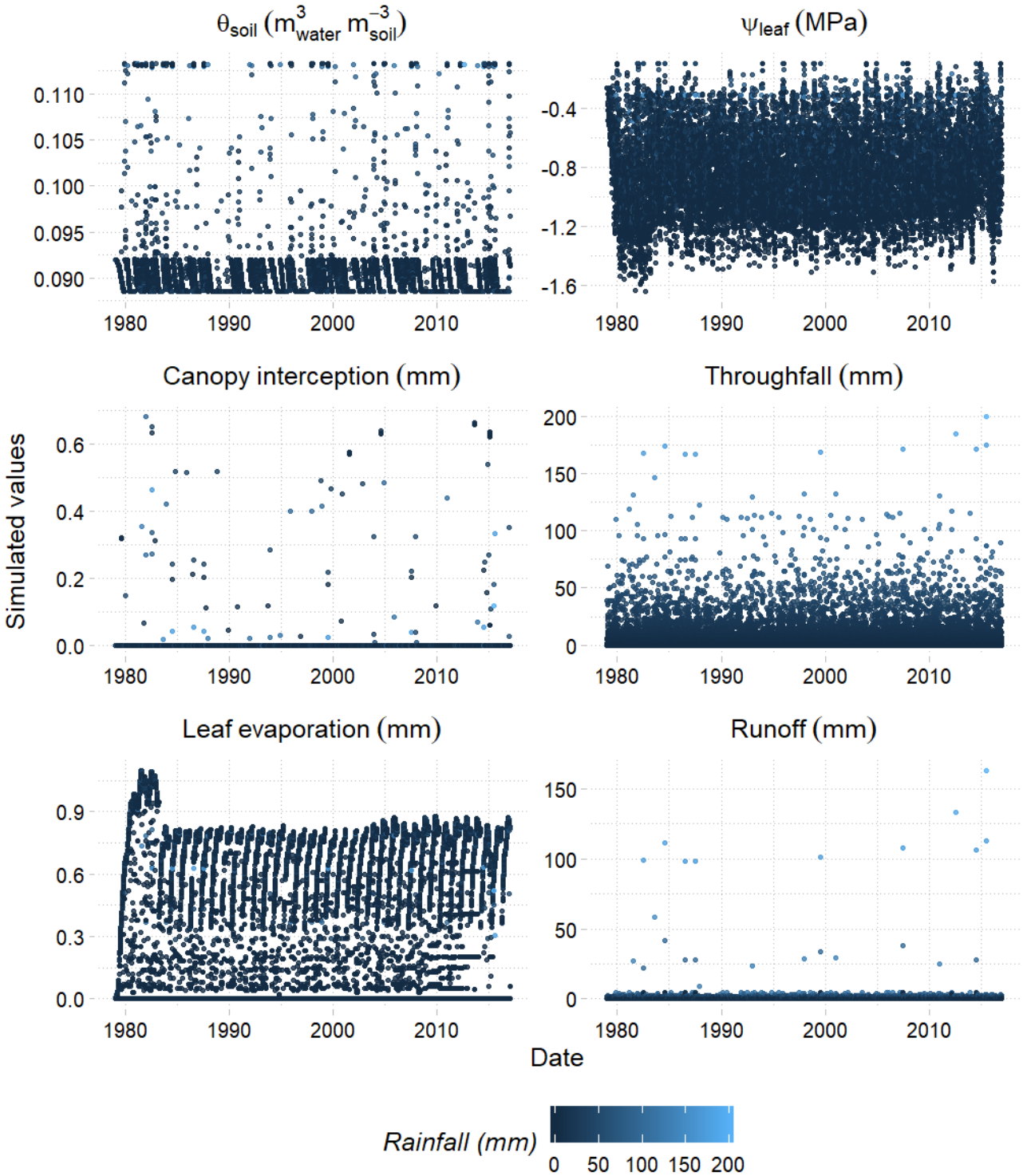


**Figure D.2. DynACof outputs related to the water balance at daily time-scale and plot scale for the period starting from 1979-01-01 until 2016-12-31. θ is the volumetric water content for the full soil profile, and .**

### 3. Temperatures

DynACof also simulates several temperatures and aerodynamic conductances in the system to better represent the microclimate experienced by the plants. DynACof uses the air temperature measured above the canopy as input, and computes the temperature of the air inside the shade tree canopy and inside the coffee canopy, and uses them to compute the leaf temperature of the shade tree and the coffee, and the soil temperature. These

computations are dependent from other computations, such as the wind extinction, the sensible heat flux of each layer (plants and soil), and a sequence of aerodynamic conductances. The temperatures simulated by the model are shown in Figure D.3.

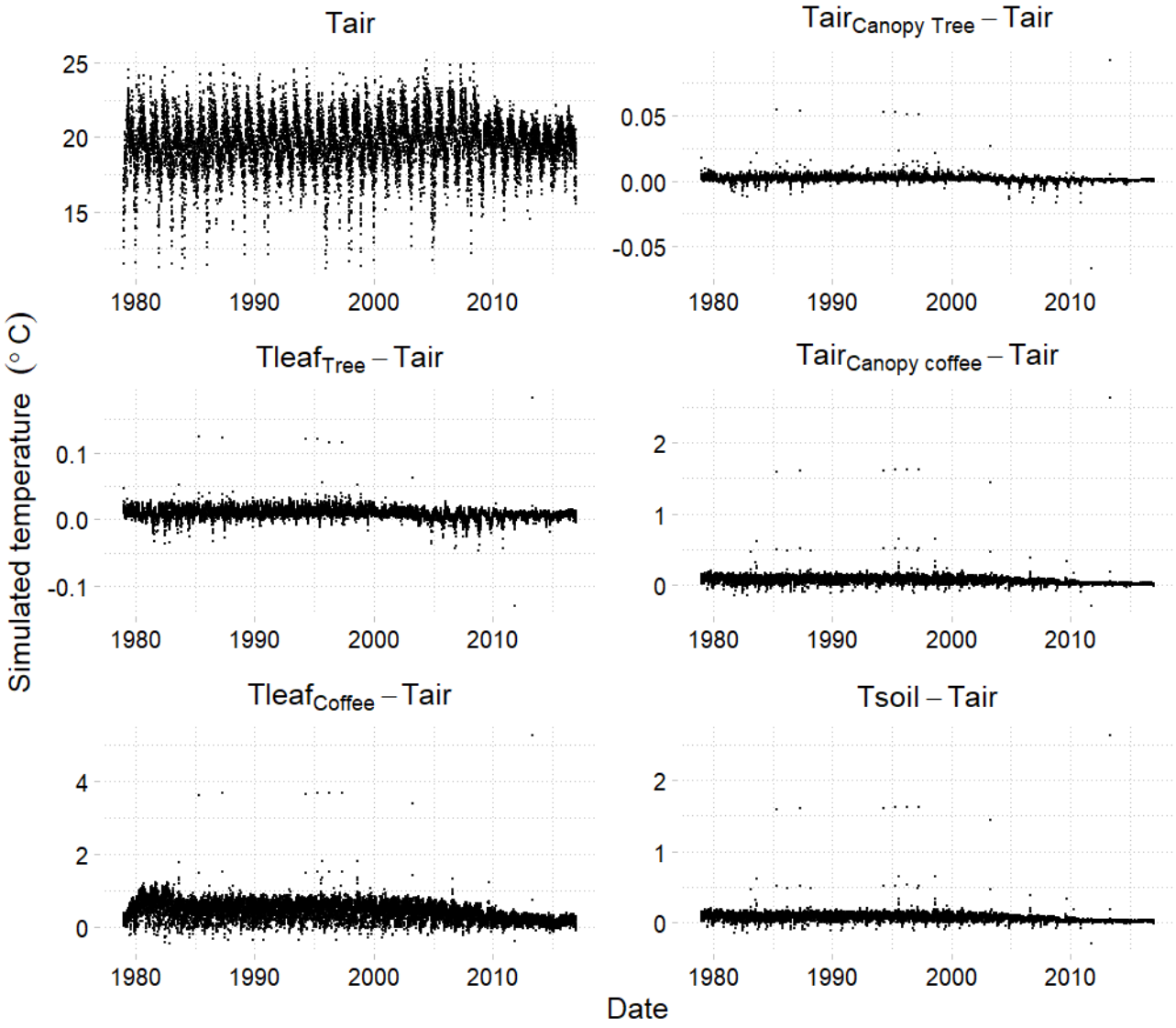


**Figure D.3. Input (Tair) and simulated temperatures. The simulated temperatures are presented as the difference between the simulated temperature and the input Tair for easier assessment.**

### 4. Litter

The litter is simulated by the model as a mortality of biomass. The coffee leaf and resprout wood litter at plot scale is mainly impacted by the pruning effect each year (Figure D.4). As expected, the fruit litter follow a seasonal variation, related to the fruit production. The shade tree leaf litter is mainly impacted by the pruning effect before 2000 (pruning management), and then by the natural seasonal variation. The simulated litterfall from the coffee (*i.e.* mortality from leaves + resprout wood + fruits) is close to the observations made by Charbonnier et al. (2017), with values of 482 $g_C$ $m^{-2}$ $year^{-1}$ compared to 450 $g_C$ $m^{-2}$ $year^{-1}$ from their measurements. The simulated tree litterfall (*i.e.* leaves + branches wood) is also close to the observations, with

a value of 96 $g_C$ $m^{-2}$ $year^{-1}$ and 110 $g_C$ $m^{-2}$ $year^{-1}$ respectively. High values of tree litterfall are explained by tree pruning before 2000.

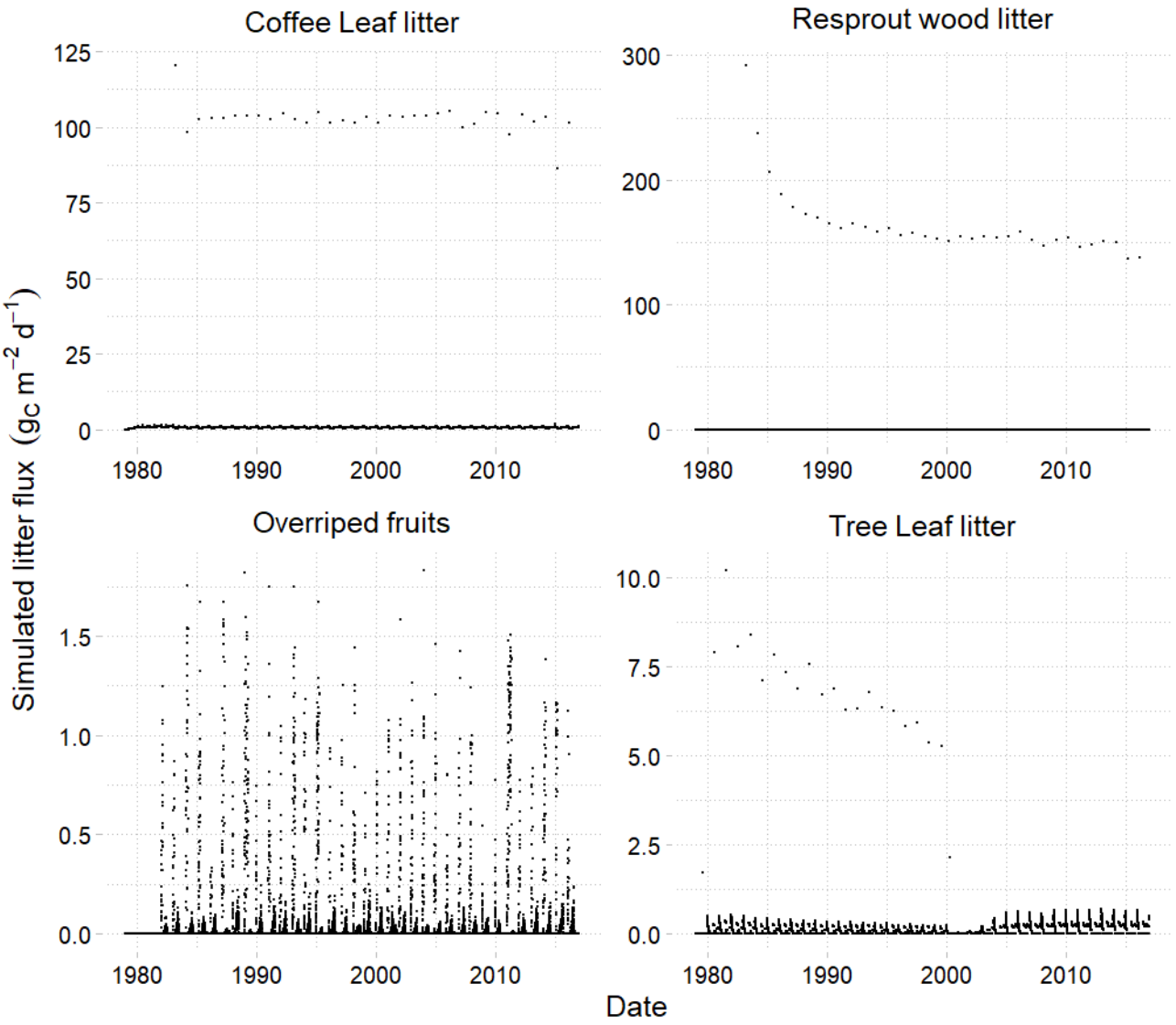


**Figure D.4. Simulated litter fluxes for the compartments with the highest contribution for the coffee (leaf, resprout wood, fruits) and the shade tree (leaf).**